Міністерство освіти і науки України
Прикарпатський національний університет ім. В. Стефаника
Вінницький національний технічний університет
Центр математичного моделювання ІППММ
ім. Я.С.Підстригача НАН України
AGH науково-технологічний університет
ім. Ст.Сташіца, Польща
Інститут кібернетики НАН України
Національний авіаційний університет
Економічна академія "Д.А.Ценов", Болгарія
Фінансово-економічний інститут Таджикистану
Лудзький університет, Польща
Новий університет Лісабона, Португалія
Азербайджанська державна нафтова академія
Об'єднаний інститут проблем інформатики НАН Білорусі
Інститут інженерів з електротехніки
та електроніки (IEEE), Українська секція
Асоціація "Інформаційні технології України"
Громадська організація "Івано-Франківський ІТ кластер"

"ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ТА КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ"

матеріали міжнародної науково-практичної конференції

15-20 травня 2017 року
Івано-Франківськ

"INFORMATION TECHNOLOGIES AND COMPUTER MODELLING"
proceedings
of the International Scientific Conference
2017, May, 15th to 20th
Ivano-Frankivsk

Івано-Франківськ - 2017

УДК (004:004.2/004.9+007):33/37+51+621
ББК 22.17 32.81
І-74 Т

Наукові редактори: докт. техн. наук, проф. **Л.Б. Петришин** (ПНУ, AGH)

Матеріали статей опубліковані в авторській редакції

"Інформаційні технології та комп'ютерне моделювання"; матеріали статей Міжнародної науково-практичної конференції, м. Івано-Франківськ, 15-20 травня 2017 року. – Івано-Франківськ: п. Голіней О.М., 2017. – 464 с.

Збірка містить матеріали статей Міжнародної науково-практичної конференції з проблем інформаційних технологій в технічних системах, в соціумі, освіті, медицині, економіці та екології; теорії інформації, кодування та перетворення форми інформації; технологій цифрової обробки інформації; захисту інформації в інформаційно-телекомунікаційних системах; математичного та імітаційного моделювання систем.

УДК (004:004.2/004.9+007):33/37+51+621
ББК 22.17 32.81
І-74 Т

ISBN

© ПНУ ім. В. Стефаника та автори, 2017

СЕКЦІЯ 1.

Інформаційні технології в технічних та системах спеціального призначення

SECTION 1.

Information technologies in technical systems and systems of special purpose

Спосіб Формування Множини Оціночних Сигналів для Дослідження Функціонального Простору Об’єкта

Юрій Стрілецький
кафедра методів та приладів контролю якості і сертифікації продукції
Івано-Франківський національний технічний університет нафти і газу
Івано-Франківськ, Україна
momental@ukr.net

Віктор Ровінський
кафедра інформатики
Прикарпатський національний університет імені Василя Стефаника
Івано-Франківськ, Україна
musicneutrino@gmail.com

Method of Evaluation Signal Set Formation for Studying Object's Functional Space

Yurii Striletskyi
methods and devices of quality control and certification
IFNTUOG
Ivano-Frankivsk, Ukraine
momental@ukr.net

Victor Rovinsky
dept. of Computer Science
Precarpathian National University
Ivano-Frankivsk, Ukraine
musicneutrino@gmail.com

***Анотація*— Розглянуто спосіб формування сигналу для дослідження функціонального простору об’єкта при його ідентифікації. Запропоновано використовувати фазоманіпульований сигнал із двох біт, що дало можливість визначити затримку поширення тестового сигналу в об’єкті за допомогою аналізу взаємокореляційних функцій отриманого сигналу із чотирма опорними сигналами, які синхронні із тестовим, але мають визначені фазові зсуви.**

***Abstract*— Was proposed method forming the signal for discovery functional space of object. A use BPSK signal of two bits, making it possible to determine the propagation delay test signal in the object by analyzing the received signal cross correlation with four reference signals and are synchronous with the test signal, but have a certain phase shift.**

Ключові слова— спектральне представлення, взаємокореляційна функція, час запізнення сигналу, огинаюча сигналу, фазоманіпольований сигнал

Keywords— spectral representation, cross correlation, time delay signal, baseband signal, BPSK

I. Вступ

В процесі пізнання невідомі об’єкти ідентифікуються шляхом побудови моделі, яка в повній мірі відображає механізм функціонування об’єкта з позиції дослідника. Ідентифікація об’єкта супроводжується етапами специфікації і оцінки. Задача специфікації полягає у виборі зручного і ефективного математичного способу опису системи. Такий опис як правило представляє собою рівняння, що містить набір оцінюваних параметрів. При ідентифікації застосовують комбінацію двох підходів – оцінку параметрів і визначення характеристик об’єкта. Описані підходи використовують дедуктивний і індуктивний методи пізнання, тому мають свої переваги і недоліки.

Згідно дедуктивного методу структурна конфігурація об’єкта вважається відомою чи відносно неї робляться певні припущення, а параметри диференційних рівняннь, що описують систему вважаться невідомими. В цьому випадку задача ідентифікації зводиться до пошуку рішення в просторі параметрів, що забезпечують мінімізацію деякої функції помилки.

Індуктивний метод побудови моделі об’єкта передбачає дослідження функціонального простору цього об’єкта. За таким методом шукається множина відкликів об’єкта, пов’язаних із його функціонуванням на множину оціночних подразнень. При цьому отримується опис об’єкта в обмеженій і заздалегідь заданій області

функціонального простору. Розширення функціонального простору об'єкту можна здійснити шляхом пошуку його додаткових функціональних особливостей. Здебільшого розглядається експериментальні дослідження поведінки об'єкта. Тому важливо отримати максимум інформації про стан об'єкта при наявності обмежень, пов'язаних із його особливостями.

Очевидно, що при повній відсутності попередніх знань про об'єкт проводити ідентифікацію простіше за відношенням множини оціночних подразнень і відкликів системи. Вибір множини оціночних подразнень проводиться із врахуванням функціонального простору об'єкту, який здебільшого передбачає зміну його стану.

Множина оціночних подразнень формується із врахуванням функціонального простору об'єкта. Для повнішої оцінки множини станів об'єкта, пов'язаних із його функціонуванням необхідно використовувати таку ж або й більшу кількість оцінок подразнень. Отримати велику кількість оцінок подразнень, які здебільшого формуються простором сигналів, можна за допомогою багатомірних сигналів.

Одним із різновидів представлення багатомірного сигналу із простору сигналів подразнень є спектр. Спектральні складові, які містять відомості про амплітуду і затримку є незалежними один від одного[1]. Тому їх аналіз дозволяє отримати значну кількість інформації для побудови функціонального простору досліджуваного об'єкту, стосовно вибраної множини оціночних подразнень. При цьому важливо виділити відклик саме на вибрану множину подразнень. Тому оцінюваний сигнал із простору вхідних сигналів має формуватися так, щоб його можна було виділити на фоні сторонніх відкликів. За умови лінійності відображення оціночного подразнення станом об'єкта зручно формувати простір сигналів подразнення із гармонійних сигналів.

Дослідження гармонійних сигналів в лінійних системах пов'язана із пошуком амплітуд і фаз. Враховуючи загальну задачу пізнання об'єкта інформація про фазу має обмежене застосування, оскільки більш загальним є знання про запізнення сигналу.

Отже завданням при дослідженні об'єкта в межах певного функціонального простору є формування множини сигналів, які можна було б вирізнити на фоні завад і знаходити їх спектральні амплітуди та запізнення.

II. Викладення основно матеріалу

Одним із способів виділення сигналів на фоні завад є використання кореляційної обробки. Загально прийнятим способом пошуку часу затримки сигналу за взаємокореляційною функцією (ВКФ) є по-точкове визначення значення ВКФ із пошуком положення її максимуму. Такий спосіб передбачає значний обсяг обробки. Зменшити цей обсяг і тим самим спростити засоби обробки можна вибравши таку форму сигналу, при якій за будь-яким виміряним значенням можна встановити часовий зсув. Пошук часового зсуву проводиться за максимумом кореляційної функції сигналу який взаємодіяв із досліджуваним об'єктом і опорним сигналом тієї ж форми. При цьому ВКФ фактично набуває виду (рис.1)

Найбільш прийнятною є така форма сигналу, часовий зсув для якої можна визначити за відносними значеннями ВКФ, оскільки при перетворенні сигналу в об'єкті абсолютні значення ВКФ будуть мінятися через вплив сторонніх процесів. Сигнал пропонується формувати за допомогою фазової маніпуляції гармонійного сигналу із досліджуваною частотою. Модулюючим сигналом є двійковий код. Критерієм вибору коду модулюючого сигналу є можливість однозначного представлення значення його авто кореляційної функції (АКФ) за допомогою двох ортогональних координат. Це дасть можливість визначити часовий зсув за допомогою коефіцієнтів кореляції вхідного сигналу із двома опорними ортогональними сигналами, один із яких синхронний із формуванням тестового.

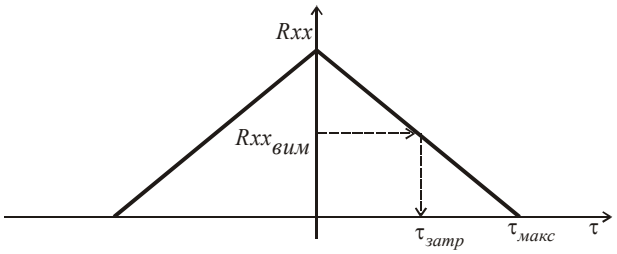


Рис. 1. Визначення часу затримки поширення сигналу за одним вимірам авто кореляційної функції

В результаті аналізу було встановлено, що найбільш прийнятна форма модулюючого сигналу має форму симетричного прямокутного сигналу. АКФ такого сигналу має форму трикутника. Тому для модуляції несучого коливання досліджуваної частоти використано сигнал симетричної прямокутної форми. В такому випадку сигнал, який використовується для подразнення об'єкта буде визначатися функцією

$$s_{ФМ}(t) = A_m \cdot sin(\omega_0 \cdot t) \cdot sign\left[1 - \frac{\omega_1}{n \cdot 2\pi} \cdot MOD\left(t, n \cdot \frac{4\pi}{\omega_0}\right)\right],$$

де A_m - амплітуда коливань, ω_0 - частота коливань, $sign\left[1 - \frac{\omega_1}{n \cdot 2\pi} \cdot MOD\left(t, n \cdot \frac{4\pi}{\omega_0}\right)\right]$ - функція, яка набуває значення +1,-1 в залежності від часу t і кількості коливань сигналу n впродовж одного біта.

При дослідженні вхідного ФМ сигналу необхідно визначити зміщення кадру відносно опорного сигналу. Зміщення кадру для періодичного кадру визначається фазовим кутом чи фазовою швидкістю поширення сигналу в об'єкті. Кадр формується періодом модулюючого сигналу. ВКФ отриманого із об'єкта відклику і синхронно утвореного опорного сигналу має інформацію про часовий зсув відклику. Форма сигналу впродовж одного кадру і ВКФ його із опорним сигналом наведені на рис.2.

Розглядаючи один кадр вхідного сигналу як період коливання можна представити сигнал ВКФ геометрично за допомогою вектора. Для побудови положення вершини вектора скористаємось квадратурною складовою ВКФ вхідного сигналу із зміщеним на 2π по відношенню до модулюючого сигналу. Положення вершини вектора ВКФ вхідного сигналу із квадратурним взірцевим сигналом впродовж тривалості кадру зображено на рис.3.

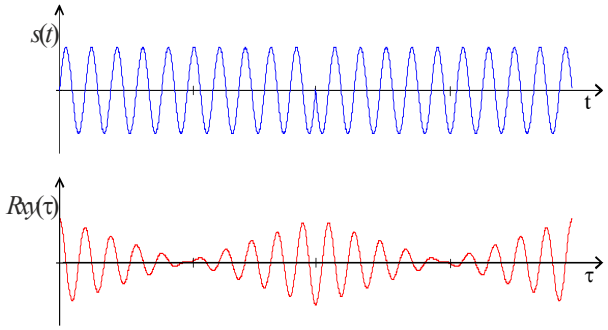


Рис. 2. ФМ сигнал і ВКФ його із взірцевим сигналом

Із наведених графіків видно, що локальні максимуми ВКФ описують форму квадрата. Це обумовлено прямокутною формою модулюючого сигналу. При цьому вектор пульсує із частотою несучого сигналу. В даному випадку, внаслідок симетричності ФМ коливання в часі операцію кореляції можна розглядати як згортку. Тому результатом кореляції буде узгоджена фільтрацію вхідного сигналу. Коефіцієнтами узгодженої фільтрації буде взірцевий ФМ сигнал, який також має гармонійну несучу.

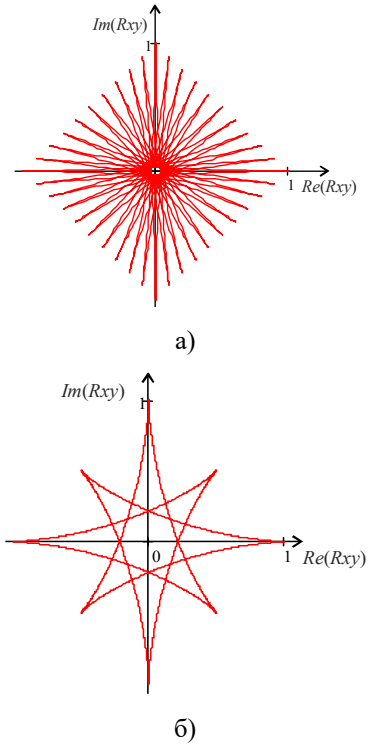


Рис. 3. Положення вершини вектора ВКФ ФМ сигналу із гармонійною несучою і різною кількістю коливань на біт

Для пошуку фазового кута, який визначає часовий зсув кадру необхідно знайти огинаючу квадратурних ВКФ по кадру ВКФ (рис.4).

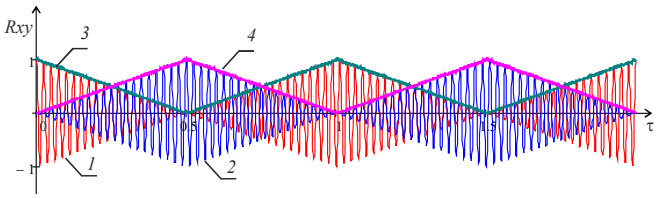


Рис. 4. ВКФ вхідного ФМ сигналу із квадратурними гармонійними сигналами 1,2 та їх огинаючі 3,4 відповідно.

Огинаючу можна знайти, перетворенням дійсного сигналу ВКФ в аналітичний сигнал за допомогою перетворення Гільберта. Однак за відсутності представлення в один момент часу всієї ВКФ будемо шукати аналітичний сигнал представлений однією точкою, яка містить дві координати.

Однією координатою є ВКФ вхідного ФМ сигналу $s(t)$ із опорним сигналом $s_{00}(t)$, який синхронний із генератором. Другою координатою є ВКФ сигналу $s(t)$ із опорним сигналом зсунутим на $\pi/2$ по відношенню до несучого сигналу $s_{90}(t)$.

Квадратурну компоненту так само шукаємо за допомогою двох координат. Однією координатою цієї компоненти є ВКФ сигналу $s(t)$ і опорного сигналу зсунутого на $\pi/2$ по відношенню до модулюючого сигналу $s_{09}(t)$. Іншою координатою є ВКФ сигналу $s(t)$ і опорного сигналу зсунутого на $\pi/2$ по відношенню до модулюючого сигналі і ще на $\pi/2$ по відношенню до несучого сигналу $s_{99}(t)$. Положення векторів сигналів, які використані при пошуку координат вектора квадратурної ВКФ наведено на рис.5

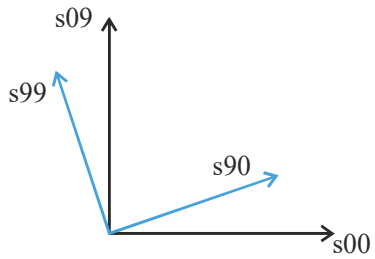


Рис. 5. Положення векторів опорного сигналу для пошуку огинаючої квадратурної ВКФ

Види опорних сигналів, які використовуються для пошуку аналітичного виду квадратурних ВКФ наведено на рис.6.

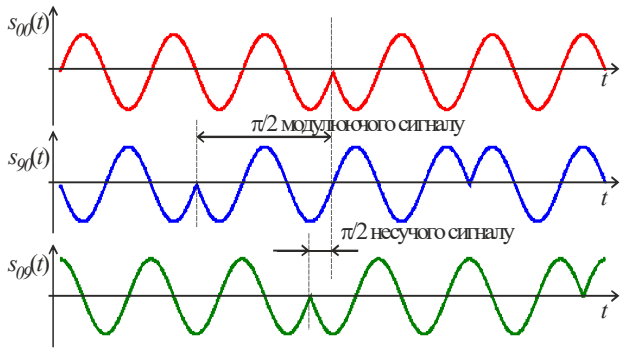


Рис. 6. Опорні сигнали для пошуку ВКФ із вхідним сигналом

Положення вершини вектора квадратурної ВКФ можна знайти за модулями аналітичного сигналу в момент вимірювання за формулами

$$Rxy_\tau = \sqrt{R00_\tau^{\,2} + R90_\tau^{\,2}} + j\sqrt{R09_\tau^{\,2} + R99_\tau^{\,2}}$$

Знак кута визначається знаками окремих складових.

$$\varphi_\tau = \begin{cases} arctg\left(\dfrac{Im[Rexy_\tau]}{Re[Rexy_\tau]}\right) & для \quad Im[Rexy_\tau] > 0 \cap Re[Rexy_\tau] > 0 \\ arctg\left(\dfrac{Im[Rexy_\tau]}{Re[Rexy_\tau]}\right) + \pi & для \quad Im[Rexy_\tau] > 0 \cap Re[Rexy_\tau] < 0 \\ arctg\left(\dfrac{Im[Rexy_\tau]}{Re[Rexy_\tau]}\right) + \pi & для \quad Im[Rexy_\tau] < 0 \cap Re[Rexy_\tau] < 0 \\ arctg\left(\dfrac{Im[Rexy_\tau]}{Re[Rexy_\tau]}\right) + 2\pi & інакше \end{cases}$$

Форму огинаючої квадратурної ВКФ наведено на рис. 7.

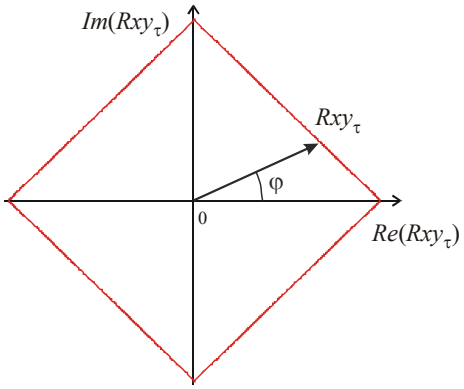
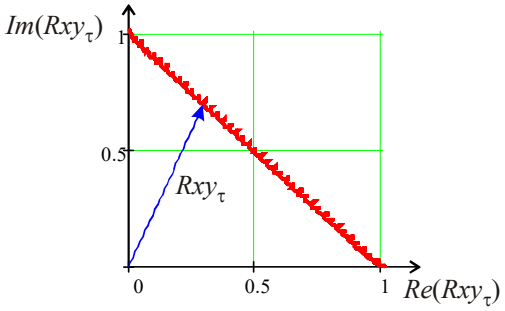


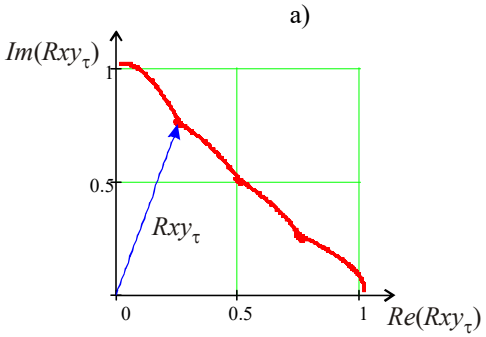
Рис. 7. Представлення вершини вектора огинаючої ВКФ вхідного і опорного сигналів

Таким чином, в залежності від того в якому квадранті знаходить вектор Rxy_τ, його кут можна визначити за результатами вимірювання чотирьох ВКФ для одного моменту часу.

При малій кількості періодів на біт кадру спостерігаються нелінійності у формі огинаючої ВКФ. На рис.8 наведено фрагмент огинаючої ВКФ вхідного ФМ сигналу і опорного сигналу із несучим сигналом гармонійної форми.



а)



б)

Рис. 8. Вигляд огинаючої ВКФ для а- 4 періодів на біт, б- 40 періодів на біт

З наведеного графіку видно, що при малій кількості періодів, які формують один кадр ФМ сигналу буде незначна похибка у визначенні фазового кута періодичної послідовності. При більшій кількості періодів похибка зменшується.

III. Висновки

Результатом пошуку є амплітуда і фаза пакету імпульсів із частотою досліджуваної частоти. Тривалість пакету більша часу поширення сигналу, тому вирішується проблема невизначеності кількості періодів, які можуть пройти поки сигнал досліджуваної частоти взаємодіяв із об'єктом дослідження.

Література References

[1] Миддлтон Д. Введение в статистическую теорию связи. ' Пер. с англ. Т. 1. М.: Сов. радио, 1961. 782 с

Технологічні Можливості Інформаційного Серверу Hyperwave

Бойко Наталія Іванівна
Кафедра інформаційних систем та мереж
Національний університет "Львівська політехніка"
Львів, Україна
nataliya.i.boyko@lpnu.ua

Копач Ольга
Кафедра інформаційних систем та мереж
Національний університет "Львівська політехніка"
Львів, Україна
olya.kopach@gmail.com

Technological Possibilities of Hyperwave Information Server

Boyko Nataliya Ivanivna
Department of Information Systems and Networks
Lviv Polytechnic National University
Lviv, Ukraine
nataliya.i.boyko@lpnu.ua

Kopach Olga
Department of Information Systems and Networks
Lviv Polytechnic National University
Lviv, Ukraine
olya.kopach@gmail.com

***Анотація*— У статті розглядаються рішення для Інтранету як глобально операційної обробки інформації компанії. Наводиться розвиток системи управління знаннями, що стосується посилань на інформаційну систему на основі інформаційного сервера Hyperwave. Представленні певні рішення від загальної моделі Hyperwave у вигляді окремих модулів. Наведенні комплексні висновки. Розглядаються можливості розширення і вдосконалення моделі Hyperwave.**

***Abstract*— In the article an intranet solution for a globally operating information processing company is considered. The development of the knowledge management system is concerned especially with the reference to the Hyperwave information system based on the Hyperwave Information Server. Two partial solutions from the general model are then presented as separate modules (metadata records and news application). Then the complex conclusions are drawn. The expansion and improvement possibilities of the model are also presented.**

Ключові слова— створення оцінки; управління знаннями; інформаційна система; обробка інформації компанії; Інтранет; інформаційний сервер Hyperwave; інформаційна система Hyperwave.

Keywords— value creation; knowledge management; information system; information processing company; intranet; Hyperwave Information Server; Hyperwave information system.

I. INTRODUCTION

The biggest challenge in today's working world: Information – namely, too much of it. Without functioning document management, it is hard to work efficiently. As well as creating and editing documents, what is important is the ability to access documents regardless of location and device, and an efficient repository and administration that allow you to find any document you are looking for quickly. The Hyperwave solution guarantees all this [1, 5].

Whether we're talking about digital or paper documents – the primary function of any document management system is to allow the publication of any type of content. Only then can content be edited and assigned attributes, or be automatically distributed or archived.

Hyperwave has made these editing options so simple and intuitive that the user can create and publish documents without any help from administrators. At the same time, we guarantee that the appropriate company processes and access rights regulations are complied with. The easy and at the same time secure handling ensures productive operation from the first day – and frees up your IT people to deal with other projects [4, 7].

The fact that knowledge is the most important source of value creation in modern organizations is nowadays recognized worldwide. The rapid growth and the increasing sharing of knowledge, as well as the virtualization of enterprises and the new possibilities of the information technology compels conglomerates to establish a systematic and active knowledge management. This is usually achieved by the means of the complex mechanisms which carry out both the modeling of the functional and structural

characteristics of the company as well as map them on the information system [3].

II. THE CHALLENGES OF THE KNOWLEDGE-BASED INFORMATION SYSTEM DEVELOPMENT

The appropriate technology should be efficient in the performing the above mentioned tasks. Further challenges in the design of the knowledge-based information systems for information processing companies include [2]:

1. Geographically distributed information as well as administration. Distributed systems do not only require appropriate technical solutions, which assure the functionality and stability of the entire information system, but also demand an elaborate strategy for the knowledge units management.
2. Diverse profiles of users that provide them with a personalized demand-oriented information system view. A company-wide, consistent user management over the highly distributed system architecture should harmonize with an effective document management.
3. Global distribution of databases that can restrict or obstruct 'fast' data transfer for large or time-critical documents. Not only the technical aspects of replication mechanisms but also the strategic decision-making criteria play a crucial part.
4. The urgent need for the development of updating and synchronization mechanism which will optimize consistency, error tolerance and performance of the storage system.

Preference should be granted to the established model approaches that offer a viable and standardized solution. The employment of standards in the information system implementation provides both a lasting and flexible, thus more cost-efficient way to expand or improve the system [6].

III. SELECTION OF THE SUITABLE TECHNOLOGY

The above analysis allows to assume that an intranet concept is a solution that represents some of the immensely important aspects that can facilitate the knowledge-based information system in the information processing company.

Intranet supports the alignment, retrieval and exchange of information within a company, therefore the knowledge base behind it is invisible and inaccessible to unauthorized users. An intranet is based on the same technologies and services as the Internet and represents the Internet at the enterprise level [2, 9].

Intranet is considered here as a tool for information management and not only as a collector of documents. That is, the focus or idea behind the use of the Intranet technology should be foremost motivated by the goals set and are not by the documents. The reason for this is obvious: people use information units to accomplish specific tasks.

Information technology companies strive to find a solution for their information system while their business processes are handled efficiently and profitably. Modern knowledge-based information systems should therefore be able to provide the right information at the right time in the right place available [8].

The technological possibilities of the Hyperwave Information Server (version 5.5 SP2) allow to establish the core architecture of the information system. Therefore the intranet environment will be implemented on the Hyperwave Information Server [11].

IV. HYPERWAVE INFORMATION SERVER

Developed in the Graz University of Technology in Austria in the early 1990s Hyperwave was first introduced in 1997 as the first commercial version of the Hyper-G technology. The application, which underlies the entire Hyperwave content management infrastructure, now goes by the name of Hyperwave IS/6.

Hyperwave is an information system focused on the document storage and service. By 'document' means any chunk of data that coexists with object record. It contains the metadata of the document – the list of attributes, open to the enhancement by user. Some of the attributes are always set by Hyperwave Server, some of them are modifiable [1, 11].

Attribute refers to the pair name/value presented as name=value. Complete object record contains the amount of pairs necessary for the user. The attribute name does not require to be equal, that is the title can appear several times throughout the object record. It makes sense in case of title specification in different languages. There is an alignment for that case – the title is preceded by a two-letter abbreviation of the language, for example: 'en: Title in English' or 'ge: Titel in deutsch'. Other attributes – description or keywords – are potential candidates. Hyperwave allows to replace language abbreviation by any other string separated from the rest of the attribute by a column [5, 13].

Each object entry has its own string representation where each name/value pair is separated by newline characters; Hyperwave extension also provides another representation – an associative array with an attribute name for a key. Multilingual entries form another associative array with a language abbreviation key.

Apart from document all its hyperlinks are stored as object records. Document hyperlinks are deleted and saved in separate objects when the document is inserted into the database. Hyperlink object record contains the information about its beginning and its end. To get the document user needs to request hyperlink-free document and insert the hyperlinks with the appropriate commands. The advantages of document-links separation are evident: if the name of linked document is modified, the link is easily changed. The link-contained document is not affected in the process at all. The link can be added to the document without modifying the last [12].

The Web-based Hyperwave infrastructure offers an integrated and configurable product family for knowledge management and eLearning, which includes several technical solution areas, including document management, content management, archiving with configuration management, special search function, extensible metadata described above,

knowledge portals, automatic link management, workflow, team workspaces, and other. The operation of the Hyperwave Information Server is versatile and comprehensive in its applicability [14].

V. TECHNICAL IMPLEMENTATION OF THE INFORMATION SYSTEM

The complex locational, functional and structural constraints within the information processing company result in the following list of the critical aspects required for consideration while introducing a knowledge-based information system [6, 13]:

1. Technology choice and the information system architecture

The technical implementation of the information system is to take place with the help of technology from the Hyperwave Information Server.

2. Availability and life cycle of data sets

Carried out in the process of knowledge development within a large and distributed system the analysis of the document nature relating to the life cycle of the company's intellectual capital allows to regard documents as explicit knowledge, therefore life cycle of documents exists.

The life cycle model for documents within a centralized information system identifies five different stages in the life cycle of a document:

- creation;
- publishing;
- organization;
- access;
- destruction.

The examination of the document lifecycle within the Information System, that is the habitat of the document, plays a crucial role.

3. Semantic enrichment of databases by metadata.

The semantic information processing by defining metadata for the individual units of the digitally captured intellectual capital proves to be very beneficial as it increases the content analysis quality and simplifies the meaning extraction from the document content.

4. Multilingual functionality of the system

Of particular importance in the management of multilingual data is the representation of the semantically identical information units. Such information units are, on the one hand, 'physically close' (e. g. via their shared storage in special container structures), and, on the other hand, managed 'virtually close' (e. g. by clustering mechanisms or Topic Maps).

5. Information units retrieval and presentation

The basic issue is the logical and technical separation of layout and contents of the dataset. In order to ensure the transparent presentation layout and behavior of interfaces follow a predefined scheme. The presentation of the scheme is set in the framework of the CI (corporate identity), the content can be adapted to correspond with portal structure. Overall six different frames are defined:

- header frame;
- navigation frame;
- search frame;
- quick links frame;
- code frame;
- content frame;

each of which displays portal specific information.

6. Company-specific role concept

A thoroughly elaborated role concept is essential for the information system of the information processing company, since it represents the general user and department administration, and therefore, exerts a direct influence on all access rights dependent operations.

7. Dynamism and modularity of the system

The term 'dynamic structures' refers to all mechanisms or modules that can be adapted to both specific strategic or operational changes regardless of the change duration (that is short, medium or long). The aspects of dynamics and modularity are considered on the all technical levels in order to overcome any future system modifications.

8. Conceptual structures creation and use

With a globally defined conceptual structure the relevancy and accuracy criteria of the knowledge access is solved on the systematic level. With the merger of companies – which are specialized for example in various fields – the structures with different duration can provide the ability to generate various departmental perspectives on the knowledge units of the system structure.

9. Increase collaboration and communication

Such demand is usually driven by the urge to improve the knowledge division. This aspect has a particularly high priority in the knowledge management and is implemented by customized solutions of the information system. It fulfills the requirement to increase collaboration and communication through the tools that provide centralized and controlled knowledge sharing. Solutions of this type are shown, for example, by the following components: eLearning platforms, discussion forums, chat systems, email applications, video conferencing, etc.

10. Specific archiving and data management criteria

Under this aspect, decisions taken determine the possible use of database or file systems (or a combination of both). Hyperwave ensures the efficient and effective data management by the combination of the following mechanisms: assigning access rights on specific object attributes, upload mechanism for all types of multimedia objects, configurable version control for document objects, editing HTML files directly to the Server (without obligatory download), specialized container objects (Cluster, Sequence, Collection, Document Classes, Search Query Object) etc..

The information system provides users with global or partial views of the dataset depending on the current role status (usually a member of a group), the activity mode (read, write, etc.) and of the – under Hyperwave – adjustable 'User Preferences' [9].

All points described so far provide an overview of the structure and functionality of the knowledge-based information system for an information processing corporation

[14]. The technical implementation of the requirements for such a system was presented based on the Hyperwave Information Server.

Not least, the Hyperwave solution doesn't care what device you use to access the document. You can access all data, without any time or location restrictions and outside of the company boundaries, edit your document and share it with colleagues or external partners

It is often referred to as the 'knowledge of power'. But no matter what you call it – the very specific knowledge that exists in your company is more or less dead capital, unless you draw on its full potential. Only then will it become a decisive factor in production. That is why it is no longer enough to simply gather and archive information and data; the pieces of information all have to be linked to each other and you have to be able to access them as quickly as possible. Hyperwave can also help your company to benefit from this kind of knowledge management [10].

Hyperwave saves all documents and related meta data in a relational database. The Hyperwave Team Workspace is the solution for web-based team working. It offers a working environment for internal and external project groups, that is both easy to set up and intuitive to use. Any user with appropriate access rights can set up a Hyperwave Team Workspace in about two minutes. And it also offers you all the document management functions you might require – with no previous knowledge you can publish, edit and replace documents easily. This allows you to start your project quickly, and also lets you bring in new employees at any stage of the project, without the need for extra training.

Using only web-based technologies, the Hyperwave Team Workspace can be deployed anywhere and immediately, without a major rollout. All internal and external project members and partners can work online or offline, regardless of their timezone or location.

An overview page lets you see any changes in one quick glance. Long drawn-out searches for modified project documents are a thing of the past. A shared calendar and task lists that can be accessed by all participants over the Web help to plan and coordinate tasks [4].

Our daily business is ruled by documents and unstructured data such as contracts, reports, calculations, presentations, design drawings and photos. Having quick and easy access to the correct and most up-to-date documents gives you a distinct competitive advantage.

A Hyperwave document management solution will guarantee that you find the documents you are looking for fast, and that you are always accessing the latest version of a document. The confidentiality of documents is assured using role-specific display of content. In this way, you are guaranteed that company employees have access to the information they require, while confidential content is protected [8].

With Hyperwave, e-mail problems, finding the correct files, and data and versioning chaos are all things of the past.

VI. CONCLUSION

The knowledge resource in large and geographically distributed information processing companies is now either explicitly or implicitly embedded in its processes. The basic objective of such corporations in this context is to generate knowledge and to increase productive and efficient use of the computer-aided knowledge systems. Therefore, the company's corporate memory requires an organized and orderly administration.

The discussed in this article model is based on the technology and architecture of the Hyperwave Information Server. The implemented knowledge-based information system provided the necessary tools for an efficient and effective solution. The technological advantages of the Hyperwave Information Server have been widely exploited for the technical implementation and provided an optimal solution for the problem at each stage of the analysis.

REFERENCES

[1] G. Bellinger, D. Castro, A. Mills, Data, Information, Knowledge, and Wisdom; 2004

[2] G. Bellinger, OutSights; Knowledge Management – Emerging Perspectives; 2004

[3] M. Ginsburg; Intranet Document Management Systems as Knowledge Ecologies; Proceedings in 33. Hawaii International Conference on System Sciences. Hawaii, USA, 2000

[4] A.J. Godbout, Filtering Knowledge: Changing Information into Knowledge Assets; Technical Report 9867, Godbout Martin Godbout & ass. 1998

[5] HIS Administrator Guide; Hyperwave Information Server, Version 5.5. Document Version HISV5.5-1.0; Hyperwave AG, München, Deutschland, 2001

[6] HIS Programmer´s Guide; Hyperwave Information Server, Version 5.5. Document Version HIS-V5.5-1.0; Hyperwave AG, München, Deutschland, 2001;

[7] Hyperwave Document Management Solutions; Hyperwave AG, München, Deutschland, 2002

[8] Liebowitz J.; Expert systems: An integral part of knowledge management; Kybernetes, 27(2) p.p.170–175. 1998

[9] R. Maier, F. Lehner, Perspectives on Knowledge Management Systems Theoretical Framework and Design of an Empirical Study; Proceedings of 8th European Conference on Information Systems, Vienna University of Economics and Business Administration, Österreich, 2000

[10] S. Mullender; Distributed Systems, Second Edition; ACM Press, New York, Addison Wesley, Kapitel 1-3,7,8. 1993

[11] D.E. O'Leary, Enterprise knowledge management; Computer, 31(3) p.p. 54–61. März 1998

[12] D.A. Stevenson, Intellectual Capital & EA (Enterprise Architecture); Extract from a paper presented to the Department of Information Systems, University of Cape Town, 1995

[13] I. Tuomi; Data is more than Knowledge: Implications of the reversed knowledge hierarchy for knowledge management and organizational memory; Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences, HICSS-32, Los Alamitos, Ca, USA, IEEE Computer Society, 1999

[14] K.M. Wiig Perspectives on introducing enterprise knowledge management; In PAKM 98. Practical Aspects of Knowledge Management. Proceedings of the Second International Conference, Zurich, Switzerland, 1998

Радіовимірювальний Перетворювач Концентрації Газу на Транзисторній Структурі з Від'ємним Опором

О.В. Осадчук, В.С. Осадчук, Я.О. Осадчук
кафедра радіотехніки
Вінницький національний технічний університет
Вінниця, Україна
osadchuk69@mail.ru

Radiomeasuring Transducer of Gas Concentration in Transistor Structure with Negative Resistance

A. Osadchuk, V. Osadchuk, I. Osadchuk
Department of Radioengineering
Vinnytsia National Technical University
Vinnytsia, Ukraine
osadchuk69@mail.ru

***Анотація* — В статті розглянуто радіовимірювальний перетворювач концентрації газу з чутливим до концентрації горючих газів MEMS резистивним елементом MiCS 6814. Розраховано та експериментально підтверджено залежності активної та реактивної складових повного опору радіовимірювального перетворювача концентрації газу від напруги живлення та керування і від зміни концентрації газу. Отримано аналітичні вирази функції перетворення і рівняння чутливості. Чутливість розробленого перетворювача в діапазоні від 1 до 1500 ppm змінюється від 645 до 175 Гц/ppm, а при збільшенні концентрації пропану чутливість значно зменшується, і в діапазоні від 1500 ppm до 9000 ppm складає від 175 Гц/ppm до 48 Гц/ppm.**

***Abstract* — In the article the radiomeasuring transducers of the gas concentration sensing the concentration of combustible gases resistive MEMS element MiCS 6814. Calculated and experimentally confirmed depending active and reactive components of the impedance transducers gas concentration of supply voltage and control and on changes in gas concentration. Analytical expressions conversion function and sensitivity equation. Sensitivity converter developed a range of 1 to 1500 ppm ranges from 645 to 175 Hz/ppm, while increasing the concentration of propane is significantly reduced sensitivity and range from 1500 ppm to 9000 ppm is from 175 Hz/ppm to 48 Hz/ppm.**

Ключові слова—радіовимірювальний перетворювач концентрації газу; MEMS сенсор; частотний вихідний сигнал; від'ємний опір.

Keywords—radiomeasuring transducer gas concentration; MEMS sensor; frequency output signal; negative resistance.

I. Вступ

Одним із нових наукових напрямків створення радіовимірювальних перетворювачів концентрації газу є використання реактивних властивостей напівпровідникових елементів з від'ємним опором. Це дозволяє перетворювати концентрацію газу в частотний вихідний сигнал, що забезпечує високу завадостійкість, а отже, і високу точність вимірювання концентрації газу. Окрім того, радіовимірювальні перетворювачі з частотним вихідним сигналом поєднують як простоту, так і універсальність, які властиві аналоговим пристроям, а також точність та завадостійкість, що характерні для перетворювачів з кодовим виходом. Вони володіють високою чутливістю до вимірюваних параметрів, малою масою, габаритами, інформаційною, конструктивною і технологічною сумісністю з мікроелектронними засобами обробки інформації, що забезпечує їх перевагу над існуючими сенсорами концентрації газу [1–3].

З метою вивчення властивостей радіовимірювального перетворювача концентрації газу необхідно розробити математичну модель, яка описує залежність частоти генерації від концентрації газу, тобто є функцією перетворення, на основі якої отримають чутливість перетворювача. В свою чергу функція перетворення залежить від активної і реактивної складових повного вихідного опору перетворювача. Розгляду цих питань присвячена дана робота.

II. Математична Модель Перетворювача

На рис.1 показана схема радіовимірювального перетворювача концентрації газу з чутливим до концентрації горючих газів MEMS резистивним елементом MiCS 6814, що дозволяє створити автогенераторний пристрій. Живлення транзисторів VT1 і VT2 здійснюється джерелами постійної напруги U1 і U2.

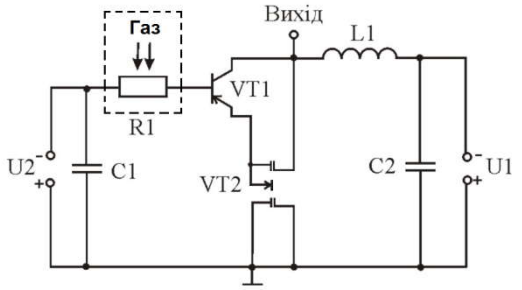


Рис. 1. Електрична схема радіовимірювального перетворювача концентрації газу

Математична модель описує залежність активної і реактивної складових повного вихідного опору структури від концентрації газу, що дозволяє отримати аналітичний вираз функції перетворення та рівняння чутливості перетворювача. Розрахунки моделі виконано на основі еквівалентних схем для змінного струму біполярного та МДН-транзистора, які складають радіовимірювальний перетворювач концентрації газу (рис. 2).

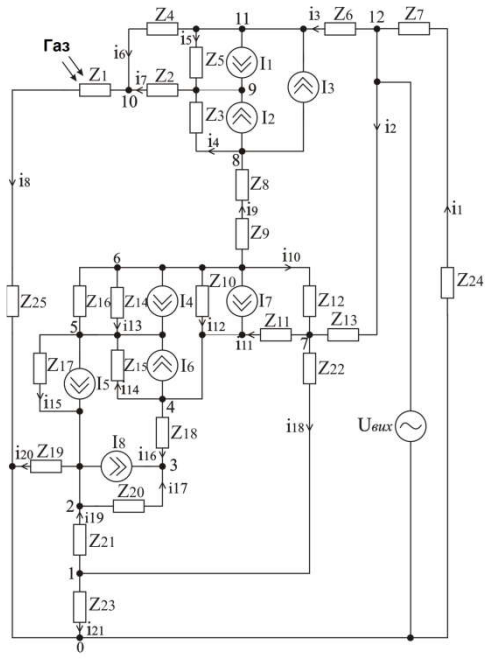


Рис. 2. Еквівалентна схема перетворювача концентрації газу для змінного струму

Вольт-амперна характеристика транзисторної структури, на основі якої створено радіовимірювальний перетворювача концентрації газу, має ділянку від'ємного опору. Від'ємний опір дозволяє компенсувати втрати в коливальному контурі, який утворений еквівалентною ємністю на електродах колектор-стік біполярного транзистора VT1 та МДН-транзистора VT2 та зовнішньою індуктивністю [3-6]. Для визначення складових повного вихідного опору перетворювача необхідно вирішити систему рівнянь Кірхгофа, складено на основі еквівалентної схеми (див. рис. 2), прийнявши вузол 0 як базисний, що має вигляд

$$\begin{cases} 0 = -\phi_1(y_{18}+y_{19}-y_{21})-\phi_2 y_{19}+\phi_7 y_{18}; \\ I_8 - I_5 = \phi_1 y_{19}-\phi_2(y_{19}+y_{20}+y_{17}+y_{15})+\phi_3 y_{17}+\phi_5 y_{15}; \\ -I_8 = \phi_2 y_{17}+\phi_3(y_{16}-y_{17})+\phi_4 y_{16}; \\ I_6 - I_1 = \phi_2 y_{16}-\phi_4(y_{16}+y_{14}+y_{12}+y_{11})+\phi_5 y_{14}+\phi_6 y_{12}+\phi_7 y_{11}; \\ I_5 - I_4 - I_6 = \phi_2 y_{15}+\phi_4 y_{14}-\phi_5(y_{15}-y_{13}-y_{14})+\phi_6 y_{13}; \\ I_4 + I_8 = \phi_4 y_{12}+\phi_5 y_{13}-\phi_6(y_{13}+y_{12}+y_{10}+y_9)+\phi_7 y_{13}+\phi_8 y_9; \\ 0 = \phi_1 y_{18}+\phi_4 y_{11}+\phi_6 y_{13}-\phi_7(y_{10}+y_{11}+y_2+y_{18})+U_{вих} y_2 \\ I_2 + I_3 = \phi_6 y_9-\phi_8(y_9+y_4)+\phi_9 y_4; \\ -(I_2+I_1) = \phi_8 y_4-\phi_9(y_4+y_7+y_5)+\phi_{11} y_5+\phi_{10} y_7; \\ 0 = \phi_9 y_7+\phi_{11} y_6-\phi_{10}(y_8+y_7+y_6); \\ I_1 - I_3 - U_{вих} y_3 = \phi_9 y_5-\phi_{11}(y_6+y_5+y_3)+\phi_{10} y_6; \\ U_{вих}(y_3+y_2+y_1) = \phi_7 y_2+\phi_{11} y_3, \end{cases} \quad (1)$$

де провідності гілок схеми визначаються за формулами

$$y_1 = 1/(Z_{24}+Z_7);\ y_2 = 1/Z_{13};\ y_3 = 1/Z_6;\ y_4 = 1/Z_3;$$
$$y_5 = 1/Z_5;\ y_6 = 1/Z_4;\ y_7 = 1/Z_2;\ y_8 = 1/(Z_{25}+Z_1);$$
$$y_9 = 1/(Z_8+Z_9);\ y_{10} = 1/Z_{12};\ y_{11} = 1/Z_{11};\ y_{12} = 1/Z_{10};$$
$$y_{13} = (Z_{16}+Z_{14})/(Z_{16}Z_{14});\ y_{14} = 1/Z_{15};\ y_{15} = 1/Z_{17};$$
$$Y_{16} = 1/Z_{18};\ y_{17} = 1/Z_{20};\ y_{18} = 1/Z_{22};\ y_{19} = 1/Z_{21};$$
$$y_{20} = 1/Z_{19};\ y_{21} = 1/Z_{23},$$

де $Z_1 = \dfrac{R_1}{1+\omega^2 R_1^2 C_3^2} - j\dfrac{R_1^2\omega C_3}{1+\omega^2 R_1^2 C_3^2}$; $Z_2 = R_2$; $Z_3 = -\dfrac{j}{\omega C_6}$;

$Z_4 = -\dfrac{j}{\omega C_4}$; $Z_5 = -\dfrac{j}{\omega C_5}$; $Z_6 = R_3$; $Z_7 = j\omega L_1$; $Z_8 = R_5$;

$Z_9 = R_6$; $Z_{10} = R_9$; $Z_{11} = -\dfrac{j}{\omega C_{11}}$;

$Z_{12} = \dfrac{R_7}{1+\omega^2 R_7^2 C_{12}^2} - j\dfrac{R_7^2\omega C_{12}}{1+\omega^2 R_7^2 C_{12}^2}$; $Z_{13} = R_8$; $Z_{14} = -\dfrac{j}{\omega C_7}$;

$Z_{15} = -\dfrac{j}{\omega C_8}$; $Z_{16} = R_{10}$; $Z_{17} = -\dfrac{j}{\omega C_9}$; $Z_{18} = R_{11}$; $Z_{19} = R_{15}$;

$Z_{20} = R_{12}$; $Z_{21} = \dfrac{R_{13}}{1+\omega^2 R_{13}^2 C_{10}^2} - j\dfrac{R_{13}^2\omega C_{10}}{1+\omega^2 R_{13}^2 C_{10}^2}$; $Z_{22} = -\dfrac{j}{\omega C_{13}}$;

$Z_{23} = R_{14}$; $Z_{24} = -\dfrac{j}{\omega C_2}$; $Z_{25} = -\dfrac{j}{\omega C_1}$,

де R_1 – опір чутливого до концентрації горючих газів MEMS резистивного елемента MiCS 6814; R_2, R_3, R_5 – об'ємний опір бази, колектора і емітера біполярного

транзистора VT1 відповідно; R_6, R_{15}, R_8 та R_{14} – об'ємні опори витоку, стоку та першого і другого затворів двозатворного МДН-транзистора VT2 відповідно; R_7 – об'ємний опір затвор-витік МДН-транзистора VT2; R_8 – об'ємний опір затвору МДН-транзистора VT2; R_9, R_{11} та R_{12} – об'ємні опори стік-витік МДН-транзистора VT2 відповідно; R_{10} – опір підкладки МДН-транзистора VT2; R_{13} – опір затвор-стік МДН-транзистора VT2; R_{15} – опір p-n переходу стоку МДН-транзистора VT2; C_1 – ємність конденсатора C_1; C_2 – ємність конденсатора C_2; C_3 – ємність чутливого до концентрації горючих газів MEMS резистивного елемента MiCS 6814; C_4 – ємність між зовнішнім виводом бази і колектора біполярного транзистора VT1; C_5, C_6 – ємності переходу база-колектор та база-емітер транзистора VT1 відповідно; C_7 – ємність підкладка-витік МДН-транзистора VT2; C_8 та C_9 – ємності підкладка-стік МДН-транзистора VT2 відповідно; C_{10} та C_{12} – ємності затвор-стік МДН-транзистора VT2 відповідно; C_{11} – ємність затвор-стік МДН-транзистора VT2; C_{13} – ємність між першим і другим затвором МДН-транзистора VT2; L_1 – зовнішня індуктивність.

На основі системи рівнянь (1) за допомогою програмного пакету Matlab 7.1. розрахована активна та реактивна складові повного опору, теоретичні та експериментальні залежності яких від концентрації газу показані на рис. 3 та 4.

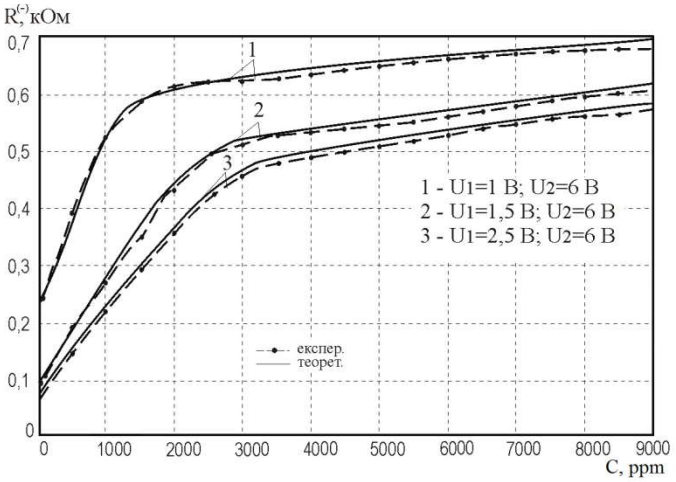


Рис. 3. Теоретичні та експериментальні залежності активної складової повного опору від концентрації газу

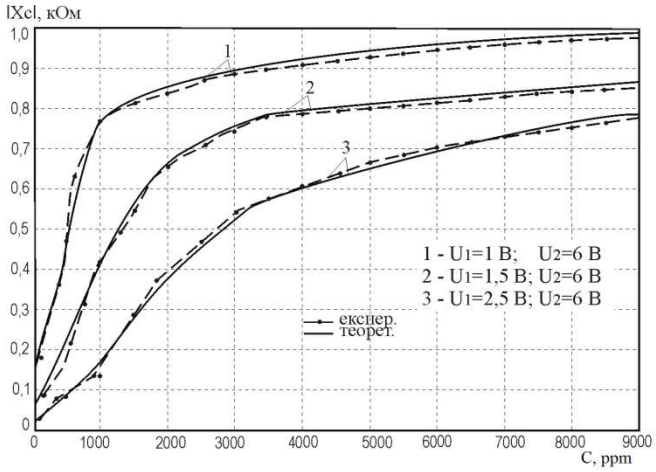


Рис. 4. Теоретичні та експериментальні залежності реактивної складової повного опору від концентрації газу

З рис. 3 та рис. 4 видно, що із збільшенням концентрації газу активна і реактивна складові повного опору збільшуються.

На рис. 5 подані експериментальні та теоретичні залежності частоти генерації перетворювача концентрації газу від зміни концентрації пропану (C_3H_8).

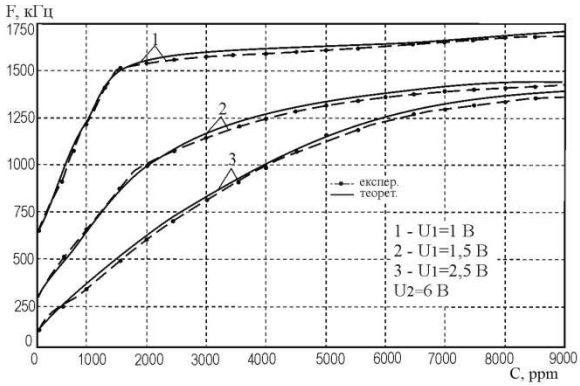


Рис. 5. Теоретичні та експериментальні залежності частоти генерації радіовимірювального перетворювача концентрації газу від зміни концентрації пропану (C_3H_8)

Як видно з рис. 5, частота генерації збільшується із збільшенням концентрації газу, а на ділянці від 0 до 2000 ppm спостерігається більша залежність частоти генерації від зміни концентрації пропану.

Залежність частоти генерації від концентрації газу, тобто функція перетворення визначається по контуру зворотного струму згідно з еквівалентною схемою (див. рис. 2) на основі теорії стійкості Ляпунова. Функція перетворення радіовимірювального перетворювача описується виразом (2)

$$F = \frac{1}{2}\frac{\sqrt{2}\sqrt{L_1C_4(-L_1C_4 + R_1^2(C)C_3^2 + R_1^2(C)C_3C_4 + A)}}{L_1C_3C_4R_1(C)}, \quad (2)$$

де $A = sqrt(L_1^2C_4^2 + 2L_1C_3^2C_4R_1^2(C) - 2L_1C_4^2C_3R_1^2(C) + +R_1^4(C)C_3^4 + D + R_1^4(C)C_3^2C_4^2;$

$$B = 2L_1C_4^2C_3R_1^2(C)\,; \qquad D = 2R_1^4(C)C_3^3C_4\,;$$

C – концентрація газу (ppm).

На основі (2) визначена чутливість генерації радіовимірювального перетворювача концентрації газу з чутливим до концентрації горючих газів MEMS резистивним елементом MiCS 6814:

$$S_E^F = \frac{1}{4}\sqrt{2}\left(2R_1(C)C_3^2\left(\frac{\partial R_1(C)}{\partial C}\right)+2R_1(C)C_3C_4\left(\frac{\partial R_1(C)}{\partial C}\right)+\right.$$
$$+\left(\frac{1}{2}\left(4L_1R_1(C)C_4C_3^2\left(\frac{\partial R_1(C)}{\partial C}\right)-4L_1R_1(E)C_4^2C_3\left(\frac{\partial R_1(C)}{\partial C}\right)+\right.\right.$$
$$+4R_1^3(C)C_3^4\left(\frac{\partial R_1(C)}{\partial C}\right)+8R_1^3(E)C_3^3C_4\left(\frac{\partial R_1(C)}{\partial C}\right)+4R_1^3(C)\times$$
$$\left.\left.\times C_3^2C_4^2\left(\frac{\partial R_1(C)}{\partial C}\right)\right)\Big/\sqrt{D_1}\right)\Big/\left(\sqrt{-L_1C_4(D_2+\sqrt{D_1})}\right)\Big)-\frac{1}{2}\sqrt{2}\times$$
$$\left.\times\sqrt{L_1C_4\left(D_2+\sqrt{D_1}\right)}\left(\frac{\partial R_1(C)}{\partial C}\right)\Big/\left(L_1C_4C_3R_1^2(C)\right)\right), \qquad (3)$$

де

$$D_1 = L_1^2C_4^2 + 2L_1C_4C_3^2R_1^2(C) - 2L_1C_4^2C_3R_1^2(C) + R_1^4(C)C_3^4 +$$
$$+2R_1^4(C)C_3^3C_4 + R_1^4(C)C_3^2C_4^2;$$
$$D_2 = -L_1C_4 + R_1^2(C)C_3^2 + R_1^2(C)C_4C_3\,.$$

Графік залежності чутливості від величини освітленості показаний на рис. 6.

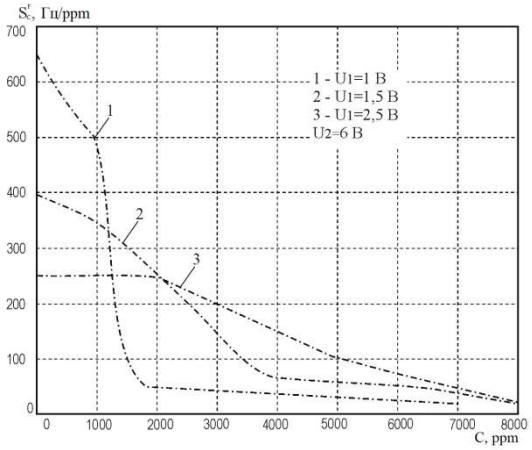


Рис. 6. Залежність чутливості радіовимірювального перетворювача від концентрації газу

З рис. 6 видно, що максимальну чутливість радіовимірювальний перетворювач концентрації газу з чутливим до концентрації горючих газів MEMS резистивним елементом MiCS 6814 має при напрузі живлення 1 В та напрузі керування 6 В. Із зміною концентрації газу від 1 до 1500 ppm чутливість змінюється від 645 до 175 Гц/ppm. При збільшенні концентрації пропану чутливість значно зменшується, і в діапазоні від 1500 ppm до 9000 ppm складає від 175 Гц/ppm до 48 Гц/ppm.

Визначимо адекватність розробленої моделі в порівнянні з експериментом [8] за формулою

$$\delta_{\text{м}} = \frac{x_M - x_e}{x_e}\cdot 100\,\%\,, \qquad (4)$$

де x_M – поточне значення параметра моделі; x_e – поточне експериментальне значення параметра.

Оптимальною напругою керування є величина 3 В, при якій існує найменша зміна частоти генерації в діапазоні від 20 °С до 80 °С.

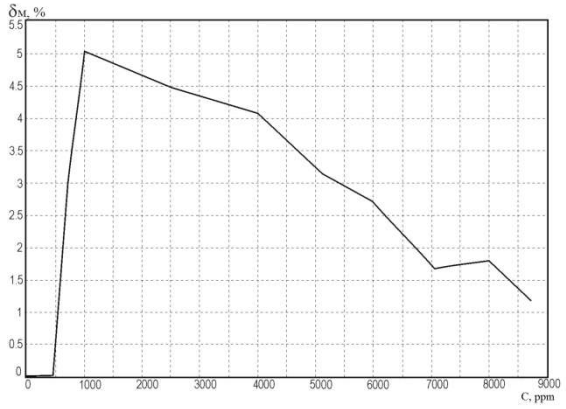


Рис. 7. Залежність відхилення теоретичної моделі від експериментальних значень від концентрації газу

III. Висновки

Запропоновано схему перетворювача концентрації газу в частоту на основі автогенераторного пристрою. Представлено дослідження параметрів радіовимірювального перетворювача концентрації газу з чутливим до концентрації горючих газів MEMS резистивним елементом MiCS 6814. Отримано аналітичні вирази функції перетворення і рівняння чутливості. Чутливість розробленого перетворювача в діапазоні від 1 до 1500 ppm змінюється від 645 до 175 Гц/ppm, а при збільшенні концентрації пропану чутливість значно зменшується, і в діапазоні від 1500 ppm до 9000 ppm складає від 175 Гц/ppm до 48 Гц/ppm.

Література References

[1] Мікроелектронні сенсори фізичних величин. За редакцією З.Ю. Готри. В 3 томах. Львів: Ліга-Прес, 2003. Т.2.

[2] Новицкий П.В., Кноринг В.Г., Гутников В.С. Цифровые приборы с частотными датчиками. Ленинград: Енергия, 1970.

[3] Осадчук О.В. Мікроелектронні частотні перетворювачі на основі транзисторних структур з від'ємним опором. Вінниця: Універсум-Вінниця, 2000.

[4] Осадчук В.С., Осадчук О.В. Реактивні властивості транзисторів і транзисторних схем. Вінниця: Універсум-Вінниця, 1999.

[5] Осадчук В.С., Осадчук О.В., Прокопова М.О. Сенсори газу. Вінниця: Універсум-Вінниця, 2008.

[6] Пат. 48600 Україна, МКІ G 01 N 27/12. Мікроелектронний давач газу/ В.С. Осадчук, О.В. Осадчук (Україна). № 2001107207; Заявлено 23.10.2001; Опубл. 15.08.2002, Бюл. № 8.

[7] Каяцкас А.А. Основы радиоэлектроники. Москва: Высшая школа, 1988.

[8] Мирский Г. Я. Электронные измерения. Москва: Радио и связь, 1986.

Застосування Частотного Критерію Інформативності Ознак в Задачах Інтелектуального Аналізу Тексту Багаторівневого Інформаційного Моніторингу

М.С. Голуб
кафедра інформаційної безпеки та комп'ютерної інженерії
Черкаський державний технологічний університет
Черкаси, Україна
Mas-golub@yandex.ru

The use of the Frequency Criterion Informative Signs in the Tasks Text Mining Multilevel Monitoring of Information

M. Holub
Department of Information Security and Computer Engineering
Cherkasy State Technological University
Cherkasy, Ukraine
Mas-golub@yandex.ru

***Анотація* — Перетворення друкованого тексту в масив чисельних характеристик дозволяє консолідувати інформацію технологією багаторівневого моніторингу. Виявлені особливості перетворення текстової інформації до форми масиву вхідних даних інформаційної системи багаторівневого моніторингу. Запропоновано новий частотний критерій інформативності ознак текстів. Розроблено метод перетворення тексту, за якого частотний критерій інформативності не залежить від розміру вікна. Експериментально доведено, що цей критерій селективний до різних задач класифікації текстів. Випробування критерію інформативності однак проведені в процесі розв'язання задач визначення місця проживання автора друкованого тесту. Експериментально підтверджено можливість та доцільність застосування описаного методу перетворення тексту в технології багаторівневого інформаційного моніторингу.**

***Abstract* — Desribed the information technology of intelligent search information technology systems. To form the deciding rules proposed use multilevel modeling techniques. Experimentally confirmed the possibility and feasibility of the technology to automate the process of finding a given text content. Offered the new frequency criterion informative signs of texts. Created a method of converting text which provided the frequency informative criterion, that does not depend on the size of the window. Experimentally proved that this criterion is selective for different classification text's problems. Criterion is informative but proved by the process of solving problems of determining residence labels of printed text. Experimentally confirmed the possibility and feasibility of the described text-tiered information technology monitoring method.**

Ключові слова—Перетворення тексту, критерій інформативності ознак, масив вхідних даних, інформаційний моніторинг, консолідація інформації, інтелектуальний аналіз тексту, багаторівневий моніторинг

Keywords—Search for text content, multilevel modeling, system engineering, information monitoring, consolidation information, text mining

I. Вступ

Процеси пошуку інформації заданого змісту, пошуку текстів, авторами яких є задані категорії осіб, виявлення текстів, що несуть в собі ворожий для України зміст та виконання багатьох інших завдань є функціями технологій інформаційного моніторингу [**3**].

Інтелектуальний аналіз друкованих текстових повідомлень є дієвим засобом забезпечення інформацією процесів прийняття рішень (ППР). З метою розширення можливостей технологій моніторингу, зокрема інформаційного моніторингу, розв'язуються традиційні завдання інтелектуального аналізу даних – класифікація, структурна та параметрична ідентифікація , прогнозування та інші. Перелік цих завдань та їх поєднання визначається потребами ППР, цілями, що досягає особа, що приймає рішення (ОПР), предметною галуззю, де використовується інформація, яка тримана із друкованих текстовий повідомлень. Найчастіше метою класифікації текстових повідомлень є їх групування за авторством, певними характеристиками стану автора (вік, стать, освіченість, фізичне та психологічне здоров'я, належність до певної спільноти).

Актуальність досліджень, пов'язаних із класифікацією текстових повідомлень та визначенням характеристик їх авторів визначається потребою у протидії методам та засобам інформаційної війни, яка зараз ведеться проти України. Технології інформаційного моніторингу є потужними інструментами, що здатні розв'язувати подібні задачі. Значна популярність моніторингових інформаційних систем (МІС) пов'язана із задоволенням інформаційних потреб військових, бізнесу, народного господарства, медицини, екології та інших галузей.

Інформаційна технологія багаторівневого моніторингу застосовується у випадках, коли необхідно забезпечити процеси прийняття рішень інформацією, яку треба здобувати «по крихтах» із багатьох різнорідних джерел [**6**]. У такому разі застосовується декомпозиція складних завдань до більш простих. Глибина декомпозиції визначає кількість рівнів перетворення інформації і зумовлена потужністю синтезатора. Формується ієрархія локальних завдань із перетворення даних [**3**].

Розв'язання кожного з них отримують у результаті синтезу багатопараметричних моделей. Ієрархічне поєднання цих моделей утворює структуру ГФЗ. На рис. 1 подана структура формування ГФЗ за технологією багаторівневого інформаційного моніторингу [**5**].

На мікрорівні моніторингу відбувається перетворення файлів із різнорідною інформацією від початкової форми тексту, відео- чи аудіо- файлів до форми масиву чисельних характеристик X. На макрорівні синтезуються моделі-класифікатори Y та відбувається їх випробування. На метарівні розробляються процедури використання цих моделей для групування вхідної інформації по класам Z, оцінюється впливовість факторів W.

Однією із складових цієї технології є методи інтелектуального аналізу текстів (Text Mining).

На сьогодні існує кілька класів задач, що розв'язуються методами інтелектуального аналізу текстів. Класифікація і кластеризація текстів [**9**] використовується для здобування ключових слів, реферування, тематичне індексування.

Поєднання в єдину технологію методів класифікації та структурно-параметричної ідентифікації дозволяє розв'язати задачі виявлення та аналізу зв'язків між поняттями, пошуку ключових фраз для навігації по текстам.

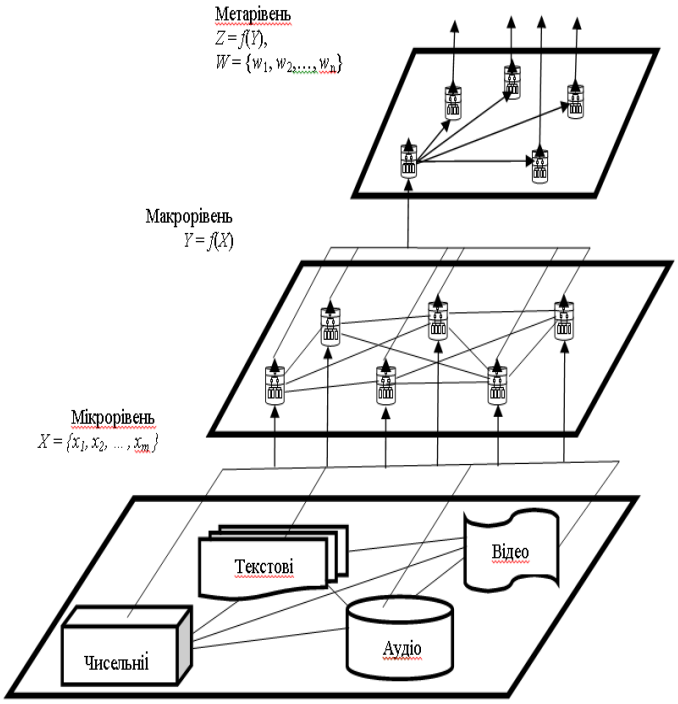


Рис. 1. Структура формування глобальної функціональної залежності системи багаторівневого інформаційного моніторингу

Використання інформаційної технології багаторівневого моніторингу для інтелектуального аналізу текстів (Text mining) дозволяє надати текстовій інформації особливої форми. Вона стає придатною для консолідації із іншими результатами моніторингу та використовується в бізнес-аналітиці, в процесі підтримки прийняття рішень та інше.

В якості характеристик тексту при автоматизованому аналізі використовуються його частотні показники: використання певних частин мови, деяких слів, розділових знаків, фразеологізмів, архаїзмів, слів, які рідко трапляються в тексті, слів іншомовного походження, довжини речення (в словах, складах, знаках) і таке інше [**1**]

II. Постановка задачі досліджень

В цій роботі досліджується застосування частотного критерію інформативності ознак в процесі розв'язання задачі класифікації тексту.

У випадку Text Mining математична постановка завдання має такий вигляд [**7**].

Нехай відомий початковий перелік текстів, що утворюють множину T:

$$T = f(t_1, t_2, \ldots, t_m) \quad (2.1)$$

і перелік властивостей їх авторів, що утворюють множину класів Z:

$$Z = f(z_1, z_2, \ldots, z_n) \quad (2.2)$$

Яка властивість автора відображена в якому тексті відомо для обмеженої кількості елементів навчальної підмножини T^n:

$$T^n = \{(t_1, z_1), (t_2, z_2), \ldots, (t_n, z_n)\} \quad (2.3)$$

Існує невідома цільова залежність – відображення

$$z^* : T \rightarrow Z \quad (2.4)$$

значення якої відоме на елементах підмножини Tn. Необхідно побудувати модель

$$a : T \rightarrow Z, \quad (2.5)$$

що здатна вірно класифікувати невідомий текст із підмножини $\{t_{n+1}, t_{n+2}, \ldots, t_m\} \in T$, тобто вірно визначити властивості автора цього тексту.

Виявляється залежність значення цього критерію від розміру вікна, відповідно до якого обраховуються числові характеристики вектору ознак МВД.

III. Дослідження частотного критерію інформативності ознак

Значна кількість завдань Text mining виконується МІС шляхом розв'язання типової задачі класифікації точок спостереження. Точка спостереження може поєднувати від одного до кількох вікон. Інколи вона поєднує всі вікна тексту, що досліджується.

Тексти перетворюються до типової форми масиву вхідних даних (МВД) – двовимірної таблиці чисельних показників, що заповнена за певними правилами. Тексти розбиваються на ділянки, що звуться вікнами. Для кожного із вікон розраховуються чисельні характеристики вектору ознак текста. Множина показників, що входять до цього вектору утворюють словник ознак.

Кожна ознака є носієм інформації – відомостей про властивості тексту. Словники ознак визначають інформативність МВД. Тому дослідження, що дозволяють підвищити інформативність МВД, є актуальними.

Наперед відомо, що існує мінімальна кількість інформації в МВД, яка дозволяє існуючими методами та засобами створити корисну модель. Ця мінімальна кількість інформації позначається як межа інформаційної достатності (МІД).

Корисність моделі визначається її здатністю розв'язувати задачі Text mining, зокрема забезпечити класифікацію текстів за заданими класами.

Для забезпечення достатньої інформативності МВД необхідно оцінити інформативність кожного із ознак тексту.

Для оцінки інформативності ознак текстів запропоновано частотний критерій (3.1):

$$= \frac{}{\sum} \quad (3.1)$$

де K_i – критерій інформативності i-ї ознаки творів автора, що ідентифікується, – частота використання і-ї ознаки, n – кількість ознак первинного опису

$$= — \quad (3.2)$$

де N_i – частість використання i-ї ознаки у вікні; R – розмір вікна.

На підставі аналізу виразу (3,1) була сформульована гіпотеза: частотний критерій інформативності ознак (ЧКІО) критерій не повинен залежати від розміру вікна R, оскільки частота використання тексту присутня і в чисельнику, і в знаменнику. Це може стати однією із вагомих переваг цього критерію інформативності перед дисперсним критерієм.

Для перевірки цієї гіпотези був проведений модельний експеримент.

Досліджувались особливості формування МВД [**8**] та процесу синтезу багатопараметричних моделей [**4**], здатних класифікувати текстові повідомлення за їх належністю до різних типів говірок, що притаманні населенню центральної, північної, південної, західної та східної частин Середньої Наддніпрянщини [**7**].

Розв'язувалась задача класифікації тестів за говірками на території Черкаської області, які використовуються в селах Макіївка та Гарбузин. Підґрунтям дослідження стали діалектні тексти, наведені у збірнику [**2**].

Основною вимогою до МВД є його інформативність. Він повинен містити достатньо інформації для того, щоб побудувати існуючими засобами синтезатора МІС корисну модель а, яка забезпечить необхідні перетворення відповідно до вимоги (6). До ПО заносяться результати першого етапу моніторингу текстового повідомлення – чисельні характеристики інформативних ознак тектсів.

Нижче запропоновано новий метод інтелектуального аналізу тексту з метою виявлення властивостей їх авторів: 1) визначення обсягу вікна; 2) визначення переліку ознак тексту; 3) визначення переліку чисельних характеристик ознак тексту; 4) розрахунок чисельних значень характеристик ознак текстів; 5) задання чисельного значення критерію класу, що поділяє точки спостереження на класи «Свій» та «Чужий»; 6) формування масиву вхідних даних. 7) синтез моделі; 8) дослідження моделі. Визначення розділяючої поверхні. 9) випробування моделі; 10) оцінка впливовості факторів.

Одними із головних завдань, що визначає успішність Text Mining – це декомпозиція тексту і визначення переліку його інформативних характеристик – ознак, що дозволять синтезатору адекватно і точно відобразити властивості автора в моделі текстового повідомлення.

В процесі дослідження застосовувати декомпозицію тексту на окремі ділянки – вікна та на окремі елементи – букви, буквосполучення, речення та інші, які на наступних етапах використовуються для розрахунку значень показників. [4].

Досліджувались значення критерію інформативності (3.1) при зміні розміру вікна в напряму від найбільшого 5000 знаків, до найменшого – 50 знаків.

Тексти перетворювались в первинний опис об'єкта (ПО) досліджень шляхом адаптивного формування переліку ознак із МІД = 1. Таким чином формувався словник із повного переліку показників, що потенційно можуть бути інформативними, і які хоча б один раз зустрічались в тексті.

Після цього ПО піддавався перетворенню шляхом розрахунку значення критерію (3.1) для вікон різного розміру. Після заміни частотних характеристик тексту на значення частотного критерію інформативності проводилось їх усереднення Таким чином ПО перетворюється в МВД. На рис. 1 і 2 подані результати досліджень.

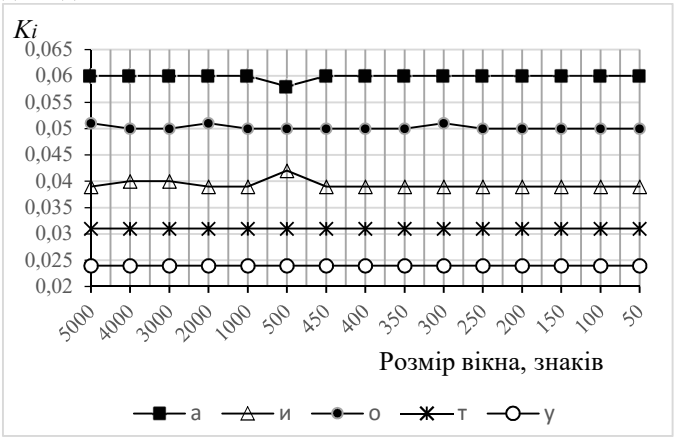


Рис. 2. Інформативність ознак текстів мешканців с. Гарбузин

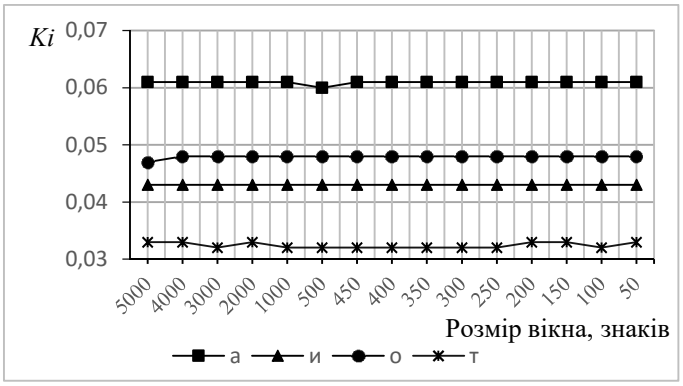


Рис. 3. Інформативність ознак текстів мешканців с. Макіївка

Значення частотного критерію інформативності (3.1) не залежить від розміру вікна після їх усереднення як для текстів с. Гарбузин, так і для текстів с. Макіївка. При цьому словник текстів с. Гарбузин на 1 інформативну ознаку більше в порівнянні із текстами с. Макіївка. Значення одних і тих же інформативних ознак для різних сіл майже завжди різняться, що доводить їх селективність.

IV. Висновки

Частотний критерій інформативності ознак друкованого тексту може бути використаний для перетворення первинного опису об'єкта в масив вхідних даних. При цьому експериментально підтверджено гіпотезу про незалежність значення критерію інформативності від розміру вікна за умови усереднення результатів. Доведено також, що ЧКІО селективний до класів. При аналізі інформативності ознак текстів повідомлень мешканців різних селищ значення ЧКІО різняться. Таким чином частотний критерій може бути використано в процесі розв'язанні задач інтелектуального аналізу текстів в технологіях багаторівневого інформаційного моніторингу.

Література References

[1] E.Y. Lynhvystycheskaya bezopasnost' rechevoy kommunykatsyy // HLƏDYS. URL: http://www.rusexpert.ru/maga¬zine/034.htm

[2] Hovirky Cherkashchyny: zbirnyk dialektnykh tekstiv / Uporyadnyky H.I. Martynova, T.V. Shcherbyna, A.A. Taran. – Cherkasy: PP Chabanenko Yu.A., 2013. – 870 s.

[3] S.V.Holub, Bahatorivneve modelyuvannya v tekhnolohiyakh monitorynhu otochuyuchoho seredovyshcha. Cherkasy: Vyd. vid. ChNU imeni Bohdana Khmel'nyts'koho, 2007. – 220 s.

[4] S.V. Holub, Vidobrazhennya vlastyvostey avtora tekstu v strukturi bahatoparametrychnoyi modeli / S.V. Holub, O.V. Konstantynovs'ka, M.S. Holub // Systemy obrobky informatsiyi: Zbirnyk naukovykh prats'. – Kh.: Kharkivs'kyy universytet povitryanykh syl imeni Ivana Kozheduba, 2014. – Vyp. 9 (125). – S. 82-87

[5] S.V. Holub, Vidobrazhennya konsolidovanoyi informatsiyi ekonomichnykh pokaznykiv rehionu u strukturi bahatorivnevykh modeley / S.V. Holub, N.O. Khymytsya // Visnyk Skhidnoukrayins'koho natsional'noho universytetu imeni Volodymyra Dalya. – 2012. – # 8 (179). – Chastyna 1. – S. 122-128.

[6] S.V.Holub, Konsolidatsiya modeley v protsesi bahatorivnevoho opratsyuvannya danykh / S.V. Holub Informatsiya, komunikatsiya, suspil'stvo 2014: materialy 3-yi Mizhnar. nauk. konferentsiyi ICS-2014. – L'viv: Vydavnytstvo L'vivs'koyi politekhniky, 2014. – S. 162-163.

[7] S.V. Holub, Modelyuvannya dialektnoho tekstu v tekhnolohiyi bahatorivnevoho informatsiynoho monitorynhu / S.V. Holub, H.I. Martynova, M.S. Holub // Matematychni mashyny i systemy. – 2016. – # 4. – S. 76-83

[8] S.V. Holub, Formuvannya pokaznykiv masyvu vkhidnykh danykh dlya identyfikatsiyi avtorstva tekstovykh povidomlen' / S.V. Holub, O.V. Konstantynovs'ka, M.S. Holub // Systemy obrobky informatsiyi: zb. nauk. prats'. – Kh.: Kharkivs'kyy universytet Povitryanykh syl imeni Ivana Kozheduba, 2014. – Vyp. 2 (118). – S. 89 – 92

[9] V.L. Pleskach, T.H. Zatonats'ka, Informatsiyni systemy i tekhnolohiyi na pidpryyemstvakh. K.: Znannya, 2011. – 718 s.

Структурно-Параметрична Ідентифікація Правил Бази Модельних Знань

Ірина Жирякова
кафедра інтелектуальних систем прийняття рішень
Черкаський національний університет
імені Богдана Хмельницького
Черкаси, Україна
irena_zh@ukr.net

Сергій Голуб
кафедра інтелектуальних систем прийняття рішень
Черкаський національний університет
імені Богдана Хмельницького
Черкаси, Україна
fpkpk@ukr.net

Structural and Parametric Identification of the Knowledge Model Base Rules

Iryna Zhyriakova
Department of intelligent decision support systems
Bohdan Khmelnytsky National University
at Cherkasy
Cherkasy, Ukraine
irena_zh@ukr.net

Serhii Golub
Department of intelligent decision support systems
Bohdan Khmelnytsky National University
at Cherkasy
Cherkasy, Ukraine
fpkpk@ukr.net

***Анотація*—Обґрунтовано застосування нового підходу до автоматичного формування правил бази модельних знань для інтелектуальних систем прийняття рішень. Запропонований підхід базується на використанні основних положень теорії ієрархічних багаторівневих систем та методу багаторівневого моделювання складних систем за даними спостережень в умовах неповноти інформації.**

***Abstract*—A new approach to automatic generation of the knowledge model base rules for intelligent decision support systems are proved. New approach is based on the main items of the hierarchical multilevel system theory and method multilevel modeling for complex systems.**

Ключові слова—база модельних знань, інтелектуальні системи прийняття рішень

Keywords—knowledge model base, intelligent decision support systems

I. Вступ

Основні напрями і тенденції автоматизації інтелектуальної діяльності людини на основі використання сучасних інформаційних технологій досить різноманітні. Та прогрес в цій галузі значною мірою залежить від успішності використання процедур моніторингу для ідентифікації стану об'єктів інтелектуальної діяльності та їх взаємодії в рамках поставленої задачі. Такі об'єкти, зазвичай, являють собою складні системи, для ідентифікації стану яких необхідно розв'язати комплекс задач в умовах невизначеності, адже, схильність до змін є невід'ємною характеристикою будь-якої предметної області.

Модель взаємодії складних систем в рамках будь-якої предметної області є основою процедурних знань (правил виводу) будь-якої інтелектуальної системи прийняття рішень. Отже, в даній роботі для ідентифікації правил виводу пропонується скористатись положеннями теорії ієрархічних багаторівневих систем [1], що дають можливість розглядати предметну область як декомпозицію етапів перетворення інформації, в результаті якої синтезуються багатопараметричні моделі, що можуть бути застосовані для представлення процедурної складової бази модельних знань. Переваги застосування ієрархічного підходу пов'язані зі зменшенням невизначеності стосовно як досліджуваних об'єктів так і середовища, в якому вони існують, на кожному етапі перетворення інформації.

II. Аналіз останніх досліджень і публікацій та постановка проблеми

Спробу повної систематизації алгоритмів і способів ідентифікації виконав один з основоположників теорії ідентифікації професор П. Ейкхоф [2]. Дана теорія є актуальною і по теперішній час при вивченні проблем побудови і всебічного аналізу моделей процесів або

систем. Вона описує поняття моделі, постановку задачі ідентифікації, базові підходи до побудови моделей для різних класів об'єктів, способи їх опису при різних підходах і алгоритмах ідентифікації.

Моделі подання знань почали розроблятися з 70-х років XX століття. Запропоновані підходи мали власну специфіку, яка найбільш помітна на синтаксичному рівні формального інструментарію для опису семантики. Застосування цих підходів досить широко використовувалось в діалогових системах типу «питання-відповідь» [3] та природно-мовних діалогових системах [4], а також в інтелектуальних розв'язувачах задач та системах управління [5-7]. Згодом проблематика даного напряму досліджень, яка розглядалась в рамках штучного інтелекту, отримала назву "системи подання знань" [8]. До найбільш відомих моделей подання знань відносяться: семантичні мережі, концептуальні мережі, ситуаційне управління, формальні логічні моделі подання знань, універсальний семантичний код, мова Ф, фреймові моделі подання знань, зростаючі пірамідальні мережі, індикативні структури та онтології для опису семантики.

Огляд та аналіз існуючих моделей подання знань дає підстави стверджувати, що їх можливості обмежені в рамках вирішення задач оптимізації затрат часу на розробку бази знань або зміну її структури (підтримку актуальності у відповідності до нових задач). Таким чином, існує протиріччя між необхідністю розв'язку комплексу задач, що стоять перед сучасними інтелектуальними системами та обмеженими можливостями науково-методичного апарату, що дозволяє формалізувати роботу зі знаннями в автоматичному режимі. Наявність цього протиріччя з одного боку та існуюча і досі практика по формуванню баз знань експертами вручну або з використанням редакторів знань типу Protégé – з іншого боку, робить дослідження нових підходів до роботи з знаннями, в тому числі ідентифікацію процедурних знань, актуальними.

III. Результати

Згідно з описаною в [9, 10] загальною концепцією розробки бази модельних знань (БМЗ), що використовує методологію перетворення інформації на основі глобальної функціональної залежності [11], де кожна модель об'єкта в свою чергу є ієрархічною структурою, що містить інші моделі цього об'єкта, отримані за завершеними алгоритмами [12] для індуктивної ідентифікації правил БМЗ на основі емпіричних даних (навчаючих виборок, які формуються на основі масиву вхідних даних (МВД)) необхідно виконати спочатку структурну ідентифікацію досліджуваних об'єктів системи, а потім застосувати алгоритми параметричної ідентифікації. Розглянемо цей процес більш докладно на прикладі інтелектуалізації моніторингу «розумного міста» (Smart Sity).

Основною метою управління Smart Sity в рамках будь-якого міста є підвищення ефективності моніторингу всіх міських служб шляхом застосування інформаційно-комунікаційних технологій, тим самим усуваючи надмірність в генерації і використанні інформації. Отже, в межах окремих ланок управління інформаційні моделі будуть являти собою локальні функціональні залежності (транспортна система, медицина, тощо), які поєднуються в єдину глобальну функціональну залежність в рамках обраного в якості об'єкта дослідження міста. Якщо ж розглядати подібну концепцію в масштабах регіону, то в процесі моделювання ми отримаємо вже цілий комплекс глобальних залежностей, які теж можуть бути поєднані в рамках системи управління країною. Отже, задача ідентифікації полягатиме у формуванні за даними моніторингу міських служб деякої множини моделей різної структури:

$$W = (X : Y), \tag{1}$$

де W – множина відображень показників стану модельованого об'єкту або змінні моделювання – вектор X, на модельований показник (вихідний сигнал), який потрібно визначити – Y.

Синтез будь-якого формального представлення моделі потребує рішення двох взаємозалежних задач: структурної ідентифікації і параметричної ідентифікації. Структурна ідентифікація передбачає визначення значущих факторів (змінних моделювання), що впливають на вихідні дані моделі (модельований показник); визначення структури, тобто виду оператора, що встановлює зв'язок між змінними моделювання і вихідними характеристиками моделі.

На першому етапі структурної ідентифікації аналізуються:

- змінні моделювання, які утворюють топологічний простір X розмірності $dim\, X = n \times m$, де n – кількість спостережень, де m – кількість змінних моделювання;
- модельований показник Y розмірності $dim Y = n \times 1$, де n – кількість спостережень;
- на основі топологічного простору формується система лінійно-незалежних векторів, що утворюватимуть матрицю МВД рангу $rank X = k, k \le m$.

В процесі формування МВД дуже важливе значення мають методи стійкого оцінювання, що враховують наявність відхилень в даних і далі на етапі параметричної ідентифікації дозволяють їх компенсувати і досить точно визначати оцінки параметрів. Такі методи, називаються робастними або стійкими. Основи робастних методів оцінки були розроблені А.Н. Колмогоровим, Н.В. Смірновим і Б.С. Ястремським, а подальший розвиток отримали в роботах американських і швейцарських математиків [13]. В даній роботі для оцінки МВД пропонується застосування L- та E- критеріїв, запропонованих Г. Трітьєном та І.Г. Муром.

L-критерій має два різновиди застосування: для виявлення грубих помилок у верхній частині L та у нижній частині L' ранжованого ряду даних:

$$L = \frac{\sum_{i=1}^{n-l}(x_i - \bar{x}_l)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}, L' = \frac{\sum_{i=l+1}^{n-l}(x_i - \bar{x}_l)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}, \quad (2)$$

де x_i – вибірка даних спостережень по одній обраній j-тій ознаці, n – обсяг вибірки, l – кількість спостережень з відхиленнями в значенні обраної ознаки, $\bar{x}$ – загальна середня для вибіркової сукупності, $\bar{x}_l$ – середнє значення для $(n-l)$ спостережень, що залишились після відкидання l грубих помилок у верхній для L (3) та у нижній частині – для L' (4) ранжованого ряду даних.

$$\bar{x}_l = \frac{\sum_{i=1}^{n-l} x_i}{n-l} \quad (3)$$

$$\bar{x}_l = \frac{\sum_{i=l+1}^{n-l} x_i}{n-l} \quad (4)$$

Е-критерій застосовується у випадку, коли у вибірці грубі помилки з найбільшими і найменшими значеннями розташовані як у верхній, так і у нижній частині ранжованого ряду даних:

$$E = \frac{\sum_{i=l+1}^{n-l}(x_i - \bar{x}_{l'})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}, \quad (5)$$

де $\bar{x}_l$ – середнє значення після відкидання найменших l і найбільших l' грубих помилок, яка обчислюється за формулою:

$$\bar{x}_{l'} = \frac{\sum_{i=l+1}^{n-l'} x_i}{n-(l+l')}. \quad (6)$$

Всі три критерії мають табличні критичні значення для заданого рівня значущості α при відомому обсязі вибірки n і передбачуваній кількості помилок l. Якщо значення критеріїв виявляються менше порогових, то помилки в даних, визнаються грубими. Інакше дані вважаються типовими для даної сукупності спостережень.

На другому етапі структурної ідентифікації для вибору класу моделей можна скористатись будь-якою з двох аксіоматичних структур: адитивною (7) або мультиплікативною (8), що визначають специфікацію моделі в математичній формі. Відповідний вибір буде цілком залежати від підсистеми управління для вибраного об'єкту моніторингу. Подібні аксіоматичні структури у вигляді поліномів і різнецевих рівнянь, які відносяться до класу моделей лінійних за параметрами, зазвичай використовують для моделювання статичних або динамічних об'єктів.

$$P(x) = \sum_{i=1}^{n} a_i x_i, \quad (7)$$

$$P(x) = \prod_{i=1}^{n} a_i x_i, \quad (8)$$

де a_i – відносні безрозмірні вагові коефіцієнти, для яких виконуються наступні обмеження:

$$0 \le a_i \le 1, \sum_{i=1}^{n} a_i = 1. \quad (9)$$

Недоліками адитивної структури є відсутність врахування наявності можливої нелінійності зв'язку. Цей недолік можна усунути при більш широких обмежуючих допущеннях про структуру моделі, наприклад, скориставшись поліномом Колмогорова-Габора:

$$P(x) = a_0 + \sum_{i=1}^{n} a_i x_i + \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} x_i x_j + \ldots \quad (10)$$

Такий підхід дозволяє описати будь-яку нелінійну залежність і не накладає ніяких апріорних обмежень на адитивність або мультиплікативність, тому що поліном (10) містить у своєму складі як перші, так і більш високі ступені змінних моделювання і всі можливі їх комбінації.

Отже, обравши відповідну аксіоматичні структуру для специфікації моделі задача її синтезу зводитиметься до параметричної ідентифікації вагових коефіцієнтів a_i.

Алгоритм синтезу моделі (АСМ), що передбачає визначення конкретних кількісних значень її параметрів, є найбільш формалізованим етапом і, як правило, зводиться до вирішення задачі оптимізації із мінімізації середньої квадратичної нев'язки для залежності з неперервним виходом чи рівня безпомилковості класифікації для залежності з дискретним виходом.

Для реалізації АСМ можна використати перебірний (комбінаторний) або ітераційний (багаторядний) метод генерації структури моделі (11). Застосування будь-якого з них, дасть можливість сформувати множину моделей-кандидатів (12), серед яких буде обрана оптимальна за визначеним критерієм якості (13):

$$\hat{Y}_f = f\left(X, \hat{\theta}_f\right), f \in \Phi, \quad (11)$$

де Φ – множина моделей-кандидатів, $\hat{\theta}_f$ – оцінка параметрів для кожної f.

$$f^* = arg\,min_{f\in\Phi}\, C\left(y, f\left(X, \hat{\theta}_f\right)\right), \quad (12)$$

де C – критерій якості моделей, f^* – оптимальна модель серед усіх $f \in \Phi$.

$$\hat{\theta}_f = arg\,min_{f\in R^{s_f}}\, Q\left(y, X, \theta_f\right), \quad (13)$$

де $Q \neq C$ – критерій якості оцінювання параметрів кожної окремої моделі при розв'язанні задачі параметричної ідентифікації, що генерується в процесі структурної ідентифікації.

Можлива різна композиція методів структурної і параметричної ідентифікації, яка буде відрізнятися лише точністю, трудомісткістю і універсальністю.

Відповідно до методу висхідного синтезу елементів, кожна глобальна функціональна залежність буде ієрархічним поєднанням локальних багатопараметричних моделей, згенерованих відповідним АСМ, та поєднаних в страти відповідно до рівнів моніторингу:

$$Z = f\left(Y\left(X\right)\right). \quad (14)$$

Така структура забезпечуватиме багаторівневе перетворення інформації від форми чисельних характеристик результатів спостереження до форми відомостей про властивості об'єктів, поданих в зручному для використання вигляді.

Кожна із страт матиме декілька недостатньо інформативних моделей, вихідні сигнали яких мають характеристики гірше заданих. Це пов'язано з тим, що задача параметричної ідентифікації є некоректною по Адамару, так як не має єдиного розв'язку. Згідно з теоремою Вейєрштрасса, будь-яку послідовність з n експериментальних даних можливо точно апроксимувати поліномом із $(n+1)$ членом. Однак, така апроксимація не означає, що синтезовано адекватну модель високої точності з гарними прогностичними властивостями. Отже, будь-який поліном високої складності апроксимує не тільки корисний сигнал, але й випадкові похибки експериментальних даних, які частково усуваються L- та E- критеріями.

Для подолання зазначеного недоліку в роботі [14] запропоновано розділяти вибірку експериментальних даних на дві підмножини: навчальну і перевірочну. Перша підмножина використовується для синтезу моделі і визначення її характеристик, а друга – для перевірки точності моделі. При цьому виявилося, що при збільшенні складності моделі точність апроксимації перевірочної послідовності експериментальних даних спочатку поліпшується, досягає деякого мінімуму і потім починає погіршуватися за рахунок обліку "шкідливих" випадкових складових. Модель мінімальної складності, що дає мінімум похибки апроксимації перевірочної послідовності, одержала назву моделі оптимальної складності.

Періодичне поновлення та доповнення БМЗ останніми результатами моніторингу та автоматичне адаптування структури БМЗ за описаними правилами дозволятиме врахувати динамічність процесів управління будь-якою ланкою Smart Sity та забезпечить інформацією процеси прийняття управлінських рішень на кожній з них.

IV. Висновки

У даній роботі запропоновано новий підхід до структурно-параметричної ідентифікації правил для баз модельних знань, який базується на наступних принципах: знання представляються у вигляді багатопараметричних моделей; зростання кількості знань досягається шляхом ієрархічного поєднання багатопараметричних моделей віповідно методу висхідного синтезу елементів в системах з багаторівневим перетворенням даних; використання положень теорії ієрархічних багаторівневих систем є основою для розв'язання задач структурної ідентифікації; консолідація знань відбувається шляхом поєднання глобальних функціональних залежностей, сформованих в рамках різнорідних предметних областей. Запропонований підхід орієнтований на автоматичне формування та подальший автоматизований супровід баз знань для будь-якої предметної області.

Література References

[1] М. Месарович, Теория иерархических многоуровневых систем / М. Месарович, Д. Мако, И. Тахакара. – М.: Мир, 1973. – 344 с.

[2] П. Эйкхофф, Основы идентификации систем управления. – М.: Мир, 1975. – 686 с.

[3] Г.Г. Белоногов, Б.А. Кузнецов, Языковые средства автоматизированных информационных систем. – М.: Наука, 1983. – 288 с.

[4] Э.В. Попов, Искусственный интеллект: Кн.1. Системы общения и экспертные системы: справочник / под ред. Э.В. Попова. – М.: Радио и связь, 1990. – 464 с.

[5] Н. Нильсон, Принципы искусственного интеллекта. – М.: Радио и связь, 1985. – 372 с.

[6] Д.А. Поспелов, Логико-лингвистические модели в системах управления. – М.: Энергоиздат, 1981. – 231 с.

[7] М. Минский, Фреймы для представления знаний. – М.: Энергия, 1979. – 151 с.

[8] Т.А. Гаврилова, В.Ф. Хорошевский, Базы знаний интеллектуальных систем. – СПБ.: Питер, 2001 – 384 с.

[9] І.А. Жирякова, Новий підхід до концептуалізації знань / І.А. Жирякова, С.В. Голуб // Технічні науки та технології. – 2015. – № 2. – С. 78-82.

[10] І.А. Жирякова, Удосконалення методу висхідного синтезу елементів в інформаційній технології багаторівневого моніторингу мобільного робота / І.А. Жирякова, С.В. Голуб, О.В. Харченко // Математичні машини і системи. – 2016. – № 3. – С. 41-47.

[11] С.В. Голуб, Багаторівневе моделювання в технологіях моніторингу оточуючого середовища. – Черкаси: Вид. від. ЧНУ імені Богдана Хмельницького, 2007. – 220 с.

[12] С.В. Голуб, Координація взаємодій локальних агрегатів в структурі систем багаторівневого перетворення моніторингової інформації // Вісник Східноукраїнського національного університету імені Володимира Даля. – 2009. – № 6(136). – Ч. 1. – С. 325-329.

[13] А.М. Шурыгин, Прикладная стохастика: робастность, оценивание, прогноз. – М.: Финансы и статистика, 2000. - 224 с.

[14] В.С. Степашко, Структурна ідентифікація моделей як задача відновлення сигналу в умовах неповноти інформації // Наукові праці ДНТУ. Серія: Обчисл. техніка та автоматизація. – Вип. 48. – Донецьк: ДНТУ, 2002. – С. 345-353.

Комп’ютерна Система Стереозору для Візуального Визначення Відстаней

Олег Прокопчук
кафедра Автоматизованих систем обробки інформації
Дніпровський національний університет імені Олеся Гончара
Дніпро, Україна
neues.system@ukr.net

Михайло Церр
кафедра Автоматизованих систем обробки інформації
Дніпровський національний університет імені Олеся Гончара
Дніпро, Україна
azimutkappa@gmail.com

Сергій Вовк
кафедра Автоматизованих систем обробки інформації
Дніпровський національний університет імені Олеся Гончара
Дніпро, Україна
vovk_s_m@ukr.net

Computer Stereo Vision System for Visual Distance Determination

Oleh Prokopchuk
dept. of Automated data processing systems
Oles Honchar Dnipro National University
Dnipro, Ukraine
neues.system@ukr.net

Mykhailo Tserr
dept. of Automated data processing systems
Oles Honchar Dnipro National University
Dnipro, Ukraine
azimutkappa@gmail.com

Sergiy Vovk
dept. of Automated data processing systems
Oles Honchar Dnipro National University
Dnipro, Ukraine
vovk_s_m@ukr.net

***Анотація*—Пропонується комп’ютерна система для візуального визначення відстаней до об’єктів, поміж об’єктами та на об’єктах, яка працює на базі програмної платформи Microsoft .NET Framework і використовує бібліотеки OpenCV та Emgu CV. Робота системи передбачає отримання зображень безпосередньо від камер стереопари, підключених до комп’ютера, або з файлів на диску. Визначення шуканих відстаней здійснюється або в автоматичному режимі через відповідні функції бібліотеки OpenCV, які пов’язані з розпізнаванням об’єктів за їх шаблонами, або в інтерактивному режимі шляхом вказування на маркерні точки та використання додатково розроблених функцій, які пов’язані зі стереоскопічними обчисленнями. Наводяться результати експериментів з визначення відстані до обличчя людини та з обчислення відстаней між окремими точками на пласкому об’єкті.**

***Abstract*—A computer system for visual determination of distances to the objects, distances between and at the objects is proposed. The system is based on the software platform of Microsoft .NET Framework and used the OpenCV library and the Emgu CV library. The proposed system provides for receiving images directly from the stereo cameras connected to the computer or the files on the disk. Determining distances is obtained either in automatic mode through the appropriate functions of the OpenCV library related to the searching of objects by their templates, or in interactive mode by using developed functions associated with the marked points at the objects. The experimental results of determining the distance to the human face and distances between the different points at the flat object are presented.**

Ключові слова—стереозір, автоматичний, OpenCV, Emgu CV.

Keywords—stereo vision; automatic; OpenCV; Emgu CV.

I. Вступ

Однією з головних задач комп'ютерного стереозору є аналіз просторових сцен з оцінюванням значень геометричних параметрів об'єктів, що спостерігаються. Розв'язок цієї задачі ґрунтується на принциповій можливості систем стереозору визначати тривимірні просторові координати об'єктів [1], що надає можливість обчислювати відстані до об'єктів та поміж об'єктами, лінійні розміри об'єктів, положення та орієнтацію пласких частин об'єктів у просторі тощо. Зважаючи на стрімкий розвиток апаратних складових комп'ютерних систем та програмного забезпечення, слід очікувати на широке впровадження комп'ютерних систем стереозору у різні галузі промисловості, у побутову техніку та в навчальний процес. Основою цього є алгоритми та програмне забезпечення, що втілені у бібліотеку функцій OpenCV.

Проте використання існуючих алгоритмів стереозору з бібліотеки OpenCV спряжено з низкою проблем, які часто призводять до зниження якості результатів та їх непостійності навіть для однотипних зображень. Пропонується використання простих та ефективних алгоритмів стереоскопічних обчислень для нескладних задач стереозору для випадків, коли швидкість роботи та передбачувана точність результату є ключовими вимогами. Запропоновані алгоритми засновані на наближенні геометричної оптики із використанням моделі камери-обскура. На їх основі було запрограмовані основні функції стереоскопічних обчислень, які згуртовані в клас StereoScopic. Клас може бути застосований в якості базового для роботи стереозору або в поєднанні із традиційними алгоритмами з бібліотеки OpenCV.

II. Алгоритм роботи системи

Звичайно система комп'ютерного стереозору повинна передбачати роботу з зображеннями стереопари незалежно від шляху їх отримання (захоплення з цифрових камер, зчитування з файлів, отримання окремих кадрів з відео), а також можливість аналізу та обробки зображень для вибору бажаного алгоритму роботи. Для розробки такої системи можна використати програмну платформу Microsoft .NET Framework 4.0, а для обробки зображень - популярну бібліотеку OpenCV 3.0 з відкритим програмним кодом. В цьому випадку в якості мови розробки можна використати C#, що передбачає використання бібліотеки-обгортки Emgu CV для можливості звертання до функцій OpenCV.

Перший крок алгоритму роботи відповідної системи полягає у зчитуванні пари зображень, отриманих за стереоскопічним принципом. Такі зображення можуть бути отримані шляхом безпосередньої передачі з камер, виділені у відеоряді або зчитані з файлів на диску. Для типової системи стереозору з двома камерами(рис. 1) важливим є одночасне отримання інформації з обох камер. Якщо сцена, що спостерігається камерами, є статичною і рух камер не передбачається, тоді одночасна зйомка не є критичним параметром. Бібліотека OpenCV надає спеціалізований клас CvCapture, який дозволяє реалізувати роботу з камерами та відео, і клас IplImage для роботи з файлами зображень.

Другий крок алгоритму роботи передбачає попередні перетворення зображень з метою покращення їх якості для подальшої роботи, наприклад, шляхом зменшення шуму на зображеннях за допомогою просторової фільтрації [2]. За допомогою класу Mat бібліотеки OpenCV можна виконувати широкий спектр різних маніпуляцій над зображеннями, використовуючи існуючі чи власні алгоритми обробки [2].

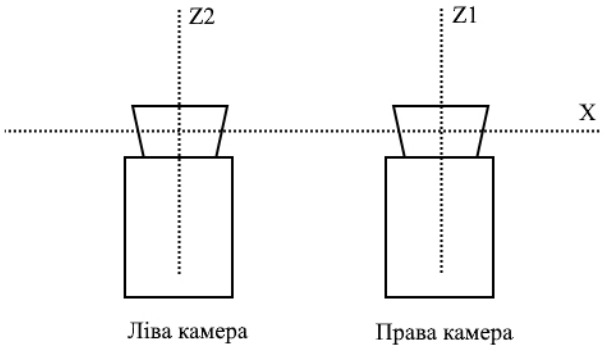


Рис. 1. Типове компонування системи стереозору

Для третього кроку алгоритму, який передбачає розпізнавання окремих об'єктів на зображенні для подальшого визначення відстані до них, застосовується клас Cascade Classification. Даний клас використовує каскад форсованих класифікаторів, що працюють з Хаар-подібними функціями. Робота Cascade Classification вимагає навчальної вибірки для формування пошукового шаблону або може використовувати готові файли з результатами навчання для аналізу зображення і виявлення об'єктів певного типу. Результати роботи формуються у вигляді масиву з інформацією про координати областей зображення, які класифіковані згідно пошукового шаблону.

На наступному кроці виконуються стереоскопічні обчислення для визначення відстані до виявлених на зображенні об'єктів. Бібліотека OpenCV надає реалізацію алгоритмів стереозору Block Matching (клас StereoBinaryBM) і Semi Global Block Matching.(клас StereoBinarySGBM) [1,3]. Для ефективної роботи даних алгоритмів необхідне виконання чотирьох послідовних етапів:

1)Стереокалібрування камер в спільній системі координат та отримання матриці проекцій для кожної з камер. Даний крок зазвичай робиться лише раз, після чого записані у файл результати використовуються повторно.

2)Захоплення і ректифікація зображень за епіполярними лініями. Даний крок передбачає використання власних матриць камер, отриманих на етапі калібрування, для обчислення фундаментальної матриці з наступним вирівнюванням зображень за епіполярними лініями. Це дозволяє мінімізувати спотворення, що пов'язані з неправильним розміщенням камер.

3)Обробка стереопари та побудова карти диспарантності. Даний етап передбачає використання алгоритму Block Matching або Semi Global Block Matching. Результатом роботи цих алгоритмів є карта диспарантності у вигляді зображення, де замість значень кольору в пікселях міститься інформація про розбіжність відповідних пікселів.

4)Побудова карти глибини для визначення відстані до точок. На основі карти диспарантності виконується тріангуляція і побудова зображення з картою глибини, де замість значень кольору кожен піксель зберігає обчислену до точки відстань.

Результати роботи зазначених алгоритмів не завжди мають коректні значення через вплив ряду негативних чинників (необхідність низки початкових припущень про зовнішнє світло, про перекриття частин об'єктів, про наявність великої кількості однакових за яскравістю областей на зображеннях тощо). Хоча дані алгоритми дозволяють достатньо швидко виконати оцінку відстані до точок на зображеннях [3], але вони не можуть бути використані для роботи в режимі реального часу на малопотужних комп'ютерах. Зіставлення даних про координати областей виявлених об'єктів та відстань до точок в цих областях дозволяє виконати оцінку відстані до об'єкта з усуненням впливу невалідних значень шляхом усереднення значень відстані, які потрапили в міжквартильний розмах. У наближенні геометричної оптики та на базі моделі камер-обскура стереоскопічні обчислення можна проводити за формулою [4]:

$$D = \frac{Bp_0}{2tg(\frac{\varphi_0}{2}) \mid p_1 - p_2 \mid}, \qquad (1)$$

де D – геометрична відстань між площиною камер та площиною, де знаходиться шукана просторова точка, B – відстань між оптичними осями камер (базова відстань), p_0 – загальна ширина зображення в пікселях, φ_0 – кут огляду камер, p_1 – горизонтальна координата точки на першому зображенні в пікселях, p_2 – горизонтальна координата точки на правому зображенні в пікселях. Уведення системи координат з її початком, наприклад, в центрі лівої камери, дозволяє отримувати абсолютні значення відповідних просторових координат.

На основі даних про просторові координати виявлених об'єктів можна визначити точки, які вказують на їх геометричні центри. Обчислення відстані до центральних точок об'єктів згідно (1) дозволяє досить точно визначати відстань, якщо об'єкт має нескладну форму. Зрозуміло, що такий підхід забезпечує високу швидкість отримання результатів через значно менші вимоги до якості вхідних зображень, хоча й не дозволяє виконувати повний аналіз зображень стереопари. Таким чином, (1) можна вважати основою для інтерактивного режиму роботи системи. Зазначені алгоритми роботи системи втілено у вигляді Windows-додатку, який надає елементи управління для вибору вхідних файлів або пристроїв, методів попередньої фільтрації та обробки зображень стереопари, інтерфейс вибору файлів з даними для класифікації об'єктів, а також дозволяє обирати один із описаних алгоритмів стереоскопічних обчислень.

III. Експерименти та результати

Для проведення експериментів було використано апаратну платформу з двома однаковими веб-камерами Logitech C120 з горизонтальним кутом огляду та роздільною здатністю 640×480 пікселів, які були розміщені на відстані 76 мм одна від одної.

Для оцінювання ефективності автоматичного режиму роботи, в якому здійснювався пошук об'єктів, в якості шаблону було використано XML-файл з даними щодо контуру людських облич. В якості об'єкта, до якого вимірюється відстань, було використано надруковане на аркуші паперу графічне зображення обличчя людини. В процесі експерименту аркуш із зображенням обличчя послідовно розміщувався на відстані від 0,6 м до 1,5 м з кроком 0,1 м відносно площини камер, після чого проводилась статична зйомка і фіксувались обчислені значення відстані. На рис.1-2 наведені приклади зображень отриманих за допомогою камер системи для відстані 0,6м. Для набору статистики зазначені вимірювання виконувалися 10 разів поспіль.

Рис. 2. Зображення лівої камери для відстані 0.6м

Рис. 3. Зображення правої камери для відстані 0.6м.

Оцінювання ефективності роботи інтерактивного режиму було виконано за допомогою зйомки статичного

плаского об'єкта, розміщеного під певними просторовими кутами відносно оптичної осі лівої камери з наступним обчисленням відстаней між окремими точками об'єкта на основі зображень стереопари. Для зручності спеціальними маркерами позначались три пари опорних точок: A1, A2, B1, B2, C1, C2 (рис.3).

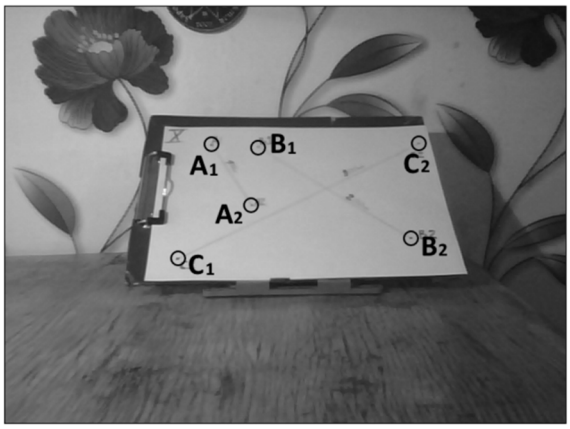


Рис. 4. Зображення лівої камери з точками позначеними маркерами.

Дані пари точок утворювали три відрізки A1A2, B1B2 та C1C2. Експеримент проводився для 10 різних конфігурацій шести точок, випадково обраних на пласкому об'єкті. Фізична відстань між точками залишалась однаковою і становила A1A2=100 мм, B1B2=200 мм, C1C2=300 мм. В процесі експерименту незначним чином змінювались кути нахилу об'єкта, а середня відстань від площини камер до його найближчої частини становила 650 мм. Під час експерименту просторові кути нахилу об'єкта не вимірювались.

Для експериментів з оцінювання ефективності автоматичного режиму роботи, в якому виконувався пошук обличчя людини та оцінювалась відстань до нього, розпізнавання умовного обличчя людини було вірним в усіх випадках як для лівої, так і для правої камери. Усереднення значень відстаней до обличчя за результатами 10 експериментів призвело до таких результатів. Для фактичної відстані 0,6 м усереднена відносна похибка обчислення відстані склала 0,3%, для відстаней від 0,7 м до 0,9 м – 0,4%, для відстані 1 м – 0,6%, для відстаней від 1,1 м до 1,2 м – 0,8%, а для відстаней від 1,3 м до 1,5 м – 1,5%.

Результати експериментів з оцінювання ефективності інтерактивного режиму роботи мали наступний зміст. Для кожної пари точок за результатами 10 експериментів обчислювались відстані між цими точками, потім отримані значення арифметично усереднювалися та обчислювалось їх середнє значення, яке для відрізка A1A2 склало 100,7 мм, для відрізка B1B2 – 201,3 мм, а для відрізка C1C2 – 301,3 мм.

Результати проведених експериментів доводять ефективність розробленої системи комп'ютерного стереозору. В цілому, реалізований варіант системи забезпечує обчислення відстані в діапазоні від 0,5м до 1,5м з відносною похибкою не більше 1,5%. Оцінювання відстаней на пласкому об'єкті забезпечується в середньому з абсолютною похибкою до 1,5 мм на відстані 0,65 м від об'єкта. Але ця похибка зростає із збільшенням відстані до об'єкта.

Керування роботою системи та візуалізація результатів здійснювались через розроблений графічний інтерфейс. Перед початком роботи заповнювалась інформація про апаратну частину системи, після чого обирався автоматичний або інтерактивний режим. Завантаження автоматичного режиму роботи відбувалось шляхом активізації програмного методу AutoMode(), після чого робота з обчислення відстані до знайдених об'єктів тривала безупинно до натискання відповідної кнопки зупинки роботи. Інтерактивний режим роботи активізувався через програмний метод InteractiveMode() і вимагав уведення інформації про обрані просторові точки та їх зміст (поодинока точка, дві точки належать прямій, три точки належать площині), після чого виводилась відповідна інформація про відстані до них та поміж ними, а також про просторові кути нахилу визначених прямих та площин.

Використання спрощених алгоритмів стереозору має певні переваги над алгоритмами побудови карт диспарантності з бібліотеки OpenCV. Це стосується тих випадків, коли повний стереоскопічний аналіз сцени, яка спостерігається, не є потрібним. Також до переваг можна віднести невисокі вимоги до зображень сцени, що дозволяє однаково успішно працювати з вхідними зображеннями, які є різними за якістю та за змістом.

IV. Висновки

На основі функціоналу бібліотеки OpenCV та з використанням бібліотеки-обгортки Emgu CV розроблена комп'ютерна система стереозору у вигляді Windows-додатку, що функціонує на персональному комп'ютері, та двох камер. Алгоритм роботи системи передбачає гнучкий вибір варіантів її функціонування залежно від потреб користувача. Для покращення якості отриманих результатів апаратна частина системи може вдосконалюватися камерами кращої якості та розширюватися додатковою кількістю камер, а програмна частина може доповнюватися новими функціями, які є необхідними.

Література References

[1] B. Tippetts, Review of stereo vision algorithms and their suitability for resource-limited systems / Beau Tippetts, Dah Jye Lee, Kirt Lillywhite, James Archibald // Journal of Real-Time Image Processing January 2016, Volume 11, Issue 1, pp. 5–25.

[2] Ju.V. Vizil'ter, Obrabotka i analiz izobrazhenij v zadachah mashinnogo zrenija: kurs lekcij i prakticheskih zanjatij / Ju.V.Vizil'ter, S.Ju. Zheltov, A.V. Bondarenko, M.V. Ososkov, A.V. Morzhin. – M.:Fizmatkniga, 2010. – 672 s.

[3] M. Stefano, Stereo Vision Algorithms for FPGAs /Stefano Mattoccia // IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2013, pp. 636-641.

[4] S.M. Vovk, Metody obrobky zobrazhen ta kompiuternyi zir / S.M. Vovk, V.V. Hnatushenko, M.V. Bondarenko. – D. : LIRA, 2016. –148 s.

Моделювання Структури Програмного Комплексу для Управління Газотранспортними Системами

Ольга Химко
Кафедра автоматизації теплових і хімічних процесів
Національний університет «Львівська політехніка»
Львів, Україна
olgakhymko@ukr.net

Василь Чекурін
Технічний факультет
Куявсько-Поморський університет в Бидгощі
Бидгощ, Польща
v.chekurin@ mail.kpsw.edu.pl

Modeling of Structure of Software Complex for Management by Gas Transmission Systems

Olga Khymko
Dep. for Automation of Thermal and Chemical Processes
L'viv Polytechnic National University
L'viv, Ukraine
olgakhymko@ukr.net

Vasyl Chekurin
Technical Department
Kujawy and Pomorze University in Bydgoszcz
Bydgoszcz, Poland
v.chekurin@ mail.kpsw.edu.pl

***Анотація*—Розглянуто модель структури інтегрованого програмно-технічного комплексу для автоматизації управління газотранспортною системою (ГТС). Комплекс включає в себе програмно-технічні системи для управління інформаційними та комерційними процесами, транспортуванням газу й експлуатацією споруд ГТС, які об'єднані єдиним інтерфейсом користувача та спільними інформаційними ресурсами.**

***Abstract*—A model of functionality and structure of integrated hardware-software complex for automation of management by a gas transmission system has been considered in the paper. The complex consists of four hardware-software systems aimed for computerization of management by informational and business processes, gas transportation and GTS infrastructure maintenance. The systems are united by single user interface and common informational resources.**

Ключові слова — газотранспортні системи; автоматизація управління; програмно-тезнічніі комплекси

Keywords — gas transmission systems; automation of management; hrdware-software complexes

I. Вступ

Газотранспортні системи (ГТС) призначені для планового переміщення природного газу від місць видобування (родовищ) до місць споживання. Комплекс ГТС об'єднує магістральні газопроводи і підземні сховища газу, оснащені технологічними елементами [1]. Складові частини ГТС утворюють в сукупності цілісний інженерний об'єкт.

Щоб ефективно керувати роботою цієї системи необхідна автоматизація технологічних процесів і виробничої діяльності підрозділів, критичних для її функціонування. Цього можна досягнути шляхом комп'ютеризації із використанням програмно-технічних комплексів (ПТК). ПТК слід розглядати як SIS-системи (Software Intensive Systems) — складні системи, функціональні можливості яких визначаються головно програмним забезпеченням [2]. Для побудови таких систем застосовують модельно-базований підхід, за яким моделі процесів, окремих компонент і системи в цілому використовуються упродовж усього циклу її життя [3]. Для побудови цих моделей застосовують різні засоби: текстуальне та графічне моделювання з використанням діаграмної техніки, уніфіковану мову візуального моделювання UML, предметно-орієнтовані мови моделювання DSL [4], засоби моделювання за методологією IDEF [4], а також формальні математичні методи теорії складних програмних систем [5].

II. Моделі предметної області

При побудові моделі SIS-системи виходять із певних моделей предметної області. Об'єктом ПТК є ГТС, а його предметною областю є процеси управління ГТС. Тож розглянемо спочатку моделі ГТС.

ГТС є гетерогенна система. Її функціонування визначають різнорідні процеси – фізичні, технологічні інформаційні та комерційні (рис.1).

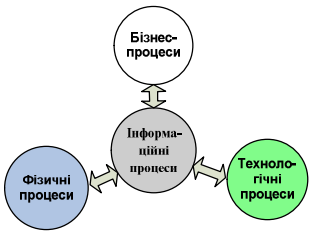


Рис. 1. Взаємодія процесів ГТС

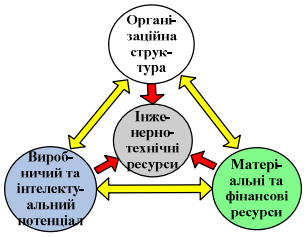


Рис. 2. Структура виробничо-комерційного комплексу ГТС

Відповідно можемо розглядати моделі ГТС з різних точок зору – як виробничо-комерційну структуру, як фізичну систему чи як інформаційне середовище.

Структура ГТС як виробничо-комерційної структури показана на рис.2.

Як фізична система ГТС являє собою відкриту термодинамічну систему, що складається із двох основних підсистем – споруд ГТС та акумульованого в них газу (рис.3). До споруд ГТС відносимо її складові, заповнені газом (трубопроводи, компресорні станції, пласти підземних сховищ газу) та інші технологічні елементи. Взаємодія цих двох підсистем між собою та ГТС із довкіллям відбувається шляхом обміну між ними масою M, імпульсом P та енергією E (рис.4)

.

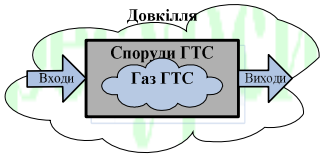


Рис. 3. ГТС як відкрита термодинамічна система

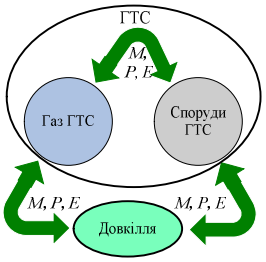


Рис. 4. Взаємодія підсистем ГТС і довкілля

Розглядаючи ГТС як об’єкт управління можна означити її як нелінійну динамічну систему з розподіленими параметрами, для якої характерні значні розміри й інерційність, дія розподілених і зосереджених зовнішніх впливів, істотна невизначеність значень внутрішніх параметрів і зовнішніх чинників.

Найпростішою моделлю ГТС є сукупність $\mathbf{G}_B = (\mathbf{W}, \mathbf{O})$, що об’єднує множини входів $\mathbf{W}$ та виходів $\mathbf{O}$. Цю систему можна представити у вигляді дводольного графа $\mathbf{G}_B = (\mathbf{N}_{\text{ext}}, \mathbf{L}_{\text{ext}})$, де $\mathbf{N}_{\text{ext}} = \mathbf{W} \cup \mathbf{O}$ – множина зовнішніх вузлових елементів ГТС, $\mathbf{L}_{\text{ext}}$ – множина ребер, які пов’язують вхідні вузли $\mathbf{W}$ графа з його вихідними $\mathbf{O}$ вузлами. Позначимо як $\mathbf{P}_{\mathbf{N}_{\text{ext}}}$ і $\mathbf{P}_{\mathbf{L}_{\text{ext}}}$ сукупності множин параметрів стану для усіх вузлових і лінійних елементів: $\mathbf{P}_{\mathbf{N}_{\text{ext}}} = \{\mathbf{P}_{N_{\text{ext}}} \ \forall N_{\text{ext}} \in \mathbf{N}_{\text{ext}}\}$, $\mathbf{P}_{\mathbf{L}_{\text{ext}}} = \{\mathbf{P}_{L_{\text{ext}}} \ \forall L_{\text{ext}} \in \mathbf{L}_{\text{ext}}\}$.

Нехай $\mathbf{P}_{\text{B}} = \mathbf{P}_{\mathbf{N}_{\text{ext}}} \cup \mathbf{P}_{\mathbf{L}_{\text{ext}}}$ — множина, тоді пара $\mathbf{M}_{\text{B}} = (\mathbf{G}_{\text{B}}, \mathbf{P}_{\text{B}})$ визначає бізнес-модель ГТС. За належного вибору множини параметрів стану $\mathbf{P}_{\text{B}}$, модель $\mathbf{M}_{\text{B}}$ є достатньою для управління ГТС на рівні бізнес-процесів.

Проте в рамках цієї моделі неможливо оптимізувати затрати на транспортування газу, оцінювати вплив системи на довкілля, як і зворотній вплив, враховувати вплив режимів транспортування на технічний стан споруд тощо. Для цього слід взяти до розгляду фізичні процеси, які протікають у системі під час її експлуатації.

Модель структури ГТС подамо у вигляді сукупності $\mathbf{M}_T = (\mathbf{C}, \mathbf{F})$, де $\mathbf{C}$ – множина споруд ГТС, $\mathbf{F}$ – накопичений в порожнинах споруд. Споруди ГТС поділяються на дві групи — вузлові та лінійні елементи (трубопроводи). Схема з’єднання споруд ГТС у цілісну мережу визначає модель конфігурації $\mathbf{G}_{\mathbf{C}}$, яку можна подати у вигляді сукупності $\mathbf{G}_{\mathbf{C}} = (\mathbf{N}, \mathbf{T})$, де $\mathbf{N}$ — множина вузлових елементів, $\mathbf{T}$ — множина трубопроводів.

Кожному елементу моделі конфігурації $N \in \mathbf{N}$ та $T \in \mathbf{T}$ поставимо у відповідність математичну модель фізичних процесів, які в ньому протікають $N \to \mathbf{P}_N$ та $T \to \mathbf{P}_T$. Сукупності математичних моделей для усіх вузлових $\mathbf{P}_{\mathbf{N}} = \{\mathbf{P}_N \ \forall \in \mathbf{N}\}$ і лінійних $\mathbf{P}_{\mathbf{T}} = \{\mathbf{P}_T \ \forall T \in \mathbf{T}\}$ елементів доповнимо сукупністю $\mathbf{P}_{\mathbf{a}}$ співвідношень, які виражають закони збереження маси, імпульсу та енергії у взаємодії споруд ГТС $\mathbf{C}$ з газом $\mathbf{F}$, який через них протікає, та довкіллям. Сукупність усіх трьох множин утворює математичну модель процесів у спорудах ГТС: $\mathbf{P}_C = \mathbf{P}_{\mathbf{N}} \cup \mathbf{P}_{\mathbf{T}} \cup \mathbf{P}_{\mathbf{a}}$, а модель споруд ГТС $\mathbf{M}_C$ визначимо як $\mathbf{M}_C = (\mathbf{G}_{\mathbf{C}}, \mathbf{P}_{\mathbf{C}})$.

У відповідність моделі конфігурації $\mathbf{G}_{\mathbf{C}}$ поставимо граф конфігурації $\mathbf{G}_{\mathbf{F}} = (\mathbf{X}, \mathbf{E})$ де $\mathbf{X}$ та $\mathbf{E}$ — множини вершин і ребер графа. Множина $\mathbf{X}$ представляє усі вузлові елементи, а множина $\mathbf{E}$ — усі лінійні елементи моделі конфігурації: $\mathbf{X} \leftrightarrow \mathbf{N}$, $\mathbf{E} \leftrightarrow \mathbf{T}$. Відтак існує взаємно-однозначна відповідність між моделями: $\mathbf{G}_{\mathbf{F}} \leftrightarrow \mathbf{G}_{\mathbf{C}}$ Вузлам $X \in \mathbf{X}$ та ребрам $E \in \mathbf{E}$ поставимо у відповідність математичні моделі $\mathbf{P}_X$ та $\mathbf{P}_E$, які описують рух газу у порожнинах відповідних вузлового $N \leftarrow X$ та лінійного $T \leftarrow E$ елементів моделі конфігурації $\mathbf{G}_{\mathbf{C}}$.

Позначимо $\mathbf{P}_{\mathbf{X}} = \{\mathbf{P}_X \ \forall X \in \mathbf{X}\}$ та $\mathbf{P}_{\mathbf{E}} = \{\mathbf{P}_E \ \forall E \in \mathbf{E}\}$ сукупності математичних моделей руху газу у всіх порожнинах вузлових та лінійних елементів ГТС, а як $\mathbf{P}_{\mathbf{FI}}$ – множину співвідношень, які визначають закони збереження маси імпульсу й енергії у процесах взаємодії газу ГТС $\mathbf{F}$ зі спорудами $\mathbf{C}$ та довкіллям. Вводячи сукупність $\mathbf{P}_{\mathbf{F}} = \{\mathbf{P}_{\mathbf{X}}, \mathbf{P}_{\mathbf{E}}, \mathbf{P}_{\mathbf{FI}}\}$, означимо модель руху газу

через ГТС як $\mathbf{M_F} = (\mathbf{G_F}, \mathbf{P_F})$, а модель ГТС як відкритої термодинамічної системи у вигляді сукупності $\mathbf{M_T} = (\mathbf{M_C}, \mathbf{M_F})$.

Моделі $\mathbf{M_B}$, $\mathbf{M_C}$ і $\mathbf{M_F}$ утворюють концептуально різні проекції ГТС як об'єкта реального світу в світ відображень. Разом із інформаційною моделлю $\mathbf{M_I}$ вони утворюють модель ГТС $\mathbf{M} = \{\mathbf{M_B}, \mathbf{M_C}, \mathbf{M_F}.\mathbf{M_I}\}$, необхідну для побудови ПТК для автоматизації її управління.

Між математичними моделями $\mathbf{M_B}$, $\mathbf{M_C}$ і $\mathbf{M_T}$ існує співвідношення $\mathbf{M_B} \subset \mathbf{M_C} \subset \mathbf{M_T}$.

Взаємозв'язок між моделями $\mathbf{M_C}$ та $\mathbf{M_F}$ проілюстровано на рис. 5. Тут зображено структуру математично-алгоритмічного забезпечення для реалізації функцій управління транспортуванням газу й експлуатацією споруд ГТС.

Рис.5. Структура математично-алгоритмічного забезпечення

III. ФУНКЦІЇ УПРАВЛІННЯ ГТС

Управління газотранспортною системою має на меті: а) гарантоване доставляння природного газу із входів системи на її виходи в заданих об'ємах; б) підтримання системи у технічному стані, який забезпечує гарантоване доставляння природного газу і заданий рівень екологічної безпеки; в) отримання максимального прибутку від роботи ГТС.

Структуру управління ГТС подамо у вигляді сукупності чотирьох напрямків: а) управління бізнес процесами, б) управління інформаційними процесами, в) управління транспортування газу та г) управління технічним обслуговуванням і експлуатацією. Виходячи з цього множину **F** функцій управління ГТС розглядатимемо як сукупність функцій управління цими групами процесів: $\mathbf{F} = (\mathbf{F}_I, \mathbf{F}_B, \mathbf{F}_F, \mathbf{F}_C)$, де $\mathbf{F}_I$, $\mathbf{F}_B$, $\mathbf{F}_F$ та $\mathbf{F}_C$ — функції управління інформаційними, комерційними, транспортними процесами й експлуатацією споруд ГТС.

Управління бізнес процесами має на меті: а) досягнення максимальної економічної ефективності роботи ГТС, б) точне виконання зобов'язань взятих перед постачальниками і споживачами природного газу, в) розвиток інфраструктури ГТС, відповідно до короткострокових і довгострокових прогнозів розвитку ринку газотранспортних послуг, г) мінімізацію екологічного та бізнес ризиків.

Управління інформаційними процесами має на меті підтримання інформаційної системи у стані, який забезпечує: а) відбір і нагромадження об'єктивних даних у обсягах, необхідних для надійного управління комерційними процесами, транспортування газу й експлуатацією споруд ГТС; б) обмін інформацією між підсистемами ГТС в реальному часі; в) оперативне надання інформаційних послуг користувачам усіх груп відповідно до їх потреб.

Управління транспортуванням газу має на меті реалізацію режимів транспортування газу, які забезпечують: а) виконання зобов'язань перед постачальниками та споживачами природного газу, б) мінімальні втрати газу та витрати паливно-енергетичних ресурсів, в) дотримання заданих рівнів інтенсивності зношування споруд ГТС, г) дотримання допустимих норм забруднення атмосферного та водного басейнів, ґрунтів і геологічних пластів, д) дотримання заданих рівнів технологічної й екологічної безпеки. Реалізація цієї мети відбувається шляхом керування конфігурацією мережі, режимами роботи компресорних станцій і підземних сховищ газу, а також управлінням виробничим та інтелектуальним потенціалами, які є у розпорядженні диспетчерської служби та інших підрозділів.

Управління експлуатацією споруд ГТС має на меті утримання споруд ГТС у технічному стані, який забезпечує: а) виконання усіх функцій транспортування газу, б) високу ефективність функціонування споруд ГТС, в) заданий рівень екологічної безпеки. Реалізація цієї мети відбувається шляхом технічного обслуговування, реконструкції та модернізації споруд ГТС, спорудження нових трубопроводів, компресорних станцій та інших споруд, управлінням виробничим та інтелектуальним потенціалами, які є у розпорядженні служби експлуатації, ремонту та реконструкції тощо.

IV. МОДЕЛЬ СТРУКТУРИ ПТК ГТС

Подамо структуру ПТК як сукупність чотирьох програмно-технічних систем (ПТС), які призначені для автоматизації управління: а) бізнес процесами (ПТС БП), б) інформаційними процесами (ПТС ІП), в) процесами транспортування газу ПТС ТГ та г) процесами експлуатації споруд ГТС (ПТС ЕС). Ці чотири складові об'єднані в інтегрований програмно-технічний комплекс єдиним інтерфейсом користувача та спільними інформаційними ресурсами (рис.6). Така структура дозволяє спростити задачі автоматизації напрямків управління ВБК ГТС

Інтерфейс користувача ПТК надає доступ до функцій усіх ПТС, обчислювальних та інформаційних ресурсів, а також зовнішніх пристроїв, залежно від потреб конкретного користувача. Інтерфейс автоматично налаштовується під зареєстрованого користувача, згідно

його потреб і ролі, яку він виконує в системі у цьому сеансі взаємодії з ПТК.

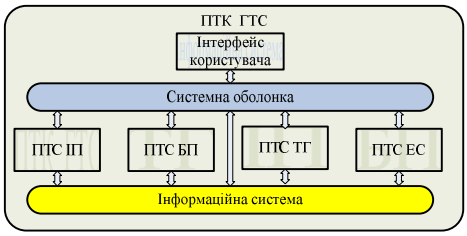


Рис.6. Структура ПТК

Системна оболонка реалізує функції керування обліковими записами користувачів (реєстрації, авторизація, ідентифікація, автентифікація), доступом прикладних процесів, які виконуються ПТК, до спільних інформаційних і обчислювальних ресурсів, а також до периферійних пристроїв, здійснює захист інформації від локальних і віддалених атак. Важливою функцією системної оболонки є формування та підтримання віртуальних обчислювальних середовищ для усіх користувачів. Ці середовища формуються з використанням системних ресурсів і функціоналів ПТС напрямків.

ПТС реалізують функції, необхідні для автоматизації відповідних напрямків управління. Їхня робота базується на математичних моделях та алгоритмічному забезпеченні, специфічних для кожного напрямку управління. До складу ПТС напрямку входять інтерфейс на частина, редактор задач, математичне ядро, бібліотека функцій, система підтримки прийняття управлінських рішень, система відображення та документування. Три складники ПТС — математичне ядро, бібліотека функцій і система прийняття рішень будуються на основі моделей $\mathbf{M}_B, \mathbf{M}_C, \mathbf{M}_F$ та $\mathbf{M}_I$.

Інтерфейсна частина ПТС відповідає за взаємодію внутрішніх її модулів із системною оболонкою.

Редактор задач дозволяє: а) завантажувати задачі з інформаційної системи, б) редагувати їх, в) формулювати нові задачі, специфічні для предметної області відповідних напрямків управління та зберігати їх у інформаційній системі, г) відлагоджувати задачі тощо.

Математичне ядро здійснює чисельну реалізацію сформульованих прикладних задач. Бібліотека функцій містить функції і процедури, які можна підвантажувати в математичне ядро і використовувати для розв'язування задач. Цю бібліотеку можна доповнювати, що дозволяє розширювати функціональні можливості ПТК впродовж усього циклу його життя;

Підсистема підтримки рішень дозволяє формувати обґрунтовані управлінські рішення, виходячи із результатів розв'язування задач, інформації, отриманої із бази знань системи та інших джерел, використовуючи методи нечіткої логіки та методи штучного інтелекту.

Підсистема відображення та документування, призначена для графічного, табличного чи математичного представлення результатів розв'язування задач і прийнятих рішень на екрані чи твердих носіях, формування відповідних документів у електронній формі і збереження їх у інформаційній системі для подальшого використання.

Інформаційна система надає ресурси, необхідні для функціонування усіх ПТС. ПТС напрямків містять функції для тематичного оброблення даних згідно потреб управління цим напрямком. Інформація, отримана в результаті такого оброблення нагромаджується у спільних інформаційних ресурсах, поповнюючи бази знань та репозиторії документів за напрямками управління.

Реалізація запропонованої системи передбачає наступні етапи: а) розроблення функційних моделей ПТС напрямків,; б) побудова моделі взаємодії підсистем ПТК з системною оболонкою та інформаційною системою; в) побудова математичних моделей процесів, визначальних для кожного із чотирьох напрямків управління, г) математичне формулювання задач, специфічних для кожного напрямку і розроблення алгоритмів для їхньої чисельної реалізації; д) проектування ПТС напрямків, інтегрованого програмно-технічного комплексу, та інформаційної системи; з) кодування; ж) інтеграція та тестування ПТС напрямків; е) інтеграція та тестування ПТК.

V. Висновки

Газотранспортна система як об'єкт управління є відкрита гетерогенна система, що поєднує в єдине ціле інженерно-технічний комплекс, організаційну структуру, виробничий та інтелектуальний потенціали, матеріальні та фінансові ресурси. Взаємопов'язані процеси, які визначають функціонування цієї системи, можна поділити на чотири групи – інформаційні, комерційні, технологічні та фізичні процеси. Управління цією системою можна автоматизувати, використовуючи інтегрований програмно-технічний комплекс, що складається із програмно-технічних систем для управління процесами транспортування газу, експлуатації споруд ГТС, інформаційних і бізнес-процесів. Єдиний інтерфейс користувача та спільні інформаційні ресурси забезпечують обмін інформацією між ПТС, ефективний захист інформації та розподіл системних ресурсів між користувачами, які діють у предметних областях різних ПТС.

Література References

[1] "The Interstate Natural Gas Transmission System: Scale, Physical Complexity and Business Model" [Online]. Available: http://www.ingaa.org/?ID=10724

[2] *V. Bonneau,* "Software intensive systems in the future" [Online]. Available: ftp://ftp.cordis.europa.eu/pub/ist/docs/ embedded/puissochet-060207_en.pdf.

[3] MSDN. Моделирование приложения. [Online]. Available: http://msdn.microsoft.com/ru-ru/library/

[4] C. Menzel , R. J. Mayer, "The IDEF family of languages" In: *Handbook on Architectures of Information Systems. International Handbooks on Information Systems.* – 2006. – P. 215-249.

[5] D. Krob, "Modeling of complex software systems: a reasoned overview." in: Najm E., Pradat-Peyre JF., Donzeau-Gouge V.V. (eds) Formal Techniques for Networked and Distributed Systems - FORTE 2006. Lecture Notes in Computer Science, vol 4229 Springer, Berlin, Heidelberg.

Принципи Побудови Систем Інформаційної Підтримки Прийняття Рішень на Авіаційних Підприємствах

Володимир Темніков
кафедра засобів захисту інформації
Національний авіаційний університет
Київ, Україна
temnikov_v@ukr.net

Андрій Темніков
кафедра засобів захисту інформації
Національний авіаційний університет
Київ, Україна
temnikoff@ukr.net

Principles of Building Information Systems for Decision Support in Aviation Enterprises

Volodymyr Temnikov
dept. of Information Security
National Aviation University
Kyiv, Ukraine
temnikov_v@ukr.net

Andrii Temnikov
dept. of Information Security
National Aviation University
Kyiv, Ukraine
temnikoff@ukr.net

***Анотація*—Наведені принципи побудови систем інформаційної підтримки прийняття оперативних рішень керівниками та операторами авіабудівних та авіатранспортних підприємств на основі інтелектуальних інформаційних технологій.**

***Abstract*—The principles of building information support systems for making operational decisions by managers and operators of aircraft building and air transport enterprises on the basis of intelligent information technologies are given.**

Ключові слова—системи підтримки прийняття рішень, інтелектуальні інформаційні технології, штучні нейронні мережі, нечітка логіка, психофізіологічний стан, авіадиспетчери, оператори

Keywords—decision support systems, intelligent information technologies, artificial neural networks, fuzzy logic, psychophysiological state, air traffic controllers, operators

I. Вступ

В даний час в авіаційній галузі усе більшої актуальності набуває створення інформаційних систем, призначених для оперативного виявлення порушень у роботі авіапідприємств. Особлива увага приділяється запобіганню порушень, пов'язаних з людським фактором. Це пояснюється тим, що, наприклад, від того, наскільки швидко і грамотно оператори будуть приймати рішення по виявленню і розпізнаванню порушень в роботі обчислювальних систем (ОС), в значній мірі залежить ефективність роботи авіабудівних підприємств, а від оперативності та обґрунтованості рішень адміністраторів диспетчерських служб аеропортів – кількість авіаційних аварій та подій, пов'язаних з людським фактором (в тому числі, помилками авіадиспетчерів, зумовленими їх перебуванням в неналежному психофізіологічному стані (ПФС)), які можуть призвести до великих людських жертв і значних матеріальних збитків.

Важливим є також те, що зазначені порушення часто стають джерелом загроз для критичної інформації, що циркулює і зберігається на підприємствах.

Цим обумовлена велика актуальність робіт, що проводяться авторами по створенню систем інформаційної підтримки прийняття рішень (СІППР) адміністраторами, операторами та іншими особами, які приймають рішення (ОПР) в процесі роботи авіапідприємств. Дослідження проводилися для двох типів авіаційних підприємств: машинобудівних (авіабудівних) і авіатранспортних (аеропортів).

В якості критеріїв ефективності роботи СІППР автори розглядають швидкодію і достовірність їх роботи.

У статті представлені вимоги до СІППР для авіапідприємств зазначених типів, а також основні

принципи побудови та способи підвищення достовірності роботи і швидкодії систем прийняття рішень ОПР, застосування яких дозволило забезпечити роботу в режимі реального часу при високій достовірності. Основну увагу у статті приділено поліпшенню характеристик СІППР, призначених для підтримки прийняття рішень ОПР при розпізнаванні і аналізі відмов ОС авіабудівних підприємств, а також при визначенні ПФС диспетчерів та операторів обох типів підприємств, що розглядаються.

При розробці СІППР враховувалися особливості їх функціонування в реальних умовах експлуатації, а саме той факт, що СІППР функціонує в умовах неповноти і невизначеності інформації, а також випадкового впливу внутрішніх факторів і зовнішнього (фізичного) середовища (внутрішніх і зовнішніх впливів випадкового характеру). Ця обставина істотно ускладнює процес розробки СІППР, обумовлює необхідність застосування методів побудови складних розподілених динамічних систем і інтелектуальних інформаційних технологій (ІТ).

Нижче наведені основні принципи використання інтелектуальних ІТ для побудови та підвищення достовірності роботи і швидкодії СІППР при їх застосуванні для розпізнавання відмов (збоїв в роботі) ОС, а також зниження впливу людського фактора на безпеку польотів в авіації.

II. Принципи побудови систем інформаційної підтримки прийняття рішень на авіабудівніх підприємствах

В основі розробки СІППР для застосування на авіабудівних підприємствах з метою підвищення ефективності роботи та забезпечення інформаційної безпеки підприємств лежать такі положення:

1. Авіабудівні підприємства розглядаються як розосереджені підприємства. При цьому зазначимо, що прагнення до включення до складу підприємства все більшої кількості об'єктів і підвищення захищеності інформації знаходяться в протиріччі, тому що розширення підприємства веде, внаслідок збільшення кількості об'єктів для атак зловмисників (а отже, і загроз для інформації), до підвищення уразливості інформації.

2. Автоматизована система (АС) розосередженого авіабудівного підприємства, в класифікації НД ТЗІ 2.5-005-99 [1], являє собою АС класу 3 і розглядається як сукупність АС класу 2, з'єднаних між собою лініями зв'язку (рис.1). АС класу 2, представлені на рис.1, суть АС, що працюють в підрозділах базового підприємства або інших підприємств, що входять до складу розосередженого, і є локальними ОС.

3. Захисту на авіабудівних підприємствах підлягає інформація, що циркулює на них у вигляді інформаційних потоків, що розповсюджуються по лініях зв’язку, і зберігається в їх базах даних/знань. АС класу 3 містить центральну базу даних/знань та локальні – у підрозділах розосередженого підприємства.

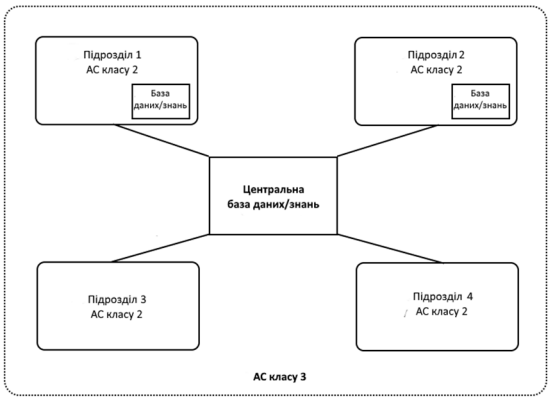


Рис. 1. Узагальнена структурна схема АС розосередженого авіабудівного підприємства

4. Відповідно до НД ТЗІ 1.1-003-99 [2], АС розглядається як організаційно-технічна система, що реалізує ІТ та об'єднує ОС, фізичне середовище, персонал і оброблювану інформацію.

5. Прийняття рішень щодо забезпечення інформаційної безпеки розосереджених авіабудівних підприємств ґрунтується на аналізі та оцінці змін ризиків при зміні параметрів АС, що мають місце внаслідок внутрішніх і зовнішніх впливів. Ризик може бути описаний кількісними або якісними (на основі методів теорії нечітких множин і лінгвістичних змінних) показниками.

6. З огляду на неповноту і невизначеність інформації, що є вихідною для прийняття рішень, а також випадковий характер впливу внутрішніх факторів і зовнішнього середовища, автори розглядають СІППР як складні розподілені динамічні системи і пропонують для їх побудови застосовувати інтелектуальні ІТ (методи теорій нечітких множин, розпізнавання образів, експертних систем, штучні нейронні мережі (ШНМ)). При цьому актуальним є комбіноване застосування різних ІТ для об'єднання їх переваг. Так, наприклад, одночасне використання ШНМ і експертних систем дозволяє забезпечити як високу функціональну гнучкість СІППР, так і швидкодію; одночасне використання методів експертних систем і нечіткої логіки – підвищити швидкодію системи і зменшити розміри баз знань.

III. Принципи побудови інформаційних систем підтримки прийняття рішень в диспетчерських службах аеропортів

Зниження впливу людського фактора на безпеку польотів в авіації автори пропонують забезпечити шляхом проведення моніторингу ПФС авіадиспетчерів в процесі виконання ними своїх функціональних обов'язків [3]. Під ПФС розуміється емоційний стан диспетчера, а також стани втоми, сонливості і тривоги.

Для надання інформаційної підтримки у прийнятті рішень особам, які здійснюють контроль за діями авіадиспетчерів (адміністратору, старшому диспетчеру), більш ефективного і швидкого запобігання порушень в роботі диспетчерів, авторами розроблена СІППР,

призначена для моніторингу ПФС авіадиспетчерів. Задачею СІППР, яка повинна функціонувати в режимі реального часу, є своєчасне подання сигналу ОПР про порушення в роботі диспетчерів; звіт про роботу системи може служити документальним підтвердженням факту порушення.

Використання голосу в якості аналізованого образу дає можливість здійснювати контроль диспетчерів безконтактно, дистанційно, не відволікаючи їх від роботи.

Основними вимогами до СІППР є забезпечення роботи системи в режимі реального часу (тобто підвищені вимоги пред'являються до швидкодії системи) і проведення контролю за безперервним злитим мовленням диспетчера.

Розробка системи проводилася з урахуванням особливостей її практичного застосування: невеликі відстані від джерел звуку до СІППР, відносно малі значення шумів в аналізованих мовних сигналах, малі кількості контрольованих осіб, наявність специфічних вимог до авіадиспетчерів (щодо застосування певної нормативно встановленої фразеології, розмірeності мови та ін.).

Розроблювана авторами система моніторингу ПФС авіадиспетчерів за їх безперервним злитим мовленням включає в себе такі підсистеми: сканування та дискретизації мовного сигналу; попередньої обробки сигналу, яка полягає в шумозаглушенні; сегментації безперервного мовлення на мовні фрагменти; пошуку в безперервному мовленні і виділення ключових мовних фрагментів (моделі яких заздалегідь включені до відповідної бази даних – словника); моніторингу ПФС авіадиспетчерів по ключовим мовним фрагментам, що проводиться на основі їх пофонемного аналізу.

У якості мовних фрагментів, за якими проводиться моніторинг ПФС, використовуються словосполучення, слова (елементи слів), що часто застосовуються авіадиспетчерами в процесі роботи, входять до складу професійної фразеології, встановленої нормативними документами.

Підставою для проведення моніторингу ПФС авіадиспетчерів по параметрам, що характеризують конкретні мовні фрагменти, послужили результати проведених авторами досліджень, які показали, що, зокрема, значення одного з основних параметрів мовного сигналу – частоти основного тону (ЧОТ) – для одних і тих же фонем (аллофонів) в різних мовних фрагментах можуть істотно відрізнятися.

Підсистема виділення ключових мовних фрагментів з безперервного злитого мовлення диспетчера створюється з застосуванням методів теорії розпізнавання образів [4]. При цьому, модуль класифікації підсистеми автори пропонують будувати на базі дикторонезалежних ШНМ, навчених на розпізнавання ключових мовних фрагментів.

Застосування ШНМ дозволяє істотно підвищити точність роботи підсистеми. Однак можливість застосування ШНМ в якості основи для побудови її модуля класифікації обмежується кількістю параметрів сигналів, що надходять на її входи. У наших розробках параметризація вказаної підсистеми здійснюється на основі методу короткотривалого аналізу [5]. У разі безпосереднього використання в якості інформативних параметрів кепстральних коефіцієнтів, що розраховуються на кожному фреймі, на які ділиться ключовий мовний фрагмент, їх кількість, а, значить, і кількість вхідних параметрів (нейронів) ШНМ, досягала б значних величин.

Для забезпечення можливості побудови модулів класифікації підсистем виділення ключових мовних фрагментів на основі ШНМ було необхідно істотно зменшити кількість інформативних параметрів мовних сигналів, що подаються на вхід ШНМ. Це було досягнуто в процесі розробки нової системи інформативних параметрів, застосування якої дозволило суттєво (більш ніж на порядок) зменшити кількість вхідних параметрів (нейронів) ШНМ без будь-якого істотного зменшення достовірності роботи.

Моніторинг ПФС за ключовими мовними фрагментами пропонується проводити на основі порівняльного аналізу контрольних і еталонних інформативних параметрів, що характеризують окремі (в першу чергу, голосні) фонеми, що входять до складу ключових мовних фрагментів, за такими етапами: сегментація ключового слова на фонеми, розпізнавання голосних фонем, параметризація голосних фонем, класифікація мовних сигналів і прийняття рішення про ПФС диспетчера. В процесі проведення моніторингу ПФС здійснюється перевірка, чи знаходяться значення ЧОТ в «коридорі», що відповідає стану авіадиспетчера «норма» (спокійний стан) – вихід значень ЧОТ за межі встановленого «коридору» свідчить про те, що авіадиспетчер знаходиться в неналежному ПФС.

Дослідження показали, що зміни ЧОТ свідчать як про зміну емоційного стану диспетчера, так і про настання станів втоми, сонливості і тривоги. Зауважимо при цьому, що стани втоми, сонливості і тривоги, на відміну від емоційного стану, не піддаються контролю з боку людини і не можуть змінюватися за її бажанням.

IV. Способи підвищення швидкодії та достовірності роботи систем підтримки прийняття рішень в диспетчерських службах аеропортів

А. Спосіб підвищення швидкодії СІППР

Простий перебір мовних фрагментів при використанні ШНМ є непродуктивним – при його застосуванні не може бути забезпечена робота підсистеми моніторингу ПФС в режимі реального часу.

Авторами розроблений алгоритм пошуку мовних фрагментів у безперервному мовленні, застосування якого дозволяє значно прискорити процес пошуку. Підвищення швидкодії роботи підсистеми досягається шляхом швидкого відкидання мовних фрагментів, моделі яких апріорі не входять до складу словника ключових мовних фрагментів, перед їх обробкою підсистемою моніторингу ПФС. Застосування ШНМ стає необхідним лише у випадку перевірки мовних фрагментів, відібраних в процесі попереднього відбору.

Значущим етапом при застосуванні розробленого способу підвищення швидкодії роботи підсистеми є складання моделей мовних фрагментів. Зазначені моделі являють собою послідовності позицій, призначених для розпізнаних фонем; при цьому, під розпізнаванням розуміється як власне розпізнавання фонем (в першу чергу, голосних), так і віднесення фонеми до певного класу (наприклад, класу глухих приголосних).

База моделей ключових мовних фрагментів створюється на етапі реєстрації диспетчерів. При застосуванні розробленого способу відкидаються мовні фрагменти, моделі яких не відповідають моделям, що складають базу моделей мовних фрагментів. Важливим є відсутність необхідності розпізнавання всіх фонем ключового слова – досить розпізнати (або віднести до певного класу) кілька фонем – їх послідовність складе гіпотезу, що підтверджується або відкидається ШНМ.

Таким чином, реалізація розробленого способу підвищення швидкодії роботи підсистеми моніторингу ПФС полягає у виконанні такої послідовності дій:

1. Складання моделі контрольного мовного фрагмента, побудованої за першими з розпізнаних фонем мовного фрагмента.

2. Пошук в базі моделей мовних фрагментів тих моделей, які відповідають мовному фрагменту, що надійшов для аналізу. У разі наявності в базі моделей відповідної моделі висувається гіпотеза, що аналізований мовний фрагмент є в словнику.

3. Перевірка гіпотези із застосуванням ШНМ.

4. При позитивному результаті перевірки гіпотези – проведення моніторингу ПФС авіадиспетчера.

Ефективність способу виділення ключових мовних фрагментів з безперервного злитого мовлення залежить від якості їх сегментації на фонеми, яку пропонується проводити із застосуванням вейвлетів за методом, який описаний в [6] стосовно шумозаглушення.

Б. Способи підвищення достовірності роботи СІППР

1. Для підвищення достовірності роботи СІППР доцільно збільшувати кількість інформативних параметрів, що аналізуються.

В якості інформативних пропонується застосовувати такі параметри ключових мовних фрагментів: ЧОТ і пов'язані з нею параметри (порізаність мелодійного контуру, дисперсія, середнє значення та ін.), формантні частоти і їхні відношення, тривалість проголошення ключових мовних фрагментів.

Проведені теоретичні та експериментальні дослідження, метою яких був аналіз параметрів, що характеризують мовні фрагменти, з точки зору їх ефективності для визначення ПФС, показали, що найбільш інформативні є друга, третя і четверта формантні частоти, а також відношення значень цих формантних частот, отриманих під час моніторингу, до їх значень, виміряних при знаходженні диктора в стані норми.

2. Для підвищення достовірності роботи підсистеми виділення ключових мовних фрагментів треба застосовувати ШНМ з декількома виходами. На кожному виході ШНМ розраховується вектор розмірності n, і-ий елемент якого є імовірнісним значенням того, що процедуру виявлення ключових мовних фрагментів проходить конкретний (і-ий) мовний фрагмент, параметри якого занесені до відповідної бази даних. Природно, застосування ШНМ з декількома виходами супроводжується ускладненням процесу навчання ШНМ.

3. Подальше підвищення достовірності визначення ступеня втоми, сонливості і тривоги авіадиспетчера може бути досягнуто на основі введення контролю стану його серцево-судинної системи під час передзмінного медико-психофізіологічного контролю та перерв в роботі. Оцінити стан серцево-судинної системи можна за пульсом і параметрами електрокардіограм, знятих з використанням стаціонарних та портативних електрокардіографів.

V. Висновки

1. Розроблено принципи побудови СІППР для забезпечення прийняття обґрунтованих рішень керівниками та операторами авіабудівних підприємств. СІППР розробляються на основі інтелектуальних ІТ, що дає можливість планувати оперативні управлінські дії в умовах неповноти та невизначеності інформації, що циркулює та зберігається на підприємствах.

2. Розроблено принципи побудови СІППР щодо доступу операторів та диспетчерів до інформаційних ресурсів ергатичних систем, реалізація якої дозволяє здійснювати моніторинг ПФС операторів і диспетчерів в режимі реального часу протягом всієї робочої зміни, що істотно знизить вплив людського фактора на ефективність роботи та інформаційну безпеку авіапідприємств.

3. Розроблено способи підвищення достовірності роботи і швидкодії СІППР в диспетчерських службах аеропортів, застосування яких дозволяє здійснювати моніторинг ПФС авіадиспетчерів в режимі реального часу з високою достовірністю (відсоток правильного розпізнавання вище 98%).

Література References

[1] НД ТЗІ 2.5-005-99 Класифікація автоматизованих систем та стандартні функціональні профілі захищеності оброблювальної інформації від несанкйіонованого доступу.

[2] НД ТЗІ 1.1-003-99 Термінологія в галузі захисту інформації в комп'ютерних системах від несанкціонованого доступу.

[3] В.А. Темников, Е.Л. Темникова, Параметризация автоматического контроля доступа операторов к ресурсам информационных систем по голосу // Вестник Восточноукраинского национального университета им.В.Даля. - №9 (151). – Ч.1. – 2010. – С.143-148.

[4] Г.С. Рамишвили, Автоматическое опознавание говорящего по голосу // М.: Радио и связь, 1981. – 224 с.

[5] Л. Рабинер, Б. Гоулд, Теория и применение цифровой обработки сигналов // М: Мир, 1978. – 848 с.

[6] В.А. Темников, Л.В. Пономаренко, Методика проведения шумоочистки речевого сигнала в процессе распознавания // Вестник Восточноукраинского национального университета им. В.Даля. - №5 (111). – Ч.1. – 2007. – С.123-127.

Відмова від Шаблонів і Перезапис Термів

Для моделювання робочих бізнес-процесів в Maude

Bartosz Zieliński
Факультет комп'ютерних наук
Університет Лодзі
Поморська 149/153, 90-236 Лодзь, Польща
bzielinski@uni.lodz.pl

Paweł Maślanka
Факультет комп'ютерних наук
Університет Лодзі
Поморська 149/153, 90-236 Лодзь, Польща
pmaslan@uni.lodz.pl

Cancellation Patterns and Term Rewriting

For Business Workflow Modeling in Maude

Bartosz Zieliński
Department of Computer Science
University of Łódź
Pomorska 149/153, 90-236 Łódź, Poland
bzielinski@uni.lodz.pl

Paweł Maślanka
Department of Computer Science
University of Łódź
Pomorska 149/153, 90-236 Łódź, Poland
bzielinski@uni.lodz.pl

***Анотація*—Зміна рангу процесів, на відміну від більшості інших моделей робочих процесів, не локальна справа. Хоча концептуально задача проста, вона, як відомо, важко моделюється з Петрі мережами - це математичний апарат, найбільш часто використовується для забезпечення семантики Workflow конструкції. В роботі показано, що використання перезапису, настільки ж ефективне, як мережі Петрі для моделювання паралельних і розподілених систем, з екваціональними специфікаціями і використовується з не локальними змінами.**

***Abstract*—Cancellation of a range of activities is, unlike most of the other workflow patterns, a non-local affair. While conceptually simple, it is notoriously difficult to model with Petri Nets – a mathematical formalism most commonly used to provide semantics to workflow constructs. Here we demonstrate that combining term rewriting, equally useful as Petri Nets to model paralell and distributed systems, with equational specifications allows dealing naturally with non-local changes. To this end, we implement in the term rewriting system Maude a toy workflow simulation framework (essentially based on token passing, similarly as in the case of Petri Nets) and then augment it with a YAWL-like cancellation region construct..**

Ключові слова— термін перезапису; бізнес-процес; робочий процес; Maude

Keywords— term rewriting; business process; workflow; Maude;

I. INTRODUCTION

Cancellation of a range of activities [1] is, unlike most of the other workflow patterns, except perhaps for a XOR split [2], a non-local affair. While conceptually simple, it is notoriously difficult [1] (though not impossible) to model with Petri Nets – a mathematical formalism most commonly used to provide semantics to workflow constructs. It requires simultanous removal of tokens from an arbitrary subset of places by a cancelling transition, and thus may require an exponential increase of complexity of the net. Using naturally distributed formalisms such as pi-calculus to model workflows (as in [3]) leads to another kind of problem with modeling cancellation patterns. Namely, the issue is with *simultanous* cancelling of a set of activities, which is obviously difficult in any system based on message passing.

Here we demonstrate that combining term rewriting, equally useful as Petri Nets to model paralell and distributed systems, with equational specifications allows dealing naturally with non-local changes. To this end we implement in the term rewriting system Maude a toy workflow simulation framework (based on token passing, similarly as in the case of Petri Nets) and then augment it with a YAWL-like cancellation region construct [1].

II. PRELIMINARIES ON MAUDE AND TERM REWRITING

Maude [4] is a language and execution system based on term rewriting [5], [6] and many sorted equational logic [7]. In

this section we recall some basic mathematical definitions to fix the notation and to make the presentation self-contained.

An algebraic signature $\Sigma=(\Sigma_S, \Sigma_F)$ consists of a finite poset of sorts Σ_S (sorts are like type names, and the order corresponds to subtyping) and a finite set of function signatures Σ_F . Each function signature is of the form $f : s_1 s_2 \dots s_n \rightarrow s$, where f is a function symbol and s_i's are sorts in Σ_S Symbols c such that $c : \rightarrow s \in \Sigma_F$ are called *constants* of sort s..

A Σ-algebra A is an assignment of a set $[[s]]_A$ to each sort s in Σ_S and a function (interpretation) $[[f]]_A : [[s_1]]_A \times \dots \times [[s_n]]_A \rightarrow [[s]]_A$ to each function $f : s_1 s_2 \dots s_n \rightarrow s$. It is required that $[[s_1]]_A \subseteq [[s_2]]_A$ whenever $s_1 \leq s_2$. We call x an element of A *if* $x \in [[s]]_A$ for some $s \in \Sigma_S$.

Let $V:=\{V_s | s \in \Sigma_S\}$ be a collection of sets where elements of V_s are variables of sort s. A term algebra $T_\Sigma(V)$ has "sort-safe" terms as elements and function symbols interpreted by themselves. We denote by T_Σ the algebra of ground Σ-terms.

In Maude, the poset of sorts Σ_S in the signature is implicitly augmented with separate top elements $[K]$ (referred to as kinds) for connected components K of Σ_S. We denote $[s]:=[K]$ for any $s \in K$. Giving a function symbol arguments of a wrong kind is treated as a syntax error. Giving a function symbol arguments of a wrong sort (but correct kind) is legal, but makes the result member of a kind but not of a sort.

Maude supports the so-called mixfix syntax - underscores in the function name correspond to consecutive arguments. Thus if Σ_F contains $_+_ : Nat\ Nat \rightarrow Nat$ and $_?_ : Bool\ Nat\ Nat \rightarrow Nat$ we can use expressions such as 1+2 or *false ? 1 : 2*.

Maude allows to define Σ-algebras as quotients of term algebras by a congruence generated by a pair (A,E) of sets of equalities. Equalities in A, specified as attributes on function symbols, define certain structural properties of binary operators such as associativity, commutativity, idempotence or unitality. E consists of actual conditional equalities interpreted as directed simplification rules. It is assumed that simplifications terminate and each term has the unique canonical (irreducible) form modulo properties from A. Maude represents equivalence classes in $T_\Sigma / \equiv_{A,E}$ with irreducible elements of $T_\Sigma / \equiv_A$. Conditions in conditional equalities are conjunctions of unconditional equalities and membership axioms.

Reductions with respect to equalities can be viewed as computations of values and as such they represent denotational aspect of the problem. The behavioural aspect is represented through rewritings. Single step rewrites are defined with a set of conditional rewriting rules of the form $T_1 \Rightarrow T_2$ *if* C, where C is a conjunction of equalities, membership axioms and rewriting axioms. A rewriting axiom $K_1 \Rightarrow K_2$ is satisfied if K_1 can be rewritten in one or more steps to K_2 .

A rewriting system consists of a signature, a set of equalities and equational attributes as well as a collection of rewriting rules. Definitions of rewriting systems are collected in Maude modules, either functional (which cannot include rewriting rules and simply define quotient algebras) or system ones which can include rewriting rules. Modules can be parametrized by theories which define properties of classes of systems with which the module can be instantiated.

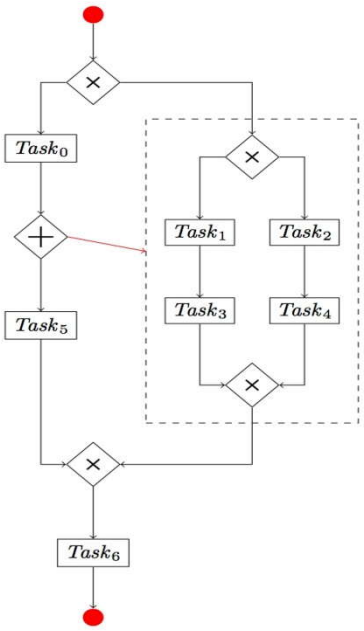


Fig. 1. Example workflow with cancellation region

III. CANCELLATION REGIONS

Cancellation refers to aborting activations of a selection of tasks within the workflow in response to some event, activity or decision by the user. As an example consider the process of preparing the travel in which separate threads of activities devoted to booking a hotel, airplane ticket, etc., might need to be simultaneously aborted when the user changes plans. An abstract example is presented in Figure 1. There, when the red path of execution is taken at the XOR decision gate, the region of the workflow enclosed in dashed lines gets cancelled, which means that whichever of the four tasks are active (we have four possibilities, e.g., {Task1, Task2}, {Task1, Task4}, etc.) becomes inactive. In case of workflow system based on Petri Nets this might be implemented through removing tokens from places inside the cancelled region, which requires special inhibitor edges (see. e.g., [1]), or an additional subnet with complicated topology. In what follows we will demonstrate how to implement cancellation within rewriting models for workflow systems. We will show this within a context of a toy rewriting theory for a workflow presented in the next section. However, it should be noted that the method is completely general.

IV. A TOY REWRITING THEORY FOR WORKFLOWS

Our toy rewriting theory will allow us to model workflows with the following elements: tasks (no distinction is made between kinds of tasks, such as user, automatic, manual, etc.), XOR and AND joins and splits, and, finally, end nodes and the single start node. Late we augment the formalism with cancellation regions.

We start by defining syntax. We represent state of the workflow as an ensamble of sort *Config* of workflow items of sort *Item* collected together using the associative and commutative (empty) operator __ . The relevant definitions are as follows:

```
sorts  Item Config .
subsort Item < Config .
op nil : -> Config .
op __ : Config Config -> Config
        [assoc comm id: nil] .
```

Note that due to the subsort declaration a single *Item* is also a *Config*. Note also the neutral element *nil* of the "__" operator. Now we define the constructors for various *Item*s constituting the state of the workflow. Each such item (except for the start node) is identified by an identifier of sort *Id* and has a state of sort *State.* Possible states of items are "ready", "active" and "next". Identifiers can be given as values of a predefined sort Qid (quoted identifiers) consiting of strings of charactes which are valid Maude identifiers preceded by a single quotation mark (e.g., 'q). Later we will also define additional identifier constructors. All other attributes of items, unique to each kind of item (and distinguishing between different kinds of them) are gathered together as terms of sort *Part.* All items, except for the start node are constructed as triples of identifier, state, and ``the rest'' with the constructor [_,_,_]:

```
sorts Id State Part Op. subsort Qid < Id.
ops ready active next : -> State .
ops xor and : -> Op .
op start : State Id -> Item .
op [_,_,_] : Id State Part -> Item  .
ops conn task :  Id -> Part .
op end : -> Part .
ops split : Id Id Op -> Part .
ops join  : Id Op -> Part .
```

We define six kinds of items: the start node and the remaining five kinds of nodes distinguished by *Part*s: end items, task, connector (conn), and, finally, split and join gates. It is assumed that tasks, gates, start and end nodes connect together with connectors. Connectors have one auxilliary attribute: the identifier of the next workflow object which the connector connects to and which is to be activated by the given connection object. The usage of the explicit connector objects simplifies the rules, as now items of other kinds are activated by and activate only connectors. This allows us to avoid the combinatorial explosion caused by considering each combination of types of consecutive items.

The start node does not have its own identifier (it is unique), it does, however, have state, as well as the identifier of the unique connector to be activated next. End nodes have no auxilliary attributes. Each task node contains identifier of the unique connector to be activated next. Both join and split gates contain identifiers of their operation of sort *Op* (either "and" or "xor"). In addition, each join (resp. split) gate contains the unique identifier of the next connector (resp. the identifiers of two next connectors.

V. REWRITE RULES OF A TOY MODEL

The rewrite rules define the operational semantics of our toy workflow model and allow it to be actually executed (for instance to test its correctness). The general idea is as follows: There are three states the workflow can be in: "ready", "active" and "next", although only tasks pass through all three. The remaining components use only "ready" and "next". The ready component can pass into "next" or "active" states when its input connectors are in the "next" state. A given component can make all or some of its output connectors transition to the "next" state when it is itself in the "next" state. The additional "active" state of tasks simulates the task being executed (perhaps for an extended amount of time). The rules are as follows. First we define variables to use in the rules:

```
vars X X' Y Y' Z : Id . var P : Part .
var S : State . var O : Op .
```

Then come the rewriting rules. For example, the following rules allows any task to transition by itself from the active state into the "next" state:

```
rl [X, active, task(Y)]
            => [X, next, task(Y)] .
```

The following rules describe the activation of tasks by an input connector and the activation of an output connector by a task:

```
rl [X, next, conn(Y)] [Y, ready, task(Z)]
    => [X, ready, conn(Y)] [Y, active,
          task(Z)] .
rl [X, next, task(Y)] [Y, ready, conn(Z)]
     => [X, ready, task(Y)] [Y, next,
           conn(Z)] .
```

As the last example consider the following rules which deal with activation of output connectors of split gates:

```
rl [Z, next, split(X, Y, and)] [X, ready,
   conn(X')] [Y, ready, conn(Y')] =>
     [Z, ready, split(X, Y, and)]
[X, next, conn(X')] [Y, next, conn(Y')] .
rl [Z, next, split(X, Y, xor)] [X, ready,
conn(X')] => [Z, ready, split(X, Y, xor)]
[X, next, conn(X')] .
rl [Z, next, split(X, Y, xor)] [Y, ready,
conn(Y')] => [Z, ready, split(X, Y, xor)]
[Y, next, conn(Y')] .
```

The END split activates both of its output rules. In case of a XOR split gate there are two rules, one for the activation of each of the output connectors. Thus, a XOR split may activate either one of its two outputs, depending on which of these rules gets chosen by the system. We omit the rest of the rules as they are similar.

VI. IMPLEMENTATION OF THE CANCELLATION REGIONS

We implement the cancellation regions utilizing the fact that Maude (except for frewrite strategy) will attempt to reduce term to the canonical form after each rewrite step. Thus, we can implement cancellation through equational reduction, which, from the point of view of rewriting, happens immediately. More precisely, we introduce a special workflow item which holds the list of identifiers of workflow objects to be cancelled. Cancellation in this contex means changing the state of all objects in the cancellation region to "ready". The relevant code is described below.

First we need operators and sorts for lists of identifiers:

```
sort IdList . subsort Id < IdList .
op nil : -> IdList .
op __ : IdList IdList -> IdList
        [assoc id: nil] .
```

The following operator represents cancellation regions in the workflow:

```
op cancel : Id IdList IdList -> Item .
```

The first argument is the identifier of the cancellation region. The remaining two are the list of identifiers of the objects within the region and the list of identifiers of workflow objects the activity of which currently remains to be cancelled. When the cancellation region is inactive the value of this second list is *nil.* In fact, both lists do not contain identifiers of connectors which are treated separately. The cancellation region can be activated through the connector object as defined with the following rule:

```
rl [X, next, conn(Y)] cancel(Y, IL, nil)
=> [X, ready, conn(Y)] cancel(Y, IL, IL).
```

Note that this rule applies only if the last argument of *cancel* is equal to nil. The main workhorse is the following equation:

```
eq cancel(X, IL, Y IL') [Y, S, P] =
    cancel(X, IL, IL') [Y, ready, P]
         ccon(Y, cnt(P)) .
```

Note that the start node cannot be the part of cancellation region. All other workflow objects in our toy model are matched by the pattern [Y, S, P]. As the lists of identifiers in *cancel* are not supposed to contain connectors' identifiers we have to deal with them separately. To this end, the equality above introduces a workflow item constructed with operator `ccon : Id Nat -> Item` which is supposed to de-activate all input connectors for workflow object Y:

```
eq [X, S, conn(Y)] ccon(Y, s N) = [X,
ready, conn(Y)] ccon(Y, N) .
eq ccon(Y, 0) = nil .
```

The second argument of *ccon* is the number of input connectors which remains to be de-activated. Its initial value is the number of incoming connectors computed for each workflow item type using the function:

```
op cnt : Part -> Nat .
eq cnt(join(X, O)) = 2 .
eq cnt(conn(X)) = 0 .
eq cnt(P) = 1 [owise] .
```

The *owise* attribute means ``in all other cases". Thus, the number of incoming connectors is 2 for join gates, 0 for connectors, and 1 in all other cases.

VII. Example

Consider the example of a workflow in Fig. 1. It can be represented in our formalism as a term

```
start(next, c(0)) ['e1,ready,end]
['g1,ready,split(c(1), c(2), and)]
['g2,ready,split(c(4), c(5), xor)]
['g3,ready,join(c(7), and)]
cancel('c1, 'g4 'g5 't1 't2 't3 't4, nil)
*** etc.
```

Here we have used a separate constructor for connector identifiers: `op c : Nat -> Id .` After some analysis, one sees hat the workflow is flawed. Indeed, if the cancellation region is invoked, then at least the left one of the branches of the last AND join gate is never invoked, and so the end node is never activated. We can easily confirm that by searching (with the search command) for irreducible (with respect to rewritings) terms reachable from the start state in which end is not activated (c is assigned to the initial workflow description):

```
search c =>! C:Config ['e1, ready, end] .
```

and verifying that Maude indeed finds several solutions.

VIII. Conclusion

We have presented a toy rewriting model for workflows and we have extended it with cancellation regions implemented through equations. While the toy model is not very useful in itself, the method for extending models with cancellations is very general and does not depend on the details of our toy model. It shows that cancellation patterns can be naturally implemented within the framework of rewriting systems.

References

[1] Van Der Aalst, Wil MP, and Arthur HM Ter Hofstede. "YAWL: yet another workflow language." Information systems 30.4 (2005): 245-275.

[2] van Der Aalst, Wil MP, et al. "Workflow patterns." Distributed and parallel databases 14.1 (2003): 5-51.

[3] Puhlmann, Frank, and Mathias Weske. "Using the π-calculus for formalizing workflow patterns." International Conference on Business Process Management. Springer Berlin Heidelberg, 2005.

[4] Clavel, Manuel, et al. "The maude 2.0 system." International Conference on Rewriting Techniques and Applications. Springer Berlin Heidelberg, 2003.

[5] Meseguer, José. "Conditional rewriting logic as a unified model of concurrency." Theoretical computer science 96.1 (1992): 73-155.

[6] Meseguer, José, and Grigore Roşu. "The rewriting logic semantics project." Theoretical Computer Science 373.3 (2007): 213-237.

[7] Meseguer, José. "Membership algebra as a logical framework for equational specification." Recent Trends in Algebraic Development Techniques (1998): 18-61.

[8] Clavel, Manuel, et al. All about maude-a high-performance logical framework: how to specify, program and verify systems in rewriting logic. Springer-Verlag, 2007.

Швидка Класифікація Статі Використовуючи Геометричні Характеристики і SVM

Piotr Milczarski
Факультет комп'ютерних наук
Університет Лодзі
Поморська 149/153, 90-236 Лодзь, Польща
piotr.milczarski@uni.lodz.pl

Zofia Stawska
Факультет комп'ютерних наук
Університет Лодзі
Поморська 149/153, 90-236 Лодзь, Польща
zofia.stawska@uni.lodz.pl

Fast Gender Classification Using Geometrical Features and SVM

Piotr Milczarski
Faculty of Physics and Applied Informatics
University of Lodz
Pomorska str. 149/153, Lodz, Poland
piotr.milczarski@uni.lodz.pl

Zofia Stawska
Faculty of Physics and Applied Informatics
University of Lodz
Pomorska str. 149/153, Lodz, Poland
zofia.stawska@uni.lodz.pl

Abstract— In the paper, we have concentrated on the problem of fast gender classification using SVM method for mobile applications. The mobile applications can have the same functionality as the desktop ones, but mobile devices have various hardware and OS specification. Often this specification is worse than for the desktop ones, so we were looking for the possibly simplest classifier giving satisfactory classification efficiency. Many authors have used SVM in the facial classification and recognition problem, but there are not many works using the facial geometry features in the SVM classification. Almost all works are based on the appearance-based methods. Choosing a minimum set of features, that can give satisfactory results, it is important problem too. The minimum set of features can simplified classification process to make it useful for the mobile applications. In the paper, we show that the classifier construct on the base of only two geometric facial features can give satisfactory (though not always optimal) results.

Анотація—розглянуть проблеми швидкої гендерної класифікації з використанням методу SVM для мобільних додатків. Мобільні додатки можуть мати таку ж функціональність як настільні ПК, але мобільні пристрої мають різні апаратні і технічні характеристики ОС. В роботі показано, що класифікатор конструкція на базі тільки двох геометричних рис осіб може дати задовільні (хоча і не завжди оптимальні) результати.

Keywords—face identification, biometrics, gender classification, Support Vector Machine, mobile applications.

Ключові слова—ідентифікації особи, біометрія, гендерна класифікація, підтримка векторних машин, мобільні додатки.

I. INTRODUCTION

The first step in the gender determination process it is finding face area on the image. Skin color pixels detecting can be used in this task. The area of the skin color pixels is checked whether it can be classified as a face e.g. using template matching methods. The several approaches to the skin color classification were presented in our previous papers [1][2]. When the face area is finding we can decide about method of the gender detection.

There are many gender classifications methods, which can be used in the recognition process. They can be divided into two groups: feature-based and appearance-based methods [3][26].

The first one requires finding the facial characteristic points as nose, mouth, eyes, ears or hair, called fiducial points. The geometric relation between these points (fiducial distances) are used as a feature vector in the classification process. Importance of these distances in the gender discrimination task is confirmed by the psychophysical studies [4][10][22-24].

Appearance based methods are based on the values of image pixels that were previously transformed on the local or global level, e.g. at the local level, the image can be divided into lower windows or specific face regions such as mouth, nose or eyes. This approach preserves natural geometric relationships which can be used as naïve features. This approach can be very computationally demanding because of very large number of features (each pixel is treated as a feature).

In our research, we decided to use geometric face features to limit computational complexity.

II. Gender Recognition Approaches

The other important element in the gender recognition process is the choice of the proper classification method. In this task many different classification methods are used, e.g.:

neural network [7],

radial basis function networks [5],

Gabor wavelets [6],

Adaboost [8],

Support Vector Machines (SVM) [9],

Bayesian classifiers [9].

Support Vector Machine (SVM) is one of the most popular classification method, often used by authors, because of good classification results [3][26]. SVM is one of the strongest classification methods useful in many applications [19-21]. This approach can be helpful also in the process of gender classification. Many authors reports very good efficiency of this method [9][15-18].

The SVM algorithm is based on subset of training samples called Support Vectors, which are used to construct the decision function [19-21]. Support Vectors are the data points that lie closest to the decision surface, which are using to maximize hyperplane margin. If the dataset in the original feature space is not linearly separable it can be mapped to a higher dimensional feature space where the same dataset is likely separable.

The SVM provides non-linear kernel function approximations by mapping the input vectors into a high dimensional feature space where a linear hyperplane can be constructed. In our research we were using Gaussian kernel as the one gives best results.

III. Facial Databases

To train a classifier we always need a set of object examples. We can prepare this dataset ourselves or use existing database. For the gender recognition problem we need the dataset of the face photos. There are several publicly available database that have been used for experiments. The most popular is the FERET database [25]. Publicly available datasets examples are AR, BioID, CAS-PEAL-R1, MORPH-2, LFW.

In our research we decided to use a part of AR face database [12] containing frontal facial images without expressions. The AR face database was prepared by Aleix Martinez and Robert Benavente in the Computer Vision Center (CVC) at the U.A.B. It contains:

over 4,000 color images;

126 people's faces of 70 men and 56 women;

images show frontal view faces with different illumination conditions and occlusions (sun glasses and scarfs);

the pictures were taken at the CVC under strictly controlled conditions;

images are of 768×576 pixel resolution and of 24 bits of depth.

We have chosen AR database subset containing 92 frontal face images: 49 women and 43 men.

IV. Results of Classification

In our research we took into account 11 facial characteristic points (Fig.1):

RO - right eye outer,

RI – right eye inner,

LI - left eye inner,

LO - left eye outer,

RS and LS – right and left extreme face point at eyes level,

MF- forehead point – in the direction of facial vertical axis defined as in [13] or in [24], see Fig.2

M – nose bottom point,

MM – mouth central point,

MC – chin point,

Oec, the anthropological face point, that has coordinates derived as an arithmetical mean value of the points RI and LI.

Points were marked on each image manually. These features were described in [24][13][14] and are only a part of facial geometric features described in [10]. The coordinates are bounded with Oec point and their values are recalculated in Muld units [13][24], where

- $1\ \text{Muld} = 10 \pm 0.5\ \text{mm}$ (1)

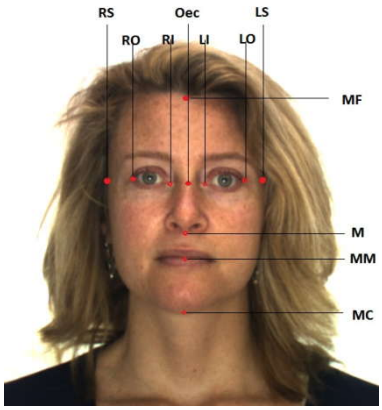


Fig. 1. Face characteristic points [13][14] (image from AR database)

The chosen points allow us to define 7 distances which are used as a features in the classification process:

MM – distance between anthropological point and mouth center.

MC – distance between anthropological point and chin point.

MC-MM – chin/jaw height.

MC-M – distance between nose-end point and chin point.

RSLS – face width at eye level.

ROLO – distance between outer eye corners.

MF-MC – face height.

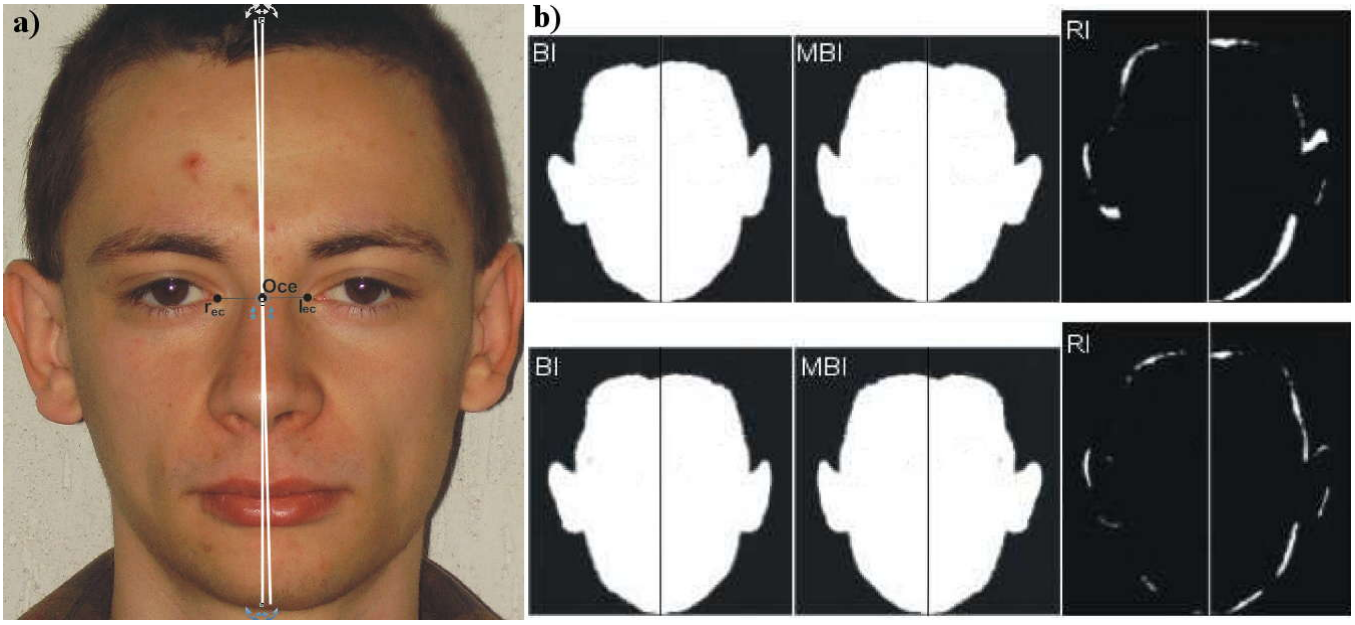


Fig. 2. Example of derivation of vertical projection axis [13]

In our experiments, we test classification efficiency using the subsets taking from the set of features described above. We look for a minimal feature subset that give the best classification results. The classification efficiency is the ratio of correctly classified test examples to the total number of test examples. We can train and test results on the different datasets, but we decided to use cross validation as a method of result testing, because our subset of AR database was not enough big to divide it into two (training and testing) subsets.

TABLE I. RESULTS FOR 2 FEATURE SETS

Features	MC	MC-MM	MC-M	RSLS	ROLO	MF-MC
MM	**72,5%**	**76,3%**	**72,5%**	**76,3%**	**73,8%**	**70,0%**
MC		**72,5%**	**75,0%**	**76,3%**	**82,5%**	**70,0%**
MC-MM			**67,5%**	**71,3%**	**53,8%**	**70,0%**
MC-M				**77,5%**	**77,5%**	**73,8%**
RSLS					**68,8%**	**73,8%**
ROLO						**65,0%**

The results of experiments presented in Table I have shown the classification error. Cross validation method was used to verified the classification results. Eight sets of 10 objects (5 women and 5 men) were used. We have tested all combinations of our 7 features. For large subsets of features we have obtained comparable results as for the best two-element sets. The best results for two-features sets we obtain using distance between eye outers (ROLO) and distance from anthropological point to the chin point (MC). For this pair error rate is 17,5 %.

It was also interesting to find out whether the two genders are evenly recognized, i.e. whether the number of misclassified women is comparable to the number of males who are misclassified. The results are shown in Table II.

We can see that for best feature pair we have visible disproportion between classification error for men and women. Women were classified more efficiently. For different feature pairs we have opposite situation men were misclassified. It suggest that some features can be characteristic for one of the gender.

V. CONCLUSIONS

It the paper, we focused on the use of SVM classifier in the gender recognition application, especially applications dedicated to mobile devices. Mobile applications may have specific requirements, often offering less computing power than desktop applications. Considering that we were looking for a classifier based on the possibly smallest set of features that will be give satisfactory classification results. We showed that it is possible to construct such classifier, using only 2 features. We obtain classification efficiency 82,5%. We showed also that the number of misclassified object depends on gender. In the feature we are going to propose classifier taking into account this relationship.

TABLE II. INDEPENDENT RESULTS FOR MALE AND FEMALE FOR 2 FEATURE SETS

Features	2 MC		3 MC-MM		4 MC-M		5 RSLS		6 ROLO		7 MF-MC	
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
1MM	72,5%	72,5%	80,0%	72,5%	80,0%	65,0%	82,5%	70,0%	67,5%	80,0%	52,5%	87,5%
2 MC			85,0%	60,0%	67,5%	82,5%	75,0%	77,5%	75,0%	90,0%	52,5%	87,5%
3 MC-MM					52,5%	82,5%	70,0%	72,5%	90,0%	17,5%	85,0%	55,0%
4MC-M							80%	75%	80,0%	75,0%	80,0%	67,5%
5RSLS									77,5%	60,0%	80,0%	67,5%
6 ROLO											87,5%	42,5%

REFERENCES

[1] Stawska Z., Milczarski P. (2013) *Algorithms and methods used in skin and face detection suitable for mobile applications*, ISIM: ISSN: 2084-5537, Vol.2, No 3, pp. 227-238

[2] Milczarski P., Stawska Z. (2014) *Complex colour detection methods used in skin detection systems,* ISIM: ISSN: 2084-5537, Vol.3, No 1, pp. 40-52

[3] Milczarski P., Stawska Z. (2016) *An automatic face detection and gender classification,* "Information Technologies and Computer Modelling" proceedings of the International Scientific Conference 2016, May, 23th to 28th, Ivano-Frankivsk

[4] Brunelli R., Poggio T. (1993) *Face recognition: features versus templates*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 15, no. 10, pp. 1042-1052

[5] Abdi H., Valentin D., Edelman B., O'Toole A.J. (1995) *More about the difference between men and women: Evidence from linear neural network and principal component approach*, Neural Comput. 7 (6), 1160–1164.

[6] Wiskott L., Fellous J. M., Krüger N., von der Malsburg C. (1997) *Face recognition by elastic bunch graph matching*, In: Sommer, G., Daniilidis, K., Pauli, J. (Eds.), 7th International Conference on Computer Analysis of Images and Patterns, CAIP'97, Kiel. Springer-Verlag, Heidelberg, pp. 456–463.

[7] Cottrell G.W., Metcalfe J. (1990) *EMPATH: Face, emotion, and gender recognition using holons,* In: Lippmann, R., Moody, J.E., Touretzky, D.S. (Eds.), Proc. Advances in Neural Information Processing Systems 3 (NIPS). Morgan Kaufmann, pp. 564–571.

[8] Shakhnarovich G., Viola P.A., Moghaddam B. (2002) *A unified learning framework for real time face detection and classification.* In: Proc. Internat. Conf. on Automatic Face and Gesture Recognition (FGR'02). IEEE, pp. 14–21.

[9] Sun, Z., Bebis, G., Yuan, X., Louis, S.J.(2002). *Genetic feature subset selection for gender classification: A comparison study*. In: Proc. IEEE Workshop on Applications of Computer Vision (WACV'02), pp. 165–170.

[10] Fellous J.M. (1997) Gender discrimination and prediction on the basis of facial metric information ,Vision Research, vol. 37, no. 14, pp. 1961–1973.

[11] Muldashev E.R. (2002) *Whom Did We Descend From?* OLMA Press, Moscow (In Russian).

[12] Martinez A.M., Benavente R. (1998) *The AR Face Database.* CVC Technical Report #24.

[13] Milczarski P., Kompanets L., Kurach D. (2010) *An Approach to Brain Thinker Type Recognition Based on Facial Asymmetry* in L. Rutkowski et al. (Eds.): ICAISC 2010, Part I, LNCS 6113, pp. 643--650. Springer, Heidelberg.

[14] Kompanets L., Milczarski P., Kurach D. (2013) *Creation of the fuzzy three-level adapting Brainthinker* in Human System Interaction (HSI), 2013 The 6th International Conference on Digital Object Identifier: 10.1109/HSI.2013.6577865, pp. 459 – 465.

[15] Moghaddam B., Yang M.H. (2002) *Learning gender with support faces,* Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 24, no. 5, pp. 707–711.

[16] Castrillon M., Deniz O., Hernandez D., Dominguez A. (2003) *Identity and gender recognition using the encara real-time face detector* in Conferencia de la Asociacin Espaola para la Inteligencia Artificial, vol. 3.

[17] Buchala S., Loomes M.J., Davey N., Frank R.J. (2005) The role of global and feature based information in gender classification of faces: a comparison of human performance and computational models International Journal of Neural Systems, vol. 15, pp. 121- 128.

[18] Jain A., Huang J., Fang S. (2005) *Gender identification using frontal facial images* in Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on, p. 4.

[19] Boser B. E., Guyon I. M., Vapnik V. N. (1992) *A training algorithm for optimal margin classifiers.* Proceedings of the fifth annual workshop on Computational learning theory – COLT '92. p. 144.

[20] Cortes C., Vapnik V. (1995) *Support-vector network.* Machine Learning. 20 (3): 273–297.

[21] Vapnik V. N., Kotz S. (2006) *Estimation of Dependences Based on Empirical Data*, Springer, ISBN 0-387-30865-2.

[22] Burton A.M., Bruce V., Dench N. (1993) What's the difference between men and women? Evidence from facial measurements, Perception, 22, pp. 153–176.

[23] Brunelli R., Poggio T. (1992) *Hyperbf networks for gender classification*, DARPA Image Understanding Workshop, pp. 311–314.

[24] Milczarski P. (2011) *A new method for face identification and determining facial asymmetry*, Semantic Methods for Knowledge Management and Communication, eds. R. Katarzyniak et al., seria Studies in Computational Intelligence vol. 381, ISBN: 978-3-642-23417-0, (pp. 329-340). Springer-Verlag Berlin Heidelberg.

[25] Phillips P.J., Moon H., Rizvi S.A., Rauss P.J. (2000) *The FERET evaluation methodology for face-recognition algorithms* Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 22, no. 10, pp. 1090-1104.

[26] Mäkinen E., Raisamo R. (2008) *An experimental comparison of gender classification methods* Pattern Recognition Letters 29, pp. 1544–56

Стохастичний Аналог Задачі Лінійного Програмування з Ефектом Галуження

Фахріддін Мірзоахмедов
кафедра математичного та інформаційного моделювання
Фінансово-економічний інститут Таджикістану
Душанбе, Таджикістан
mirfakh@mail.ru

Stochastic Analog Linear Programming Problem the Effect of Branching

Fakhriddin Mirzoahmedov
Department of Mathematical and Information Modeling
Finance and Economics Institute of Tajikistan
Dushanbe, Tajikistan
mirfakh@mail.ru

***Анотація*—У статті проаналізовано ефективний алгоритм симплексного методу для вирішення загальної задачі лінійного програмування, який також використовується при вирішенні стохастичного варіанту цієї ж задачі. Доведено збіжність алгоритму рішення стохастичного варіанту задачі з урахуванням ефекту розгалуження, що підтверджено тестовим прикладом.**

***Abstract*—The article provides an efficient algorithm simplex method for solving the general problem of linear programming, which is also used in solving stochastic version of the same problem. The convergence of the algorithm for solving the stochastic version of the problem with allowance for the branching effect is given and is confirmed by a test example.**

Ключові слова—детермінована і стохастична задача; розгалуження; квазіградієнт; алгоритм рішення; теорема збіжності; чисельний експеримент

Keywords—deterministic and stochastic task; branching; kvazihradient; solution algorithm; convergence theorem; numerical experiment

I. ВСТУП

В сучасних умовах з точки зору новизни та прикладного застосування, необхідними умовами математичного моделювання є наявність таких факторів, як невизначеність, нелінійність, неоднозначність, критичні точки, розгалужені процеси тощо.

У даній роботі досліджено стохастический аналог задачі лінійного програмування (ЗЛП) з урахуванням біфуркаційного ефекту та алгоритму її рішення.

II. ДЕТЕРМІНОВАНА ПОСТАНОВКА ЗЛП

Класична детермінована ЗЛП в канонічній формі представлена наступним чином [1]:

$$x_j^* = \arg\ \min\Big\{ \sum_{j=1}^{n} c_j x_j / x_j \in X \Big\} \to min, \qquad (1)$$

де

$$X = \Big\{ x_j : \sum_{j=1}^{n} a_{ij} x_j = b_i, b_i \ge 0, i = \overline{1,m} \Big\}. \qquad (2)$$

Параметри економічної постановки задачі можуть бути інтерпретовані наступним чином: n - кількість видів продукції, що випускається; b_i – запаси i-го виду ресурсів; a_{ij} - технологічні коефіцієнти, які вказують, скільки i-го ресурсу потрібно для виробництва продукту j-го виду; x_j - обсяг виробництва j-го продукту; c_j - витрати, пов'язані з виробництвом одиниці продукту j-го виду.

Нижче проаналізовано ефективний алгоритм вирішення задачі (1) - (2) симплексним методом [1], який використовується при наближених рішеннях стохастичної задачі на кожній ітерації.

Алгоритм 1. Завдання (1) - (2) вирішується за допомогою симплекс-методу, заданого оператором:

$$x_j^{s+1} = \psi(x_j^s), \quad x_j^0 = 0.$$

Тут x_j^0 - початкове наближення для симплекс-методу, що є базисним рішенням, вектори x_j^{s+1} і x_j^s являють собою наступне і попереднє рішення, тобто

$$x_j^{s+1} = \min_{x_{ij}^s}[x_i^s / x_{ij}^s] = x_l^s / x_{lk}^s,$$

$$x_i^{s+1} = x_i^s - (x_{ij}^s - x_j^{s+1}), \quad i \in I_{s+1}; \quad i \neq j,$$

$$x_j^{s+1} = 0, \quad j \notin I_{s+1}. \qquad (3)$$

Обчислення елементів нової симплекс-таблиці на $(s+1)$ -ій ітерації здійснюється згідно формули:

$$x_{ik}^{s+1} = \begin{cases} x_{ik}^s - (x_{ij}^s - x^s{}_{lk}) / x_{lj}^s, & якщо \quad i \neq l, \\ x_{lk}^s / x_{lj}^s, & якщо \quad i = l. \end{cases}$$

III. СТОХАСТИЧНА ПОСТАНОВКА ЗЛП

Додатково до розглянутих параметрів при постановці детермінованої ЗЛП (1) - (2) введемо наступні позначення: ω_j - незалежні випадкові величини, наприклад, попит на j-ту вироблену продукцію; α_i і β_i - питомі витрати, пов'язані зі зберіганням залишків і втрати від дефіциту j -ї продукції x_j. У зв'язку з випадковістю попиту в процесі прийняття рішень щодо обсягу виробництва продукції виникають такі ефекти розгалуження:

- якщо обсяг виробленої продукції x_j в потрібний момент перевищує обсяг попиту ω_j, тобто $x_j \geq \omega_j$, то фірма-виробник понесе витрати на зберігання надлишково виробленої продукції, що рівні $\alpha_j (x_j - \omega_j)$, j=1, 2, ..., n, та пов'язані з профіцитом у розмірі x_j - ω_j;
- якщо обсяг попиту ω_j в потрібний момент перевищує обсяг виробленої продукції x_j, тобто $\omega_j > x_j$, то фірма виробник терпить збитки, рівні $\beta_j (\omega_j - x_j)$, що пов'язані з дефіцитом в розмірі ω_j - x_j.

Функція витрат $f_j(x_j, \omega_j)$, пов'язана з обсягом виробництва продукції x_j, з урахуванням випадкового попиту ω_j, може бути представлена у вигляді наступної опуклої разгалуженої (кусочно-лінійної) функції витрат:

$$f_j(x_j, \omega_j) = \begin{cases} \alpha_j(x_j - \omega_j), якщо \quad x_j \geq \omega_j, \alpha_j > 0, \\ \beta_j(\omega_j - x_j), якщо \quad x_j < \omega_j, \beta_j > 0. \end{cases} \qquad (4)$$

На рис. 1 показано функцію $f_j(x_j, \omega_j)$ для випадку $c_j < \beta_j$.

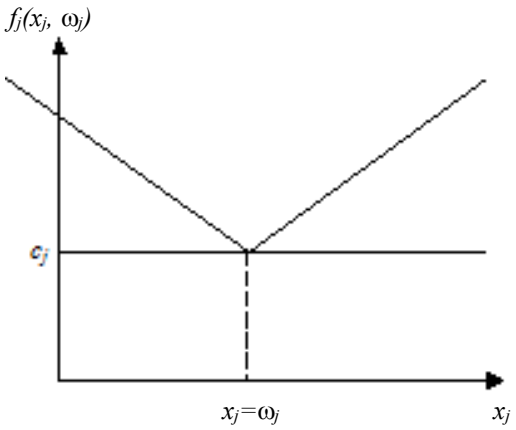


Рис. 1. Виникнення ефекту розгалуження при $x_j \neq \omega_j$.

Таким чином, стохастичний аналог ЗЛП (1) - (2) набуде вигляду

$$x_j^* = arg\min\{F(X)/X\}, \qquad (1a)$$

де

$$F(x) = M(\sum_{j=1}^n c_j x_j + \sum_{j=1}^n f_j(x_j, \omega_j)) = \sum_{j=1}^n c_j x_j +$$

$$+ M\sum_{j=1}^n max\{\alpha_j(x_j - \omega_j); \beta_j(\omega_j - x_j)\} =$$

$$= \sum_{j=1}^n c_j x_j + \sum_{j=1}^n (\int_0^{x_j} \alpha_j(x - \omega_j) P(d\omega_j) +$$

$$+ \int_{x_j}^{\infty} \beta_j(\omega_j - x_j) P(d\omega_j)).$$

Тут M - оператор математичного очікування; $f(x_j, \omega_j)$ - вимірна функція при кожному ω_j; ω_j - елементарна подія імовірнісного простору (Ω, Z, P), де Ω - множина елементарних подій; Z є σ-алгебра вимірних множин з Ω, P - певна в Ω імовірнісна міра, тобто $P(\Omega)$=1. У свою чергу, Z можна інтерпретувати як перелік всіх трьох груп конкретних результатів, які можна приписати для всіх груп з Z (наприклад, це може бути аналітична функція тощо).

Цільова функція $F(x)$ крім витрат на виробництво у виразі (1), також містить очікувані витрати, пов'язані з *дефіцитом* і *профіцитом* виробленої продукції. Завдання (1a) - (2) є окремим випадком загальної задачі стохастичного програмування [2]. Нижче розглянемо алгоритм вирішення цієї задачі.

Алгоритм 2. Нехай на *s*-му кроці (ітерації) отримано наближення x_i^s, $s = 0,1,...,$ (x^0 -задане довільне початкове наближення). Тоді:

1. У відповідності з апріорним розподілом $\varphi(\omega)$ отримуємо спостереження ω_j^s над реалізацією випадкової величини ω_j. Тут ω_j^s - незалежне спостереження над величиною ω_j может бути отримане в результаті компютерного експерименту (імітаційного моделювання) або ж в якості ω_j^s може бути взятий *s*-й елемент з набору статистичних спостережень (даних) величин в різні періоди (роки, місяці, декади тощо).

2. Обчислимо вектор-стохастичний квазіградієнт функції, компоненти якого визначаються наступним чином:

$$\xi_j^s = c_j + \begin{cases} \alpha_j, \text{якщо} \quad x_j^s \ge \omega_j^s, \quad s = 0,1,..., \\ -\beta_j, \text{якщо} \quad x_j^s < \omega_j^s, \quad j = 1,...,n. \end{cases} \qquad (5)$$

Нове наближення знаходимо згідно наступного варіанту стохастичного методу лінеаризації:

$$x_j^{s+1} = x_j^s + \rho_s(\bar{x}_j^s - x_j^s),\ 0 < \rho_s \le 1,$$

$$z_j^{s+1} = z_j^s + \gamma_s(\xi_j^s - z_j^s),\ \gamma_s > 0,\ j = \overline{1,n},\ s = 0,1,.... \qquad (6)$$

Тут $\bar{x}_j^s$, $j = \overline{1,n}$, є рішенням (згідно алгоритму 1) наступної ЗЛП:

$$\bar{x}_j^s = arg\,min\left\{\sum_{j=1}^{n} z_j^s x_j / X\right\},\ s = 0,1,.... \qquad (7)$$

де z_j^s, $j = \overline{1,n}$, визначається згідно (6), для якого ξ_j^s, $j = \overline{1,n}$, обчислюється за (5).

Важливою особливістю цього алгоритму є простота його реалізації на ЕОМ, оскільки напрям «спуску» в ньому побудовано на основі випадкового вектора-стохастичного квазіградієнта ξ_j^s, який є незміщеною оцінкою узагальненого градієнта $\hat{F}_x(x^s)$ цільової функції $F(x)$. Іншими словами, умовне математичне очікування визначається як:

$$M(\xi^s / \chi_s) = \hat{F}_x(x^s).$$

Тут вектор $\hat{F}_x(x^s)$ задовольняє нерівності

$$F(x) - \hat{F}_x(x^s) \ge (\hat{F}_x(x^s), x - x^s),$$

оскільки, за визначенням квазіградієнта

$$f(x,\omega) - f(x^s, \omega^s) \ge (\hat{f}_x(x^s, \omega^s), x - x^s).$$

Отже, отримавши математичне очікування з обох частин цієї нерівності, визначаємо квазіградієнт функції $F(x)$ для будь-яких x.

Теорема. *Нехай виконуються умови*

$$\left\|\xi_j^s\right\| \le d < \infty,\ \rho_s > 0,\ \gamma_s > 0,\ s = 0,1,...,$$

$$\sum_{s=0}^{\infty} \rho_s = \infty,\ \frac{\rho_s}{\gamma_s} \xrightarrow[s\to\infty]{} 0,$$

$$\sum_{s=0}^{\infty} \rho_s^2 < \infty,\ \sum_{s=0}^{\infty} \gamma_s^2 < \infty.$$

Тоді з ймовірністю 1 граничні точки послідовності $\{x_j^s\}_{s=0}^{\infty}$, *визначається* (5) - (7)*, що належать множині рішень задачі* (1а) - (2).

Доказ теореми і вибір параметрів методу (6) детально досліджені в [3].

У загальному випадку для вектора стохастичного квазіградієнта, маємо:

$$M(\xi^s / \chi_s) = \hat{F}_x(x^s) + b^s,$$

де b^s - зміщення, яке для збіжності викладених методів має в певному сенсі прямувати до 0 при $s \to \infty$.

IV. РЕЗУЛЬТАТ ЧИСЕЛЬНОГО ЕКСПЕРИМЕНТУ

Опишемо використання методу стохастичної лінеаризації (5) - (7) для вирішення окремого випадку завдання (1а) - (2) з додатковими обмеженнями типу $0 \le x_i \le d_i,\ i = \overline{1,5}$, коли витрати на виробництво не враховуються. Підпрограма, яка реалізує алгоритм 2, написана на мові VBA.

Нехай потрібно мінімізувати функцію

$$F(x) = M\sum_{i=1}^{n} max\left\{\begin{matrix} \alpha_i(x_i - \omega_i); \\ \beta_i(\omega_i - x_i) \end{matrix}\right\} \to min, \qquad (1в)$$

при обмеженнях

$$x_1 + x_2 + 2x_3 + 3x_4 + x_5 = 200,$$

$$0 \le x_1 \le 50;\ 0 \le x_2 \le 7;\ 0 \le x_3 \le 7;$$

$$0 \le x_4 \le 0;\ 0 \le x_5 \le 25. \qquad (2а)$$

Тут ω_j - випадкові величини, рівномірно розподілені на відрізках $[l_j, q_j]$, $i = 1, ..., 5$. Вектори $l = (l_1, ..., l_5)$; $q = (q_1, ..., q_5)$; $\alpha = (\alpha_1, ..., \alpha_5)$; $\beta = (\beta_1, ..., \beta_5)$; задані у вигляді

l = (0, 0, 0, 0, 0); q = (60, 15, 17, 90, 40);
α = (1, 0, 3, 1, 2); β = (3, 4, 1, 2, 3).

При вказаних значеннях аналітичний вираз функції мети (1в) має вигляд:

$$F(x) = \frac{1}{3}x_1^2 + \frac{2}{15}x_2^2 + \frac{2}{17}x_3^2 + \frac{1}{60}x_4^2 + \frac{1}{16}x_5^2 - \quad (1c)$$
$$- 3x_1 4x_2 - x_3 - 2x_4 - 3x_5 + 278{,}5$$

Стохастичний квазіградієнт $\xi^s = (\xi_1^s, ..., \xi_5^s)$ функції (1в) обчислюється аналітично згідно (5) за формулою

$$\xi_i^s = \begin{cases} \alpha_j, & \text{если} \quad x_j^s \geq \omega_j^s, \\ -\beta_j, & \text{если} \quad x_j^s < \omega_j^s, \quad i = 1, 2, ..., 5. \end{cases}$$

Для порівняння, отримано наступне точне вирішення задачі (1с) - (2а) одним з методів квадратичного програмування [3]:

x^* = (41.88057; 7.00000; 2.48092; 41.27456; 22.33456),
$F(x^*)$ = 98.10089.

Для вирішення стохастичної задачі (1в) - (2а) використовуємо алгоритм 2, для якого початкове наближення

x^0 = (10,59375; 0,59375; 5,18750; 58,78125; 2,09375),

крокові множники ρ_s і δ_s визначалися наступним чином:

$$\rho_s = 1/(s + 1), \delta = 1/(s+1)^{2/3}.$$

Наведемо результати усереднених значень шуканих змінних і функцій на ітераціях з 161-ї по 170-у:

$\bar{x}_1$ =41.24893;
$\bar{x}_2$=7.00000;
$\bar{x}_3$=2.22827;
$\bar{x}_4$= 42.2829;
$\bar{x}_5$=20.47934;

$$\bar{x}_i = \frac{1}{10}\sum_{s=164}^{170} x_i^s;$$

$$F(\bar{x}) = \frac{1}{10}\sum_{s=164}^{170} f(x^s, \omega^s) = 103.8945.$$

Слід зауважити, що значення $\bar{x} = (\bar{x}_1, ..., \bar{x}_5)$ близьке до оптимального рішення детермінованої задачі (1с) - (2а), а значення $F(\bar{x})$ мало відрізняється від значення $F(x^*)$.

Підвищення точності рішення вимагає збільшення кількості ітерацій. При неточності апріорної інформації про параметри завдання відпадає необхідність в отриманні високої точності рішення.

Чисельний експеримент виконаний за допомогою програми, розробленої автором на мові VBA.

Література References

[1] Ф. Мирзоахмедов, Противоречивые модели линейного программирования в производственно-экономических системах. - Душанбе: ТГНУ, 2002, 240с.

[2] Ю.М. Ермольев, Методы стохастического программирования. – М.: Наука,1976, 240с.

[3] Ф. Мирзоахмедов, Математические модели и методы управления производством с учетом случайных факторов. -Киев, Наукова думка, 1991, 220с.

СЕКЦІЯ 2.

Інформаційні технології в соціумі, освіті, медицині, економіці, управлінні, екології та юриспруденції

SECTION 2.

Information technologies in society, education, medicine, economics, management, ecology and jurisprudence

Пошук Облич з Допомогою Поділу Зображення та Параметричних Статистичних Ознак

Р. Мельник
кафедра програмного забезпечення
національний університет «Львівська політехніка»
Львів, Україна
ramelnyk@lp.edu.ua

Ю. Каличак, С. Дахній
кафедра програмного забезпечення
національний університет «Львівська політехніка»
Львів, Україна
yurii.i.kalychak@lpnu.ua[1], seryoga.dah@gmail.com[2]

Face Searching by Image Partitioning and Parametric Statistical Features

R. Melnyk
Software Department
Lviv Polytechnic National University
Lviv, Ukraine
ramelnyk@lp.edu.ua

Yu. Kalychak, S. Dachniy
Software Department
Lviv Polytechnic National University
Lviv, Ukraine
yurii.i.kalychak@lpnu.ua, seryoga.dah@gmail.com

***Анотація*—Представлено реалізацію алгоритмів поділу зображення на частини та обчислення статистичних ознак цих частин. Алгоритми застосовано до пошуку обличчя в базі даних. Наведено порівняння з різними способами формування ознак зображень.**

***Abstract*—The algorithms for image partitioning and parametric statistical features calculation is considered. The proposed features used for the face recognition approach. For the same data base the comparison results with different algorithms are presented.**

Ключові слова—параметричні статистичні ознаки; сітка; фрагемент; середні координати пікселів

Keywords—parametric features, intensity; grid; fragment; pixel coordinates mean

I. Introduction

Nowadays there is a great number of articles, books and surveys considering a face recognition problem. Also many software products realizing different algorithms are described in different publications and Internet resources. Here we can mark only a small part of the algorithms and works given by the Google search. Many surveys and classification of face recognition approaches are given in the works [1-4] paying attention on knowledge based methods, neural networks, template-based methods and many others. Among them we can select statistical approaches with their important representatives: 1) Principal Component Analysis (PCA) estimates 2-D facial image as 1-D vector by concatenating each row (or column) into a long thin vector [5]; 2) Linear Discriminant Analysis in which every component has a different discrimination for identifying a person [6]. Linear discriminant collects images of the same class and separates images of different classes; 3) Support Vector Machine is based on the theory of statistical learning [7].

The majority of the above-mentioned approaches are quite complicated and time-consuming. The software of algorithms is not available. The present work suggests a much simpler technique for the retrieval of face image by the feature vectors. The technique is based on the fragmentation of the three-dimensional image intensity space forming fragments and segments.

II. Parametric statistical features for one level partitioning

A. Algorithm "cell - number of means" (A1)

In order to obtain the three-dimensional surface of the image, first the color image is converted into the shades of grey. Every elementary cell, a pixel, assumes a value ranging from black to white, which we label as b – intensity (brightness). The range of all possible intensity values is within 0÷255.

Let us divide the image intensity space into n fragments by the horizontal planes XOY with an interval (intensity fragmentation step) d =255/n

Fragments or segments contain a large number of pixels which could be characterized by different distributed features. For example, the mean of the fragment or segment pixel coordinates:

$$\bar{x}(s) = 1/k_s \sum X(s) \quad \bar{y}(s) = 1/k_s \sum Y(s),$$

$$x_i \in X(s), y_i \in Y(s) \qquad (1,2)$$

where s – fragment or segment number, X, Y – sets of pixel coordinates.

To get a parametric feature connected with pattern coordinates, we divide a pattern surface by a grid with some steps by OX and OY coordinates (fig.1). The grid cells are numbered as 1, 2, …, Lx*Ly from left to right and down from a grid top. In every pixel fragment or segment the mean coordinates $\bar{x}(s)$, $\bar{y}(s)$ are associated with the cell number and its coordinates by corresponding formula:

$$C_n = (i-1)*L_x + j,$$

$$i = (\bar{x}(s)/L_x)_{ceil} + 1, \; j = (\bar{y}(s)/L_y)_{ceil} + 1 \qquad (3,4)$$

. Some mean coordinates of pixels in different segments or fragments have close values. By this reason they belong to the same cell. As a result the cell feature C_f has value equal to the number of segment (fragment) having the same cell coordinates. In fig. 1 we can observe the face covered by a grid with some numbers in cells. Numbers and cells create parametric features of a dace covered by the grid 10x10.

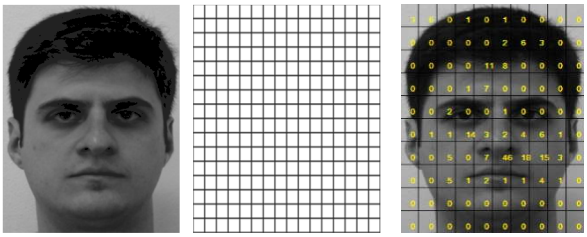
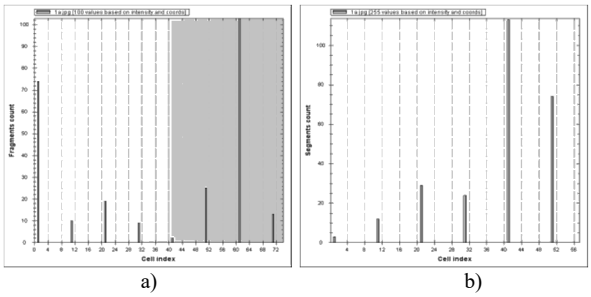


Fig. 1. Face, grid and mean numbers

Another representation of parametric features is realized with a chart by two coordinate axis: 0-C_f and 0-C_n. For one face decomposition by fragments and segments two examples of charts are given in fig. 2.



a) b)

Fig. 2. Cell value by cell numbers in fragments (a) and segments (b)

In both representations features and their numbers are the same cell position and cell value. A grid dimension defines a number of features.

B. Algorithm "cell—mean intensity" (A2)

In order to improve previous statistical features we correlate every coordinate mean with its intensity. So we get so called "cell – mean intensity feature. We can call them also "parametric" as depending from a step of the covering grid. For four faces in Fig. 3 we have fragment and segment parametric features for grid 10x10 presented in fig. 4.

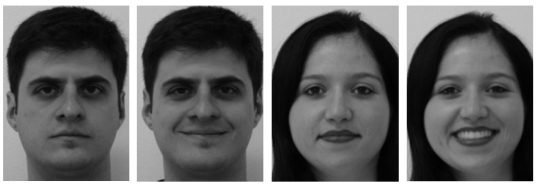


Fig. 3. Fragment and segment parametric features (5x5)

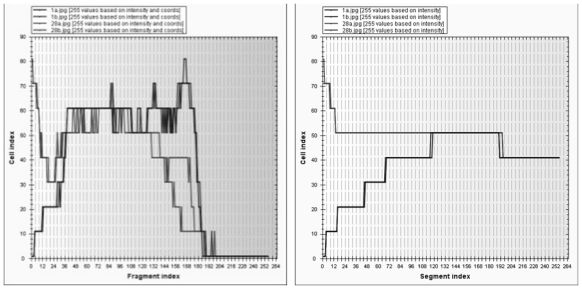


Fig. 4. Fragment and segment parametric features (10x10)

The less step of the grid the more precise parametric features are. Also from presented charts in fig. 4 we can conclude that the fragment features are more accurate and could be used for face and emotion recognition.

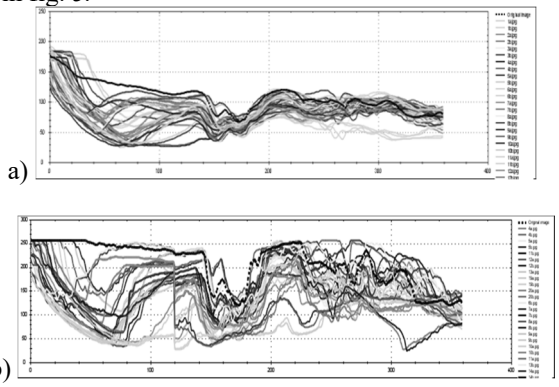
C. Algorithm "mean intensity - row'" (A3)

To compare the methods effectiveness one more statistical feature was developed. We consider an image as of pixel intensity. So for every row or column we calculate the mean intensity by following formula:

$$\bar{B}_j = 1/H \sum_{i=1}^{H} b_{i,j}, 1 \le j \le W, \; \bar{B}_i = 1/W \sum_{j=1}^{W} b_{i,j} \quad 1 \le i \le H, \qquad (5)$$

where $b_{i,j}$ – pixel intensity in the i-th row and j-th column.

The charts (5) of 50 faces of column and row features are shown in fig. 5.



a)

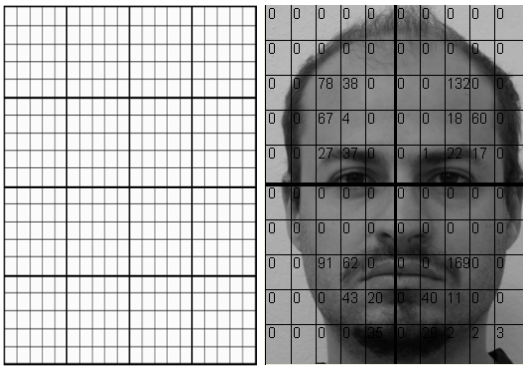
b)

Fig. 5. Distributed mean intensity values by columns (a) and rows (b)

III. TWO LEVELS OF FACE PARTITIONING

A. Algorithm "two levels - mean" (A4)

As we can see from Fig.1,2 the cell value feature has many null elements. So, to get more information we cover a face image by a grid with two levels of partitioning. For example, grid on Fig.6 has basic division with 20x20 cells and additional one with 4x4 (or 2x2) cells. So, we consider every bigger part covered by its own grid separately. The resulting grid has much less null elements doing the majority of cells informative (fig. 6).



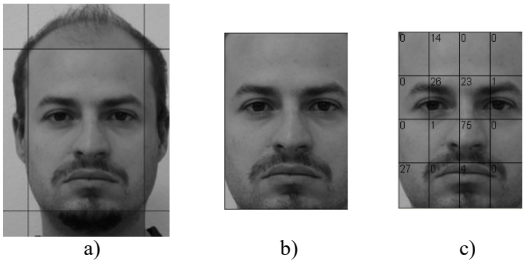
a) b)

Fig. 6. Grid with two level (a) and features in cells (b)

B. Preprocessing of frame extraction

One more approach considers the grid covering not a full face but its central part – so called frame. It contains the most essential and essential information: eyes, mouth and nose. External zones including background are cut.

One example of face cutting to get frame including covered by grid and features are shown in fig. 7.



a) b) c)

Fig. 7. Face with shadows (a), frame (b) and features on cells (c)

IV. COMPARISON OF METHODS

For testing, the database of 50 images of face pairs with different expressions from [8] has been chosen. 14 searching experiments were held for the 6a face: by differents features of one level partitioning and two levels partitioning, parametrica and mean features of full faces and frames. The experiments were held for fragments and segments of intesity. To compare the proposed method and having no software of world known algorithms we have chosen as refference points the searching results of our "mean intensity – row" (A3) algorithm. To precise these points we use three algorithms: full image and full intensity - mean intensity in rows, frame image and full intensity - mean intesity in rows, frame image and image segment (t=80) – mean intesity in rows. An example of a spread of the features given by the last algorithm is shown in fig. 8. In this figure we can see how charts (280 points) differ by intensity and shapes.

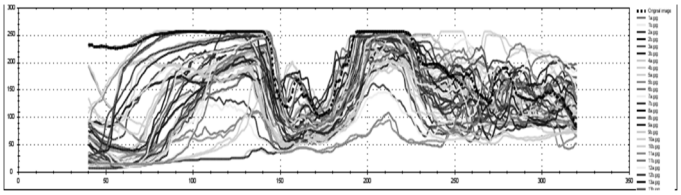
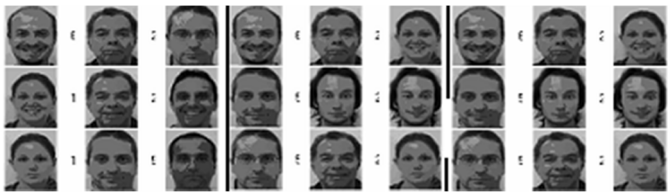


Fig. 8. Mean intensity in rows for segmented frames of 50 faces

Nine first representatives (a lack of article space is the reason) from 50 resulting faces given by three algorithms are shown in fig. 9.



a b) c)

Fig. 9. Resulting faces (9) for full images (a), for frames (b) and segmented frames (c) by mean intensity in row algorithm

Every face in the list has its own position, i.e. place number. As resulting refference points of these faces (algorithms) we use mean place number:

$$f_o(x) = 1/3 \sum_{x=1}^{3} |f_i(x)| \qquad (6)$$

where $f_0(x)$ – a place number of the face (x), $f_i(x)$ – a place number of the same face in reference algorithms.

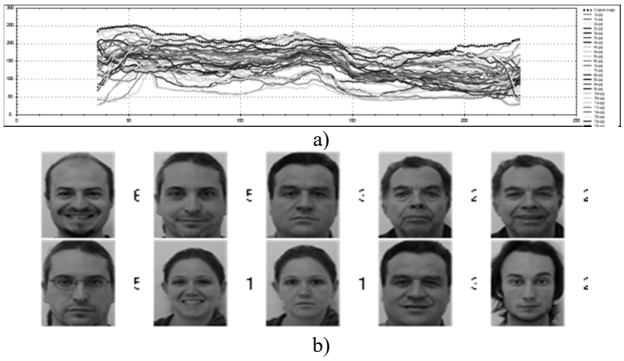
All faces taking part in experiment are being signed by numbers got by searching and sorting. So, the evaluation function between etalon and comparative methods we calculate as follows

$$E_m = \sum_{x=1}^{W} |f_0(x) - f_m(x)| \ , \qquad (7)$$

where $f_m(x)$ – a place number of the face (x) in comparative algorithm.

Having evaluation function we can compare results of all developed algorithms with the etalon $f_0(x)$ features. For similarity we use the formula (7) only for 9 first faces.

There are three groups of experiments based on algorithms A3, A1, A2. In the first group we do six experiments: by columns and rows and for full images, frames and segmented frames. Evaluation function values are being inserted into the table 1. One experiment – the mean intensity in columns of segmented frames and first 10 faces – is illustrated in fig. 10.



a)

b)

Fig. 10. Mean intensity in columns for segmented frames of 50 image (a) and closest faces to the image 6a (b)

From the second group of experiments we illustrate by two groups of resulting faces got by the algorithm A1 for the segment and fragment intensity features. In fig. 11 we can see ten faces found for one level grid (12x15).

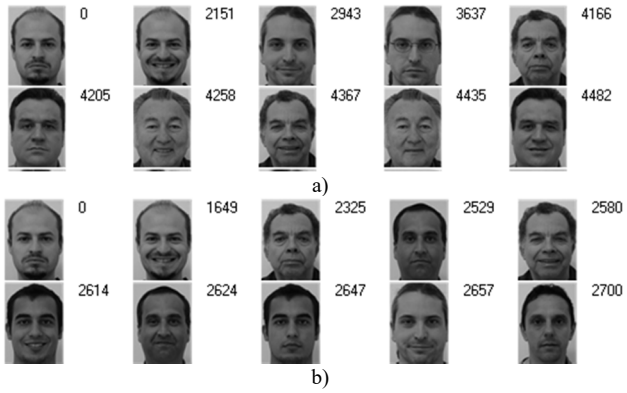
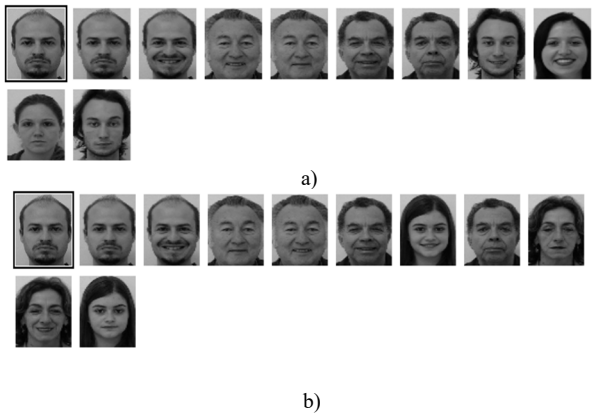


Fig. 11. Searching results segment features (a) and fragment features (b) by grids of 12x15

All evaluation values of four experiments fith A1 are inserted into the table 1. Also four experiments were held with algorithm A2. On fig. 12 we can see ten first faces found for the grid 12x15 by intensity segments and fragments.



a)

b)

Fig. 12. Results for segments (a) and fragments (b) with grid 12x15

TABLE I. COMPARATIVE RESULTS OF FACE SEARCHING

Methods		**Evaluation function**		
		Full image	***Frame***	***Frame segm.***
A3 -rows		22	10	10
A3 -columns		19	15	18
A1	frag.	37	34	-
	segm.	33	31	-
A2	frag.	41	36	-
	segm	38	34	-

The experiments have confirmed that: 1) first results from different features are not the same, 2) some pictures that are close to the request are being repeated in different experiments, 3) results are sensitive to parameters i.e. grid dimension, mode of intensity decomposition etc. Maybe partly responsible for such results are cumulative histograms of images which differ between themselves essentially. So, preliminary preprocessing is needed particularly for grid methods.

V. CONCLUSION

The statistical parametric features of image and its intensity fragments and segments have been proposed. Some results of face image searching and comparison of approaches are presented.

REFERENCES

[1] Stan Z. Li, Anil K. Jain, "Handbook of Face Recognition", *Springer,* p.395, 2005.

[2] W. Zhao, R. Chellappa, P. J. Phillips, A. Rosenfeld "Face Recognition: A Literature Survey", *ACM Computing Surveys*, vol.35, No.4, December 2003 pp.399-460.

[3] Rabia Jafri, Hamid R. Arabnia, "A Survey of Face Recognition Techniques", *Journal of Information Processing Systems*, vol.5, No.2, pp.41-68, June 2009.

[4] Kavita, Ms. Manjeet Kaur. "A Survey paper for Face Recognition Technologies", *International Journal of Scientific and Research Publications*, vol.6, Issue 7, pp.441-445, July 2016.

[5] Martinez A.M. and Kak A.C., "PCA versus LDA", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.23, No.2, pp.228-233, 2001.

[6] G. Guodong, S. Li, and C. Kapluk, "Face recognition by support vector machines", *In Proc. IEEE International Conference on Automatic Face and Gesture Recognition*, pp.196–201, 2000.

[7] Toygar Onsen and Acan Adnan, "Face recognition using PCA, LDA and ICA approaches on colored images", *Journal of electrical and Electronics Engineering*, vol.3, No.1, pp.735 – 743, 2003.

[8] FEI data base, http://fei.edu.br

Моделювання Плану Ремонту Доріг в Межах Територіальної Громади на Основі Модифікації Алгоритму Пріма

В. Литвин
кафедра інформаційних систем та мереж
Національний університет «Львівська політехніка»
Львів,Україна
Vasyl.V.Lytvyn@lpnu.ua

Д. Угрин
кафедра інформаційних систем
Чернівецький факультет Національного технічного університету «Харківський політехнічний інститут»
Чернівці,Україна
ugrind@mail.ru

З. Рибчак
кафедра інформаційних систем та мереж
Національний університет «Львівська політехніка»
Львів,Україна
zoriana.l.rybchak@lpnu.ua

Modelling of a Road Repair Plan Within the Territorial Community Based on Prima's Modifications Algorithm

V. Lytvyn
Department of Information Systems and Networks
National Lviv Polytechnic University
Lviv, Ukraine
Vasyl.V.Lytvyn@lpnu.ua

D. Ugryn
Department of Information Systems
National Technical University “Kharkiv Polytechnic Institute”
Chernivtsi, Ukraine
ugrind@mail.ru

Z. Rybchak
Department of Information Systems and Networks
National Lviv Polytechnic University
Lviv, Ukraine
zoriana.l.rybchak@lpnu.ua

***Анотація*—У роботі розглянуто моделювання планування заходів щодо ремонту доріг в межах територіальної громади в залежності від наявних коштів та стану доріг. Для моделювання запропоновано використати метод пошуку мінімального остовного дерева на основі модифікації алгоритму Пріма.**

***Abstract*—The paper describes modeling of planning measures to repair roads within a territorial community depending on available funds and roads condition. The minimum spanning tree search method based on modified Prima algorithm is proposed.**

Ключові слова—граф; остовне дерево; алгоритм Пріма; територіальна громада; населений пункт

Keywords— graph; spanning tree; Prima algorithm; local community; town

I. Вступ

Метою законопроекту «Про добровільне об’єднання територіальних громад» є створення потужних територіальних громад. Таке об’єднання дає змогу заощадити бюджетні кошти шляхом скорочення держапарату на місцях [1]. В Україні створено вже ряд територіальних громад. В межах територіальної громади є адміністративні будівлі для розміщення органів управління місцевого самоврядування, органу правопорядку, пожежної частини, пункту швидкої допомоги, центру надання адміністративних послуг, державного

казначейства тощо. Такі будівлі як правило розміщуються в центрі територіальної громади. Складовою територіальних громад (*ТГ*) є населений пункт (*НП*), тобто $ТГ = \{НП_1, НП_2, ..., НП_N\}$. Виникає проблема дорожнього зв'язку між центром *ТГ* та її *НП*. Очевидно, що окремі *НП* зв'язані між собою дорогами (у цій роботі будемо лише говорити про автомобільні дороги), однак на даний час стан 95% доріг України є незадовільним [2]. У 2016 році було відремонтовано лише 1% автомобільних доріг, у 2017 році планується аналогічний відсоток [3]. Тому виникає задача ефективного розподілу коштів для ремонту доріг на кількарічний період.

II. Постановка задачі

Дорогу між двома різними *НП* будемо розглядати як окрему дорогу. Якщо є розвилки доріг поза межами *НП*, то в місці такої розвилки вводитимемо фіктивні *НП* (*ФНП*). Дороги мають різне підпорядкування, та стан. Підпорядкованість визначає за який бюджет ця дорога буде ремонтуватись. Як правило, існує 3 бюджети: B_1 – загальнодержавний бюджет, B_2 – обласний бюджет, B_3 – бюджет *ТГ*. Нехай за бюджетом B_j було виділено W_j коштів. Стан дороги визначає необхідний кошторис на ремонт відповідної дороги. Як правило, є таких 5 станів: 1) видимі незначні дефекти і, необхідність ремонту до 5 %; 2) шелушіння, окремі нерівності покриття, частково присутні тріщини та невеликі вибоїни, необхідність ремонту до 25 %; 3) викрішування, раковини, зсуви, просідання, незначно виражена колійність, руйнування кромок дорожнього покриття, граней бетонного покриття, бордюрів, необхідність ремонту до 50 %; 4) вибоїни, проломи, великі ями, значна колійність, місцями пересування значно ускладнено, необхідність ремонту до 75 %; 5) базовий тип покриття практично відсутній, явно виражена колійність, пересування значно ускладнено, необхідність ремонту до 100 %. Вартість ремонту доріг також залежить від типу покриття. Розрізняють 6 різних типів покриття: цементно-бетонне, залізобетонне, бруківка, асфальтобетонне, гравійне, ґрунтове.

При розрахунку вартості ремонту доріг визначалась ціна ремонту 100 метрів дороги. Врахування стану дороги забезпечується множенням на певний коефіцієнт. Розроблений нами калькулятор знаходиться за адресою: http://www.roadcost.96.lt/. Приклад обчислення вартості доріг для *Ходорівської ТГ* наведено на рис. 1. Розроблений модуль дає змогу фільтрувати вартість ремонту доріг в межах *ТГ* за такими параметрами: тип дороги, тип покриття, стан.

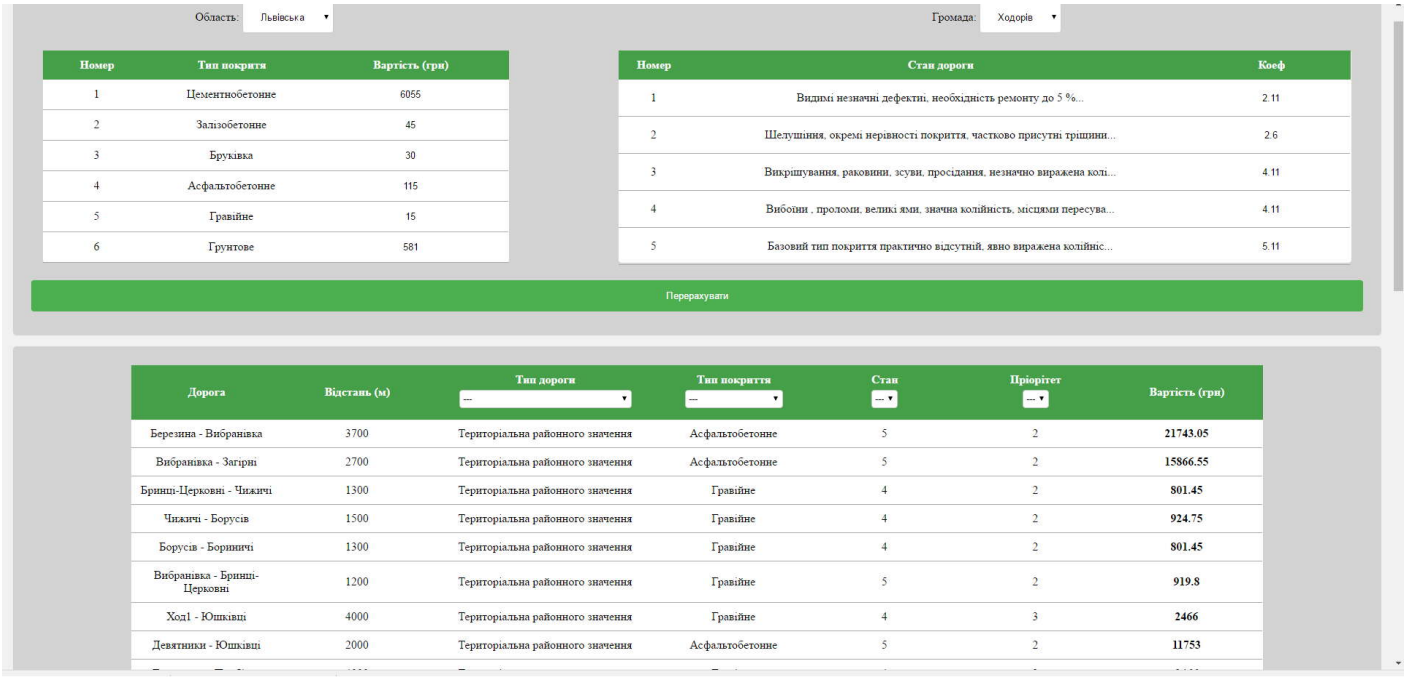


Рис. 1. Модуль обчислення вартості ремонту доріг в межах територіальної громади

Отримуємо зважений граф $G = (НП, E)$, вершинами якого є *НП* та *ФНП* (надалі всі позначатимемо *НП*), а ребра $E = \{e_1, e_2, ..., e_n\}$ задають кошторис $w(e_i)$, необхідний на ремонт дороги між відповідними *НП* [4-5]. Тобто, маємо N *НП*, які необхідно об'єднати дорогами. Очевидно, що для цього достатньо відремонтувати $N-1$ доріг між *НП*. Виникає задача: як об'єднати *НП* між собою в межах *ТГ*, щоб сумарна вартість ремонту доріг була мінімальна й вартість ремонту доріг не перевищувала відповідний бюджет. Тобто всі ребра поділимо на 3 підмножини в залежності від виду бюджету $E = E_1 \cup E_2 \cup E_3$, $E_j = \{e_1^j, e_2^j, ..., e_{n_j}^j\}$, $j = 1,2,3$.

Тобто задача полягає в знаходженні такого зв'язаного ациклічного підграфу $T \subset G$, який містить всі вершини, щоб сумарна вага всіх його ребер була мінімальною при умові, що сумарна вага ребер, які належать до одного бюджету не буде перевищувати розмір цього бюджету.

Оскільки *T* зв'язаний й не містить циклів, він є деревом й називається остовним деревом (spanning tree) [6-8]. Остовне дерево *T*, в якого сумарна вага його ребер $w(T) = \sum_{e_i \in T} w(e_i)$ мінімальна, називається мінімальним остовним деревом (minimum spanning tree). Частину дерева *T*, що містить ребра з підмножини E_j позначатимемо T_j, $T = T_1 \cup T_2 \cup T_3$. Тим самим отримаємо таку математичну модель задачі: знайти $T \subset G$, щоб

$$w(T) = \sum_{e_i \in T} w(e_i) \to \min , \quad (1)$$

$$w(T_j) = \sum_{e_s^j \in T_j} w(e_s^j) \le W_j , \; j = 1,2,3 . \quad (2)$$

III. Модифікація алгоритму Пріма

Як відомо, для знаходження мінімального остовного дерева використовується алгоритми Пріма або Крускала [5-6]. Для розв'язування задачі (1)-(2) нами модифіковано алгоритм Пріма. Модифікація полягає в ітераційному використанні алгоритму Пріма, поки не виконається умова (2). Якщо умова (2) не виконується, то із початкового графа *G* вилучаємо ребро із максимальною вагою серед підмножини ребер для яких не виконується (2) із максимальною різницею між необхідним коштами на ремонт доріг й наявним бюджетом. Ребро вилучаємо із умовою, що граф *G* залишається зв'язним, тобто не має ізольованих вершин. Якщо такого ребра немає, то задача (1)-(2) немає розв'язку.

Отримаємо такий алгоритм визначення доріг, які необхідно відремонтувати в межах *ТГ*:

1) Утворити граф *G*, вершинами якого є *НП ТГ*, а ребра задають вартість ремонту доріг між *НП ТГ*. Відомі кошти, які закладені у бюджетах на ремонт доріг W_1 , W_2 , W_3 .

2) Запустити алгоритм Пріма для графу *G*:

2.1. Утворимо дерево T_1 з одним ребром:

– виберемо його вершиною $НП_0$ центр ТГ;

– виберемо ребро e_1 з найменшою вагою серед тих, що мають вершину $НП_0$;

– покладемо $k = 1$.

2.2. Якщо існують вершини початкового графа *G* зовні останнього побудованого дерева T_k з ребрами $e_1, e_2, ..., e_k$, то робимо таке:

– вибираємо ребро e_{k+1} з найменшою вагою серед тих, у яких одна вершина належить до T_k , а інша вершина не належить;

– утворюємо дерево T_{k+1} долученням до T_k вибраного ребра e_{k+1} і його вершин;

– збільшуємо величину k на 1;

2.3. Якщо всі вершини початкового графа *G* належать до дерева T_k , то припиняємо побудову мінімального остовного дерева, інакше переходимо на пункт 2.2

3) Для дерева T_k знаходимо вартості ремонту доріг за 3-ма підмножинами: $w(T_j) = \sum_{e_s^j \in T_j} w(e_s^j)$, $j = 1,2,3$.

4) Обчислюємо значення $\Delta_j = W_j - w(T_j)$, $j = 1,2,3$.

5. Якщо всі $\Delta_j \ge 0$, то дерево (план ремонту доріг) знайдено, інакше серед множини E_l, де $l = \arg\max_{\Delta_j < 0} |\Delta_j|$, вилучаємо ребро з максимальною вагою серед підмножини ребер E_l, після вилучення яких граф *G* залишається зв'язним й переходимо до п.2. Якщо ребер, після вилучення яких граф *G* залишається зв'язним, немає, то плану ремонту доріг для таких початкових значень не існує. Необхідно збільшувати бюджети для яких $\Delta_j < 0$.

Тобто в алгоритм Пріма додано ще один ітераційний процес, мета якого зменшувати розмірність початкового графу за рахунок відкидання ребер про які можна припустити, що вони не будуть відображені у кінцевому дереві. Суттєво перевіряти, щоб у графа при такому відкиданні нез'являлись ізольовані вершини, бо інакше в *НП*, який задає така вершина, добратись буде неможливо.

Складність запропонованого алгоритму в n разів більша за алгоритм Пріма, оскільки додається ще один цикл, тобто становитиме $O(n^2 \cdot \ln n)$.

IV. Приклад моделювання плану ремонту доріг в межах Давидівської територіальної громади

Розглянемо приклад ТГ й роботу запропонованого алгоритму. 17 серпня 2016 року у Львівській обласній державній адміністрації підписали розпорядження про створення *Давидівської територіальної громади.* Добровільно у цю громаду увішли такі НП: Пасіки-Зубрицькі, Горішній, Кротошин, Чишки, Бережани, Соснівка, Волиця, Виннички, Гончари, Дмитровичі, Давидів, Черепин. Граф доріг цієї ТГ наведено на рис. 2. Чорними точками позначено НП, білими – ФНП (перехрестя доріг поза НП).

Дороги належать до 3 підмножин: E_1 – загальнодержавного підпорядкування (червоний колір, ремонт здійснено ще у 2011 році перед ЧЄ з футболу), E_2 – обласного підпорядкування (синій колір) E_3 – районного підпорядкування (чорний колір).

На графі задана вартість доріг (дані є експериментальними). Взявши, що $W_1 = 30$, $W_2 = 30$,

$W_3 = 22$, отримаємо пропонований план ремонту доріг, який наведений на рис. 3. Зауважимо, що якщо б не враховувалась те, що підмножини ребер різні, то за алгоритмом Пріма у Виннички пропонувалось б доїзжати із Гончари. Однак через обмеження на бюджет, рекомендується ремонтувати дорогу із множини E_2 і у Виннички їздити із перехрестя дороги, що йде на Дмитровичі. Загалом за множинами виходить така вартість ремонту доріг: $w(T_1) = 0$, $w(T_2) = 28 < 30$, $w(T_3) = 21 < 22$.

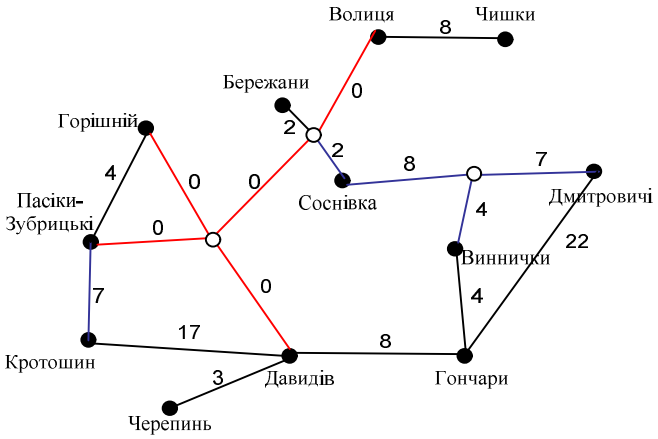


Рис. 2. Граф доріг Давидівської територіальної громади

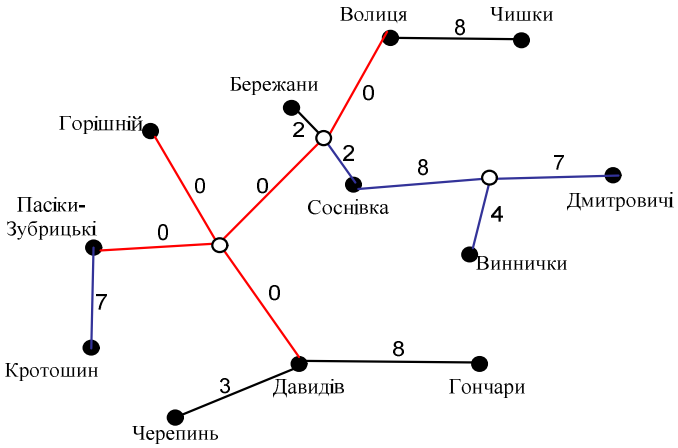


Рис. 3. Пропонований план ремонту доріг Давидівської територіальної громади.

V. Висновки

Запропоновано використати модифікацію алгоритму Пріма пошук мінімального остовного дерева для задачі моделювання планування ремонту доріг в межах територіальної громади. Задача стає актуальною на даний час, оскільки процес формування таких громад є важливим завданням сьогодення. В межах територіальної громади знаходяться важливі адміністративні будівлі, лікарня тощо. Тому доїзд до цих закладів в межах громади є суттєвою проблемою, зважаючи на стан доріг. Щоб ефективно використати наявні кошти на ремонт доріг з врахуванням їх стану та важливості, запропоновано розв'язувати класичну задачу пошуку мінімального остовного дерева з врахуванням специфіки підпорядкування доріг.

Наведено приклад використання запропонованого підходу в межах Давидівської територіальної громади.

У подальших дослідженням планується враховувати ще один показник, який суттєво впливає на визначення доріг, які підлягають якнайшвидшому ремонту, а саме – пріоритет дороги (1 – дорога має стратегічно важливе значення, з трафіком понад 500 автотранспортних засобів за добу; 2 – дорога має важливе значення, з трафіком руху від 100 до 500 автотранспортних засобів за добу; 3 – дорога має другорядне значення, з трафіком руху менше 100 автотранспортних засобів за добу). Тим самим дослідження полягатимуть в побудові інтегральної функції оцінки ребра початкового зваженого графу G, який зв'язує певні два населені пункти, де буде враховуватися не лише вартість ремонту дороги, але й її пріоритет.

Література References

[1] Закон України Про добровільне об'єднання територіальних громад [Електронний ресурс]. – Режим доступу: http://zakon5.rada.gov.ua/laws/show/157-19

[2] Стан українських доріг [Електронний ресурс]. – Режим доступу: http://censor.net.ua/news/430063/95_ukrainskih_dorog_ostayutsya_v_neprigodnom_sostoyanii_omelyan

[3] Ремонти доріг [Електронний ресурс]. – Режим доступу: http://censor.net.ua/news/430172/ukravtodoru_nado_organizatsionno_pravilno_rabotat_95_dorog_nahodyatsya_v_ujasnom_sostoyanii_groyisman

[4] В.В. Литвин, Д.І. Угрин, А.М. Фітьо, «Моделювання процесу формування територіальних громад як задачі розбиття графу», Східно-Європейський журнал передових технологій, №1/4(79), С.47-52, 2016

[5] В.В.Литвин, Д.І.Угрин, А.М.Фітьо «Формалізація задачі формування територіальних громад», 11 Міжнародна науково-практична конференція „Математичне та імітаційне моделювання систем МОДС" (27 червня – 1 липня 2016). – Жукин. - С. 290-292.

[6] Макконелл Дж., «Основи сучасних алгоритмів: 2-е доповнене видання», М.: Техносфера, 2004, 368с.

[7] В. А. Евстигнеев, «Применение теории графов в программировании», М.: Наука, 1985, 352 с.

[8] М. Свами, К. Тхуласираман «Графы, сети и алгоритмы», М.: Наука, 1984, 256с.

Оптична Система для Фототерапії з Контролем Поглинутої Енергії

Олексій Яненко
кафедра радіо конструювання
НТУУ «КПІ»
м. Київ, Україна
op291@meta.ua

Роман Ткачук
кафедра біотехнічних систем
ТНТУ ім. І. Пулюя
м. Тернопіль, Україна
vmp@tu.edu.te.ua

Костянтин Шевченко
кафедра радіо конструювання
НТУУ «КПІ»
м. Київ, Україна
autom1@meta.ua

Василь Кузь
кафедра світлотехніки та електротехніки
ТНТУ ім. І. Пулюя
м. Тернопіль, Україна
vasylkuz1992@gmail.com

Optical System of Phototherapy to Control the Flow of Absorbed Energy

Oleksiy Yanenko
dept. of Radio construction
NTUU "KPI"
Kyiv,Ukraine
op291@meta.ua

Roman Tkachuk
dept. of biotechnical Systems
Ivan Puluj TNTU
Ternopil,Ukraine
vmp@tu.edu.te.ua

Kostyantyn Shevchenko
dept. of Radio construction
NTUU "KPI"
Kyiv, Ukraine
autom1@meta.ua

Vasyl Kuz
dept. of electrical and lighting
Ivan Puluj TNTU
Ternopil,Ukraine
vasylkuz1992@gmail.com

***Анотація*—Проаналізовані поширені конструкції опромінювання шкіри людини різної складності, які використовуються для світлотерапії. Визначено, що їх основним недоліком є відсутність можливості вимірювання дози поглинутої енергії біооб'єктом і відповідно, неможливість оперативного прогнозування результату лікування. Запропоновано пристрій для світлотерапії та досліджено алгоритм перетворення параметрів світлового потоку, які усувають виявлений недолік. Це дозволяє оперативно визначати дозу поглинутої енергії опромінюваної ділянки тіла пацієнта при змінах вибраних режимів калібрування та вимірювання параметрів опромінення.**

***Abstract*—Analyzed the most common design of varying difficulty devices, which are used for light therapy. Determined that their main drawback is the lack of the possibility of measuring the absorbed light energy dose, and therefore, the inability to operational forecasting outcomes. A device and investigated transformation algorithm parameters flux, eliminating deficiencies. This allows you to quickly determine the absorbed dose of energy to the irradiated areas of the patient's body in the modes of calibration and measurement tanning signal.**

Ключові слова—пристрій для світлотерапії, світловий потік, потужність, опромінюючий сигнал, доза поглинутої енергії, комутаційно-модуляційне перетворення.

Keywords—light therapy device, light output, power, tanning signal dose absorbed energy switching modulation-conversion.

I. Вступ

Відомо, що роль світла в процесах метаболізму живих організмів є дуже важливою. При порушенні цих механізмів перспективним для корекції і відновлення функцій організму, з метою покращення самопочуття пацієнтів, є застосування регульованих джерел світла [1]. Серед відомих напрямків такого використання є світлотерапія, яка знаходить все більше застосування в практичній медицині для лікування, профілактики та реабілітації з використанням потоків енергії різної довжини хвиль, шляхом впливу на біологічно активні точки чи зони шкіри людини. Лікування різноманітних уражень шкіри, бронхолегеневих та ортопедичних захворювань, деяких новоутворень, опромінення крові в процедурі фотофереза - це далеко не повний перелік ефективного застосування світлотерапії [2-3].

Важливим моментом при проведенні терапевтичних процедур є встановлення оптимальних параметрів опромінення – потужності джерела світла, експозиції, довжини хвиль опромінення. Зазвичай рівні вихідної потужності встановлюється в межах одиниць мВт, а ефективність регулюють часовою експозицією [2]. В той же час відомо, що біологічний ефект пропорційний поглинутій дозі електромагнітного опромінення, яку бажано було б вимірювати протягом процесу проведення процедури.

II. Постановка задачі дослідження

Відсутність зворотнього зв’язку є причиною недостатнього застосування фототерапевтичних технологій, тому виникає завдання удосконалення їх шляхом пошуку нових підходів у виборі джерела випромінювання та реалізації на їх основі нових оптико-електронних пристроїв із неперервним контролем параметрів процесу.

Розробка та дослідження пристрою для світлотерапії і алгоритму перетворення сигналів, які б забезпечували можливість вимірювання інтенсивності та контролю поглинутої енергії в процесі проведення процедури і, тим самим, прогнозувати часові параметри досягнення лікувального ефекту.

III. Матеріали та методи досліджень

На сьогоднішній день широкого застосування набули напівпровідникові джерела світла, які відомі своїми якісними характеристиками, а саме високою світловіддачею, механічною стійкістю, безпечністю, ціною, тривалістю служби та спектром випромінювання.

Використання напівпровідникових елементів для оптико-електронних пристроїв, що можуть застосовуватися у фототерапевтичних технологіях є високо перспективним, оскільки вони володіють покращеним коефіцієнтом корисної дії та мають можливість регулювання просторового розподілу енергії в малих площинах для різного спектру випромінювання. Також вони характеризуються незначною величиною робочої напруги та струму, малою інерційністю при роботі в імпульсному режимі, високою надійністю, низькою вартістю тощо. На відміну від традиційних розрядних джерел випромінювання, такі пристрої забезпечують надійну роботу в імпульсному режимі, допускають керування випромінюванням в заданому діапазоні енергетичних та часових параметрів.

Світлодіоди завдяки відповідністю медичним вимогам [2] дозволяють створення на їх основі світлодіодних матриць, які забезпечують можливість реалізувати стаціонарні і стимуляційні режими опромінення по всій його площі за допомогою програмно керованих комутацій світлодіодів різних характеристик. Підключення світлодіодних матриць до мікроконтролера дасть змогу включати тільки необхідні частини опромінювальної установки і зменшить негативний ефект впливу опромінювання на здорову поверхню шкіри.

Таким чином, за допомогою фототерапії, тканини, опромінювані світлом певної довжини хвилі, благотворно використовують світлову енергію для більш ефективного функціонування. У тканинах підвищується синтез колагенових і еластинових волокон, що поліпшує її якість і зовнішній вигляд.

На основі проведених літературних досліджень визначено основні критерії для виявлення вимог, щодо характеристик елементів для структури пристроїв контролю динамічного опромінення, аналізу їх можливостей і розробки рекомендацій по оптимальному вибору елементів пристроїв сумісних з пристроями фототерапії. При вирішенні даного завдання широка номенклатура промислових типів електронних компонентів, зокрема СВД і фотоприймачів, дозволяє вибрати потрібні компоненти. Перш за все, це стосується відповідності їх спектральних характеристик певним спектральним смугам поглинання. Беручи до уваги недостатність, а в окремих випадках - відсутність інформації, щодо формулювання зазначених вище вимог, порівняльного аналізу відповідності їм існуючих складових, по рекомендацій їх вибору і експериментальної перевірки відповідності, постановка і рішення даного завдання важливі і актуальні.

Завдяки плідній співпраці вчених двох університетів КПІ та ТНТУ створено корисну модель під назвою «Пристрій для світлотерапії» під номером заявки u201608572, яка увібрала в себе найнеобхідніші елементи для досягнення якісного світлолікування. Було запропоновано нову структуру оптико-електронної системи безперервного контролю при фототерапії, яка повинна складатися, як мінімум, з трьох частин: випромінюючої, приймальної, передавальної.

В основу цієї корисної моделі була поставлена задача створити такий пристрій для світлотерапії, в якому б шляхом введення нових елементів та зв’язків забезпечувалась би можливість вимірювання інтенсивності світлового потоку та контроль поглинання

енергії світлового потоку в процесі проведення процедури.

Поставлена задача вирішується тим, що в пристрій для світлотерапії, що включає в себе матрицю світлодіодів, з'єднану з блоком контролю та керування, датчики фізіологічного стану людини, підключені через модуль зв'язку до блоку контролю та керування, вхід-вихід якого підключений до входу-виходу мікроелектронно-обчислювальної машини, додатково введені світлочутливий елемент, розміщений на матриці світлодіодів, виходом підключеним до перетворювача струм-напруга, вихід якого підключений до першого входу керованого перемикача, до другого входу якого підключений генератор опорної напруги, а вихід підключений до каналу перетворення, який складається із послідовно з'єднаних вибіркового підсилювача частоти комутації, синхронного детектора, інтегратора та аналого-цифрового перетворювача, вихід якого підключений до другого входу мікроелектронно-обчислювальної машини, а управляючі входи керованого перемикача та синхронного детектора підключені до другого виходу блоку контролю та керування.

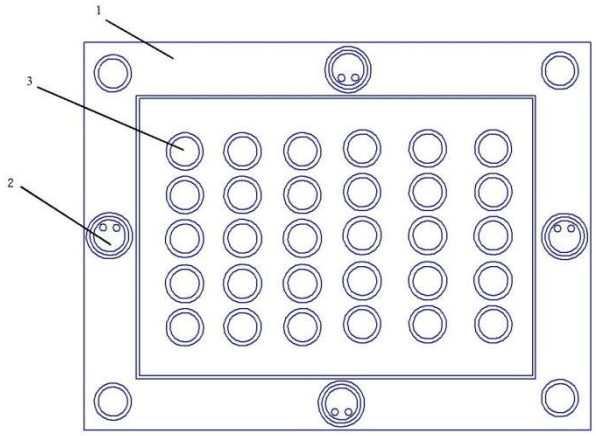


Рис. 1. Структурна схема світлодіодної матриції. 1–корпус світлодіодної матриці, 2–фотоприймач, 3–світлодіод;

На рис. 2 представлена розроблена авторами функціональна схема пристрою світлотерапії для світлотерапії, яка реалізує поставлену мету.

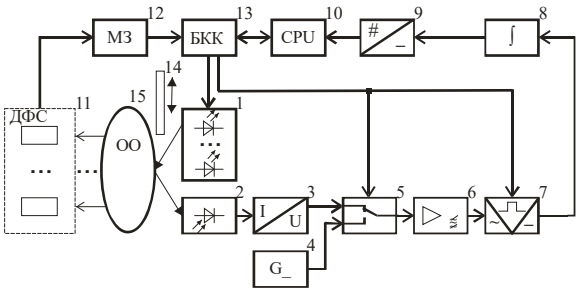


Рис. 2. Структурна схема пристрою світлотерапії.

При цьому на виході перетворювача «струм-напруга» 3 реєструється відповідне значення напруги:

$$U_1 = I_1 R_N \quad (1)$$

.

де I_1 - струм через світлочутливий елемент 2 при максимальній інтенсивності світлового потоку світлодіодної матриці 1; R_N - вхідний опір перетворювача 3.

Напруга U_1 подається на перший вхід керованого перемикача 5. На другий вхід керованого перемикача 5 подається напруга U_0 із генератора опорної напруги 4.

Це значення напруги U_0 вибирається зі співвідношення:

$$U_0 = U_T = I_T R_N \quad (2)$$

.

де I_T - темновий струм через світлочутливий елемент 2; U_T - напруга, що пропорційна I_T .

На управляючий вхід керованого перемикача 5 із другого виходу блоку комутації та керування 13 з частотою Ω подаються прямокутні імпульси, які забезпечують періодичне перемикання контактів керованого перемикача 5.

В один напівперіод комутації на виході керованого перемикача 5 утворюється напруга (1), а в другий напівперіод комутації – відповідно напруга (2).

В результаті періодичного перемикання перемикача 5 з послідовності імпульсів напруг U_1 і U_0 вибірковим підсилювачем частоти комутації 6 виділяється і підсилюється змінна складова напруги амплітудою, що визначається за формулою

$$U_2 = K_1 \frac{U_1 - U_0}{2}, \quad (3)$$

.

де K_1 - коефіцієнт підсилення вибіркового підсилювача 6 частоти комутації.

Змінна напруга після підсилення випрямляється синхронним детектором 7, який керується напругою частоти комутації керованого перемикача 5. Випрямлена напруга накопичується в інтеграторі 8, в якому придушуються низькочастотні шуми перетворювальних і підсилювальних блоків. На виході інтегратора 8 виділяється постійна різницева напруга, яка визначається за формулою

$$U_3 = \frac{K_1 K_2 K_3}{2}(U_1 - U_0) \quad (4)$$

.

де K_2 - коефіцієнт випрямлення синхронного детектора 7; K_3 - коефіцієнт передачі інтегратора 8.

При подальшому перетворенні в аналого-цифровому перетворювачі 9 формується код, що відповідає максимальній інтенсивності світлового потоку світлодіодної матриці 1.

$$N_1 = \alpha U_3 = \frac{SK_1K_2K_3}{2}(U_1 - U_0) \quad (5)$$

.

де α – сумарний коефіцієнт перетворення вимірювального каналу; S - масштабний коефіцієнт аналого-цифрового перетворення. Код (5) записується в пам'ять мікроЕОМ 10.

На другому етапі дзеркало прибирається і опромінюється вибрана ділянка шкіри об'єкту 15 із заданою інтенсивністю.

При виконанні перетворень, аналогічних першому етапу, на виході аналого-цифрового перетворювача формується код, що відповідає інтенсивності світлового потоку світлодіодної матриці 1, відбитого від об'єкта опромінення 15:

$$N_2 = \alpha U_3' = \frac{SK_1K_2K_3}{2}(U_1' - U_0) \quad (6)$$

.

де U_1' - напруга, обумовлена інтенсивністю відбитого випромінювання від об'єкта опромінення 15; U_3' - значення різницевої напруги, яке відповідає вимірюванню сигналу на другому етапі. Отриманий код (6) записується в пам'ять мікроЕОМ 10.

Різниця між результатами вимірювань на першому та другому етапах, яку вираховує мікроЕОМ 10 характеризує поглинуту об'єктом опромінення 15 відповідну частину світлового потоку світлодіодної матриці 1:

$$N_3 = N_1 - N_2 = \alpha(U_3 - U_3') \quad (7)$$

.

Інформація про здатність досліджуваної ділянки шкіри про поглинання вноситься в базу даних, наприклад, в електронну карту пацієнта.

IV. Результати та їх обговорення

Таким чином, запропонований пристрій для світло терапії забезпечує можливість вимірювання потужності світлового потоку, а за рахунок розробленого алгоритму вимірювання - контроль за рівнем поглинання оптичного опромінювання в процесі терапевтичної процедури і, відповідно, можливість оцінки біологічного ефекту та прогнозування результату опромінення. Важливим є також те, що використання комутаційно-модуляційного методу перетворення, дозволяє значно підвищити чутливість аналізу та проводити обробку сигналів низької інтенсивності потужністю $10^{-6} – 10^{-10}$ Вт.

V. Висновки

- Результати вимірювання поглинальної здатності, визначених для лікування ділянок шкіри, біологічно активних зон пацієнта фіксуються в пам'яті комп'ютера і можуть повторно використовуватися лікарем на наступних сеансах лікування, що значно спрощує процес лікування.
- Окрім того цей параметр може бути використаний як діагностичний при повторних сеансах із зафіксованими змінами.
- Запропонований пристрій для світлотерапії значно спрощує процес установки лікувальних параметрів, підвищує його об'єктивність і забезпечує можливість визначення кореляційної залежності між рівнем поглинання та ефективністю світлотерапії.

Література References

[1] А.М. Коробов, Фототерапевтические апараты Коробова серии «Барва»,.А.М.Коробов, В.А. Коробов, Т.А. Лесная. – Харьков, ХНУ им. В.Н. Каразина. – 2008. – 175 с.

[2] О.Т Кожухар, Можливості електронної системи із зворотнім зв'язком для лікування вушних шумів / М. Зазуляк, О.Т. Кожухар, Є.Р. Косий // Матеріали III конф. з міжнарод. участю «Актуальні проблеми біомедичної інженерії, інформатики, кібернетики і телемедицини», Київ, 11-13 березня 2010. - С. 53-56.

[3] М.С Івах, Підвищення ефективності фотомедичних технологій біоінформативною компонентою/ О.Т. Кожухар, М.А. Костів, М.С. Івах, А.М. Зазуляк, І.В. Мельник // "Медична та біологічна інформатика і кібернетика: Віхи розвитку" : конф. з міжнарод. участю, 20-23 квіт. 2011 р. : Збірник праць. - К.: НМАПО імені П.Л. Шупика, 2011. - С. 32.

[4] В.А. Андрійчук, Аналітичний метод світлотехнічного розрахунку для опромінювальних пристроїв з несиметричним світловим розподілом /Андрійчук В.А., Зеленков І.А.// Матеріали IV Міжнародної науково-технічної конференції "АВІА-2002" 23-25 квітня 2002 року, м.Київ: Національний авіаційний університет. – 2002р. – Т.2 Аерокосмічні системи моніторингу та керування, Секція: Авіоніка. - С. 23.79-23.82

[5] Л.П. Круковская, Ультрафиолетовое излучение: - его биологическая воздействие, приемники: Методическое пособие. - СПб.: СПбТПУ, 2009. - 26 с.

[6] М.С Івах, Підвищення ефективності фотомедичних технологій біоінформативною компонентою/ О.Т. Кожухар, М.А. Костів, М.С. Івах, А.М. Зазуляк, І.В. Мельник // "Медична та біологічна інформатика і кібернетика: Віхи розвитку" : конф. з міжнарод. участю, 20-23 квіт. 2011 р. : Збірник праць. - К.: НМАПО імені П.Л. Шупика, 2011. - С. 32.

Моделювання Елементів Мікросистем-на-Кристалі зі Структурою КНІ для Неінвазивних Глюкометрів

Ігор Когут
кафедра комп'ютерної інженерії та електроніки
Прикарпатський національний університет
Івано-Франківськ, Україна
mikotyk@mail.ru

Віктор Довгий
кафедра комп'ютерної інженерії та електроніки
Прикарпатський національний університет
Івано-Франківськ, Україна

Віктор Голота
кафедра комп'ютерної інженерії та електроніки
Прикарпатський національний університет
Івано-Франківськ, Україна

Михайло Котик
кафедра комп'ютерної інженерії та електроніки
Прикарпатський національний університет
Івано-Франківськ, Україна
mikot@i.ua

Simulation of Elements of System-on-Chip with the SOI-Structures for Non-Invasive Blood Glucose Meters

Igor Kogut, Victor Holota
dept. of Computer Engineering and Electronics
Precarpathian National University
Ivano-Frankovsk, Ukraine
mikotyk@mail.ru

Victor Dovhij
dept. of Computer Engineering and Electronics
Precarpathian National University
Ivano-Frankovsk, Ukraine

Victor Holota
dept. of Computer Engineering and Electronics
Precarpathian National University
Ivano-Frankovsk, Ukraine

Mykhailo Kotyk
dept. of Computer Engineering and Electronics
Precarpathian National University
Ivano-Frankovsk, Ukraine
mikot@i.ua

***Анотація*—В роботі наведено результати проектування та моделювання високочутливих елементів схем для систем неінвазивного контролю рівня глюкози в крові з використанням аналітичної мікросистеми-на-кристалі (АМнК) зі структурами кремній-на-ізоляторі (КНІ). Такі компоненти необхідні для забезпечення високого рівня чутливості, що є критично важливим для пристроїв такого типу, і водночас для зменшення похибки вимірювань.**

***Abstract*— In this paper design and simulation of highly sensitive circuit elements for non-invasive monitoring of blood glucose using analytical microsystem-on-Chip (AMSoC) with the structures of silicon-on-insulator (SOI) are presented. Such components are necessary for ensuring a high-sensitivity, which is critical for devices of this type, and at the same time to reduce measurement error.**

Ключові слова—аналітична мікросистема-на-кристалі, високочутливі елементи, кремній-на-ізоляторі, неінвазивні глюкометри, операційний підсилювач

Keywords—analytical microsystem-on-chip, highly sensitive elements, silicon-on-insulator, non-invasive blood glucose meters, operational amplifier

I. Вступ

В сучасній медицині набувають все більшої актуальності оптико-електронні неінвазивні методи «in

vivo» моніторингу показників крові людини: зокрема кисню, глюкози, холестерину та інш. Проте, оптико-електронні неінвазивні методи і пристрої мають вагомий недолік, порівняно з інвазивними – значно нижчу точність вимірювання (для неінвазивних вимірювачів рівня глюкози в крові допустима похибка точності результатів знаходиться на рівні 20 %). Також такі пристрої мають вищу складність будови і складніші безконтактні оптичні механізми отримання даних про рівень концентрації компонент крові.

Важливими первинними чутливими елементами (ЧЕ) для неінвазивних пристроїв контролю рівня глюкози в крові людини є фотодіоди і фотоприймачі, які функціонують на довжині хвилі λ=940 нм, на якій спостерігається максимальне значення поглинання глюкози та мінімальний вплив інших факторів, наприклад, таких як вплив поглинання водою з максимальною інтенсивністю на довжині хвилі 960 нм. Для портативних систем неінвазивного моніторингу рівня глюкози в крові з автономним джерелом живлення важливим критичним параметром є споживана потужність.

Для реалізації схем первинної обробки інформації від ЧЕ розроблено функціональну схему з використанням АМнК та проведено моделювання її складових інтегральних елементів на основі КНІ-структур. Зокрема, промодельовано електричні та часові характеристики елементів тракту проходження сигналу від сенсорного елемента неінвазивного пристрою вимірювання глюкози в крові, операційні підсилювачі, перетворювачі рівнів сигналів тригерного типу та ін. [1,2].

II. ФУНКЦІОНАЛЬНА СХЕМА ОПРАЦЮВАННЯ ІНФОРМАЦІЇ ДЛЯ НЕІНВАЗИВНИХ ГЛЮКОМЕТРІВ З ВИКОРИСТАННЯМ АМнК

Розроблена функціональна схема опрацювання інформації для неінвазивних глюкометрів з використанням АМнК зі структурами КНІ зображена на рис. 1. Використання КНІ-структур є найбільш перспективними в цьому плані завдяки їх переважаючим властивостям щодо електричних характеристик, стійкості до зовнішніх впливів, споживаної потужності та швидкодії, які є суттєво кращими порівняно з відомими на об'ємному кремнії. [3].

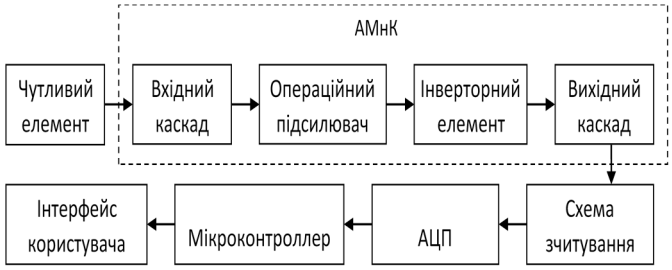


Рис. 1. Функціональна схема опрацювання інформації для неінвазивних глюкометрів з використанням АМнК

Для проектування систем неінвазивного контролю рівня глюкози в крові розроблено відкриту програмовану апаратну платформу на основі мікроконтролера ATMega328p, а реалізацію схем первинного опрацювання сигналів від сенсорів – на основі елементів аналітичної мікросистеми-на-кристалі (АМнК). Така платформа може бути використана також для роботи з іншими фізичними та біомедичними об'єктами обробки сигналів. Наприклад, для проектування джерела широтно-імпульсного модульованого (ШІМ) сигналу, який є складовою частиною систем неінвазивного контролю рівня глюкози в крові. Сигнал, отриманий від оптичних сенсорів - випромінюючого світлодіода та чутливого фотоелемента є імпульсним, тому важливим є також зменшення тривалості фронтів імпульсів, перед подачею їх на аналоговий вхід схеми зчитування мікроконтролера.[2].

III. ОПЕРАЦІЙНИЙ ПІДСИЛЮВАЧ

Одним з основних елементів схеми підсилення та покращення точності чутливого елемента є операційний підсилювач. Спроектована топологія базового елемента операційного підсилювача на основі фрагмента базової матричної комірки згідно схеми рис. 2, а, зображено на рис. 2, б. Було розроблено і досліджено два варіанти топологій: з використанням стандартної КМОН-технології на основі об'ємного кремнію та на основі КНІ КМОН-технології. [3].

Для первинної аналогової обробки сигналів одним із базових елементів є інтегральний операційний підсилювач (ОП). Такий елемент являє собою схему порівняння двох сигналів та підсилення різниці їх напруг. Схема електрична принципова базового елемента ОП в інтегральному виконанні зображена на рис. 2 а.

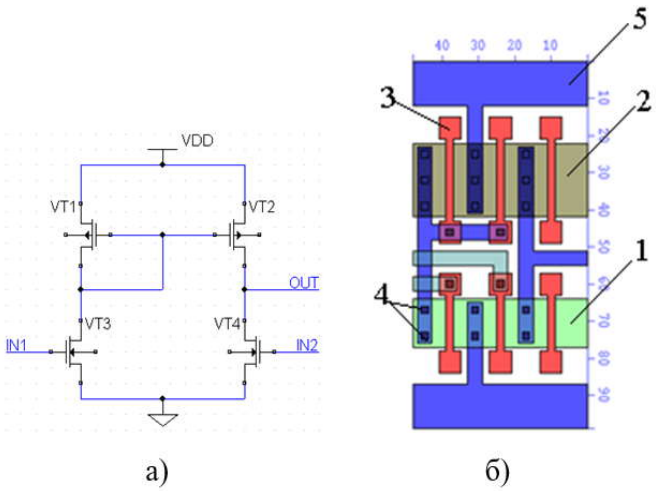


Рис. 2. Електрична схема а) та топологія б) інтегрального ОП на основі фрагмента базової матричної комірки зі структурою КНІ: 1 – стік-витокові області n-канальних транзисторів; 2 – стік-витокові області p-канальних транзисторів; 3 – полікремнієві затвори; 4 – контактні вікна; 5 – шари металізації та результат.

Об'ємне зображення інтегрального базового елемента операційного підсилювача зі структурою КНІ зображено на рис. 3, а, а порівняльні результати схемотопологічного моделювання базового елемента ОП безпосередньо із топологій, спроектованих на основі структур за стандартною об'ємною КМОН-технологією та на основі КМОН КНІ-структур, наведено на рис. 3.б. Моделювання

проводились при напрузі живлення 1,5 В, вхідний сигнал IN1 являв собою змінний сигнал частотою 0,167 МГц, часом наростання та спадання 1 нс, тривалістю імпульсу 2 нс та амплітудою 0,04 В а IN2 – постійний сигнал з амплітудою 0,7 В.

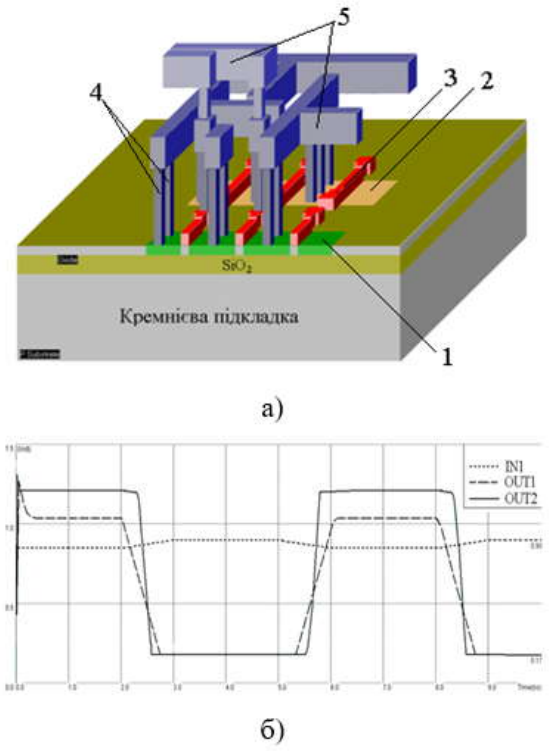


Рис. 3. Об'ємне зображення а) інтегрального ОП на основі фрагмента базової матричної комірки зі структурою КНІ: 1 – стік-витокові області n-канальних транзисторів; 2 – стік-витокові області p-канальних транзисторів; 3 – полікремнієві затвори; 4 – контактні вікна; 5 – шари металізації та результати моделювання б) ОП на основі КНІ та об'ємних КМОН-структур: 1 – вхідний сигнал IN1; OUT1 – вихідний сигнал ОП, спроектованого за об'ємною КМОН технологією; OUT2 – вихідний сигнал.

IV. РЕЗУЛЬТАТИ МОДЕЛЮВАННЯ

Результати схемотопологічного моделювання, безпосередньо із топології з врахуванням паразитних зв'язків та впливу взаємного розміщення елементів топології базового ОП показують, що вихідні сигнали для схеми зі структурами КНІ порівняно з об'ємними КМОН мають суттєво кращу в середньому на 30% крутизну фронтів та більший коефіцієнт підсилення, завдяки чому вихідний сигнал має рівень амплітуди на 20% вищий. Такі параметри дозволять значно зменшити споживану потужність під час перехідних процесів (менша тривалість фронтів імпульсів), і як наслідок, підвищити ступінь інтеграції елементів в АМнК. Окрім цього, реалізація ОП на основі КНІ-структур дозволить розширити температурний діапазон їх використань.

Аналогічно спроектовано і промодельовано перетворювач рівнів сигналів тригерного типу з низького рівня до високого (з 2.5В до 5 В), універсальний JK-тригер та вихідний формувач сигналів з контактною площадкою.

Аналогічним чином було промодельовано КМОН КНІ перетворювач рівнів сигналів з високого рівня до низького (з 5В до 2В). Такі перетворювачі потрібні для зменшення амплітуди вхідних сигналів від зовнішніх вхідних інтерфейсних схем і їх наступної передачі у внутрішні елементи АМнК, які функціонують при малих напругах живлення. Результати схемотехнічного моделювання такого перетворювача зображено на рис. 4, б. Суцільною лінією зображено вхідний сигнал з амплітудою 5 В, а штрих-пунктирною – перетворений вихідний сигнал з амплітудою 2 В

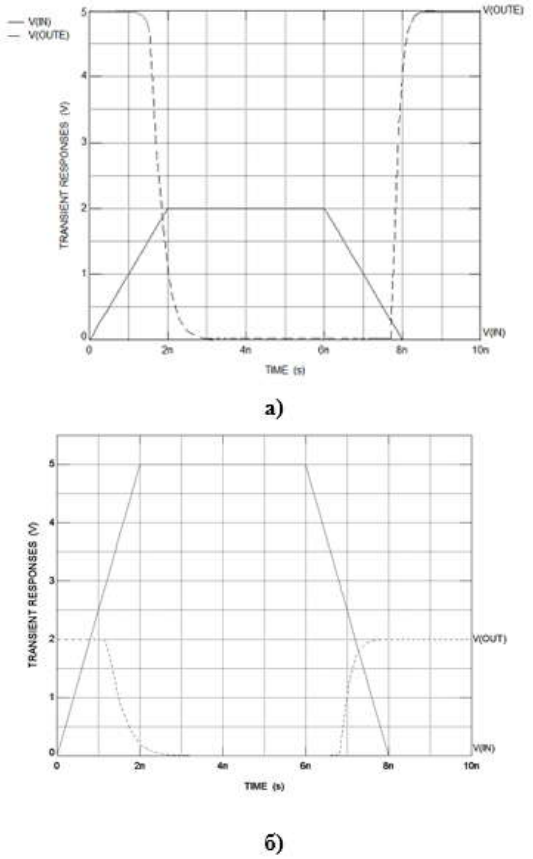


Рис. 4. Часова діаграма вхідного (суцільна лінія) та вихідного (пунктирна) сигналів перетворювачів рівнів з низького на високий (а) та з високого на низький (б).

В ІС з високим рівнем інтеграції і які мають декілька напруг живлення на одному кристалі існує необхідність перетворення сигналів з менш нижчих рівнів і подальшого використання таких перетворювачів в одній ІС або мікросистемі-на-кристалі. Для цього було спроектовано і досліджено шляхом моделювання топологію даного перетворювача рівнів сигналу відповідно до електричної схеми рис. 6, а, але не на базовій матричній комірці, як у попередньому випадку, а на спеціалізованій топології і з меншими 180 нм проектними нормами та двома рівнями металізації для перетворення сигналів з вхідною амплітудою 1,2 В і вихідною 2,5 В. Спеціалізована топологія є більш трудоємкою, але більш оптимальною за

площею на кристалі, узгодженістю та симетрією параметрів. Топологію цього перетворювача зображено на рис. 5.

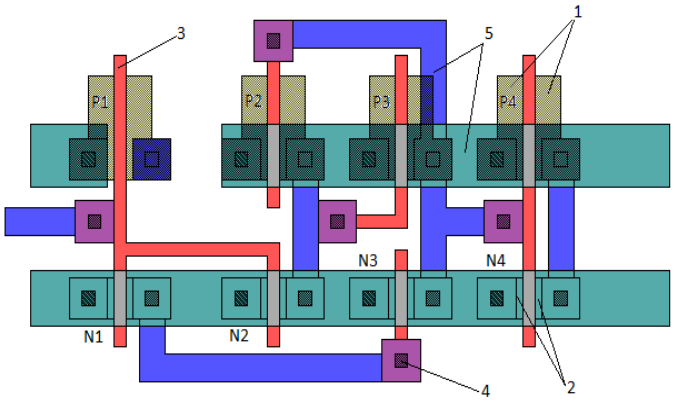


Рис. 5. Спеціалізована топологія перетворювача рівнів сигналів

Такий перетворювач рівнів забезпечує перетворення невеликих перепадів рівнів сигналів, а затримка вихідного сигналу відносно вхідного на рівні 0,5 амплітуди по передньому фронту становить не більше 20 пс.

Для побудови цифрових елементів обробки сигналів, як складових аналітичних мікросистем-на-кристалі зі структурою КНІ для неінвазивних глюкометрів було розглянуто особливості схемотехнічного і топологічного проектування JK-тригера на основі матричної комірки та проведено моделювання його електричних, часових, температурних та енергетичних характеристик. Цей тип тригера було обрано тому, що він належить до розряду універсальних тригерів, і на його основі можна побудувати тригери RS-, D- і T-типу та інші, більш складні схеми цифрової обробки сигналів.

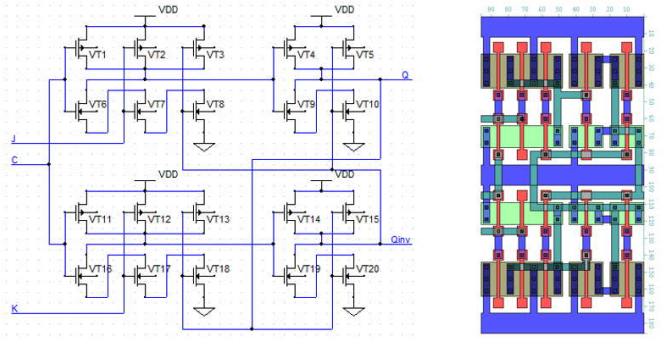


Рис. 6. Електрична схема JK-тригера на основі логічних елементів I-НЕ (а) та топологія JK-тригера на основі базових матричних комірок (б).

Для дослідження характеристик JK-тригера з врахуванням паразитних зв'язків і особливостей топології відповідно були розроблені спеціалізована топологія з використанням об'ємних КМОН транзисторів, та топологія на основі базової комірки зі структурою КНІ. Топологію JK-тригера на основі базової комірки зображено на рис. 6, б. Ширини каналів для всіх p-канальних транзисторів становлять 20 мкм, n-канальних транзисторів – 10 мкм, довжини каналів всіх транзисторів є однаковими і становлять 2 мкм. Топологічна реалізація JK-тригера здійснена змінними програмованими шарами 2-х рівнів металізації та 2-х рівнів контактних вікон.

V. ВИСНОВКИ

Розроблено функціональну схему програмованої апаратної платформи для портативних неінвазивних глюкометрів на основі мікроконтролера ATMega328p, яка опрацьовує первинну інформацію з АМнК зі структурами КНІ. Розроблено і промодельовано інтегральні вхідний та вихідний каскади з контактними площадками, операційний підсилювач, перетворювач рівнів тригерного типу. Результати моделювання показують, що запропоновані елементи схем опрацювання сигналів для неінвазивних глюкометрів зі структурами КНІ мають покращену швидкодію, в середньому на 30% крутіші фронти та більший коефіцієнт підсилення, завдяки чому вихідний сигнал має рівень амплітуди на 20% вищий, порівняно з об'ємними КМОН-структурами. Це відкриває додаткові можливості для зменшення споживаної потужності під час перехідних процесів (менша тривалість фронтів імпульсів), і як наслідок, підвищення ступені інтеграції елементів в АМнК та розширення температурного діапазону їх використань. Спроектовано еквівалентні між собою топології вхідних каскадів аналітичної мікросистеми-на-кристалі, як на основі стандартної об'ємної КМОН-технології, так і на основі КМОН КНІ-структур. Проведено їх схемо-топологічне моделювання. Показано, що вихідні каскади на КНІ-структурах мають меншу затримку вихідного сигналу відносно вхідного (4 пс та 7 пс відповідно) та меншу споживану потужність (6,89 мВт та 8,88 мВт відповідно) порівняно з об'ємною КМОН-технологією. Споживана потужність є важливим критичним параметром для портативних систем неінвазивного моніторингу рівня глюкози в крові з автономним джерелом живлення.

ЛІТЕРАТУРА REFERENCES

[1] А. В. Белов, Микроконтроллеры AVR от азов программирования до создания практических устройств, М:, Академия. -2008.

[2] Котик М. В. Розробка і моделювання високочутливих компонентів схем для систем неінвазивного контролю рівня глюкози в крові / М. В. Котик, І. Т. Когут, В. В. Довгий, В. М. Грига // V-та міжнар. науково-практ. конф. «Фізико-технологічні проблеми передавання, оброблення та зберігання інформації в інфокомукаційних системах». – Чернівці, 2016. – С. 130-131.

[3] Druzhinin A. Electrical and layouts simulation of analytical microsystem-on-chip elements for high frequence and low temperature applications / A. Druzhinin, Yu. Khoverko, V. Dovhij, I. Kogut, V. Holota // UkrMiCo'2016. – Kyiv, 2016. – P. 29-32

[4] Довгий В. В. Розробка і моделювання елементів аналітичної мікросистеми-на-кристалі зі структурами "кремній-на-ізоляторі" / В. В Довгий, І. Т. Когут, В. І. Голота // "Фізика і хімія твердого тіла". – Івано-Франківськ, 2016. – № 2. – С. 275-280.

Побудова Системи Моніторингу та Керування Станом Навколишнього Середовища на Основі Технології IoT

Н. Бондаренко
кафедра системного проектування
Національний технічний університет України "КПІ імені І.Сікорського"
Київ, Україна
nataliya.bondarenko17@gmail.com

Designing of System for Environment State Monitoring (Based on Internet of Things Technology)

N. Bondarenko
Computer aided design department
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"
Kyiv, Ukraine
nataliya.bondarenko17@gmail.com

***Анотація*—У даній роботі запропоновано підхід до побудови системи моніторингу навколишнього середовища, а саме – посівних угідь. Проаналізовано доцільність використання технології Інтернет Речей у сукупності з пристроєм Raspberry Pi із набором потрібних сенсорів у якості автономної станції. Також досліджено можливості використання мікросервісної архітектури як основної для проектування масштабованого серверного додатку.**

***Abstract*— This paper proposes an approach to building a system for monitoring the state of environment, crop lan, in particular. Feasibility of using the Internet of Things technology in conjunction with a device Rapsberry Pi with a set of sensors required in a stand-alone station was analyzed. The possibility of using architecture based on microservices as a basic design principle for a scalable server application was investigated.**

Ключові слова—Інтернет речей, мікросервіс, пристрій, датчик, Docker, Raspberry Pi.

Keywords—Internet of Things, microservice, device, sensor, Docker, Rapsberry Pi.

I. Вступ

Комп'ютерні науки вже давно перестали забезпечувати лише проблеми, пов'язані з військовою галуззю, як це було на початкових етапах їх розвитку більш ніж півстоліття тому. Наразі вони все більше інтегруються в інші прикладні галузі, пропонуючи свій інструментарій для полегшення праці людей. Аграрна промисловість стала однією з тих галузей, вплив на які інформаційних та інженерних розробок залишається досить слабким, особливо в Україні. Різноманітні автоматизовані ферми, "розумні" теплиці все ще залишаються поодиноким явищем на теренах України. Подальший розвиток сільського господарства вимагає використання сучасних технологій і обладнання, які економлять ресурси і підвищують ефективність виробництва. В усьому світі зростає попит на послуги, що дають змогу зібрати й опрацювати зібрані дані в інтересах точного землеробства. Такі дані можуть бути використані для планування та контролю етапів сільськогосподарського виробництва, економії посівного матеріалу та добрив, що дозволить ефективно використати час і грошові ресурси, а також покращити стан навколишнього середовища.

Однією з іновацій, що починає активно захоплювати аграрний ринок (принаймні у США) є безпілотні технології. Використання дронів дозволяє аграріям моніторити стан якість оранки й посіву, прогнозувати врожай, захищати ділянки від пожеж та розкрадань і навіть «точково» поліпшувати стан ґрунту, вносячи мікродобрива. Поки що ця технологія має низку недоліків: труднощі у закладенні маршруту для дрона, залежність від погодних умов тощо.

Дана робота присвячена дослідженню можливості проектування системи наземного моніторингу середовища, що буде надавати дані про його стан, а також аналізувати дані по мірі їх накопичення. Труднощі і водночас

іноватівність поляють у тому, що подібні системи не використовувались раніше для моніторингу на великих площах.

II. Технологія Internet of Things та її використання в контексті робробки системи моніторингу стану середовища

Інтернет речей (IoT) – це мережева парадигма, що передбачає наявність взаємопов'язаних, “розумних” об'єктів, які безперервно генерують дані і передають їх через Інтернет. Стратегія роботи з даними для IoT має декілька етапів: збір, обробку, розповсюдження та використання даних. [1]

Концепція IoT не є новою. Програмований логічний контролер (ПЛК) з 1970-их років є мікро-моделлю системи IoT, яка широко використовувалась для управління верстатами і процесами на заводах. ПЛК є системою, що складається з входів (датчиків, приводів, вимикачів), виходу (дані у цифровому або аналоговому вигляді), процесора, і включає зв’язок між ними. Такі системи знаходилися в межах заводу і не були підключеними до мережі Інтернет. Наразі існує цілий ряд технологій, які роблять втілення концепції Інтернету Речей у життя можливим. [2]

ТАБЛИЦЯ I. Технології, що є підгрунтям IoT систем (відповідно до досліджень, проведених компанією Deloitte

Технологія	**Визначення**
Сенсор	Пристрій, який генерує електричний сигнал від фізичного стану або події.
Мережа	Механізм для передачі електронного сигналу.
Механічно доповнений інтелект	Аналітичні інструменти, які покращують здатність описувати, передбачати і використовувати відносини між явищами.
Аргументована поведінка	Технології та методи, які дозволяють виконувати дії, базуючись на типовій поведінці людини у даній ситуації.

Технології, зазначені у таблиці стають з часом все більш доступними. Нова версія Інтернет-протоколу (IPv6), що підтримує адреси розміром 128-біт, що відповідає 340 трильйонам адрес ($3{,}4 \times 10^{38}$), а це 5×10^{28} адрес на кожну людину, дозволяє практично необмежену кількість пристроїв, підключених до мереж. Ціни датчиків суттєво знизилися за останні десятиліття. Завдяки закону Мура розмір та ціна інтегрованих процесорів знизилися, а їх можливості достатньо зросли, щоб задовольнити вимірювання з великою точністю та навіть первинну обробку вимірювань. Ціни модулів підтримки бездротових мереж (Bluetooth, Wi-Fi, 3G) також знижуються в геометричній прогресії. Така доступність технологій дозволяє розробити достатньо потужну IoT мережу з мінімальними витратами.

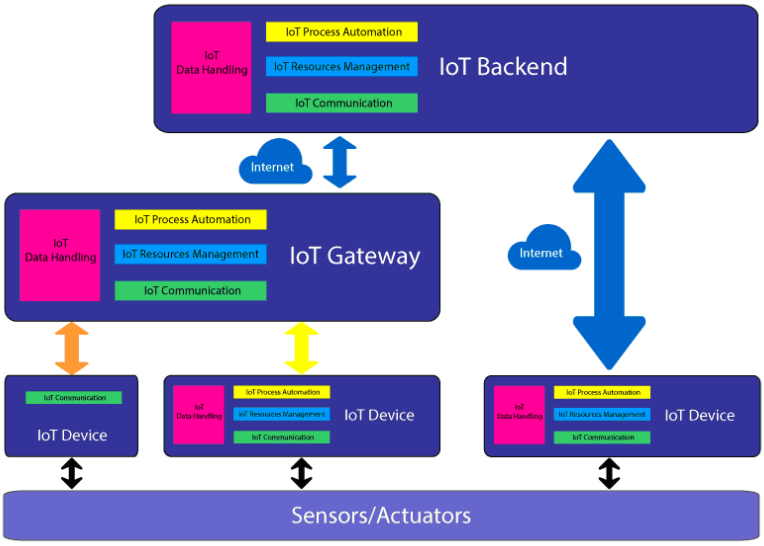


Рис. 1. Схема системи побудованої відповідно до технології Internet of Things

Система, побудована за допомогою технології Internet of things може мати різну архітектуру в залежності від ролей, відведених її компонентам. [3]

На рис. 1 зображено чотири основні компоненти, зв'язки між ними та функції, які вони можуть виконувати у рамках загальної мережі.

A. Складові системи IoT

Сенсори. В залежності від цілі та призначення мережі сенсори можуть відрізнятися за часом життя, чутливістю, часом відповіді, рівнем енергоспоживання, вартістю, часом життя та призначенням.

Пристрій (device, thing). Електронний пристрій: мікрокомп'ютер, платформа з мікроконтролером, телефон, смартфон, натільний пристрій, який часто безпосередньо поєднується з сенсорами і отримує від них дані.

Пристрій-шлюз (gateway). Використовується в деяких IoT системах для проміжного збору та обробки даних. Прикладом може слугувати система "розумного будинку", в якому всі пристрої (телефони, холодильники, мікрохвильовки тощо) посилають дані в локальній мережі на проміжні пристрої, наприклад смартфони, які, в свою чергу, посилають оброблені дані з певною систематичністю на сервер через мережу Інтернет.

Сервер, хмара, сховище даних (backend). В залежності від призначення та об'ємів IoT мережі дані можуть оброблятися та зберігатися в різних середовищах. Це можуть бути, наприклад, високопотужні сервери з встановленими системами управління нереляційними базами даних у випадку, якщо обсяги даних дійсно є дуже великими та потребують особливих методів зберігання.

B. Різновиди мереж в залежності від завдань, що виконуються складовими мережі

Пристрій-Сервер. Система, в якій пристрій, що збирає дані, має достатню обчислювальну здатність, щоб виконувати первинну обробку отриманих від сенсорів показань. Після обробки пристрій відправляє дані на сервер або тримає їх у тимчасовій пам'яті до запиту сервера.

Передавач-Шлюз-Сервер. Система, у якій пристрій, що збирає дані з датчиків, не виконує ніяких функцій, крім передачі даних на шлюзовий пристрій, який вже потім виконує попереднє очищення, обробку та впорядкування даних. Приклад такої системи був наведений в описі пристрою-шлюзу.

Пристрій-Шлюз-Сервер. Кожна складова такої системи може виконувати обробку даних на різному рівні та контролювати процес комунікації з іншими складовими.

Гібридна система. Така мережа будується з використанням декількох видів зв'язків між пристроями, але з дотриманням ієрархії, що є для них загальною.

C. Вибір виду мережі для проектування системи моніторингу та керування навколишнього середовища.

Для розробки системи моніторингу оберемо різновид IoT мережі Пристрій-Сервер. Роль пристрою буде виконувати мікрокомп'ютер Raspberry Pi. Так як даний мікроком-п'ютер має бібліотеки та може бути настроєний таким чином, щоб виконувати процес збору та упорядкування даних для подальшого відправлення на сервер. Система не передбачає створення локальних мереж, а лише передачу даних за допомогою GSM-мережі, тому введення проміжних пристроїв є недоцільним.

III. Планування серверної архітектури для розроблюваної системи

Для планування архітектури серверу перш за все потрібно визначити задачі, які він має виконувати.

A. Задачі, які має виконувати сервер.

Збір даних з усіх пристроїв системи.

Зберігання даних про стан кожного пристрою та його показання.

Представлення даних у вигляді, зрозумілому для користувача.

Побудова на основі отриманих даних висновків про стан навколишнього середовища та генерування рекомендацій стосовно ведення певного виду господарської діяльності на ділянці.

B. Вибір архітектури.

Коло задач, що виконуються сервером, є досить великим і може бути доповнене іншими задачами в майбутньому. Також для зберігання великих обсягів даних може бути замало однієї бази даних.

Враховуючи вищезазначені вимоги до сервера можна зробити припущення, що монолітна архітектура є недоцільною в даному випадку і треба віддати перевагу мікросервісній архітектурі. Монолітне програмне забезпечення завжди будується як автономна одиниця.

Основною проблемою такої архітектури є те, що всі цикли обробки інформації дуже тісно зв'язані один з одним. Зміни, внесені в одну невелику частину програми, можуть вимагати повторної збірки та встановлення на серверах нової версії ПЗ. За умов масштабування певної частини додатку, доведеться використовувати збільшення кількості процесів цілої програми замість окремих компонент. У таких випадках вирішенням усіх проблем може стати використання мікросервісної архітектури.

Мікросервісна архітектура або, мікросервіси - це один із підходів розробки програмного забезпечення(ПЗ), який набув широкого розповсюдження за останні роки. Мікросервісна архітектура розглядається як набір незалежно розміщених невеликих сервісів, кожен з яких має свій процес і які спілкуються через чітко визначений легкий протокол для обслуговування потреб. [4]

Контейнери є одним із засобів побудови модульної архітектури. Контейнер являє собою самодостатнє середовище виконання програми, що ізольоване від операційної системи. Найпоширенішою реалізацією контейнерів є Docker контейнери. Docker дозволяє автоматизувати їх створення та дає інструменти для управління ними.

Для реалізації архітектури ми також виберемо Docker. Схематично архітектуру можна представити таким чином як на Рис. 2.

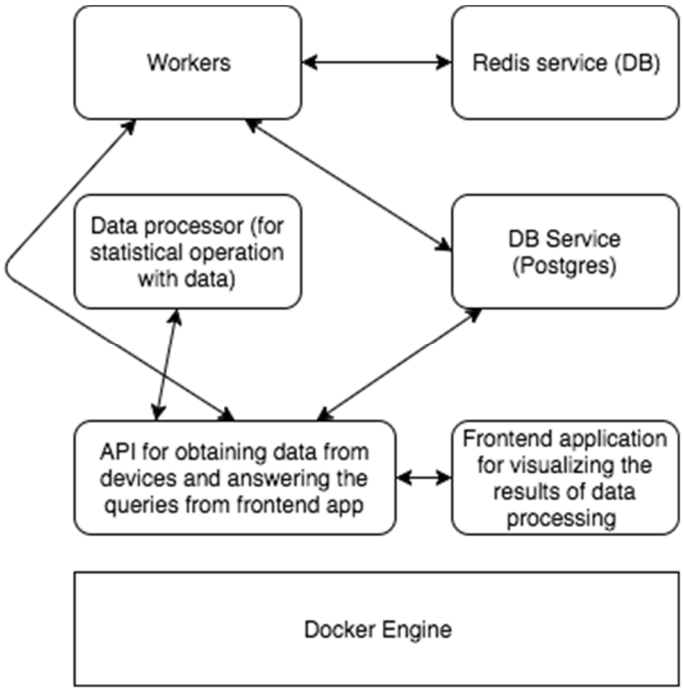


Рис. 2. Схема архітектури серверного додатку, що базується на мікросервісному підході

IV. Планування пристрою для розроблюваної системи

У якості пристрою ми обрали мікрокомп'ютер Raspberry Pi. Визначимо набір сенсорів, які будуть корисними у визначенні стану ґрунту та повітря.

A. Сенсори для Raspberry Pi

1) DHT2. Здійснює вимірювання температури та вологості від *0 до 100%* із точністю *2%.*

2) BMP180. Здійснює вимірювання тиску повітря. На Raspberry контролюється I2C модулем.

3) MQ-138 Gas Sensor. Здійснює вимірювання рівня різних газових сполук у повітрі. [5]

4) Moisture Sensor. Здійснює вимірювання вологості ґрунту.

5) MCP3008. Потрібен для перетворення аналогових сигналів у цифрові, так як два останні датчики з вищезазначених подають дані у аналоговому форматі.

Даний набір сенсорів дасть нам змогу побудувати початкову тестову модель пристрою для оцінки релевантності системи.

Зв'язок із сервером може здійснюватись за допомогою будь-якого 3G-модема з підключенням бібліотеки Sakis3g.

V. Висновки

У даній роботі було проаналізовано доцільність використання таких технологій як Internet of Things та мікросервіси для побудови системи моніторингу та контролю стану навколишнього середовища.

Також була запропонована модель серверу, спроектована з урахуванням використаних технологій. У роботі був описаний пристрій, який має виконувати збір та початкову обробку даних відповідно до технології Internet of Things. Проведені дослідження дозволяють розробити тестову модель системи і випробувати її у реальних умовах.

Література References

[1] Paul Kurchina, “The 2016 Mandate: Master The Internet Of Things Before Everything Is Connected” [Online]. Available: http://www.digitalistmag.com/digital-economy/2015/12/03/2016-mandate-master-nternet-of-things-before-everything-connected-03830995

[2] Mervat Abu-Elkheir, Mohammad Hayajneh, Najah Abu Ali, “Data Management for the Internet of Things: Design Primitives and Solution” [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3871070/

[3] Jayavardhana Gubbi, Rajkumar Buyya, Slaven Marusic, Marimuthu Palaniswami, “Internet of Things (IoT): A Vision, Architectural Elements, and Future Directions”, [Online]. Available: https://arxiv.org/pdf/1207.0203.pdf

[4] Nate Slater, “Using Containers to Build a Microservices Architecture” [Online]. Available: https://medium.com/aws-activate-startup-blog/using-containers-to-build-a-microservices-architecture-6e1b8bacb7d1 - .9gk2xjoyo

[5] MQ Gas sensors. [Online]. Available: http://playground.arduino.cc/Main/MQGasSensors - list

Формалізація Етапів Діагностично-Лікувальних Заходів при Проектуванні Систем Підтримки Прийняття Рішень в Медицині

Анатолій Поворознюк
кафедра обчислювальної техніки та програмування
Національний технічний університет
«Харківський політехнічний інститут»
Харків, Україна
ai.povoroznjuk@gmail.com

Ганна Філатова
кафедра обчислювальної техніки та програмування
Національний технічний університет
«Харківський політехнічний інститут»
Харків, Україна
filatova@gmail.com

Formalization of the Stages of Diagnostic and Treatment Activities in the Design of Decision Support Systems in Medicine

Anatoliy Povoroznjuk
Department of Computer Science and Programming
National Technical University
"Kharkiv Polytechnic Institute"
Kharkiv, Ukraine
ai.povoroznjuk@gmail.com

Anna Filatova
Department of Computer Science and Programming
National Technical University
"Kharkiv Polytechnic Institute"
Kharkiv, Ukraine
filatova@gmail.com

***Анотація*—Формалізовано етапи діагностично-лікувального процесу при проектуванні комп'ютерних систем підтримки прийняття рішень в медицині. Розроблено математичну модель процесу діагностики та лікарських дій з метою підвищення ефективності надання медичних послуг та мінімізації ризиків лікарських помилок.**

***Abstract*—The stages of the diagnostic and therapeutic process in the design of computer decision support systems in medicine are formalized. The mathematical model of the process of diagnostics and medical actions for the purpose of increasing the efficiency of providing medical services and minimizing the risks of medical errors is developed.**

Ключові слова—комп'ютерна система; діагностика; лікарська дія; математична модель; лікарська помилка.

Keywords— computer system, diagnostics, medical action, mathematical model, medical error.

I. Вступ

Комплекс діагностично-лікувальних заходів (ДЛЗ) складається з двох взаємопов'язаних етапів: діагностики захворювань і лікування виявлених патологій, причому після діагностики і призначення лікувальних процедур необхідний моніторинг поточного стану пацієнта з метою оцінки ефективності процесу лікування і, при необхідності, його корекції. Для лікування того чи іншого захворювання необхідно надання певних лікарських дій (ЛД) на організм (хірургічне втручання, фармакологічний, лікувально-терапевтичний вплив, реабілітаційні заходи).

На кожному із зазначених етапів лікар, як особа, яка приймає рішення, виробляє управлінське рішення в умовах дефіциту вихідних даних і суттєвої апріорної невизначеності, ґрунтуючись на своїй кваліфікації, досвіді і інтуїції. При цьому прийняття неправильного рішення (лікарська помилка) як на етапі діагностики, так і на етапі лікування може мати катастрофічні наслідки для здоров'я пацієнта. Термін «лікарська помилка» (ЛП) визначає неправильну діагностику хвороби або неправильні ЛД, які обумовлені добросовісною помилкою лікаря, при цьому виключається недбалість і несумлінність при виконанні своїх обов'язків. Причиною помилки діагностики є недостатній обсяг діагностичних даних, або їх невірна інтерпретація (особливо при суб'єктивному аналізі якісних показників). Причиною неправильних ЛД при правильному діагнозі є недостатнє врахування індивідуальних особливостей пацієнта (алергічні реакції на

певні препарати, список хвороб, якими вже хворів пацієнт, які ліки приймав і т.д.).

В даний час є широкий спектр комп'ютерних діагностичних систем в різних предметних областях медицини [1, 2], інформатизація ЛД обмежується медичними довідниками, в тому числі у вигляді інформаційно-пошукових систем [3]. В сучасних комп'ютерних системах підтримки прийняття рішень в медицині задачі діагностики і ЛД розглядаються незалежно один від одного, тому актуальною є задача мінімізації ризику ЛП при комплексній оцінці діагностичних і лікарських дій (ДЛД).

Метою роботи є формалізація етапів та розробка математичної моделі ДЛД з метою підвищення їх ефективності та мінімізації ризику лікарських помилок.

II. Розробка математичної моделі діагностично-лікувального процесу

На основі аналізу етапів ДЛД [2] пропонується наступна математична модель M_{da}:

$$M_{da} = \left\langle X, S, B, \tilde{S}, \tilde{B}, D, M_a, \tilde{D}, f, Q_f \right\rangle,$$

де $X = \{x_i(\cdot) \mid i \in \{1,2,\ldots,n_X\}\}$ – множина реєструємих біомедичних сигналів/зображень БМС/З (n_X – кількість БМС/З). Так як БМС є деякою функцією одного аргументу, а БМЗ – двох аргументів, то для уніфікації розгляду процесу обробки біомедицинських сигналів і зображень аргументи функції, що задає сигнал або зображення, вилучимо, тобто *i*-й БМС/З позначаємо $x_i(\cdot)$;

$S = \{s_j(\cdot) \mid j \in \{1,2,\ldots,n_S\}\}$ – множина діагностичних ознак (n_S – кількість діагностичних ознак);

B – множина можливих комбінацій діапазонів значень діагностичних ознак S;

$\tilde{S} = \langle \tilde{s}_1, \ldots, \tilde{s}_{n_S} \rangle$ – упорядкована множина значень діагностичних ознак, отриманих для конкретного пацієнта, включаючи інструментальне обстеження;

$\tilde{B} = \langle \tilde{b}_1, \ldots, \tilde{b}_{n_S} \rangle$ – упорядкована множина значень діапазонів для $\tilde{S}$;

$D = \{d_k \mid k \in \{1,2,\ldots,n_D\}\}$ – множина можливих діагнозів у заданій предметній області, тобто алфавіт діагнозів (n_D – кількість діагнозів);

$M_a = \{m_{a_k} \mid k \in \{1,2,\ldots,n_{M_a}\}\}$ – множина лікарських дій (n_{M_a} – кількість лікарських дій);

$\tilde{D} = \{\tilde{d}_j \mid j \in \{1,2,\ldots,n_D\}\}$ – нечітка множина діагностичних висновків;

f – множина відображень відповідних множин X, S, B, $\tilde{B}$, D, M_a;

Q_f – множина критеріїв відображень f;

$n_X, n_S, n_S, n_D, n_{M_a} \in N$; N – множина натуральних чисел.

Розглянемо більш детально елементи запропонованої математичної моделі M_{da}, які формалізують відповідні етапи ДЛЗ. У процесі інструментального обстеження реєструється множина БМС/З $X = \{x_i(\cdot) \mid i \in \{1,2,\ldots,n_X\}\}$, передбачених протоколом обстеження в даній предметній області. Наприклад, у результаті електрокардіографічного обстеження формується множина відведень ЕКГ, у результаті мамографічного обстеження – множина проекцій лівої й правої молочних залоз і так далі.

В результаті обробки кожного БМС/З $x_i(\cdot)$ визначається підмножина діагностичних ознак $S_i = \{s_{ij}\}, j = \overline{1,n_i}$, де n_i – кількість діагностичних ознак для *i*-го БМС/З, причому $S_i \cap S_j = \varnothing$, якщо $i \neq j$, і $\bigcup_{i=1}^{n_X} S_i = S$ – множина діагностичних ознак для всіх БМС/З, тобто для множини X.

Для кожної з діагностичних ознак s_j ($j = \overline{1,n_S}$) визначені діапазони значень $B_j = \{b_{jk}\} \mid k = 1, n_{B_j}$, де n_{B_j} – кількість діапазонів значень для діагностичної ознаки s_j. Розбивка припустимого інтервалу значень числової ознаки s_j на діапазони b_{jk} відповідає значенням типу «норма», «нижче/вище норми», «граничні значення», «небезпечні значення», «критичні значення» і так далі, прийнятим у заданій предметній області медицини. Число діапазонів n_{B_j} може бути різним у різних ознак.

Множина можливих комбінацій діапазонів діагностичних ознак це прямий (декартовий) добуток множин B_j, тобто $B = B_1 \times \ldots \times B_j \times \ldots \times B_{n_S} = \{\langle b_1, \ldots, b_j, \ldots, b_{n_S} \rangle \mid b_j \in B_j, j \in \{1,2,\ldots,n_S\}\}$.

В результаті реєстрації й аналізу БМС/З формується кортеж значень діагностичних ознак $\tilde{S} = \langle \tilde{s}_1, \ldots, \tilde{s}_{n_S} \rangle$ для конкретного пацієнта. Тоді кортеж оцінок відповідних діапазонів $\tilde{B} = \langle \tilde{b}_1, \ldots, \tilde{b}_{n_S} \rangle \in B$, це реалізація B для даного пацієнта, тобто представлення результатів вимірів точкою в просторі B.

Проекції B на можливі комбінації координатних осей утворюють множину можливих симптомокомплексів $C = \{pr_i B, pr_{i_1,\ldots,i_j,\ldots,i_q} B \mid i, i_j, q \in \{1,\ldots,n_S\}\}$. Із цієї множини виділяється підмножина інформативних симптомокомплексів $C' = \{c'_i : i = \overline{1,n_{C'}}\} \subseteq C$ ($n_{C'}$ – кількість інформативних симптомокомплексів) щодо системи діагнозів D.

Проекції $\tilde{B}$ на можливі комбінації координатних осей утворюють множину можливих кортежів, які складаються із можливих комбінацій діагностичних ознак

$$\tilde{C} = \left\{ pr_i \tilde{B}, pr_{i_1,\ldots i_j,\ldots,i_q} \tilde{B} \mid i,i_j,q \in \{1,\ldots,n_S\} \right\} = \left\{ \tilde{c}_i : i = \overline{1,n_{\tilde{C}}} \right\},$$

де $n_{\tilde{C}}$ – кількість комбінацій діагностичних ознак.

Тоді множину інформативних комбінацій ознак щодо системи діагнозів *D* можна визначити як

$$\tilde{C}' = \left\{ \tilde{c}_i \mid \tilde{c}_i \in C'_i,\ i = \overline{1,n_{C'}}\ j = \overline{1,n_{C'}} \right\} = \left\{ \tilde{c}'_i \mid i = \overline{1,n_{\tilde{C}'}} \right\}.$$

Для реалізації системи підтримки прийняття рішень в медицині (СППРМ) необхідною умовою є те, що множина можливих діагнозів $D = \{ d_k \mid k \in \{1,2,\ldots,n_D\} \}$, повинна утворювати повну групу несумісних подій, тобто

$$P\left(\bigcup_{k=1}^{n_D} d_k \right) = \sum_{k=1}^{n_D} P(d_k) = 1 \cdot$$

Нехай множина $D = \{ d_k \mid k \in \{1,2,\ldots,n_D\} \}$ є універсальною множиною діагнозів. Нехай задано сімейство характеристичних функцій $\mu_j(d_k)$, $j = \overline{1,n_{\tilde{D}}}$, що показують належність *k*-го діагнозу *j*-му діагностичному висновку. Тоді нечітка підмножина $\tilde{D}_j = \{ (d_k, \mu_j(d_k)) \mid d_k \in D \}$ відповідає нечіткому поняттю «*j*-й діагностичний висновок». Позначимо множину нечітких понять як $\tilde{D} = \{ \tilde{D}_j \mid j \in \{1,\ldots,n_{\tilde{D}}\} \}$,де $n_{\tilde{D}}$ – число діагностичних висновків.

Характеристична функція $\mu_j(d_k)$ приймає значення з лінійно упорядкованої множини належностей *M*. Якщо для діагностики застосовується детерміністична логіка, то $M = \{0,1\}$, і нечітка множина розглядається як чітка, причому $\tilde{D}_j = \{ d_k \mid \mu_j(d_k) = 1 \}$.

При застосуванні імовірнісної логіки кожний діагноз d_k задається умовною ймовірністю $p(d_k \mid \tilde{c}'_j)$, яка виступає в якості характеристичних функцій $\mu_j(d_k)$, тобто $\tilde{D}_j = \{ (d_k, \mu_j(d_k)) \mid d_k \in D, \mu_j(d_k) = p(d_k \mid \tilde{c}'_j) \}$.

Крім розглянутих вище множин у математичній моделі M_{da} задається множина відображень $f = \{ f_{X\tilde{S}}, f_{\tilde{S}\tilde{B}}, f_{\tilde{B}\tilde{C}'}, f_{\tilde{C}'\tilde{D}}, f_{\tilde{D}M_a} \}$ і множина критеріїв $Q = \{ Q_{X\tilde{S}}, Q_{\tilde{B}\tilde{C}'}, Q_{\tilde{C}'\tilde{D}}, Q_{\tilde{D}M_a} \}$ для цих відображень. Розглянемо докладніше компоненти цих множин.

Функція $f_{X\tilde{S}}$ у відповідність із підмножиною критеріїв $Q_{X\tilde{S}}$ і множиною діагностичних ознак *S* задає відображення множин БМС/З *X* на множину значень діагностичних ознак $\tilde{S}$, тобто $f_{X\tilde{S}} : X \to \tilde{S}$. Дане відображення реалізується процедурою морфологічного аналізу БМС/З [4,5].

Функція $f_{\tilde{S}\tilde{B}}$ задає бієктивне відображення множини значень діагностичних ознак $\tilde{S}$ на множину оцінок діапазонів діагностичних ознак $\tilde{B}$, тобто $f_{\tilde{S}\tilde{B}} : \tilde{S} \to \tilde{B}$.

Функція $f_{\tilde{B}\tilde{C}'}$ у відповідність із підмножиною критеріїв $Q_{\tilde{B}\tilde{C}'}$ задає відображення множини оцінок діапазонів діагностичних ознак $\tilde{B}$ на множину оцінок симптомокомплексів $\tilde{C}'$, тобто $f_{\tilde{B}\tilde{C}'} : \tilde{B} \to \tilde{C}'$.

Інформативність (корисність) ознак, що входять у симптомокомплекс, визначається експертними оцінками або розрахунками інформативності ознак по навчальній вибірці. Причому симптомокомплекси повинні мати не тільки інформативність, але й повнотою опису, компактністю й несуперечністю ознак, що входять у нього.

Функція $f_{\tilde{C}'\tilde{D}}$ у відповідність із підмножиною критеріїв $Q_{\tilde{C}'\tilde{D}}$ задає відображення множини оцінок симптомокомплексів $\tilde{C}'$ на множину діагностичних висновків, тобто $f_{\tilde{C}'\tilde{D}} : \tilde{C}' \to \tilde{D}$. При цьому виконується синтез ієрархічної структури діагностуємих станів в заданій предметній області медицини (бінарне дерево рішень D_S), процедурою ієрархічної кластеризації множини діагностуємих станів $\{D_i\}_{n_D}$ за критерієм мінімуму помилки кластеризації в просторі ознак *S*.

Результатом кластеризації є бінарне дерево рішень D_S, коренем якого є повна множина діагнозів $\{D_i\}_{n_D}$ в заданій предметній області, в гілках розташовуються кластери діагнозів, а листками є окремі діагнози. В процесі діагностики *i*-го пацієнта при відомому векторі ознак $\tilde{S}_i$ відбувається рух по дереву рішень, в кожній *j*-й вершині якого виконується диференційна діагностика станів D_q та D_l, шляхом обчислення вирішального правила (ВП) і прийняття рішення на користь D_q або D_l.

Для реалізації імовірнісного ВП в роботі реалізується метод синтезу уточнюючого діагнозу [2], який є модифікацією методу послідовного аналізу (методу Вальда). На кажному *k*-му етапі ВП ($k = \overline{1,n_k}$), при диференційній диагностиці між двома станами D_q і D_l, аналізується чергова ознака $\tilde{s}_k$ і вираховується відношення правдоподібності

$$\Theta = \prod_i P(\tilde{s}_k / D_q) / P(\tilde{s}_k / D_l),$$

яке порівняється з порогами $\Theta > A$, $\Theta < B$.

При виконанні однієї з умов приймається рішення про D_q або D_l відповідно й виконується перехід на більш низький рівень ієрархії D_S з метою уточнення діагнозу. При невиконанні обох нерівностей додається наступна $i+1$ ознака й процедура повторюється.

Ризики неправильного прийняття рішення на етапі діагностики: α – помилка першого роду і β – помилка

другого роду, визначаються розташуванням еліпсоїдів розсіювання об'єктів навчальної вибірки в просторі ознак *S* без урахування їх впливу на етап вибору необхідних ЛД і їх подальшої реалізації.

Функція $f_{\tilde{D}M_a}$ у відповідність із критерієм $Q_{\tilde{D}M_a}$ задає відображення множини діагностичних висновків $\tilde{D}$ на множину лікарських дій M_a, тобто $f_{\tilde{D}M_a} : \tilde{D} \to M_a$.

ЛД представляються моделлю $M_a = \langle Ta, Pa, SI \rangle$, де M_a – множина ЛД; $Ta = \{ta_i \mid i \in \{1,2,\ldots,n_{ta}\}\}$ – множина терапевтичних дій (ТД), $Pa = \{pa_i \mid i \in \{1,2,\ldots,n_{pa}\}\}$ – множина фармакологічних дії (ФД), $SI = \{si_i \mid i \in \{1,2,\ldots,n_{si}\}\}$ – множина видів хірургічного втручання.

Призначення ЛД при відомому діагностичному висновку $\tilde{D}_i$ складається з визначення їх типу (*Ta*, *Pa*, *SI*, або їх комбінацій) і переліку конкретних дій. Вибір типу ЛД є задачею багатокритеріального вибору альтернатив, для вирішення якої використовується метод аналізу ієрархій (МАІ). Для кожного з допустимих для даного діагностичного висновку $\tilde{D}_i$ типу ЛД формується підмножина необхідних ЛД $ta_{D_i} \in Ta$, $pa_{D_i} \in Pa$, $si_{D_i} \in SI$, після чого визначається їх реалізація з урахуванням індивідуальних особливостей пацієнта, протипоказань до окремих ЛД і багатокритеріального вибору аналогів.

Для комплексної оцінки етапів ДЛД і мінімізації ризику ЛП шукається залежність між помилкою діагностики (D_q замість D_l) і її наслідків при реалізації ЛД. Так як для переважної кількості патологій в різних областях медицини ЛД реалізуються медикаментозним шляхом, розглянемо такий варіант реалізації ЛД, при якому кожному діагностичному висновку $\tilde{D}_i$ відповідає множина необхідних ФД pa_{D_i}, на підставі якого формується комплекс лікарських препаратів (КЛП) Y_i, який забезпечує реалізацію pa_{D_i}, з урахуванням непереносимості *i*-го пацієнта до окремих препаратів і багатокритеріального порівняння препаратів-аналогів. Так як $Y_q \to pa_{Dq}$, а $Y_l \to pa_{Dl}$, то ризик ЛП визначається розбіжністю компонентів множин pa_{D_q} та pa_{D_l}, і для його мінімізації виконується перехід від традиційного простору ознак *S* в простір ФД *Pa*, компонентами якого є бінарні змінні (0 – відсутня ФД, 1 – присутня), а кожен стан $\tilde{D}_i$ представляється *i*-ю вершиною гіперкуба. В якості міри близькості в просторі *Pa* вибрана зважена відстань Хеммінга:

$$r_{ij} = \sum_{k=1}^{g} w_{ij} \left| pa_{ki} - pa_{kj} \right|,$$

де $pa_{ki}, pa_{kj} \in \{0,1\}$ – *k*-та ФД *i*-го та *j*-го діагнозів відповідно; *g* - розмірність простору *Pa*; w_{ij} – коефіцієнт, який забезпечує збільшення відстані, в разі наявності конфліктуючих ФД.

Застосування ієрархічної кластеризації за критерієм мінімуму сумарної зв'язку (мінімальний розріз *R*) в просторі *Pa* забезпечує мінімум ризику прийняття рішення на етапі формування КЛП при синтезі дерева рішень D_S на етапі діагностики. Крім того, в роботі пропонується метод корекції порогів *A* і *B* в ВП, враховуючи помилки, які виникають на етапі призначення КЛП. Отримано залежності між α, β і нормованим значенням мінімального розрізу $\overline{R_i}$: $\alpha_n = 0{,}5(1 - \overline{R_{ql}})$, $\beta_n = 0{,}5(1 - \overline{R_{lq}})$. Визначені таким чином α і β задають пороги $A = (1-\beta)/\alpha$, $A = \beta/(1-\alpha)$ в ВП, що забезпечує врахування ризиків ЛП при призначенні КЛП в діагностичному ВП.

Розглянута технологія комплексної оцінки ДЛД з метою мінімізації ризиків ЛП адаптується до інших видів ЛД, при цьому виконується перехід з простору ознак *S* у простір відповідних ЛД (*Ta* або *SI*).

Виконана програмна реалізація системи і її тестування на реальних медичних даних з використанням навчальної вибірки з 400 пацієнтів. В якості основної платформи обґрунтований вибір Java. В архітектурі системи виділено три основних модуля: модуль взаємодії з користувачем, базу даних, що включає в себе базу знань і модуль побудови знань.

III. Висновки

Виконано формалізацію етапів проведення ДЛД. Розроблено математичну модель ДЛД, яка враховує комплексну оцінку її складових, що дозволяє мінімізувати ризики лікарських помилок, підвищити достовірність і обґрунтованість рішень. Архітектура програмного забезпечення системи дозволяє легко адаптуватися до різних предметних областей медицини.

Література References

[1] Об инфраструктуре информационной поддержки клинической медицины / В.А. Лищук, А.В. Гаврилов, Г.В. Шевченко и др. // Медицинская техника. – М.: 2003. – № 4. – С. 36-42.

[2] Поворознюк А.И. Системы поддержки принятия решений в медицинской диагностике. Синтез структурированных моделей и решающих правил / А.И. Поворознюк – Saarbrücken Germany: LAP LAMBERT Academic Publishing GmbH & Co. KG, 2011. – 314 с.

[3] Компендиум 2015 – лекарственные препараты /Под ред. В.Н. Коваленко, А.П. Викторова. [Електронный ресурс] http://www.compendium.com.ua.

[4] Пытьев Ю.П. Косые проекторы и относительные формы в морфологии / Ю.П. Пытьев // Журнал вычислительной математики и математической физики. – 2013. – № 53 (12). – С. 2100-2122

[5] Филатова А.Е. Морфологическая фильтрация полутоновых изображений на основе локальных статистик / А.Е. Филатова // Автомат. технологии и производства. – 2016. – №3(13). – С. 33-39.

The Approach to Drawing up a List of Professional Competences in the Field of Software Engineering for Harmonization of Educational Standards

Igor Turkin, Vitaliy Narozhnyy, Yuliya Vykhodets
Chair of software engineering
National Aerospace University "Kharkiv Aviation Institute"
Kharkiv, Ukraine
nikrutrogi@mail.ru

Підхід до Формування Переліку Професійних Компетенцій в Галузі Інженерії Програмного Забезпечення для Гармонізації Освітніх Стандартів

Ігор Туркін, Віталій Нарожний, Юлія Виходець
Кафедра інженерії програмного забезпечення
Національний аерокосмічний університет ім. М.Є. Жуковського «ХАІ»
Харків, Україна
nikrutrogi@mail.ru

Анотація—Статтю присвячено підходу, що розроблений для обгрунтування вибору цільових професійних компетенцій для навчання за спеціальністю «Програмна інженерія». Перелік цільових професійних компетенцій повинний відповідати вимогам працедавців, підтримувати позитивний досвід національних стандартів вищої освіти та галузевих професійних стандартів, відповідати рамці e-Competence Європейського Союзу та стандарту SWECOM, що оснований на компетенціях. В статті розглянуто підхід до розробки та гармонізації переліку професійних компетенцій із застосуванням формальної математики та теорії графів. Верифікацію рішення проведено за допомогою експертного порівняння обраного переліку компетенцій для спеціальності «Програмна інженерія» зі стандарту e-CF (що базується на профілях) з компетенціями стандарту SWECOM.

Abstract—The article is devoted to approach designed to justify selection of target professional competencies for training in the specialty "Software Engineering". The list of targeted professional competence must meet the requirements of employers; maintain a positive experience of national higher education and professional industry standards; conform to the European Union e-Competence Framework and competence-based standard SWECOM. The paper presents an approach to the development and harmonization of the list of professional competence with application of formal mathematics and using graph theory. Verification of the solution was made by application of expert comparison of selected profile-based standard e-CF list of competencies for specialty "Software Engineering" with competencies of the SWECOM standard.

Ключові слова—програмна інженерія, дескріптор, компетенція, e-CF, SWECOM, SWEBOK, навчальний план, інформаційні та комунікаційні технології

Keywords—software engineering, descriptor, competence, e-CF, SWECOM, SWEBOK, curriculum, information and communication technologies

I. INTRODUCTION

Development of IT education and industry at whole in Ukraine has two pressing problems. First, Ukraine has an outflow of ICT specialists. This can significantly hit the country's potential in the industry. This issue has been considered in the article [1] by Mr. I.V.Sergienko, Director of the Institute of Cybernetics named by V.M. Glushkov, National Academy of Sciences of Ukraine. The second problem is the outflow of graduates of secondary schools to foreign universities [2]. Although many universities of Ukraine provide education of much higher quality.

Thus, the main task of higher education in Ukraine is to prepare highly qualified specialists in the field of ICT

(information and computer technology) able to compete internationally. This is achievable by providing educational services as close to European standards. This approach will increase the attractiveness of receiving ICT education for graduates of secondary schools in Ukraine and give opportunity to attract foreign students.

II. Analysis of the latest publications

General Standard for ICT competences of the European Union is an e-Competence Framework (e-CF). European standard e-CF EN 16234-1 [3] was adopted in July 2016. It provides links to 40 ICT competencies demanded by employers [4 - 7]. There are five levels of competencies (skills, knowledge and competences), applicable across Europe and overall ICT sector.

The overall structure of e-CF framework is four-dimensional. It is based on areas of competencies (descriptor 1) and competence (descriptor 2), and not on official duties. The competence-based approach is more flexible and personalized. Descriptor 3 is to describe the level to which qualifications may respond. Number of levels that may relate to a particular competency may vary depending on the content and variety of the job activities. Descriptor 4 is a brief description of examples of knowledge and skills, but it is not exhaustive. Definitions of competence, knowledge, skills and attitudes have been worded in terms of organizations, not individuals. However, individual competencies can be included in descriptor 3, which ensures comparability and understanding of the connection between the concept of competence on the part of the organization and the individual. Determination of the descriptor 4 is consistent with the general definition of the knowledge, skills and competences in EQF (European Qualifications Framework, a meta-system intended to ensure the transparency and comparability in the recognition of qualifications, diplomas, certificates of education, to promote academic and labor mobility of citizens). There we see a generalized model, which describes information systems (IS) business - process, which includes the following phases of the IS life cycle: planning, implementation, commissioning, adjustment and control. Stages of adaptation and management permeate all stages of the IS life cycle. This model has been applied for heuristic correlation of business processes with competences. To determine the appropriate level of competence the e-CF uses the following indicators EQF: «context complexity", "independence", and «behavior". This made possible to agree the e-CF educational levels with EQF.

Standard e-CF is common to the whole ICT sector in the European Union, but there are international models of knowledge, skills and competences specifically for professionals in the field of Software Engineering: Software Engineering Body of Knowledge (SWEBOK) and Software Engineering Competency Model (SWECOM). These documents are based on assumptions, that software development is a key skill for students in the chairs/departments that teach and graduate professionals in the field of software engineering (SE). However, this skill is only the foundation for the entire field of software engineering because the term "software engineering» is interpreted as the use of technology for the design, development, implementation, testing and maintenance of software in a systematic method [7].

Released SWEBOK, as a set of knowledge [8] is the international standard ISO / IEC TR 19759: 2005 [9]. In late 2013, SWEBOK V3 was approved for publication and published officially [10]. SWEBOK was created through cooperation between several professional bodies and industry and supported by IEEE Computer Society. According to SWEBOK V3 specialist in software engineering should have 15 major disciplines. SWECOM, unlike SWEBOK, describes competences of software developers. The knowledge and skills within the field of training and employment is an additional description of the required competencies. The model SWECOM is designed for managers, staff members and developers of training programs [11]. SWEBOK and SWECOM both describe the requirements for an ICT professional who has practical experience as SWECOM agreed with SWEBOK. The fundamental difference in drawing up these requirements is the following: SWEBOK is knowledge based but SWECOM is competence based. The second difference is that the SWECOM is subordinate to the model of IT competencies, which is implemented by the United States Department of Labor to determine the knowledge, skills and abilities required for employees to work successfully in the field of information technology.

The following guidelines for teaching SE at universities are the basis for the practice of education in software engineering (Curriculum Guidelines for Undergraduate [12, 13] and Graduate [14] Degree Programs in Software Engineering). They are also having been applied to major existing standards and recommendations. In-depth elaboration of these recommendations vary by level of education. In 2002, a specially created working group started developing the first version of recommendations for Bachelor of Software Engineering (level of higher education that meets the sixth level of qualification National Qualifications Framework Ukraine). The work finished in 2004. In 2014, the second version of the recommendations was prepared. At the master's level, we have only one version of the recommendations dated 2009. A guidance for PhD level does not exist. Recommendations for Bachelor of Software Engineering is mature, logical and consistent guidance to practical educational decisions and actions. At the same time, to the authors' opinion, guidelines for teaching masters is only one and not quite successful attempt to create a standard.

III. Unsolved part of the general problem. The article Purpose

Thus, there are four international standards and recommendations, such as e-CF, SWEBOK, SWECOM and Curriculum. They define the leading vector and restrictions in problems of formulation of requirements for training SE specialists. Documents containing these standards and

recommendations are of high quality due to attracting a large number of leading experts from science, industry and education. The documents are consistent to each other in matters of principle and compatible with previous versions. The dynamic development of the ICT sector causes changes in standards and guidelines, so there may be differences in not significant issues (this is concerning latest versions of these standards and their compliance with each other).

Unsolved problem is the adaptation of educational standards and curriculum of software engineering at universities in Ukraine to constantly renewed international standards, guidelines e-CF, SWEBOK, SWECOM and Curriculum. This will be the basis for the approval requirements of employers to the qualifications and competencies. This problem is not specific only for Ukraine but for the most post-Soviet countries. This fact allows us to bring their expertise available in the standards [15], monographs [16] and guidelines [17].

The aim of this paper is to present the approach to drawing up a list of targeted competencies for training specialists in software engineering that meet the three-tier system of higher education in Ukraine, taking into account international standards.

IV. RESEARCH METHODOLOGY. FORMAL STATEMENT OF THE PROBLEM

Scientific and practical significance of the paper is that the approach to the selection and justification of targeted professional competence in the field of "Software Engineering" allows meeting the requirements of employers, using a positive experience of national higher education (expressed in professional industry standards) and EU compliant e- Competence Framework. The task of choosing and justification is solved mathematically using graph theory.

Input assumptions and limitations:

target competencies for training specialists in software engineering should comply as fully e-CF, and be as close to current industry standards for higher education in Ukraine;

target competencies for training specialists in software engineering must provide training for profiles "developer" and "test specialist" because these profiles fully meet the qualifications of Degree "Specialist in the development and testing of software" by National classification of professions Ukraine . Since the profile "Digital media specialist" is almost identical for the necessary competences to profiles "developer" and "specialist test", we assume that these three profiles will form the core for the specialty "Software engineering" (European ICT Professional Profiles, p. 5, ftp: // ftp.cen.eu/CEN/Sectors/List/ICT/CWA-s/CWA%2016458.pdf);

The core profiles are not enough to determine the competencies targeted training for the Software engineering profession under the three-tier higher education system of Ukraine, as according to e-CF and EQF these profiles do not require the level of competence PhD.

To form national educational standards for software engineering it is necessary to go all the way (fig. 1), which is uncertain for two principal reasons:

1. We may define the endpoint path partially because there are no generally accepted guidelines for teaching software engineering for the Masters and PhD.

2. Starting point is also not certain, since there is a contradiction between the defined competences: general, employer-oriented competences of the European model (e-CF), and SWECOM defined for concrete specialty Software engineering, but subordinate to competence model of the US Department of Labor. In addition, e-CF is a profile-based approach as the standard profile description, i.e. job assignment descriptions that include the required qualification, tasks, skills and competence. SWECOM approach is limited by definition of competencies, which are followed by further description required skills and knowledge.

The formal statement of the problem. The initial data is the interrelation of the "profile" - "competence" set in the e-CF in tabular form. In the mathematical representation, this table defines the mapping of the 23 profiles $P = \{P_1, P_2, ..., P_{23}\}$ in the set of 40 competences $C = \{C_1, C_2, ..., C_{40}\}$.

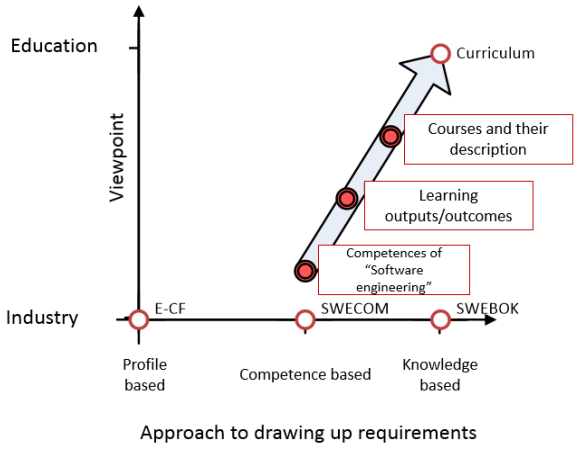


Fig. 1. The roadmap for elaborating educational standarts in the field of Software engineering

$$\phi : P \to C \tag{1}$$

According to this table for e-CF display, each profile corresponds to three to five competencies:

$$\forall P_i \in P, 3 \le \left\| \left\langle P_i, C_j \right\rangle \right|, C_j \in C \left\| \le 5 \right. . \tag{2}$$

We need to define this subset of target profiles $P_{aim} \subset P$, for which the power of the set of target competencies $\|C_{aim}\|, C_{aim} \subset C$ will meet the guidelines for developing higher education standards approved by the Scientific and

Methodological Council of the Ministry of Education and Science of Ukraine (Minutes of 29.03.2016 number 3). "The estimated number of special competences usually does not exceed 10-20 competences taking into account the level of education." An additional limitation is the requirement for mandatory including profiles "developer",» test specialist" and " digital media specialist" to the target subset of profiles P_{aim} .

Given the constraints, the task of identifying the target profile subset P_{aim} is optimization:

$$P_{aim}^{*} = \arg\max\left(\left\|P_{aim}\right\|\right), \tag{3}$$

with constraints:

$$\left\{\begin{array}{l}\text{“Developer”}\\ \text{“Test specialist”}\\ \text{“Digital media specialist”}\end{array}\right\} \subseteq P_{aim} \quad , \tag{4}$$

$$\left\|\phi\left(P_{aim}\right)\right\| \le 20 . \tag{5}$$

V. Description of the Proposed Methodological Approach

Step 1. Interpretation of the problem in terms of graph theory. In the above formulation the problem (problem formulation) (3-5), it is solvable and the subset of target profiles exists. That means that a solution can be found for a limited calculation time, as this formulation provides a consistent analysis only of 220 variants by the method of exhaustive search. As a result there will be received not one solution, but a set of subsets of targeted profiles, which will satisfy both the objective function (3) and limitation (4, 5). In other words, the formal solution is not enough, we need to make meaningful analysis and justify one set of target profiles.

We take into consideration the weighted undirected graph profiles:

$$G = \left\langle P, V \right\rangle, \tag{6}$$

when each edge of the graph $v_{i,j} \in V$ has a corresponding value (weight of edge) $w_{i,j}$, which is calculated as the number of shared competencies of i-th and j-th profiles.

Step 2. Selection and justification of the method of clustering profiles. By this, we mean the process of ordering structures into relatively homogeneous groups. A common feature of these groups is a structural proximity between profiles within the group (strong link inside) and weak links among groups. One of the most common metrics to calculate proximity, with the aim of clustering in graph theory is modularity. Modularity is calculated as the proportion of edges that are within the specified groups, excluding edges that would be used within specified groups in random distribution.

For the proposed weighted graph (6), we calculate modularity taking into account the weight of edges:

$$Q = \sum_{i}\left(e_{ii} - a_i\right), \tag{7}$$

where e_{ii} – relative sum of weights of all edges connecting profiles within the i-th cluster, calculated as the ratio of the sum of the weights of edges in the cluster to the sum weights of all edges vertices i-th cluster; $a_i = \sum_{j,\, j\neq i}\left(e_{ij}\right)$ - relative sum of weights of all edges connecting cluster c_i with other clusters.

Modularity metric realizes calculation of the degree of "similarity" between any two clusters A and B according to the Jaccard similarity coefficient:

$$k = \frac{\left\|A \cap B\right\|}{\left\|A \cup B\right\|} = \frac{\left\|A \cap B\right\|}{\left\|A\right\| + \left\|B\right\| - \left\|A \cap B\right\|} . \tag{8}$$

Modularity value by definition lies in the range [-1, 1]. Modularity is positive if the sum of the weights of edges within groups exceeds than their sum a random distribution of vertices by groups. In other words, modularity reflects the concentration of edges' weights within modules compared to the random distribution of the weights between all competences regardless of their competencies into clusters. It is known from the research of social networks communities that if the modularity value exceeds 0.3, it is a sign of existing community in the network. Estimated value of modularity for the graph of profiles proximity by e-CF is 0.423, from which we can conclude about the validity of the hypothesis of the existence of groups of profiles with close requirements to the competences.

Step 3. Clustering of the graph into relatively homogeneous groups with the help of Gephi [18]. It has multi interface and a rich set of instruments for networks visualization. Gephi written in the programming language Java and OpenGL project based on the modular NetBeans platform. This program is one of the most popular and allows importing data directly from a database MySQL. It supports extensions for different algorithms and clustering, gives statistics of proximity, development over time and in 3-dimensional space, allows you to update data via API (Application Program Interface) mode flow. System architecture and open source involves the addition of new features through the development of plug-ins. To perform clustering a graph into relatively homogeneous groups Gephi uses the method of medoids when only one of the available vertices may act as the center of the cluster. Therefore, with using this system we may build a graph of profiles proximity according to the e-Competence Framework (e-CF) version 3.0 (edge thickness is proportional to the number of common competencies for top-profiles). Experimental visualization has provided a clear argument that among e-CF profiles it is necessary to distinguish just 3 profiles groups which are

closely related by the list of required competence (unlike to 6 families profiles correlated with business processes for e-CF). One of these three groups is fully consistent to the formulated optimization problem (3-5) to determine the target profiles cluster, which contains profiles "Developer", "Test specialist " and "Digital Media Specialist".

Step 4. Determine the profiles (one or more) to add to the target profiles cluster to allow delivering educational programs of PhD level in Software engineering. To the authors' opinion, it is enough to add a profile "information systems architect» since this profile is closely linked with common competencies with other target profiles in Software engineering, and also has its own demands to the level of training.

Step 5. The final step is to check possibilities for matching e-CF competencies with the SWECOM competences. We have described above the existing ideological differences in the approaches to the formation of these standards, so the only way of comparison is the expert evaluation. The results of such testing now confirm the possibility to formulate a list of targeted professional competences in Software engineering in terms of SWECOM definitions. This facilitates the further development of educational standards in the specialty Software engineering complying with the SWEBOK standards and Curriculum recommendations.

VI. CONCLUSIONS

To our mind, there are two principal directions for developing presented results: "in depth" and "in breadth". For the «in depth" development it is necessary to draw up a list of the program learning outcomes. The methodological basis may be taken from the project "The use of learning outcomes in engineering education» (Engineering observatory on Competence based Curricula for job Enhancement) [19] in which the model was developed applying the concept "learning outcomes" of Bologna process in higher engineering education. The development «in breadth" means differentiation by target profiles according to e-CF specialties which belong to the field of knowledge "Information Technology":

Computer Science and Information Technology;

Computer Engineering;

System analysis;

Cyber security;

Information systems and technologies.

Implementation of development "in breadth" will allow the following:

- to determine the "point of intersection" of different specialties (common needs in training),
- to articulate differences in requirements for training specialists in different specialties in the knowledge domain «Information Technology",

This direction supports the interests of different stakeholders - universities (provides the basis for optimization of educational programs), students (improves understanding and justification of choice of their future profession), employers (interlink learning outcomes for different specialties to professional requirements).

REFERENCES

[1] І. Сергієнко Держава повинна зупинити відтік ІТ-шників [Online]. Available: http://ukurier.gov.ua/uk/articles/derzhava-povinna-zupiniti-vidtik-it-shnikiv/

[2] Отток научного потенциала из Украины приобрел масштабы стихийного бедствия [Online]. Available: http://news.finance.ua/ru/news/-/373147/ottok-nauchnogo-potentsiala-iz-ukrainy-priobrel-masshtaby-stihijnogo-bedstviya-ekspert

[3] EN 16234-1. e-Competence Framework (e-CF) - A common European Framework for ICT Professionals in all industry sectors - Part 1: Framework. [Online]. Available: https://standards.cen.eu/dyn/www/f?p=204:110:0::::FSP_PROJECT,FSP_ORG_ID:41798,1218399&cs=17B0E0F8CABCDBDDB8066A46FA937510B

[4] A common European framework for ICT [Online]. Available: http://www.ecompetences.eu/

[5] European e-Competence Framework [Online]. Available: http://www.ecompetences.eu/e-cf-overview/

[6] ICT profile European e-Competence Framework [Online]. Available: http://wp1087322.server-he.de/

[7] Программная инженерия [Online]. Available: https://en.wikipedia.org/ wiki/Software_engineering

[8] Software Engineering Body of Knowledge [Online]. Available: https://en.wikipedia.org/wiki/Software_Engineering_Body_of_Knowledge

[9] ISO/IEC TR 19759:2005 [Online]. Available:http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=33897

[10] Board of Governors Approves New Version of Guide to the Software Engineering Body of Knowledge [Online]. Available: https://www.computer.org/web/pressroom/board-of-governors-approves-new-version-of-guide-to-the-software-engineering-body-of-knowledge

[11] Software Engineering Competency Model (SWECOM) [Online]. Available: https://www.computer.org/web/peb/swecom

[12] SE 2004 - Curriculum Guidelines for Undergraduate Degree Programs in Software Engineering. – 2004 [Online]. Available: http://sites.computer.org/ccse.

[13] SE 2014 - Curriculum Guidelines for Undergraduate Degree Programs in Software Engineering [Online]. Available: http://www.acm.org/binaries/content/assets/education/se2014.pdf.

[14] GSwE2009 - Graduate Software Engineering 2009. Association for Computing Machinery and Computer Society of IEEE [Online]. Available:http://www.acm.org/binaries/content/assets/education/gsew2009.pdf

[15] Федеральный государственный образовательный стандарт высшего образования. Уровень высшего образования – бакалавриат. Направление подготовки 09.03.04 Программная инженерия [Online]. Available: http://www.osu.ru/docs/fgos/vo/bak_09.03.04.pdf

[16] Проектирование секторальных рамок квалификаций в области Информатика: Учебно-методическое пособие / [Н. С. Вольпян, В. В. Тихомиров, А. В. Разгулин та ін.]. – М.: Издательский отдел факультета ВМиК МГУ им. М.В. Ломоносова, 2015. – 218 с.

[17] Розроблення освітніх програм. Методичні рекомендації / В. М.Захарченко, В. І. Луговий, Ю. М. Рашкевич, Ж. В. Таланова / За ред. В.Г. Кременя. – К.: ДП «НВЦ «Пріоритети», 2014. – 120 с.

[18] The Open Graph Viz Platform [Online]. Available: https://gephi.org/.

Методи Розпізнавання Облич: Короткий Огляд

І. Голуб'як
кафедра інформаційних технологій
ДВНЗ «Прикарпатський національний університет
імені Василя Стефаника»
Івано-Франківськ, Україна
i.holubiak@gmail.com

An overview of Methods for Face Recognition

I. Holubiak
Department of Informational Technologies
Vasyl Stefanyk Precarpathian
National University
Ivano-Frankivsk, Ukraine
iholubiak@gmail.com

***Анотація*—Проведено аналіз переваг та недоліків методів розпізнавання облич. Запропоновано алгоритм отримання зображення для уникнення обману системи.**

***Abstract*—The advantages and disadvantages of face recognition methods are analysed. The author proposes an algorithm of obtaining image for avoid cheating the system.**

Ключові слова—розпізнавання облич; характерні ознаки;алгоритм;методи розпізнавання

Keywords—face recognition;characteristic features; algorithm;recognition methods.

I. ВСТУП

В даний час широко використовується ідентифікація особи за допомогою біометричних ознак. Суттєвого поширення набули системи, які використовують для розпізнавання обличчя людини, пристрої як уже давно не є новими на ринку техніки, а саме за допомогою відеокамер. Дана технологія використовується, як у великих корпораціях з певною обмеженістю доступу так і прости прикладом може бути система безпеки аеропорту, в якій уже розпізнавання здійснюється при обробці масиву зображень облич отриманого з відеопотоку. Навіть такі пристрої, як ноутбук чи смартфон, які присутні у повсякденному житті дозволять скористатися можливістю даної технології.

Розпізнавання облич є актуальним та використовується в багатьох сферах життя.

Системи не потребую дороговартісного обладнання, для певної якості роботи достатньо застосувати веб камеру. Звичайно самої веб камери не достатньо, потрібна програма, яка оброблятиме отримані зображення та "вирішуватиме" згідно заданого алгоритму та методу роботи чи потрібно надати доступ. Алгоритм роботи таких систем часто повторюються, а от методи зазвичай суттєво відрізняються. На даний час відомо та використовується велика кількість методів та їх модифікацій. Поширеними є методи та їх похідні такі як метод Віоли-Джонса, еластичних графів, метод головний компонент, методи принцип яких базується на геометричному методі розпізнавання, методи засновані на нейронних мережах такі як метод прихованої Маркової моделі, метод згорткової нейронної мережі, метод локальних бінарних шаблонів та ін.

Кожен із методів має свої переваги та недоліки, які проявляються при тих чи інших ситуаціях, що призводить до непередбачуваних випадків "пропустити чужого чи відмова у доступі для свого". Фактори, що впливають на якість роботи методів зазвичай одні і ті ж, та все ж таки кожен із методів має свою стійкість до певних із них. Такими факторами являється рівень та кут освітлення, відстань від камери, стан міміки чи здоров'я людини, настрій, кут нахилу обличчя, кут самого обличчя відносно камери, вікові зміни наявність бороди чи вус та ін.

II. МЕТОДИ РОЗПІЗНАВАННЯ ОБЛИЧ ПЕРЕВАГИ ТА НЕДОЛІКИ

Геометричний метод розпізнавання обличчя є одним із перших серед використовуваних методів розпізнавання обличчя [1, с.117]. В методах цього виду розпізнавання полягає у виділенні набору ключових точок (або областей) особи і наступному формуванні набору ознак. Серед ключових точок можуть бути куточки очей, губ, кінчик носа, центр ока тощо. Даний метод задає високі вимоги до зйомки зображень і потребує надійного алгоритму знаходження ключових точок для загального випадку. Приклад побудови геометричних ліній на обличчі людини зображено на рис. 1.

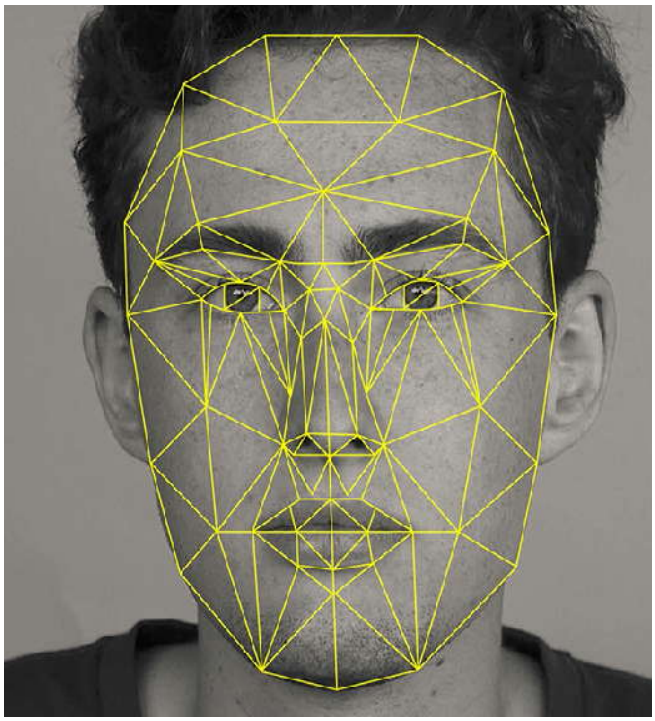
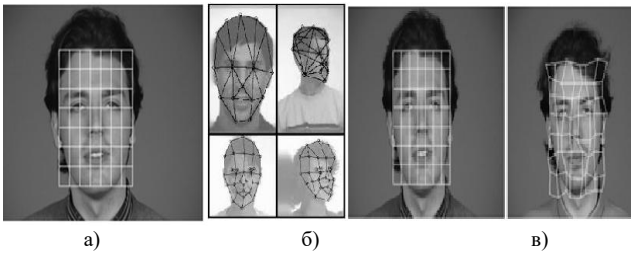


Рис. 1. Приклад побудови геометричних ліній на обличчі

До переваг методу можна віднести використання недорогого обладнання; при відповідному обладнанні є можливість розпізнавання зображень отриманих на значних відстанях. Недоліки наступні: низька статистична достовірність, високі вимоги до освітлення, обов'язкове фронтальне зображення особи, з невеликими відхиленнями. Не враховують можливі зміни міміки обличчя.

Метод гнучкого порівняння на графах, суть якого зводиться до порівнянні графів, що описують зображення обличчя особи. В окремих публікаціях вказується 95-97% ефективність розпізнавання навіть при наявності різних емоційних виразів і зміні ракурсу при формуванні зображення особи до 15 градусів. Проте для порівняння вхідного зображення особи з 87 еталонними витрачається приблизно 25 секунд при роботі на паралельній ЕОМ [2]. Приклад структури графа для розпізнавання осіб: а) регулярна решітка б) граф на основі антропометричних точок обличчя в) деформації графа у вигляді регулярної решітки, наведено на рис. 2.



а) б) в)

Рис. 2. Приклад структури графа для розпізнавання осіб: а) регулярна решітка б) граф на основі антропометричних точок обличчя в) деформації графа у вигляді регулярної решітки.

Іншим недоліком такого підходу є низька технологічність при запам'ятовуванні нових еталонів, що загалом призводить до нелінійної залежності часу роботи від розміру бази даних осіб. Основною перевагою є низька чутливість до рівня освітленості обличчя та до зміни кута обличчя, але сам по собі цей підхід має нижчі показники за достовірністю розпізнавання [3], ніж методи, побудовані із використанням нейромереж.

Метод головних компонент (МГК) зводить процес розпізнавання чи класифікації до побудови для вхідного зображення певної кількості головних компонент зображень [4, с. 88]. Приклад перших десяти власних векторів (власних осіб), отриманих на навчаємому наборі осіб наведено на рис 3. Приклад побудови (синтезу) людського обличчя за допомогою комбінації власних осіб та головних компонент наведено на рис. 4.

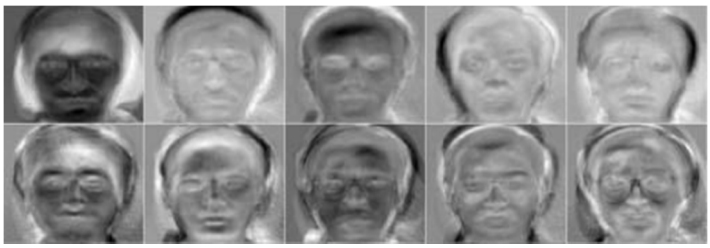


Рис. 3. Приклад перших десяти власних векторів (власних осіб), отриманих на навчаємому наборі осіб.

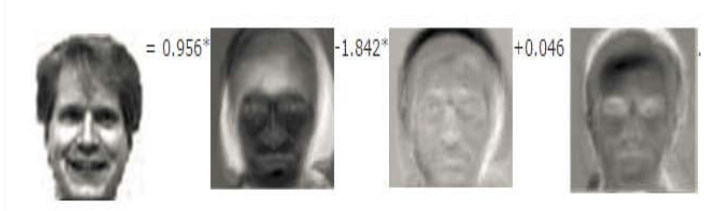


Рис. 4. Приклад побудови (синтезу) людського обличчя за допомогою комбінації власних осіб та головних компонент

МГК добре зарекомендував себе в практичних додатках. Однак, у тих випадках, коли на зображенні обличчя присутні значні зміни в освітленості або виразі обличчя, ефективність методу значно зменшується

Метод Віоли-Джонса, запропоновано в [5] та дозволяє виявляти об'єкти на зображеннях в реальному часі. Метод добре працює при спостереженні об'єкта під невеликим кутом, приблизно до 30°. Точність розпізнавання з використанням даного методу частково досягає понад 90%, що є хорошим результатом. Однак при куті відхилення понад 30° ймовірність розпізнавання різко падає. Враховуючи дану особливість унеможливлюється детектування особи під довільним кутом [6, 3].

Використання нейронних мереж. Одні з найкращих результатів в області розпізнавання осіб досягається за допомогою використання згорткових нейронних мереж (ЗНМ), які є логічним розвитком таких архітектур як когнітрон і неокогнітрон. Успіх обумовлений можливістю обліку двовимірної топології зображення, на відміну від багатошарового перцептрона. На исунку 5 зображено приклад роботи нейромережі.

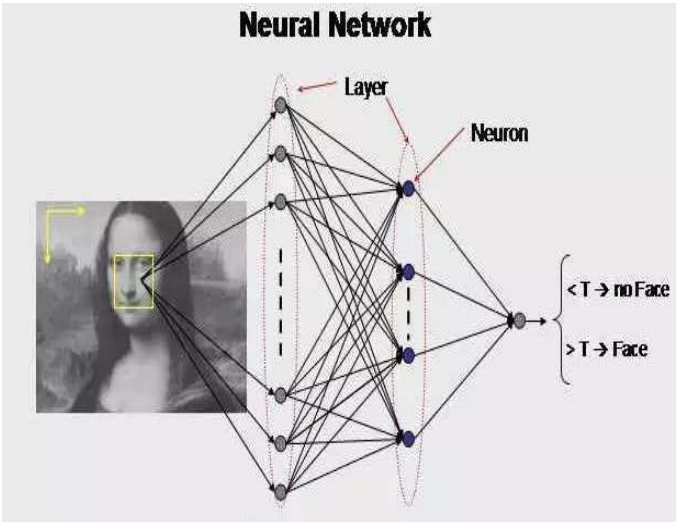


Рис. 5. Приклад роботи нейромережі.

Завдяки цим нововведенням ЗНМ забезпечує часткову стійкість до змін масштабу, зсувам, поворотам, зміні ракурсу та іншим спотворень [7-10]. Тестування ЗНМ на базі даних ORL [11], що містить зображення осіб з невеликими змінами освітлення, масштабу, просторових поворотів, положення і різними емоціями, показало 96% точність розпізнавання. До недоліків методів, які побудовані на основі нейронних мереж можна віднести додавання нового еталонного особи в базу даних, що вимагає повного перенавчання мережі на всьому наявному наборі, а це досить тривала процедура, яка залежно від розміру вибірки вимагає годин роботи а то і декількох днів. Також їм властиві проблеми математичного характеру, пов'язані з навчанням: попадання в локальний екстремум, вибір оптимального кроку оптимізації, перенавчання, тощо.

Локальні бінарні шаблони (ЛБШ) вперше були запропоновані в 1996 році для аналізу текстури півтонових зображень рис.5.

Рис. 6. Локольні бінарні шаблони

Дослідження показали, що ЛБШ інваріантні до невеликих змін в умовах освітлення і невеликим поворотам зображення [12, 13]. Методи на основі ЛБШ добре працюють при використанні зображень облич із різною мімікою, різним освітленням, поворотами голови. Серед недоліків - необхідність якісної попередньої обробки зображень через високу чутливість до шуму, оскільки за його присутності зростає кількість помилкових бінарних кодів.

Приховані моделі Маркова. Прихована марківська модель — статистична модель, що імітує роботу процесу, схожого на марківський процес з невідомими параметрами. Згідно моделі ставиться задача знаходження невідомих параметрів на основі інших параметрів, за якими ведеться спостереження. Отримані параметри можуть бути використані в подальшому аналізі для розпізнавання облич. З точки зору розпізнавання – зображення це двомірний дискретний сигнал. Важливу роль в побудові моделі зображення грає вектор спостереження. Для того, щоб уникнути розбіжностей в описах, зазвичай використовують прямокутне вікно для розпізнавання. Щоб не втрачати області даних, прямокутні вікна мають перекривати одне інше. Значення для перекривання, як і області розпізнавання підбираються експериментально. Після зняття блоку виконують його перетворення в цифровий блок за одним з двох методів [13].

Карунена-Лоева (КЛП);

дискретне косинусне перетворення (ДКП).

В роботі А. Нефіан; М. Хаєс основна увага приділяється використанню ПММ де в якості ознак розпізнавання використовуються двохвимірні вектори, такий підхід значно зменшує обчислювальну складність методу [13].

Активні моделі зовнішнього вигляду до задачі розпізнавання обличчя, що означає новий підхід до інтерпретації зображень запропоновано у [14]. Також їх застосовують для розпізнавання артикуляції на обличчі [15]. Активні моделі зовнішнього вигляду - це статистичні моделі зображень, які шляхом різного роду деформацій можуть бути підігнані під реальне зображення. В моделях цього виду описуються два типи параметрів: параметри, пов'язані з формою (параметри форми), і параметри, пов'язані зі статистичною моделлю зображення або текстурою (параметри зовнішнього вигляду). Перед використанням модель повинна бути навчена на безлічі заздалегідь розмічених зображень. Кожна мітка має свій номер і визначає характерну точку, яку повинна буде знаходити модель під час адаптації до нового зображення [16]. Приклад подібної розмітки показаний на рис. 7.

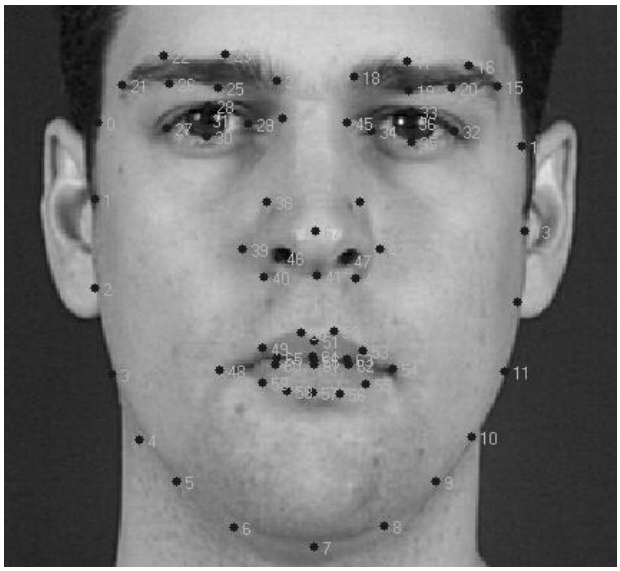


Рис. 7. Адаптаційна розмітка обличчя

III. Обхід системи розпізнавання

Все частіше інформаційні бази потребують більш кращого рівня захисту від несанкціонованого доступу. Говорячи про системи які працюють як за відбитком пальця чи за ідентифікацією по голосу, чи за допомогою електронного ключа, так і системи ідентифікації за допомогою обличчя можливо обійти використавши скопійований елемент який потрібний для ідентифікації, в даному випадку буде достатньо навіть фотографії. Для уникнення таких проблем, потрібно змінити метод отримання порівнюваного ображення. Можливо використати підказки в програмі при ідентифікації які дадуть можливість людині змінити міміку. Пропонується використати новий алгоритм отримання зображень при порівнянні з еталонним фронтальним а саме використання двох джерел отримання зображення обличчя під певним кутом, тобто два профільних зображення які не корелюють один з одним і водночас відтворюють певні елементи фронтального зображення особи , яке внесене попередньо, що дозволить збільшити точність розпізнавання.

IV. Висновки

Проведено аналіз існуючих методів розпізнавання, їх переваги та недоліки. На основі проведеного аналізу можна стверджувати про відсутність абсолютної переваги окремого напрямку. Більш чіткий вибір методу має бути продиктований умовами застосування, тобто властивостями біометричної системи, яка застосовується для отримання того чи іншого виду доступу та можливість альтернативного використання уже відомих характерних ознак обличчя.

Література References

[1] L.Juwei, N. P.Konstantinos, A. Venetsanopoulos, “Face recognition using kernel direct discriminant analysis algorithms”, IEEE Transactions On Neural Networks, vol.14, no. 1, pp.117–126, January 2003.

[2] M. Lades, J. Vorbruggen, J. Buhmann, “Distortion invariant object recognition in the dynamic link architecture”, IEEE Transactions on computers, 1993, vol. 42, no. 3, pp. 300 -310, March 1993.

[3] P. Viola, “Robust realtime face detection”, International Journal of Computer Vision, 2004, vol. 57, no. 2, pp. 137-154, 2004 .

[4] А. М. Лисенко, “Застосування біометричних систем для ідентифікації особи”, Вісник Київського нац. ун.-ту ім. Т.Шевченка, Юридичні науки, 2004, №60/62, с. 87-91 .

[5] Метод Виолы-Джонса (Viola-Jones) как основа для распознавания лиц, Электронный ресурс, Режим доступа:https://habrahabr.ru/post/133826/

[6] P. Viola, “Rapid object detection using a boosted cascade of simple features”, IEEE Conf. on Computer Vision and Pattern Recognition, V. 1, Kauai, Hawaii, USA., pp. 511–518, 2001.

[7] S.Lawrence, C.L. Giles., C. Tsoita, “Face Recognition: A Convolutional Neural Network Approach”, IEEE Transactions on Neural Networks, Special Issue on Neural Networks and Pattern Recognition, vol. 8, no 1, pp.98–113, 1997.

[8] Y.Taigman, M.Yang, M.Ranzato, “DeepFace: Closing the Gap to Human-Level Performance in Face Verification”, [Online]. Available at : https://www.cs.toronto.edu/~ranzato/publications/taigman_cvpr14.pdf

[9] Joo Er Meng, W.Chen, Wu Shiqian, “High-speed face recognition based on discrete cosine transform and RBF neural networks”, IEEE Transactions on Neural Networks, vol. 16, no. 3, pp. 679 – 691,2005.

[10] H.A. Rowley, S. Baluja, T.Kanade (1998), “Rotation invariant neural network-based face detection”, Computer Vision and Pattern Recognition, Proceedings. IEEE Computer Society Conference, [Online]. Available at : http://ieeexplore.ieee.org/abstract/document/698585/

[11] T. Mäenpää (2003), “The local binary pattern approach to texture analysis – extensions and applications”, Finland ,Oulu University Press.

[12] T.Ahonen, A. Hadid, M.Pietikainen, “Face Description with Local Binary Patterns: Application to Face Recognition”,IEEE Transactions on Pattern Analysis and Machine Intelligence,2006, vol. 28, no 12, pp. 2037 – 2041.

[13] Ю. Лифшиц, Методы распознавания лиц, Електронний ресурс, Режим доступу: http://yury.name/modern/08modernnote.pdf

[14] T. F.Cootes, G. J. Edwards, C. J. Taylor, “Active appearance models”, Computer Vision ECCV’98, vol.. 14, no. 7 of the series Lecture Notes in Computer Science, pp. 484-498, 2006.

[15] М. В.Давидов, Ю. В. Нікольський, С. М. Тиханський, “Алгоритм визначення форми губ під час артикуляції для української жестової мови”, Вісник Національного університету "Львівська політехніка", no. 673 : Інформаційні системи та мережі, сс. 267-273, 2010.

[16] T. F. Cootes, G. J. Edwards, C. J. Taylor, “Active appearance models”, IEEE Trans. on Pattern Recognition and Machine Intelligence, vol. 23, no.6 , pp. 681–685, 2001.

Формування Масиву Вхідних Даних для Мамографічних Досліджень

С. Палаш
Кафедра інтелектуальних систем прийняття рішень
Черкаський національний університет імені Богдана Хмельницького
Черкаси, Україна
s_palash@ukr.net

О. Харченко
Кафедра інтелектуальних систем прийняття рішень
Черкаський національний університет ім. Богдана Хмельницького
Черкаси, Україна
alexandrharchenko106@gmail.com

Formation of an Array of Input in the Mammographic Research

S. Palash
Department of intelligent systems of acceptance of decision
Cherkasy National UniversityBohdan Khmelnytsky
Cherkasy, Ukraine
s_palash@ukr.net

A.Kharchenko
Department of intelligent systems of acceptance of decision
Cherkasy National University Bohdan Khmelnytsky
Cherkasy, Ukraine
alexandrharchenko106@gmail.com

***Анотація*—представлено порівняння методів формування масиву вхідних даних для дослідження мамографічних знімків. Експерементально підтверджено кращі результати розпізнавання змішаним методом формування масиву вхідних даних.**

***Abstract*—presented a comparison of methods for forming an arrays of input data for the mammography image analysis. Experimentally proved that the better method for mammography recognition and analysis is a mixed method.**

Ключові слова—масив вхідних даних, мамограмирозпізнавнання.

Keywords—arrays of input data, mammography image, recognition.

I. Вступ

Аналіз мамографічних зображень актуальна задача, яка потребує постійного вдосконалення. Для дослідження мамограм, формуються масиви вхідних даних. Встановивши вид матриці вихідних даних, приступають до формування інформаційного масиву. На цьому етапі визначають перелік змінних і об'єктів спостереження. Відбір ознак (змінних)і об'єктів спостереження є дуже важливим етапом роботи і виконується в суворій відповідності з метою дослідження. Якість інформації тут значною мірою залежить від знань про об'єкт і ознак, які мають найбільш істотну інформацію про цей об'єкт. Важливими критеріями відбору при формуванні масиву вхідних даних (МВД) є інформативність показників і точність відображення їх чисельних характеристик.

II. Постановка задач досліджень

Метою даної роботи є дослідження процесу формування МВД в умовах різної інформативності показників. Передбачається досягнути мету за рахунок використання різних алгоритмів збору даних. Мета буде досягнута тоді, коли інформативність МВД буде достатньою для синтезу корисних моделей. Моделі повинні розв'язати задачу аналізу знімків.

III. Формування масиву вхідних даних

Масив вхідних даних – це матриця чисельних характеристик елементів зображення, отриманих в процесі його декомпозиції.

Процес формування МВД має принципово важливе значення для успішного виконання завдань машинного навчання. Часто завдання машинного навчання виконуються саме завдяки правильного формування масиву даних для навчання моделей – навчальної множини. Помилки у формуванні навчаючої множини зазвичай виявляються критичними і здатні звести нанівець ефективність самих алгоритмів навчання. Серед фахівців по машинному навчання загально визнаним вважається, що наявність хороших навчальних даних набагато важливіше якості алгоритму навчання.

У зв'язку з активним розвитком систем багаторівневого перетворення інформації в останнє десятиліття питання формування множини навчальних даних приймають особливо важливе значення, оскільки в багатьох задачах система багаторівневого перетворення даних демонструює якість, істотне покращення ніж інші алгоритми машинного навчання, однак, щоб отримати подібний виграш як, необхідно використовувати навчальну множину дуже великого розміру (до декількох мільйонів зображень, при цьому навчання вимагає великого обсягу обчислювальних ресурсів і може займати кілька тижнів на багатопроцесорному кластері).

Використання методу групового урахування аргументів дає кращий результат в умовах обмеженої інформативності.

Представлення навчаючої вибірки, можна описати так:

Нехай маємо множину об'єктів , множина відповідей P, і існує цільова функція $^*: G \to P$, значення якої $= {}^* * (\ )$ відомі тільки для підмножини об'єктів $\{1, .., \} \subset$. Сукупність пар $= (\ ,\ )$ називається вибіркою. Задача навчання полягає в тому, щоб по навчаючій вибірці відновити залежність *, тобто побудувати вирішуючу функцію a: $G \to P$, яка б наближувала б функцію $^*(\ )$, причому не тільки на навчаючій вибірці, але ні на всій вибірці G [1]. Метод навчання – це відображення $:(\ \times\ ) \to$, яке у довільній конечній навчаючій вибірці $= (\ ,\ )$ ставить у відповідність вирішуючу функцію a: G $\to$ P. Метод навчання μ будує вирішуючу функцію a по навчаючій вибірці [1].

Правильне формування навчальної вибірки часто має вирішальне значення в задачах машинного навчання, що визнається більшістю фахівців в даній області. Найчастіше рішення задач машинного навчання зводиться до грамотного формування навчальної вибірки. Незважаючи на це, в сучасній літературі по машинному навчання питань формування навчальної вибірки майже не приділяється увага, теоретична база практично відсутня.[1].

Способи генерації навчальної множини описані нижче.

Програмна генерація - генерація навчальних даних по деякому алгоритму. Алгоритм генерації і його параметри, в даному випадку, визначають отримані на виході розподілення у просторі об'єктів. Доцільно варіювати якомога більше параметрів в процесі генерації. Однак метод може виявитись неефективним та згенерувати велику кількість не потрібних даних.

Якщо реально можлива лише мала частина об'єктів з простору параметрів генерації або найбільш типові об'єкти мають складну конфігурацію.

Спосіб покликаний подолати недоліки попереднього методу – семплірування. Алгоритм намагається згенерувати вибірку з даного розподілу, здається деякий апріорний розподіл в просторі об'єктів. Семплірування по Monte Carlo, Гиббсу, схема Метрополіс – Гастингса, застосовуються методи вибірки з відхиленням. Застосування даних методів використовується для того, щоб досліджувати найбільш осмислені частини простору об'єктів (зазвичай це лише мала частина потенційно можливих об'єктів, коли простий перебір об'єктів з перевіркою їх коректності може зайняти невиправдано багато часу).

Закономірна модифікація базового об'єкта. В даному випадку маємо набір базових об'єктів, навчальна множина формується шляхом багаторазової модифікації їх параметрів (як правило, зовнішніх змінних). Прикладом може бути вибірка кадрів з відеопослідовності, вибірка поз людини з mocap-а. Метод доцільно застосовувати в тих випадках, коли немає можливості аналітично задати розподіл можливих об'єктів в просторі об'єктів, або не підходить використання синтетичних даних. При використанні даного методу генерації навчальної множини зберігається велика кількість фонових закономірностей. Крім того, виникають проблеми, якщо отриману таким чином множину розбити в деякому відношенні на навчальну і тестову і використати отриману тестову множину для контролю за перенавчанням. Оскільки об'єкти результуючого мають сильну взаємозалежність, низький рівень помилки алгоритму на тестовій вибірці не гарантує відсутності «заучування» навчальних даних. При використанні метричних алгоритмів класифікації найефективніше в даному випадку - завжди відносити об'єкт до того самого класу, до якого належить найближчий навчальний об'єкт. Приклад помилки подібного роду при використанні алгоритму k-NN наведено в роботі [4] (при налаштуванні гіперпараметрів алгоритму методом ковзаючого контролю за багатьма даних, отриманого вищеописаним методом, завжди вибирався параметр k = 1). Тому при використанні даного методу формування множини даних не варто застосовувати метод ковзаючого контролю в чистому вигляді.

Ще один спосіб – вибірка з бази об'єктів. Формування множини зображень з довільної множини зображень обличь людини погано описуються наведеними вище моделями. Є фіксований набір об'єктів із заздалегідь визначеними параметрами і ми не можемо отримати об'єкт з будь-якими заданими внутрішніми параметрами. Якщо це зображення місцевості, розташування цих об'єктів підпорядковується строгим обмеженням. Об'єкти можна розбити на групи, причому об'єкти всередині групи будуть дуже схожі, а об'єкти з різних груп - відрізнятися (можна порівняти зображення осіб злочинців і зображення облич політиків). Об'єкти, географічно розташовані близько один до одного, зазвичай більш схожі, ніж об'єкти, що знаходяться далеко один від одного (наприклад, будинки в одному і в різних містах). Якщо це колекція фотографій, слід звертати увагу на переваги фотографа, тому, що люди набагато частіше фотографують красиві речі (наприклад, пам'ятки), ніж одноманітний ліс або пустелю, різні люди роблять знімки різних типів.

При генерації навчальної множини даними способом – важко гарантувати наявність усіх принципово важливих типів об'єктів в множини даних. В даному випадку може допомогти перехід до автоматичної генерації. Також слід

робити вибір об'єктів з якомога більш широкого набору груп об'єктів або місць зйомки (детектор особи, навчений на базі осіб злочинців, можливо, буде не дуже добре працювати в загальному випадку).

IV. Способи додавання даних в навчальну множину

Додавання даних є одним з найпростіших і ефективних способів поліпшити якість навчальної множини. При цьому, незавжди ефективно просте додавання даних довільного виду часто потрібно додати дані певного різновиду для підняття якості розпізнавання. Розглянемо деякі способи додавання даних.

Програмна генерація. У разі використання синтетичних навчальних даних зручніше всього згенерувати відсутні навчальні приклади. Однак не у всіх завданнях допустимо використання програмно згенеровані дані. У таких випадках доводиться застосовувати складніші методи додавання даних.

Data augmentation. Модифікація наявних зображень з метою розширити навчальну вибірку. Активно застосовується в системах багаторівневого перетворення даних [5], а також в умовах дефіциту розмічених даних. Застосовуються стиснення/розтягування, горизонтальне відображення, поворот, випадковий зсув в колірному просторі, випадкове або закономірна зміна деяких пікселів. Вважається, що додавання повністю випадкового шуму неефективно, слід додавати шум, обумовлений даними (тільки потенційно можливі в реальних даних спотворення). Істотний недолік даного методу - зберігається більшість фонових закономірностей.

Hard samples mining [6]. Класична проблема в задачах пошуку об'єктів на зображенні - потреба в підтримці достатнього числа hard negative samples (навчальних прикладів, що схожі на об'єкт інтересу, але такими не є) в навчальній множині. Складність виникає через те, що в природних умовах такі об'єкти зустрічаються рідко, тому застосовуються спеціальні методи для їх пошуку і додавання в навчальну множину (hard samples mining). Ключове припущення в даних методах - цікавлять нас об'єкти позитивних але схожі між собою. Зазвичай застосовуються data augmentation, адаптивний пошук, пошук по шаблонах, методи на основі машинного навчання. Цікавим є застосування методів тематичного моделювання для пошуку складних негативних прикладів.

Імітація додавання даних. При навчанні систем багаторівневого перетворення інформації можливе використання методу dropout [6]: випадкове обнулення активацій деяких моделей в мережі при подачі їй на вхід чергового тренувального зображення (зазвичай в кожному шарі випадково вибирається 20-50% моделей). Без застосування даної техніки система багаторівневого перетворення інформації може «заучувати» велику кількість фонових закономірностей через те, що складність моделі перевищує обсяг доступних даних. По суті, dropout - це імітація додавання даних в навчальну вибірку. Усередині алгоритму навчання ми імітуємо мінливість даних - на вхід більш глибоких рівнів мережі надходить випадковим чином змінена версія реального зображення (мається на увазі зображення, що не викликало б активацію обнулення моделей), хоча таких даних насправді немає в навчальній множині. Недолік даного методу - може бути імітований додавання таких даних, яких в принципі не може бути в реальності, через чого може страждати точність розпізнавання. Цікавим є створення модифікацій даного методу, що враховують природу даних.

Краудсорсинг. Оскільки для систем багаторівневого перетворення інформації потрібні величезні об'єми вручну розмічених навчальних даних, для формування навчальної множини активно використовуються сервіси краудсорсингу - користувачі сервісу за невелику плату створюють розмітку «сирих» даних (наприклад, вказують, які об'єкти є на даному зображенні і де вони розташовані). Найпопулярніший з таких сервісів - Amazon Mechanical Turk [7].

Проблема даного методу - велика кількість помилок в розмітці, так як користувачі не завжди роблять свою роботу сумлінно, а іноді просто помиляються, тому потрібні спеціальні методи контролю помилок в розмітці. Зазвичай одну і ту ж картинку дають розмітити декільком користувачам, а потім вибирають той варіант розмітки, який вибрало найбільше число користувачів. Також вводяться спеціальні метрики «сумлінності» користувача.

V. Результати роботи алгоритмів формування МВД

Для проведення аналізу формування МВД використовувалися такі методи:

Ручне формування МВД.

Програмна генерація МВД.

Data augmentation.

Hard samples mining.

Формування МВД шляхом ковзаючого вікна.

Змішані методи.

В таблиці 1 представлені порівняння результатів роботи даних методів та показники стійкості та адекватності моделей. Для тестування було відібрано зображення розмічене експертом. В результаті експерт виявив 10 ділянок з підозрою на ущільнення тканин, 2 ділянки з мікрокальцинатами та 2 ділянки з підозрілою поведінкою. Показники стійкості та адекватності визначалися в межах від 0 до 1, де 0 – абсолютно стійка модель, тобто та, що розпізнає образ в умовах зашумленості зображення. За даною таблицею легко побачити, що формування масивів вхідних даних змішаним методом дає найкращий результат.

ТАБЛИЦЯ I. ПОРІВНЯННЯ МЕТОДІВ

Назва методу	*Кількість розпізнаних зображень*	*Стійкість*	*Адекватність*
Ручне формування МВД	8	0,7	0,75
Програмна генерація МВД	8	0,72	0,75
Data augmentation	9	0,8	0,7
Hard samples mining	10	0,82	0,72
Формування МВД шляхом ковзаючого вікна	10	0,85	0,76
Змішані методи	13	0,9	0,9

Нижче наведено результати роботи програми у вигляді графічних зображень.

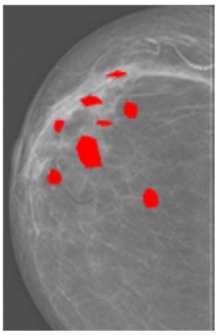


Рис. 1. Результати роботи методу Hard samples mining

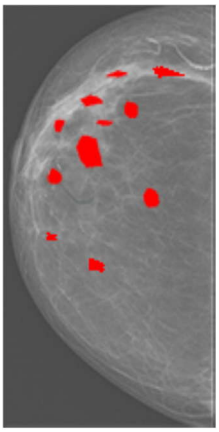


Рис. 2. Результати роботи методомко взаючого вікна

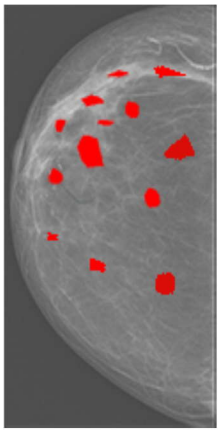


Рис. 3. Результати роботи змішаного методу

Змішаний метод формування масиву вхідних даних дав кращий результат. Оскільки було розпізнано 13 зображень із 14.

VI. ВИСНОВОК

Дана робота демонструє способи формування масиву вхідних даних різними методамби. Результати досліджень дозволяють зробити висновок, що формування масиву вхідних даних змішаними методами показують кращий результат. Даним методом було вірно розпізнано 93% ділянок.

ЛІТЕРАТУРА REFERENCE

[1] И.Л. Кафтанников, А.В. Парасич Проблемы формирования обучающей выборки в задачах машинного обучения Южно-Уральский государственный университет, г. Челябинск, Вестник ЮУрГУ. Серия «Компьютерные технологии, управление, радиоэлектроника». 2016. Т. 16, № 3. С. 15–24

[2] Mangalova E., Petrun'kina I. [Prediction Capacity of Wind Power Plants Based on NonParametric Algorithm, K Nearest to Neighbors]. Doklady vserossiyskoy nauchnoy konferentsii AIST'2013 [Reports of the All-Russian Scientific Conference AIST'2013]. Ekaterinburg, 2013, pp. 1–8. (in Russ.) J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73.

[3] Canavet O., Fleuret F. Efficient Sample Mining for Object Detection. Proceedings of the Asian Conference on Machine Learning (ACML), 2014, pp. 48–63.

[4] . Srivastava N. Hinton G.E., Krizhevsky A., Sutskever I., Salakhutdinov R.R. Dropout: A Simple Way to Prevent Neural Networks From Overfitting. The Journal of Machine Learning Research, 2014, vol. 15, no. 1, pp. 1929–1958

[5] Голуб С.В. Багаторівневе моделювання в технологіях моніторингу оточуючого середовища / С.В. Голуб. – Черкаси: Вид. від. ЧНУ імені Богдана Хмельницького, 2007. – С.220.

[6] Amazon Mechanical Turk. Avaliable at: https://www.mturk.com/mturk/welcome (accessed September 2015).

[7] R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.

Використання Фрактальних Характеристик Щільності Хроматину Ядра Клітини для Визначення Стану Інтерфази

О. Шумейко
Д.т.н., проф.
Дніпровський державний технічний університет
Кам'янське, Україна
Shumeiko_a@ukr.net

Г. Шевченко
К.т.н., доц.
Науковий центр Noosphere
Дніпро, Україна
she-genn@yandex.ua

С. Герасименко
Науковий центр Noosphere
Дніпро, Україна
gerasimenko-stan1slav@yandex.ru

Using the Fractal Characteristics of the Density of the Chromatin of the Cell Nucleus to Determine the Interphase State

O.Shumeyko
Prof.
Dniprovs`k State Technical University
Kamenske, Ukraine
Shumeiko_a@ukr.net

G.Shevchenko
PhD.
Noosphere Research Center
Dnipro, Ukraine
she-genn@yandex.ua

S.Gerasimenko
Noosphere Research Center
Dnipro, Ukraine
gerasimenko-stan1slav@yandex.ru

Анотація—Запропоновано метод використання характеристик зміни щільності розподілу хроматину в ядрі клітини, що дозволяє отримати ефективні алгоритми виділення на цифрових зображеннях клітин в стані інтерфази.

Abstract—Research fractal properties of the distribution of chromatin in the nucleus of human cells is important for early diagnosis of cancer. The study of the density distribution of chromatin in the cell nucleus is complicated by the fact that this value is different in different periods of the life of the cell. In studying the problems of the state of chromatin during interphase important life cycle of cells.

The results obtained using the characteristics of changes in the density distribution of chromatin in the cell nucleus, yield efficient algorithms for digital images highlight cells in interphase state.

Ключові слова—мітоз клітин; медичні зображення; фрактали.

Keywords—cell mitosis; medical images; fractals.

I. Вступ

Дослідження (в тому числі, з метою ранньої діагностики раку) фрактальних властивостей розподілу хроматину в ядрах клітин людини, останнім часом, набуло інтерес у фахівців [1,2,3]. Інтерес до даної тематики обумовлений, перш за все, можливістю автоматизації процесу діагностування, який зводиться до аналізу

цифрових зображень відповідних клітин людини [4,5]. Таким чином, одним з ключових елементів даного підходу є виділення області цифрового зображення, що містить ядро клітини, з метою подальшого аналізу виділеного фрагмента. Вивчення розподілу щільності хроматину в ядрі клітини ускладнене тим фактом, що ця величина є різною в різні періоди життя клітини (див., наприклад, [9]). В період інтерфази в клітині відбувається процес біосинтезу білка, подвоюються всі найважливіші структури клітки. Уздовж вихідної хромосоми з наявних у клітині хімічних складових, синтезується її точна копія, подвоюється молекула ДНК. Подвоєна хромосома складається з двох половинок- хроматид. Кожна з хроматид містить одну молекулу ДНК. Інтерфаза в клітинах рослин і тварин в середньому продовжується на протязі 10-20 годин. Потім настає процес розподілу клітини - мітоз. Мітоз є неперервним процесом, але для зручності його розбивають на чотири стадії в залежності від того, як виглядає в цей час хроматин ядра. У мітозі виділяють профазу, метафазу, анафазу і телофазу. У профазі відбувається вкорочення і потовщення хромосом внаслідок їх спіралізації. У метафазі завершується утворення веретена поділу, яке складається з мікротрубок двох типів: хромосомних, які зв'язуються з центромерами хромосом, і центросомних (полюсних), які тягнуться від полюса до полюса клітини. В анафазі хроматиди подвоєні, в інтерфазі хромосоми розходяться до полюсів клітини. У цей момент в клітині знаходяться два диплоїдних набору хромосом. У телофазі відбуваються процеси, зворотні тим, які спостерігаються в профазі: починається деспіралізація (розкручування) хромосом. Навколо хромосом у кожного полюса з мембранних структур цитоплазми формується ядерна оболонка, в ядрах виникають ядерця. Руйнується веретено ділення. На стадії телофази відбувається поділ цитоплазми (цитотомія) з утворенням двох клітин. В результаті мітозу з однієї клітини виникають дві дочірні з тим же набором хромосом, що і в материнській клітині.

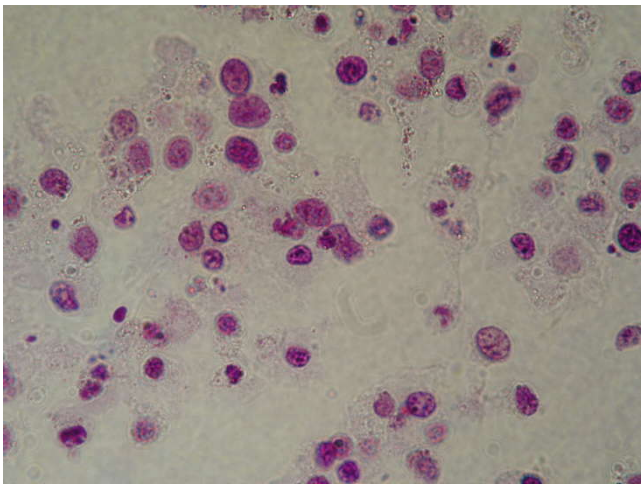


Рис. 1. Цифрове зображення соскобу епітеліцітов буккального епітелію, взяті з середньої глибини шипуватого прошарку, зафіксовані в суміші Нікіфорова і пофарбовані по Фельгену.(Архів відділу онкогематології ІЕПОР ім. Р.Є.Кавецького НАН України).

На кожному етапі мітозу розподіл щільності хроматину істотно змінюється. Якщо під час інтерфази хроматин рівномірно розподілений в ядрі клітини, то під час профази від збирається в «брили», які під час менафази і анафази розбиваються на частини і притискаються до протилежних сторін клітини.

У задачах вивчення стану хроматину нас цікавить тільки період інтерфази життєвого циклу клітини. Таким чином метод дослідження розподілу щільності хроматину ядер клітин повинен враховувати цю особливість. Вихідні дані представлені в вигляді повнокольорових зображень, отриманих в результаті цитологічного дослідження (рис.1).

II. ОГЛЯД ІСНУЮЧИХ МЕТОДІВ

Ефективним методом дослідження такого роду є фрактальний аналіз напівтонових зображенні. Перш за все це такі методи.

Метод BOX COUNTING. Цей метод застосовується до зображення будь-якої структури на площині і може бути пристосований до 3D-об'єктів [9,10], він пов'язаний з розмірністю подібності, і дозволяє визначити фрактальну розмірність не строго самоподібних об'єктів [11]. Для оцінки box-розмірності, евклідовий простір, що містить зображення об'єкта, поділяють сіткою з кроком розміру r і підраховуються непусті, зайняті досліджуваним об'єктом, квадрати N(r). Потім розмір r зменшують, і знову підраховують число непустих полів N(r). Нахил графіка в логарифмічному масштабі N (r) від 1 / r відповідає величині розмірності. Як видно з визначення, для цього методу попередньо потрібно провести сегментацію зображення, що істотно погіршує отриману фрактальну характеристику.

Метод 2D Variation [12] призначений для обробки бінарних (binary) і напівтонових (grayscale) зображень. В цьому випадку знаходяться мінімальний і максимальний рівні сірого в межах кожної квадратної комірки зі стороною r. Таким чином, визначаються двомірна мінімальна і максимальна функції для кожного кроку сітки.

Потім для всього зображення в цілому визначається різниця об’єму між мінімальним та максимальним значеннями. Тоді $V(r) = r^S const$, де V - об’єм різнички , r - крок сітки, S- фрактальна розмірність.

Подібна процедура повторюється для всіх пікселів зображення, досить віддалених від границі, потім r збільшується, і процедура повторюється. Нахил s графіка V (r) в логарифмічному масштабі використовується для визначення фрактальної розмірності: $D_{2D} = 3 - \frac{S}{2}$. Даний метод дозволяє оцінювати як розмірність кожного колірного каналу RGB окремо, так і розмірність "яскравості" зображення.

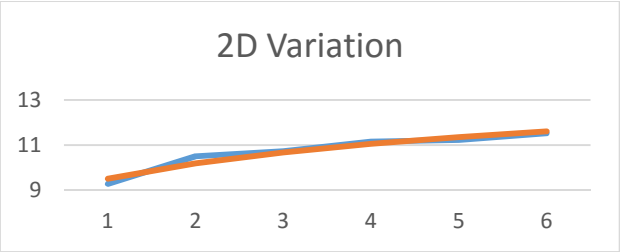


Рис. 2. Розрахунок фрактальної розмірності методом 2D Variation

Ще одним методом оцінювання фрактальної розмірності напівтонових зображень є броунівський розмірність [9], яка пропорційна експоненті Херста (Hurst`s exponent) D = 3-H. Параметр H можна обчислити статистично із закону дисперсії, використовуючи властивості броунівського руху: $H = -\frac{\log(\sigma_{\Delta R}(\Delta I))}{\log(\Delta R)}$, $\sigma_{\Delta R}(\Delta I)$- середньо-квадратичне відхилення різничок, тобто різничка освітлення в точках $R + \Delta R$ та R, $\Delta I = I(R + \Delta R) - I(R)$), R - координати точки.

Алгоритм обчислень виглядає наступним чином: для всіх точок зображення (досить віддалених від краю) обчислюються збільшення освітлення ΔI в ΔR- околиці, для отриманого масиву збільшень обчислюється середньоквадратичне відхилення $\sigma_{\Delta R}(\Delta I)$, потім змінюємо ΔR та повторюємо попередні кроки. Нахил графіку залежності $\sigma_{\Delta R}(\Delta I)$ від ΔR в логарифмічних координатах дає нам величину H.

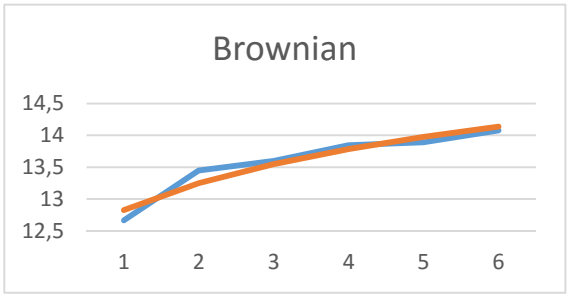


Рис. 3. Розрахунок фрактальної розмірності методом Brownian

III. ПОБУДОВА ХАРАКТЕРИСТИК ЩІЛЬНОСТІ РОЗПОДІЛУ ХРОМАТИНУ В ДАНОМУ НАПРЯМКУ

Складність використання даних характеристик для розв'язуваної задачі в тому, що отримані значення характеризують фрактальність зображення в цілому, а для вирішення поставленого завдання важливо визначити локальну нерівномірність розподілу кольору (щільності хроматину), або хоча б нерівномірність в якомусь напрямку, що характерно для стану мітозу клітини.

Нехай дано монохроматичне зображення ядер клітин $K_k = \{(x_i^k, y_i^k) | i = 0,1,...,n_k - 1\}(k = 1,2,...,N)$.

Використовуючи значення кольору пікселя в якості маси відповідної матеріальної точки, знайдемо центр мас

$$X_k = \frac{1}{n_k}\sum_{i=0}^{n_k-1} x_i^k, Y_k = \frac{1}{n_k}\sum_{i=0}^{n_k-1} y_i^k.$$

і через цю точку проведемо пучок прямих $y = tg\varphi(x - X_k)$, де $\varphi \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$.

Для кожного фіксованого φ, маси точок множини K_k, які попали на пряму $y = tg\varphi(x - X_k)$ представляють собою одновимірний масив $P(\varphi) = \{P_0, P_1, ..., P_{N(\varphi)}\}$.

Нехай $\Delta P(\varphi) = \frac{1}{N(\varphi)}\sqrt{\sum_{i=0}^{N_v-1} |P_i - P_{i+1}|^2}$. Якщо хроматин в ядрі клітини розташований рівномірно, то $\Delta P(\varphi)$ зміна щільності в будь-якому напрямку φ приблизно однакова, тобто, якщо $\Pi(\Delta P(\varphi))$ перестановка Харді $\Delta P(\varphi)$, то вона містить велику горизонтальну (або майже горизонтальну) ділянку. Наблизимо $\Pi(\Delta P(\varphi))$ ламаною $\ell(\varphi, \{\tau_i\}_{i=0}^3, t)$ з трьома ланками і граничними умовами (щоб прибрати викиди, як великі, так і маленькі)

$$\ell(\varphi, \{\tau_i\}_{i=0}^3, \tau_0) = 0, \ \ell(\varphi, \{\tau_i\}_{i=0}^3, \tau_3) = 1,$$

де $\tau_0 = 0, \tau_3 = N_v$ Так, щоб для

$$\left\| \Pi(\Delta P(\varphi)) - \ell(\varphi, \{\tau_i\}_{i=0}^3) \right\|_2 \to \min \,|\, \tau_1, \tau_2.$$

Для знаходження вирішення цього завдання нам знадобляться додаткові означення.

Нехай $\Delta_n : 0 = t_0 < t_1 < ... < t_{n-1} < t_n = T$ - довільне розбиття відрізку $[0,T]$ і $S_r(\Delta_n)$ - множина всіх сплайнів порядку r мінімального дефекту за розбиттям Δ_n, тобто множина функцій з неперервною $r-1$-ю похідною на $[0,T]$, які співпадають на кожному проміжку $(t_{i,n}, t_{i+1,n})(i = 0,1,..., n-1)$ з алгебраїчним поліномом ступеню не вище r.

Через $P(\Delta_n)$ позначимо оператор, який відображає C^ρ в $S_r(\Delta_n)$, $\rho \geq r - k$. При фіксованому r послідовність $\{P^*(\Delta_n^*)\}_{n=1}^\infty$ будемо називати асимптотично оптимальною сплайн-регресією для функції $x(t)$, якщо при $n \to \infty$

$$\int_a^b (x(t) - P^*(\Delta_n^*, t))^2 dt = \inf_{\Delta n} \inf \left\{ \int_a^b (x(t) - P(\Delta_n, t))^2 dt \,\middle|\, P(\Delta_n) \in S_r(\Delta_n) \right\} (I + o(I)). \quad (1)$$

Нехай на проміжку $[0,T]$ у вузлах розбиття Δ_n дані значення функції $x_i\ i=0,1,...,n$. Кубічний сплайн $s_3(x,\Delta_n)=s_3\left(\{x_i\}_{i=0}^n,\Delta_n\right)$ будемо називати інтерполяційним з граничними умовами

$$s_3^{(i)}(x,\Delta_n,0)=x(0),i=0,1\ \geq,\ s_3^{(i)}(x,\Delta_n,T)=x(T),i=0,1,$$

якщо $s_3(x,\Delta_n,t_i)=x_i,i=0,1,...,n.$

В роботі [13] доведено наступне твердження.

Теорема 1. Нехай $x\in L_2^3$ така, що x'' дорівнює нулю на обмеженому числі відрізків або точок з $[0,T]$ і $X(t)$ друга первісна функції $x(t)$ така, що $X'(0)=0$ і $X(0)=0$. Виберемо M_i з умов

$$\lambda_i M_{i-1}+2M_i+\mu_i M_{i+1}=\frac{6}{h_{i-1}+h_i}\left(\frac{X_{i+1}-X_i}{h_i}-\frac{X_i-X_{i-1}}{h_{i-1}}\right),\ i=1,2,...,n-1,$$

де $\mu_i=\frac{h_{i-1}}{h_{i-1}+h_i},\lambda_i=1-\mu_i.$

Тоді $\inf\left\{\left\|x-s_1(\Delta_n)\right\|_2\middle|s_1(\Delta_n)\in S_1(\Delta_n)\right\}=\left\|x-s_3''\left(\{M_i\}_{i=0}^n,\Delta_n\right)\right\|_2$.

Виберемо вузли t_i^* розбиття Δ_n^* з умови

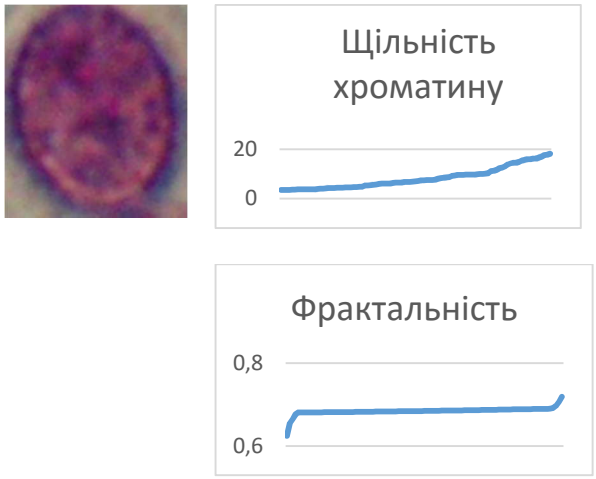


Рис. 4. Характеристики клітини в стані інтерфази

$$\int_0^{t_i^*}\left(\left|s_3''\left(\{M_i\}_{i=0}^n,t\right)\right|+\frac{1}{n^\gamma}\right)^\alpha dt=\frac{i}{n}A_n,$$
$$A_n=\int_0^T\left(\left|s_3''\left(\{M_i\}_{i=0}^n,t\right)\right|+\frac{1}{n^\gamma}\right)^\alpha dt,$$

тоді при $n\to\infty$

$$\inf_{\Delta_n}\inf\left\{\left\|x-s_1(\Delta_n)\right\|_2\middle|s_1(\Delta_n)\in S_1(\Delta_n)\right\}=$$
$$\left\|x-s_3''\left(\{M_i\}_{i=0}^n,\Delta_n^*\right)\right\|_2(1+o(1))=\frac{\theta}{n^2}\|x''\|_\alpha+o\left(\frac{1}{n^2}\right),$$

де $\theta=\frac{1}{2!}\left(\int_0^1 t^2(1-t)^2\,dt\right)^{1/2}$ і $\alpha=2/5.$

Використовуючи результати теореми 1, знайдемо розв'язок задачі (1). Рішенням будуть значення τ_1 та τ_2, які реалізують екстремум. Тоді величина

$$\delta=\frac{\ell\left(\varphi,\{\tau_i\}_{i=0}^3,\tau_2\right)-\ell\left(\varphi,\{\tau_i\}_{i=0}^3,\tau_1\right)}{\tau_2-\tau_1}$$

буде характеризувати рівномірність розподілу хроматину, чим ближче δ до нуля, тем рівномірніше (незалежно від напрямку) він розподілений, а величина $\theta=\ell\left(\varphi,\{\tau_i\}_{i=0}^3,\tau_1\right)$ характеризує величину щільності, тобто, чим більше θ, тим більше перепади кольору хроматину, а величина $\vartheta=\tau_2-\tau_1$ характеризує наявність або відсутність викидів, чим ближче ϑ до $N(\varphi)$ тим менше викидів даних.

Зауважимо, що велике значення δ характерно для стану менафази і анафази, мале значення θ характерно для клітин в стані профази, а маленьке значення ϑ може говорити про те, що ми спостерігаємо телофазу ядра.

Таким чином, використання отриманих характеристик дозволяє, з одного боку, відкинути ті клітини, які знаходяться в стані мітозу, а, з іншого, отримати характеристику щільності хроматину.

Крім того, для оцінки щільності, можна використовувати фрактальну розмірність в даному напрямку. Для цього використовуємо таку конструкцію.

Нехай $r<<N(\varphi)$ и $P_i=0$ при $i>N(\varphi)$, тоді для $i=0,1,...,[N(\varphi)/r]$

$m_i=\min\{P_{ir},P_{ir+1},...,P_{(i+1)r}\}$, $M_i=\max\{P_{ir},P_{ir+1},...,P_{(i+1)r}\}$ і $V_i(r)=(M_i-m_i)r.$

Далі, нехай $V(r)=\sum_{i=0}^{[N(\varphi)/r]}V_i(r)$. Подібна процедура повторюється для всіх пікселів зображення, досить

віддалених від кордону, потім r збільшується, і процедура повторюється. Нахил s графіка V (r) в логарифмічному масштабі використовується для визначення фрактальної розмірності.

Якщо хроматин розташований в ядрі рівномірно, то і фрактальна розмірність в будь-якому напрямку однакова.

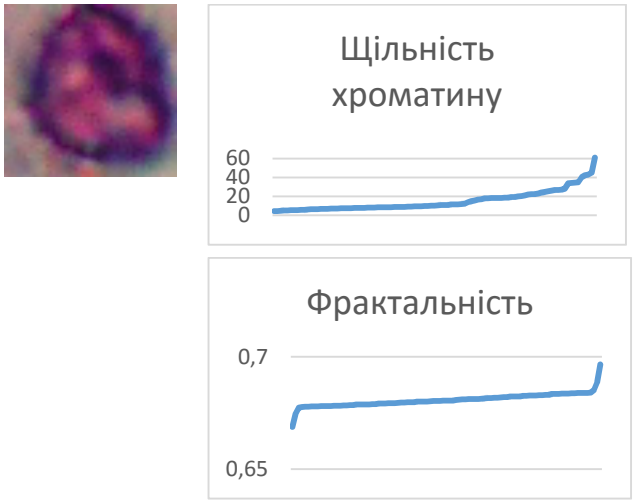


Рис. 5. Характеристики клітини в стані мітозу

Побудова характеристик визначення клітин в стані інтерфази.

У разі, якщо потрібно виділити клітини в стані інтерфази (що є найбільш часто розглядаються випадком), то завдання дещо спрощується.

Як видно з поведінки графіка щільності хроматину, якщо ці дані описати квадратичної регресією

$$\sum_{i=0}^{100}\left(\Pi(\Delta P(\varphi_i))-ai^2-bi-c\right)^2 \to \min_{a,b,c}$$

тоді коефіцієнти a,b,c будуть характеризувати зміну щільності хроматину.

Рис.4	0.0015835	0.2163586	1.5100646
Рис.5	0.0037519	0.1356285	4.4715336

Кращою характеристикою є старший коефіцієнт. Для тестової вибірки це значення лежить в межах від 0.000804 до 0.0061442.

Перестановка коефіцієнтів, що характеризують фрактальність, добре описується кубічною моделлю. Для прикладів на зображенні 5 і 6 коефіцієнти кубічної регресії будуть наступні

Рис.4	0.000000284	0.0000076194	0.0109206	-0.2296844
Рис.5	0.0000003088	0.0000083126	0.0117783	-0.2497191

Як і в попередньому випадку, найкращою характеристикою є старший коефіцієнт, для тестової вибірки його значення лежать в межах від 0.0000002736 до 0.0000003524.

Таким чином, визначення стану інтерфази клітин може спиратися на значення старших коефіцієнтів квадратичної регресії для перестановки Харді зміни щільності хроматину і кубічної регресії коефіцієнтів фрактальності.

IV. Висновки

Використання характеристик зміни щільності розподілу хроматину в ядрі клітини, дозволяє отримати ефективні алгоритми виділення на цифрових зображеннях клітин в стані інтерфази.

Автори статті висловлюють подяку науковим співробітникам відділу онкогематології Інституту експериментальної патології, онкології і радіобіології ім. Р.Є. Кавецького НАН України за наданий для аналізу і обробки фото- та інформаційний матеріал.

ЛІТЕРАТУРА REFERENCES

[1] Диагностическая онкогематология. / Д. Ф. Глузман [и др.] ; ред. Д. Ф. Глузман. - Киев: ДИА, 2011. - 256 с.

[2] Acharya, Rajendra, et al. "Analysis of cardiac health using fractal dimension and wavelet transformation." ITBM-RBM 26.2 (2005): 133-139.

[3] Smith, T. G., et al. "A fractal analysis of cell images." Journal of neuroscience methods 27.2 (1989): 173-180.

[4] Einstein, Andrew J., et al. "Fractal characterization of chromatin appearance for diagnosis in breast cytology." The Journal of pathology 185.4 (1998): 366-381.

[5] Klyushin, D. A., et al. "Screening of breast cancer using Peano curve." Журнал обчислювальної та прикладної математики 4 (2012): 117-121.

[6] Клюшин, Д. А., et al. "Новий метод скринінгу раку молочної залози на підставі фрактального аналізу інтерфазних ядер букального епітелію." Журнал обчислювальної та прикладної математики 1 (2011): 104.

[7] Клюшин Д.А. и др. Фрактальный анализ опухоль-ассоциированных изменений интерфазных ядер буккального эпителия. Журнал обчислювальної та прикладної математики 4 (2013): 114.

[8] https://uk.wikipedia.org/wiki/Поділ_клітини

[9] Лигун А.А., Шумейко А.А. Асимптотические методы восстановления кривых.-Киев, Институт математики НАН Украины, 1997.-358 с.

[10] Кроновер Р.М. Фракталы и хаос в динамических системах. - М. ПОСТМАРКТ, 2000. - 350 с.

[11] Falconer K. Fractal Geometry. Mathematical foundations and applications. - New York: John Wiley & Sons, 1990.

[12] Kraft R. Estimating the Fractals and Dimensions // Journal of agricultural, nutritional and environmental sciences - 1996. - №10 //

[13] Шумейко А.А., Шумейко Е.А. О построении асимптотически оптимальной кусочно-линейной регрессии.// Інформатика та математичні методи в моделюванні. Том 1, №2,2011, с.99-106.

Моделювальний Комплекс для Симуляції Дефіциту Енергії в Клітинах Людини

Т. Аксьонова
Інститут програмних систем НАН України
Київ, Україна
akstanya@ukr.net

Modeling Complex Which Simulates the Energy Deficit in Human Cells

T. Aksonova
Institute of Software Systems
Kyiv, Ukraine
akstanya@ukr.net

***Анотація*—Розглянуто проблему дефіциту енергії в клітинах людини. Розроблений програмний моделювальний комплекс реалізує основні моделі механізмів клітини енергетичної теорії адаптації і дозволяє здійснювати тестування теоретичних концепцій інтегративної фізіології людини.**

***Abstract*—The problem of the energy deficit in human cells is considered. Software modeling complex implements basic models of energetically reversible adaptation theory and allows researcher to test the theoretical concepts of integrative physiology.**

Ключові слова—математичні моделі; мітохондрії; глюкоза; інтегративна фізіологія.

Keywords—mathematical models; mitochondria; glucose; integrative physiology.

I. Вступ

Дефіцит енергії (ДЕ) пригнічує усі форми біологічної роботи в клітині і може провокувати розвиток таких нетривіальних захворювань як слабкість серця, гіпертрофія, артеріальна гіпертонія. Дефіцит енергії традиційно повязують з дефектами генів і патологією мітохондрій, однак, нещодавні дослідження вказують на існування інших причин послаблення функцій мітохондрій та появи тривалого дефіциту енергії у великої кількості клітин людини.

При розпаді фосфорного зв’язку молекул аденомзинмонофосфату (АМФ), аденозиндифосфату (АДФ) виділяється енергія, але основним джерелом енергії для виконання всіх різновидів біологічної роботи є молекула аденозинтрифосфату (АТФ) [1]. В статиці швидкість синтезу молекул АТФ приблизно дорівнює швидкості використання [2]. Для вивчення зв'язку між систезом АТФ та клітинною дисфункцією були запропоновані математичні моделі. Існують моделі, які описують лише окремі ланки комплексного механізму, що забезпечує синтез АТФ [3-8]. Більш загальним підходом є енергетична теорія адаптації [9], що відображає інтегральне функціонування комплексу клітинних та багатоклітинних механізмів, які забезпечують ефективну боротьбу клітин проти хронічної нестачі молекул АТФ. Для для встановлення конкретних закономірностей взаємодії усіх механізмів, які забезпечують актуальну потребу клітин в енергії, необхідне проведення значної кількісті експериментів з моделями клітинніх механізмів, які дозволили б експериментально визначити оптимальні значення параметрів моделей. Потреба в адекватному дослідницькому інструменті, який базується на енергетичній теорії адаптації, обумовлює актуальність розробленого моделювального комплексу.

II. Моделі фізіологічних механізмів

Розроблений моделювальний комплекс – це спеціалізована інформаційна технологія, яка грунтується на математичних моделях часткових механізмів [10-12], моделі взаємодії батареї механізмів клітини і батареї додаткових механізмів організму [10-13], а також на програмній технології [14]. Часткові моделі описувють такі фізіологічні механізми:

- Модель батареї механізмів клітини – запезпечує приблизну рівність середньої швидкості споживання АТФ і аеробної швидкості синтезу АТФ у віртуальній аеробній клітині [10];
- Модель гомеостазу глюкози крові [11]. Описує зв’язки між прийомом вуглеводів (глюкози), динамікою інсулину, запасанням надлишків глюкози в глікоген

печінки, а також зворотнє перетворення глікогена в глюкозу при фізичних навантаженнях;

- Модель функціонування гемодинамічно замкненої серцево-судинної системи людини [12];
- Модель збільшення провідності кровоносних судин шляхом локальної вазодилятації і ангіогенеза [13];
- Модель збільшення вентиляції легенів [13];
- Модель збільшення кількості гемоглобіну в крові завдяки мобілізації крові з органів-депо і активації еритровоеза [13];

Сукупність описаних моделей можна розглядати як комплексну математичну модель, що описує фізіологію реагування організму людини на ДЕ в клітинах.

III. Структура моделювального комплексу

Моделювальний комплекс включає в себе наступні модулі:

- Модуль відображення загальної системної інформації;
- Модуль створення та параметризації моделей;
- Модуль кофігурування ієрархії моделей;
- Модуль проведення експериментів з моделями;
- Модуль моніторингу стану виконання обчислень;
- Модуль виведення та збереження результатів.

З точки зору користувача, що проводить імітаційні експерименти в даній системі, найбільш часто використовуваним модулем є підсистема виведення та збереження результатів. Дана підсистема містить в собі такі основні компоненти:

1. Компонент збереження результатів;
2. Компонент графічного відображення результатів моделювання;
3. Компонент налаштування графіка результатів;
4. Компонент табличного представлення результатів моделювання.

Слабко-зв'язана архітектура і багатопоточність дозволяє моделювальному комплексу виконувати розрухунки кожної моделі незалежно, щоб максимально ефективно задіяти процесор. Базовим елементом організації потоків обчислень є конфігурація моделей. Кожна модель у конфігурації може мати свій крок розрахунку, тому загальним часом розрахунку є максимальний час, за який кожна модель конфігурації вже розрахувала всі свої змінні. Синхронізація моделей реалізована методом бар'єрної синхронізації (рис. 1) – визначаються моменти часу, коли моделі будуть обмінюватись даними. Якщо модель дісталась до бар'єру швидше, вона чекає на решту, і тільки коли всі досягнуть бар'єра, відбувається синхронізація. Через те, що у моделей можуть застосовуватись різні значення кроків обчислень (в залежності від структури моделі та необхідної точності), бар'єри можуть припадати на середину кроку. В такому випадку модель зупиниться лише після перетину бар'єра. Це потрібно враховувати при розрахунку точності обчислень. В якості інтервалу синхронізації бажано брати проміжок часу, не менший ніж найбільший крок обчислення моделей, так як в противному випадку точність обчислень не збільшиться, а ефективність розрахунку знизиться через зупинки для синхронізації та пересилки даних. Взагалі інтервал синхронізації треба вибирати, виходячи із компромісу бажаної точності обчислень та часом цих обчислень. Тому для зменшення часу очікування моделей в якості інтервалу синхронізації обирається проміжок часу, не менший ніж найбільший крок обчислення моделей. Моделювальний комплекс підтримує точки розширення, реалізувавши які користувач може створювати власні моделі, кастомізувати і перевизначати поведінку існуючих моделей.

Рис. 1. Синхронізація моделей

При проведенні короткотривалого моделювання результати всіх обчислень зберігаються у вигляді неперервного масиву даних в оперативіній пам'яті. Даний спосіб є найшвидшим з можливих, оскільки навіть зміна розміру масиву в процесі обчислень вже призводить до втрат ресурсів ЕОМ. При цьому для відображення результатів досліджень та при збережнні їх для подальшого аналізу результатів експериментів вже немає потреби зберігати всі дані. Щоб відобразити лише ті точки графіка, що в стані розрізнити людське око, достаньо виводити 1 точку на секунду. Таким чином для зменшення технічних вимог до моделювального комплексу було реалізовано алгоритм, який для кожної моделі в залежності від кроку розрахунку даної моделі h зберігає в пам'яті 1 точку на 1/h розрахованих значень (при h<1).

При проведенні довготривалого моделювання навіть із врахуванням оптимізації в процесі відображення та збереження результатів постає проблема недостатньої кількості операційної пам'яті. В таких випадках необхідним є вкорочення масивів та вивантаження розрахованих даних на жорткий диск. Масиви вкорочуються, так як для розрахунку наступного кроку використовується лише декілька попередніх кроків (за умови вокристання методу Ейлера – один попередній крок), тому тримати в оперативній пам'яті повністю всі масиви результатів за умов її дефіциту не доцільно. Достатньо визначити деякий початковий розмір та тримати розмір масивів постійним, забезпечуючи доступ до попередніх кроків. При цьому треба враховувати, що процес запису даних на жорсткий диск вимагає значних

витрат часу. Було експериментально з'ясовано, що оптимальним є збережнення даних кожної секунди із скороченням масів відповідно до кроку розрахунків кожної моделі, та розроблено спеціальний алгоритм, що зберігає в пам'яті лише останній «фрейм» даних, а всі попередні розрахунки записуються на жорсткий диск. При такому підході важливою є організація доступу до отриманих результатів. Якщо навіть лише виконувати перевірку того, де знаходяться дані, це вже сильно позначиться на продуктивності. Тому використовується підхід на кшталт реалізації віртуальної пам'яті операційною системою. Але на відміну від операційної ситеми, ми оперуємо не сторінками, а частинами масивів. Модель отримує запити, у яких вказується ім'я потрібної змінної моделі та момент часу (або номер кроку), аналізує ці запити та повертає потрібні значення. Процес аналізу починається з визначення місця розташування потрібної інформації – в оперативній пам'яті чи на жорсткому диску. Якщо в оперативній пам'яті, то одразу повертається результат, в противному випадку частина масиву завантажується в оперативну пам'ять. Даний підхід дозволяє знайти баланс між оптимізацією використання оперативної пам'яті та продуктивністю моделювального комплексу.

У випадках, коли ми маємо справу із моделями, що описують принципово різні за часом перебігу процеси, є потреба проводити довготривалі експерименти із можливістю дослідження швидкопливних процесів у певні проміжки часу. Для забезпечення цією можливості моделювальний комплекс дозволяє повністю зберегти стан кожної з моделей та продовжити розрахунки із заданого стану не перериваючи хід занального експерименту. Таким чином, маючи попередні знання щодо характеру описаних процесів, можна оптимізувати як використання оперативної пам'яті, так і швидкість проведення обчислень.

IV. Приклад симуляції реакцій системи на дефіцит енергії

Проілюструємо реакції системи на появу показників клітинного неблагополуччя в крові. У вихідному режимі моделі були налаштовані на стабільне функціонування віртуальної клітини при постійному рівні сумарних навантажень на клітину. При початковому значенні ефективного кисню в клітині на рівні 1.53 в крові відсутні показники клітинного неблагополуччя, середній артеріальний тиск становить 95 мм рт.ст., локальний кровотік дорівнює 8.1 мл/сек, локальний судинний опір - 11.47 мм рт.ст.*сек/мл. Далі імітувалося зростання навантаження на віртуальну клітину: навантаження лінійно зростала до досягнення максимуму і залишалося на цьому значенні до кінця симуляції.

Розглянемо процеси, які відбуватимуться за умови, що задіяно допоміжний механізм додаткового збільшення еритроцитів крові за рахунок збільшення вентиляції легенів (швидкий процес) та еритропоезу (повільний процес) для автономної клітини із внутрішньоклітинною регуляцією без навантажень та з навантаженнями показано на рис. 2. Вертикальна пунктирна лінія проведена через різні масштаби часу.

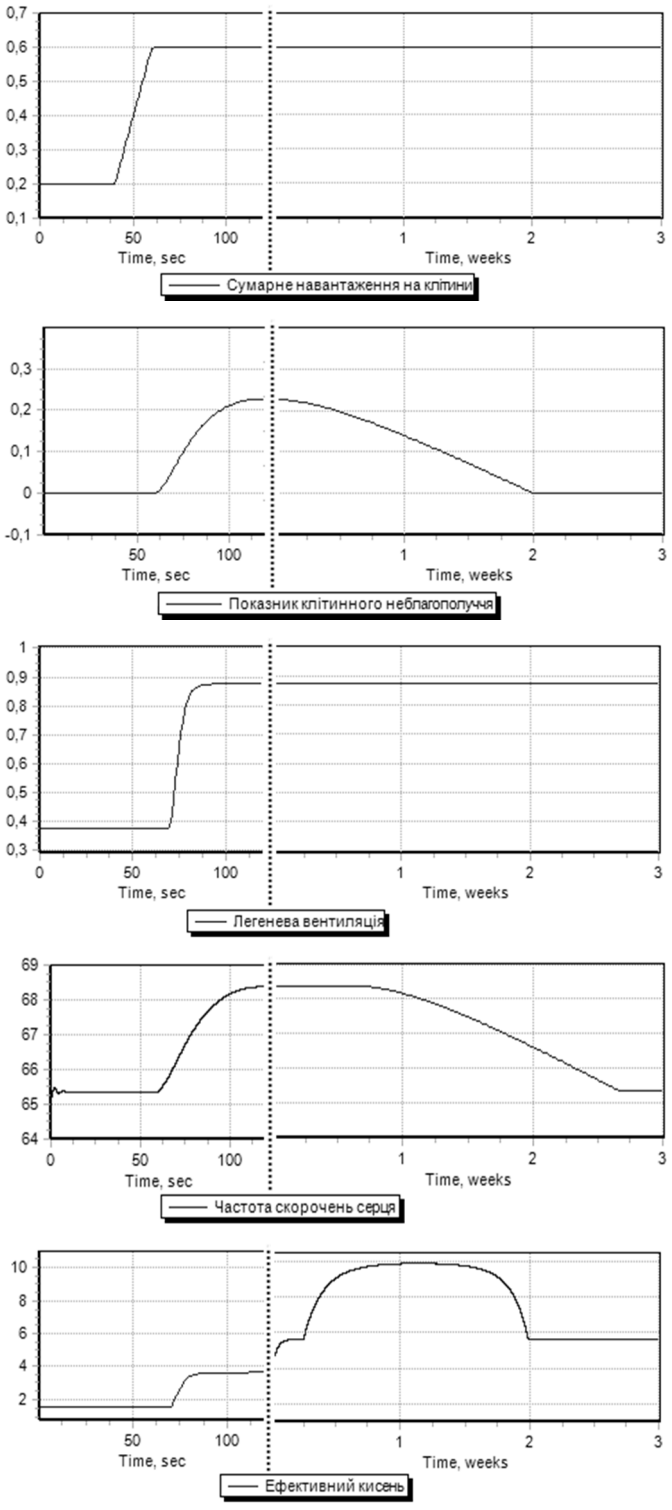


Рис. 2. Ефекти додаткового збільшення еритроцитів крові для клітини із внутрішньоклітинною регуляцією під час навантажень

Як бачимо, при зростанні навантажень з'являються показники клітинного неблагаполуччя, що в свою чергу призводить до розвивитку швидких компенсаторних реакцій системи зовнішнього дихання. При цьому відбувається зростання притоку кисню та, відповідно, рівня АТФ, але даного зростання недостатньо, щоб

здолати навантаження і позбутися факторів неблагополуччя. Отже, далі за умови наявності показників клітинного неблагополуччя еритропоез призводить до поступових змін вмісту ефективного кисню із результатом, подібним до дії додаткового збільшення вентиляції легенів, при чому за кілька тижнів додаткове збільшення еритроцитів призводить до зникнення факторів неблагополуччя. При постійно діючому або довготривалому навантаженні збільшення еритроцитів крові без інших допоміжних мехінізмів не в змозі вивести клітину на рівень відсутності показників клітинного неблагополуччя.

Даний приклад ілюструє можливості використання лише одного компенсаторного механізму. Так само підвищення артеріального тиску є одним із способів збільшення притоку кисню та вуглеводів до клітини, тому його ефективність, як і ефективність кожного іншого способу, має свої обмеження. Попри неможливість ілюструвати графіки всіх можливих сценаріїв симуляцій необхідно відзначити, що однакове збільшення швидкості аеробного синтезу молекул АТФ може бути забезпечено різними, але енергетично нееквівалентними способами. Комплекс адаптивних реакцій організму на дефіцит енергії в клітинах формується в залежності від характеру дефіциту енергії (локальній, регіональний або тотальний) та стану компенсаторних механізмів. При цьому з достатньо великими навантаженнями може впоратися лише та комбінація механізмів, яка включає також паралельну активацію механізмів біогенезу мітохондрій.

V. Висновки

Орієнтований на фізіолога моделювальний комплекс дозволяє проводити експерименти, що симулюють основні реакції здорової людини на різноманітні причини і сценарії розвитку дефіциту енергії. В розробленому моделювальному комплексі було реалізовано і експериментально налаштовано основні моделі енергетичної теорії адаптації: модель батареї механізмів клітини, модель гомеостазу глюкози крові, модель функціонування гемодинамічно замкненої серцево-судинної системи людини, модель збільшення провідності кровоносних судин шляхом локальної вазодилятації і ангіогенеза, модель збільшення вентиляції легенів, модель збільшення кількості гемоглобіну в крові завдяки мобілізації крові з органів-депо і активації еритровоеза. Реалізовано можливість перевизначення існуючих і пікдлючення нових моделей механізмів клітини. Прикладна цінність розробленого моделюючого комплексу полягає у тому, що фізіолог-дослідник отримав сучасний інструмент для тестування теоретичних концепцій та моделей інтегративної фізіології людини. Окрім того, даний моделюючий комплекс може використовуватися як сучасна технологія навчання студентів в вузах біологічного та медичного профілю. Після додаткової доробки моделюючий комплекс також може використовуватися фахівцями-кардіологами для попереднього аналізу діагнозу та пошуку оптимальних сценаріїв корекції стану пацієнту, а також в спортивній медицині.

Література References

[1] Skulachev V.A., Bogachev A. V., Kasparinsky F. O. Principles of Bioenergetics. Springer-Verlag, Berlin Heidelberg, 2013. – 436 p.

[2] Michiels C. Physiological and Pathological Responses to Hypoxia // Am J Pathol., 2004. –164(6) . –P. 1875–1882.

[3] De Gaetano A., Arino O. Mathematical modeling of the intravenous glucose tolerance test. // J. Math. Biol. – 2000. – № 40. – P. 136–168.

[4] Cortassa S., Aon M.A., O'Rourke B.et al. An integrated model of cardiac mitochondrial energy metabolism and calcium dynamics // Biophys. J., 2003,84. – P. 2734–2755.

[5] Pokhilko A.V., Ataullakhanov F.I., Holmuhamedov E.L. Mathematical model of mitochondrial ionic homeostasis: three modes of Ca2+ transport // Journal of Theoretical Biology.— 2006.—243(1).—P.152-169.

[6] Bertram R., Pedersen M., Luciani D., Sherman A. A simplified model for mitochondrial ATP production // Journal of Theoretical Biology.— 2006. —243(4).— P.575–586.

[7] Nazaret C., Heiske M., Thurley K., Mazat J.P. Mitochondrial energetic metabolism: a simplified model of TCA cycle with ATP production.// Journal of Theoretical Biology.—2009.— V.258. I.3 — P.455-464.

[8] Serena S., Nirody J.A., Racz Z.M. Analysis of unstable behavior in a mathematical model for erythropoiesis. Journal of mathematical biology 2013. –66(3) . – P.595-625.

[9] Rafik D. Grygoryan The energy basis of reversible adaptation. Nova Science Publishers, Incorporated, Sep 1, 2012 - Adaptation (Biology) - 262 p.

[10] Р.Д. Григорян, И.И. Дериев, Т.В. Аксенова Программный симулятор реакций аэробной клетки на дисбаланс энергии // Проблеми програмування. – 2014. – № 1. – С. 90–98.

[11] Р.Д. Григорян, Т.В. Аксенова, Р.А. Маркевич Программный симулятор поджелудочной железы // Проблеми програмування. – 2013. – №1. – С. 100-106.

[12] Григорян Р.Д., Аксенова Т.В., Дегода А.Г. Моделирование механизмов и гемодинамических эффектов гипертрофии сердца // Кибернетика и вычислительная техника. – 2016. – Вып.184. – С. 72-83.

[13] Григорян Р.Д., Аксьонова Т.В. Моделювання боротьби механізмів організму з нестачею енергії в клітинах // Вісник університету «Україна». Серія: Інформатика, обчислювальна техника та кібернетика. – 2016. – С. 91-99.

[14] Аксьонова Т.В. Програмна технологія для проведення імітаційних експериментів з математичними моделями фізіологічних систем // Проблеми програмування. – 2012. – №1. – С. 110-120.

Застосування Дискретної Оптимізації для Автоматизації Вибору Курсів Студентами

Г. Добровольський, О. Тодоріко, Є. Кушнір
кафедра комп'ютерних наук
Запорізький національний університет
Запоріжжя, Україна
gen@znu.edu.ua

Application of Dicrete Optimization to Automating of Elective Courses Enrollment

H. Dobrovolskyi, O. Todoriko, E. Kushnir
computer science department
Zaporizhzhya National University
Zaporizhzhya, Ukraine
gen@znu.edu.ua

***Анотація*—Розглядається призначення студентам навчальних курсів з урахуванням їх персональних уподобань. Формулюється відповідна задача дискретної оптимізації з обмеженнями на мінімальну кількість учасників навчальної групи, кількість призначених кожному студенту дисциплін та забороною їх повторного призначення. Студент задає пріоритети наданим курсам, впорядковуючи їх перелік. Навчальні групи формуються за допомогою запропонованого в роботі алгоритму. Якщо студент не зробив свій вибір до визначеного часового обмеження, розстановка пріоритетів від його імені створюється методом симуляції попарного порівняння.**

***Abstract*—Considered the enrollment of students for elective courses taking into account their personal preferences. Formulated discrete optimization problem is accompanied with restrictions of the minimal number of the study group participants, the number of courses assigned to each student, and prohibition of re-appointment. Solution of the problem requires that each student ranks all suggested courses then study groups are collected with the proposed algorithm. Crucial point of the algorithm is a prediction of the course ratings if a student hasn't made his/her choice. The prediction is performed with the simulation of pairwise comparison.**

Ключові слова—задача дискретної оптимізації, пріоритети курсів, метод попарного порівняння, алгоритм формування навчальних груп.

Keywords—discrete optimization problem, course priorities, the method of pairwise comparison , algorithm of study groups forming

I. Вступ

Згідно з законом України «Про вищу освіту» [1], студентам надано право на вибір навчальних дисциплін у межах відповідної освітньої програми та робочого навчального плану, тому актуальною є програмна реалізація та впровадження процедури, яка дозволить викладачам пропонувати свої дисципліни студентам, які мають обґрунтовано вибирати дисципліни, навчальному відділу формувати навчальні групи відповідно до особистих уподобань кожного студента. Але за відсутності засобів автоматизації неможливо повною мірою врахувати особисті вподобання кожного студента – потрібно взяти до уваги занадто багато параметрів, тому вибір дисциплін відбувається тільки в найпростіших варіантах [2,3,4] - з мінімальною кількістю альтернатив та в кілька раундів.

Недолік спрощеної процедури, під час якої студент вибирає із наданого переліку кілька навчальних дисциплін і не вказує свої побажання щодо інших, полягає в тому, що непопулярні курси вибираються занадто малою кількістю студентів. Внаслідок цього процес опитування студентів доводиться або спрощувати або повторювати кілька разів, послідовно виключаючи найменш популярні дисципліни, що вимагає значних організаційних зусиль. В той же час, перевагою найпростішого способу є його зрозумілість, що зменшує кількість непорозумінь.

З 2014 року Запорізький національний університет використовує для інформаційної підтримки навчання безкоштовне програмне забезпечення Moodle [5], у якому зареєстровані всі викладачі, студенти ЗНУ та більшість навчальних дисциплін. Тому, було прийнято рішення автоматизувати вибір дисциплін за допомогою модуля Moodle. З метою мінімізації втрат робочого часу на організацію опитування студентів, було запропоновано виконувати вибір дисциплін наступним шляхом: студент заходить до системи Moodle та встановлює власні пріоритети навчальним курсам, доступним для вибору.

Пробна експлуатація модуля протягом двох років показала, що значною проблемою є наявність студентів, які з різних причин не встановили пріоритети курсам. Це ускладнює формування груп, тому що існує імовірність помилково виключити дисципліну, яка майже зібрала достатню кількість студентів. Таким чином виникає підзадача автоматичного призначення пріоритетів з урахуванням вибору інших студентів.

Задачею студента під час вибору є впорядкування переліку курсів відповідно до своїх уподобань. Пріоритети

можна змінювати багато разів до призначеної дати закінчення. Перелік курсів має вигляд таблиці, в якій студент може змінювати порядок рядків (Рис.1). Чим вище курс знаходиться у списку, тим вищий у нього пріоритет. У кожному рядку міститься назва курсу та посилання, перейшовши за яким, студент може переглянути публічні матеріали, якими викладач презентує свій курс.

Рис. 1 Сторінка студента у модулі Moodle «Вибір курсів ЗНУ»

Після виставлення кожним студентом персональних пріоритетів, автоматично формуються навчальні групи, і студентам призначається визначена навчальним планом кількість дисциплін з урахуванням усіх додаткових обмежень.

II. Формування навчальних груп як задача оптимізації

Вхідними даними алгоритму формування навчальних груп є перелік N запропонованих різних курсів $\{K_1,...,K_N\}$, відображення курсів на перелік навчальних дисциплін $D(K_i) \in \{D_1,...,D_L\}, L \le N$, перелік із M студентів $\{S_1,...,S_M\}$, мінімальний розмір навчальної групи G_{min} та S_r - кількість курсів, які має прослухати студент S_r.

Вважається, що більшість студентів встановили курсам пріоритети, призначивши максимальний пріоритет найбажанішому курсу, тобто майже для кожного студента S_r існує множина пар $(K_i, P_{ri}), i = \overline{1,N}, r = \overline{1,M}$, де пріоритет P_{ri} - ціле число, і чим більшим є пріоритет, тим бажанішим є курс для студента. Кожен курс K_i відповідає одній навчальній дисципліні з навчального плану, але одній дисципліні може відповідати кілька курсів, наприклад, від різних викладачів.

Потрібно максимізувати суму пріоритетів на підмножині призначених курсів:

$$F = \sum_{i=1}^{N} \sum_{r=1}^{M} P_{ri}\delta(r,i) \to max$$

де функція $\delta(r,i) \in \{0,1\}$ відображає факт зарахування студента S_r на курс K_i:

$$\delta(r,i) = \begin{cases} 1, \text{якщо студента зараховано} \\ 0, \text{якщо студента не зараховано} \end{cases}.$$

Відомі наступні додаткові обмеження:

- Кількість студентів G_i у навчальній групі курсу K_i має бути не меншою, ніж G_{min}:

$$\sum_{r=1}^{M} \delta(r,i) \ge G_{min}, i = \overline{1..N}$$

- Студент S_r має бути зарахований на Z_r курсів:

$$\sum_{i=1}^{N} \delta(r,i) = Z_r$$

- Не можна зарахувати студента на дисципліну більше одного разу, наприклад, якщо студент вже вивчав дисципліну в минулому семестрі або вибрав інший курс для вивчення дисципліни.
- Якщо у деякому варіанті зарахування курсові i^* не вистачає студентів, тобто

$$\sum_{r=1}^{M} \delta(r,i^*) < G_{min}$$

то курс i^* видаляється зі списку пропозицій, а студентам призначаються інші курси відповідно до пріоритетів кожного студента.

Описана вище задача належить до класу проблем дискретної оптимізації [6]. Її особливістю є необхідність спочатку прогнозувати вибір студентів, які не розставили пріоритети, а потім формувати навчальні групи. В іншому випадку деякі курси будуть виключені зі списку пропозицій занадто рано. Щоб передбачити вибір студентів, у роботі пропонується прогнозувати розстановку пріоритетів методом попарного порівняння.

III. Прогнозування пріоритетів

Метод попарного порівняння [7] є одним із відомих у соціології та психології способів опитування, призначених для впорядкування деякого набору об'єктів a_1, a_2, ..., a_n (n - кількість об'єктів). Кожному респонденту пропонуються всі пари (a_i, a_j), складені з розглянутих об'єктів. Він повинен вказати, який об'єкт з цієї пари йому подобається більше. Нижче замість виразів виду "об'єкт a_i подобається респонденту більше об'єкта a_j" використовується вираз "$a_i > a_j$".

Отримані таким чином дані записуються в квадратну матрицю $n \times n$, яку називають матрицею парних порівнянь. У загальному вигляді матрицю парних порівнянь для респондента r_q (q = 1, ..., Q, де Q - кількість респондентів) позначимо через $\| (q)_{ij} \|$, де $(q)_{ij}$ =1, якщо респондент r_q відповів, що $a_i > a_j$, і $(q)_{ij}$ = 0, якщо респондент r_q відповів, що $a_j > a_i$. Рейтинг об'єкта a_i обчислюється як

$$\sum_{j=1}^{n} \varphi(q)_{ij} .$$

Приклад матриці парних порівнянь показано в Таблиці 1.

ТАБЛИЦЯ I. ПРИКЛАД МАТРИЦІ ПАРНИХ ПОРІВНЯНЬ, ОТРИМАНОЇ ВІД ОДНОГО РЕСПОНДЕНТА.

	a_1	**a_2**	**…**	**a_n**
a_1	×	1	…	1
a_2	0	×	…	1
…	…	…	…	…
a_n	0	0	…	×

Значення елементів матриці показують, що респондент вибрав $a_1 > a_2$, $a_1 > a_n$ та $a_2 > a_n$.

На головній діагоналі матриці проставлені хрестики, оскільки сам із собою об'єкт не порівнюється. Результати парних порівнянь заслуговують на більшу довіру, ніж ранжування, тому що впорядкувати об'єкти іноді буває досить важко, в той час як попарно їх порівняти набагато легше.

Для моделювання розстановки пріоритетів використовується той факт, що порівняння елементів пари (a_i, a_j) може дати тільки два результати: "$a_i > a_j$" або "$a_i < a_j$". Тому припускається, що імовірність того, що елемент $(q)_{ij}$ матриці попарних порівнянь набуде значення 1, підпорядковується біноміальному розподілу з параметром

$$\varphi_{ij} = \frac{1}{N} \sum_{q=1}^{Q} \varphi(q)_{ij} , \qquad (1)$$

Числа φ_{ij} є середніми значеннями, і сформовану з них матрицю можна назвати середньою матрицею попарних порівнянь. На наступному етапі, маючи параметри біноміальних розподілів, можна згенерувати відповідну їм вибірку матриць попарних порівнянь та на основі кожної матриці обчислити пріоритети курсів.

IV. АЛГОРИТМ ФОРМУВАННЯ НАВЧАЛЬНИХ ГРУП

Формування навчальних груп відбувається наступним чином:

1. Ініціалізація:
   - 1.1. з бази даних прочитуються всі пріоритети курсів, включаючи пусті значення;
   - 1.2. на основі непустих значень пріоритетів знаходиться середня матриця попарних порівнянь (1);
   - 1.3. з допомогою середньої матриці попарних порівнянь генерується вибірка матриць порівнянь та на її основі заповнюються відсутні значення пріоритетів;
   - 1.4. створюється початковий варіант призначення курсів студентам без урахування вимог до мінімального розміру навчальних груп;
   - 1.5. у черзі зберігається початковий варіант призначення із статусом "невизначений";
2. Для кожного варіанта призначення із черги, який має статус "невизначений", виконуються кроки:
   - 2.1. Якщо всі курси у варіанті призначення мають достатню кількість студентів, то обчислюється цільова функція та зберігається разом із варіантом зарахування, якому встановлюється статус "дійсний";
   - 2.2. Якщо не всі курси у варіанті призначення мають достатню кількість студентів,
     - 2.2.1. Для кожного курсу, який має недостатню кількість студентів, в чергу із статусом "невизначений" додається новий варіант призначення, у якому курс виключено, а студенти розподілені на інші курси відповідно до їх пріоритетів.
     - 2.2.2. Варіанту призначення встановлюється статус "недійсний".
3. Завершення алгоритму: серед варіантів призначення із статусом "дійсний" вибирається один із максимальним значенням цільової функції.

Наведена вище послідовність кроків відрізняється від повного перебору тим, що склад студентських груп змінюється тільки за рахунок курсів, які набрали найменше студентів. Названа особливість обумовлена як необхідністю спростити алгоритм, так і потребою зробити його зрозумілішим для всіх студентів і викладачів, зменшивши тим самим кількість претензій після публікації результатів.

Незважаючи на спрощення, наведений вище алгоритм залишається доволі складним для розуміння, що спричиняє труднощі при роз'ясненні викладачам результатів призначення. Тому для попередньої оцінки популярності курсів модуль містить звіт, у якому наведено середні показники пріорітетів, що дозволяє приблизно оцінити кількість студентів для курсів з найвищими пріорітетами.

V. ПРАКТИЧНЕ ЗАСТОСУВАННЯ АЛГОРИТМУ

Нижче наводиться простий приклад призначення курсів студентам. Дані про встановлені студентами та згенеровані пріоритети показані в таблиці 2.

Обчислена на основі виставлених студентами пріоритетів середня матриця попарного порівняння показана на таблиці 3, а згенеровані випадкові елементи матриці порівняння - в таблиці 4.

ТАБЛИЦЯ II. ВСТАНОВЛЕНІ СТУДЕНТАМИ ПРІОРИТЕТИ КУРСІВ. (ЗІРОЧКОЮ ПОЗНАЧЕНІ ПРОГНОЗОВАНІ ЗНАЧЕННЯ.)

Студенти	Курси	
	А. Вступ до спеціальності	***Б. Роль фізики металів у сучасному виробництві***
В***** С. О.	2	1
В***** В. О.	1*	2*
З***** В. В.	1	2
К***** О. В.	2*	1*
М***** К. О.	2	1
Р***** А. В.	2	1
С***** Д. В.	1	2
Ф***** А. І.	2*	1*

ТАБЛИЦЯ III. СЕРЕДНЯ МАТРИЦЯ ПОПАРНОГО ПОРІВНЯННЯ

	А	**Б**
А	×	0.6
Б	0.4	×

ТАБЛИЦЯ IV. ЗГЕНЕРОВАНІ ВИПАДКОВІ ЕЛЕМЕНТИ МАТРИЦІ ПОРІВНЯННЯ

№	**(А>Б)**	**(А<Б)**
1	0	1
2	1	0
3	1	0
...	...	...
98	1	0
99	1	0
100	1	0
Середні значення	0.57142	0.42857

В результаті роботи алгоритму всім студентам призначається курс “А. Вступ до спеціальності”, тому що його пріоритет виявився вищим.

В показаному прикладі група студентів містить лише 8 осіб, 5 із яких встановили пріоритети курсам і 3 - не встановили. Мала кількість респондентів призводить до ненадійної оцінки параметрів і великого відхилення від середніх значень під час генерації. Тому для надійнішого передбачення пріоритетів варто або розглядати більші групи студентів або накласти на автоматично генеровані пріоритети додаткові обмеження, що і реалізується на практиці.

VI. ВИСНОВКИ

У роботі формулюється задача дискретної оптимізації для призначення студентам навчальних курсів відповідно до індивідуальних уподобань кожного студента. Для її успішного розв’язання необхідно, щоб студент замість простого вибору кількох курсів встановив власні пріоритети всім запропонованим курсам. Цільовою функцією, максимум якої потрібно знайти, є сума пріоритетів на підмножині призначених студентам курсів. Додатковими обмеженнями задачі є мінімальна кількість студентів у групі, задана навчальним планом кількість дисциплін, заборона повторного призначення дисципліни.

Пропонується алгоритм для формування студентських груп, який комбінує повний перебір із простою евристикою. Особливістю алгоритму є необхідність знати пріоритети, встановлені всіма студентами, включаючи осіб, які з різних причин не брали участь у процедурі вибору. Щоб передбачити вибір студентів, пропонується прогнозувати розстановку пріоритетів методом симуляції попарного порівняння.

Запропонований алгоритм реалізований у вигляді модуля інформаційної системи Moodle та з 2015 року використовується в Запорізькому національному університеті для формування навчальних груп.

ЛІТЕРАТУРА REFERENCES

[1] Закон України "Про вищу освіту", Відомості Верховної Ради (ВВР), 2014, № 37-38, ст.2004

[2] Procedure: registration for elective courses - bachelor/master http://studerende.au.dk/en/studies/subject-portals/business-communication/teaching/electives/procedure-registration-for-elective-courses-bachelormaster/

[3] Elective Courses | CBS - Copenhagen Business School http://www.cbs.dk/en/study/single-elective-and-supplementary-courses/elective-courses

[4] Elective Courses "College of Dentistry" University of Florida http://dental.ufl.edu/education/dmd-program/elective-courses/

[5] Moodle - Open-source learning platform | Moodle.org, http://moodle.org

[6] О.А. Емец, Т.Н. Барболина, Комбинаторная оптимизация на размещениях, К.:Наукова думка. - 2008, 162 с, http://dspace.uccu.org.ua/handle/123456789/473

[7] Ю.Н Толстова.. Измерение в социологии : учебное пособие / Ю.Н. Толстова. — М.: КДУ. — 288 с. 2007

Геоінформаційна Підсистема Моніторингу Концентрації Радіоактивних і Небезпечних Хімічних Речовин Внаслідок Надзвичайної Ситуації

Петро Гучек
кафедра ІТ
ХНТУ
Херсон, Україна
phuchek@gmail.com

Олена Литвиненко, Ігор Астіоненко
кафедра ВМ і ММ
ХНТУ
Херсон, Україна
mmkntu@gmail.com

Geoinformation Subsystem Monitoring the Concentration of Radioactive and Hazardous Chemicals as a Result of an Emergency

Petro Guchek
dept. of IT
KNTU
Kherson, Ulraine
phuchek@gmail.com

Olena Litvinenko, Igor Astionenko
dept. of HM and MM
KNTU
Kherson, Ukraine
mmkntu@gmail.com

***Анотація*—У разі виникнення надзвичайної ситуації, пов'язаної з виливом (викидом) у довкілля радіоактивних чи небезпечних хімічних речовин, потребується наявність інформації як про саму речовину, площу зараження, так і про величину приземної концентрації токсичних чи радіоактивних речовин в зонах зараження. В рамках дослідження авторами запропоновано застосування експрес-діагностики з використанням серендипових скінченних елементів та геоінформаційної підсистеми моніторингу.**

***Abstract*—In the emergency situation related outpour (emissions) into the environment of radioactive or hazardous chemicals required availability of information as the same substance, the area of infection and about the size of the surface concentration of toxic or radioactive substances into areas of infection. The study authors suggested use of rapid diagnostic using serendipity finite elements and geoinformation subsystem monitoring.**

Ключові слова—метод скінченних елементів; серендипові скінченні елементи; функції форми; надзвичайна ситуація; небезпечні хімічні речовини;

Keywords—finite element method; serendipity finite elements; form functions; emergency situation; hazardous chemicals;

I. Постановка проблеми

В Україні налічується 931 об'єкт, де зберігаються або використовуються у виробничій діяльності 308,07 тис. т. небезпечних хімічних речовин (НХР), зокрема 4,08 тис. т. хлору, 202,66 тис. т. аміаку та 101,33 тис. т. інших небезпечних хімічних речовин. У зонах вірогідного хімічного забруднення мешкає майже 10,24 млн осіб, що складає 22,6% від загальної кількості населення країни. Близько п'яти тисяч закладів та підприємств використовують понад 22000 джерел іонізуючого випромінювання, в тому числі: радіонуклідних джерел – 10 084; генеруючих пристроїв – 12 644[1].

Аварії на таких об'єктах, як правило, супроводжуються забрудненням навколишнього середовища отруйними та радіоактивними речовинами, а також пожежами та вибухами.

Географічне положення України, наявність розвинутої мережі залізниць, автомобільних доріг, морських портів і трубопроводів створює умови для транзиту вантажів у великих обсягах, значна частина з яких є небезпечними вантажами (НВ), що, зокрема, містять радіоактивні та НХР. Крім того, велику загрозу транспортуванню НВ становить збройне протистояння на сході України, не

контрольованість окремих територій і частини державного кордону та загальне зростання тероризму у світі[1].

Тому в умовах надзвичайних ситуацій дуже важливо швидко і правильно прийняти рішення по ліквідації її наслідків. Процес прийняття рішень по ліквідації надзвичайної ситуації характеризується браком часу, неповнотою і поганою якістю подання інформації, необхідної для прийняття рішень.

Таким чином, завчасне створення нових моделей, методів, систем моніторингу та прогнозування розповсюдження небезпечних речовин у повітряному середовищі є актуальними на сьогоднішній день.

II. Аналіз останніх досліджень і публікацій

Математичному моделюванню забруднення атмосферного повітря присвячені роботи Згуровського М.З., Берлянда М.Е., Бизової Н.Л., Колмогорова А.Н., Марчука Г.И., Гаргера Е.К., Іванова В.Н., Яглома А.Н., Моніна А.С., Соловей В.В., Прохача Е.Ю., Бєляєва М.М тощо[2-5].

Розробці інформаційного забезпечення і математичних моделей прогнозування розвитку аварій, що супроводжуються пожежами на різноманітних об'єктах народного господарства, оцінювання обстановки і вироблення управлінських рішень з метою створення автоматизованих систем підтримки та прийняття рішень керівниками ліквідації аварій і надзвичайних ситуацій присвячені результати досліджень, які описані в роботах М.М. Брушлинського, П.П. Клюса, І.Ф. Кимстача, В.Г. Теряєва, Ю.О. Абрамова, О.А. Громовенка, С.М. Мінаєва, Б.Ф. Турніна, П.М. Бортнічука, С.В. Юхимчука, Р.В. Котельникова, В.Ф. Шостака, В.А. Геловані, А.В. Матюшкіна, У. Бейкера, В. Маршалла та інших вітчизняних і зарубіжних авторів.

У методі скінченних елементів (МСЕ) важливу роль відіграють скінченні елементи серендипової сім'ї [6]. Для побудови базису скінченного елемента традиційно використовують *матричну процедуру* [6]. *Процедура систематичного генерування базису*, що була запропонована Тейлором у 1972 році, привела до вже відомих стандартних моделей на елементах серендипової сім'ї [7]. У 70-х роках минулого століття завдяки роботам Уачспресса з'явився метод *"product of planes"* для конструювання базисних функцій СЕ [8]. Але цей метод не застосовувався на серендипових скінченних елементах (ССЕ). На початку 80-х років був запропонований *ймовірнісно-геометричний метод* конструювання базисів скінченних елементів різноманітної конфігурації [9]. Переваги цього метода найбільш виразно проявилися саме на серендипових моделях. Для конструювання серендипових елементів використовувався *геометричний метод*[10]. Це модифікація методу *" product of planes "*, яка використовує техніку перемноження рівнянь площин і поверхонь другого порядку. Нові методи значно спрощують процедуру побудови базису (не виникає потреби розв'язувати СЛАР відповідного порядку на елементі) і дозволяють отримати альтернативні моделі ССЕ. Наявність " позавузлових " параметрів у моделях, що отримані за допомогою нових методів, дає можливість позбутися недоліків, які притаманні стандартним моделям (наприклад, від'ємних значень навантажень у вузлах).

Але побудова систем базисних функцій цими методами не дає можливості будувати базиси на ССЕ з наперед заданими характеристиками. Для розв'язання на ССЕ задачі інтерполювання з умовами [11] був запропонований *аналітичний метод побудови ієрархічних форм базисних функцій* [12].

III. Виклад основного матеріалу

За допомогою матричних методів на серендипових елементах отримані базиси, які називаються *стандартними.* Матричний метод був розвинений в працях Вандермонда, Коші, Лагранжа стосовно задач інтерполяції функцій одного аргументу. У МСЕ він вважається основним методом побудови інтерполяційних поліномів для функцій двох і трьох аргументів. Наприклад, процедура побудови базису для СЕ з 4-ма вузлами (рис.1) починається з вибору 4-параметричного поліному з двома аргументами:

$$\varphi(\xi,\eta)=\alpha_1+\alpha_2\xi+\alpha_3\eta+\alpha_4\xi\eta \text{ , де } |\xi|\le 1, |\eta|\le 1 \qquad (1)$$

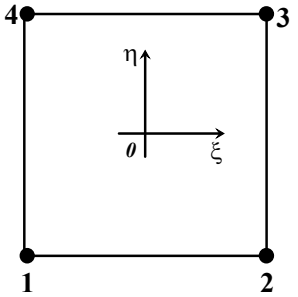


Рис. 1. Елемент з білінійним базисом (СЕ-4)

Базисні функції білінійної інтерполяції $N_i(\xi,\eta)\ \ (i=1,2,3,4)$ повинні задовольняти інтерполяційній гіпотезі

$$N_i(\xi_k,\eta_k)=\delta_{ik}\ ,\quad i,k=\overline{1,4}\ , \qquad (2)$$

де δ_{ik} – символ Кронекера.

Використання (2) приводить до системи лінійних алгебраїчних рівнянь 4×4 відносно параметрів α_j :

$$\alpha_1+\alpha_2\xi_k+\alpha_3\eta_k+\alpha_4\xi_k\eta_k=\delta_{ik}\ , \qquad (3)$$

де i – номер базисної функції; k – номер вузла.

Знайдемо невідомі параметри α_j і запишемо інтерполяційний поліном (1) у вигляді [6]:

$$\varphi(\xi,\eta)=N_1(\xi,\eta)\Phi_1+N_2(\xi,\eta)\Phi_2+N_3(\xi,\eta)\Phi_3+N_4(\xi,\eta)\Phi_4=[N_i]\{\Phi_i\}, \qquad (4)$$

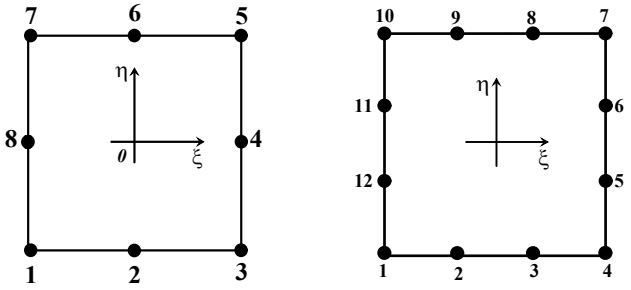
де $[N_i]$ – матриця-рядок базисних функцій СЕ ($i=1,2,3,4$); $\{\Phi_i\}$ – матриця-стовпець граничних значень функції, що інтерполюється.

Отримуємо білінійний базис на СЕ-4:

$$N_i = \frac{1}{4}(1+\xi_i\xi)(1+\eta_i\eta), \quad i=1,2,3,4; \quad \xi_i,\eta_i = \pm 1. \qquad (5)$$

Функції форми (базисні функції) $N_i(\xi,\eta)$ виконують важливу роль в методі скінченних елементів – апроксимують поле в середині елемента. Від вибору апроксимуючих функцій та їх властивостей в значній мірі залежить точність розв'язку. Слід зауважити, що серендипові моделі є унікальним прикладом одночасної інтерполяції та апроксимації – вони інтерполюють функцію на границі елемента та апроксимують всередині його.

У роботі розглядаються скінченні елементи серендипової сім'ї з біквадратичною та бікубічною інтерполяцією (рис. 2, рис. 3) [6,19].



ССЕ-8 ССЕ-12

$(|\xi| \le 1, |\eta| \le 1)$ $(|\xi| \le 1, |\eta| \le 1)$

Рис. 2. Використовуючи аналітичний метод побудови ієрархічних форм базисних функцій, множину базисів ССЕ-8 можливо отримати додаванням до поверхні стандартного базису [12]:

$$N_1^{(S)}(\xi,\eta) = \frac{1}{4}(1-\xi)(1-\eta)(-1-\xi-\eta), \qquad (6)$$

$$N_2^{(S)} = \frac{1}{2}(1-\xi^2)(1-\eta) \qquad (7)$$

гіперболічного параболоїда з коефіцієнтом K:

$$N_1 = \frac{1}{4}(1-\xi)(1-\eta)(-1-\xi-\eta+K(1+\xi)(1+\eta)), \qquad (8)$$

$$N_2 = \frac{1}{2}(1-\xi)(1-\eta)\left((1+\xi)-\frac{1}{2}K(1+\xi)(1+\eta)\right) \qquad (9)$$

де $N_1^{(S)}, N_2^{(S)}$ - стандартні базисні функції ССЕ-8,

N_1, N_2 - модифіковані базисні функції ССЕ-8.

Запишемо узагальнені формули для побудови альтернативних базисних функцій біквадратичного елемента:

$$N_i = \frac{1}{4}(1+\xi_i\xi)(1+\eta_i\eta)(-1+\xi_i\xi+\eta_i\eta+K(1-\xi_i\xi)(1-\eta_i\eta)) \qquad (10)$$

,

$$i=1,3,5,7; \quad \xi_i = \pm 1; \quad \eta_i = \pm 1.$$

$$N_i = \frac{1}{2}(1-\xi)(1+\eta_i\eta)\left((1+\xi)-\frac{1}{2}K(1+\xi)(1-\eta_i\eta)\right), \qquad (11)$$

$$i=2,6; \quad \eta_i = \pm 1.$$

$$N_i = \frac{1}{2}(1-\eta)(1+\xi_i\xi)\left((1+\eta)-\frac{1}{2}K(1+\eta)(1-\xi_i\xi)\right), \qquad (12)$$

$$i=4,8; \quad \xi_i = \pm 1.$$

Аналогічно використовуючи аналітичний метод побудови ієрархічних форм базисних функцій отримуємо множину базисів ССЕ-12 додаванням до поверхні стандартного базису [12]:

$$N_1^{(S)} = \frac{1}{32}(1-\xi)(1-\eta)(9\xi^2+9\eta^2-10), \qquad (13)$$

$$N_2^{(S)} = \frac{9}{32}(1-\xi^2)(1-\eta)(1-9\xi), \qquad (14)$$

гіперболічного параболоїда з коефіцієнтом K:

$$N_1 = \frac{1}{4}(1-\xi)(1-\eta)\left(\frac{1}{8}(9\xi^2+9\eta^2-10)+K\cdot\frac{9}{4}\cdot(1+\xi)(1+\eta)\right), \qquad (15)$$

$$N_2 = \frac{3}{8}(1-\xi)(1-\eta)\left(\frac{3}{4}(1+\xi)(1-3\xi)-K\cdot\frac{3}{4}\cdot(1+\xi)(1+\eta)\right), \qquad (16)$$

де $N_1^{(S)}, N_2^{(S)}$ - стандартні базисні функції ССЕ-12,

N_1, N_2 - модифіковані базисні функції ССЕ-12.

Запишемо узагальнені формули для побудови альтернативних базисних функцій бікубічного елемента:

$$N_i = \frac{1}{4}(1+\xi_i\xi)(1+\eta_i\eta)\times$$
$$\times\left(\frac{1}{8}(9\xi^2+9\eta^2-10)+K\cdot\frac{9}{4}\cdot(1-\xi_i\xi)(1-\eta_i\eta)\right), \qquad (17)$$
$$i=1,4,7,10; \quad \xi_i,\eta_i = \pm 1;$$

$$N_i = \frac{3}{8}(1+3\xi_i\xi)(1+\eta_i\eta)\times$$
$$\times\left(\frac{3}{4}(1-3\xi_i\xi)(1+9\xi_i\xi)-K\cdot\frac{3}{4}(1-3\xi_i\xi)(1-\eta_i\eta)\right), \qquad (18)$$
$$i=2,3,8,9, \quad \eta_i = \pm 1, \quad \xi_i = \pm\frac{1}{3}.$$

$$N_i = \frac{3}{8}(1+3\eta_i\eta)(1+\xi_i\xi)\times$$
$$\times\left(\frac{3}{4}(1-3\eta_i\eta)(1+9\eta_i\eta)-K\cdot\frac{3}{4}(1-3\eta_i\eta)(1-\xi_i\xi)\right), \qquad (19)$$
$$i=5,6,11,12, \quad \xi_i = \pm 1, \quad \eta_i = \pm\frac{1}{3}.$$

Отримані базисні функції відповідають всім властивостям, які притаманні функціям форми в МСЕ [6,12].

У [13,17-18] авторами була розроблена автоматизована підсистема дослідження моделей скінченних елементів, яка дає користувачеві можливість подальшого дослідження отриманих моделей, вирішувати практичні прикладні задачі, проводити порівняльну характеристику альтернативних моделей і дозволяє приймати рішення щодо подальшого застосування та оптимізації обчислювальних властивостей моделей.

На відміну від традиційної поелементної процедури метода скінченних елементів у роботі показано, що поле приземної концентрації небезпечної радіоактивнооі чи хімічної речовини (рис.1-3) можна побудувати лише за допомогою одного елемента серендипової сім'ї (експрес-методика), що буває дуже важливо при визначенні нульового наближення розв'язку задачі, яка досліджується.

Концентрація C в будь-якій внутрішній точці при використанні ССЕ-8 визначається за допомогою аналітичної залежності:

$$C = \sum_{i=1}^{8} N_i(\xi,\eta)\cdot C_i \,. \qquad (20)$$

Значення концентрації небезпечної речовини у вузлових точках C_i можуть бути отримані безпосередньо за допомогою датчиків, на прикладі системи [14,16]. Датчики розташовуються як на стаціонарних постах збору інформації, так і на оперативних пересувних лабораторіях, які обладнанні відповідними каналами передачі даних на сервер автоматизованої системи моніторингу.

Підсистема обробки та аналізу інформації (рис.4) за допомогою спеціальної служби через певні інтервали часу виконує запити до постів збору та передачі інформації для зчитування, обробки та зберігання в базі даних актуальної інформації про стан та концентрацію небезпечних речовин.

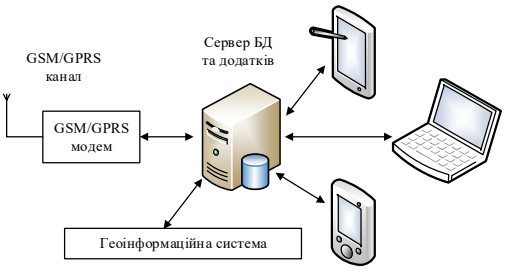


Рис. 3. Підсистема обробки та аналізу інформації

Крім того, комплекс має у своєму складі метеостанцію, яка призначена для вимірювання метеорологічних параметрів, таких як швидкість і напрямок вітру, температура повітря, вологість, атмосферний тиск.

Розроблена підсистема моніторингу (рис. 5) дозволить надати наступні переваги в порівнянні зі звичайними механізмами та методами:

- отримання об'єктивних даних в реальному часі про концентрацію небезпечних речовин в місцях встановлення відповідних датчиків;
- проведення експрес-діагностики при застосуванні запропонованої моделі для скінченно-елементної апроксимації концентрації небезпечних речовин на область, яку покривають датчики;
- зв'язок даної підсистеми з ГІС системою дозволяє виносити на певний шар відповідні значення концентрації та візуалізувати зону надзвичайної ситуації і прогнозувати можливі наслідки;
- проводити моніторинг наявних та задіяних в ліквідаціях надзвичайних ситуацій сил та засобів на базі інформації, що надходить від GPS трекерів.
- надавати оперативну інформацію мобільним підрозділам про стан, зону надзвичайної ситуації та задіяні підрозділи за допомогою web-сервісів.

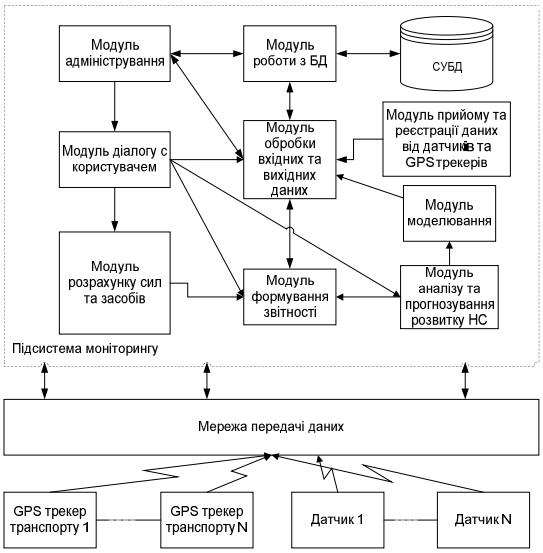


Рис. 4. Підсистема моніторингу

Наприклад, застосовуючи розглянуті методи та моделі, проведемо експрес-діагностику концентрації небезпечних речовин в приземному шарі м. Херсона за даними, які поступають від датчиків, розташованих з географічними координатами, представленими на рис.6.

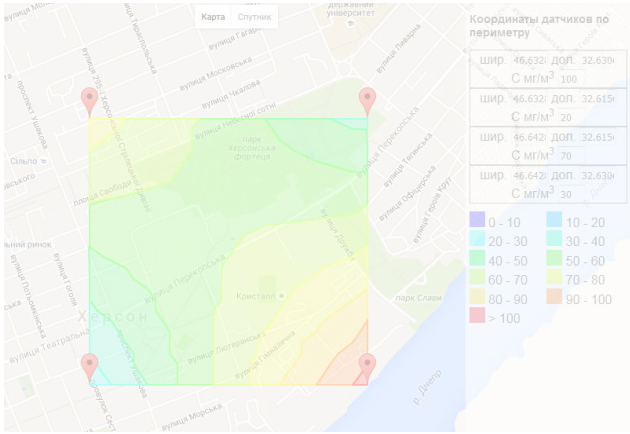


Рис. 5. Поширення концентрації небезпечних речовин

Для проведення експрес-діагностування застосуємо стандартний базис білінійного скінченного елемента (5). Концентрація C в будь-якій внутрішній точці при використанні СЕ-4 визначається за допомогою аналітичної залежності(21):

$$C = \sum_{1}^{4} N_i(\xi,\eta)\cdot C_i \,. \qquad (21)$$

Білінійний елемент успішно використовується для апроксимації функцій з малим градієнтом. Для функцій, що швидко змінюються, отримати більш точну апроксимацію у МСЕ можна послідовним подрібненням сітки білінійних елементів. Але це призводить до зростання обчислювальної похибки. З іншого боку, збільшення точності наближення може бути досягнуто послідовним збільшенням степеню базисних функцій. При цьому розміри елементів залишаються без змін, а покращують представлення полів у елементі за допомогою зростаючої кількості вузлів, наприклад, використовуючи ССЕ-8, ССЕ-12, ССЕ-16.

Для визначення концентрації не тільки в приземному шарі необхідно застосовувати просторові скінчені елементи. При цьому датчики розташовують як в приземному шарі, так і у просторі – на висотних будівлях, антенах мобільного зв'язку, а також на різноманітних літальних апаратах, які в динамічному режимі передають дані до ГІС.

Методика, яка на сьогоднішній день використовується державною службою з надзвичайних ситуацій, дозволяє лише експрес розрахунок зони можливого поширення небезпечних речовин під час виникнення надзвичайної ситуації (рис.7) [15].

Рис. 6. Зона поширення небезпечних речовин

Тому для визначення концентрації небезпечних речовин пропонується використовувати дискретні методи відновлення гармонічної функції та геоінформаційні технології.

IV. Висновки

У роботі розглянуто скінченно-елементну модель для експрес-діагностики концентрації небезпечних радіоактивних та хімічних речовин під час викиду (виливу) внаслідок надзвичайної ситуації. Пропонується розробка геоінформаційної системи безперервного моніторингу та прогнозування зони поширення небезпечних речовин.

Перспективним є подальше дослідження з використанням різноманітних моделей, та їх порівняльна характеристика для даного класу задач.

Література References

[1] Аналітичний огляд стану техногенної та природної безпеки в Україні за 2015 рік. – Режим доступу: http://undicz.dsns.gov.ua/ua/Analitichniy-oglyad-stanu-tehnogennoyi-ta-prirodnoyi-bezpeki-v-Ukrayini.html

[2] Марчук Г.И. Математическое моделирование в проблеме окружающей среды / Г.И. Марчук. – М.: Наука, 1982.–320 с.

[3] Берлянд М.Е. Современные проблемы атмосферной диффузии загрязнения атмосферы / М.Е. Берлянд. – Л.: Гидрометеоиздат, 1975. – 448 с.

[4] Бызова Н.Л. Экспериментальные исследования атмосферной диффузии и расчеты рассеяния примесей / Н.Л. Бызова, Е.Г. Гаргер, В.Н. Иванов. – Л.: Гидрометеоиздат, 1991. – 273 с.

[5] Семенчин Е.А. Стохастические методы решения обратных задач в математической модели атмосферной диффузии / Е.А. Семенчин, М.В. Кузякина. – М.: Физматлит, 2012. – 176 с.

[6] Зенкевич О. Метод конечных элементов в технике / О. Зенкевич. — М. : Мир, 1975. — 541 с.

[7] Taylor R. L. On the completeness of shape functions for finite element analysis / R. L. Taylor // Internat. J. Numer. Methods Eng. — 1972. —V.4. — № 1. — P. 17-22.

[8] Wachspress E.I. A rational finite element basis / E.I. Wachspress. — Academic Press. — New York , 1975. — 344 p.

[9] Хомченко А.Н. О вероятностном построении базисных функций МКЭ / А. Н. Хомченко // Ивано-Франковс. ин-т нефти и газа. — Ивано-Франковск , 1982. — 5 с. — Деп. в ВИНИТИ 21.10.1982, №5264.

[10] Хомченко А.Н. Геометрия серендиповых аппроксимаций / А.Н. Хомченко, Е.И. Литвиненко, П.И. Гучек // Прикл. геом. и инж. графика. — К. : Будівельник, 1996. — Вып. 59. — С. 40-42.

[11] Попов Б.А. Приближение функций для технических приложений / Б.А. Попов, Г.С. Теслер. — Киев : Наукова думка, 1980. — 352 с.

[12] Хомченко А.Н. Новый подход к построению базисов серендиповых элементов / А.Н. Хомченко, Е.И. Литвиненко, И.А. Астионенко //Геометричне та комп'ютерне моделювання. – 2009. – Вип. 23. – С. 90-95.

[13] Астионенко И.А. Автоматизированная подсистема исследования моделей конечных элементов / И.А. Астионенко, П.И.Гучек, Е.И. Литвиненко // Проблеми інформаційних технологій. – 2014. – № 1. – С. 222-228.

[14] Гучек П.Й. Автоматизація моніторингу та прогнозування розповсюдження небезпечних хімічних речовин внаслідок надзвичайних ситуацій / П.Й. Гучек // Інформаційні технології та комп'ютерна інженерія: матеріали V міжнародної науково-практичної конференції, м. Івано-Франківськ – Ворохта – Вінниця, 27-29 травня 2015 р. – Івано-Франківськ: ПНУ, 2015. – С.25-27.

[15] Методика прогнозування наслідків виливу(викиду) небезпечних хімічних речовин при аваріях на промислових об'єктах і транспорті (затверджена спільним наказом МНС України, Міністерства аграрної політики, Міністерства економіки, Міністерства екології і природних ресурсів від 27.03.2001 N 73/82/64/122). – Режим доступу: http://zakon2.rada.gov.ua/laws/show/z0326-01.

[16] Гучек П.Й. Скінченно-елементна модель для оцінки та прогнозування хімічної обстановки внаслідок надзвичайних ситуацій / П.Й. Гучек // Комунальне господарство міст. – 2015. – № 1. – С. 6-9.

[17] Гучек П.Й. Система управління базою моделей скінченних елементів / П.Й. Гучек, О.І. Литвиненко // Проблеми інформаційних технологій . – 2012. – №2. – С. 54-58.

[18] Гучек П. Й. Інтерактивна процедура візуалізації функцій форми на серендипових елементах / П. Й. Гучек, О. І. Литвиненко, А. Н. Хомченко // Вестник Херсонского национального технического университета. – 2012. – Вып. 1 (44). – С. 274-280.

[19] Астионенко И.А. Многопараметрические серендиповы элементы в 2D /И.А. Астионенко, П.И. Гучек, Е.И. Литвиненко, А.Н. Хомченко // Вестник Херсонского национального технического университета. – Херсон, 2010.–№3(39) – С.40–47.

Застосування Зростаючої Пірамідальної Мережі для Вирішення Задачі Ідентифікації Людини по її ЕКГ

Віталій Вишневський
Інститут проблем математичних машин і систем
НАН України
Київ, Україна
vit.vizual@gmail.com

Лілія Кізуб
Інститут проблем математичних машин і систем
НАН України
Київ, Україна
lilimmsp@gmail.com

Віталій Величко
Інститут кібернетики ім. В.М. Глушкова НАН України
Київ, Україна
velychko@aduis.com.com

Applications of Growing Pyramidal Network for Solving to the Problem of Humen Identification with Using Electrocardiogram

Vitaliy Vishnevskey
Institute of Mathematical Machines and
Systems of NASU
Kyiv, Ukraine
vit.vizual@gmail.com

Liliya Kizub
Institute of Mathematical Machines and
Systems of NASU
Kyiv, Ukraine
lilimmsp@gmail.com

Vitalii Velychko
V. M. Glushkov Institute of Cybernetics of NASU
Kyiv, Ukraine
velychko@aduis.com.com

***Анотація*—У документі розглядається вирішення задачі ідентифікації людини за її електрокардіограмою, з допомогою автоматичного класифікатора. При наявності відповідних ознак – координат точок канонічного сплайну, що апроксимує тривимірну криву QRS - комплексу, з допомогою зростаючих пірамідальних мереж може бути проведена класифікація ЕКГ.**

***Abstract*—In this paper, a solution to the problem of human identification by human's electrocardiogram has been considered with using an automatic classifier. ECG can be classified by the growing pyramidal networks when there are corresponding signs such as coordinates of the canonical spline points which approximate the three - dimensional curve of the QRS - complex.**

Ключові слова—електрокардіограма (ЕКГ); зростаюча пірамідальна мережа; ідентифікація людини; QRS-комплекс; вектор ознак

Keywords—electrocardiogram (ECG); growing pyramidal network; human identification; QRS-complex; characteristic vector

I. Вступ

Ідеї щодо використання спеціальних методів обробки електрокардіограми в якості біометричної ідентифікації особистості (автентифікації) сформульовані відносно недавно [1-3]. Однак розвиток телемедичних технологій робить ідею такої автентифікації все більш привабливою. Річ у тому, що завдяки тотальному розповсюдженню

смартфонів та мобільного інтернету на фоні стагнації ринку медичних приладів для кабінетів функціональної діагностики, експертами вже зараз фіксується бурхливий розвиток малогабаритних мобільних медичних приладів, зокрема електрокардіографів, що використовуються в концепції мобільної та домашньої телемедицини. Оскільки діагностичні послуги для пацієнта на основі його електрокардіограми (6-ть або 12 відведень) виконуються на відстані, актуальність біометричної ідентифікації (автентифікації) набуває особливого значення, адже навіть для аналізу динаміки змін в кардіосигналі, лікар має бути впевнений, що аналізує сигнал тієї ж самої людини. Не менш актуальною, є вирішення задачі автентифікації за електрокардіограмою і в умовах впровадження страхової медицини та парадигми персоналізованої діагностики.

Отже, ідентифікація (автентифікація) людини за її електрокардіограмою в 6-ти або 12-ти стандартних відведеннях є актуальною саме зараз, на сучасному етапі розвитку мобільних медичних пристроїв, для ефективної організації телемедичних консультацій.

В роботах [4-7] нами наведені алгоритмічні рішення щодо формування системи інформаційних ознак у вигляді керуючих точок параметричного сплайну, які відповідають QRS-комплексу електрокардіограми в трьох ортогональних відведеннях або перших трьох відведеннях. Наступний етап роботи с цією системою ознак повинен використовувати автоматичний класифікатор, який власне і буде проводити ідентифікацію пацієнта. Ця стаття присвячена використанню в якості такого класифікатора «зростаючої пірамідальної мережі».

II. Пірамідальна мережа та система CONFOR

Пірамідальною мережею [8] називається ациклічний орієнтований граф, в якому немає вершин, що мають одну дугу, що заходить. Вершини, що не мають дуг, що заходять, називаються рецепторами, інші вершини - концепторами. При побудові мережі вхідною інформацією служать набори значень ознак, що описують деякі об'єкти (матеріали, стани агрегату, ситуації, хвороби). Рецептори відповідають значенням ознак. У різних завданнях це можуть бути імена властивостей, стосунків, станів, дій, об'єктів або класів об'єктів. Концептори відповідають описам об'єктів в цілому і перетинам описів. У початковому стані мережа складається тільки з рецепторів. Концептори формуються в результаті роботи алгоритму побудови мережі.

Концептори мережі відповідають поєднанням значень ознак, що визначають кон'юнктивні класи об'єктів. При включенні вершин в піраміду об'єкту здійснюється прив'язка об'єкту до класів, визначення яких представлені цими вершинами. Таким чином при побудові мережі формуються кон'юнктивні класи об'єктів, тобто здійснюється класифікація без учителя. Класифікуючі властивості пірамідальної мережі мають велике значення для моделювання середовищ і ситуацій. Алгоритм побудови мережі забезпечує автоматичне встановлення асоціативної близькості між об'єктами на основі спільних елементів їх описів. Усі процеси, пов'язані з побудовою мережі, при обробці опису одного об’єкту локалізуються у відносно невеликій частині мережі - піраміді, що відповідає цьому опису.

Пірамідальні мережі зручні для виконання різних операцій асоціативного пошуку. Наприклад, можна вибрати усі об'єкти, що включають задану множину значень ознак, простежуючи шляхи, що виходять з вершини мережі, яка відповідає цій множині. Для вибірки усіх об'єктів, описи яких перетинаються з описом заданого об'єкту, достатньо простежити шляхи, що виходять з вершин, які утворюють його піраміду.

Важливою властивістю пірамідальних мереж є їх ієрархічність, що дозволяє природним чином відображати структуру складених об'єктів і родовидові зв'язки.

В даній роботі для діагностики і прогнозування використовувалась система CONFOR [9]. В основу системи CONFOR [10] покладено метод індуктивного формування понять (CONcept FORmation), який базується на використанні зростаючих пірамідальних мереж. CONFOR призначений для виявлення знань, виділення закономірностей, властивих даним, використання закономірностей для вирішення задач класифікації, діагностики і прогнозування. Система CONFOR працює в трьох режимах: навчання, розпізнавання, екзамен.

Підсистема навчання здійснює виділення закономірностей (знань) які характеризують класи об’єктів [11]. Вхідними даними для підсистеми навчання служить вибірка для навчання, яка включає приклади об’єктів, що характеризуються різними класами. Кожен клас об’єктів має бути представлений деякою кількістю прикладів, достатнім для того, щоб на підставі їх аналізу була виділена закономірність, що характеризує цей клас. Приклади об’єктів задаються ознаковими описами. Ознаковий опис об’єкту повинен включати ім'я, клас, до якого об’єкт належить, і набір значень ознак, які характеризують об’єкт. Вихідні дані для підсистеми навчання – закономірності які характеризують класи об’єктів.

Процес навчання складається з наступних стадій:

- представлення початкових даних у вигляді пірамідальної мережі;
- формування понять на основі пірамідальної мережі;
- побудова логікових виразів, що відповідають сформованим поняттям.

Блок побудови пірамідальної мережі реалізує першу стадію процесу навчання, коли внутрішнє представлення об'єктів вибірки для навчання перетворюється в пірамідальну мережу.

Результатом роботи CONFOR'а в режимі навчання є узагальнена модель класів об'єктів, що включає найбільш характерні властивості цих об'єктів. Важливою особливістю алгоритму є можливість включати в поняття ознаки, які не належать об'єктам досліджуваного класу. В результаті поняття, сформовані алгоритмом, мають більш досконалу логікову структуру, що дає змогу підвищити точність діагностики та прогнозування.

Підсистема іспиту призначена для тестування якості навчання комплексу. Вхідними даними для підсистеми іспиту служить екзаменаційна вибірка що включає приклади об'єктів, які не увійшли до вибірки для навчання, але відносно яких відомо, до якого класу вони належать. Вихідні дані підсистеми - її відповідь відносно класу, до якого віднесена кожен об'єкт, і статистичні дані про кількість правильних, неправильних і невизначених відповідей підсистеми. Отримані результати зіставляються з інформацією про реальну приналежність об'єктів до відповідних класів, на основі чого можна судити про те, наскільки добре навчений комплекс.

Підсистема розпізнавання призначена для ідентифікації класу до якого належить новий об'єкт.

Вхідними даними для підсистеми розпізнавання є ознакові описи нових об'єктів. Розпізнавання виконується на підставі зіставлення опису нового об'єкту і понять, сформованих на етапі навчання. Вихідні дані підсистеми - її відповідь відносно класу, до якого віднесений об'єкт, що розпізнається..

III. Використання пірамідальної мережі для класифікації електрокардіограм

В нашому дослідженні у режимі навчання вихідними даними для системи являються об'єкти вибірки для навчання – QRS-комплекси, які належать класу, який досліджується (певному пацієнту). Об'єкти вибірки для навчання представляються наборами значення ознак, тобто координатними точками канонічного сплайну, що апроксимує тривимірну криву. У режимах розпізнавання або іспиту вихідними даними є описи пацієнтів, які повинні бути розпізнані.

Для експерименту були зареєстровані електрокардіограми різних людей, функціонально здорових, віком від 25 до 55 років, чоловічої і жіночої статі. Дослідження проводились у різний час доби, пацієнти знаходились у різному емоційному та фізичному стані. Для кожної персони були зареєстровані по 10 електрокардіограм і виділено з кожної від 90 до 100 QRS-комплексів. Відповідно до кожної з ЕКГ отримано вектор ознак.

Вихідні дані у вигляді таблиці об'єкт - властивість було необхідно підготувати для передачі в систему CONFOR. На цьому етапі виконувалась дискретизація даних. Дискретний аналіз розподілень об'єктів на шкалах ознак відіграє важливу роль як процес попередньої обробки інформації в різних задачах атрибутивного аналізу. Щоб виконати будь-яке логікове узагальнення для об'єктів, що описуються числовими ознаками, необхідно вміти відповідати на запитання: на скільки повинні відрізнятись числа, які характеризують деяку ознаку, для того, щоб їх можна було розглядати як одне значення ознаки. Щоб підвищити ефективність методів пошуку логікових закономірностей потрібно спрощувати простір ознак, які використовуються для опису об'єктів. Одним з прийомів такого спрощення є використання ознак тільки номінального типу. Для цього діапазони значень числових ознак поділяють на інтервали, що не перетинаються. Кожному виділеному інтервалу присвоюється індивідуальне ім'я. Таким чином, неперервні шкали числових ознак трансформуються в дискретні, а самі числові ознаки стають номінальними.

Інтервали, які виділяться в результаті дискретизації, повинні в найбільшій мірі характеризувати класи об'єктів, що мають однакові значення цільової властивості, тобто властивості, відносно значень якої досліджуються існуючі функціональні залежності між значеннями інших властивостей об'єктів. Проведення дискретизації не повинно приводити к виникненню нових функціональних залежностей, які відсутні у просторі числових ознак об'єктів. Якщо залежність між значеннями нецільових властивостей та значеннями цільової властивості об'єкта існує, але не задана явно, необхідно щоб така залежність максимально була виявлена на етапі дискретизації. Номінальні ознаки, сформовані в результаті дискретизації шкал нецільових властивостей об'єктів, повинні забезпечувати розділимість об'єктів по їх значенням цільової властивості в просторі номінальних ознак. Наведеним вимогам відповідає алгоритм дискретизації, який розглянемо нижче.

Нехай задана система ознак $A = \{A_1, A_2, \ldots, A_n\}, n = |A|$, а для кожної з ознак заданий домен D_i - множина значень ознаки A_i. Тоді кожний об'єкт q можна представити множиною $D_q = \left\{ d_i^q \middle| d_i^q \in D_i, i \in \overline{I,n} \right\}$. Для кожної i властивості виконується розбиття діапазону її значень на κ інтервалів, рівних за довжиною та отримуємо множину $\tilde{D}_i = \left\{ I_1^{(l)}, I_2^{(l)}, \ldots, I_n^{(l)}, \ldots, I_k^{(l)} \right\}$, $k = |\tilde{D}_i|$.

Для кожного інтервалу $I_n^{(l)}$ по кожній властивості V формується множина ваг $\left\{ w_1^n, w_2^n, \ldots, w_j^n, \ldots, w_l^n \right\}$: $w_j^n = Card\left\{ q \middle| d_i^q \in I_n^{(l)} \right\}, j = \overline{I,l}$, $l = Card(V)$ – кількість значень властивості, яка досліджується.

Далі визначається тип кожного з інтервалів $I_n^{(l)}$:

$I_n^{(l)}$ - порожній, якщо для всіх ***j*** виконується w_j^n =0;

$I_n^{(l)}$ - чистий, якщо існує одне w_j^n таке, що w_j^n >0;

$I_n^{(l)}$ - змішаний, якщо інтервал не належить до порожнього або чистого.

До кожного не порожнього інтервалу приєднується по обидва боки задану кількість порожніх інтервалів. Інтервали одного типу, які розташовані поруч, об'єднуються. В результаті отримуємо множину інтервалів, на яку поділені значення певної числової ознаки. Дискретизація виділяє найбільш характерні інтервали в розподілах об'єктів на шкалах числових ознак, чим забезпечує компактність логікових виразів, які

формуються системою CONFOR для класифікації об’єктів та підвищує якість розпізнавання.

Далі відбувається побудова мережі (рис. 1.), після цього виконується процес формування понять. Під поняттям будемо розуміти елемент системи знань, що представляє собою узагальнену логікову ознакову модель класу об'єктів, за допомогою якої реалізуються процеси розпізнавання.

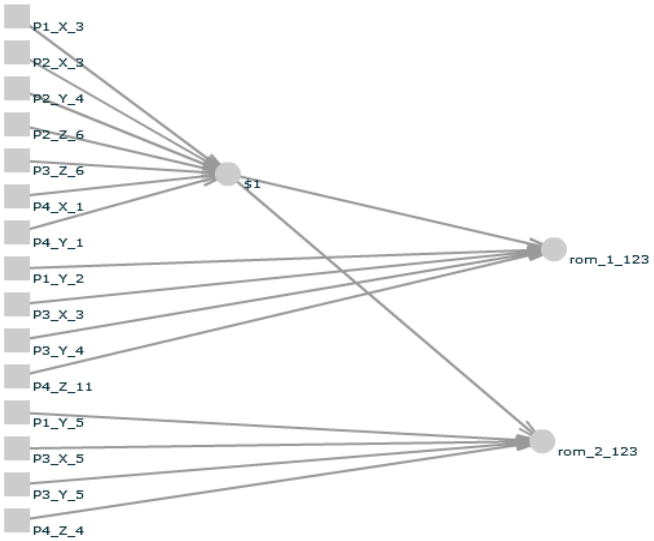


Рис. 1. Фрагмент побудованої пірамідальної мережі

Результатом роботи CONFOR'a в режимі навчання є узагальнена модель класів - пацієнтів, що включає найбільш характерні властивості цих об'єктів. В режимі розпізнавання, використовуючи експериментальні дані, було проведено два варіанти експерименту. В першому експерименті виконувалась попарна класифікація. Одного пацієнта розрізняли з-поміж електрокардіограм другого пацієнта.

Для двох пацієнтів, у яких на вході в систему кількість апроксимованих QRS-комплексів дорівнювала 100 і 100, при попарній класифікації правильно ідентифіковано 100 об’єктів і 90 об’єктів. При класифікації другої пари пацієнтів з такою ж кількістю QRS-комплексів, правильно ідентифіковано всі об’єкти, тобто 100% ідентифікація.

В другому експерименті приймали участь 5 пацієнтів. Кожному з них належало по 10 апроксимованих QRS-комплексів. За результатами класифікації, всі 5 класів - пацієнтів були розпізнано правильно.

В даний час продовжуються модельні експерименти на великій кількості даних за участю більшої кількості пацієнтів.

IV. Висновки

Використовуючи систему CONFOR, що базується на використанні зростаючих пірамідальних мереж в режимі розпізнавання можливо вирішення завдання - ідентифікація. При наявності відповідних ознак проведено серію експериментів, які показують ефективність застосування цього інструменту в якості класифікатора для електрокардіограм. В проведеному експерименті точність ідентифікації була досить високою.

Література References

[1] L.Fainzilberg, T. Potapova, “Computer Analysis and Recognition of Cognitive Phase Space Electrocardiographic Image”, Proceeding of 6th International Conference on Computer analysis of Images and Patterns, Prague, 1995, pp 668-673.

[2] L. Biel., O. Petersson, L. Philipson, P. Wide, “ECG analysis: a new approach in human identification,” IEEE Transactions on Instrumentation and Measurement, 2001, vol 3, pp. 808-812.

[3] S. Yogendra Narain, P., “Biometrics Method for Human Identification Using Electrocardiogram,” Proceedings of third International Conference, ICB. Alghero, Italy, 2009, pp. 1270-1279.

[4] V. Vishnevskey, “Telemedicine technologies and scientific research,” Ukrainian Journal of Telemedicine and Medical Telematics, 2006. - vol.4, №1, pp. 9-13.

[5] V. Vishnevskey, “Grid system for mass storage and processing of digital electrocardiograms,” Ukrainian Journal of Telemedicine and Medical Telematics, 2013, vol. 11, №. 1, pp. 202-208.

[6] V. Vishnevskey, V. Kalmykov, T. Romanenko, “Approximation of one-, two- and three-dimensional curves arcs by parametric splines,” Mathematical Machines and Systems, 2015, No. 4, pp. 160.

[7] V. Vishnevskey, T. Romanenko, L. Kizub, “Experimental Verification of Possibility of Human Identification by the Electrocardiogram,” Proceedings of 5 TH International Conference on Application of Information and Communication Technology and Statistics in Economy and Education, Sofia, Bulgaria, 2015, pp. 318.

[8] V. Gladun, “Growing pyramidal networks - the organization of memory of intellectual systems,” Artificial intelligence, 2003, No. 3, pp. 70-77.

[9] http://confor.inhost.com.ua/confor2/

[10] The Association of Developers and Users of Intelligent Systems [Online]. Available: http://www.aduis.com.ua.

[11] V. Gladun, V.Velichko, “Instrumental Complex of Decision-Making Support on the Basis of the Network Model of the Subject Object Field,” Decision Support Systems. Theory and practice, Kiev, 2012, pp. 128.

Using Gamification in Mobile Activation Learning System

V. Emets, J. Rogowski, J. Krasiukianis
Institute of Information Technology
Lodz University of Technology
Wólczańska Str. 215, Lodz, Poland
volodymyr.yemyets@p.lodz.pl, jan.rogowski@p.lodz.pl, jacek.krasiukianis@gmail.com

Використання Гейміфікації в Мобільній Системі Активізації Навчання

В. Емець, Я. Роговський, Я. Красюкянис
Інститут інформаційних технологій
Технологічний університет м. Лодзь
вул. Вульчаньська 215, м. Лодзь, Польща
volodymyr.yemyets@p.lodz.pl, jan.rogowski@p.lodz.pl, jacek.krasiukianis@gmail.com

***Abstract*—This paper describes the application of gamification in the mobile and web applications serving as a gamified learning system for use both at home and in a classroom. The use of the system can lead to the pupils activation during the whole process of learning.**

***Анотація*—У даній статті описується застосування гейміфікації в мобільних і веб-додатках, які служать в якості гейміфікованої системи навчання для використання як вдома, так і в класі. Використання системи може призвести до активізації учнів протягом усього процесу навчання.**

Keywords—gamification; pupils activation

Ключові слова—гейміфікація; активізація учнів

I. INTRODUCTION

The purpose of the paper is to present the idea behind gamification as an innovative method of motivation and to prove the viability of a teaching system that is taking advantage of gamification of learning and persuasive technologies. Studies show that game-based learning not only can improve motivation to learn [1], but also keep students' concentration at a higher level [2].

Gamification describes a broad trend of applying of game design elements: mechanics and experience in a non-game context in order to improve user digital engagement and motivation to achieve one's goals. Key terms of the definition consist of *mechanics* (badges, points, leaderboards and other elements common to many games), *experience* (playing space, a storyline etc.), *digital engagement* (interaction with digital devices as opposed to people), *motivation* (to change behaviors, develop skills, stimulate innovation), g*oals* (those of the player should be aligned with those of the organization).

Gamification attempts to increase engagement by the following means:

Accelerated feedback cycles Real-world feedback loops, such as annual appraisals, are too slow and too far apart in time. Gamification rises the frequency of feedback.

Clear goals and rules of play Real-world goals are often ill-defined and rules are applied selectively. Gamification sets clear goals and enforces clear rules of play.

A compelling narrative Most real-world activities are hardly compelling. Gamification provides a narrative to fill that hole.

Challenging yet achievable tasks Most real-world challenges are simply too complex and long-term. Gamification offers the player more lesser challenges to keep them interested.

Techniques used in gamification include:

Rewards for accomplishing desired tasks and friendly competition were the first techniques used in gamification. Rewards can be points, levels, badges, the filling of progress bars, virtual currency, a place on a leaderboard, unlocking a higher difficulty level. A simple way of encouraging competition is to make each player's progress and results visible to the others, for example by means of a leaderboard [3].

Actual gameplay elements. Another rather direct approach to gamification is to incorporate elements of actual gameplay to make virtually everything seem like a game. Such elements include: meaningful choices, informative tutorials, adaptive increases in challenge, narrative.

Big data can play a big part in a gamified system. Its main purpose is to capture the data on the behavior of users, such as employees and clients. Having this data allows an organization to adapt by offering more engaging experience, which results in better performance (in case of employees) and increased loyalty (in case of customers).

The concept of motivating people through fun and play is not recent by any means. In 1982, Thomas W. Malone published a paper on "Heuristics for designing enjoyable user interfaces" [4]. In 1999, Stepher W. Draper published "Analysing fun as a candidate software requirement" [5]. Further research on fun and play in serious scenarios was a natural follow-up, leading to coining the term *funology*, which is an idea that software interfaces can be actually fun to use, going as far as eliciting emotions and feelings in users. The noughties saw the first attempt to gamify software. The actual term *gamification* was coined in 2002 by Nick Pelling, but attracted little attention. In 2008, blogger Bret Terrill mentioned the word *gameification* used to describe the application of game mechanics to web design [6]. The term gained popularity approximately two years later, not before being simplified by dropping the 'e'. 2010 saw the foundation of Badgeville, a SaaS (Software as a Service) company oriented around gamification, a gamified to-do list application Epic Win, a book on *Games-Based Marketing* by Gabe Zichermann and a TED Talk by Jane McGonigal [5]. Finally, in 2011, the gamification become widely recognized in industry and academia alike, with the establishment of the annual Gamification Summit and the Gamification Research Network. In 2012, gamification began to lose its initial momentum. The idea of merely accumulating badges seemed to gradually get old. Gartner predicted that in two years time, 80% of gamified applications will not meet business goals mainly because of their poor design [7]. However, Barry Kirk of Bunchball defied their prediction claiming that even bad gamification is still viable [8]. The claim was incidentally backed up by Badgeville, which raised $25 million [9].

II. GAMIFICATION IN SCHOOLS

This section will focus on the incorporation of gamification into school curricula, specifically, the examples of successful deployments.

A. Quest to Learn

Quest to Learn is a public school in New York for children between six and twelve years old. A defining feature of the school is its innovative approach to education, developed by teachers cooperating with game designers of the Institude of Play and funded by the MacArthur Foundation [10]. Every trimester, each teacher teams up with a game designer and a curriculum designer. They prepare gamified curricula which are relevant in modern day, establish concrete learning goals and put special emphasis on subjects which students tend to have difficulty with [11]. Not only the students are learning at Quest to Learn. The teachers and game designers gather feedback from the courses and improve them.

Quest to Learn does not distinguish between traditional school subjects, and integrates new technologies into every course. For example, math is taught in two units: The Way Things Work, together with physics, chemistry and biology, and Codeworlds, mostly concerned with the relationship between language and code, intertwined with programming. In every unit, especially Sports for the Mind, students design games and use digital media. Through interdisciplinary courses students become researchers, historians and biologists, by means of podcasts, movies and games.

B. Gaming the Classroom

Gaming the Classroom is a blog created by Lee Sheldon, a teacher at Indiana University, and Jenna Hoffstein, a graduate. The blog describes a game design course taught by Sheldon that has been gamified to imitate an MMORPG (massively multiplayer online role-playing game). The whole course resembles a multiplayer game. Every student needs to choose a handle, an avatar (a portrait of the in-game character) and decide what kind of game they wish to design. Students then join into groups, called guilds, based on that preference and their skill, and sit together in the classroom. Taking tests and exams is referred to as fighting monsters (also known as mobs), preparing reports and analyses is classed as crafting, and research and presentations are the equivalent of completing quests.

C. Programming class at Lincoln Lutheran

Lincoln Lutheran is a middle and high school in Missouri. Lloyd Sommerer is a programming teacher who gamifies his class by introducing game mechanics to it. He first thought of experimenting with gamification in August 2013, in order to provide an external motivation for students who lacked it internally [12]. Leaderboards were the first gamification mechanics incorporated by Sommerer. The leaderboards were arranged in such a way that the home page showed the top 5 students in each category, and individual pages showed top 10 students. Everybody could also view their positions on each leaderboard on their personal page. Students gain levels by reaching certain experience points milestones.

III. SYSTEM PROJECT

Incorporation of gamification into school curricula is already happening. However, what they usually lack is computerized, standard infrastructure which would allow for much wider adoption. A system designed to aid in lecturing and motivating students. The system is supposed to be adopted by schools, and as such must meet several criteria: be simple enough to be easily deployable, provide a clear, user-friendly interface as to require as little additional staff training as possible, support several platforms to reduce deployment costs.

The system can be used in two ways: in a classroom and at home. In a classroom, the system allows to assign the tasks to students and ensure that students actually do them. At home, students use the system to access assigned tasks remotely. The system also saves teachers time by checking the homework automatically.

The system usage scenarios are as follows.

1) Student

The student is supposed to only access the mobile application and can perform the following actions: receive

tasks to solve, solve the tasks in application, submit solutions to the server, receive points for solving the tasks.

The actions are connected and executed in sequence, as presented in Figure 1.

2) Parent

From the perspective of the system, the role of a parent is mainly connected with the points students collect. Although merely accumulating them might feel somewhat rewarding, the main purpose of the points is to be exchanged for rewards. What rewards are available and how many points they cost is intentionally not covered by the system, but instead left at the sole discretion of a parent. Moreover, taking advantage of the automation and integration of the system it is very easy to distribute email notifications to parents about the progress of their children. Parent's interactions with the system are presented in Figure 2.

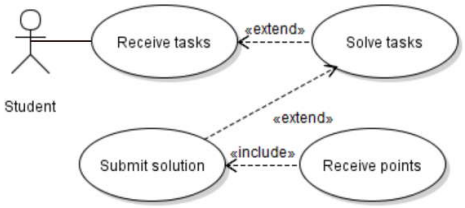


Fig. 1. Use cases of a student

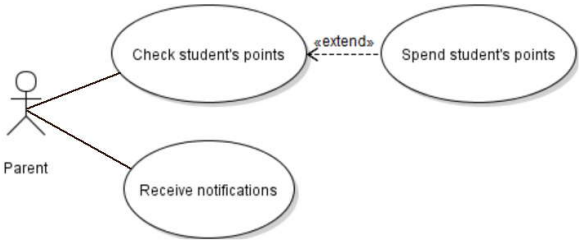


Fig. 2. Use cases of a parent

3) Teacher

Instead of presenting in real-time on a blackboard, the teachers are supposed to prepare video and text content beforehand and distribute it using the system. They also have the access to the leaderboards, which rank students who have successfully completed assigned tasks. Teacher's usage scenarios are presented in Figure 3.

The system is composed of two modules: the mobile application and the web component. The mobile application is intended to be used primarily by students, but also contains features designed for parents. A teacher is supposed to manage the course by using the web application. The applications do not communicate directly, but instead rely on the common server to pass the data between them. This ensures modularity and extensibility.

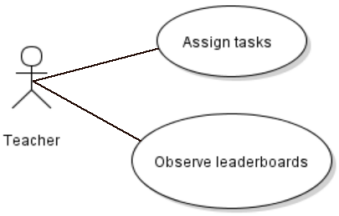


Fig. 3. Use cases of a teacher

The web component is written in Java programming language and runs on the Google App Engine platform. Google App Engine is a cloud computing solution, used for both development and hosting of web applications (platform as a service). It has been selected because of its ease of use, as well as freemium pricing model.

The web component serves two primary purposes: providing teachers with a simplified CMS (content management system), currently implemented using the standard combination of HTML5, CSS3 and JavaScript, and exposing a RESTful API, which allows clients to consume content (in form of learning material and problems to solve) and to submit solutions. REST (representational state transfer) is a set of specifications describing the interactions between web-enabled components. It has become the de facto standard architectural style for web communication, often viewed as a successor of service-oriented architecture. Data is exchanged using JSON (JavaScript Object Notation) format, due to its superiority to XML in terms of simplicity and bandwidth consumption.

Given the actual system is a prototype, no persistent storage mechanism has been implemented, with the server holding all data in memory. A possible solution to this would be to incorporate Objectify, a simple data access API for the Google App Engine, or a separate, fully fledged database (seeing that separation is a desirable trait in system design, whereas vendor lock-in is not).

The other component, the mobile application, is also written in Java and developed for the Android operating system. Android has been chosen as the target platform because of its popularity, accessibility, flexibility and openness - most of its components are free and open source software. The application supports all Android versions from 4.4.2 onwards, thus being available to virtually all Android users.

The primary purpose of the application is to present up-to-date learning material to students. The API exposed by the web component is consumed by Retrofit, a popular open source REST handling library, paired with Gson for mapping JSON submissions and responses.

The storage mechanism of the mobile application is greatly simplified. Because the learning materials are not persisted on the device, only the SharedPreferences storage is used. The Android operating system offers native support for SQLite3 databases, making them an obvious choice for data persistence. Currently, learning material consist of only text and URLs to video content, making binary file storage unnecessary. In a case when video content is to be directly fetched from the server and stored on the device, it should be saved separately to

external storage, as binary files should not be placed directly inside an SQLite database.

Because the mobile application is supposed to be used much more extensively by many more users, significantly more focus has been put on designing it. The source code is organized according to the Clean Architecture. All software dependencies are managed by Dagger 2, a dependency injection framework developed by Google as a successor to Dagger. View layer is bound to the logic layer using ButterKnife, a lightweight boilerplate code generator utilizing custom annotations. Such design offers maximum compliance with SOLID principles and as a result makes the code significantly more reliable, maintainable and extensible.

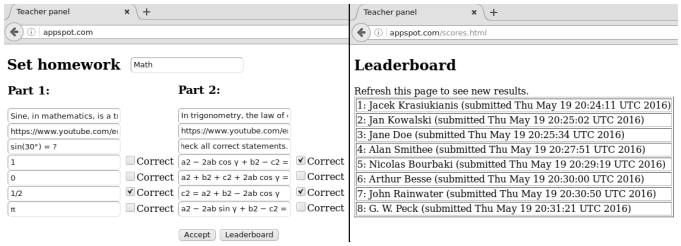


Fig. 4. Task assignment panel and Leaderboard presented in web application

For the system to be of any use to a student, that student must be registered. As such, the students are best registered before the start of a course, their system accounts linked to the school accounts from the very beginning. This also allows the mobile application to skip the registration step, making it simpler and thus easier to use.

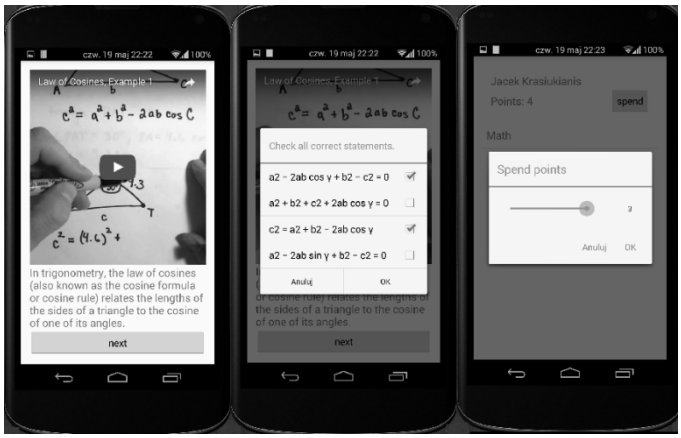


Fig. 5. Screens of the mobile application

In order to assign the tasks for students, a teacher visits a website hosted over the Internet or the school intranet and accesses the panel presented in Figure 4. The tasks are grouped in homeworks. Each task consists of: textual information, URL to video content, multiple-choice type question.

To access the application a student logs in providing a username. After that, main screen of the mobile application is presented. A student provides the name of a homework and initiates the process of solving it. Then, a separate screen is presented for each task the homework is composed of (see Figure 5). After answering the last question, the solutions provided by a student are uploaded to the server. The server verifies the answers and rates the submission. Beside the feedback, a student receives a number of points as a reward. The screen also feature the option to share the achievement on Facebook. If the submission is 100% correct, the server places it on the leaderboard.

IV. Conclusion

The system may offer several advantages. The first, most clear advantage is that a significant part of the menial work a teacher must do (checking homework, managing grades, etc.) is eliminated. The delivery of the content is also automated - a teacher needs to prepare the content once, and then, during classes, merely answer questions (which can serve as a feedback as to which parts of the content need clarification). Naturally, this workflow remains simply a suggestion and a teacher remains in complete control over the course. The system is very modular in nature and highly extensible. It can be fitted with support for any school intranet or content management system already in place, or act as a module for it.

References

[1] T.Y. Liu, Y.L. Chu, "Using ubiquitous games in an English listening and speaking course: Impact on learning outcomes and motivation", Computers & Education, vol. 55(2), pp. 630–643, 2010.

[2] K. Squire, "Video Games in Education", International Journal of Intelligent Games and Simulation, vol. 2(1), pp. 49–62, 2003.

[3] Badgeville wiki, https://badgeville.com/wiki/.

[4] T. W. Malone, "Heuristics for designing enjoyable user interfaces: Lessons from computer games", ACM Digital Library, 1982.

[5] Z. Fitz-Walter, "A brief history of gamification", http://zefcan.com/2013/01/a-brief-history-of-gamification/, 2013.

[6] B. Terrill, "My Coverage of Lobby of the Social Gaming Summit. Bret on Social Games", http://www.bretterrill.com/2008/06/my-coverage-of-lobby-of-social-gaming.html, 2008.

[7] C. Pettey, R. Meulen, "Gartner Says by 2014, 80 Percent of Current Gamified Applications Will Fail to Meet Business Objectives Primarily Due to Poor Design", http://www.gartner.com/newsroom/id/2251015, 2012.

[8] B. Kirk, "Defying the Odds of Gartner's Gamification Prediction", http://www.bunchball.com/blog/post/944/defying-odds-gartners-gamification-prediction, 2012.

[9] R. Herring, "Gamification Platform Badgeville Raises $25M", http://www.redherring.com/internet/gamification-platform-badgeville-raises-25m/, 2012.

[10] About Q2L, http://www.q2l.org/about/.

[11] B. Waniewski, "Meet The Game Designers Who Are On A Quest To Make NYC Public School More Fun", Fast Company, http://www.fastcompany.com/3003920/meet-game-designers-who-are-quest-make-nyc-public-school-more-fun, 2012.

[12] L. Sommerer, "Gamification of our Programming Class. Lincolnlutheran.org", http://www.lincolnlutheran.org/programming/.

[13] V.F. Emets, J. Rogowski, J. Krasiukianis, "Mobile Activation Learning System Using Gamification Approach", in Advances in Intelligent Systems and Computing, vol. 512, N. Shakhovska, Ed. Springer, 2017, pp. 101-114.

The Comparison of the Web Application Development Frameworks

J. Rogowski
Institute of Information Technology,
Lodz University of Technology,
Branch of the University of Lodz in Tomaszow Mazowiecki,
jan.rogowski@uni.lodz.pl

Порівняння Каркасів для Створення Веб-Додатків

Я. Роговський
Інститут інформаційних технологій,
Технологічний університет м. Лодзь,
Університет в Лодзі, Філія Томашів-Мазовецький
jan.rogowski@p.lodz.pl

***Abstract*—The main goal of this work is to evaluate the web application development frameworks in context of productivity, powerful features, security, flexibility scale, performance and size of the community. The chosen technologies are Ruby on Rails written in Ruby, Django written in Python, and Grails written in Groovy.**

***Анотація*—Основною метою даної роботи є оцінка можливостей каркасів для створення веб-додатків в контексті продуктивності, потужних функцій, безпеки, гнучкості та розміру спільноти. Обрані технології то Ruby on Rails створений в Ruby, Django створений в Python і Grails створений в Groovy.**

Keywords—web programming; frameworks comparison

Ключові слова—веб-програмування; порівняння каркасів веб-додатків

I. INTRODUCTION

The career path of basic designers of static websites is pretty straightforward. Problems really begin for so-called back-end programmers. They have to choose from many web technologies available on the market, and what is worse, there is currently no recommendation system for the best decision. What is pretty obvious nowadays, however, is that nobody creates web services without using frameworks. It would be like reinventing the wheel because each and every framework helps with the most common things we meet in everyday web application.

Usually, developers choose the most favourite frameworks by coincidence, for example, the company where they are hired, using the same technology for years, old acquaintance recommendation, a program of study focused on one of the technologies, and so on. To be excellent in a particular field requires spending many weeks of effort, and humans inherently do not like doing something which is not enjoyable or does not see promising results fairly quickly. That is why, once chosen, a specific technology most often becomes a favourite. Every framework website advertises themselves as the best, with a lot of possibilities, so the problem is how to choose a framework which we would like to work within over the next few years. What is crucial and what is worthy of our attention?

The main goal of this work is to evaluate a few of the most promising and so-called cutting-edge technology frameworks and find out which ones support the criteria in place for present-day services. The technologies chosen for appraising are: Ruby on Rails [1] written in Ruby, Django [2] writen in Python, and Grails [3] written in Groovy.

II. WEB FRAMEWORKS

Web applications realize almost the same activities, and it does not make sense to write similar code every time. That is why frameworks provide a predefined structure and all needed mechanisms. Frameworks can be seen as a set of ready pre-written code or bunch of miscellaneous libraries with one general purpose – working in the Internet based on the client - serve architecture. The difference between framework and a set of libraries is their focusing range. Libraries often solve a narrow scope of a problem, whilst frameworks give a broader span of functionalities. Framework should aid web developers in their work. Except the code, frameworks provide the design patterns and principles. Thanks to that, developers are able to implement more unified, nicer and more reusable code, which leads to better quality of application. Nowadays, so-called 'full stack' frameworks are the most popular. They form a connected software stack, that is useful for every aspect of web development. These frameworks may be considered as an extension of the programming language.

The list of benefits for using web frameworks is long, and what is more - nowadays there are various available web frameworks. Most of them are free to use and have an open

source. Anyone can use some part of web framework or even contribute to framework development, like propose improvements, review a code or find some bugs and security vulnerabilities. Service *java-source.net* presents almost 70 different open-source java web frameworks, along with short descriptions and links to them. For the purposes of this work there were selected only three web frameworks.

A. Ruby on Rails

Ruby on Rails is built in Ruby (open source, dynamic, interpreted, object-oriented programming language, created in Japan by Yukihiro Matsumoto) under MIT License. Ruby on Rails uses the Model-View-Controller architecture where models map tables in a database, views are interpreted and converted to HTML at the run-time, and controllers are the server-side component for responding to external requests from the web server. Sometimes people say that RoR is not only a framework, but also a way of thinking about web application. It can be true, because nowadays exist few great frameworks, which officially mention about inspiration of Rails. One of them is Grails. Convention over configuration is the strongest point in Ruby on Rails and along with DRY (Don't Repeat Yourself) are the key concepts of Ruby on Rails.

B. Django

Django is a framework based on Python that was written in 2003. Python is open source, dynamic, interpreted, object-oriented programming language, which implementation was started in December 1989 in Netherlands. The framework was developed to meet the needs of a fast-paced online newsroom environment. Django as well as RoR follows Model-View-Controller but with a significant difference. Django models define the data of the web application, views handle requests and make a response for them, and templates are used for rendering response. So, in the world of Django MVC does not work as an acronym, and is better to call it Model-Template-View. The view describes the data that gets presented to the user (which data is presented, not how is displayed) and is the Python callback function for an appropriate URL.

C. Grails

Grails is a web framework, for the Java platform built in Groovy. Groovy is object-oriented programming language for the Java platform, released on January 2, 2007 in England. Groovy is dynamically compiled to Java Virtual Machine bytecode. First name of Grails was 'Groovy on Rails'. In March 2006, that name was changed with regard to a request by David Heinemeier Hansson, the founder of the Ruby on Rails. This situation clearly shows the influence of RoR. However, what distinguishes Grails from the above frameworks, is a fact that it takes a set of other frameworks literally. Grails bundles Spring, Hibernate, H2, Tomcat and remove the complexity of it, and thanks to that, using them is simple and provide functionality out of the box. Grails should allow to smoothly integrate with any other library running on JVM, which provides powerful features of the enormous world of Java tools.

III. Frameworks evaluation

To achieve the stated goals, three web applications have been created in the above-specified technologies. They have the same database scheme, sample dataset and front-end interface. In addition, in order to perform the evaluation of the frameworks in an effective way, the process is split into three phases.

A comparison between frameworks is presented, showing the features, helpers, and methods in a few examples.

The results of the benchmarks for the most popular actions in web application are presented. All results of benchmarks have been counted via the application written in Java. This application uses HTTP requests to get access to the tested applications and trigger examined actions. Each test is executed many times in order to reject outlier results.

During the process of web application frameworks evaluation the following features have been evaluated: internationalization, URL mapping, Cross-Site Scripting attack defense, Cross-Site Request Forgery attack defense and SQL Injection attack defense.

Each framework provides an effective mechanism against the most popular vectors of attack, like SQL Injection, Cross Site Scripting or Cross-Site Request Forgery. These types of attacks are in the top of ten most critical web application security risks. All of the three frameworks bring the same approach. Security mechanisms against SQL Injection and XSS are turned on by default. The developer is fully conscious when he is doing an action which is possibly vulnerable. In case of CSRF, only Grails needs some extra action to do. Django and RoR do it again by default. For internationalization, Ruby on Rails came off the best, providing the most powerful features out of the box.

IV. Benchmarks

This section presents the real data of performance comparison of executing fundamental tasks such as JSON serialization, database access, and server-side template composition. All tests were done several times to reject incidental cases, on the same machine and with the identical algorithm. The comparison does not take into account server stress, multithreading, number of simultaneous requests or any other server features.

Each experiment provides specified REST API and has a stated JSON response. Due to these facts it was possible to execute all tests for each framework with the same input and expected formatted results. Another thing common to all tests is returning time in milliseconds. For the purpose of this work three almost identical web applications have been written. Models and relation between them are the part of a real financial application.

A. Persisting data in a database

This test is called simple insert. The function which is responsible for handling requests, takes the value of parameter *n* from the query, then it obtains time from a system clock, and opens transaction *n* times, creates a sample user object, saves it and closes the session. Finally, the algorithm obtains system time again, and is able to calculate passed time needed for creation of *n* user objects. The last necessary thing is to prepare specified JSON response. Experiment started from n = 10 to n

= 498 with step 2, from n = 500 to n = 1000 with step 5, where every step was repeated 3 times.

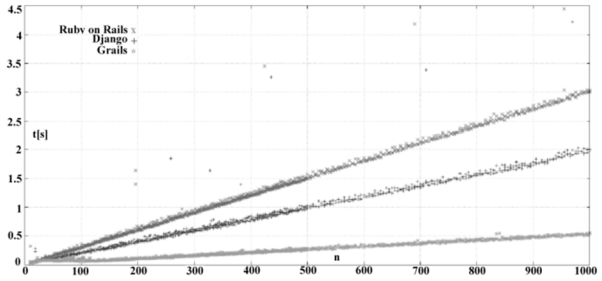


Fig. 1. Simple entity persisting results (time versus number of entities)

It is easy to notice that the dependency is linear (see Fig. 1). Ruby on Rails has the worst time needed to persist object in database. Django has a bit better result while Grails has the best score.

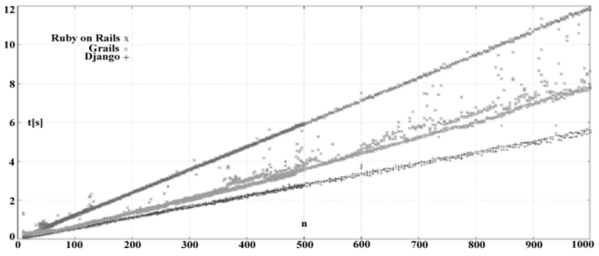


Fig. 2. Complex entities persisting results (time versus number of entities)

In the second experiment called complex insert, frameworks had to persist more complex objects, because this time the user had two kinds of relations. Each user entity has one of its project and three of its rights. This time Django had the best score, which came as a surprise, because Grails comes second in this competition. The longest time to persist connecting object was needed for Ruby on Rails. Making persisting harder did not change linear character of dependency (see Fig. 2), but just increased the time needed to insert all entities in the database.

B. Reading from a database

Reading data from a database is also the most needed task for a web application. In the beginning, the function takes the value of parameter *n* from the URL query, and prepares the data set. The method counts whether there is enough rows in the database and eventually inserts the desired number of rows. From this point, the time is 'started', the function gets the limited number of objects from the database using query set. Then the time is 'stopped', the elapsed time is calculated, and the response in JSON is prepared. This time Django has the worst times needed to read and map entities. Rails took second place but with small loss to the Grails framework (see Fig. 3).

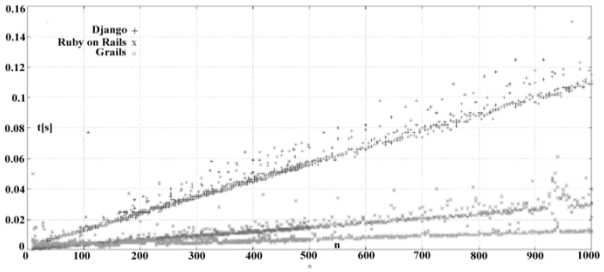


Fig. 3. Reading entities results (time versus number of entities)

C. Serializing objects to JSON

Sometimes web applications return JSON instead of HTML view. It takes place when the final client is not a user's browser but another application or service, for example, Facebook API which integrates other services with it.

In this experiment, like in the previous one, it is needed to have a given number of project entities in the database. After getting the value of parameter *n* from the query, the function optionally prepares missed rows. Next, when the objects are downloaded from the database, the time starts to be counted. The idea of this experiment is to evaluate the time needed to serialize objects to JSON.

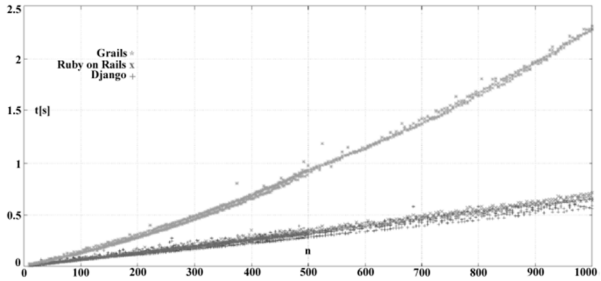


Fig. 4. Serialization results (time versus number of entities)

The result of serialization experiment, when it is needed to serialize more complicated object, for example, list of *Project* entities with two deeper levels of relations, is presented in Fig. 4. The plot indicates that this time Grails is the worst framework for serializing complex objects to JSON. The best time was achieved by Django and was only a little better than RoR.

D. Handling request

MVC frameworks function more or less in the same way. The request from the client goes to the server which forwards it to the *Dispatcher*. *Dispatcher* is a primary object which is responsible for dispatching requests to the appropriate controller. Then, the controller action could use all available framework features to do the job, and return some response. This experiment tested the average time needed to take care of client requests, and dispatching the requests to the controller.

Contrary to the other experiments, this one needs *startTime* parameters. Initially, after stepping into the controller action, a current time is obtained. After that it is possible to count the time, which passes between sending the request and stepping into the appropriate controller.

TABLE I. HANDLING REQUEST TIMES

Framework	Time
Ruby on Rails	13.53 ms
Django	1.8 ms
Grails	1.06 ms

The results are presented in Table 1. The time unit is presented in milliseconds, and this is the average value from one hundred tests. The difference between Grails and Django is minimal for the benefit of the first one. The average request handle time for Ruby on Rails is about ten times larger.

E. Sending response

Handling request experiments tested the time needed for dispatching the request to the appropriate controller. This test counts the time passed between the execution of the last instruction of the controller and the moment when the response is received by the final client. The research application has to count the passed time.

TABLE II. SENDING RESPONSE TIMES

Framework	Time
Ruby on Rails	0.54 ms
Django	1.05 ms
Grails	1.1 ms

The results of this test are presented in Table 2. Ruby on Rails is the fastest framework in this comparison in preparing response. Django and Grails have twice worse results, although one millisecond is still a very small value, almost unnoticeable.

F. Rendering view

Rendering experiment has to get a list of project entities from the database, map them to the object and pass to the template engine, which renders HTML view. The web application usually returns in response the HTML content ready to be displayed via the browser, and this test assesses the performance of the template engines in three frameworks.

First, the algorithm checks if the database has the required number of projects. If it does not, than the missing entities are inserted before counting the time. Next, the required number of projects is downloaded from the database, and after that the time starts to be counted. When the function returns HTML content, the end time is received and elapsed time is calculated. Now the function has only to prepare JSON response.

The experiment was done more than one thousand times with a different number of rendered rows. Again, the results show linear dependency (see Fig. 5). Django has the worst time, with Ruby on Rails three times and Grails almost ten times quicker.

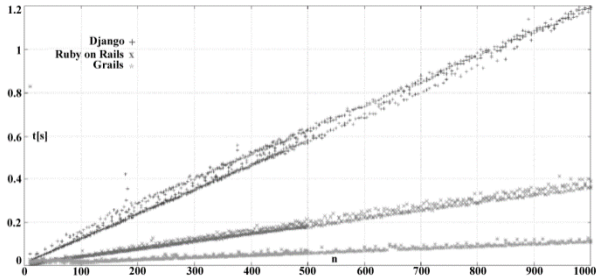


Fig. 5. Rendering template results (time versus number of entities)

V. CONCLUSION

Concluding the results of the benchmarks, the best results are obtained by Grails, which merely confirms a speed advantage of compiled language in comparison with interpreted language. Just in one test Grails scored the worst time, and it happened in serializing objects to JSON with two-level depth relations. Comparing Django and Ruby on Rails, they are competitive with each other. Hence, it would be fair to determine the tie. In the size of the community category, a job market, an amount of available books and resources, Ruby on Rails wins, while Grails has the worst score. Presented results are reflected in development process, because the least problems and the biggest help in documentation and the Internet was for RoR. All examined frameworks provide security features and support developers in writing non-vulnerable applications.

The conclusion is that the choice of the web framework is of secondary importance. In other words, any given framework is not the ultimate blocker in the development path. The most important thing in the decision process is to take into account programming language preferences, the size of the community and a job market.

REFERENCES

[1] Ruby on Rails. The website of ruby on rails framework, March 2017, URL http://rubyonrails.org/.

[2] Django. The website of django framework, March 2017, URL https://www.djangoproject.com/.

[3] Grails. The website of grails framework, March 2017, URL https://grails.org/.

[4] D. Heinemeier Hansson, S. Ruby, D. Thomas, "Agile Web Development with Rails 4", The Pragmatic Programmers, LLC, 2013.

[5] V. Kushner, O. Fernandez, K. Faustino, "The Rails 4 Way", Addison-Wesley, 2014.

[6] W. Chun, J. Forcier, P. Bissex, "Python Web Development with Django", Pearson Education Inc., 2009.

[7] B. Klein, D. Klein "Grails 2: A Quick-Start Guide", The Pragmatic Programmers, LLC, 2013.

Застосування Багатомодельного Класифікатора в Інформаційних Системах Кризового Моніторингу

А. Авраменко
кафедра інтелектуальних систем прийняття рішень
Черкаський національний університет
імені Богдана Хмельницького
Черкаси, Україна
RedStar929@gmail.com

С. Голуб
кафедра інтелектуальних систем прийняття рішень
Черкаський національний університет
імені Богдана Хмельницького
Черкаси, Україна
fpkpk@ukr.net

Application of Multimodel Classifier in the Information Systems for Crisis Monitoring

A. Avramenko
Department of intelligent decision support systems
Bogdan Khmelnitsky Cherkassy National University
Cherkassy, Ukraine
RedStar929@gmail.com

S. Golub
Department of intelligent decision support systems
Bogdan Khmelnitsky Cherkassy National University
Cherkassy, Ukraine
fpkpk@ukr.net

***Анотація*—Запропоновано покращення методу класифікації масивів вхідних даних в інформаційних системах кризового моніторингу. Зменшення часу оптимізації досягнуто шляхом використання багатомодельних класифікаторів. Класифікація проводиться за допомогою групового рішенням сукупності класифікаторів. Експериментально підтверджено ефективність покращеного методу. Час перебудови структури зменшується в 2-3 рази, в порівнянні з не модернізованим методом. Похибка моделювання значимо не погіршується.**

***Abstract*—Improvement of method for classification of the input data arrays in the crisis monitoring information systems is proposed. Reducing time of optimization achieved by using multimodel cassificators. Classification is conducted by the group decision of classifiers. Effectiveness of the improved method was experimentally confirmed. Time for restructuring of models reduced by 2-3 times compared to unimproved method. Errors of modelling is not significantly worse.**

Ключові слова—кризовий моніторинг, багаторівневі класифікатори багаторівневе моделювання, час перебудови системи, похибка моделювання.

Keywords—crisis monitoring, multimodel cassificators, multilevel modeling, group decision, system restructuring time,modelling error

I. Вступ

Використання та розробка моніторингових систем є надзвичайно важливим питанням як наукових досліджень, так і практичного застосування. Моніторингові системи дозволяють робити висновки про певні події, базуючись при цьому лише на попередні дані, отримані шляхом спостереження. Саме тому дані системи досить успішно використовують у якості систем діагностики у різних сферах життя.

Склад системи моніторингу визначається наступними чинниками: функціональне призначення, область застосування та цільові установки з переліку вирішуваних задач, функції обробки інформації, покладені на систему моніторингу і визначені користувачем.

Основним завданням моделювання в моніторингових системах є забезпечення інформацією процесу прийняття рішень. Інформація отримується в результаті моделювання властивостей об'єкта моніторингу на основі даних, отриманих в процесі вимірювання чисельних характеристик цього об'єкта.

Саме зараз такі системи використовуються і для моніторингу надзвичайних ситуацій. Проте прийняття рішень в умовах надзвичайних ситуацій накладає ряд обмежень на технології, що забезпечують ці процеси інформацією.

Розвиток надзвичайних ситуацій є ланцюговим лавиноподібним динамічним процесом. Він полягає в

різкому погіршенні стану сукупності територій, які можуть містити об'єкт чи об'єкти економіки і житлових комплексів, що призводить до катастрофічних для цих об'єктів та їх оточення наслідків.

Рішення в даних ситуаціях повинні отримуватися як найшвидше. Також незадовільна прогнозованість та динамічність надзвичайних ситуацій генерує велику кількість параметрів для моделювання, частина з яких може виникнути вперше, що означає велику можливість помилки в попередньо отриманих моделях.

На сьогодні одними з перспективних систем моніторингу є системи засновані на технології багаторівневого перетворення даних, яка реалізована у вигляді інформаційної системи з ієрархічним поєднанням багатопараметричних моделей [1].

Такі моделі можуть бути синтезовані за допомогою індуктивних алгоритмів, нейронних мереж, генетичних алгоритмів та інших. В даній технології цій системі сценарій вибору кращого алгоритму синтезу багатопараметричних моделей (АСМ) реалізовано шляхом послідовного їх випробування та вибору кращого [2].

З синтезованих моделей формується ієрархія. Моделі на кожному рівні ієрархії розв’язують локальні задачі із перетворення даних. В таких ієрархічних структурах можуть поєднуватися велика кількість моделей, від п’ятдесяти і більше (рис. 1).

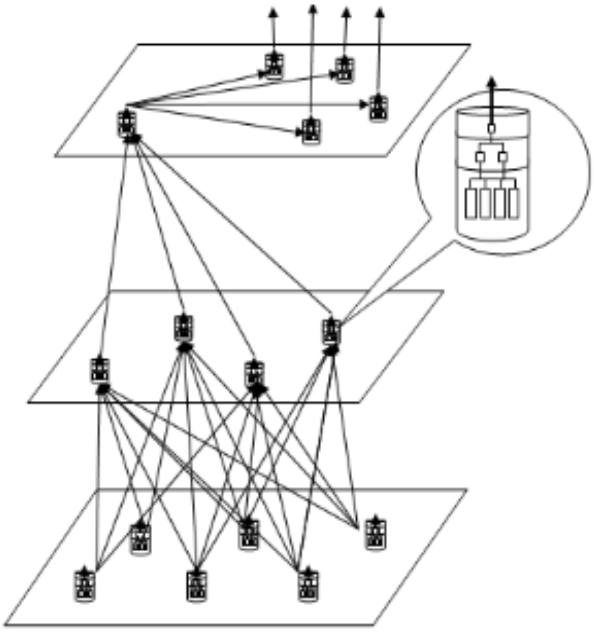


Рис. 1. Структура системи перетворення інформації

В процесі моніторингу кризових ситуацій властивості масиву вхідних даних постійно змінюються. Тому існує велика можливість, що одна або декілька моделей можуть почати видавати не адекватні результати. Для виправлення таких «пошкоджень» проводиться заміна цих моделей з їх повторним синтезом та синтезом усіх моделей, що з ними пов’язані.

Процес синтезу усіх «пошкоджених» моделей та моделей, що з ними пов’язані займає тривалий час, приблизно 40 хвилин. Тривалість залежить від кількості моделей в структурі.

Зважаючи на те, що в умовах кризового моніторингу на обґрунтування рішень виділяється не більше 2-3 хвилин, а властивості МВД змінюються динамічно. Тому є необхідність в зменшенні часу перенавчання ієрархічної системи моделей.

В попередній роботі [**3**] було запропоновано покращення існуючого алгоритму шляхом класифікації. А саме, на відміну від послідовного випробування кожного із наявних АСМ із наступним вибором кращого алгоритму, було запропоновано побудувати вирішальне правило, яке забезпечить для кожного МВД вибір найбільш придатного АСМ.

Вибір найбільш придатного АСМ проводиться на основі таких інформативних параметрів таблиць первинного опису [6]:

кількість спостережень;

кількість незалежних змінних;

кількість параметрів максимально суміщених з функцією мети;

кількість не суміщених параметрів;

середній коефіцієнт кореляції незалежних змінних;

середній коефіцієнт кореляції незалежних змінних та функції мети;

середній коефіцієнт детермінації незалежних змінних;

середній коефіцієнт детермінації незалежних змінних та функції мети;

визначник нормованої таблиці первинного опису;

визначник нормованої матриці значень незалежних змінних;

власне число нормованої таблиці первинного опису;

власне число матриці значень незалежних змінних.

максимальне сингулярне число нормованої таблиці первинного опису;

максимальне сингулярне число матриці значень незалежних змінних.

Дане дослідження ґрунтується на вивченні результатів попередньої роботи та аналізу сучасних методів, які дозволяють покращити результати роботи АСМ та навчання ієрархічних систем моніторингу.

За результатами дослідження методу запропонованого в [**3**], виявлено, що метод має певні недоліки, а саме - погіршення показників якості змодельованих моделей. Зважаючи на те, що в структурі інформаційної системи багаторівневого перетворення даних міститься від п’ятдесяти моделей і більше, вдається досягнути значного скорочення часу адаптації структури системи до зміни

властивостей МВД. Проте, у той самий час, ми не маємо можливості заздалегідь виявляти, які моделі будуть мати погіршення або покращення в порівнянні з повним перебором.

Середній результат моделювання всієї ієрархічної структури при використанні методу класифікації може навіть, у деяких випадках, бути кращим за повний перебір, а у випадку втрат точності було запропоновано компенсувати їх за допомогою збільшення складності багатошарової ієрархічної структури системи. Це досягається завдяки ефекту емерджентності властивостей багаторівневих ієрархічних систем.

Отже результати моделювання кожної моделі є не передбачуваними і можуть як поліпшитись так і погіршитись, в порівнянні з повним перебором у стандартній реалізації.

Об'єктом нашого моніторингу є кризові ситуації, це накладає ряд обмежень на отримані моделі. Зокрема, для правильної інтерпретації результатів ми повинні мати довіру до результатів роботи кожної отриманої моделі, а не тільки до кінцевого результату. А так як метод запропонований у [**3**] отримує свої результати при не передбачувані втраті точності моделей, тому з'являться необхідність його покращення з метою зменшення втрат точності.

Було розглянуто метод що дозволяє покращити результати моделювання шляхом використання особливостей формування МВД у багаторівневих ієрархічних системах [4].

Даний метод створює дублікати рівнів у ієрархії, що при особливостях утворення МВД дозволяє покращувати результати вже змодельованих моделей створивши багатошарову систему з дублікатів цих моделей.

II. Мета роботи

Метою цієї роботи є зменшення часу перебудови структури ієрархії моделей без значної втрати точності результатів моделювання на виході системи.

Для досягнення поставленої мети було запропоновано вдосконалити метод сформульований у попередній роботі [3] та усунути його недоліки.

При вирішення даної задачі було сформульовано наступні гіпотези:

Спостереження, з якими класифікатор не справився можливо використати для навчання іншого класифікатора, що буде впізнавати саме ці дані

Можливо отримати кілька класифікаторів, які виконують класифікацію одного і того самого АСМ

Набір класифікаторів що виконують класифікацію МВД для одного і того ж АСМ можливо використовувати як експертів у питанні класифікації для даного АСМ, а результати поєднувати за допомогою методів прийняття групових рішень [5]

Моделі класифікаторів можливо поліпшити використавши метод адаптивного дублювання рівнів [4]

III. Результати досліджень

При поєднанні висловлених гіпотез було сформульовано наступний алгоритм навчання класифікаторів.

Після проведення навчання класифікаторів на етапі тестування ми отримуємо новий МВД для навчання наступного класифікатора, якщо це можливо. Навчаємо його, а результати їх роботи поєднуємо за допомогою методу простого голосування [5].

Далі проводиться тестування поєднаного рішення класифікаторів та отримується новий МВД, для якого поєднаний результат не розпізнав найкращий АСМ. Всі ці дії повторюються до отримання нового МВД, за допомогою якого вже неможливо навчити наступний класифікатор. Результатом виконання даного алгоритму є отримання сукупності моделей та особливості їх групової взаємодії.

Кожна модель отримана за даним алгоритмом є рівноправною, а отже ієрархічно вони знаходяться на одному рівні. Тому ці моделі можливо подальше покращити за допомогою методу адаптивного дублювання рівнів [4].

Для проведення дослідження було модернізовано програмний інструмент створений у [3], в якому реалізовані запропоновані покращення.

Метою дослідження було:

- Перевірка доцільності запропонованих покращень
- Подальше тестування роботи методу класифікації

Для синтезу моделей даного дослідження використані результати моніторингу захворюваності населення Черкаської області впродовж 2000-2014 років [1].

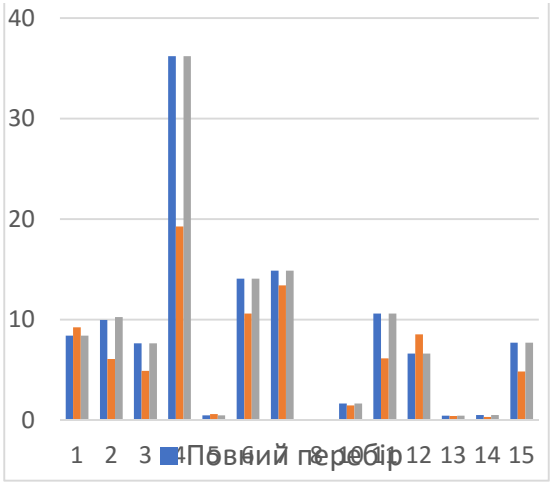


Рис. 2. Порівняння СКО

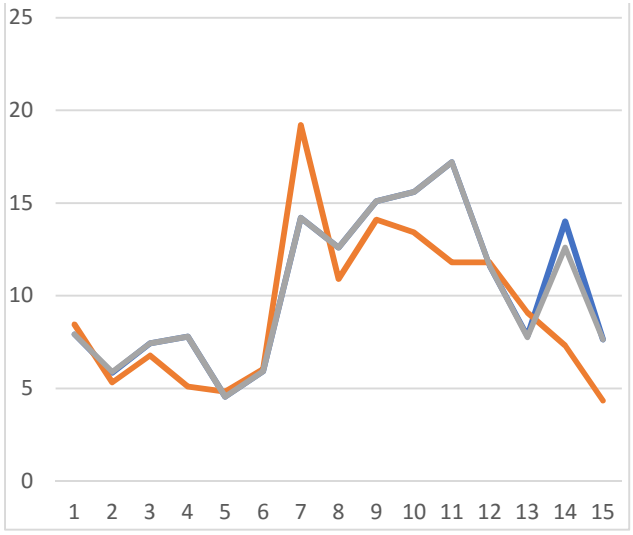


Рис. 3. Порівняння абсолютних похибок

На даних діаграмах рис.2 та рис.3 зображено відповідно порівняння стандартного та абсолютного відхилення між алгоритмом повного перебору, алгоритмом з використанням одномодельного класифікатора та запропонованим у даній роботі багатомодельним варіантом класифікатора.

В результаті можна побачити, що у деяких випадках одномодельний класифікатор має кращі результати, а у деяких гірші, ніж повний перебір АСМ. Але в середньому одномодельний класифікатор має кращий результат. Якщо порівнювати з багатомодельним варіантом, то він має результат що майже повністю сходиться з алгоритмом повного перебору.

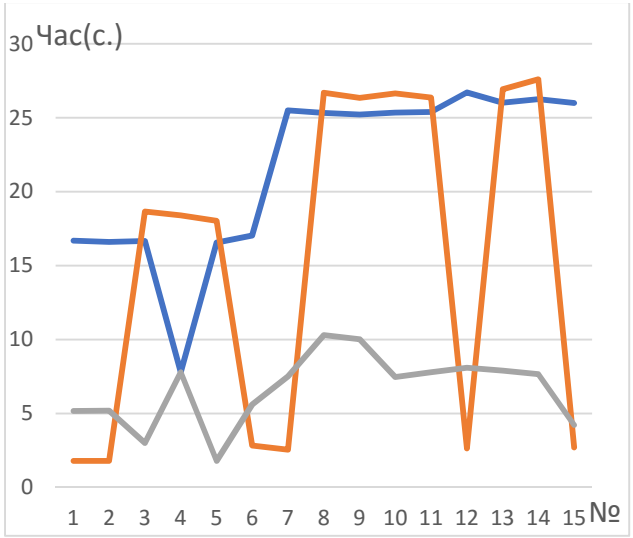


Рис. 4. Порівняння часу навчання

При порівнянні часу навчання рис.4 можливо побачити як саме нестабільно працює одномодельний класифікатор. Покращений метод в середньому працює швидше, а отже призводить до швидшого навчання системи.

IV. Висновки

В результаті порівняння роботи повного перебору, методу із [3] та його покращений багатомодельний варіант виявилося, що запропонований метод майже сходиться з повним перебором по таким критеріям якості як абсолютна та стандартна похибка, різниця між похибками не перевищує 1%. В той самий час при використанні звичайного одномодельного класифікатора різниця відхилень сягає 30-40%. Схожі результати було отримано і при порівнянні часу навчання ієрархічних систем.

В середньому час навчання системи з використанням нового покрашеного класифікатора зменшився в два рази при незначних втратах якості моделей. Зважаючи на те, що якість моделей гірша тільки в порівнянні з методом запропонованим у [3] можливо стверджувати, що поставлена мета дослідження була досягнута.

Дані результати можуть свідчити про досягнення системності при використанні багатомодельних класифікаторів.

Отже, за даними дослідження запропоновані покращення методу класифікацій алгоритмів синтезу моделей дійсно дозволяють усунути знайдені недоліки та демонструють можливість подальшого вдосконалення що потребує вивчення.

Література References

[1] Голуб С.В. Багаторівневе моделювання в технологіях моніторингу оточуючого середовища / С.В. Голуб. – Черкаси: Вид. від ЧНУ імені Богдана Хмельницького, 2007. – 218 с.

[2] Голуб С.В. Застосування стратегії оптимальності при виборі алгоритмів синтезу моделей у системах багаторівневого соціоекологічного моніторингу / С.В. Голуб, П. О. Колос // Математичні машини і системи. - 2010. - № 4. - С. 127 - 134.

[3] Avramenko A.: Classification models in information systems for social and environmental crisis monitoring / Avramenko A., Golub S. // Engineer of XXI Century – We design the future: Wydawnictwo Naukowe Akademii Techniczno-Humanistycznej w Bielsku-Bialej - Bielsku-Biala : ATH, 2016. – 928 p.

[4] Адаптивне формування дублюючих рівнів в структурі ієрархічних систем багаторівневого соціогігієнічного моніторингу / С.В. Голуб, В.Ю. Немченко // Індуктивне моделювання складних систем: Зб. наук. пр. — К.: МННЦ ІТС НАН та МОН України, 2011. — Вип. 3. — С. 41-48. — Бібліогр.: 8 назв. — укр.

[5] Барабаш Ю.Л. Коллективные статистические решения при распознавании. — М.: Радио и связь, 1983, —224 с., пл.

[6] Колос П.О. Визначення множини інформативних параметрів таблиці первинного опису об'єкта моделювання./ Вісник Черкаського університету, випуск 173. – Черкаси: Вид. ЧНУ, 2009. – С. 121-128.

Технологія Проектування Електронних Дидактичних Ресурсів

Олеся Власій
кафедра інформатики
ДВНЗ «Прикарпатський національний університет імені Василя Стефаника»
Івано-Франківськ, Україна
olesia_vlasii@comp-sc.if.ua

Ольга Дудка, Орест Гейко
кафедра інформатики
ДВНЗ «Прикарпатський національний університет імені Василя Стефаника»
Івано-Франківськ, Україна
olga_dudka@comp-sc.if.ua, ifgo69@gmail.com

Technology of Designing Electronic Teaching Resources

Olesia Vlasii
Dept. of Computer Science
Vasyl Stefanyk Precarpathian National University
Inano Frankivsk, Ukraine
olesia_vlasii@comp-sc.if.ua

Olga Dudka, Orest Geiko
Dept. of Computer Science
Vasyl Stefanyk Precarpathian National University
Inano Frankivsk, Ukraine
olga_dudka@comp-sc.if.ua, ifgo69@gmail.com

***Анотація*—В статті проведено аналіз можливостей використання електронних дидактичних ресурсів у навчальному процесі, здійснено огляд інструментальних засобів для їх створення, виділено основні етапи створення електронних дидактичних ресурсів; запропоновано використання ігрових Scratch-проектів в якості дидактичного ресурсу.**

***Abstract*—In the article the possibilities of using electronic teaching resources in the educational process are analyzed, there is given reviewing instrumental tools for creating such recources, the main stages of creating electronic teaching resources are considered; using game Scracth-project as a key electronic teaching resource is proposed.**

Ключові слова—електронний дидактичний ресурс, Scratch, навчально-ігрова програма,

Keywords—electronic didactic resource, Scratch, gaming teaching computer programm

I. Вступ

Сучасний етап розвитку освіти характеризується впровадженням інформаційно-комунікаційних технологій (ІКТ) в освітній процес з метою розвитку особистості учня/студента, інтенсифікації його навчальної діяльності, залучення до сучасних методів опрацювання інформації. Інформаційно-технічний ринок пропонує сьогодні вдосталь програмних продуктів, готових до використання в навчальних закладах.

Оскільки успіх навчального процесу може бути досягнений тільки за умови гармонічної єдності між метою заняття, методичним задумом педагога й індивідуальними потребами тих, хто навчається, то виникає необхідність в дидактичних засобах, призначених для реалізації такої єдності в конкретних педагогічних ситуаціях [1]. Доцільною в таких випадках є не лише можливість педагога самостійно обирати готові електронні ресурси навчального характеру, а й спроможність розробити необхідний дидактичний засіб з урахуванням соціально-психологічних умов навчально-виховного процесу.

Зазначимо, що американський національний технологічний стандарт для вчителів ISTE (http://www.iste.org/STANDARDS) зобов'язує вчителів володіти уміннями розробляти, проектувати і адаптувати авторські електронні ресурси навчання для того, щоб максимізувати зміст навчання, розвивати знання, уміння школярів, підтримувати індивідуальність кожного школяра, стимулювати активність у досягненні власних освітніх цілей, в управлінні навчанням і оцінюванні особистого прогресу. Тому проектування та реалізація електронних дидактичних ресурсів стає важливою складовою професійної діяльності педагога, особливо це стосується вчителів інформатики. Зазначимо, що така діяльність тривалий час не розглядалася в цьому аспекті і була прерогативою професійних програмістів та компаній-розробників.

II. Електронні дидактичні ресурси, їх суть та місце у сучасній системі засобів навчання

Впровадження в освітній процес новітніх ІКТ призвело до суттєвих змін у системі дидактичних засобів [2]. Насамперед, традиційні друковані засоби навчання (навчальні підручники, посібники, робочі зошити тощо)

поступово витісняються електронними. Змінилася не тільки форма зберігання навчальної інформації – відбулися якісні зміни у сутності та структурі засобів навчання. Сучасні електронні дидактичні ресурси вже не можуть бути зведені до друкованих видань без втрати їх функціональності. Змінилася також якість ілюстративного матеріалу, який використовується для демонстрації тих чи інших теоретичних положень. Сучасні репродукції, фотографії, графічні зображення, відеоматеріали, які зберігаються у цифровому форматі, створюють ілюзію присутності, дозволяють розглянути об'єкт у різних ракурсах і при цьому не втрачають своєї якості при багаторазовому використанні. Педагог отримав можливість самостійно створювати ілюстративний матеріал відповідно до потреб конкретного заняття. Змінилася якість засобів схематичного подання інформації – схем, таблиць, графіків.

За допомогою сучасних технологій необхідні матеріали можуть бути створені разом зі школярами в межах заняття, відредаговані та збережені для подальшого використання вдома. Безумовно, змінилася частка дидактичних засобів, які педагог самостійно створює для потреб конкретного заняття. Для підготовки необхідних дидактичних засобів можна скористатися сучасними цифровими пристроями – відео- та фотокамерами, аудіо пристроями, готовими фрагментами наявних дидактичних ресурсів, програмними засобами, які надають можливість швидко виокремити, відредагувати та скомпонувати потрібні фрагменти зображень, відеозаписів, презентацій тощо. Варто зазначити, що постійно оновлюється і вдосконалюється інструментарій для проведення експериментальних досліджень. Лабораторне устаткування, підключене до комп'ютера, надає нові можливості для здійснення реальних експериментів; стало можливим використання віртуальних лабораторій; шкільний експеримент наблизився до реального наукового дослідження. Звісно, змінилися засоби подання і відтворення інформації. Педагог отримав можливість скористатися новим інструментарієм – технічними пристроями для відтворення інформації в електронній формі (проекторами, інтерактивними дошками, документ-рідерами тощо) та програмними засобами – віртуальними дошками, картами знань тощо.

На основі проведеного дослідження можемо констатувати, що розвиток ІКТ стимулював суттєве розширення спектру дидактичних засобів, які можуть бути використані для потреб навчально-виховного процесу [3]. Завдяки появі й розвиткові інструментальних засобів особливо відчутно поповнився спектр дидактичних засобів у електронній формі – навчальних посібників, електронних курсів, мультимедійних фільмів тощо. У зв'язку з цим сучасний навчально-виховний процес не обмежується тими засобами, що розробляються спеціально для потреб шкільної освіти, отримують офіційне визнання й постачаються централізовано. На сьогодні навчання школярів відбувається в умовах надлишку дидактичних засобів, оскільки активними їх розробниками є не тільки спеціалізовані фірми та компанії-виробники програмного забезпечення, але й педагоги, які створюють власні ресурси, зорієнтовані на досягнення визначених ними конкретних цілей. Таким чином, навчально-виховний процес стає насиченим дидактичними засобами, які можна використовувати для забезпечення тієї або іншої дидактичної функції.

III. Можливості інструментальних засобів щодо створення електронних дидактичних ресурсів

З метою створення електронних дидактичних ресурсів потрібно проаналізувати можливості інструментальних засобів, за допомогою яких можна розробляти такі ресурси, та вибрати серед них ті, які найкращим чином дадуть змогу реалізувати педагогічні задуми, забезпечити функціональність складових елементів, компонентів контролю, діагностики тощо. Саме завдяки наявності потужних інструментальних засобів педагог, який може бути непрофесіоналом в галузі ІКТ і, зокрема, програмування, отримав можливість створювати електронні дидактичні ресурси, які відповідають педагогічним цілям конкретного заняття [4].

Зауважимо, що в практиці зручними є не потужні навчальні системи, які охоплюють весь курс, а невеликі посібники, призначені для використання на певному занятті. Для створення таких ресурсів педагогу потрібні інструменти інтегрованого характеру, наприклад PowerPoint, Google Presentation, Freelance Graphics, Corel Presentations, Harvard Graphics, та багато інших. Також у практиці навчання великого значення набувають електронні наочні засоби, за допомогою яких можна демонструвати реальні процеси й об'єкти, створювати й досліджувати моделі. Підготовка таких дидактичних електронних ресурсів здійснюється за допомогою інструментальних засобів для створення образних і знакових моделей об'єктів (наприклад, графічні редактори, засоби опрацювання фотознімків та відеоматеріалів, засоби створення двовимірних та тривимірних моделей, засоби створення ментальних карт тощо).

З розвитком інформаційних технологій серед дидактичних ресурсів особливого поширення набули тренувальні системи. Для самостійного створення електронних тренажерів педагог може скористатися сучасними інструментальними засобами створення інтерактивних вправ та комп'ютерних дидактичних ігор (Classtools.net, Zondle, Learningapps, Studystack).

Традиційно для контролю рівня знань та умінь учнів застосовуються програмні засоби, які здатні процедуру контролю перетворити на гру, змагання, поєднати контроль з навчанням, з наданням фіксованої й дозованої допомоги учневі (наприклад, Google Forms, On-lineTestPad, Tester, MyTestPro тощо).

Кожна з перелічених вище груп включає деякий перелік різних засобів, серед яких є популярні і менш відмі, безкоштовні й платні, прості та більш складні у використанні. Отже, кожного разу, коли педагог має намір розробити той або інших програмний ресурс конкретного дидактичного призначення, перед ним постає проблема вибору найбільш ефективного і зручного інструменту для його реалізації.

Вважаємо, що одним із таких інструментальних засобів, що дозволяють педагогу самостійно розробляти електронні дидактичні ресурси, можна розглядати і середовище візуального програмування Scratch, яке дає можливість створювати інтерактивні мультимедійні навчально-ігрові проекти [5][3]. З даним візуальним середовищем програмування учні знайомляться вже з 2-го класу, тому робота з ним їм буде знайомою та цікавою.

Середовище програмування Scratch дає можливість вільно маніпулювати з різноманітними об'єктами на екрані комп'ютера. В ході маніпуляцій користувач виконує різноманітні дії з об'єктами на екрані, що спонукає його до активних дій. Користувач може, задаючи поведінку об'єкта й обмірковуючи його реакцію, осмислювати свої кроки до здійснення мети. В той же час, користувач вчиться керувати процесом навчання: ставити собі задачу й знаходити шляхи її вирішення. За допомогою навчальної комп'ютерної гри, створеної засобами Scratch можна проводити різні експерименти, повторюючи імітацію до того часу, поки не буде досягнутий бажаний результат.

IV. Особливості проектування електронних дидактичних ресурсів

Хочемо відзначити важливу роль навчально-ігрових дидактичних ресурсів в освітньому процесі, адже ігрова діяльність – це багатоскладовий і продуктивний процес, структура якого включає інтелектуальну, пізнавальну, пошукову, проектну та інші форми прояву творчої активності. Ігри формують у дитини мотиваційну та інтелектуальну готовність використання комп'ютерних засобів для здійснення своєї діяльності. Дидактичні ресурси у формі комп'ютерної гри тренують пам'ять, логіку, координацію рухів, уміння планувати свою діяльність, знаходити інформацію, необхідну для досягнення поставленої мети. Оскільки комп'ютерні ігри стали невід'ємною складовою розвитку сучасних дітей, то постає нагальна потреба створення якісних навчально-ігрових програмних продуктів для заохочення школярів до процесу навчання та спрямованих на розвиток їх інтелектуальних і особистісних якостей.

Аналіз наукових джерел свідчить, що процес проектування дидактичних електронних ресурсів містить певні обов'язкові етапи. Такими етапами є: цілевизначення, конструювання, впровадження. Разом з тим можлива трансформація їх складових і виокремлення нових етапів [1].

Проектування електронного ресурсу розпочинається з визначення його цільового призначення. До прикладу, Scratch-проект може бути спрямований на досягнення пізнавальних цілей (ознайомитись з навчальним матеріалом, оволодіти навичками розв'язання завдань, повторити певні поняття тощо), розвиваючих (формування прийомів розумової діяльності, розвиток навичок самостійної роботи тощо), виховних (формування упевненості у власних силах, моральних якостей тощо).

Проектування електронного дидактичного ресурсу потребує ретельної попередньої аналітичної діяльності, яка зорієнтована на з'ясування традиційних складностей, що виникають у школярів при опануванні навчального матеріалу, виявлення труднощів, пов'язаних з індивідуальними особливостями учнів, з наявними прогалинами в їх знаннях і уміннях. Крім того, потрібно передбачити ті складнощі, з якими учень може зіштовхнутися в процесі роботи з програмним засобом і визначити призначення та місце компонентів системи допомоги, спрямованих на їх усунення.

При проектуванні навчально-ігрового дидактичного ресурсу потрібно чітко визначити місце й сутність ігрової компоненти, яка має бути цілком природною і сприяти досягненню дидактичної мети. Навчально-ігровий Scratch-проект може містити різні рівні ігрових компонент – від подання завдань в нестандартній формі до реалізації повноцінної ігрової ситуації на екрані комп'ютера. В проекті можуть бути присутні елементи новизни, сюрпризності, засоби заохочення, що так полюбляють діти. Продумуючи діяльність школярів в Scratch-проекті, потрібно турбуватися про те, щоб вона не була одноманітною. Це дає змогу підтримати інтерес до виконання завдань, запобігти втомі школяра. Наприклад, діяльність учня, спрямована на розглядання зображення і пошуку окремих об'єктів на ньому, змінюється на заповнення кросворду та введення літер з клавіатури тощо. У навчально-ігровому Scratch-проекті доцільно закладати особистісно-орієнтовану траєкторію просування навчання й розвитку кожного школяра, залежну від його індивідуальних особливостей. Наприклад, учень може самостійно вибирати послідовність виконання завдань у довільному порядку, розподіляти час на виконання завдань різної складності, повторно проходити етапи гри для покращення результатів і т.п.

Варіант шляхів просування школяра необхідно продумати у структурній моделі. При створенні структурної моделі дидактичного електронного ресурсу важливо передбачити певну часову рівномірність кожного варіанту просування школяра в середовищі ресурсу. Кожний учень, незалежно від особливостей його темпераменту, темпу опанування навчального матеріалу, рівня наявних умінь тощо, за відведений час повинен досягти поставленої мети на власному рівні і завершити роботу.

Важливим і трудомістким моментом проектування дидактичного електронного ресурсу є методична діяльність, яка полягає у розробці змістового наповнення ресурсу та конкретизації вмісту кожного компоненту структурної моделі. Йдеться про послідовність завдань для школярів, текстових пояснень, зміст і об'єм текстової інформації, яка буде на екрані, послідовність і характер практичних завдань, зміст контрольних завдань і запитань у кожний момент взаємодії школяра та комп'ютера, спосіб демонстрації та доцільність пояснювальних елементів. Крім того, необхідно ретельно продумувати реакцію програмного засобу на результат виконання учнем певної кількості завдань. Наприклад, швидке і правильне виконання перших кількох завдань служить сигналом для зміни рівня тренувальних завдань, підвищення їх складності. При формуванні тренувальних вправ слід також продумати спосіб схвалення успіхів школярів, відображення правильності кожного виконаного завдання

– у вигляді збільшення балів, надання словесного коментаря, реакції персонажу тощо.

При проектуванні електронного дидактичного ресурсу важливо приділити увагу створенню комфортного і гармонійного середовища, продумуванню візуального оформлення ресурсу в цілому та кожного його структурного компоненту. Особливої уваги заслуговує процес попередньої розробки дизайну електронного засобу, який дає можливість спростити і стандартизувати процес конструювання – наповнення середовища навчальними матеріалами. Необхідно продумати як естетичний вигляд середовища, в якому працюватиме школяр, так і логічність та цікавість компонент цього середовища для дитини, які повинні утворювати єдине ціле [1].

Після створення дизайну відбувається конструююча діяльність, пов'язана з підготовкою електронних матеріалів, наповненням середовища ними, а також оформленням всіх матеріалів для забезпечення його цілісності і комфортності. Конструююча діяльність заслуговує особливої уваги, оскільки в результаті повинно бути створено комфортне для школяра середовище, наповнене видовищними елементами, які привертають увагу; інструментами діяльності, призначення яких зрозуміло дитині; анімованими персонажами, які супроводжують учня, координують його діяльність, надають своєчасну допомогу тощо. Конструююча діяльність перш за все полягає у підготовці електронних матеріалів, що відбувається шляхом переведення підібраних раніше матеріалів в електронну форму, або пошуку та відбору потрібних матеріалів із бібліотек готових зображень, моделей, відеофрагментів тощо.

Наступний етап роботи з електронними матеріалами полягає в їх адаптації до створеного середовища. Зокрема, необхідно скоригувати пояснювальні матеріали, формули, зображення, вправи так, щоб вони відповідали як загальному стилю середовища, так і логіці їх викладу. Для ефективного сприйняття матеріалу важливо правильно вибрати його шрифт і колір літер,вірно розташувати основні й другорядні об’єкти, звернути увагу на зрозумілість умовних позначень. Необхідно продумати й визначити також зміст коментарів, підказок, керівництв та інших компонентів системи допомоги. Проектування електронного дидактичного ресурсу вимагає значної уваги до підготовки ілюстративних матеріалів (фотографій, рисунків, схем), які будуть розкривати зміст понять, що вивчаються, демонструвати зв’язки із раніше вивченим матеріалом і допомагатимуть у систематизації та узагальненні знань. Отже, потрібно потурбуватися про відповідність ресурсу ергономічним вимогам, принципам комфортного сприйняття інформації з екрану комп’ютера.

Авторський електронний дидактичний ресурс обов’язково повинен пройти попередню експертизу, результати якої дають змогу спланувати й здійснити удосконалення спроектованого ресурсу до його впровадження у практику навчання. Зокрема, на цьому етапі потрібно уважно проаналізувати функціональність ресурсу, працездатність кожної його компоненти, комфортність і зручність роботи з викладеним матеріалом, коректність і вчасність виклику допоміжних елементів тощо. Перш, ніж ресурс буде використовуватися в навчальному процесі, необхідно виявити і усунути помічені недоліки та, за потреби, покращити дизайн. Разом з цим, адекватне уявлення про якість спроектованого ресурсу можна одержати на основі аналізу й оцінювання результатів апробації в конкретних умовах навчального процесу. Результати апробації електронного дидактичного ресурсу дають змогу проаналізувати й оцінити якість розробки ресурсу, визначити ступінь його ефективності для вирішення поставлених педагогічних завдань, спрогнозувати шляхи його покращення, внести за такої потреби корективи. Однією із переваг дидактичних засобів в електронній формі є можливість їх багаторазового використання без втрати якості та функціональності. Авторські електронні дидактичні ресурси можна удосконалювати, покращувати, адаптувати для інших умов і використовувати у подальшому.

V. ВИСНОВКИ

Проектування електронних дидактичних ресурсів потрібно розглядати як невід’ємну складову професійної діяльності вчителя інформатики, яка дає можливість реалізувати творчі задуми вчителя з метою адаптації технологій навчання до конкретних соціально-психологічних особливостей навчально-виховного процесу. Одним зі шляхів реалізації такого підходу є проектування та реалізація навчально-ігрових Scratch-проектів як дієвих педагогічних інструментаріїв «з відкритою навчальною архітектурою».

ЛІТЕРАТУРА REFERENCES

[1] Дидактичні електронні ресурси у системі сучасних засобів навчання / Н. В. Олефіренко “*Науковий вісник Ужгородського національного університету. Серія : Педагогіка. Соціальна робота*”, 2014, Вип. 33, С. 129-133. - Режим доступу: http://nbuv.gov.ua/UJRN/Nvuuped_2014_33_39.

[2] G. Kozhuharova, D. Ivanova Didactic Models for applying ICT in education in “*Trakia Journal of Sciences”,* Vol. 13, Suvppl. 1, 2015, P. 462-467.

[3] E. Stemposz, A. Jodłowski, A. Stasiecka: Zarys metodyki wspierającej naukę projektowania systemów informacyjnych, Warszawa, Wydawnictwo PJWSTK, 2013, 560 s..

[4] О. Рибалко Флеш-технології як засіб створення комп’ютерних дидактичних ігор для дітей “*Вісник Інституту розвитку дитини. Сер.: Філософія, педагогіка, психологія*”, 2014, № 31, С. 99-103

[5] Scratch : Programming for All / Mitchel Resnick, John Maloney, Andres Monroy Hernandez, Natalie Rusk, Evelyn Eastmond, Karen Brennan, Amon Millner, Eric Rosenbaum, Jay Silver, Brian Silverman, Yasmin Kafai in “*Communications of the ACM*”, 2009, Vol. 52, No. 11. – P. 60 – 67.

Інформаційна Технологія Моніторингу Якості Загальноосвітнього Навчального Закладу

Володимир Лужецький
кафедра захисту інформації
Вінницький національний технічний університет
Вінниця, Україна
lva.zi2002@gmail.com

Богдан Білусяк
Манявський навчально-виховний комплекс
Манява, Україна
Manyava_school@meta.ua

Олег Білик
Вінницька академія неперервної освіти
Вінниця, Україна
Bilyk.oleg2012@gmail.com

Information Technology for Monitoring the Quality of a General Educational Institution

Volodymyr Luzhetskyi
department of Information Security
Vinnytsia National Technical University
Vinnytsia, Ukraine
lva.zi2002@gmail.com

Bohdan Bilusak
Maniava School
Maniava, Ukraine
Manyava_school@meta.ua

Oleg Bilyk
Vinnytsia Academy of Continuing Education
Vinnytsia, Ukraine
Bilyk.oleg2012@gmail.com

***Анотація*—Запропонована інформаційна технологія моніторингу якості загальноосвітнього навчального закладу, яка відрізняється від відомих технологій тим, що процес збирання інформації розглядається як сукупність трьох процесів: введення інформації, її вимірювання та оцінювання, а також наявністю процесів візуалізації даних і розповсюдження результатів моніторингу.**

***Abstract*—The information technology for monitoring the quality of a general educational institution is offered, which differs from known technologies in that the process of collecting information is viewed as a set of three processes: inputting information, measuring and evaluating, and the availability of process data visualization and dissemination of results of monitoring.**

Ключові слова—інформаційна технологія; моніторинг; якість освіти; кваліmивні моделі.

Keywords—information technology; monitoring; quality of education; quality model.

I. Вступ

У Національній доктрині розвитку освіти України [1] визначено нову стратегію реформування освіти, яка спрямована на забезпечення державних гарантій рівного доступу до якісної освіти на різних етапах навчання та організацію науково-аналітичного супроводу всіх управлінських рішень. Реалізацію даної стратегії можуть забезпечити лише адекватні сучасні засоби вимірювання досягнутої якості освіти, виявлення факторів впливу, оцінювання ефективності освітніх програм і реформ. Такі засоби є складниками системи моніторингу освітнього процесу та ефективного управління освітою. Організація моніторингу в сучасних умовах неможлива без використання комп'ютерних технологій. Їхнє застосування стає життєво важливою потребою для управління, оскільки прискореними темпами зростають потоки інформації і звичні паперові форми і звітність уже помітно гальмують ефективне управління. В цих умовах застосування комп'ютера в школі має еволюціонувати від простого використання для підготовки документів і

здійснення елементарних розрахунків в електронній таблиці до створення локальних баз даних і автоматизації діловодства і, далі, до створення на основі телекомунікаційних мереж автоматизованих систем інформаційного забезпечення [2,3].

II. Кваліативні моделі загальноосвітнього навчального закладу

В основу побудови моделей якості (кваліативних моделей) ЗНЗ покладено методологічні принципи TQM-ідеології [4] та кваліметрії [5,6].

Одним із основних принципів TQM-ідеології є принцип процесного підходу. Концептуальне положення процесного підходу полягає у тому, що якість процесу визначає якість результатів. Тому для забезпечення якісного результату має бути забезпечена якість ресурсів і якість самого процесу.

Другим базовим принципом TQM-ідеології є принцип системного підходу до управління. Згідно з ним контроль якості має бути впровадженим у кожен процес і будь-які помилки, що виявляються всередині процесу, мають бути відомі тим, хто бере участь у наступному процесі з метою покращення якості.

Процес функціонування навчального закладу має багато аспектів і може бути поданий у вигляді сукупності різнорідних процесів, які спрямовані на досягнення одної цілі. Водночас кожен процес може бути поданий як сукупність окремих елементарних процесів (видів діяльності).

Виходячи з цього, пропонується узагальнена процесна модель ЗНЗ, яку наведено на рис.1.

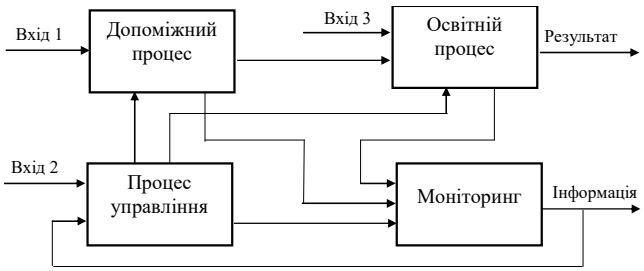


Рис. 1. Узагальнена процесна модель ЗНЗ

Тут «Вхід1», «Вхід2» і «Вхід3» відображають зовнішні ресурси, які використовуються процесами. Наприклад, «Вхід1» - це фінанси, «Вхід2» - законодавчі та нормативні акти, «Вхід3» - учні, які поступили до першого класу.

Кожен з цих процесів є складним і охоплює різні види діяльності (процеси), які й утворюють множину об‘єктів моніторингу. Наочною моделлю якості об‘єкта є дерево якостей [7] (рис.2.).

Якість загальноосвітнього навчального закладу $Q_{\text{ЗНЗ}}$ має три складові: якість ресурсів $Q_{\text{рес.}}$, якість процесів $Q_{\text{проц.}}$ і якість результатів $Q_{\text{рез.}}$, які, у свою чергу, мають певні складові.

Оскільки оцінювання якості здійснюється для потреб внутрішніх і зовнішніх споживачів, тому потрібні різні моделі якості ЗНЗ. Якість може розглядатися або як сукупність складових

$$Q_{\text{ЗНЗ}} = \left\langle Q_{\text{рес.}}; Q_{\text{проц.}}; Q_{\text{рез.}} \right\rangle , \quad (1)$$

або як функціональна залежність від них

$$Q_{\text{ЗНЗ}} = k_1 Q_{\text{рес.}} + k_2 Q_{\text{проц.}} + k_3 Q_{\text{рез.}} , \quad (2)$$

де k_1, k_2 і k_3 - вагові коефіцієнти.

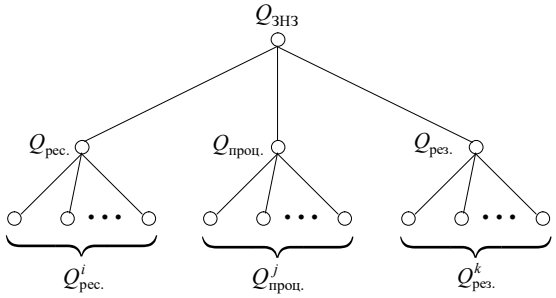


Рис. 2. Узагальнене дерево якості ЗНЗ

Представлення якості ЗНЗ у вигляді (1) надає інформацію окремо про стан ресурсів, якість процесів і рівень результатів, що забезпечує можливість спрямувати зусилля ЗНЗ на покращення конкретної складової якості.

Комплексна оцінка якості (2) спрощує порівняння ЗНЗ зовнішніми споживачами, але приховує від них складові якості.

У свою чергу, кожна із трьох складових якості ЗНЗ має свої складові, які можуть розглядатися як сукупність і як елементи функції.

III. Узагальнені моделі процесів інформаційної технології моніторингу якості

Виходячи з найуживанішого означення моніторингу [8], інформаційна технологія моніторингу якості ЗНЗ має складатися з процесів збирання, оброблення, зберігання та розповсюдження інформації про об’єкт моніторингу, а також процесів аналізу станів об’єкта і прогнозування його розвитку. Орієнтуючись на вимогу можливості розвитку системи моніторингу, що реалізує таку технологію, пропонується процес збирання інформації розглядати як сукупність трьох процесів: введення, вимірювання та оцінювання. Також доцільним є введення процесу візуалізації даних, оскільки аналіз з пошуком закономірностей часто складно реалізувати за допомогою тільки обчислювальних процедур, і тому в таких випадках аналітичні дослідження здійснюються з використанням графічних методів.

З урахуванням вище наведеного пропонується інформаційна технологія моніторингу якості ЗНЗ, схему якої зображено на рис.3.

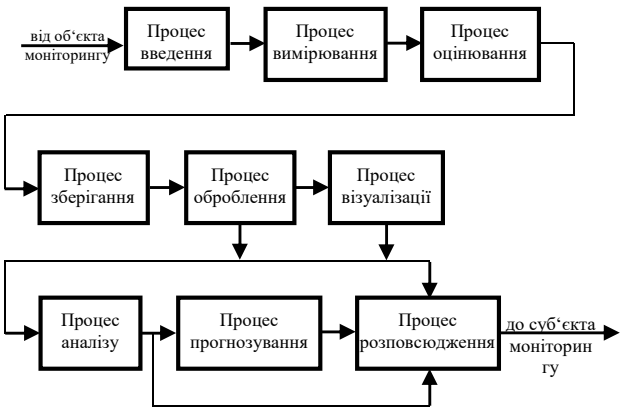


Рис. 3. Схема інформаційної технології моніторингу якості

Об'єкт моніторингу може бути представлений одним чи кількома елементами. При цьому об'єкту моніторингу відповідає об'єктна множина чи множина показників якості **Q** зі структурою відношень між ними λ_Q. Згідно з факторно-критеріальною моделлю якості кожній якості відповідає сукупність факторів **F** зі структурою відношень між ними λ_F і, в свою чергу, кожному фактору відповідає сукупність критеріїв **K** із власною структурою відношень λ_K. Таким чином, об'єкту моніторингу відповідає така модель:

$$\mathbf{O}_b = \left\langle \left\langle \mathbf{Q}, \lambda_Q \right\rangle, \left\langle \mathbf{F}, \lambda_F \right\rangle, \left\langle \mathbf{K}, \lambda_K \right\rangle \right\rangle .$$

Об'єктами моніторингу є ресурси, процеси і результати. Для кожного з цих об'єктів можна здійснювати незалежний моніторинг, однак, виходячи з концептуальної моделі якості, потрібно забезпечити моніторинг ресурсів, які беруть участь у даному процесі, моніторинг самого процесу і результатів, що одержуються на виході даного процесу.

Моніторинг відбувається за участю суб'єкта моніторингу та з урахуванням його потреб. Суб'єкт моніторингу $\mathbf{S}_b$ може бути одною особою чи їх групою. Формалізацією суб'єкта є множина осіб, що здійснюють моніторинг, зі структурою відношень між ними λ_S, тобто: $\mathbf{S}_b = \left\langle \mathbf{S}, \lambda_S \right\rangle$.

Процес введення забезпечує введення даних, що відображають атрибути об'єкта та параметри його стану. Атрибутам відповідає множина атрибутів $\mathbf{A}_o$ зі структурою відношень між ними λ_A. Стан об'єкта описується множиною параметрів **Pr** і структурою їх відношень λ_P. Введення даних може здійснюватися або вручну людиною-оператором за допомогою клавіатури, або напівавтоматично шляхом сканування паперового документа, або автоматично шляхом зчитування з електронного носія інформації. Всі ці дії вимагають використання певного набору алгоритмів $\mathbf{L}_I$. Виходячи з цього, маємо таку модель процесу введення даних:

$$\mathbf{I}_D = \left\langle \left\langle \mathbf{A}_o, \lambda_A \right\rangle, \left\langle \mathbf{Pr}, \lambda_P \right\rangle, \mathbf{L}_I \right\rangle .$$

Процес вимірювання забезпечує одержання значень критеріїв шляхом порівняння стану об'єкта моніторингу з базою порівняння **B**, яка конкретизується в вигляді системи еталонів, системи нормативів якості та ін. Алгоритм вимірювання формується з операторів порівняння і йому відповідає множина операторів порівняння $\mathbf{C}_{mp}$. Відзначимо, що порівняння може здійснюватися як людиною (наприклад, для усної відповіді), так і комп'ютером (комп'ютерний тест). Отже модель процесу вимірювання має вигляд:

$$\mathbf{M}_S = \left\langle \mathbf{K}, \mathbf{B}, \mathbf{C}_{mp}, \lambda_K \right\rangle .$$

Процес оцінювання якості об'єкта моніторингу здійснюється на основі результатів вимірювань, ієрархічної структури показників якості і кваліметричних шкал. Алгоритм оцінювання складається із множини операторів оцінювання і йому відповідає множина операторів оцінювання. Виходячи з цього, маємо таку модель процесу оцінювання:

$$\mathbf{O}_q = \left\langle \left\langle \mathbf{Q}, \lambda_Q \right\rangle, \left\langle \mathbf{F}, \lambda_F \right\rangle, \left\langle \mathbf{K}, \lambda_K \right\rangle, \left\langle \mathbf{H}, \lambda_H \right\rangle, \mathbf{\Theta} \right\rangle,$$

де **H** – множина кваліметричних шкал; λ_H – структура відношень у множині **H**; $\mathbf{\Theta}$ – множина операторів оцінювання.

Процес оцінювання якості складається з таких операцій: згортання шкал $\left(\mathbf{\Theta}_{зг.шк}\right)$; згортання критеріїв $\left(\mathbf{\Theta}_{зг.к}\right)$ за моделлю, що відповідає λ_K; згортання факторів $\left(\mathbf{\Theta}_{зг.ф}\right)$ за моделлю, що відповідає λ_F; визначення комплексної оцінки якості $\left(\mathbf{\Theta}_{KO}\right)$ за моделлю, що відповідає λ_Q.

Процес оброблення даних пропонується описувати такою моделлю:

$$\mathbf{C} = \left\langle \mathbf{D}_T, \mathbf{DI}, \mathbf{DO}, \mathbf{L}_C \right\rangle,$$

д $\mathbf{D}_T$ – множина типів даних; **DI** – множина вхідних даних; **DO** – множина результатів; $\mathbf{L}_C$ – множина алгоритмів обчислень.

Процес зберігання даних має реалізувати керування реляційною базою даних, архівування даних і відновлення даних у разі їх часткової втрати. Отже, модель цього процесу має вигляд:

$$\mathbf{M}_E = \left\langle \mathbf{D}_R, \lambda_R, \mathbf{L}_R \mathbf{L}_A, \mathbf{L}_r \right\rangle,$$

де $\mathbf{D}_R$ – дані реляційної бази даних; λ_R – структура відношень в базі даних; $\mathbf{L}_R$ – множина функцій керування базою даних; $\mathbf{L}_A$ – множина функцій архівування; $\mathbf{L}_r$ – множина функцій відновлення даних.

Форми візуалізації даних пропонується описати за допомогою множини візуальних об'єктів $\mathbf{V}_{обj}$. Відображення даних D_i у відповідний їм візуальний

об'єкт $V_{обj}^{i}$ здійснюється за алгоритмом візуалізації L_V^i. Сукупність таких алгоритмів утворює множину алгоритмів візуалізації $\mathbf{L}_V$. Виходячи з цього, пропонується така модель процесу візуалізації даних:

$$\mathbf{V} = \left\langle \mathbf{V}_{обj}, \mathbf{DO}, \mathbf{L}_V \right\rangle.$$

Процес аналізу результатів оброблення інформації може бути здійснений або автоматично, або суб'єктом, який приймає рішення. В першому випадку, комп'ютер має формулювати висновки і рекомендації за формальними правилами. У другому випадку, комп'ютер має забезпечувати підтримку прийняття рішень згідно із запитами суб'єкта. Враховуючи сказане, пропонується така модель процесу аналізу:

$$\mathbf{A} = \left\langle \mathbf{DO}, \mathbf{V}_{обj}, \mathbf{B}_C, \mathbf{Z}_A, \mathbf{L}_M, \mathbf{L}_{ZA} \right\rangle,$$

де $\mathbf{B}_C$ – множина висновків і рекомендацій; $\mathbf{Z}_A$ – множина запитів суб'єкта щодо підтримки аналізу; $\mathbf{L}_M$ – множина алгоритмів, що реалізують формальні правила; $\mathbf{L}_{ZA}$ – множина алгоритмів, які реалізують запити суб'єкта щодо аналізу.

Процес прогнозування розвитку описується такою моделлю:

$$\mathbf{Ф} = \left\langle \mathbf{Q}_Ф, \mathbf{Z}_Ф, \mathbf{L}_{ZФ} \right\rangle,$$

де $\mathbf{Q}_Ф$ – множина показників якості, що прогнозуються; $\mathbf{Z}_Ф$ – множина запитів суб'єкта щодо прогнозування розвитку; $\mathbf{L}_{ZФ}$ – множина алгоритмів, за якими реалізуються запити суб'єкта щодо прогнозування розвитку.

Результати моніторингу якості потрібні в першу чергу суб'єкту моніторингу для здійснення ефективного управління якістю об'єкта, але зацікавленими в них можуть бути й інші суб'єкти, наприклад, батьки учнів. Тому одним з процесів інформаційної технології має бути процес розповсюдження даних, який забезпечить доступ до результатів моніторингу всіх зацікавлених суб'єктів, однак з урахуванням наданих їм прав щодо отримання інформації певного виду. При цьому доступ може відбуватися як із локальної мережі, так і з Інтернету. Виходячи з цього, пропонується така модель процесу розповсюдження даних:

$$\mathbf{T}_r = \left\langle \mathbf{S}_L, \mathbf{S}_I, \mathbf{D}, \mathbf{M}_D, \mathbf{L}_D \right\rangle,$$

де $\mathbf{S}_L$ – множина користувачів локальної мережі; $\mathbf{S}_I$ – множина користувачів мережі Інтернет; $\mathbf{D}$ – результати моніторингу; $\mathbf{M}_D$ – матриця прав доступу користувачів до даних; $\mathbf{L}_D$ – множина алгоритмів, що реалізують політику безпеки.

Реалізація запропонованої інформаційної технології моніторингу ЗНЗ передбачає побудову автоматизованої системи моніторингу, яка однак не виключає участі людини в деяких процесах. Наприклад, оцінювання творчих робіт учнів чи аналіз результатів моніторингу неможливо здійснювати без участі людини, тобто повністю автоматично.

Процеси введення даних, вимірювання властивостей об'єктів моніторингу, оцінювання якості, аналізу даних моніторингу, прогнозування та розповсюдження можуть виконуватися як за участю людини, так й автоматично за допомогою комп'ютера, тоді як процеси зберігання, оброблення і візуалізації повністю реалізуються комп'ютером.

IV. Висновки

Запропонована інформаційна технологія моніторингу якості ЗНЗ відрізняється від відомих технологій тим, що процес збирання інформації розглядається як сукупність трьох процесів: введення інформації, її вимірювання та оцінювання, а також наявністю процесів візуалізації даних і розповсюдження результатів моніторингу. Відокремлення процесів введення, вимірювання та оцінювання забезпечує можливість розвитку автоматизованої системи, що реалізує цю технологію, разом із розвитком інструментарію (математичного, програмного і технічного). Доцільність введення процесу візуалізації даних, обґрунтовується тим, що аналіз з пошуком закономірностей часто складно реалізувати за допомогою тільки обчислювальних процедур, і тому в таких випадках аналітичні дослідження здійснюються з використанням графічних методів. Для розповсюдження результатів моніторингу з'явилися нові можливості завдяки значному поширенню Інтернету. Учні та їх батьки повинні мати вільний доступ до певних результатів моніторингу, знаходячись за межами ЗНЗ.

Література References

[1] Про Національну доктрину розвитку освіти / Указ Президента України від 17. 04. 2002 р. № 347 // Освіта України. – 2002. - № 33. – С.4-6.

[2] Добровольський В. Створення єдиної інформаційно-аналітичної локальної мережі у школі // Директор школи (Шкільний світ). – 2006. – №9. – С. 11-14.

[3] Лужецький В.А., Білик О.О. Підходи щодо проектування автоматизованих систем моніторингу загальноосвітніх навчальних закладів // Вісник ЧДТУ. - 2007. - № 3/4. - С. 158-168.

[4] Джордж С., Ваймерскирх А. Всеобщее управление качеством: стратегии и технологии, применяемые сегодня в самых успешных компаниях. (TQM). – М.: Виктория-плюс, 2002. - 256 с.

[5] Субетто А.И. Квалитология образования. (Основания, синтез). Изд. второе, исправлен. и дополненное. – СПб. – М.: Исследовательский центр проблем качества подготовки специалистов, 2000. – 223 с.

[6] Яковлев Е.В. Квалиметрический подход в педагогическом исследовании: Новое видение // Педагогика. - 1999. - №3. - С.49-54.

[7] Лужецький В.А., Білик О.О., Заячковський В.М. Квалітивні моделі загальноосвітнього навчального закладу // Інформаційні технології та комп'ютерна інженерія. - 2007. - №1(8). – С. 153-

[8] Майоров А.Н. Мониторинг в образовании: Изд. 3-е. – М.: Интеллект-Центр, 2005. - 424 с.

Старовікові Насадження Природного Заповідника “Ґорґани”, їх Просторова Структура та Ідентифікація Супутниковими Знімками

Анна Цуняк
Відділення землевпорядкування та дизайну
Екологічний коледж Львівського національного аграрного університету
Львів, Україна
anna.cunyak.72@mail.ru

Олег Часковський, Микола Король
Кафедра лісової таксації та лісовпорядкування
Національний лісотехнічний університет України
Львів, Україна
oleh.chaskov@gmail.com, nikkorol@ukr.net

Ольга Токар
Кафедра міжнародної інформації
Національний університет «Львівська політехніка»
Львів, Україна
tokarolya@gmail.com

Virgin and Old-Growth Forests of the Nature Reserve «Gorgany», Their Spatial Structure and Identification Using Satellite Images

Anna Cunjak
Department of Land Use and Design
Ecological College of Lviv National Agricultural University
Lviv, Ukraine
anna.cunyak.72@mail.ru

Oleh Chaskovskyy, Mykola Korol
Department of Forest Inventory and Management
Ukrainian National Forestry University
Lviv, Ukraine
oleh.chaskov@gmail.com, nikkorol@ukr.net

Olha Tokar
Department of International Information
Lviv Polytechnic National University
Lviv, Ukraine
tokarolya@gmail.com

***Анотація*—Описано застосування космічних знімків для виявлення площ старовікових насаджень. Використано сучасні підходи та новітні технології для дослідження просторової структури, що дало змогу провести адекватну оцінку їх формуванню та встановити тип розміщення дерев по площі, а також визначити основні лісівничо-таксаційні показники у розрізі ярусів. Побудовано відповідні тематичні карти для заповідника за переважаючими породами і віком.**

***Abstract*—Experience of application of satellite images for identification of old-growth and virgin forests is described. Application of modern approaches and technologies for studying spatial structure of forests that allowed assessing of the process of forming of the forests and determine the type of tree allocation and main forest inventory parameters for the forest layers. Thematic maps by prevailing tree species and age classes are developed for the nature reserve.**

Ключові слова—лісова рослинність; супутниковий знімок; растр, старовікові насадження; тематична карта.

Keywords—forest vegetation; satellite image; raster; virgin and old-growth forests; thematic map.

I. Вступ

Сучасні лісові угруповання Ґорґан формувалися протягом тривалого часу і мають різні стадії сукцесій. На значних площах зростають складні за будовою та різні за віком насадження, де їх середній вік становить понад 120 років. На території Українських Карпат поширені букові, ялинові та ялицеві праліси, які охоплюють значні площі до 35 тис. га. [1]. У Ґорґанах на висоті 700-1500 м.н.р.м. на великих площах збереглись буково-ялицеві ялинові, ялицево-ялинові та ялинові старовікові насадження та праліси. Дані насадження мають важливе значення для регіону, оскільки ефективно виконують захисні, охоронні та соціальні функції. Площа лісів Ґорґан, за літературними даними, складає близько 374,3 тис.га [2], де більшу частку тут займають хвойні ліси [3, 4]. Дослідження старовікових та пралісових формацій необхідно проводити із метою вивчення їх стабільності, особливості формування і застосування даних результатів досліджень для ефективного використання лісових ресурсів, вдосконалення підходів збалансованого, зрівноваженого ведення лісового господарства на лісових ділянках, зокрема Карпатського регіону [5].

Для отримання інформації про розміщення таких ділянок слугують у першу чергу матеріали лісоінвентаризації та супутникові знімки.

Мета дослідження – на підставі сателітних знімків встановити розміщення старовікових насаджень та вивчити методами наземної таксації їх просторову структуру на прикладі природного заповідника (ПЗ) «Ґорґани».

II. Методика дослідження

Для визначення структури насаджень можна використати дворівневі дослідження:

- на основі матеріалів дистанційного зондування Землі (ДЗЗ);
- на основі наземних досліджень.

Наземні дослідження є найточнішими та найзатратнішими. У роботі заплановані комплексні дослідження на основі високороздільних матеріалів ДЗЗ. Даний підхід дозволяє зробити аналіз структури насаджень за даними ДЗЗ на основі матеріалів структури насаджень на тестових ділянках. Відповідні дослідження дають можливість виділити подібні структури старовікових насаджень на значних площах [6]. Для забезпечення повного перекриття території дослідження було використано супутникові знімки Landsat. Під час вибору знімків керувались доступністю для користувачів, достатньою кількістю якісних знімків з мінімальною хмарністю та зроблених у незначному часовому проміжку. Знімки містять сім спектральних каналів. Крім цього для верифікації класів використано високо роздільні супутникові знімки QuickBird. Методика дешифрування космічних знімків описана у працях ряду авторів [7, 8].

Для вибору місць закладання пробних площ здійснено рівномірну растрову сітку з кроком 500 м у вертикальному та горизонтальному напрямі території ПЗ «Ґорґани» (Рис. 1). У подальшому, на підставі матеріалів лісовпорядкування за 2011 рік, відібрано ті ділянки (виділа), де середній вік насадження у них становив більше 120 років. Даний тип розміщення пробних площ (тестові ділянки) забезпечує об'єктивний характер і випадковість охоплення території, що є важливим при статистичних дослідженнях. Растрова сітка була обчислена та створена за допомогою програмного комплексу QGIS. Встановлення основних лісівничо-таксаційних показників та тип розміщення дерев проводили відповідно до методик, які описані у роботах [9, 10].

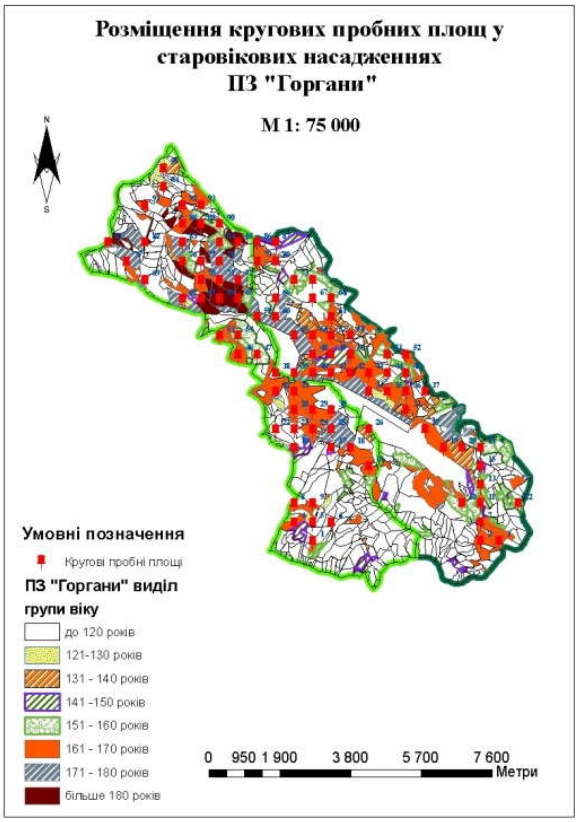


Рис. 1. Розміщення тестових ділянок на території дослідження

III. Оцінка результатів дослідження

За результатами дешифрування сателітних знімків та проведених досліджень створено карту наземного вкриття для ПЗ «Ґорґани» (Рис. 2).

Аналіз рослинного покриву за знімком вказує, що територія природного заповідника у більшості покрита хвойним лісом (coniferous forest), мішаний ліс (mixed forest) становить невелику частку, а самі вершини території вкриті кам'яними розсипами (bare rock).

Отримана інформація на підставі сателітного знімка, а саме, карта земного вкриття для ПЗ «Ґорґани», підтверджується лісовими картами лісовпорядкування, що на території переважають хвойні насадження ялини (Рис. 3).

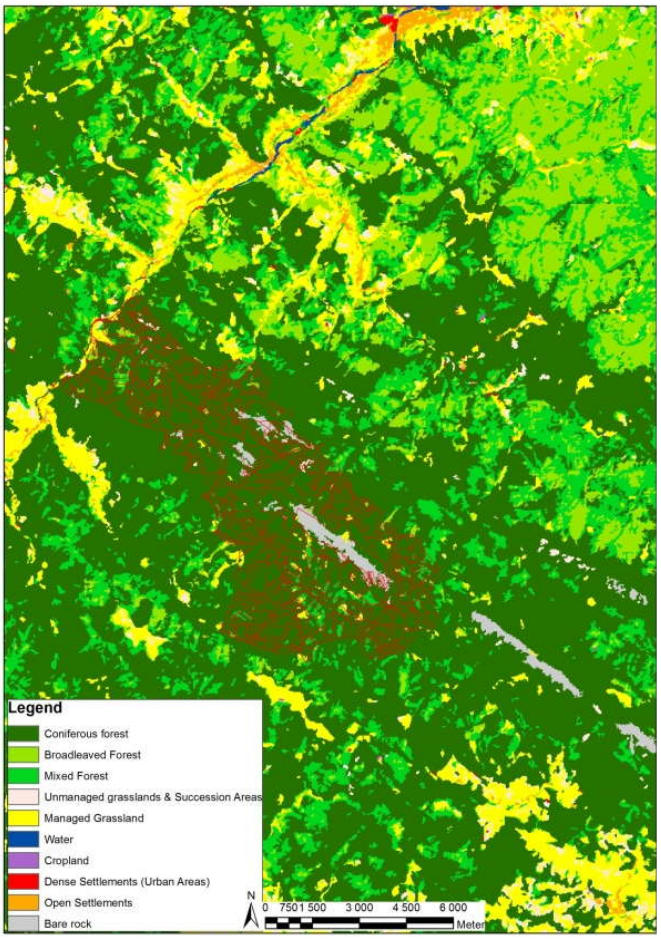


Рис. 2. Рослинний покрив ПЗ «Ґоргани» на основі зображення космічного знімка Landsat

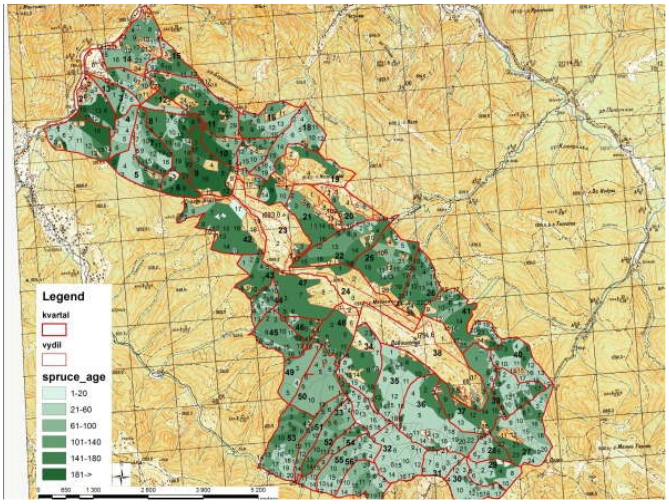


Рис. 3. Розподіл смерекових насаджень за групами віку

Площа насаджень, вік яких становить більше 120 років становить до 42 % від загальної території заповідника. Динаміка середнього діаметра залежно від висоти над рівнем моря для старовікових насаджень за основними лісотвірними породами і матеріалами пробних площ наведено у табл. 1.

ТАБЛИЦЯ 1. СЕРЕДНІЙ ДІАМЕТР НАСАДЖЕНЬ У РОЗРІЗІ ДЕРЕВНИХ ПОРІД НА ПРОБНИХ ПЛОЩАХ ЗАЛЕЖНО ВІД ВИСОТИ НАД РІВНЕМ МОРЯ

Висота над рівнем моря	Порода (середній діаметр, $D_{1.3}$, см)						
	береза повисла	*бук лісовий*	*горобина звичайна*	*клен-явір*	*смерека*	*сосна кедрова*	*ялиця біла*
805		20,8		15,1	19,9		
833		26,5	13,4		13,9		24,5
925		33,9			36,7		24,4
950					21,7		12,7
980					28,5		35,6
1060		26,6			38,0		44,6
1096					17,3		
1110			8,8		19,3		12,1
1122					40,1		
1171	31,9				23,9		
1196					34,4		
1210	35,1				26,0		23,4
……..	…..	…	…	…	….	….	……
1450					17,9	29,9	
1462					22,1		
1470					20,7		

Аналіз дослідження вказує на те, що до висоти 1100 м.н.р.м зростають змішані насадження, де присутні такі листяні породи як бук лісовий, клен-явір, береза повисла та горобина звичайна. Вище даної висоти поширені хвойні насадження – ялицево-смерекові ліси до висоти 1250 м.н.р.м. і смерекові ліси із домішкою сосни кедрової, що завершують даний вертикальний пояс.

Диференціація дерев за ступенями товщини у досліджуваних насадженнях в основному описується бімодальним розподілом (Рис. 4.) [11].

Даний тип розподілу притаманний насадженням пралісового типу [12, 13]. Більшу частку дерев становлять молоді дерева діаметром на висоті грудей до 24 см. Тут представлені всі три яруси, де розподіл за кількістю дерев становить: перший ярус – до 25 %, другий – до 35 % і третій – до 45 % (Рис. 5). Розподіл запасу у розрізі ярусів є іншим, більший запас деревини у першому ярусі до 75 %, другий ярус – 15 % і третій – до 10 %. Запас сухостійних дерев у середньому становить до 15 % і відпадають дерева в основному шляхом конкурентного впливу менших діаметрів. Будову за діаметром типового деревостану, який зростає у Ґорґанах до висоти над рівнем моря 850 м, у розрізі деревних порід ілюструє наступний рисунок (Рис. 6). Основними лісотвірними породами для даного типу насадження є бук лісовий, клен-явір, смерека та ялиця біла, де 33 % від запасу становить смерека, 30 % – ялиця біла, 18 % – бук лісовий, 3 % – клен-явір і відпад до 16 %.

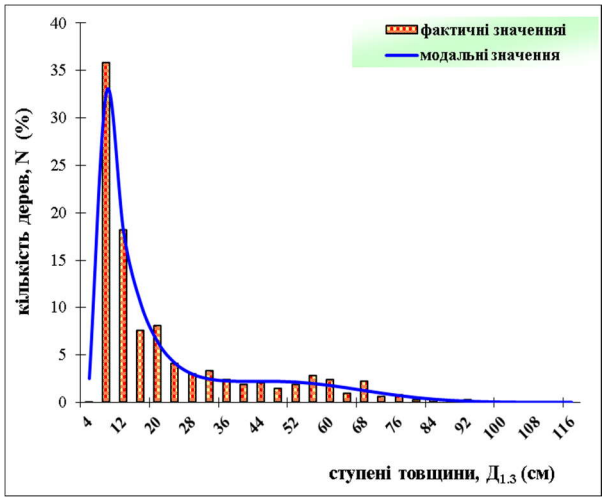


Рис. 4. Диференціація дерев за ступенями товщини смерекового деревостану Ґорган

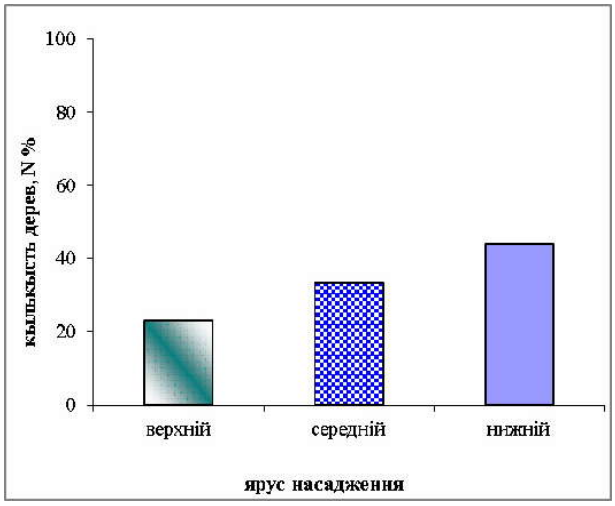


Рис. 5. Розподіл дерев за ярусами мішаного насадження Ґорган

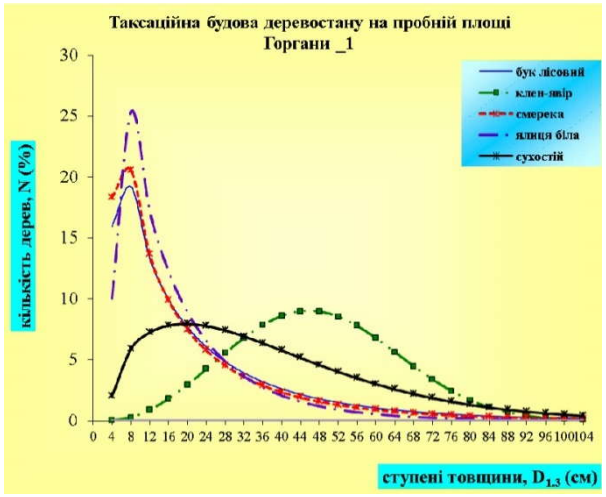


Рис. 6. Диференціація дерев за ступенями товщини мішаного деревостану Ґорган у розрізі деревних порід

IV. Висновки

У результаті виконання дослідження було застосовано на практиці нову методику ідентифікації старовікових насаджень за допомогою сателітних знімків, інформації даних лісовпорядкування та тестових ділянок. Встановлення вікової та просторової структури, а також основних лісівничо-таксаційних показників відповідних насаджень необхідно проводити шляхом інструментальних замірів на підставі растерної сітки, як регулярної, так і нерегулярної. Основою для вибору кількості дослідних ділянок повинна слугувати карта рослинного вкриття регіону дослідження.

Література References

[1] С.М. Стойко, "Характеристика пралісів Українських Карпат та їх значення для формування близького до природного лісового господарства", *Природні ліси в помірній зоні Європи – цінності та використання*: Матер. міжнарод. конф., Бірменсдорф–Рахів, 2003, С. 141.

[2] І.С. Круглов, "Делімітація, метризація та класифікація морфогенних екорегіонів Українських Карпат", *Український географічний журнал*, 2008, Вип. 3, С. 59-68.

[3] М.М. Король, "Особливості формування ялинових деревостанів у Ґорганах (Українські Карпати)", автореф. дис. канд. с.-г. наук: 06.03.02 / Микола Михайлович Король; Націон. аграрн. ун-тет, Київ, 2004, 20 с.

[4] Ju. Klimuk et al., "Nature Reserve "Gorgany", 2006, Plant World. Kyiv, Phytosociocentre.

[5] М.П. Горошко, С.І. Миклуш, О.Г. Часковський, "Дистанційні спостереження за гірськими територіями басейну Верхньої Тиси", *Наук. вісн. УкрДЛТУ: Зб. наук.-техн. пр.*, Львів: УкрДЛТУ, 2003, Вип. 13.1, С. 58-61.

[6] А.М. Цуняк, О.Г.Часковський, М.М.Король, "Розподіл наземного вкриття Стрийсько – Сянської Верховини на основі супутникових знімків LANDSAT", *Науковий вісник НЛТУ України*, 2013, Вип. 23.14, С. 27 – 32.

[7] С.І. Миклуш, С.А. Гаврилюк, "Виділення "лісової маски" для Західного Лісостепу України", *Лісівництво і агролісомеліорація*, 2006, Вип. 110, С. 60-66.

[8] С.І. Миклуш, С.А. Гаврилюк, "Інформативність каналів космічних знімків Landsat - 7 ETM+ для дешифрування рослинності", *Науковий вісник НЛТУ України*, 2006, Вип. 16.7. – С. 8-13.

[9] Ю.Й. Каганяк, М.П. Горошко, М.М. Король, О.Г. Часковський, "Інвентаризація садово-паркових об'єктів: навчальний посібник", Львів: *Камула*, 2014, 220 с.

[10] K. Streit, B. Commarmot, C. Temperli, P. Brang, "Stichprobeninventur in schweizerischen Naturwaldreservaten. Anleitung zur Methode". Birmensdorf, 2008, Eidg. Forschungsanstalt WSL. – 86 s, unveröff.

[11] T. Pukkala and K. von Gadow, "Continuous Cover Forestry", in *Managing Forest Ecosystems*, Springer Netherlands, vol. 23, 2012. DOI 10.1007/978-94-007-2202-6 2.

[12] J. Puumalainen, "Die Beta-Funktion und ihre analytische Parameterbestimmung für die Darstellung von Durchmesserverteilungen" in *Modelling Forest Development*, Georg-August-University Göttingen. Arbeitspapier 15-96, 1996.

[13] C. Westphal et. al., "Is the reverse J-shaped diameter distribution universally applicable in European virgin beech forests?" in *Forest Ecology and Management,* 2006, vol. 223 (1), pp. 75-83.

Про Квантове Моделювання Інформації в Фінансовій Математиці

Анатолій Прикарпатський
кафедра економічної кібернетики та інновацій
Державний педагогічний університет імені Івана Франка
м. Дрогобич, Україна
AGH науково-технологічний університет
Краків, Польща
pryk.anat@cybergal.com

On Quantum Information Modeling in Financial Mathematics

Anatolij Prykarpatski
Dept of Economical Cybernetics and Innovations
Ivan Franko State Pedagogical University of Drohobych, Ukraine
AGH University of Science and Technology
Krakow, Poland
pryk.anat@cybergal.com

***Анотація*—В доповіді пропонується застосування квантової інформатики до моделювання та аналізу даних фінансової математики. На основі некомутативної теорії операторів встановлена теорема про ціноутворення активів, в якій коротко говориться, що стохастичний процес $S = (S_t)$ не дозволяє арбітражні можливості, якщо і тільки якщо існує *еквівалентна міра ймовірності*, при якій $S = (S_t)$ є мартингалами.**

***Abstract*—In the report there is proposed application of quantum informatics to modeling and data analysis in financial mathematics. Based on noncommutative operator theory there is stated a theorem about option pricing, which says that stochastic process $S = (S_t)$ does not allow arbitrage if and only if there exists an equaivalent probability measure for which $S = (S_t)$ is martingale.**

Ключові слова—квантова інформатика; стохастичний процес; тензорний добуток; мартингал; арбітраж

Keywords—quantum informatics; stochastic process; tensor product; martingale; arbitrage

I. Вступ

Як добре відомо, дивлячись у ретроспективі, - наука моделювання інформації та її аналізу в фінансовій математиці зазнала значного розвитку після знакових і видатних праць F. Black і M. Scholes [2] та R. Merton [13] (нагороджених Шведською академією наук спеціальною премією з економіки імені Альфреда Нобеля, і яка, як відомо, не має жодного стосунку до так званих Нобелівських премій - правдиві Нобелевські премії надає, згідно із заповітом короля динаміту Альфреда Нобеля, - особисто Король Швеції спільно із Нобелівським комітетом, що складається із Нобелівських лауреатів, а премії з економіки, котрі мають назву «Премії з економіки імені Альфреда Нобеля», - надає Шведська Академія Наук в особі вчених-економістів, членів Академії, котрі і приймають рішення про нагородження премією, а сам преміальний фонд складається із фінансових вкладів майже виключно від крупних приватних Американських та Європейських банківських трастів), в котрих була отримана знаменита «формула оцінки опціонів Блека-Шоулза». Ідея розробки «формули» для вартості опціону насправді сягає своїми коренями ще 1900 року, коли L. Bachelier написав докторську дисертацію [1] під керівництвом H. Poincare з назвою "Theorie de la speculation". Це був власне L. Bachelier, котрий вперше висунув інноваційну ідею використовувати стохастичний процес в якості моделі для еволюції цін на акції. Для випадкового процесу $(St)_{0<t<T}$ він зробив природний і далеко йдучий вибір використати модель браунівського руху, що в даному контексті інтерпретується наступним чином: So сьогоднішня (тобто відома) ціна запасу (скажімо, частка компанії XYZ, для прикладу), а для часу $t > 0$ ціна St є нормально (за Гауссом) розподіленою випадковою величиною.

Як виявилось, праця L. Bachelier не була належно оцінена в сучасній економічній літературі. Тільки в 1965 році відомий економіст P.R. Samuelson [16] знову зайнявся темою розробки відповідної моделі для процесу оцінки акцій: для цього він запропонував геометричну версію того ж самого броунівського руху. Ця модель геометричного браунівського руху сьогодні стала стандартною, а може і еталонною моделлю для опису еволюції ціни акції, хоча зініційованою Самуельсоном, тепер її часто називають моделлю Блека-Шоулза або навіть «стохастичним світом Блека-Шоулза». Прийнято вважати, що геометрична модель браунівського руху зі знесеннями (чи зсувами, «drifts») є економічно більш адекватною від початкового вибору L. Bachelier, але на питання про те, чи є геометрична модель броунівського руху «кращою моделлю» не можна відповісти простими «так» чи «ні»: це суттєво залежить від контексту і самої мети моделювання. Основна проблема як моделі L. Bachelier, а також сучасних математичних моделей у галузі фінансів в цілому є проблема призначення ціни для умовних, так званих контингентних запитів. L. Bachelier з цією метою використовував класичну умову рівноважності або стійкості. І власне заслугою Блека та Шоулза [2] і Мертона [13] було те, щоб замінити цю умову на так званий принцип «відсутності арбітражу», який має і принципове і важливе значення для всієї теорії. Грубо кажучи, арбітражем є безризиковий спосіб отримати прибуток з нульовим вкладом чистих інвестицій. Вважається вельми економічно розсудливим припущенням щодо фінансового ринку вимога про те, чи немає можливості для арбітражу. Примітним є той факт, що цей простий і ніби примітивний "принцип відмови арбітражу" дозволяє вже визначити унікальну ціну опціону в моделі Блека-Шоулза. Це є темою так званої фундаментальної теореми ціноутворення активів, в якій коротко говориться, що процес S = (St) не дозволяє арбітражні можливості, якщо і тільки якщо існує еквівалентна міра ймовірності, при яких S = (St) є мартингалами. Історія теореми фундаментального ціноутворення активів сходить до видатної роботи Harrison, Kreps і Pliska [9, 10, 11]. Після їх новаторської праці багато авторів внесли свій вклад, поступово покращуючи розуміння цієї фундаментальної теореми, наприклад, Duffie і Huang [8], Stricker [19], Dalang, Morton і Willinger [5, 6] і т.д. В праці [7] була доведена справедливість цієї теореми для досить загальних, але комутативних стохастичних процесів, для яких пізніше вдалось довести [3] важливий не комутативний аналог цієї теореми.

Власне, у даній праці теж маємо справу із некомутативним аналогом цієї теореми на основі математичної теорії гільбертових просторів та їх тензорних добутків.

Автором подається зформалізоване визначення ***квантового арбітражу і квантових ринкових стратегій***. Встановлюються відповідні теореми ціноутворення на основі принципу відсутності арбітражу і опціонального розкладу. Зокрема, дається також характеристика повних ринкових моделей для випадку некомутативного опису фінансової моделі. Також встановлено, що для так званої біноміальної ринкової моделі із некомутативними даними процесу одноступенева модель із не комутативними параметрами має бути обов'язково неповною.

II. Квантові (Некомутативні) Ринкові Стратегії та Стохастичні Процеси

Нехай H буде певним сепарабельним гільбертовим простором із нормою ||...||, а **A** буде фільтрованою параметром $t \in R_+$ унітальною сімєю слабо *-замкнених в H алгебр операторів, тобто таких (A_t), $t \in R_+$, що $A_s \subset A_t$ для всіх $s \le t$, причому $A_0 := \mathbf{C}I$. Приймаємо також, що на **A** є задана операторна trace-норма

$$\| A \|_p := \left[\mathrm{trace}(A^* A)^p\right]^{1/p} \qquad (1)$$

для довільного оператора $A \in \mathbf{A}$ та $p \in [1,\infty)$, котра визначає банахів простір $L_p(\mathbf{A})$, ***тобто повний нормований простір операторів.*** Оскільки для будь-якого $A \in \mathbf{A}$ коректно визначена функція

$$L_p(A) \ \ni A \to \mathrm{trace}\ A \in C \qquad (2)$$

то існує єдине *умовне математичне сподівання* $E := \mathrm{trace} : L_p(\mathbf{A}) \to \mathbf{C}$ на алгебрі **A**, для якого можна визначити таку важливу характеристику як ***мартингал*** щодо вказаної вище фільтрації $A_s \subset A_t$ для всіх $s \le t$

$$E(M_t \mid A_s) = E(M_s). \qquad (3)$$

Його некомутативне узагальнення дається таким визначенням.

Визначення 1. Нехай $\rho : L_p(\mathbf{A}) \to C$ є довільним невідємно визначеним *станом* на алгебрі A, тобто для певного оператора $\hat{p} \in L_p(\mathbf{A})$ виконується нерівність $trace(\hat{p}A) \ge 0$ для всіх невідємно-визначених операторів $A \in \mathbf{A}_+$. Тоді відображення

$$R_+ \ni t \to M_t \in L_p(A)\ , \qquad (4)$$

для якого має місце властивість

$$\rho(AM_tA^*) = \rho(AM_sA^*) \qquad (5)$$

для кожного оператора $A \in A_s$ та всіх $s \le t$ називається ***некомутативним мартингалом***. Легко зауважити, що у випадку $\hat{p} = I$ вираз (5) із визначення 1 переходить у вираз (3). Тепер ми можемо перейти до побудови так званого біпроцесу як відображення

$$R_+ \ni t \to H_t \in A \otimes A, \qquad (6)$$

де позначено, як звичайно, через "$\otimes$" тензорний добуток алгебри операторів A. Вважатимемо також, що біпроцес є ***адаптованим***, тобто існує такий набір часів $0 = t_0 \le t_1 \le t_2 \le ... \le t_N$, що має місце наступна адитивна факторизація

$$H_t = \sum_{j=\overline{1,m}} A_{j,t} \otimes B_{j,t} \in A \otimes A, \qquad (7)$$

причому $A_{j,t} \otimes B_{j,t} := A_{j,t_k} \otimes B_{jt_k}$ для всіх $j = \overline{1,m}$ та $k = \overline{0, N}$, а також $A_{j,t} = 0 = B_{j,t}$ для всіх $t \ge t_N$, $j = \overline{1, m}$.

Визначення 2. Нехай $R_+ \ni t \to H_t \in A \otimes A$ є адаптованим біпроцесом. Тоді стохастичний інтеграл щодо процесу $S := (S_t)_{t \in R_+}$ задається стандартним виразом

$$\int_{t \in = R_+} H_t \circ dS_t := \sum_{j=\overline{1,N}} H_{t_k} \circ (S_{j,t_k+1} - S_{j,t_k}) = \\ = \sum_{j=\overline{1,N}} \sum_{j=\overline{1,m}} \sum_{j=\overline{1,m}} A_{j,t_k} (S_{j,t_k+1} - S_{j,t_k}) B_{j,t_k} \in A, \qquad (8)$$

незалежним від розкладу (7).

Тепер в стосунку до стохастичного процесу $S := (S_t)_{t \in R+}$ можна дати визначення **мартингального стану** $p : L_p(A) \to C$.

Визначення 3. Квантовий стан $p : L_p(A) \to C$ називається ***мартингальним***, якщо стохастичний процес $S := (S_t)_{t \in R+}$ є мартингалом на фільтрованій алгебрі $(A_t)_{t \in R+}$.

Додатково впровадимо важливе поняття ***квантових ринкових стратегій*** для стохастичного процесу $S := (S_t)_{t \in R+}$.

Визначення 4. Елементи $H := \{H : (H_t)_{t \in R+} \in A \otimes A\}$ будемо називати набором квантових ринкових стратегій для стохастичного процесу $S := (S_t)_{t \in R+}$, якщо $(H_t)_{t \in R+}$ є простим симетричним біпроцесом вигляду

$$H_t = \sum_{j=\overline{1,m}} \xi_j A_{j,t} \otimes A^*_{j,t} \in A \otimes A, \qquad (9)$$

де $\xi_j \in R, j = \overline{1, m}$.

III. Оптимальна Стратегія Безарбітражної Оцінки Активів для Квантового Стохастичного Процесу

Якщо для заданого стохастичного процесу $S := (S_t)_{t \in R+}$ розглянути опуклий конус $C^s(S;H)$ всіх самоспряжених елементів $A \in A$, що задовольняють умову $A \le \overline{H} \circ S$ для деякої ринкової стратегії $\overline{H} \in H$, то можна ввести так званий безарбітражний, або "***no-free-lunch***"-принцип

$$C^s(S;H) \cap (\cup_{t \in R+} A_t) = 0, \qquad (10)$$

сформульований у 1997 році David Wolpert та William Macready в праці *"No Free Lunch Theorems for Optimization"* [20, 14], і який полягає в тому, що коли заданий стохастичний процес передбачає оптимальну ціну акцій на заданій множині квантових ринкових стратегій, тоді цей процес є обовязково затратним на залишковій множині стратегій. Оскільки арбітражем є ***безризиковний спосіб отримати прибуток з нульовим вкладом чистих інвестицій***, то вказаний принцип «**відсутності арбітражу**» має принципове значення для всієї теорії фінансової математики.

Стосовно заданого квантового стохастичного процесу ми додатково позначимо через $M_f(S)$ множину всіх введених вище квантових мартингальних станів $\rho : L_p(A) \to C$. Якщо $M_f(S) \ne \{0\}$, то кажемо, що квантовий стохастичний процес є властивим.

Грунтуючись на введених вище поняттях можна сформулювати основне твердження про оцінювання акцій в фінансовій математиці. А саме, має місце наступна теорема.

Теорема 5. Заданий некомутативний (квантовий) самоспряжений стохастичний процес $S := (S_t)_{t \in R+}$ задовільняє принцип безарбітражності (10) тоді і лише тоді, коли він є властивим, тобто множина $M_f(S)$ квантових мартингальних станів $\rho : L_p(A) \to C$ є нетривіальною.

Нарис доведення.

Лема 6. Нехай квантова стратегія $H_t \in H$ і $\rho : Lp(A) \to C$ є квантовим станом на алгебрі A. Якщо процес $S := (S_t)_{t \in R+}$ є мартингалом щодо стану ρ, то відображення $R_+ \ni t \to (H \circ S)_t \in A$ є також мартингалом щодо стану ρ.

$\lhd$ Покладемо $H_t := A \otimes A * 1_{[t1,t2)}(t)$ де $A \in A_{t1}$ і візьмемо довільний елемент $B \in A_s$ для $s \le t$. Треба показати, що

$$\rho(B \int_s^t H_T \circ dS_T B^*) = 0. \qquad (11)$$

Оскільки $\int_s^T H_\tau \circ dS_\tau = A(S_{\min(\max(t,t_1),t_2)} - S_{\max(\min(s,t_2),t_1)})A^*$ і за визначенням $S := (S_t)_{t \in R+}$ є мартингалом, легко отримуємо справедливість рівності (11).

Для доведення теореми 5 тепер достатньо припустити, що квантовий процес $S := (S_t)_{t \in R+}$ є мартингалом щодо стану ρ. Тоді на основі Леми 6 робимо висновок, що

$$\rho((H \circ S)_\infty) = \rho((H \circ S)_0) = 0 \qquad (12)$$

для кожної квантової стратегії $H \in H$. Навпаки, якщо $s \le t$ і довільне $B \in As$, тоді для $H_\tau := B \otimes B^* 1_{[s,t)}(\tau)$ легко отримати, що $(H \circ S)_\infty = B(S_t - S_s)B^*$, звідки слідує $\rho(B(St - Ss)B^*) = 0$, що завершує доведення теореми.

Вказана теорема є апріорним критерієм вибору оптимальної квантової ринкової стратегії при заданому некомутативному стохастичнму процесі $S := (S_t)_{t \in R+}$. Так, при умові арбітражу існує така квантова ринкова стратегія, при котрій початкова нульова інвестиція приводить до кінцевої ненульової вартості акції.

Зокрема, легко показати, що для так званої біноміальної ринкової моделі (9) із некомутативними даними квантового стохастичного процесу його одноступенева модель із некомутативними параметрами має бути обов'язково неповною, тобто існування мартингального стану $\rho : L_p(A) \to C$ може не існувати.

Література References

[1] L. Bachelier, Theorie de la speculation, Ann. Sci. Ecole Norm. Sup., 17, 21-86 (1900). English translation in: The Random Character of Stock Market Prices, P.Cootner ed., MIT Press, Cambridge, Mass., 1964, pp 17-78

[2] F. Black, M.Scholes, The pricing of options and corporate liabilities, J. Political Econ., 81(1973), 637-659

[3] Z. Chen, A non-commutative version of the fundamental theorem of asset pricing, Preprint (quant-ph/0112159), 2001

[4] J.C. Cox, S.A. Ross, M. Rubinstein, Option pricing: A simplified approach, J. Finance Econ., 7(1979), 229-263

[5] R.C. Dalang, A. Morton, W. Willinger, Equivalent martingale measures and noarbitrage in stochastic securities market model, Stochastics Stoch. Rep., 29(1990), 185-201

[6] F. Delbaen, W. Schachermayer, A general version of the fundamental theorem of asset pricing, Math. Ann., 300(1994), 463-520

[7] F. Delbaen, W. Schachermayer, The fundamental theorem of asset pricing for unbounded stochastic processes, Math. Ann., 312(1998), 215-250

[8] D. Duffie, C.F. Huang, Multiperiod security markets with differential information: martingales and resolution times, J. Math. Econom., 15(1986), 283-303

[9] J.M. Harrison, D.M. Kreps, Martingales and arbitrage in multiperiod securities markets, J. Econ. Theory, 20(1979), 381-408

[10] J.M. Harrison, S.R. Pliska, Martingales and stochastic integrals in the theory of continuous trading, Stoch. Proc. Appl., 11(1981), 215-260

[11] D.M. Kreps, Arbitrage and equilibrium in economies with infinitely many commodities, J. Math. Econ., 8(1981), 15-35

[12] B.Yu. Kyshakevych, A.K. Prykarpats'kyi, and I.P. Tverdokhlib. Analysis of optimal strategies for a competing stock market portgolio model with a polyvariant profit function. Cybernetics and Systems Analysis, Vol. 47, No. 2, 210-227

[13] R.C. Merton, Theory of rational option pricing, Bell J. Econ. Manag. Sci., 4(1973), 141-183 P92 K.R.Parthasarathy, An Introduction to Quantum Stochastic Calculus, Birkhauser Verlag, Basel, 1992

[14] https://en.wikipedia.org/wiki/No_free_lunch_theorem

[15] K.R. Parthasarathy, An Introduction to Quantum Stochastic Calculus, Birkhauser Verlag, Basel, 1992

[16] P.A. Samuelson, Proof that properly anticipated prices fluctuate randomly, Industrial Manage. Review, 6(1965), 42-49

[17] W. Schachermayer, Introduction to the Mathematics of Financial Markets, (St. Flour summer school 2000), to appear in Lecture Notes in Math. Sch66 H.H.Sch?afer, Topological Vector Spaces, Springer-Verlag, Berlin, 1966

[18] C. Stricker, Arbitrage et lois de martingale, Ann. Inst. H. Poincare, Probab. Statist., 26(1990), 451-460

[19] D.H. Wolpert, W.G. Macready, "No Free Lunch Theorems for Optimization IEEE Transactions on Evolutionary Computation, 1(1997), p. 67

Гейміфікація Навчання

Віталій Горєлов
кафедра інформатики
Прикарпатський національний університет
Івано-Франківськ, Україна
vitaliy.goryelov@pnu.edu.ua

Даріуш Саля
кафедра управління
Університет AGH
Краків, Польща
dsala@zarz.agh.edu.pl

Gamification of Education

Vitaliy Goryelov
dept. of Computer Science
Precarpathian National University
Ivano-Frankivsk, Ukraine
vitaliy.goryelov@pnu.edu.ua

Dariusz Sala
dept. of Enterprise Management
AGH University
Krakow, Poland
dsala@zarz.agh.edu.pl

***Анотація*—Розглянуто гейміфікацію як сучасний процес навчання, складові процесу гейміфікації, його елементи. ДВНЗ «Прикарпатський національний університет імені Василя Стефаника» та науково-технологічний університет Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie – учасники консорціуму проекту GameHub. Показано діяльність університетів у проекті.**

***Abstract*—The gamification as a modern education process, its components and elements are analysed. Vasyl Stefanyk Precarpathian National University and AGH University of Science and Technology are GameHub Consortium members. Their project activity is illustrated.**

Ключові слова—гейміфікація; гейміфікація навчання; елементи гейміфікації; GameHub

Keywords—gamification; gamification in education; gamification elements; GameHub

I. Вступ

Підхід до набуття нових знань, при якому використовують прямий вплив на зацікавленість людини у процесові, мотивування її продовжувати процес та формування задоволення від досягнутого, дозволяє суттєво підвищити ефективність, у тому числі швидкість засвоєння та якість набутих знань.

В іграх люди виступають акторами, наділеними новими якостями, часом недоступними у реальному житті. Такі якості можуть бути рафінованими, що надає можливість формувати модель поведінки гравця із рисами, серед яких відсутні небажані та наявні бажані.

Таким чином, гейміфікація навчання полягає у створенні елементів розваги, котрі знаходять у іграх, і перенесенні їх у реальне життя при формуванні знань та навичок, необхідних у виробничій діяльності (Yu-kai Chou). Ray Wang (CEO & Principal Analyst у Constellation Research, Inc) визначає гейміфікацію як набір принципів конструювання, процеси та системи, котрі використовують для зацікавлення, мотивування та впливу на осіб, груп та спільнот для управлення їхньою поведінкою та досягнення бажаного результату.

II. Застосування гейміфікації у навчанні

Перетворення процесу навчання у гру можна розділити на наступні елементи.

Рис. 1. Складові процесу гейміфікації навчання.

Ключем до ефективного засвоєння знань є розуміння цільової аудиторії.

Важливими можуть бути вікова категорія, соціальний статус, попередній досвід, схильність до навчання тощо.

Така інформація не є достатньою без умов, за яких повинна відбутися навчальна гра. Умови можуть як сприяти засвоєнню нових знань, так і слугувати фактором послаблення концентрації гравців та зменшення ефективність гри.

Умовами можуть бути тривалість гри, місце проведення, групова гра чи індивідуальна, кількість учасників у випадкові групової гри тощо.

Розуміння цільової аудиторії та умов проведення гри повинні приводити до визначення інструктором (особою, що проводить тренінг) підводних каменів, що працюють на зменшення позитивного ефекту від навчальної гри. Такими факторами, наприклад, можуть бути неможливість суміщати роботу в офісі та навчання, відсутність мотивації тощо.

Наступним кроком є визначення мети та завдань навчання.

Завданням інструктора є здобуття слухачами (учасниками гри) певних досягнень. Такими можуть бути складання іспиту, написання підсумкової праці, розробка проекту тощо.

Конкретними завданнями навчання, зазвичай, є розуміння слухачами засад навчальної дисципліни, вміння виконувати певні завдання чи виконання навчальної програми.

Важливим є зрозуміле формулювання мети та завдань навчання.

Знайомство із цільовою групою, умовами проведення навчання та формулювання мети та завдань навчання дозволяють перейти до оцінки ресурсів.

Успішна реалізація ігрового проекту вимагає наявності наступних елементів.

Механізм трекінгу – застосовують для моніторингу перебігу навчання.

У процесі гри може існувати еквівалент грошової одиниці, здатний збільшуватися чи зменшуватися залежно від дій гравця (успішні або помилкові дії) або незалежно від нього (плин часу).

Певна кількість еквіваленту грошової одиниці, отриманої гравцем за вірні дії (вивчення навчального матеріалу) робить можливим перехід на новий рівень гри. Цікавість до отримання нових вражень (ігрового досвіду) мотивує до вивчення навчального матеріалу. Таким чином, поділ гри на рівні є одним із мотиваційних факторів у ігровому навчанні.

Ігри, зазвичай, мають правила. Правила у ігровому навчанні надають гравцям рівних початкових можливостей, що слугує додатковим мотиваційним рушієм.

Виконання завдань гри із дотриманням правил гарантують вивчення навчального матеріалу.

При цьому гравець отримує можливість переходу на новий рівень гри (перехід до нової теми тощо).

Зворотній зв’язок надає можливість інструкторові чи викладачеві отримувати інформацію про перебіг процесу навчання. Студенти отримують звіти про проходження ними тестів, що сприяє ефективному засвоєнню матеріалу.

Оскільки гейміфікація – перенесення ігрової механіки у реальне життя, слід визначити, які елементи ігор будуть застосовані у процесі навчання. Такі елементи можна поділити на притаманні гравцеві, та – соціальні.

Індивідуальні – це бали, віртуальні товари, часові обмеження тощо. Соціальні – дошки лідерів, інтерактивна взаємодія з іншими гравцями і т. п.

Використання певного типу елемента передбачає певну реакцію аудиторії, тому підбір ігрових елементів повинен бути осмисленим та збалансованим.

III. Приклад ігрового підходу до навчання

Розглянемо реалізацію ігрового підходу до вивчення дисципліни «Програмування», мова – Java. Тема: ООП. Нехай існують класи A та B, а також інтерфейси C та D. Мета: створення класу X з певними властивостями та полями. Завдання гравця полягає у виборі потрібних батьківських класів та імплементації необхідних інтерфейсів (рис. 2).

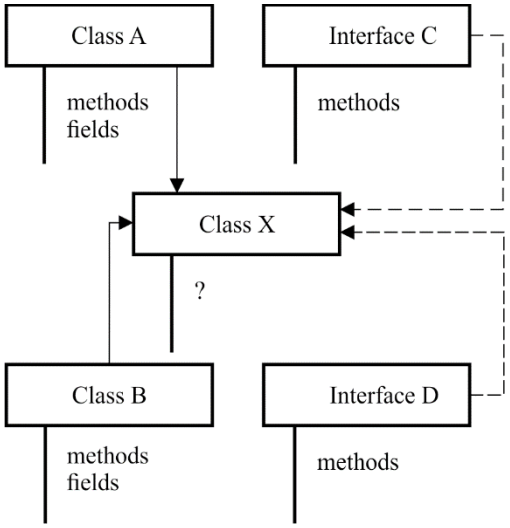


Рис. 2. Приклад гейміфікації вивчення ООП.

Навчальний матеріал реалізовано у вигляді гри, де ігровий персонаж – новостворений клас – може переміщуватися у межах ігрового поля, на якому розташовані можливі батьківські класи та інтерфейси, котрі можуть бути імплементовані.

Дотик до класів A/B та інтерфейсів C/D означає успадкування чи, відповідно, імплементацію. При цьому клас X набуває нових властивостей.

Гравець за певний час повинен отримати клас X із заданими властивостями. Збільшення кількості класів та інтерфейсів, можливість імплементації ряду інтерфейсів, також реалізація каскаду успадкувань роблять процес не очевидним та вимагають від студента певних знань. Разом з тим, візуалізація успадкувань та імплементацій

інтерфейсів робить такі процеси наочними та сприяють практичному засвоєнню матеріалу, що дає гравцеві можливість перейти від невдалих комбінацій до виконання завдання.

IV. Проект GameHub

ДВНЗ «Прикарпатський національний університет імені Василя Стефаника» та науково-технологічний університет Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie залучені до реалізації міжнародного грантового проекту Erasmus+ KA2 CBHE №561728-EPP-1-2015-1-ES-EPPKA2-CBHE-JP «Співробітництво між університетами та підприємствами в сфері ігрової індустрії в Україні — GameHub», який полягає у забезпеченні студентів, безробітних, тимчасово переселених та ветеранів антитерористичної операції знаннями та навичками, необхідними для роботи в цифровій ігровій галузі, що підвищить їх можливості працевлаштування та самореалізацію [1 – 3].

Університети увійшли до консорціуму GameHub.

Учасниками консорціуму також є: University of Deusto, Іспанія (координатор проекту); FH Joanneum Gesellschaft M.B.H., Австрія; Quality Austria – Trainings, Zertifizierungs und Begutachtungs GmbH, Австрія; Fundación Virtualware Labs, Іспанія; Донецький національний технічний університет, Україна; Херсонський національний технічний університет, Україна; Київський національний університет будівництва і архітектури, Україна; Національний технічний університет "Харківський політехнічний інститут", Україна; Одеський національний політехнічний інститут, Україна; Українська асоціація фахівців інформаційних технологій, Україна.

Очікувані результати проекту: формування профілів компетенцій співробітників на всіх професійних рівнях ринку цифрових ігор; навчання 180 викладачів вищих навчальних закладів, 500 студентів і 150 безробітних (у тому числі ветеранів АТО) для підвищення рівня їх кваліфікації та компетентності; структура і засоби GameHub в кожному українському університеті-партнері; новостворені ігрові лабораторії та 18 навчальних модулів, що забезпечують відкриті і інноваційні навчальні ресурси для нових предметів, які будуть модернізувати українську інженерну освіту; Міжнародна Game Design Конференція, організована консорціумом в Україні; підвищення актуальності навчальних програм університетів на українському ринку праці.

Цілі та завдання функціонування GameHub у Прикарпатському університеті: розробка і створення нових курсів, навчальних матеріалів для програм підготовки IT-спеціалістів ігрової галузі; впровадження розроблених курсів у навчальний процес при підготовці фахівців на рівні бакалаврів та магістрів зі спеціальності «Інформатика» та «Комп'ютерні науки»; розробка і впровадження навчальних курсів для ветеранів АТО та безробітних; побудова міцної співпраці між університетами та підприємствами ігрової галузі через надання консультаційних послуг; створення Game-лабораторії для практичної реалізації навчальних проектів і завдань студентами університету та слухачами курсів.

Пріоритетні напрямки діяльності: підготовка бакалаврів та магістрів спеціальностей «Інформатика» та «Комп'ютерні науки» для роботи в ігровій галузі; підготовка та перепідготовка ветеранів АТО та безробітних для професійної діяльності у сфері розробки ігор; розробка ігор у рамках міні-проектів для вивчення STEM школярами.

Цільова аудиторія: викладачі університету; студенти; ветерани АТО; безробітні; працівники центрів зайнятості; представники IT-компаній.

Критерії відбору до цільових груп: зацікавленість, мотивація, попередні досягнення, володіння теоретичною базою та практичні навички.

Методи відбору до цільових груп: тестування, співбесіди.

Форми навчання: лекції, практичні, семінарські, лабораторні заняття, самостійна робота студентів; навчальні конференції; консультації; індивідуальні заняття; навчальна виробнича практика; курсові, дипломні, магістерські роботи.

Форми контролю, оцінювання та обліку знань, умінь і навичок студентів: колоквіуми, заліки, іспити, захист курсових, дипломних і магістерських робіт.

Форми організації науково-дослідної роботи студентів: науково-дослідні гуртки, проблемні групи, об'єднання.

У рамках реалізації проекту GameHub будуть розроблені та впроваджені у навчальний процес наступні модулі: Computer game development (Cocos 2DX engine); Advanced computer game development. (Cocos 2DX engine); Introduction to computer game audio engineering.

Вимоги до викладацького складу: педагогічна складова проекту забезпечується наявністю висококваліфікованих фахівців, які мають практичний досвід діяльності, пов'язаної із заданою тематикою, та володіють на достатньому рівні інформаційними технологіями, необхідними для його реалізації.

Викладачі дисциплін, дотичних до тематики проекту, повинні володіти мовами C++/Java/Objective C, вміти використовувати системи контролю версій, вміти формувати команду, розподіляти обов'язки та здійснювати менеджмент процесу розробки на різних рівнях, а також розуміти суть бізнес-процесів при проектуванні та розробці програмного забезпечення.

Викладацький склад повинен відповідати вимогам вищої школи до вищого навчального закладу IV рівня акредитації.

Механізми комунікацій: співпраця викладача і студента в процесі наукової і навчально-виховної діяльності; безпосереднє спілкування викладача зі студентом не лише в аудиторії, а й поза навчальним процесом; регулярне проведення індивідуальних, групових і масових форм наукової і науково-виховної діяльності; використання

можливостей Інтернету, зокрема при організації дистанційного навчання та здійснення поточного контролю знань (з обов'язковим очним проміжним та підсумковим контролем).

Методологічна складова регіонального GameHub спрямована на формування компетенції фахівця у галузі. Реалізація цієї складової передбачає проведення методичних семінарів кафедри; впровадження компетентністно-орієнтованого підходу до створення модулів навчальних дисциплін; організацію самостійної роботи слухачів.

Вивчення наведених вище модулів передбачає наступні форми навчання: а) аудиторні (лекційні заняття, практичні, лабораторні, семінари), б) самостійну робота. Форми організації навчання: а) фронтальна; б) нефронтальна; в) парна: г) групова. Види навчання вибирають у залежності від цільової аудиторії.

Студенти навчального закладу навчаються за звичною для них системою (лекційні заняття, лабораторні роботи тощо). Для осіб, що не мають постійної роботи, ветеранам АТО використовують методику проблемного та програмоване навчання.

Слухачі курсів, що належать до групи безробітних та ветеранів АТО, слухають курси за модулями: Computer game development. Cocos 2DX engine; Advanced computer game development. Cocos 2DX engine; Introduction to computer game audio engineering.

Програми підготовки студентів університету визначено робочим навчальним планом, що передбачає вивчення згаданих та наступних дисциплін у формі спеціальних курсів.

Цикл фундаментальної та природничо-наукової підготовки: математичний аналіз; алгебра та геометрія; дискретна математика; теорія ймовірностей та математична статистика; теорія алгоритмів і математична логіка; диференціальні рівняння.

Цикл професійної та практичної підготовки: іноземна мова; алгоритми і структури даних; архітектура обчислювальних систем; бази даних інформаційних систем; інтелектуальні інформаційні системи; комп'ютерні мережі; методи оптимізації та дослідження операцій; методика викладання інформатики; обробка зображень та мультимедіа; операційні системи та системне програмування; організація та обробка електронної інформації; програмування; паралельні та розподілені обчислення; платформи корпоративних інформаційних систем; програмування та підтримка веб-застосувань; системний аналіз та теорія прийняття рішень; теорія програмування; цифрова техніка; мікропрограмування; мікропроцесорні системи; аналіз даних; теорія інформації та кодування; перетворення форми інформації; цифрова обробка інформації; безпека інформаційних технологій; БЖД і цивільний захист; алгоритми і структури даних; розпаралелювання в обчислювальних системах; мультимедійні додатки в моделюванні та візуалізації; компоненти систем перетворення інформації; візуальне програмування; структури даних та алгоритми їх перетворення; технології побудови систем перетворення інформації.

Методичне забезпечення: навчальні плани, навчальні програми нормативних і вибіркових навчальних дисциплін; програми навчальної, виробничої та інших видів практик; підручники і навчальні посібники; інструктивно-методичні матеріали до семінарських, практичних і лабораторних занять; індивідуальні навчально-дослідні завдання; контрольні роботи; текстові та електронні варіанти тестів для поточного і підсумкового контролю, методичні матеріали для організації самостійної роботи студентів та слухачів курсів, виконання індивідуальних завдань, курсових і дипломних робіт.

Політика сертифікації: ДВНЗ «Прикарпатський національний університет імені Василя Стефаника» здійснює видачу сертифікату слухачеві курсів на підставі виконання навчального плану та виконання підсумкового навчального проекту.

V. Висновки

Проект «Співробітництво між університетами та підприємствами в сфері ігрової індустрії в Україні — GameHub» передбачає розробку дидактичного та методичного забезпечення міждисциплінарних курсів, метою вивчення яких є набуття слухачами теоретичних знань та практичних навичок, необхідних для розробки ігрового програмного забезпечення як індивідуально, так і у складі команди розробників.

Розробка програмного забезпечення при вивченні шкільних дисциплін у ігровій формі та проведення навчання серед вчителів сприятимуть поглибленню зв'язків між ВУЗом та середньою освітою.

Тісні зв'язки із провідними IT-компаніями дозволяють виконувати поставлені освітні задачі на високому рівні, що повинно знайти свій відбиток у потенційному працевлаштуванні слухачів курсів та студентів.

Інша, не менш важлива мета проекту – психологічна реабілітація учасників АТО.

Досягнення обидвох складових мети матиме важливе значення і після завершення проекту.

Література References

[1] Співробітництво між університетами та підприємствами в сфері ігрової індустрії в Україні — GameHub [Online]. Available: http://gamehub-cbhe.eu/ua/

[2] Команда AGH University of Science and Technology [Online]. Available: http://gamehub-cbhe.eu/ua/project-consortium/agh-team/

[3] Команда ДВНЗ «Прикарпатський національний університет імені Василя Стефаника» [Online]. Available: http://gamehub-cbhe.eu/ua/project-consortium/vasyl-stefanyk-precarpathian-national-university-team/

Інформаційні Технології в Діяльності Керівника Навчального Закладу

Юрій Іляш, Олеся Власій
кафедра інформатики
Прикарпатський національний університет
Івано-Франківськ, Україна
yurii.iliash@pu.if.ua, olesia_vlasii@comp-sc.if.ua

Богдан Білусяк
Манявський навчально-виховний комплекс
Манява, Україна
Manyava_school@meta.ua

Information Technologies in of School Director Activity

Yurii Iliash, Olesia Vlasii
dept. of Computer Science
Precarpathian National University
Ivano-Frankovsk, Ukraine
yurii.iliash@pu.if.ua, olesia_vlasii@comp-sc.if.ua

Bohdan Bilusak
Maniava School
Maniava, Ukraine
Manyava_school@meta.ua

***Анотація*—проведено аналіз можливостей використання інформаційно-комунікаційних технологій в управлінській діяльності керівника навчального закладу, запропоновано шляхи впровадження інтернет-технологій, хмарних технологій та технологій візуалізації інформації в різних сферах управлінської діяльності керівника навчального закладу.**

***Abstract*—Possibilities of using information and communication technologies in administrative activity of a school director are analyzed, the ways of implementation of internet technologies, cloud based technologies and technologies of visualization in different areas of administrative activity are proposed.**

Ключові слова—інформаційно-комунікаційні технології (ІКТ); інтернет-технології; хмарні технології; візуалізація інформації; управлінська діяльність; навчальний заклад (НЗ)

Keywords—information and communications technologies; internet technologies; cloud based technologies; vizualization of information; management activities; school

I. Вступ

Згідно національної стратегії розвитку освіти в Україні на 2012–2021 роки, одним із пріоритетних напрямків розвитку є впровадження в освітній процес сучасних інформаційно-комунікаційних технологій (ІКТ), які забезпечують доступність та ефективність освіти, підготовку молодого покоління до життєдіяльності в інформаційному суспільстві. Реалізація таких пріоритетів досягається, зокрема, шляхом створення системи інформаційно-аналітичного забезпечення у сфері управління навчальними закладами (НЗ), вдосконалення інформаційно-комунікаційного забезпечення навчальних закладів та постійним підвищенням кваліфікації керівників НЗ відповідно до вимог сучасного інформаційного суспільства [1].

Стрімкий розвиток ІКТ та входження їх у всі сфери людської діяльності змушує змінювати усталені підходи до організації діяльності навчального закладу з метою як вдосконалення інформаційно-правової бази знань, так і недопущення шкідливого інформаційно-психологічного впливу засобів медіаіндустрії на дітей. В той же час, швидке інтегрування ІКТ у навчально-виховний і управлінський процеси вимагають від керівника НЗ, перш за все, володіння новітніми ІКТ, а також підтримку їх впровадження і використання. Однією з оцінок рівня професійності керівників НЗ став рівень володіння ІКТ [2].

На сьогодні помітно зросла кількість досліджень можливостей використання ІКТ в освітньому процесі. Цій темі в Україні присвячені дослідження багатьох науковців (наприклад, А. Т. Ашеров, В. Ю. Биков, О. М. Бондаренко, Я. В. Булахова, Ю. О. Дорошенко, Ю. О. Жук, В. Ф. Заболотний, Г. О. Козлакова, Ю. І. Машбиць, О. А. Міщенко, Н. В. Морзе, О. П. Пінчук, О. М. Торубара, О. В. Шестопал, О. Г. Яцюк та ін.) [3, 4]. Практично усі автори відзначають, що використання ІКТ дозволяє підвищити інтенсивність та ефективність освітнього процесу, створюючи умови для самоосвіти і тим самим дозволяючи здійснювати перехід до безперервної освіти. Однак, на нашу думку, питання використання ІКТ саме в діяльності керівників НЗ вимагає глибшого дослідження. Тому метою даної роботи є аналіз можливостей використання сучасних ІКТ в діяльності керівника навчального закладу та пошук шляхів їх впровадження з метою вдосконалення організації освітнього процесу.

II. Використання інтернет-технологій як важливий елемент діяльності навчального закладу

A. Ведення сайтів та блогів

Інтернет став невід'ємною частиною життя сучасного покоління. Поряд з небезпекою інтернет-залежності, він надає необмежені можливості доступу до інформаційних ресурсів, які можна використовувати для самоосвіти. У зв'язку з цим, все більше навчальних закладів підключається до всесвітньої мережі. З одного боку, НЗ задовольняє свою потребу залучення до відкритого інформаційного простору, з іншого –починає шукати можливості для реалізації освітніх цілей і завдань [2, 5]. Необхідною умовою функціонування НЗ в умовах сьогодення є наявність його офіційного сайту, на якому відображається нормативно-правова, навчально-методична та інформаційна складові його діяльності. Підтримку ведення сайту чи блогу НЗ повинен забезпечувати весь педагогічний колектив, а координацію та контроль цього процесу покладається на керівника НЗ. На рис. 1Рис. 5 наведено приклад сторінки блогу Манявського навчально-виховного комплексу. В блозі розміщується актуальна інформація про методичну та виховну роботу, співпрацю із партнерами, можливості професійно-технічного навчання тощо.

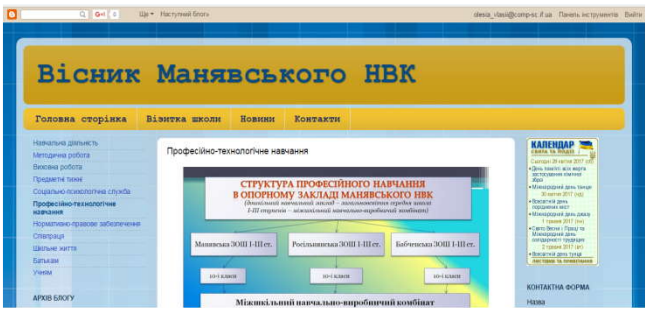


Рис. 1. Сторінка блогу навчального закладу

Дослідження свідчать, що використання мережі Інтернет сприяє розвитку мислення, надає нові засоби для розв'язання творчих завдань, змінює сам стиль розумової діяльності [6][6]. Тому використання інтернет-технологій вчителями-предметниками також розглядається як одна із складових організації ефективного навчального процесу. Наприклад, кожен із вчителів Манявського НВК веде свій блог, розміщуючи в ньому цікаву інформацію для заохочення школярів до навчання та саморозвитку. Приклад блоuy вчителя наведено на рис. 2.

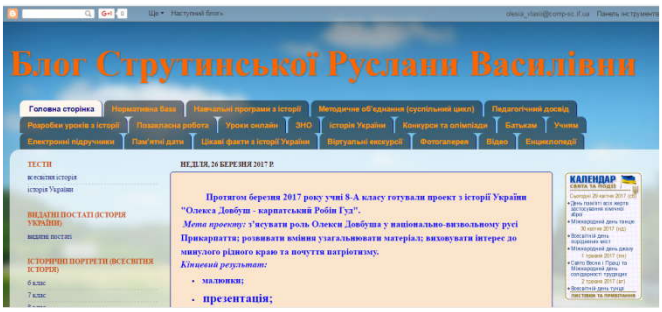


Рис. 2. Приклад блогу вчителя історії Манявського НВК

B. Використання хмарних технологій

У зв'язку із необхідністю зберігання та опрацювання великої кількості інформації, наявності доступу до електронних ресурсів багатьох користувачів одночасно, все більшої популярності набуває використання хмарних технологій в різноманітних сферах людської діяльності, зокрема і освітньої [7]. Хочемо, зокрема, зупинитися на можливостях використання пакету сервісів G Suite for Education, розроблених компанією Google спеціально для освіти. Пакет містить як загальновідомі сервіси (Gmail, Google Calendar, Google Documents, Google Drive, Google Sites, Google Groups), так і специфічні освітянські сервіси (Google Classroom, Vault) (див. рис. 3).

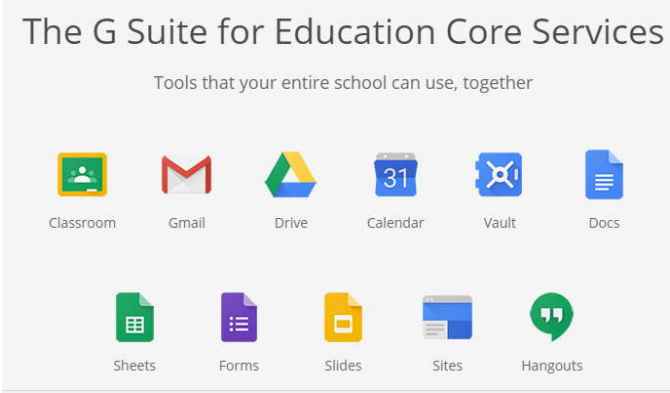


Рис. 3. Сервіси пакету G Suite for Education

Наведемо коротку характеристику основних сервісів для аналізу можливості їх використання в діяльності керівника НЗ.

Gmail працює на будь-якому комп'ютері або мобільному пристрої з можливістю передавання даних, а підтримка роботи в режимі офлайн дозволяє продовжувати роботу навіть за відсутності з'єднання з мережею. Налаштування груп контактів та можливість фільтрування дозволяє ефективно налагодити спілкування за допомогою електронної пошти.

Google Calendar дозволяє організувати робочий день і отримувати нагадування про події на телефон або в папку "Вхідні", створювати спільні події, прикріплювати до них файли, щоб усі потрібні матеріали до початку зустрічі були під рукою у всіх учасників зустрічі.

Пакет Google Documents (Docs, Sheets, Slides, Forms) призначений для створення мультимедійних документів із зображеннями, таблицями, формулами, посиланнями. Можливість спільного редагування, коментування, рецензування полегшує роботу над загальними доповідями, звітами, презентаціями тощо.

За допомогою Google Sites можна розробляти проекти сайтів, не пишучи жодного рядка програмного коду. Оновлена версія Google Sites дає можливість створювати сайти так само просто, як звичайні документи. Для заощадження часу пропонується сотні доступних шаблонів. Такі сайти зручно використовувати з навчально, виховною та розвиваючою метою.

Google Groups можна використовувати як списки розсилки й надавати на їх основі швидкий доступ до календарів, документів, сайтів і відео своїм колегам. Користувачі можуть легко створити групи, а власники – керувати членством у кожній із них.

Google Drive – хмарний диск на комп'ютері під керуванням ОС Mac, ПК або мобільному пристрої (чи у веб-переглядачі), який утворює централізоване місце зберігання найновіших версій файлів, доступ до яких не залежить від місцезнаходження користувача. Зауважимо, що пакет G Suite for Education надає навчальному закладу у користування необмежений Google Диск. Таким чином можна забезпечити доступ до необхідної документації усіх членів колективу.

Google Classroom – система управління навчанням, організована за принципом технологій Веб 2.0. Доступний інтерфейс та автоматична організація документів на Google Диску робить цей сервіс надзвичайно зручним для організації освітнього процесу як у межах навчального закладу, так і для дистанційного навчання.

Отже, можна виділити наступні переваги використання пакету G Suite for Education: мінімальні вимоги до апаратного забезпечення (єдина умова – наявність доступу до Інтернет); відсутність витрат на придбання й обслуговування спеціального програмного забезпечення; підтримка всіх операційних систем та клієнтських програм, що використовуються в НЗ; можливість роботи з документами з будь-якого мобільного пристрою, що підтримує доступ до Інтернет; всі сервіси пакету безкоштовні.

Таким чином, керівник НЗ, який сприяє впровадженню G Suite for Education, має можливість краще налагодити спільну діяльність усього колективу.

III. СПЕЦИФІКА ВІЗУАЛІЗАЦІЇ ІНФОРМАЦІЇ В УПРАВЛІНСЬКІЙ ДІЯЛЬНОСТІ

У зв'язку з інформаційним вибухом значного поширення набув візуальний спосіб подачі інформації, який дозволяє представити значний обсяг даних в організованому вигляді, зручному для перегляду та засвоєння. Очевидно, що візуалізація інформації – необхідний інструмент в діяльності керівника НЗ, одним із завдань якого є донесення інформації до шкільного колективу, керівництва та партнерів. Володіння сучасними способами візуалізації стала однією із складових інформаційної компетентності сучасного управлінця.

Науковцями доведено, що людина схильна обробляти найкраще саме візуальну інформацію. Окрім швидшого сприйняття та обробки, візуалізація має ще декілька переваг: акцентування уваги на різних аспектах даних, аналіз великого набору даних зі складною структурою, зменшення інформаційного перевантаження сприймача інформації і утримування його уваги, однозначність і ясність виведених інформації. Важливою рисою візуалізованих даних є їхня переконливість, тому необхідно уникати викривлення інформації в процесі візуалізації. В той же час, в хорошій візуалізації наочність загальної картини не заважає сприйняттю деталей [8-10].

Найпростіші способи візуалізації – це таблиці, кругові діаграми, графіки, карти, логічні схеми, однак доцільне комбінування текстових блоків, фотографій, карт, таблиць, діаграм тощо [9]. В літературних джерелах можна зустріти використання терміну інфографіка як синоніму до візуалізації інформації, однак такі терміни слід розрізняти. Візуалізація інформації – термін, пов'язаний з дизайном, розробкою та застосуванням комп'ютерного графічного чи інтерактивного представлення інформації. Інфографіка – це оригінальна, створена дизайнером графіка для візуалізації не занадто великої кількості інформації. Сьогодні під цим терміном частіше розуміють графіку для друкованих та Інтернет–видань. Таким чином, візуалізація має загальний характер, інфографіка – специфічний та вузькоспрямований. Візуалізація є вільною від контексту, інфографіка ж залежить від контексту. Візуалізація в основному генерується автоматично, інфографіка – результат ручної роботи [11]. Приклад візуалізації наведено на рис. 4Рис. 4, а інфографіки – на рис. 5.

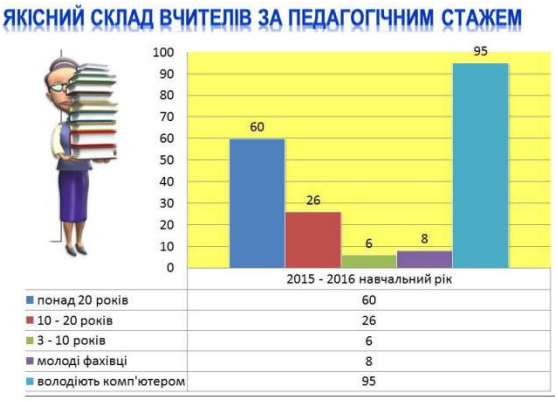


Рис. 4. Приклад візуалізації подачі інформації

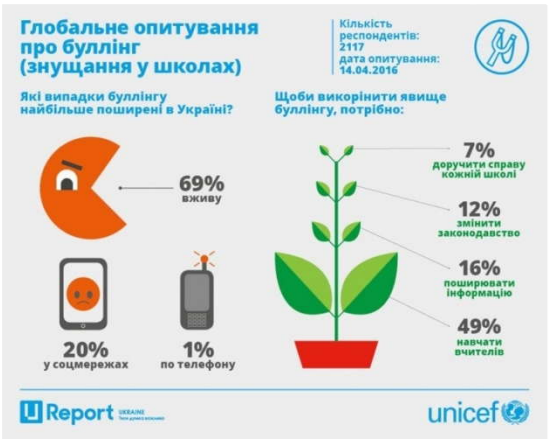


Рис. 5. Приклад представлення інформації засобами інфографіки

Створити якісну інфографіку непросто. Те, що виглядає логічно, красиво й функціонально, частіше за все формується шляхом перебору множини варіантів викладу, потребує наявності базових знань подання інформації, високого рівня розуміння предмету. Розробник інфографіки повинен вникнути в задачу, дослідити дані на вході, проаналізувати їх, інтерпретувати та показати зрозуміло, доступно, яскраво та логічно.

Інфографіка відрізняється сучасним яскравим дизайном, тому здається, що створити її може тільки дизайнер-професіонал, використовуючи потужні програмні середовища Adobe Photoshop, Corel Draw тощо. Виявляється, що ні – наявна значна кількість сервісів, що надають можливість створювати інфографіку самому [3][6]. Існує ряд безкоштовних онлайн-ресурсів, які допоможуть створити інфографіку на потрібну тему, намалювати інтерактивні графіки та навіть візуалізувати резюме:

Infogr.am (простий 3-кроковий сервіс: оберіть шаблон, графічно унаочніть дані та поділіться інфографікою із друзями та знайомими, найбільше підходить для створення схем, графіків і карт);

Easel.ly (десятки безкоштовних шаблонів для створення інфографіки на будь-яку тему, максимальна простота у використанні);

Vizualize.me (допомагає перетворити резюме на цікаву інфографіку, діє за принципом конструктура, присутня функція синхронізації кар'єрних даних із профілю у LinkedIn);

Venngage (простий у користуванні ресурс, у наборі додаткових можливостей якого − використання анімації та завантаження фонів і картинок із зображенням бренду);

Dipity (створення інфографіки з можливістю розміщення готового продукту на сторонніх сайтах, зручна для розробки інфографіки-хроніки, доступна функція додавання схем, фото, а також відеороликів, аудіозаписів, календарних позначок і навіть гіперпосилань);

Draw.io (один із найзручніших і найпростіших безкоштовних ресурсів для створення діаграм і блок-схем. На сайті доступні різноманітні шаблони структур і форм. Також можна використовувати для створення графіків та UML-моделей. Підтримка функції експорту готових схем, а також опції синхронізації з Google Диском);

Spritesapp (можливість створення інфографічних презентацій, а також опція CSS-анімації. Ресурс орієнтований на мобільну інфографіку);

Stat Planet (створення інтерактивної візуалізації та статичних зображень, доступ до різногалузевих даних з усього світу);

Wordle (створення словесних візуалізацій, десятки способів унаочнити дані)

Наведемо приклади використання інфографіки на різних рівнях професійної комунікації в процесі діяльності керівника НЗ:

на особистому рівні (розстановка пріоритетів у цінностях, уміннях, діях та планах на майбутнє, прийняття усвідомленого рішення при багатоваріантному виборі),

на рівні спілкування з колегами (логічно, структуровано доносити інформацію з визначенням понять і термінів у формі, зрозумілій іншим, що дозволяє швидше приходити до єдиної думки при обговоренні питань, а також легко мультиплікувати ідеї,

на рівні спілкування з підлеглими (доступно пояснювати процеси, які відбуваються в навчальному закладі, проводити порівняльний аналіз ефективності складових того чи іншого напрямку діяльності),

на рівні спілкування з партнерами (оптимізувати «бюджет» переговорів, створюючи єдине бачення теми, яка обговорюється, чітко доносячи аргументацію, а також фіксуючи результат з єдиним баченням домовленостей),

на рівні спілкування з керівництвом (оптимально і з доступною аргументацією доносити свої ідеї, створювати звіти про вашу діяльність з логічними аргументами до подальших дій).

IV. Висновки

Володіння сучасними інформаційно-комунікаційними технологіями та впровадження їх у освітній процес дозволяє керівнику навчального закладу ефективно покращити системи інформаційно-аналітичного та інформаційно-комунікаційного забезпечення навчального закладу.

Література References

[1] Національна стратегія розвитку освіти в Україні на 2012-2021 роки, [Електронний ресурс], Режим доступу: https://goo.gl/WVk7oR.

[2] Т. Дадонова, Роль керівника навчального закладу в упровадженні інноваційних технологій в управлінську діяльність [Електронний ресурс] Режим доступу: https://goo.gl/V8mDic.

[3] І. В. Ставицька, Інформаційно-комунікаційні технології в освіті. [Електронний ресурс], Режим доступу: https://goo.gl/XxuGZz.

[4] Р. І. Сулейманов, Е. Р. Шаріпова, Використання інформаційних технологій і інформаційно-комунікаційних технологій у навчальному процесі підготовки інженерів-педагогів, "*Комп'ютерно-інтегровані технології: освіта, наука, виробництво*", Луцьк, 2013. Випуск №11, С. 139-144.

[5] Н. О. Ласкова, Використання ІКТ в управлінській діяльності загальноосвітнього навчального закладу, [Електронний ресурс], Режим доступу: https://goo.gl/V8mDic.

[6] І. З. Адамова, М. І. Ураk, Використання інтернет-технологій у навчальному процесі, "*Вісник Чернівецького торговельно-економічного інституту. Економічні науки*", 2014, Вип. 1, С. 374-379, Режим доступу: http://nbuv.gov.ua/UJRN/Vchtei_2014_1_50

[7] Моделювання й інтеграція сервісів хмаро орієнтованого навчального середовища : монографія, за заг. ред. С. Г. Литвинової, К. : ЦП «Компринт», 2015, 163 с.

[8] Дзв. Новакивська "Новый взгляд на визуализацию информации", *HRMagazine*, 2010, №5-6, Режим доступу: https://goo.gl/B8kms6.

[9] Н. В. Житєньова, (2016) Technologies of visualization in modern educational trends Відкрите освітнє е-середовище сучасного університету (2). pp. 144-157.

[10] Як і для чого використовувати візуалізацію даних? (2014), [Електронний ресурс], Режим доступу: https://goo.gl/EPc6ey.

[11] І. С. Сакун, "Інформаційна графіка: до питання термінології." Теорія та практика дизайну 2 (2012): 113-119.

СЕКЦІЯ 3.

Теорія інформації, кодування та перетворення форми інформації

SECTION 3.

Information theory, coding and transformation of information form

Апаратно-Мінімальні Паралельні Коди CRC

Василь Семеренко
кафедра обчислювальної техніки,
Вінницький національний технічний університет
Вінниця, Україна
vpsemerenko@ukr.net

Hardware-Minimum Parallel CRC Codes

Vasyl Semerenko
Department of Computer Technique
Vinnytsia National Technical University
Vinnytsia, Ukraine,
vpsemerenko@ukr.net

***Анотація*—Запропоновано паралельні коди CRC (Cyclic Redundancy Code) для виявлення та виправлення помилок в каналах з паралельним надходженням вхідних даних. Як математичний апарат таких кодів використано теорію лінійних послідовнісних схем (ЛПС). Досліджено властивості запропонованих паралельних кодів CRC і показана їх апаратна реалізація**

***Abstract*—Parallel CRC codes for detecting and correcting the errors in channels with parallel input data are proposed. The theory of linear finite state machine (LFSM) is used as a mathematical tools of such codes. The properties of the proposed parallel CRC codes are investigated and their hardware implementation are shown**

Ключові слова—коди CRC; лінійна послідовнісна схема; паралельні коди; завадостійке кодування

Keywords—CRC codes; linear finite state machine; parallel codes; error correcting coding

I. Вступ

Коди CRC (Cyclic Redundancy Code) належать до найбільш розповсюджених кодових методів контролю в різноманітних системах передачі, збереження та архівування даних [1].

З часу своєї появи в 1961 році ці коди були орієнтовані на виявлення помилок при послідовному надходженні вхідних даних. В сучасних інформаційних системах традиційний побітовий спосіб обчислення став головним недоліком CRC, тому стали активно розвиватись паралельні реалізації цих кодів.

Перехід до паралельних CRC вимагає розробки і відповідного теоретичного базису кодів. Більшість дослідників таким базисом вибрали теорію лінійних систем [2] і теорію автоматів [3]. Кодери та декодери послідовних кодів CRC з породжувальним поліномом

$$g(x) = g_0 + g_1 x + \cdots + g_{r-1} x^{r-1} + x^r, \quad GF(2) \qquad (1)$$

стали представляти у вигляді кінцевого автомату на основі так званої матриці F, а паралельних кодів CRC – на основі матриці F^r [4].

Основним недоліком такого підходу стали значні витрати пам'яті при програмній реалізації та велика кількість логічних елементів XOR при апаратній реалізації CRC.

Оскільки основною апаратною технологією були обрані ПЛІС (FPGA), тому основні зусилля зараз спрямовані на схемну оптимізацію паралельних CRC на базі ПЛІС: зменшення апаратних витрат і підвищення швидкодії. Задача мінімізації таких схем є NP-складною задачею, а різними евристичними методами можна оптимізувати схему в межах 10-20%. Навіть для нескладних 32-розрядних кодів CRC необхідно декілька сотень елементів XOR, що суттєво уповільнює роботу CRC [4, 5].

Як було показано в [6], перехід від матриці F до матриці F^r означає перехід від послідовного надходження вхідних даних до кортежно-паралельного способу реорганізації цих даних. Цей спосіб надходження даних прискорює обробку даних в рамках традиційної парадигми кодування-декодування, однак з пропорційним ускладненням апаратури.

Очевидно, необхідно шукати принципово інші варіанти паралельних кодів CRC.

II. Теоретичний базис паралельних кодів CRC

Як паралельний (n, k, ρ)-код CRC над полем $GF(2)$ будемо розуміти код CRC, який використовується для передачі по ρ паралельних каналах [7]. Такий код

складається із ρ кодових слів z_i ($i = 1 \ldots \rho$), об'єднаних в кодову матрицю:

$$Z_{(\rho)} = \begin{bmatrix} z_1 \\ z_2 \\ \ldots \\ z_\rho \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} & \ldots & z_{1n} \\ z_{21} & z_{22} & \ldots & z_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ z_{\rho 1} & z_{\rho 2} & \ldots & z_{\rho n} \end{bmatrix}, \quad GF(2). \qquad (2)$$

Можна розрізняти два типи паралельних кодів CRC: складені та інтегровані. В подальшому будемо говорити лише про інтегровані коди CRC.

Для розуміння суті кодів CRC необхідно повернутись до основ циклічних кодів, до яких ці коди належать. Традиційні способи представлення циклічних кодів (матричне, поліноміальне, алгебраїчне) відіграли важливу роль в становленні цього класу кодів. Однак практика формує нові проблеми і для їх розв'язання найбільш придатним математичним апаратом є теорія лінійних послідовнісних схем (ЛПС) [8].

Традиційна ЛПС з одним входом і одним виходом є кінцевим автоматом лінійного типу (лінійним автоматом), який над полем Галуа $GF(2)$ описується функцією станів (переходів)

$$S(t+1) = A \times S(t) + B \times U(t), \quad GF(2) \qquad (3)$$

і функцією виходів.

$$Y(t) = C \times S(t) + D \times U(t), \quad GF(2),$$

де t – дискретний час; A, B, C, D – характеристичні матриці ЛПС; $S(t)$ – слово стану; $U(t)$ – вхідне слово; $Y(t)$ – вихідне слово.

Оскільки вхідні дані паралельного коду CRC поступають по паралельних каналах, тому необхідно використати багатовходову ЛПС. В [8] представлено багатоканальний аналог традиційної ЛПС, математична модель якої базується на функції станів (переходів)

$$S(t+1) = A_{(\rho)} \times S(t) + B_{(\rho)} \times U_{(\rho)}(t), \quad GF(2)$$

і функції виходів

$$Y_{(\rho)}(t) = C_{(\rho)} \times S(t) + D_{(\rho)} \times U_{(\rho)}(t), \quad GF(2).$$

Характеристичні матриці $A_{(\rho)}$, $B_{(\rho)}$, $C_{(\rho)}$, $D_{(\rho)}$ такої ЛПС мають значно складнішу структуру, ніж відповідні матриці A, B, C, D одновходової ЛПС в (3). Наприклад, матриця $A_{(\rho)}$ дорівнює степені ρ від A: $A_{(\rho)} = A^\rho$. Вхідні та вихідні дані представляються як матриці $U_{(\rho)}(t)$ і $Y_{(\rho)}(t)$, які містять масиви слів $U(t)$ і $Y(t)$.

Оскільки матриця $A_{(\rho)}$ описує внутрішню структуру ЛПС, тому апаратна реалізація багатоканальної ЛПС також буде складною. Саме ця обставина є причиною громіздкої структури відомих варіантів паралельних CRC.

Розглянемо можливість створення апаратно простих кодерів та декодерів паралельних кодів CRC [7].

Теоретичною основою таких кодів CRC може бути паралельна ЛПС, функціонування якої описується функцією станів (переходів)

$$S(t+1) = A \times S(t) + B_{(\rho)} \times U_{(\rho)}(t), \quad GF(2) \qquad (4)$$

і функцією виходів

$$Y_{(\rho)}(t) = C \times S(t) + D_{(\rho)} \times U_{(\rho)}(t), \quad GF(2),$$

Функція станів в (4), як і в (3), базується на одній і тій же самій матриці A, тобто внутрішня апаратна реалізація паралельної ЛПС буде такою ж, як і в одновходової ЛПС.

Паралельна ЛПС може бути побудована на основі трьох різних типів послідовних ЛПС:
рекурсивної ЛПС типу 1 з матрицями

$$A = \begin{bmatrix} 0 & 0 & 0 & \ldots & g_0 \\ 1 & 0 & 0 & \ldots & g_1 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 1 & \ldots & g_{r-2} \\ 0 & 0 & 0 & 1 & g_{r-1} \end{bmatrix} \quad B = \begin{bmatrix} 1 \\ 0 \\ \ldots \\ 0 \\ 0 \end{bmatrix}, (5)$$

рекурсивної ЛПС типу 2 з матрицями

$$A = \begin{bmatrix} g_{r-1} & 1 & 0 & \ldots & 0 \\ g_{r-2} & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & 0 \\ g_1 & 0 & 0 & \ldots & 1 \\ g_0 & 0 & 0 & \ldots & 0 \end{bmatrix}, \quad B = \begin{bmatrix} g_{r-1} \\ g_{r-2} \\ \ldots \\ g_1 \\ g_0 \end{bmatrix}, (6)$$

та рекурсивної ЛПС типу 3 з матрицями

$$A = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \\ g_0 & g_1 & g_2 & \ldots & g_{r-1} \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ 0 \\ \ldots \\ 0 \\ 1 \end{bmatrix}. (7)$$

Елементи останнього стовпця матриці A із (5), першого стовпця матриці A із (6), і елементи останнього рядка матриці A із (7), представляють собою коефіцієнти породжувального поліному (1).

Паралельна ЛПС відрізняється від вказаних послідовних ЛПС лише матрицею B, яка матиме вигляд одиничної матриці.

Помилки в паралельному циклічному коді мають решіткову конфігурацію і математичною моделлю помилок може бути матриця помилок

$$\mathrm{E}_{(\rho)} = \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \dots & \varepsilon_{1n} \\ \varepsilon_{21} & \varepsilon_{22} & \dots & \varepsilon_{2n} \\ \dots & \dots & \dots & \dots \\ \varepsilon_{\rho 1} & \varepsilon_{\rho 2} & \dots & \varepsilon_{\rho n} \end{bmatrix},$$

в якій ненульові елементи відповідають спотвореним позиціям в кодовій матриці (2). В результаті впливу завад в паралельних каналах в приймачі буде отримано кодову матрицю з помилками:

$$Z_{(\rho)err} = Z_{(\rho)} + \mathrm{E}_{(\rho)}, \quad GF(2)$$

На Рис. 1 представлена загальна схема r-розмірної паралельної ЛПС типу 1 з r входами і одним виходом.

III. Кодування паралельного коду CRC

Для отримання CRC як контрольної суми достатньо обчислити протягом k тактів роботи стан $S(k)$ паралельної ЛПС згідно (4).

Якщо необхідно не тільки виявляти, але і виправляти деякі класи помилок, тоді робота паралельної ЛПС продовжується ще протягом r тактів для отримання CRC як циклічного коду, тобто здійснюється процес кодування.

При систематичному кодуванні на основі інформаційної $(\rho \times k)$-матриці $I_{(\rho)}$ обчислюється контрольна $(\rho \times r)$-матриця $\Psi_{(\rho)}$ і їх об'єднання утворює кодову матрицю $Z_{(\rho)}$.

Для інтегрованого паралельного (n, k, ρ)-коду CRC матриця $\Psi_{(\rho)}$ формується на основі однієї паралельної ЛПС, тому знадобиться лише один паралельний кодер. На боці приймача необхідно мати паралельний декодер на основі такої ж паралельної ЛПС.

Основна стратегія кодування для паралельного коду CRC залишається такою ж, як і для звичайних циклічних кодів [7].

Кодування циклічних кодів на основі їх автоматного представлення базується на властивості r-керованості паралельної ЛПС, тобто існуванні вхідної послідовності довжини r, яка переводить ЛПС з одного заданого стану в інший. Для (n, k, ρ)-коду CRC ЛПС буде r-керованою, якщо ранг матриці

$$L_{(\rho),r} = \begin{bmatrix} l_1 & l_2 & \dots & l_r \end{bmatrix},$$

де $l_j = \sum_1^r a_i$; a_i – i-й стовпець матриці $A_{(\rho)}^{r-i} \times B_{(\rho)}$, $j = 1 \dots r-1$; $l_r = \sum_1^r a_i$; a_i – i-й стовпець матриці $B_{(\rho)}$ над полем $GF(2)$ буде дорівнювати r [8].

Нехай задана r-вимірна паралельна ЛПС типу 1 і деяка інформаційна $(r \times k)$-матриця $I_{(r)}$ початкових даних, яка призначена для передачі по каналу зв'язку ($r = \rho$). Якщо ($\rho < r$), тоді $(r - \rho)$ значень можна прийняти нульовими.

Внаслідок подачі $I_{(r)}$ на r входів ЛПС через k тактів відбудеться її перехід з початкового стану $S_{beg}(0)$, зазвичай нульового, в деякий стан $S(k)$. Аналітично стан $S(k)$ можна визначити з рівняння (4).

Завдяки r-керованості паралельної ЛПС можна не більше ніж за r тактів під дією $(r \times r)$-матриці $\Psi_{(r)}$ вхідних даних перейти із стану $S(k)$ в стан $S_{end}(n)$:

$$S_{end}(n) - A_{(r)}^r \times S(k) = L_{(r),r} \times \Psi_{(r)}, \quad GF(2). \qquad (8)$$

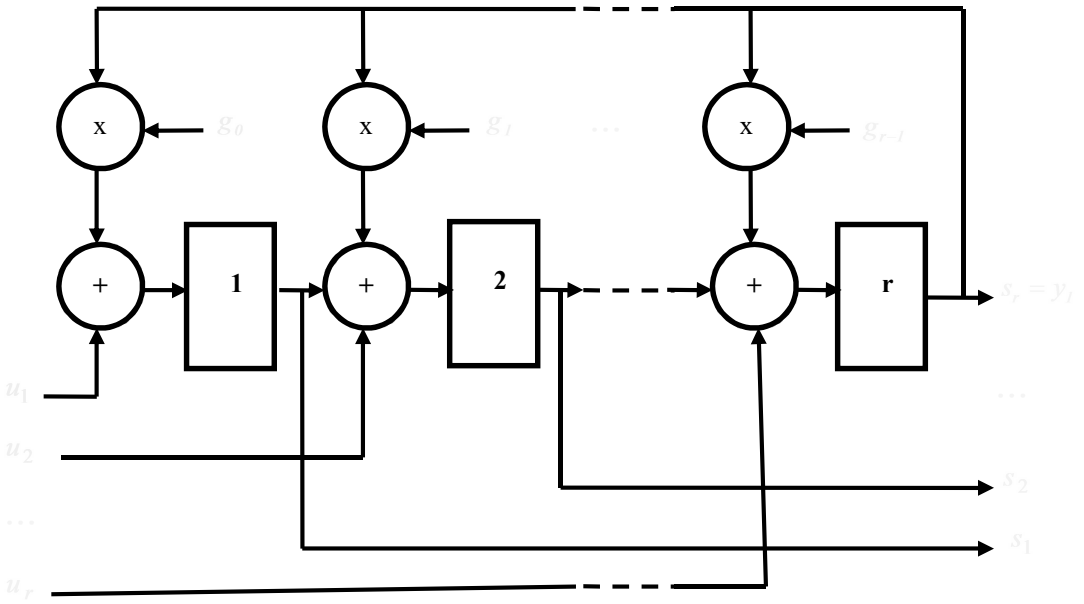


Рис. 1. Загальна схема r- розмірної паралельної ЛПС типу 1 з r входами і одним виходом (тригери позначені прямокутниками, елементи XOR – колом з символом '+')

Приймаючи кінцевий стан $S_{end}(n)$ ЛПС нульовим, запишемо рівняння (8) у вигляді

$$L_{(r),r} \times \Psi_{(r)} = S(n), \quad GF(2), \tag{9}$$

де $S(n) = A^r_{(r)} \times S(k), \quad GF(2)$.

Підставляючи в (9) значення матриці $L_{(r),r}$ із (8), отримаємо систему r рівнянь для знаходження невідомих значень матриці $\Psi_{(r)}$. Але для однозначного знаходження ($r \times r$) невідомих компонент матриці $\Psi_{(r)}$ недостатньо r рівнянь системи (9). Тому необхідно прийняти однаковими значення всіх рядків матриці $\Psi_{(r)}$. В підсумку отримуємо таку систему r рівнянь для обчислення значень матриці $\Psi_{(r)}$:

$$\begin{bmatrix} l_{1,1}\psi_1 & + & l_{1,2}\psi_2 & + & \dots & + & l_{1,r}\psi_r \\ l_{2,1}\psi_1 & + & l_{2,2}\psi_2 & + & \dots & + & l_{2,r}\psi_r \\ \dots & & \dots & & \dots & & \dots \\ l_{r,1}\psi_1 & + & l_{r,2}\psi_2 & + & \dots & + & l_{r,r}\psi_r \end{bmatrix} = \begin{bmatrix} s_1 \\ s_2 \\ \dots \\ s_r \end{bmatrix}.$$

В результаті розв'язання цієї системи рівнянь отримаємо систему перевірок, яка зв'язує компоненти контрольної матриці $\Psi_{(r)}$ і слова стану $S(n)$. Така система перевірок і застосовується в процесі кодування.

Обчислене для конкретного стану $S(n)$ слово

$$\Psi_{(r)} = [\psi_1 \quad \psi_2 \quad \dots \quad \psi_r]$$

і представляє собою CRC як код.

Далі формується контрольна ($r \times r$)-матриця $\Psi_{(r)}$ систематичного інтегрованого паралельного (n, k, r)-коду CRC.

Підкреслимо, що зазначена система рівнянь розв'язується тільки один раз для конкретного інтегрованого паралельного коду CRC.

Таким чином, при апаратній реалізації кодера інтегрованого паралельного коду CRC, (паралельної ЛПС), додатково знадобиться схема обчислення вказаної системи перевірок.

ПРИКЛАД 1. Для інтегрованого паралельного (15, 11, 4)-коду CRC з породжувальним поліномом $g(x) = 1 + x + x^4$ виконати систематичне кодування інформаційної матриці (в канал дані поступають справа)

$$I_{(4)} = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

на основі рекурсивної ЛПС типу 1.

Заданому коду відповідають такі характеристичні матриці паралельної ЛПС типу 1:

$$A_{(4)} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}; \quad B_{(4)} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

В результаті кодування отримаємо код CRC та контрольну матрицю $\Psi_{(4)}$:

$$\text{CRC} = [1 \quad 0 \quad 1 \quad 1], \ \Psi_{(4)} = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix}.$$

IV. Висновки

Введено поняття (n, k, ρ)-коду CRC за допомогою кінцевих автоматів в полях Галуа – лінійних послідовнісних схем (ЛПС). Розглянуто три типу ЛПС, які відрізняються своїми характеристичними матрицями.

Основна перевага запропонованого підходу – суттєве спрощення апаратної реалізації CRC. По-перше, зменшується кількість логічних елементів XOR, оскільки матриця A має значно більше нульових елементів, ніж матриця A^r. По-друге, елементи XOR рівномірно розміщені між тригерами, що дозволяє збільшити тактову частоту і, відповідно, швидкодію всієї схеми.

Література References

[1] V. P. Semerenko. "Theory and Practice of CRC Codes Based on Automaton Models," (in Russian), *Eastern-European Journal of Enterprise Technologies*, vol. 4, issue 9 (76), pp. 38–48, 2015. doi: 10.15587/1729-4061.2015.47860

[2] G. Campobello, G. Patane, M. Russo. "Parallel CRC realization," *IEEE Transactions on Computers*, vol. 52, issue 10, pp. 1312–1319, 2003.

[3] K.V. Krishna Reddy. "An Optimization Technique for CRC Generation," *International Journal of Computer Trends and Technology (IJCTT)*, vol. 4, issue 9, pp. 3260–3265, 2013.

[4] M. Grumel and S. B. Furber. "Novel Programmable Parallel CRC Circuit," *IEEE Transactions on Very Large Scale Integration (VLSI) Circuit*, vol. 4, issue 9, pp. 3260–3265, 2010.

[5] S. R. Ruckmani and P. Anbalagan. "High Speed Cyclic Redundancy Check for USB," *DSP Journal*, vol. 6, issue 1, September, pp. 45–49, 2006.

[6] V. P. Semerenko. "The Theory of Parallel CRC Codes Based on Automaton Models," *Eastern-European Journal of Enterprise Technologies*, vol. 6, issue 9 (84), pp. 45–55, 2016. doi: 10.15587/1729-4061.2016.85603

[7] В. П. Семеренко. "Паралельні циклічні коди," *Вісник ВПІ*, – 2014. – № 6. – С. 65–72.

[8] A. Gill. *Linear Sequential Circuits. Analysis, Synthesis and Application,* New York, London: McGraw-Hill Book Company, 1967.

On some Class of Quantum Finite Automata

Volodymyr G. Skobelev
Dept. of Theory of Digital Automata
V.M. Glushkov Institute of Cybernetics of NAS of Ukraine
Kyiv, Ukraine
skobelevvg@mail.ru

Про Один Клас Квантових Скінченних Автоматів

Володимир Г. Скобелєв
Відділ теорії цифрових автоиатів
Інститут кібернетики імені В.М. Глушкова НАН України
Київ, Україна
skobelevvg@mail.ru

***Abstract*—Classes of languages accepted either with given probability, or with given mistake by basic models of finite quantum automata are characterized under supposition that unitary operators satisfy to commutative law.**

***Анотація*—Охарактеризомано класи мов, які розпізнають з заданою ймовірністю, обо з заданою помилкою скінченні квантові автомати за умови, що унітарні оператори задовільняють закону комутативності.**

Keywords—quantum finite automata; unitary operators; commutativity; accepted languages.

Ключові слова—квантові скінченні автомати; унітарні оператори; комутативність; розпізнаваємі мови.

I. Introduction

Success in the development of quantum algorithms theory has stimulated intensive formation of quantum automata (QA) theory. Very extraordinary situation takes place for finite QA theory (since we deal only with finite QA the word "finite" in word combination "finite QA" is omitted). A variety of different QA models intended to recognize languages in the given alphabet (i.e. all these models are acceptors) has been proposed. This means, at least, that QA models can be used for the decision of problems, in which this or that subtask can be reduced to recognition of some language. It is worth to note that all proposed QA models differ by complexity, as well as by recognizing capacities.

Each QA model is defined in terms of quantum Turing machine (QTM) in the following way. The set of the states of QTM is the set of unit vectors in the complex n-dimensional Euclidean space, in which some orthonormal basis is fixed. Elements of this orthonormal basis are the basic states for QA. Besides, some set of accepting basic states is fixed. Initially QTM is either in some fixed initial state, or in some fixed initial mixed state, which is some finite set of ordered pairs of the type (state, its probability), such that the sum of probabilities equals to 1. There is infinite to the right input tape partitioned into cells, enumerated by positive integers.

The number of heads of QTM is equal to some fixed positive integer k. Finite input alphabet is fixed. With each input string, which length does not exceed the integer k, it is associated some unitary operator acting in the complex n-dimensional Euclidean space. Analyzed input string is written, letter by letter, in the leftmost fragment of the tape, and initially k heads of QTM observe the first k cells of the tape.

When QTM heads observe current fragment of the tape, then only a single action of the following two types of actions can be performed. The first type of actions is measurement. It transforms current state (or current mixed state) of QTM into some basic state (or some set of basic states) of QA. The probability for QTM to be transformed into this or the other basic state depends only on the current state (or current mixed state) of QTM. The second type of actions is transformation of current state (or current mixed state) of QTM by unitary operator, which corresponds to the current fragment of the input string, and simultaneous shifting all k heads one cell to the right (i.e. we deal only with 1-way QTM).

Probability of QTM to accept analyzed input string is computed at the final stage. This computing is reduced to projection of final state (or final mixed state) of QTM into the subspace spanned by the accepting basic states. It is worth to note that any projection of current state of QTM is also considered as some variant of measurement. Since we deal with probability of accepting analyzed input string, two problems of language recognition for QA models can be considered, namely, recognition with given probability and recognition with given mistake.

The main problems that have been investigated for QA are the following ones: to describe in the explicit form the class of languages recognized by this or the other QA model, to

compare recognizing capacities of different QA models, and to find criteria of states equivalence for this or the other QA model.

Unfortunately, the situation with the first problem is very far from its complete decision. Because of this situation it seems reasonable to investigate QA models under supposition that there are some additional requirements, which are natural, at least, from the mathematical point of view. Obviously, that this class of requirements includes the supposition that unitary operators satisfy to commutative law. The aim of the given paper is to investigate QA models under this assumption.

II. Basic QA Models

Let X ($|X|=m$) be fixed input alphabet, C^n be complex n-dimensional Euclidean space, Q be fixed basis in it, $Q_{ac} \subseteq Q$ be the set of accepting states, and P_{ac} be the projection operator on the subspace spanned by the accepting basic states.

The following well known approach, which has been used in the theory of probabilistic automata, can be applied for definition of languages recognized by QTM:

- A language $L \subseteq X^+$ is accepted by QTM with given probability p $(0.5 < p \le 1)$, if any string $w \in L$ is accepted with probability not less then p, while any string $w \notin L$ is accepted with probability not exceeding $1-p$.

- A language $L \subseteq X^+$ is accepted by QTM with given mistake (p_1, p_2) $(0 < p_1 < p_2 < 1)$, if any string $w \in L$ is accepted with probability not less then p_2, while any string $w \notin L$ is accepted with probability not exceeding p_1.

A. Models of QA Defined via 1-Head QTM

We deal with 1-head QTA, which head at each instant move one cell to the right. With each input letter $x \in X$ it is associated some unitary operator U_x, acting in the complex n-dimensional Euclidean space C^n. Thus, for each input string $w = x_1 \ldots x_l$ it is uniquely defined unitary operator $U_w = U_{x_l} \ldots U_{x_1}$.

Firstly we consider basic QA models that start in fixed initial state $\mathbf{s}_0 \in C^n$, which is some unit column vector.

The model MO-1QFA [1] is 1-head QTM, with single measurement, carried out only on the final stage. An input string $w = x_1 \ldots x_l$ is accepted by this model with probability

$$P_{\text{MO-1FQA}}(\mathbf{s}_0, w) = \| P_{acc} U_w \mathbf{s}_0 \|^2 .$$

It has been established in [2] that the class of languages accepted with given mistake by the model MO-1QFA equals to the class of languages accepted by group finite automata.

The model MM-1QFA [3] differs from the model MO-1QFA in the following way. Two non-intersecting subsets Q_{rj} and Q_{nt} $(Q_{rj} \cup Q_{nt} \subseteq Q \setminus Q_{ac})$, correspondingly, of rejecting and of non-accepting basic states are fixed.

At any intermediate instant it is applied the relevant unitary operator, possibly followed by measurement in the basis Q. If the result of measurement is some state $\mathbf{s} \in Q_{nt}$ then the next instant starts, while if the result of measurement is some state $\mathbf{s} \in Q_{ac} \cup Q_{rj}$ then QTM halts. Analyzed input string is accepted, if $\mathbf{s} \in Q_{ac}$, and rejected, if $\mathbf{s} \in Q_{rj}$.

It is worth to note that the model MM-1QFA is based on well known "Decide and Halt" approach, which is intended to solve the class of problems of recognition called "promise problems". The class of languages accepted by the model MM-1QFA is not defined in explicit form, till now. It is only known that this class of languages includes properly the class of languages accepted by group finite automata, and it is included properly in the class of regular languages. Moreover, it has been established in [4] that the class of languages accepted by the model MM-1QFA is not closed under Boolean operations.

The model N-QFA [5] at any instant admits application of any sequence of unitary operators and projective measurements. Thus, exploring of ancilla qubits can be easily presented in this model. The class of languages accepted by the model N-QFA is not defined in explicit form, till now. It is only known that this class of languages satisfies to strict inclusions pointed above for the model MM-1QFA.

The model CL-QFA [6] is some generalization of the model N-QFA. In this model some regular language in the alphabet of igenvalues of Hermitian operators is used to determine projective measurements at intermediate instants. The class of languages accepted by the model CL-QFA is not defined in explicit form, till now. It is only known that any regular language can be accepted with given mistake by the model CL-QFA.

Now we consider basic QA model, called L-QFA [7]. This model is 1-head QTM, with single measurement carried out only on the final stage, and which starts in some fixed initial mixed state. Any mixed state is some set $\{(\mathbf{s}_i, p_i) \mid i = 1, \ldots, r\}$, where unit vectors $\mathbf{s}_i$ $(i = 1, \ldots, l)$ are pair-wise different, $p_i > 0$ for each $i = 1, \ldots, l$, and $p_1 + \ldots + p_r = 1$. An input string $w = x_1 \ldots x_l$ is accepted by the model L-QFA with probability

$$P_{\text{L-FQA}}\{(\mathbf{s}_i, p_i) \mid i = 1, \ldots, r\} =$$
$$= p_1 P_{\text{MO-1FQA}}(\mathbf{s}_1, w) + \ldots + p_r P_{\text{MO-1FQA}}(\mathbf{s}_r, w) .$$

This mixed state $\{(\mathbf{s}_i, p_i) \mid i = 1, \ldots, r\}$ can be presented by density operator

$$\rho = p_1 \mathbf{s}_1 \mathbf{s}_1^T + \ldots + p_r \mathbf{s}_r \mathbf{s}_r^T .$$

Thus, initial state of the model L-QFA is some density operator. Each unitary operator U transforms mixed state ρ

into the mixed state $U^{\dagger}\rho U$. It is known that the class of languages accepted by the model L-QFA equals to the class of languages accepted by invertible probabilistic finite automata.

By analogy with above considered generalizations of the model MO-1QFA, similar generalizations can be constructed for the model L-QFA. However, as far as I know, such generalizations have not been investigated in detail.

B. Model of QA Defined via k -Heads QTM

We deal with k -heads QTM, which heads at each instant, move simultaneously one cell to the right. It is supposed that with each input string $u = x_1 \ldots x_r$ $(1 \le r \le k)$ it is associated some unitary operator U_u acting in the space C^n. Thus, for each input string $w = x_1 \ldots x_l$ it is uniquely defined unitary operator $U_w = U(l) \ldots U(1)$, where $U(j) = U_{x_j \ldots x_{\min\{k+j-1,l\}}}$ for all $j = 1,\ldots,l$.

The model k QFA [8] is k -heads QTM, with single measurement carried out only on the final stage. This model starts in some fixed initial state $\mathbf{s}_0 \in C^n$, which is some unit column vector. The class of languages accepted by the model k QFA is not defined in explicit form, till now. It is only known that this class of languages is closed under Boolean operations, but it is not closed under concatenation and iteration. Moreover, it has been established that for all integers $k \ge 1$ the class of languages accepted by the model k QFA is strictly included into the class of languages accepted by the model $(k+1)$ QFA.

By analogy with above considered generalizations of the model MO-1QFA, similar generalizations can be constructed for the model k QFA. However, as far as I know, such generalizations have not been investigated in detail.

III. Main Results

A. Analysis of Models of QA Defined via 1-Head QTM

Let us characterize languages accepted by the models MO-1QFA and L-QFA under the supposition that unitary operators associated with input letters satisfy to commutative law. We remind that MO-1QFA and L-QFA are the models in which measurement is carried out only at a final stage.

Let $X = \{x_1,\ldots,x_m\}$ be the input alphabet of QA. Unitary operator associated with input letter $x_i \in X$ is denoted U_i. Since it is supposed that $U_i U_j = U_j U_i$ for all $i, j = 1,\ldots,m$, then to define the unitary operator for analyzed input sequence, it is sufficient to know only the number of occurrences of each letter in this input sequence, and unitary operators associated with these letters. Based on this factor we can present uniquely the unitary operator defined for any input string $w \in X^+$ in the standard form

$$sfuo(w) = U_1^{r_1} \ldots U_m^{r_m},$$

where r_i $(i = 1,\ldots,m)$ is the number of occurrences of the letter x_i in the input string w.

Let $\equiv$ be the equivalence relation on the set X^+ defined as follows:

$$w_1 \equiv w_2 \Leftrightarrow sfuo(w_1) = sfuo(w_2).$$

As a matter of fact, this equivalence relation is a congruence on the set X^+, since $w_1 \equiv w_2$ $(w_1, w_2 \in X^+)$ implies that $w_1 x \equiv w_2 x$ for all $x \in X$. We deal with the factor-set $X^+/\equiv$, as with a partition of the set X^+. It is evident that the following theorem holds.

Theorem. For each of the models, MO-1QFA and L-QFA, any language, accepted either with given probability, or with given mistake, is the union of some blocks of the partition $X^+/\equiv$.

It is of special interest the situation, when the equivalence relation $\equiv$ is finitary, i.e. when the partition $X^+/\equiv$ consists of finite number of blocks. Such situation arises, in particular, in the following important case.

Let us suppose that for each unitary operator U_i $(i = 1,\ldots,m)$ there exists some positive integer a_i, such that $U_i^{a_i} = I$, where I is the identity map acting in C^n.

Let a_i $(i = 1,\ldots,m)$ be the minimal positive integer, such that identity $U_i^{a_i} = I$ holds. The unitary operator defined for any input string $w \in X^+$ can be present uniquely in the reduced standard form

$$rsfuo(w) = U_1^{r_1 (\mathrm{mod}\, a_1)} \ldots U_m^{r_m (\mathrm{mod}\, a_m)},$$

where r_i $(i = 1,\ldots,m)$ is the number of occurrences of the letter x_i in the input string w. This factor implies that the equivalence relation $\equiv$ on the set X^+ can be defined in the following way:

$$w_1 \equiv w_2 \Leftrightarrow rsfuo(w_1) = rsfuo(w_2).$$

It is evident that the number for blocks of the partition $X^+/\equiv$ equals to the product $a_1 \cdot \ldots \cdot a_m$.

The following example justifies non-triviality of proposed constructions.

Example. Let us consider 1-qubit MO-1QFA. In this case any special unitary operator can be presented as some composition of rotations about the x -, y -, and z -axes of the Bloch sphere (see [9], for example). The word "special" means that the determinant of a matrix that defines unitary operator equals to 1. It is well known that to investigate QA, it is sufficient to deal only with special unitary operators.

Let us suppose that unitary operator associated with each input letter x_i $(i = 1,\ldots,m)$ is the rotation through the angle

$0.5\alpha_i$ $(\alpha_i \in (0,4\pi))$ around the y -axe of the Bloch sphere. It is evident that these unitary operators satisfy to commutative law. Thus, any language recognized by considered model of QA, either with given probability, or with given error, is the union of some blocks of the partition $X^+/\equiv$. This partition is finitary if and only if each number $\alpha_i\pi^{-1}$ $(i=1,\ldots,m)$ is some rational one. End of Example.

Let us characterize models of QA with measurements at intermediate instants (the models MM-1QFA, N-QFA, and CL-QFA), under supposition that unitary operators associated with input letters satisfy to commutative law.

In this case, when any of the models MM-1QFA, N-QFA, and CL-QFA is processing any given input string $w \in X^+$ of the length $l \geq 2$, different variants for computing, provided by QTM, are possible.

Each variant for computing, provided by QTM, can be determined in the following way. The number r $(0 \leq r < l)$ of measurements at intermediate instants is selected with given probability. Instants for intermediate measurements are fixed by inserting r symbols $|$ on a segment of the analyzed input string w between the first and last letters with given probability.Each maximal fragment of the analyzed input string, which does not contain the symbol $|$, is replaced with the standard form of the unitary operator defined for this fragment. Each symbol $|$ is replaced with measurement admissible on this intermediate step, with given probability, and measurement only on the final stage is also inserted.

It is evident, that if the partition $X^+/\equiv$ is finitary, then each maximal fragment of the analyzed input string, which does not contain the symbol $|$, can be replaced with the reduced standard form of the unitary operator defined for this fragment. Thus, for any of the models MM-1QFA, N-QFA, and CL-QFA, the set of all different variants for computing, provided by QTM, when processing any given input string $w \in X^+$ of the length $l \geq 2$, can be presented by some labeled rooted oriented tree.

B. Analysis of Models of QA Defined via k -Heads QTM

We consider two different cases.

The first case takes the place when either $k=1$, or $k \geq 2$ and all unitary operators associated with input string of the length not exceeding the integer k satisfy to commutative law. In this case we get the situation similar to the one that has been analyzed previously for the model MO-1QFA. It means that the theorem holds in this case.

The second case takes the place when $k \geq 2$ and only unitary operators associated with input string of the length k satisfy to commutative law. In this situation the standard form of the unitary operator defined for any input string $w = x_1 \ldots x_l$ can be defined in the following way.

If $l \leq k$, then

$$sfuo(w) = U(l)\ldots U(1),$$

where $U(j) = U_{x_j \ldots x_{\min\{k+j-1,l\}}}$ for all $j = 1,\ldots,l$.

If $l > k$, then

$$sfuo(w) = U'U'',$$

where the unitary operator U' is the standard form for the product $U(l)\ldots U(l-k+1)$ constructed by the rules, similar to the ones that has been defined for the model MO-1QFA, and $U'' = U(l-k)\ldots U(1)$.

Since there is only finite number of input strings of the length less than k, the presence of the unitary operator U'' have no influence on the property "to be finitary" for the partition $X^+/\equiv$ of the set X^+.

IV. Conclusions

In the given paper basic models of finite quantum automata have been considered under the supposition that unitary operators satisfy to commutative law. The classes of languages accepted by these models either with given probability, or with given mistake have been characterized in terms of constructed partition of the set of all input strings.

Detailed analysis of these classes of languages is tightly connected with investigation the structure of the set of all finitely generated commutative semigroups of special unitary operators, acting in the complex n -dimensional Euclidean space. This is the subject for future research.

It is worth to note that this problem is sufficiently hard even in the case when $n = 2$, i.e. for 1-qubit finite quantum automata.

References

[1] G. Moore, and J. Crutchfield, "Quantum automata and quantum grammars," in Theor. Comput. Sci., vol. 237, pp. 257-306, 2000.

[2] A. Brodsky, and N. Pipenger, "Characterizations of 1-way quantum finite automata" in Journal of Computing, No. 5, pp. 1456-1478, 2002.

[3] A. Kondacs, and J. Watrous, "On the power of quantum finite state automata", in Proceedings of the 38th IEEE Symposium on Foundations of Computer Science, pp. 66-75, 1997.

[4] A. Ambainis, A. Kikusts, and M. Valdats, "On some class of languages recognizable by 1-way quantum automata", in LNCS, vol. 2010, pp. 75-86, 2001.

[5] A. Nayak, "Optimal lower bounds for quantum automata and random access codes", in Proceedings of the 40th Annual IEEE Symposium on Foundations of Computer Science, pp. 369-377, 1999.

[6] A. Bertoni., C. Mereghetti, and B. Palano, "Quantum computing: 1-way quantum automata", in LNCS, vol. 2710, pp. 1-20, 2003.

[7] A. Ambainis, M. Beaudry, M. Golovkins, A. Kikusts, M. Mercer, and D. Therian, "Algebraic results on quantum automata", in LNCS, vol. 2996, pp. 93-104, 2004.

[8] A. Belovs, A. Rosmanis, and J. Smotrovs, "Multi-letter reversible and quantum finite automata", in LNCS, vol. 4458, pp. 60-71, 2007.

[9] C.P. Williams. Explorations in Quantum Computing. Second Edition. Springer-Verlag London Limited, 2011.

Багатоканальні Системи Збору Даних

Марія Семаньків
кафедра інформатики
Прикарпатський національний університет
імені Василя Стефаника
Івано-Франківськ, Україна
dlyamarii@gmail.com

Даріуш Саля
кафедра управління
Університет AGH
Краків, Польща
dsala@zarz.agh.edu.pl

Multi-Channel Data Acquisition System

Maria Semankiv
dept. of Computer Science
Vasyl Stefanyk Precarpathian
National University
Ivano-Frankivsk, Ukraine
dlyamarii@gmail.com

Dariusz Sala
dept. of Enterprise Management
AGH University
Krakow, Poland
dsala@zarz.agh.edu.pl

***Анотація*—розглянуто питання застосування аналого-цифрових перетворювачів в системах збору даних; визначено поняття системи збору даних; розглянуто характеристики і архітектури побудови аналого-цифрових перетворювачів; запропоновано метод зсуву області інтегрування для АЦП Монте-Карло.**

***Abstract*—the application of analog-digital converters in data acquisition system is presented; concept of data collection is defined; architecture and construction characteristics of analog-digital converters are considered; method bias field for integrating ADC Monte Carlo is proposed.**

Ключові слова—аналого-цифровий перетворювач; системи збору даних; метод Монте-Карло

Keywords—analog-to-digital converter; data collection system; Monte Carlo

I. Вступ

При побудові пристроїв, що використовують обчислювальну техніку для управління і контролю складними процесами, виникає питання обробки в реальному масштабі часу швидкозмінні сигнали, які будуть надходити одночасно від багатьох джерел і давачів, та відновлювати аналогову інформацію і розподіляти її між різними виконавчими пристроями. Для цього слід використовувати аналого-цифрові перетворювачі багатоканальної інформації, що відрізняються наступними характеристиками: швидкодією, собівартістю, точністю перетворення і т.д. Пристрої, які здійснюють нормалізацію і аналого-цифрове перетворення сигналів з подальшим введенням інформації в комп'ютер, називають системами збору даних (СЗД) [1].

Системи збору даних здійснюють функцію перетворення первинних вхідних сигналів від одного або декількох вимірювальних перетворювачів в еквівалентні цифрові сигнали, які придатні для подальшого оброблення, відображення інформації або використання в системах управління. Діапазон застосувань систем збору даних широкий — починаючи з простого поточного контролю значень однієї аналогової змінної і закінчуючи контролем та управлінням сотнею параметрів в складних технічних системах. Відповідно існують як досить дешеві системи збору даних прямого перетворення, так і складні багатоканальні системи, що забезпечують дуже високу точність та надійність[2].

II. Складові системи збору даних

Для здійснення збору даних в сучасних автоматизованих системах управління технологічним процесом використовують персональні комп'ютери, програмовані логічні контролери, аналого-цифрові перетворювачі (АЦП) та інше спеціалізоване обладнання. Вибір комплексу технічного забезпечення для системи збору даних проводиться на основі відомостей про технологічний процес і обладнання, значень параметрів технологічного процесу та їх відхилення, вимог до системи автоматизації та типу давачів і виконавчих механізмів, з врахування швидкості передачі інформації, та вимог, що ставляться до точності вимірювання параметрів. Коректність отриманих результатів визначається компонентами системи збору даних, до яких входять:

давачі;

виконавчі механізми;

пристрої узгодження сигналів;

шини передачі даних;

обладнання для збору даних: програмовані логічні контролери, аналого-цифрові перетворювачі (АЦП) та інше спеціалізоване обладнання);

персональний комп'ютер;

спеціалізоване програмне забезпечення.

При виборі складу технічного забезпечення автоматизованої системи управління технологічним процесом враховують їх вартість, тому структура простою системи збору даних може складатись з давачів, виконавчих елементів, системи узгодження сигналів, що підключені безпосередньо до одного з портів персонального комп'ютера.

Основу вище вказаної системи складає АЦП. Існують три основні типи АЦП: інтегруючий, послідовного наближення і паралельний. Інтегруючий АЦП усереднює вхідний сигнал за часом, є найточнішим, але і "повільним". Час перетворення інтегруючого АЦП лежить в діапазоні від 0,001 до 50 с і більше, похибка становить 0,1-0,0003%. Похибка АЦП послідовного наближення дещо більша (0,4-0,002%), але час перетворення - від ~ 10мкс до ~ 1 мс. Паралельні АЦП - найшвидкодіючі, але і найменш точні: їх час перетворення порядку 0,25 нс, похибка - від 0,4 до 2% [2].

В системах збору даних не ставлять такі високі вимоги до частоти дискретизації, як в цифрових осцилографах, і до роздільної здатності, як в цифрових мультиметрах. Тому в даних системах переважно використовуються АЦП послідовного наближення та сигма-дельта АЦП, що належать до класу інтегруючих перетворювачів.

АЦП паралельного типу (прямого перетворення) володіють найвищою швидкодією, яка визначається швидкодією компараторів і затримками в логічному дешифраторі. Недоліком їх є необхідність великої кількості компараторів. Кількість можливих кодових комбінацій (а отже, і загальна кількість компараторів) дорівнює 2^m-1, де m — кількість розрядів АЦП.

АЦП послідовного наближення має дещо меншу швидкодію, але істотно більшу розрядність (роздільну здатність). В ньому використовується тільки один компаратор, максимальна кількість спрацьовувань якого за один цикл вимірювання не перевищує кількості розрядів перетворювача. АЦП послідовного наближення розширили діапазон частот дискретизації до мегагерц [3].

Першими для високоточної реалізації алгоритмів аналого-цифрового перетворення вимірювальних сигналів використовувалися інтегруючі АЦП (ІАЦП). Поєднання високої точності і завадостійкості ІАЦП дозволили вирішити ряд проблем боротьби з перешкодами у вимірювальних каналах. Серед завадостійких ІАЦП, що поєднують аналогові і цифрові методи, лідерами є сигма-дельта АЦП, які реалізують алгоритми аналого-цифрової фільтрації для зменшення шуму квантування.

Для підвищення функціональних можливостей АЦП запропоновано аналого-цифрові перетворювачі Монте-Карло. Метод Монте-Карло дозволяє вирішувати завдання моделювання випадкових процесів шляхом статистичних досліджень на основі аналітичної формалізації об'єкту та здійснювати перетворення інтегралу вхідної функції в цифровий код в реальному часі. Для цього за допомогою генератора випадкових чисел формуються випадкові відліки, значення яких порівнюються із значенням вхідного параметру, внаслідок чого обчислюється сукупність позитивних результатів статистичного дослідження вхідного параметру. Така реалізація полягає у проведенні статистично незалежних циклічних досліджень із накопиченням вказаних результатів порівняння. Характеристичною для методу Монте-Карло є оцінка маточікування дискретизованих значень випадкової величини вхідного параметру, яка визначає інтеграл функції перетворення за обчисленою сумою відліків позитивних результатів порівняння, внаслідок чого формується результат аналого-цифрового перетворення в області визначення, заданій відомим рівномірним імовірнісним розподілом псевдовипадкового опорного сигналу.

III. Характеристики аналого-цифрових перетворювачів

Від характеристик АЦП залежить кінцевий вибір користувача і їх область застосування. Для цифрової частини АЦП важливим є швидкість передачі даних, яка може перебувати в діапазоні від декількох десятків гц до декількох ГГц. Розрядність даних на виході АЦП – у сучасних пристроях від 8 до 32 біт даних. З боку аналогових характеристик важливе відношення сигнал-шум (SNR) і динамічний діапазон, який вільний від паразитних складових, що виражаються, як правило, в децибелах (типові рівні SNR лежать в межах 70-80 дБ). Важливими параметрами при виборі АЦП є кількість каналів в одному корпусі, інтерфейс обміну з цифровим вузлом і, зрозуміло, вартість компонента [2-3].

Точність перетворення залежить від:

– похибки квантування і похибок, що створюються електронними вузлами, такими, як ЦАП, компаратор, джерело опорної напруги;

Точність вимірювання, як і точність встановлення будь-якого процесу в електронних вузлах, зменшується по мірі збільшення швидкості перетворення. Похибка при цьому збільшується за логарифмічною залежністю, оскільки вона в основному залежить від сталих часу встановлення сигналу.

Похибка перетворення і швидкодія аналого-цифрового перетворювача послідовного наближення визначається в основному параметрами ЦАП (роздільною здатністю, лінійністю, швидкодією) і компаратора (порогом чутливості, швидкодією). Точність АЦП з паралельним перетворенням обмежується точністю і стабільністю кожного компаратора. Похибка інтегруючих перетворювачів визначається нестабільністю порогу спрацювання компаратора. Похибки перетворення однотактного інтегруючого АЦП з пилоподібною напругою поділяються на два основні класи: похибки, що

пов’язані з параметрами пилоподібної напруги, та похибки, що викликані неточністю тактування. Точність перетворення двотактного інтегруючого АЦП не залежить від ємності конденсатора, ні від частоти тактового генератора, оскільки їх вплив однаково позначається на час повернення до нуля. Істотною перевагою методу двотактного інтегрування полягає в тому, що він придатний для створення перетворювачів, що забезпечують зменшення вхідних завад при деяких частотах. Похибка АЦП Монте-Карло визначається кількістю згенерованих на області інтегрування множини випадкових чисел та якістю рівномірності їх розподілу на даній області [4].

– динамічної похибки.

Виникнення динамічних похибок викликано дискретизацією сигналів, що змінюються в часі. Можна виділити наступні параметри АЦП, що визначають її динамічну точність: максимальна частота дискретизації, час перетворення.

АЦП двотактного інтегрування мають високу точність і високу роздільну здатність, а також мають порівняно просту структуру. Це дає можливість виконувати їх у вигляді інтегральних мікросхем. Основний недолік таких АЦП - великий час перетворення.

В АЦП послідовного наближення час перетворення визначається кількістю тактів, які в свою чергу залежать від розрядності АЦП. Але слід відмітити, що робота АЦП послідовного наближення має особливість, яка пов'язану з перехідними процесами у внутрішньому ЦАП. Теоретично, напруга на виході ЦАП для кожного з n тактів перетворення має встановлюватися за однаковий проміжок часу. Але насправді цей проміжок в перших тактах значно більше, ніж в останніх. Тому час перетворення 12-розрядного АЦП послідовного наближення більш, ніж в два рази перевищує час перетворення 6-розрядного АЦП даного типу.

Для паралельного АЦП цифровий код, що відповідає вхідній аналоговій напрузі, появляється на виході майже миттєво (якщо не враховувати невеличкої затримки, що пов’язана з роботою компараторів).

Для АЦП Монте-Карло частота перетворення визначається динамічними характеристиками ЦАП і компаратора. Час перетворення являється сталою величиною та дорівнює

$$t_{np} = 2^n T \qquad (1)$$

де n – розрядність АЦП, T – період слідування тактових імпульсів, що відповідає часу вибірки одного кванта [5].

IV. Метод зсуву області інтегрування в АЦП Монте-Карло

Розроблено ряд архітектурних рішень аналого-цифрового перетворення Монте-Карло. архітектурного вирішення та апаратурних і грошових витрат.

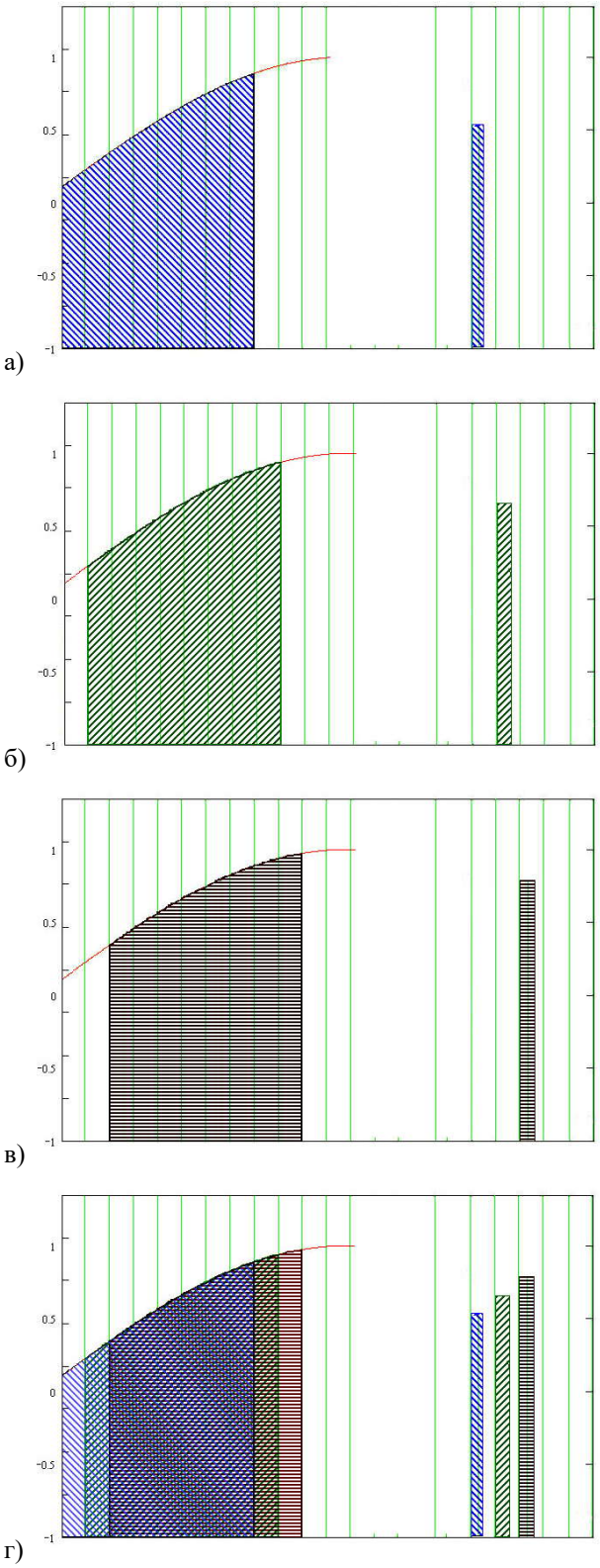


Рис. 1. Графічне подання методу зсуву області інтегрування та усереднені обчисленні значення для кожної області.

Вони реалізують багатовимірне інтегрування за умови співрозмірності архітектурного вирішення та апаратурних і грошових витрат.

Проведено дослідження статистичних характеристик випадкових чисел, які утворені відомими генераторами випадкових чисел за допомогою пакету статистичних тестів (Diehard, пакет NIST STS, критерії Колмогорова-Смірнова). Результати проведених досліджень на визначення типу розподілу послідовності псевдовипадкових чисел, сформованих методом генерування псевдовипадкових чисел Ґалуа на основі циклічних зсувів, показали найвищу якість випадковості отриманих чисел. Рівномірність розподілу сформованих послідовностей псевдовипадкових чисел вказує на ефективність використання даного генератора в складі аналого-цифрового перетворювача Монте-Карло. Використання методу генерування псевдовипадкових чисел Ґалуа на основі циклічних зсувів дозволить регуляризувати структуру аналого-цифрового перетворювача Монте-Карло, підвищити швидкодію та технологічність виробництва і зменшити вартість продукту.

Для підвищення точності та швидкодії АЦП даного типу запропоновано метод зсуву області сканування-інтегрування (рис. 1.)

Якщо на періоді перетворення (область сканування-інтегрування) сформувати 2^n значень псевдовипадкового характеру і відповідно здійснити стільки ж порівнянь, то даний тип АЦП відповідатиме за швидкодією та простотою технічної реалізації до послідовних, проте зі значно вищою точністю. АЦП Монте-Карло сканує сигнал на всій області інтегрування порівнюючи вхідний сигнал з генерованими випадковими числами та визначаючи усереднене значення вхідного сигналу на всій області дослідження.

Якщо область сканування зміщувати на один такт, зберігаючи попередні значення порівняння, то даний тип АЦП за швидкодією в один такт відповідатиме паралельним АЦП, проте зі значно простішою технічною реалізацією та розрядністю (меншою похибкою квантування за рівнем). В даному випадку використано один компаратор, на відміну від 2^n компараторів для паралельного АЦП.

Якщо область сканування зміщувати на n тактів, то даний тип АЦП за швидкодією відповідатиме АЦП послідовного наближення, проте з вищою точністю. Адже зміна вхідного сигналу буде врахована внаслідок усереднення вхідного сигналу [5-6].

На рис. 1. графічно зображено метод зсуву області інтегрування для АЦП Монте-Карло. Подано графіки областей сканування, які зміщенні на один такт (рис 1.а-в.). На рис. 1.г всі три області показано одночасно на одній площині. До кожної області інтегрування сигналу показано усереднене значення – результат інтегрування.

Однієї з складових частин аналого-цифрових перетворювачів даного типу є генератори випадкових чисел, якість рівномірності розподілу яких визначаються точність перетворення для перетворювачів даного типу.

V. ВИСНОВКИ

При виборі АЦП враховують область їх використання та технічні параметри, які б задовольняли користувача. Для багатоканальних систем збору даних запропоновано АЦП Монте-Карло, які забезпечують достатню швидкість та точність перетворення для вибраної сфери використання.

При застосуванні методу зсуву області сканування-інтегрування з фіксованою шириною 2^n та величиною кроку зміщення в розмірі одного такту, даний тип АЦП забезпечить час перетворення як для паралельних АЦП, проте зі значно простішою технічно реалізацією. Для АЦП Монте-Карло за такт перетворення формується результат порівняння вхідного та опорного сигналів. Зсув області інтегрування на n тактів віднесе даний тип аналого-цифрових перетворювачів за часом перетворення до перетворювачів послідовного наближення, проте з відслідкуванням зміни вхідного сигналу, що позитивно впливає на точність перетворення.

ЛІТЕРАТУРА REFERENCES

[1] М.М. Гельман “Системные аналого-цифровые преобразователи и процессоры сигналов,” М.: Мир, 1999.

[2] В.А. Романов “Параметры АЦП общего применения и быстродействующих АЦП,” *Электронные компоненты и системы*, 2001, № 8, с. 24-25.

[3] В.В. Островерхов “Динамические погрешности аналого-цифровых преобразователей,” Л.: Энергия, 1975.

[4] М.В. Семаньків “Переваги аналого-цифрових перетворювачів Монте-Карло,” *Методи та засоби кодування, захисту й ущільнення інформації*: п’ята міжнар. наук.-практ. конф., Вінниця, 2016. – С.140-142.

[5] M. Semankiv, D. Dumenchuk “Comparative characteristics of ADC’s Monte Carlo,” *52 Konferencje Studenckich Kół Naukowych*, Kraków, 2015. - 295 s.

[6] М.В. Семаньків “Підвищення технічних характеристик багатоканальних аналого-цифрових перетворювачів,” *Інформаційні технології та комп’ютерна інженерія*: матеріали статей п’ятої міжнар. наук.-практ. конф., м. Івано-Франківськ, 2015 р.–С.102-103.

Оптимізація Пошукового Алгоритму Боєра-Мура-Хорспула для Бінарних Даних

Ігор Завадський
Факультет кібернетики і комп'ютерних наук
КНУ ім. Шевченка
Київ, Україна
ihorza@gmail.com

Василь Остапчук
Факультет кібернетики і комп'ютерних наук
КНУ ім. Шевченка
Київ, Україна
vasylysk@gmail.com

The Boyer-Moore-Horspool Pattern Matching Algorithm Optimization for Binary Data

Igor Zavadskyi
Dept. Of Cybernetics and Computer Science
T. Shevchenko Kyiv National University
Kyiv, Ukraine
ihorza@gmail.com

Vasyl Ostapchuk
Dept. Of Cybernetics and Computer Science
T. Shevchenko Kyiv National University
Kyiv, Ukraine
vasylysk@gmail.com

***Анотація*—У статті описано метод пошуку підрядка у двійковому рядку. Метод являє собою адаптацію та оптимізацію алгоритму пошуку підрядка Боєра-Мура-Хорспула для бінарних даних. Середня величина зсуву вікна пошуку у запропонованому методі суттєво перевищує аналогічний показник для інших відомих адаптацій алгоритмів сімейства Боєра-Мура для бінарних даних.**

***Abstract*—The method of exact pattern matching on binary alphabet is presented. It represents the adaptation and optimization of well-known Boyer and Moore pattern matching algorithm for binary alphabet. The average search window shift length is essentially greater than in other known adaptation of Boyer and Moore algorithm.**

Ключові слова—алгоритм Боєра-Мура; пошук патерна; бінарні дані; зсув вікна пошуку

Keywords—Boyer-Moore; Pattern matching; binary data; search window shift

I. Вступ

Однією з найактуальніших задач кібернетики протягом останніх десятиліть є розробка ефективних алгоритмів пошуку підрядка (патерна) в тексті. Більша частина відомих на сьогодні алгоритмів в цій галузі в тій чи іншій мірі використовує ідеї алгоритму Боєра-Мура [1]. Йдеться про три основні ідеї:

1. Сканування тексту зліва направо, порівняння символів вікна пошуку справа наліво.
2. Зсув вікна пошуку за евристикою стоп-символу.
3. Зсув вікна пошуку за евристикою його суфікса.

Алгоритм Боєра-Мура-Хорспула [2] являє собою спрощений варіант алгоритму Боєра-Мура. Він не включає в себе евристику суфікса. Цей алгоритм і дотепер залишається одним з найефективніших для текстів на великих алфавітах (64 символи і більше).

Однак чи не найголовніший недолік сімейства алгоритмів Боєра-Мура полягає в тому, що на малому алфавіті величина зсуву за евристикою стоп-символу є в загальному випадку невеликою. Розглянемо більш конкретний приклад, а саме алфавіт, що складається з 0 та 1, тобто бінарні дані. Використання евристики стоп-символу на таких даних, крім крайніх випадків (патерн складається з довгих послідовностей вигляду 0…0 та 1…1), даватиме мінімальну величину зсуву, тобто 1 символ.

Цю проблему зазвичай вирішують, розглядаючи при порівнянні замість конкретного біта групу бітів, тобто машинне слово (надалі ми вважатимемо машинним словом один байт, що складається з 8 бітів, хоча загалом його величину можна варіювати як завгодно). Так, у [3] розглянуто варіант алгоритму пошуку патерна за алгоритмом Боєра-Мура для бінарних даних, який не використовує ніякі бітові маніпуляції, а працює лише з байтами. Тим самим значно зменшується кількість необхідних порівнянь. Однак запропонований у [3] метод недостатньо оптимізований і в цій статті ми опишемо, як виправити його вади.

Слід зауважити, що є кілька алгоритмів пошуку підрядка, які на бінарному алфавіті ефективніші за алгоритм Боєра-Мура-Хорспула, зокрема алгоритм Shift-And [4] для коротких патернів та Hashq [5] для патернів середньої довжини. Однак у них, як і в оригінальних версіях алгоритмів [1] і [2], не враховується можливість одночасної обробки байтів тексту. Більше того, виконання побітових операцій є невід'ємною складовою алгоритмів [4]–[5], і тому існування їхніх «байтових» версій видається проблематичним.

Ще один підхід може полягати в опрацюванні байтів як елементів 256-символьного алфавіту, однак його очевидним недоліком є неможливість пошуку патернів, що включають частини байтів.

II. Ідея методу

Нехай *t[i]* і *p[i]* - *i*-ий біт тексту та патерна відповідно. Позначатимемо *i*-ий байт тексту через text[i], а *j*-ий байт патерну після відступу на *sh* біт вправо – через *pat[sh, j]*. Оскільки довжини патерну й тексту можуть бути не кратними байту (рис. 1.), для патерна використовується маска *mask[sh,j]*, як і в [3].

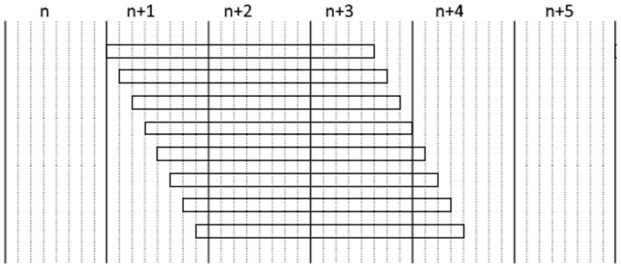


Рис. 1. Зсуви патерна відносно меж байта тексту

Маска показує, які біти конкретного байта патерна при порівнянні є значущими. Наприклад, *mask[0,3] = 11111000, mask[0,2] = 11111111, mask[2,1] = 00111111.* Операція перевірки збігу *i*-го байту тексту з *j*-м байтом патерна з відступом *sh* виглядає так:

(text[i] XOR pat[sh,j] AND mask[sh,j]) == 0.

Згідно з алгоритмом Боєра-Мура, величина зсуву вікна пошуку за евристикою стоп-символу залежить від позиції найбільш правого входження останнього символу вікна пошуку в патерн, а в середньому – від частоти входження певного символу алфавіту в патерн. Для бінарних алфавітів символом слід вважати байт. Однак при порівняні останнього байта в патерні ми обмежені його маскою і порівнюємо насправді не 8 бітів, а, залежно від розміщення патерна, від 1 до 8. Середня величина зсуву за такою евристикою є набагато меншою, ніж у разі порівняння 8 бітів. Це і є головним недоліком методу [3].

Основна ідея даної роботи полягає в тому, щоб базувати евристику стоп-символу не на останньому байті, який може бути неповним, а з передостанньому, який гарантовано є повним. Оскільки неповний останній байт до перевірки не залучається, використання евристики суфікса є не можливим. Тому наш алгоритм є модифікацією для бінарних текстів алгоритму Боєра-Мура-Хорспула, у якому не використовується евристика суфіксу, на відміну від [3], де до бінарних даних адаптовано саме класичний алгоритм Боєра-Мура.

Пояснимо переваги нашого методу детальніше. Ймовірність того, що в конкретній позиції випадкового тексту зустрінеться конкретний 8-бітовий байт випадковим чином згенерованого патерну становить *1/256*. Якщо припустити, що всі зсуви останнього байта патерну відносно байта тексту відбуваються з однаковою ймовірністю, то ймовірність того що останній не повний байт зустрінеться в тексті дорівнюватиме *1/8(1/2+1/4+1/8+…+1/256) = 255/256*8 = 31,875/256.* Тобто при порівнянні неповного байта ймовірність того, що він трапиться в певній позиції патерна збільшується, а отже, зменшується середня величина зсуву.

Однак у випадку, коли розмір патерна менший або рівний двом байтам, середня величина зсуву для нашого алгоритму є меншою, ніж для алгоритму [3], оскільки передостанній байт може бути неповним і сенс використання його як бази для евристики зсуву втрачається.

III. Алгоритм

На стадії передобчислень будуються таблиці *delta1[sl,b]* (*sl* ≤*bytesize i* $0 \le b < 2^{sl}$) та *Correction[sh,j]*. Перша з них – це таблиця довжин зсуву вікна пошуку за евристикою стоп-символу. *delta1[sl, b] = paternLength−l*, де *l* це індекс останнього біта найправішого входження в патерн *sl*-бітового значення *b.* Для прикладу, нехай *sl* = 4 і *P = 0010101011101101*, найправіше входження *b=5=0101* підкреслене, тоді *delta1[*4,5*] = 16-9=7.* Але слід розглянути ще один варіант, коли *b* не трапляється в патерні, але суфікс *b* є префіксом патерна. Для прикладу при *sl=5* і *b=17=10001*, *b* не входить в патерн, але його суфікс є префіксом патерна. Тоді патерн не може бути зсунутий на свою максимальну довжину, тобто *delta1[*5,17*]=16–3=13*, а не *16.* Алгоритм ініціалізації таблиці *delta1* наведено

нижче (значення по замовчуванню для *delta1[sl,b] = paternLength*).

Для збігу з префіксом патерна:

```
for i=1 to sl-1
  for всі значення b такі, що p[1]…p[i] є суфікс b
    delta1[sl,b] = paternLength-i
  В усіх інших випадках:
for i =1 to paternLength-sl
  b=p[i]…p[i+sl-1]
  delta1[sl, b] = paternLength–i-sl+1.
```

Також позначимо через *last[sh]* номер останнього байту патерна після зсуву *sh*. Формально *last[sh] = (sh+ patternLength -1)/bytesize +1*. Для прикладу на рис 1. *last[sh] = 3* для *sh=0,1,2,3* і *last[sh] = 4* для *sh = 4,5,6,7*.

Через *sl* (рис. 2.) позначимо кількість біті в останьому байті патерна після зсуву *sh*: *sl = (sh+patternLength-1) mod bytesize + 1*. *sl* буде обчислюватися в ході виконання програми, тому для цієї величини таблиця не створюється.

Рис. 2. Схематичне зображення *sl* та *sh*

В оригінальному алгоритмі Боєра-Мура [1] таблиця *delta1* застосовувалася незалежно від того, який байт вікна пошуку не дорівнює відповідному байту в патерні, тому що *delta1* показувало насправді величину, на яку поточний покажчик на текст *i* міг би бути збільшений, а не кількість символів патерну, на які можна зсунути вікно пошуку. У бінарному варіанті *delta1* дорівнює максимально можливому зсуву для вікна пошуку, якщо рахувати від кінця вікна пошуку. Тому потрібно ввести ще одну таблицю констант *Correction[sh,j],* яку будемо використовувати, якщо не співпадіння відбулося не на останньому байті. *Correction[sh,j]* показує на скільки далеко від кінця патерна відбулося не співпадіння і на скільки вліво (тобто назад) ми повинні зсунути патерн, якщо не співпадіння відбулося на *j*-ому байті:

Correction[sh,j] = sl +(last[sh]−1−j) bytesize.*

Наведемо приклад, коли потрібно використовувати таблицю *Correction*. Припустимо, відбувається порівняння *i*-ого байта вікна пошуку та відповідного патерна, як показано на рис. 3. Позначимо цей байт тексту через *x*. У випадку, коли байт *x* не збігається з відповідним байтом патерна, а ми повинні зсунути вікно пошуку.

Рис. 3. Приклад використання таблиці *Correction*

Для цього ми використовуємо таблицю *delta1*, яка показує величину зсуву вікна пошуку у випадку, коли останній байт вікна пошуку не збігається з останнім байтом патерна. Ця величина дорівнює відстанню від кінця патерна до найправішого входження байта *x* в патерн, або довжині патерна, якщо такого входження немає. Припустимо, що байт *z* патерна є найправішим входженням байта *x* в патерн, тоді *delta1[sl,x] = dist(z), dist(z)* – це відстань від кінця патерна до *z*. Але якщо зсунути патерн на таку відстань, байт *x* не збігатиметься з байтом *z*, тому потрібно скоректувати зсув таким чином, щоб байт *x* опинився над байтом *z*. Зрозуміло, що величина зсуву тоді дорівнює *dist(z) – dist(x)*. Але *x* – це *i*-ий байт поточного вікна пошуку, тоді *dist(x) = Correction[sh, i]*, звідси величина зсуву можна обчислити так:

$$dist(z) - dist(x) = delta1[sl,x] - Correction[sh, i] = delta1[sl,text[i]] - Correction[sh, i]$$

Зрозуміло, що за такою формулою можливий варіант коли величина зсуву буде від'ємною, наприклад, коли байт *y* дорівнює *x.* У цьому випадку відстань до байта *y* є меншою за відстань до байта *x* і для збігу *x* з *y* потрібно патерн назад. Однак даний варіант розміщення патерна був уже відкинутий або розглянутий на попередньому етапі алгоритму. Тому, щоб уникнути цього небажаного зсуву, назад в алгоритм була додано перевірка (*d1* – величина зсуву після корекції):

```
if(d1<=0)
      d1=1
```

Згідно з основною ідеєю даної роботи, на стадії пошуку байти вікна пошуку звіряються з патерном від передостаннього до першого і лише в разі збігу їх усіх перевіряється збіг останнього неповного байта вікна пошуку із відповідним байтом патерна. Таким чином, ми не оперуємо значенням останньому байту, крім випадку, коли ми перевірили уже всі байти від передостаннього до першого і повернулися до останнього.

Наведемо формальний опис алгоритму за допомогою псевдокоду.

```
sh = 0
i = last[0]
lastbit = patternLength
// Довжина тексту в байтах
while i <= (textLength-1)/byteSize+1
   // Починаємо перевірку з передостаннього байта
   j = last[sh]
   oldj = j, oldi = i
   j = j – 1, i = i - 1
```

```
sl=1+(patternLength+sh-1)mod byteSize
// indic вказує, чи збігається поточний байт тексту і патерна
// перевіряємо байти від передостанього до першого
indic = (text[i] XOR pat[sh, j])
AND mask[sh, j]
while j > 0 AND indic == 0
   i = i – 1
   j = j - 1
   if (j != 0)
      indic = (text[i] XOR
      pat[sh, j]) AND mask[sh, j]
// якщо indic = 0, то збігаються всі байти, від передостанього до першого
if indic == 0 // тоді перевіряємо останній байт
   indic = (text[oldi] XOR
   pat[sh,oldj]) AND mask[sh, oldj]
   if(indic==0)
      report match (byteSize*i+sh)
      lastbit++
   else
                  // інакше обчислюємо зсув за евристикою стоп-символу
                  i = oldi
                  j = oldj
                  b = (text[i]/1<<(byteSize - sl)) mod (1<<K)
      d1 = delta1[min(sl, K), b]
      if(d1<=0)   d1=1
      lastbit = lastbit + d1
else
   // якщо немає збігу, обчислюємо зсув за евристикою стоп-символу
   d1 = delta1[K, text[i] mod
      (1<<K)]– Correction[sh, j]
   if(d1<=0)
      d1=1
   lastbit = lastbit + d1
i = ((lastbit - 1) / byteSize) + 1
sh = (lastbit - patternLength)      mod byteSize
```

Обсяг пам'яті, необхідної для зберігання таблиці *delta1*, можна обчислити за формулою (1).

$$\sum \quad 2 \quad = 2 \quad - 2 \qquad (1)$$

Тобто він збільшується експоненційно, однак залишається допустимим для значень *byteSize* 8 або 16. Але вже при *byteSize* = 32 зберігати такий обсяг даних може виявитися складно. Тому було введено додаткову змінну *K*, яка за замовчуванням дорівнює *byteSize*, але може бути зменшена. Вона показує, скільки бітів із кожного байту беруть участь в обчисленні максимально можливого зсуву, що дає змогу під час обчислення *delta1* не обчислювати значення *sl*, які більші за *K*, і суттєво зекономити пам'ять ціною зменшення величини зсуву.

IV. Обчислювальний експеримент

Експеримент було проведено на випадковим чином згенерованих текстах довжиною 40000 бітів та патернах довжиною від 10 до 1000 бітів. На рис. 4 зображено середнє значення величини зсуву залежно від довжини патерна. Продемонстровано дані для 3 алгоритмів:

1. ABM – алгоритм Боєра-Мура, адаптований для бінарних даних, згрупованих у байти [3].
2. AHBM – алгоритм Боєра-Мура-Хорспула, адаптований для бінарних даних, згрупованих у байти.
3. AHBM-ZV – запропонований нами алгоритм.

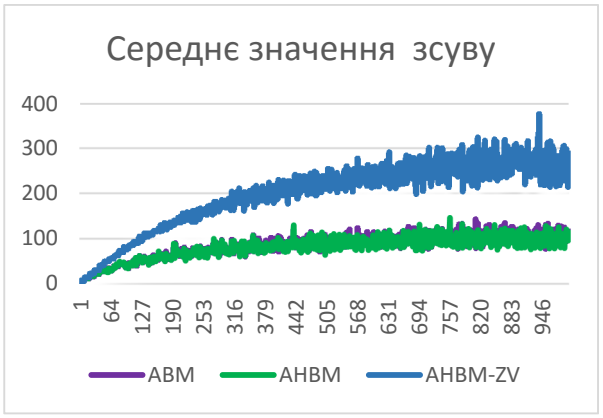


Рис. 4. Середня величина зсуву вікна пошуку

V. Висновки

Як свідчить комп'ютерний експеримент, середня величина зсуву вікна пошуку в нашому алгоритмі для достатньо довгих патернів перевищує цю величину в алгоритмі [3] більш ніж удвічі. Отже, запропонований алгоритм є ефективною адаптацією алгоритму Боєра-Мура-Хорспула для пошуку підрядка у двійковому тексті, якщо довжина підрядка перевищує 32 біти.

Зазначимо, що запропонований алгоритм не оптимальний і є кілька можливих шляхів його вдосконалення. Найбільш перспективним із них видається використання лише тих зсувів патерна, які вирівнюють його початок на межу байтів. Це дасть змогу уникнути операцій, що визначають положення початку вікна пошуку всередині байта.

Література References

[1] R.S. Boyer, J.S. Moore, A fast string searching algorithm, Comm. ACM 20 (1977) 762–772.

[2] R.N. Horspool, Practical fast searching in strings, Software — Practice and Experience 10 (1980) 501–506.

[3] S.T. Klein, M. Ben-Nissan. (2007) Accelerating Boyer– Moore Searches on Binary Texts. Proc. 12th Int. Conf. Implementation and Application of Automata (CIAA 2007), Lecture Notes in Computer Science 4783, 130–143.

[4] S. Wu, U. Manber Fast text searching allowing errors. Commun. ACM, 35(10):83–91, 1992.

[5] T. Lecroq Fast exact string matching algorithms. Inf. Process. Lett., 102(6):229–235, 2007.

New Probabilistic Model of Stochastic Context-Free Grammars in Greibach Normal Form

Yevhen Hrubiian
Student, group FI-33, Institute of Physics and Technology
NTUU "Igor Sikorsky Kyiv Polytechnic Institute"
Kyiv, Ukraine
grubian.euhen@gmail.com

Імовірнісна Модель Стохастичних Контекстно-Вільних Граматик в Нормальній Формі Грейбах

Євген Грубіян
Студент групи ФІ-33, Фізико-технічний інститут
НТУУ "Київський політехнічний інститут ім. Ігоря Сікорського"
Київ, Україна
grubian.euhen@gmail.com

***Abstract*—In this paper we propose new probabilistic model of stochastic context-free grammars in Greibach normal form. This model might be perspective for future developments and researching stochastic context-free grammars.**

***Анотація*—В роботі запропоновано нову імовірнісну модель стохастичних контекстно-вільних граматик в нормальній формі Грейбах. Ця модель може бути перспективною для подальших досліджень стохастичних контекстно–вільних граматик.**

Keywords—stochastic context-free grammar; Greibach normal form; hidden Markov models

Ключові слова—стохастична контекстно-вільна граматика; нормальна форма Грейбах; приховані моделі Маркова.

I. INTRODUCTION

Today, there is a huge interest in developing new methods of natural language processing due to a lot of issues from cryptanalysis, biology and artificial intelligence. Several ways for investigation had been proposed since works of Noam Chomsky in the mid 50s that originated development of modern formal theory of syntax. His conception of generative grammars and context-free grammars became one of the most discussing and controversy in linguistics but these ideas lay behind development of formal theory of programming languages and compilers later. Context-free grammars were proposed firstly for describing grammar of English or any other language but this approach does not cover all features of language such as semantics. Later, head-driven grammars were introduced to describe some valuable properties of languages such as semantic dependencies between words. Probabilistic models of languages were proposed by Claude Shannon to describe informational aspects of languages such as entropy but they became very relevant and actual in modern linguistics. N-gram model is one of the most accurate statistical models of language but the main drawback is that for precious learning this model for high values of N one needs a huge amount of texts in different thematic, but unfortunately, even if we have got a lot of texts it turns out to be harmful for accurate and precise modeling whole language. Because of this reason stochastic context-free grammars were introduced for more precious description of languages because context-freeness guarantees grammatical pattern of language and probability distributions on grammatical rules approach semantic structure of phrases. Stochastic context-free grammars are also useful in modeling DNA sequences, image recognition and modeling plain texts in cryptography so that investigation of stochastic context-free grammars in very important and actual in today's science.

In this paper we consider stochastic context-free grammars in Greibach normal form although significant contributions in natural language processing was associated with algorithms of learning stochastic context-free grammars in Chomsky normal form but complexity of these methods is quite high. We propose new idea of investigation of stochastic context-free grammars based on our conception of hidden Markov models with stack.

II. STOCHASTIC CONTEXT-FREE GRAMMARS

Conception of stochastic context-free grammars (SCFG) came from several problems in natural language processing such as finding the most probable sequence of the words,

learning unknown grammar from corpora of texts etc. But NLP is not only sphere of applications of SCFG, several biologists and geneticists proposed usage SCFG for modeling DNA sequences, modern pattern recognition, speech processing, image processing also use this conception.

We say that G = <N, Σ, R, S, P> is *stochastic context-free grammar* if and only if Q = <N, Σ, R, S> is context-free grammar (N – alphabet of nonterminal symbols or variables, Σ – alphabet of terminal symbols, R – set of production rules, S – start symbol) and one defines family of probability distributions P on R so that for every production in R there is probability of applying this production rule when forming some sequence of symbols such that sum of probabilities of productions starting with same symbol must be equal to 1.

There are several normal forms of grammars. The most popular are Chomsky normal form and Greibach normal form.

We say that context free grammar G = <N, Σ, R, S> is in *Chomsky normal form* (CNF) if and only if all production rules have form:

$$A \rightarrow BC$$

$$A \rightarrow a$$

Our major interest forms Greibach normal form.

We say that context free grammar G = <N, Σ, R, S> is in *Greibach normal form* (GNF) if and only if all production rules have form:

$$A \rightarrow aA_1A_2\ldots A_n$$

Where A is nonterminal, a is terminal symbol and $A_1A_2\ldots A_n$ is a probable empty sequence of nonterminal symbols except S.

The definition of SCFG in GNF might get stronger if we set number of nonterminal symbols in right hand side of any production to be not larger than 2. Productions of such grammar will be in any of the following forms:

$$A \rightarrow aA_1A_2$$

$$A \rightarrow aA$$

$$A \rightarrow a$$

It could be shown that every SCFG in GNF might be transformed into equivalent grammar in the stronger GNF. For this reason in this paper we suppose that every Greibach normal form will be in the form described above.

Every context-free grammar could be transformed to equivalent grammar in Greibach normal form in polynomial time such that languages they form are equal. The same idea is behind transforming SCFG to SCFG in Greibach normal form. We obtained algorithm for this transformation based on existing algorithm of transforming an ordinal CFG in Chomsky Normal form to context-free grammar in Greibach normal form adding several steps of rebalancing probability distributions P.

Here is the algorithm:

1. Eliminate null productions, unit productions and useless symbols from the grammar G and then construct a $G_0 = (V_0, T, R_0, S)$ in Chomsky Normal Form (CNF) generating the language $L(G_0) = L(G) - \{\varepsilon\}$.
2. Rename the variables like $A_1, A_2, \ldots A_n$ starting with $S = A_1$.
3. Modify the rules in R so that if $A_i \rightarrow A_j\gamma \in R_0$ then $j > i$
4. Starting with A_1 and proceeding to An this is done as follows:
   a) Assume that productions have been modified so that for $1 \le i \le k$, $A_i \rightarrow A_j\gamma \in R_0$ only if $j > i$
   b) If $A_k \rightarrow A_j\gamma$ is a production with $j < k$, generate a new set of productions substituting for the A_j the body of each A_j production. The transfer probabilities should split uniformly from base probability.
   c) Repeating (b) at most $k - 1$ times we obtain rules of the form $A_k \rightarrow A_p\gamma$, $p \ge k$
   d) Replace rules $A_j \rightarrow A_k\gamma$ by removing left-recursion as stated above. Transfer probabilities should also be splitted.
5. Modify the $A_i \rightarrow A_j\gamma$ for $i = n - 1, n - 2, ., 1$ in desired form at the same time change the Z production rules.

III. Scoring problem for regural grammars

Suppose we have to find probability of sequence of symbols GNF generated: $Pr(w_{1:n} \mid G)$ i.e. solve the scoring problem.

Trivial situation come up when our grammar in GNF is regular, i.e. all rules have form:

$$A \rightarrow aB$$

Regular grammars are equivalent to finite state machines (FSM). This equivalence could be set if we assume that set of states of FSM is equal to set of nonterminal symbols and every transition is described by corresponding rule of grammar with emission of terminal symbol.

We say that bivariate stochastic process $\{X_i, O_i, i=1,2,\ldots\}$ is *hidden Markov model of order n by emissions (HMMn)* if

$$Pr(X_k \mid X_{1:k-1}) = Pr(X_k \mid X_{k-1})$$

$$Pr(O_k \mid X_{1:k}) = Pr(O_k \mid X_{k-n+1:k})$$

X_k are called latent variables or hidden states of HMM and O_k are called observations.

Stochastic regular grammars are equivalent to hidden Markov models of order 2 by emissions. It follows from the fact that set of observed values corresponds to set of terminal symbols and set of latent variables corresponds to set nonterminal symbols of grammar. Second order of models is followed from fact that observed word depends on previous 2 states, however order by transitions is still first. This equivalence also requires independence of appliance every grammatical rule. Because of this fact we could apply *forward* algorithm for solving scoring problem for stochastic regular grammar. Complexity of this algorithm is $\Theta(nm^2)$, where n is

length of observed sequence and m is number of nonterminal symbols.

IV. MARKOV CHAINS WITH STACK

The case of general SCFG in GNF is much more complicated. If we try to build HMM out of SCFG in GNF we will face with some uncertainty with description of latent variables set. Naïve approach is based on setting equivalence between set of all possible sequences of nonterminal symbols that grammar could produce during inference all possible sentences. But in that case number of latent states could be very large and even unbounded because of all recursions in grammatical rules.

Let's try to overcome these difficulties. It is known fact that every context-free grammar is equivalent to some finite state machine with stack (pushdown automaton). Hence we could propose conception of *hidden Markov model with stack* for emulation SCFG in GNF.

We say that stochastic process $\{S_t,\ t=1,\ldots\}$ is a *Markov chain with stack M* if and only if:

$$Pr(S_t \mid S_{1:t-1},\ M_t) = Pr(S_t \mid S_{t-1},\ M_t)$$

Where M_t is element on the head of the stack at time moment *t*.

Finally, we say that bivariate stochastic process $\{S_i, O_i,\ i=1,2,\ldots\}$ is *hidden Markov model with stack M of order n by emissions* if and only if for any k>0 holds:

$$Pr(S_k \mid S_{1:k-1},\ M_k) = Pr(S_k \mid S_{k-1},\ M_k)$$

$$Pr(O_k \mid M_k,\ S_{1:k}) = Pr(O_k \mid M_k,\ S_{k-n+1:k})$$

As we can see, this definition differs from the origin definition of hidden Markov model because of M_k – head of the stack in conditional part of emission and transition probabilities.

In fig. 1. there is example of *Markov chain with stack*:

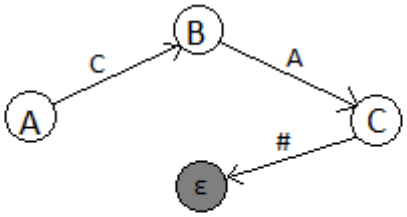


Fig. 1. Markov chain with stack

Here A,B,C denote ordinary states and ε denotes special symbol used for sudden transition to state pointed by top of the stack *M*. Letters above arrows denote states to push into stack. Note that for every transition not more than one state could be pushed into stack. This fact corresponds to conception of the *stronger GNF* described in the second chapter. Symbol # denotes popping out head of stack to which chain jumps after. Transitional probabilities are omitted because the chain in example is deterministic. If start state is A than chain will produce sequence ABCABCABC…

One could notice that stack size in this example is growing unrestrictedly but it is not common case for natural language's grammars. The model terminates when stack is empty and model reaches # symbol.

Our goal than will be setting up equivalency between stochastic context-free grammars in Greibach normal form and hidden Markov models of order 2 with stack (HMM2S).

V. SCFG IN GNF AND HMM2S EQUIVALENCY

In previous paragraph we could see that Markov chains with stack work just similar to the way stochastic context-free grammars in Greibach normal form infers sequences of nonterminal symbols. Indeed, if we consider state space of Markov chain with stack to be the set of nonterminal symbols in SCFG in GNF and transitions between states would show action of the rules of grammar in such way that first nonterminal symbol in right hand side of the rule would be state that the Markov chain with stack jumps to and letters above arrows are the rest of right hand side of the rule than we could negotiate equivalency between nonterminal inferences in stochastic context-free grammars in Greibach normal form and Markov chains with stack.

A bit tricky step is construction of Hidden Markov model with stack that considers terminal productions in each step of inference some sentence. We need to deal with ambiguity of grammatical rules that could produce different terminal symbols but transfer to the same state. This problem could be solved by adding dependence to emitted terminal symbol by current state and previous. Hence we must build Hidden Markov Model of the second order. Also we should take to consideration that in general case of stochastic context-free grammar in Greibach normal form the first letter of each production would not identify uniquely state the system jumps. To solve this we propose to associate leftmost different sequences of nonterminal symbols to states of HMM. For example if we have rules:

$$A \rightarrow aCD$$

$$B \rightarrow bCE$$

We will than have got states {A, B, CD, CE} for our HMM. In this case for each emission there would be no ambiguity to determine which grammatical rule has produced it. Note that in every step of production there is emission of terminal symbol. If some production in the form:

$$A \rightarrow a$$

is applied the system will pop out the head state in the stack and will jump to that state. Size of the stack will reduce than.

Thus we have obtained the hidden Markov model of the second order.

VI. Open problems, ways of solving and motivation for this model

As we could see, stochastic context-free grammars in Greibach normal form are equivalent to Hidden Markov models of the second order with stack. So that we could develop algorithms for finding particular sentence probability, probability of sequence of latent states (grammatical rules applied) and even algorithms for learning stochastic context-free grammars. The main idea why construction proposed in this article might helpful is the fact that there is a lot of algorithms developed for ordinary hidden Markov models and they are quite scalable and easily transformable for appliance for more complex Markov models.

For example for finding most probable latent state sequence there is Viterbi algorithm. For finding observed sequence probability there is forward algorithm. For learning model parameters there is Baum-Welch algorithm based on forward-backward algorithm. Complexity of these algorithms is $O(nm^2)$ which is quite better than complexity of algorithms for stochastic context-free grammar in Chomsky normal form. For example the algorithm inside-outside for learning grammar in Chomsky normal form has complexity $O(n^3m^3)$. This is the reason why equivalence of stochastic context-free grammars in Greibach normal form and hidden Markov models of the second order with stack may be very perspective. However, there might be a lot of work in developing theory of hidden Markov models with stack we believe that this ideas will be helpful for further researching in natural language processing, biology and computer science.

VII. Acknowledgment

Author conveys a big thanks to his supervisor Andriy Fesenko, collective of cathedra of Mathematical methods of informational security in Institute of Physics and Technology and Georgii Riabov, fellow researcher at the Department of the theory of stochastic processes, Institute of Mathematics, NAS of Ukraine.

References

[1] http://www.stat.berkeley.edu/%7Eterry/Classes/s246.2002/Week9/week9b.pdf

[2] Rafael C. Carrasco and Jose Oncina, “Learning deterministic regular grammars from stochastic samples in polynomial time”, Departamento de Lenguajes y Sistemas Inform´aticos, Universidad de Alicante, E-03071 Alicante

[3] Zoubin Ghahramani “An Introduction to hidden Markov models and Bayesian networks”, Gatsby Computational Neuroscience Unit, University College London, London WC1N 3AR, England

[4] K. Lari; S. J. Young, (1990). "The estimation of stochastic context-free grammars using the inside-outside algorithm". Computer Speech and Language.

[5] K. Lari; S. J. Young, (1991). "Applications of stochastic context-free grammars using the inside-outside algorithm". Computer Speech and Language.

[6] J. E. Hopcroft; J. D. Ullman, (1979). Introduction to Automata Theory, Languages, and Computation. Addison-Wesley.

[7] C. de la Higuera. Ten open problems in grammatical inference. In Sakakibara et al. [54], pages 32–44

[8] R.L. Kashyap: Syntactic decision rules for recognition of spoken words and phrases using stochastic automaton. IEEE Trans. on Pattern Analysis and machine Intelligence 1(1979) 154–163

[9] Y. Sakakibara, M. Brown, R. Hughley, I. Mian, K. Sjolander, R. Underwood, D. Haussler: Stochastic context-free grammars for tRNA modeling. Nuclear Acids Res. 22 (1994) 5112–5120

[10] Diego Linares, Jose-Miguel Bened, and Joan-Andreu Sanchez “A Hybrid Language Model based on Stochastic Context-free Grammar“, Pontificia Universidad Javeriana – Cali Calle 18 No. 118-250 Av. Canasgordas, Cali (Colombia), DSIC- Universidad Polítecnica de Valencia Camino de Vera s/n, 46022 Valencia (Spain)

[11] Jose L.Verd Verdu-Mas, Jorge Calera-Rubio, and Rafael C. Carrasco “Offspring-annotated probabilistic context-free grammars”, Departamento de Lenguajes y Sistemas Informaticos Universidad de Alicante, E-03071 Alicante, Spain

[12] http://www.iitg.ernet.in/gkd/ma513/oct/oct18/note.pdf

[13] R. Lawrence, A. Rabiner. “Tutorial on hidden Markov models and selected applications in speech recognition”. Proceeds of the IEEE, vol.77, no.2, February1989.

[14] Noam Chomsky, (1956). "Three models for the description of language.". IRE Transactions on Information Theory.

[15] Baum, L. E.; Petrie, T.; Soules, G.; Weiss, N. (1970). "A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains". The Annals of Mathematical Statistics.

[16] Chomsky, Noam (June 1959). "On certain formal properties of grammars". Information and Control

[17] Sheila Greibach, (January 1965). "A New Normal-Form Theorem for Context-Free Phrase Structure Grammars". Journal of the ACM. 12

[18] Norbert Blum, Robert Koch, (1999). "Greibach Normal Form Transformation Revisited". Information and Computation. 150

[19] Michael A. Harrison (1978). Introduction to Formal Language Theory. Addison-Wesley.

[20] L.E. Baum, T. Petrie, (1966). "Statistical Inference for Probabilistic Functions of Finite State Markov Chains".

[21] D. Jurafsky, J.H. Martin, (2009). Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition. Upper Saddle River, N.J: Pearson Prentice Hall.

СЕКЦІЯ 4.

Цифрова обробка та ущільнення інформації

SECTION 4.

Digital information processing and compression

Arbitrary Meshes to Simulate Natural Video Objects

Olena Osharovska
Television and Radiobroadcasting Department
O.S. Popov's ONAT
Odesa, Ukraine
osharovskaya@gmail.com

Нерівномірні Сітки для Моделювання Натурних Відео Об'єктів

О.В. Ошаровська
кафедра телебачення та радіомовлення,
ОНАЗ ім. О. С. Попова
Одеса, Україна
osharovskaya@ gmail.com

***Abstract*—The paper considers the option of creating three-dimensional mesh objects from images natural scenes. The image contains both areas with a smooth change in brightness and chromaticity, and areas of detail. To divide the scene into sections with different details, the gradient method was used. A comparison is made with respect to the digital stream being created when the luminance component is represented by vertices of an equidistant uniform mesh and with an adaptive no equidistant mesh. For the case of an adaptive mesh, Delaunay triangulation is used for the threshold values of the gradient module. It is shown that the gain in compression of a luminance signal by an adaptive mesh increases when coding images of high and ultra-high definition.**

***Анотація*—У статті розглядається можливість створення тривимірних сітчастих об'єктів із зображень природних сцен. Зображення містить обидві області з плавним зміною яскравості і колірності і області дуже щільною детальністю. Щоб розділити сцену на ділянки з різними деталями, був використаний метод градієнта. Проведено порівняння щодо цифрового потоку, коли компонент яскравості представлений вершинами рівномірної сітки і з адаптивним не рівновіддаленою сітки. Для випадку адаптивної сітки, тріангуляція Делоне використовується для порогових значень модуля градієнта. Показано, що виграш в стисненні сигналу яскравості за допомогою адаптивної сітки збільшується при кодуванні зображень з високої і надвисокої чіткості.**

Keywords—Image; scene; meshl; triangulation; detail; gradient; bitrate

Ключові слова—Зображення; сцена; сітка; тріангуляція; детальність; градієнт; швидкість цифрового потоку

I. INTRODUCTION

In this paper, we will focus on methods for creating a high-quality image on the transmitting side. Object-based approach MPEG 4 and JPEG 2000 audiovisual codecs operate in a frame- and block-based fashion. A scene is composed and rendered at the sender site. Video frames and audio are coded, multiplexed, and transmitted to the receiver. At the receiver site, the transport stream is demultiplexed, and video and audio data are decoded, synchronized, and presented as defined by the sender site. MPEG-4, however, codes objects separately, using different coding tools for video, speech, or music. A description of the scene, called binary format for scenes (BIFS), dynamically describes the scene and allows objects to be animated. Thus, scene composition is shifted to the decoder. This object-based approach also permits choosing an appropriate coding tool for each audio, video, and graphics object. From the point of view of transmission systems, stereo is still used with an additional depth signal [1, 2].

II. THE TWO-DIMENSIONAL MESH MODEL

The advantage of mesh-based processing is that it can always provide better performance for both spatial and temporal interpolation. Therefore, it has been found being more effective to be used for inter-view prediction since it can provide a coarse representation of the disparity field with adequate models degenerated from three-dimensional mesh. On the other hand, mesh-based video coding is a new technique for motion estimation and compensation.

There are several assumptions for the two-dimensional mesh model. First, there is no disparity discontinuity, that is, the surface of the object is continuous. Second, no occlusion occurs. With these two assumptions, the two-dimensional mesh model can provide the same performance as that of the three-dimensional mesh model. There are two kinds of two-

dimensional mesh: regular mesh and arbitrary mesh [2D mesh]. With arbitrary mesh, the delicate surface structure can be represented with small patches, and the smooth surface can be represented with large patches. Therefore, the performance of arbitrary mesh is better [3, 4].

Consider the creation of three-dimensional mesh of objects from the real scene. The image of dolphins in the sea contains practically flat areas without a shallow structure, as well as ripple structures on the sea and clouds in the sky (Fig. 1).

Fig. 1. The image of dolphins in the sea

Select from the original image individual objects: dolphins, sea spray, smooth sea surface. There are several methods for selecting objects; they are based on splitting the entire image into separate sections or clusters, differing either in color, or in texture or semantic content. We used a gradient method for isolating the contours of individual clusters. Fig. 2 shows the image of clusters.

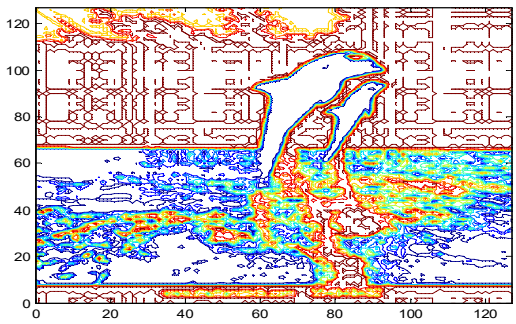


Fig. 2. The contours of individual clusters

Using program Matlab, a portion of mesh model was selected and vertices of these clusters were identified.

Analyzing the gradient mask, one can conclude that there are objects with different details. Our task is to build a three-dimensional image. In accordance with the details of individual regions, we apply a grid partition, and on even areas the grid will be uniform and with a large step, and on regions with detailed use of a non-equidistant grid. A three-dimensional model is obtained based on the orthogonal grid and the values of the gradient of the luminance signal at each point. In the next step, each three-dimensional object is approximated by the Delaunay triangulation. The number of triangles on objects of different detail is differ hundreds of times.

Let us examine the dependence of the number of mesh vertices on the choice of gradient threshold steps. Table 1 shows the dependencies for the lowest resolution image obtained because of thinning with interpolation of the original image to a size of 128 by 128 pixels. Denote the number of vertices for a uniform mesh as T1 and for an uneven mesh as T2.

TABLE I. DEPENDENCE OF THE NUMBER OF MESH VERTICES ON THE CHOICE OF GRADIENT THRESHOLD STEPS 128×128.

Gradient threshold step	T1	T2
0.1	32 258	3 256 800
0.25	32 258	524 228
0.5	32 258	232 562
0.75	32 258	131 072
1	32 258	57 800

Tables II – IV show the simulation results for the images 256×256, 512×512, 1024×1024.

TABLE II. DEPENDENCE OF THE NUMBER OF MESH VERTICES ON THE CHOICE OF GRADIENT THRESHOLD STEPS 256×256.

Gradient threshold step	T1	T2
0.1	130 050	13 107 200
0.25	130 050	2 097 152
0.5	130 050	524 288
0.75	130 050	232 562
1	130 050	131 072

TABLE III. DEPENDENCE OF THE NUMBER OF MESH VERTICES ON THE CHOICE OF GRADIENT THRESHOLD STEPS 512×512.

Gradient threshold step	T1	T2
0.1	522 242	52 428 800
0.25	522 242	8 328 608
0.5	522 242	2 097 152
0.75	522 242	930 248
1	522 242	524 288

TABLE IV. DEPENDENCE OF THE NUMBER OF MESH VERTICES ON THE CHOICE OF GRADIENT THRESHOLD STEPS 1024×1024.

Gradient threshold step	T1	T2
0.1	2 093 080	211 226 044
0.25	2 093 080	33 554 438
0.5	2 093 080	8 328 608
0.75	2 093 080	3 726 450
1	2 093 080	2 097 152

For example, in the area of an equidistant mesh where the gradient remains practically unchanged, 373 triangles have left for the description of image objects with a gradual change in the gradient of the image. Moreover, in a non-equidistant mesh where brightness jumps occur, 6275 triangles, a minimum triangle size of 3 picture elements, and a maximum triangle of 126 picture elements were required.

We estimate the digital luminance signal stream at the standard progressive scan the original image. Based on the image size M×N ultra-high definition, progressive scan with

number of frames in a second n, we will take a bit digital luminance signal is equal to 10. Then the upper cut-off frequency can be determined by the formula (1)

$$f_{max}=M \times N \times n/2, \quad (1)$$

$$f_{max}=3840 \times 2160 \times 50/2=297.4\ MHz$$

The sampling frequency of the luminance signal of UHDTV is $fs_Y = 432$ MHz, and the speed of the digital stream will be 4320 Mbps for the luminance signal. The number of vertices and the bit depth of the codes for the three coordinates of each vertex can estimate the flow created by triangulated objects. We can assume that the number of vertices of V is approximately twice as large as the triangles of T. In our example, the minimum dimension of the side of the triangle is two decomposition elements and we can limit ourselves to the twelve-digit code for the vertices for all three coordinates. In this case, a rough estimate of the digital flow can be represented by the formula (2)

$$V_{\Sigma}=V_{sm}+V_{det}=[3 \times m \times n \times 2 \times (T_{sm}+T_{det})] \quad (2)$$

$$V_{\Sigma}=[3 \times 12 \times 50 \times 2 \times (373+6275_t)] = 23.93\ Mbit/s.$$

where V_{sm} – the bit rate for the polygons on flat areas;
V_{det} – the bit rate for polygons in areas of great detail;
T_{sm} – number of triangles on flat areas;
T_{det} –number of triangles;
m – bit code for vertices;
n – number of frames transmitted per second.

Moreover we found values of bitrate in each case. The result you can see in the Table V.

TABLE V. BITRATE OF DIFFERENT IMAGE SIZE AND MESH CODING METHODS

Image size	Bitrate, Mbit/s
128×128	0,38
256×256	0,38
512×512	1,50
1024×1024	6,00
2048×2048	18,76
4096×4096	23,93

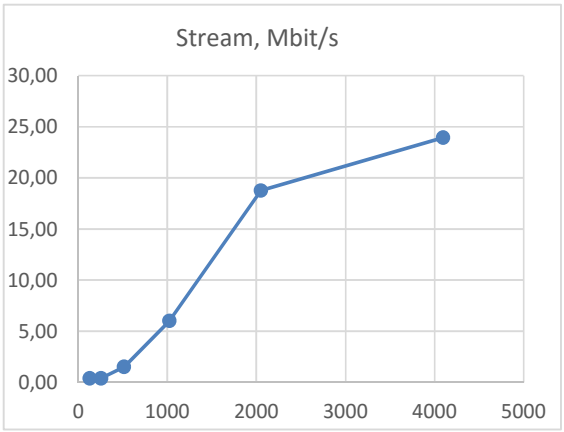


Fig. 3. A plot of the bit rate of the image size for mesh coding methods

Even with this approximate analysis it is shown that the transition from the progressive scan, when information is transmitted about each picture element to an object representation allows you to get the original digital stream which is much less for ultra-high-definition images. From the graph in Fig. 3, the dependence of the digital bit stream on the size of the images it is seen that the gain from the transition to object-oriented representation of scenes appears only for images with the size larger than 2K.

III. THE SOME ASPECTS OF ASSESSING THE QUALITY OF MESH MODELS

There is no doubt that the introduction of 3D TV can have a lasting success only if perceived image quality and viewing comfort, at least comparable to a conventional high-definition TV.

The main emphasis in our work is made on the concept of an image quality model, which is a key component of the development of an object-oriented television system. This concept can be defined on the basis of formally perceived attributes related to this system, and how these attributes can be integrated to predict image quality.

To date, the most accurate ways to assess the perceived quality of 3D are subjective assessment methods that are obtained by statistical processing of the results of examinations. The disadvantage is that these studies are time consuming. In addition, universal procedures for assessing the quality of stereoscopic images do not yet exist; they adapt the existing 2D standards that should be used correctly for subjective evaluation of 3D content.

In the subjective examination of TV stereo images, along with resolution, color reproduction and other parameters, which are usually checked in monoscopic TV systems, other characteristics specific to stereo systems are also analyzed, including resolution of the depth of the image, movement in its depth, distortion of the type of effect miniaturization (unnatural size of objects) and the effect of "cardboard" (unnatural thickness of objects), the comfort of viewing images, etc.

In multimedia applications, such as games and simulators, object-oriented methods for creating the bulk effect have become more widespread. Such objects require painstaking work to create, and, for the most part, artificially created objects, or even transformed from actually shot stereo. When creating objects from two cameras, there are many difficulties in matching the two types for the left and right eyes, since it is difficult to ensure the identity of the cameras, and the angle of incidence of the illumination is different. The latter circumstance leads to the need to identify points on two images, and by identifying by color, brightness, shadows; there is no possibility of achieving high reliability.

In search of an ideal 3D model of quality, on the one hand, there were numerous studies that tried to determine the quality of stereoscopic 3D only with respect to 2D components of stereoscopic content, as in the case of images and video. On the other hand, there were also studies in which stereoscopic 3D quality was regarded as part of a more

complex multidimensional concept. The factors listed, such as reproduction depth, stereoscopic deterioration and visual comfort, must be taken into account when developing a model for evaluating 3D quality.

We accept as input polygon-mesh models of 3D-objects, which is a set of triangles in 3D-space. A mesh consists of a set of vertices V. In addition to three-dimensional coordinates, each vertex contains information about the color represented in the RGB space; the point cloud in space specifies a discrete source object. Each image corresponds to an associated tensor. For each vertex, we want to estimate the color and the color difference from the neighboring vertices. The collection of colors of peaks is a color model to be investigated. According to our estimates, nonlinear spline approximation problems that minimize the color divergence compared to the original continuous object obtained from the camera for each vertex are applicable to the color model. A particular task, to be optimized, is to solve the problem of metamerism of colors.

When using polygon-mesh models of objects, especially those created on the basis of real ones, it is necessary to pay attention to the effect on the overall quality of the change in the reproduction of textures. The establishment of the type of the reflecting surface plays an important role, namely: diffuse or specular. In the case of specular reflection, glare, ghosting, significantly masking other distortions.

The manifestation of the aggregate of distortions in the subjective assessment is the subject of investigations, since deterioration stack with different impact on the final score. One should also take into account the masking effect of some distortions on others. Neither of the existing image bases included information on the spectral composition of the primary colors, nor does it include the characteristics in which color system the primary colors are represented.

One of the fundamental problems of image analysis is the creation of an adequate mathematical description of images that convey their content, meaning. In other words, this description should reflect only significant (from the point of view of the problem being solved) image features, and not depend on non-essential details. The methods of morphological analysis are, therefore, a step towards solving the problem of describing the semantics of images

At the heart of the methods of morphological analysis lies the mathematical concept of form. The form of the image is understood as the maximum invariant of image transformations, to which it undergoes changes in observation conditions, changes in the parameters of recording equipment,

Contextual or content description of images is a certain set of features that characterize this image. Therefore, it is important what source of descriptions is used to obtain them, since it determines what methods and means of analysis and representation of images can be applied. The main problem of content image retrieval is the translation of the content of images from the low-level representation of features (color, texture, shape, position, etc.) into the semantic plane, that is, the system's understanding of the content of the image at the level of the image, scene, situation and events.

The task of creating the most objective description in the form of a semantic annotation involves enumerating all possible objects potentially interesting to the user with support for objective data of different types, in particular: - the actual relative size of the object in the image, for example, the part of the image area occupied by the object in relation to The area of the whole image or other object; - perspective - reducing the size of objects when moving away from the observer; - knowledge about the device, relative dimensions (for example, in comparison with the size of a person) and the functionality of objects. The concepts of ontology are elements of the semantic annotations of images represented in the XML formalism.

IV. CONCLUSION

The development of the means of production of television programs is faced with the tasks of an adequate assessment of the quality of images and sound.

Large screens allow viewers to view the smallest details. For their unmistakable transfer, it is suggested to use the division of the scene into separate objects. Depending on the details of these objects, a different number of polygons represents them.

It is proposed to use 2D polygon-mesh models, and the size of the triangles and their number depend on the detail. Investigations of the model based on the gradient of the luminance signal have shown the possibilities of winning in the volume of the transmitted data due to the redistribution of the reference points for areas with different details.

Further research is aimed at developing a criterion for assigning images to a particular detail group.

REFERENCES

[1] Draft New Recommendation ITU-R BT.[3DTV-REQS] Performance requirements for the production, international exchange and broadcasting of 3DTV Programmes.

[2] Recommendation H.264 (10/16): Advanced video coding for generic audiovisual services

[3] ISO/IEC, JTC1/SC29/WG11, “Coding of Moving Pictures and Audio - MPEG-4 Visual Standard, ISO / IEC FIS 14446-2, 2000.

[4] A. Murat. Tecalp and all Two-Dimensional Mesh-Based Visual-Object Representation for Interactive Synthetic/ Natural Digital Video // Proceedings of the IEEE vol 86, No 6, june 1998, pp. 1029 – 1051.

[5] K. Mamou, N. Stefanoski, H. Kirchhoffer K. Muller, T. Zaharia, F. Preteux, D. Marpe, J. Ostermann. The New MPEG-4/FAMC Standard For Animated 3D Mesh Compression // 3DTV-CON'08, May 28-30, 2008, Istanbul, Turkey. pp. 97-100

Multivariate Outlier Detection Technique Based on Normalizing Transformations for Non-Gaussian Data

Sergiy Prykhodko
Department of Software of Automated Systems
Admiral Makarov National University of Shipbuilding
Mykolaiv, Ukraine
E-mail: sergiy.prykhodko@nuos.edu.ua

Natalia Prykhodko
Finance Department
Admiral Makarov National University of Shipbuilding
Mykolaiv, Ukraine
E-mail: natalia.prykhodko@nuos.edu.ua

Lidiia Makarova
Department of Software of Automated Systems
Admiral Makarov National University of Shipbuilding
Mykolaiv, Ukraine
E-mail: lidiia.makarova@nuos.edu.ua

Kateryna Pugachenko
Department of Information Systems and Technologies
Admiral Makarov National University of Shipbuilding
Mykolaiv, Ukraine
E-mail: kateryna.puhachenko@gmail.com

Метод Визначення Багатовимірних Викидів, що Базується на Нормалізуючих Перетвореннях для Негаусівських Даних

Сергій Приходько
кафедра програмного забезпечення автоматизованих систем
Національний університет кораблебудування імені адмірала Макарова
Миколаїв, Україна
sergiy.prykhodko@nuos.edu.ua

Наталія Приходько
кафедра фінансів
Національний університет кораблебудування імені адмірала Макарова
Миколаїв, Україна
natalia.prykhodko@nuos.edu.ua

Лідія Макарова
кафедра програмного забезпечення автоматизованих систем
Національний університет кораблебудування імені адмірала Макарова
Миколаїв, Україна
lidiia.makarova@nuos.edu.ua

Катерина Пугаченко
кафедра інформаційних управляючих систем та технологій
Національний університет кораблебудування імені адмірала Макарова
Миколаїв, Україна
kateryna.puhachenko@gmail.com

***Abstract*—The statistical technique for detecting outliers in multivariate non-Gaussian data on the basis of normalizing transformations and a test statistic for the Mahalanobis squared distance (MSD), which has an approximate *F* distribution, is proposed. Application of the technique is considered for detecting outliers in the data of four-variate measurements.**

***Анотація*—Пропонується статистичний метод для виявлення викидів в багатовимірних негаусівських даних на основі нормалізуючих перетворень і тестової статистики для квадрата відстані Махаланобісу (MSD), який має наближений *F* розподіл. Розглянуто застосування цього методу для виявлення викидів за даними вимірювань чотирьох величин.**

Keywords—outlier; normalizing transformation; multivariate non-Gaussian data; Mahalanobis squared distance; F distribution

Ключові слова—викид; нормалізуюче перетворення; багатовимірні негауссівськи дані; квадрат відстані Махаланобісу; F-розподіл

I. INTRODUCTION

A very important step in data analysis and processing is the outlier detection. Today the problem of outlier detection in a multivariate data set is solved with different methods including statistical [1, 2]. However, well-known statistical methods (for example, multivariate outlier detection based on a test statistics for MSD, which have an approximate the Chi-Square distribution or *F* distribution) are used to detect outliers in a multivariate data set under the assumption that the data is generated by a Gaussian distribution. And this assumption is valid only in particular cases.

In [3] a statistical outlier detection technique for multivariate non-Gaussian data on the basis of normalizing transformations and MSD, which has an approximate the Chi-Square distribution, was proposed. However, the quantile of the Chi-Square distribution does not depend from the number of data. We propose a statistical outlier detection technique for multivariate non-Gaussian data on the basis of normalizing transformations and a test statistic for MSD, which has an approximate *F* distribution. At that the quantile of the *F* distribution depends from the number of data. The technique consists of two steps. In the first step, multivariate non-Gaussian data is normalized using a multivariate normalizing transformation. In the second step, MSD and a test statistic for MSD are calculated and compared with a quantile of the *F* distribution. The data values for which a value of test statistic for MSD is greater than the quantile of the *F* distribution are considered as outliers and these values are cut off. Two steps should be repeated for the data after outlier cutoff until all values of test statistic for MSD will be less than or equal to the quantile of the *F* distribution.

II. THE STATISTICAL TECHNIQUE

The outlier detection technique for multivariate non-Gaussian data is based on normalizing transformations and a test statistic for MSD, which has an approximate *F* distribution. Consider bijective multivariate normalizing transformation of non-Gaussian random vector $\mathbf{X} = \{X_1, X_2, \ldots, X_m\}^T$ to Gaussian random vector $\mathbf{Z} = \{Z_1, Z_2, \ldots, Z_m\}^T$ is given by

$$\mathbf{Z} = \psi(\mathbf{X}). \tag{1}$$

The values of the sample observations or multivariate data points $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_N$ are normalized using the transformation (1).

The Mahalanobis squared distance for each multivariate data point *i*, $i = 1,2,\ldots,N$, is denoted by d_i^2 and given by

$$d_i^2 = (\mathbf{Z}_i - \overline{\mathbf{Z}})^T S_N^{-1} (\mathbf{Z}_i - \overline{\mathbf{Z}}), \tag{2}$$

where $\overline{\mathbf{Z}}$ is the sample mean vector and S_N is the sample correlation matrix

$$S_N = \frac{1}{N} \sum_{i=1}^{N} (\mathbf{Z}_i - \overline{\mathbf{Z}})(\mathbf{Z}_i - \overline{\mathbf{Z}})^T. \tag{3}$$

A test statistic (TS) for d_i^2 can be created as follows [4]

$$(N-m) N d_i^2 / (N^2 - 1) m, \tag{4}$$

which has an approximate *F* distribution with *m* and $N - m$ degrees of freedom.

A test statistic for MSD (4) is compared with a quantile of the *F* distribution, which is noted as $F_{m,N-m,\alpha}$. Here α is significance level. We take α as 0.05. The data values for which a value of TS (4) is greater than the quantile of the *F* distribution are considered as outliers and these values are cut off. After outlier cutoff the reduced number of multivariate data points are normalized using the transformation (1) again until all values of TS (4) will be less than or equal to the quantile of the *F* distribution.

III. MULTIVARIATE NORMALIZING TRANSFORMATIONS

Some transformations have been proposed for normalizing multivariate non-Gaussian data, such as, transformation on the basis of the decimal logarithm, the Box-Cox transformation, the Johnson translation system and others. However, only a few normalizing transformations are bijective. Such bijective transformation is the transformation of S_U family of the Johnson translation system. The Johnson normalizing translation is given by [5]

$$\mathbf{Z} = \boldsymbol{\gamma} + \boldsymbol{\eta} \mathbf{h}\left[\boldsymbol{\lambda}^{-1}(\mathbf{X} - \varphi)\right] \sim N_m(0_m, \Sigma), \tag{5}$$

where Σ is the correlation matrix; $\boldsymbol{\gamma}$, $\boldsymbol{\eta}$, φ and $\boldsymbol{\lambda}$ are parameters of the Johnson normalizing translation; $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_m)^T$; $\boldsymbol{\eta} = diag(\eta_1, \eta_2, \ldots, \eta_m)$; $\varphi = (\varphi_1, \varphi_2, \ldots, \varphi_m)^T$; $\boldsymbol{\lambda} = diag(\lambda_1, \lambda_2, \ldots, \lambda_m)$; $\mathbf{h}[(y_1, y_2, \ldots, y_m)] = \{h_1(y_1), h_2(y_2), \ldots, h_m(y_m)\}^T$; $h_i(\cdot)$ is one of the translation functions

$$h = \begin{cases} \ln(y), & \text{for } S_L \text{ (log normal) family;} \\ \ln[y/(1-y)], & \text{for } S_B \text{ (bounded) family;} \\ \mathrm{Arsh}(y), & \text{for } S_U \text{ (unbounded) family;} \\ y & \text{for } S_N \text{ (normal) family.} \end{cases}$$

Here $y = (x - \varphi)/\lambda$; $\mathrm{Arsh}(y) = \ln\left(y + \sqrt{y^2 + 1}\right)$.

IV. EXAMPLES

We consider the examples of detecting outliers in multivariate non-Gaussian data for two cases: the first, under the assumption that the data set is generated by a Gaussian distribution and, the second, this assumption is not valid. In [2]

example of detecting outliers in the lumber stiffness data (example 4.15, p.190) is presented. Table I contains the data on lumber, MSD and a test statistic for MSD for standardized measurements, which are

$$Z_{ki} = \left(X_{ki} - \overline{X}_k\right) \Big/ S_{X_k} \ , \ k = 1,2,3,4 \ ; \ i = 1,2,\ldots,30 . \quad (6)$$

TABLE I. MSD for the standardized data

i	X_1	X_2	X_3	X_4	d_i^2	TS for MSD
1	1889	1651	1561	1778	0.61	0.13
2	2493	2048	2087	2197	7.48	1.62
3	2119	1700	1815	2222	7.85	1.70
4	1645	1627	1110	1533	5.23	1.13
5	1976	1916	1614	1883	1.44	0.31
6	1712	1712	1439	1546	2.27	0.49
7	1943	1685	1271	1671	5.17	1.12
8	2104	1820	1717	1874	1.34	0.29
9	2983	2794	2412	2581	12.37	2.68
10	1745	1600	1384	1508	0.78	0.17
11	1710	1591	1518	1667	2.05	0.44
12	2046	1907	1627	1898	0.5	0.11
13	1840	1841	1595	1741	2.78	0.60
14	1867	1685	1493	1678	0.12	0.03
15	1859	1649	1389	1714	1.11	0.24
16	1954	2149	1180	1281	17.53	3.80
17	1325	1170	1002	1176	3.63	0.79
18	1419	1371	1252	1308	4.08	0.89
19	1828	1634	1602	1755	1.48	0.32
20	1725	1594	1313	1646	1.48	0.32
21	2276	2189	1547	2111	10.25	2.22
22	1899	1614	1422	1477	4.8	1.04
23	1633	1513	1290	1516	0.8	0.17
24	2061	1867	1646	2037	2.64	0.57
25	1856	1493	1356	1533	4.33	0.94
26	1727	1412	1238	1469	3.37	0.73
27	2168	1896	1701	1834	2.13	0.46
28	1655	1675	1414	1597	2.93	0.64
29	2326	2301	2065	2234	6.76	1.47
30	1490	1382	1214	1284	2.67	0.58

These data consist of four different measures of stiffness X_1, X_2, X_3 and X_4, on each $N = 30$ of boards. For significance level which equals to 0.05 the last column in Table I reveals that specimen 16 is multivariate outlier, since $F_{m,N-m,\alpha}$ =2.74.

Table II contains the normalized data on lumber, MSD and a test statistic for MSD for normalized data. These data is normalized by S_U family of the transformation (5).

TABLE II. MSD FOR THE NORMOLIZED DATA

i	Z_1	Z_2	Z_3	Z_4	d_i^2	TS for MSD
1	0.1984	-0.1910	0.3966	0.3539	0.77	0.17
2	1.4562	1.0898	1.5843	1.2985	2.88	0.63
3	0.8508	0.0237	1.0712	1.3430	4.23	0.92
4	-0.9978	-0.3027	-1.7482	-0.4821	9.07	1.97
5	0.4882	0.7632	0.5598	0.6343	0.78	0.17
6	-0.6296	0.0734	-0.0412	-0.4297	2.31	0.50
7	0.3859	-0.0401	-0.8259	0.0249	4.54	0.98
8	0.8177	0.4724	0.8404	0.6117	1.13	0.25
9	1.9326	2.1752	2.0416	1.8902	5.26	1.14
10	-0.4518	-0.4328	-0.2728	-0.5859	0.59	0.13
11	-0.6406	-0.4770	0.2530	0.0116	1.99	0.43
12	0.6799	0.7381	0.5977	0.6715	0.58	0.13
13	0.0031	0.5407	0.5029	0.2456	1.54	0.33
14	0.1139	-0.0401	0.1644	0.0480	0.14	0.03
15	0.0819	-0.2002	-0.2507	0.1630	0.92	0.20
16	0.4209	1.2978	-1.3383	-1.7120	18.95	4.11
17	-2.2232	-2.2878	-2.3454	-2.3094	6.33	1.37
18	-1.9595	-1.5575	-0.9289	-1.5624	5.40	1.17
19	-0.0488	-0.2697	0.5241	0.2872	1.17	0.25
20	-0.5587	-0.4622	-0.6080	-0.0594	1.68	0.36
21	1.1474	1.3723	0.3511	1.1373	4.25	0.92
22	0.2352	-0.3648	-0.1103	-0.7201	3.97	0.86
23	-1.0620	-0.8721	-0.7256	-0.5522	1.50	0.32
24	0.7172	0.6214	0.6518	0.9874	1.16	0.25
25	0.0697	-0.9741	-0.4000	-0.4821	4.34	0.94
26	-0.5479	-1.3717	-1.0063	-0.7558	3.78	0.82
27	0.9523	0.7069	0.7996	0.5081	1.42	0.31
28	-0.9435	-0.0836	-0.1435	-0.2339	3.52	0.76
29	1.2270	1.5616	1.5482	1.3640	3.14	0.68
30	-1.7150	-1.5089	-1.1418	-1.6952	3.96	0.86

In these case the parameters are such: $\gamma_1 = -0.76817$, $\gamma_2 = -0.81796$, $\gamma_3 = -1.52027$, $\gamma_4 = -2.11559$, $\eta_1 = 1.01985$, $\eta_2 = 1.29026$, $\eta_3 = 1.58452$, $\eta_4 = 1.87154$, $\varphi_1 = 1686.87$,

$\varphi_2 = 1523.59$, $\varphi_3 = 1148.94$, $\varphi_4 = 1209.93$, $\lambda_1 = 184.403$, $\lambda_2 = 252.185$, $\lambda_3 = 269.810$ and $\lambda_4 = 327.027$.

The sample correlation matrix (3) of the $\mathbf{Z}$ is used as the approximate moment-matching estimator of correlation matrix Σ

$$S_N = \begin{pmatrix} 1.00000 & 0.85039 & 0.79001 & 0.82173 \\ 0.85039 & 1.00000 & 0.71725 & 0.73854 \\ 0.79001 & 0.71725 & 1.00000 & 0.87227 \\ 0.82173 & 0.73854 & 0.87227 & 1.00000 \end{pmatrix}.$$

In Table II the last column reveals that only specimen 16 is multivariate outlier, since $F_{m,N-m,\alpha}$ =2.74. We note the value of TS for MSD for the normalized data from Table II is greater than the same value for the data from Table I in this case.

Table III contains the normalized data on lumber and MSD for normalized data after outlier cutoff of specimen 16 (according Table II). These data is normalized by S_U family of the transformation (5). After outlier cutoff the parameters are such: $\gamma_1 = -0.70625$, $\gamma_2 = -0.62410$, $\gamma_3 = -0.70133$, $\gamma_4 = -21.2264$, $\eta_1 = 0.33081$, $\eta_2 = 0.23699$, $\eta_3 = 0.21906$, $\eta_4 = 6.20266$, $\varphi_1 = 1700.55$, $\varphi_2 = 1587.45$, $\varphi_3 = 1304.12$, $\varphi_4 = -181.883$, $\lambda_1 = 10.202$, $\lambda_2 = 2.123$, $\lambda_3 = 1.495$ and $\lambda_4 = 124.038$.

The sample correlation matrix (3) of the $\mathbf{Z}$ is such

$$S_N = \begin{pmatrix} 1.00000 & 0.70699 & 0.74789 & 0.77134 \\ 0.70699 & 1.00000 & 0.72856 & 0.75745 \\ 0.74789 & 0.72856 & 1.00000 & 0.70002 \\ 0.77134 & 0.75745 & 0.70002 & 1.00000 \end{pmatrix}.$$

The last column in Table III reveals there are no multivariate outliers in the data, since $F_{m,N-m,\alpha}$ =2.76. We note, if the parametric anomaly detection technique [6] based on the Grubb test applies for detecting outliers in the normalized data from Table II and Table III then all data sample units do not appear to be an outlier in each of the univariate distributions.

Following [7] Mardia's multivariate kurtosis β_2 is calculated for the data from Table I, Table II and Table III as

$$\beta_2 = \frac{1}{N}\sum_{i-1}^{N}\left\{\left(\mathbf{Z}_i - \overline{\mathbf{Z}}\right)^T S_N^{-1}\left(\mathbf{Z}_i - \overline{\mathbf{Z}}\right)\right\}^2.$$

It is known that $\beta_2 = m(m+2)$ holds under multivariate normality. The given equality is a necessary condition for multivariate normality. In our case $\beta_2 = 24$. The values of multivariate kurtosis equal, respectively, 30.71, 23.84 and 23.78 for the data from Table I, Table II and Table III. These values indicate that the necessary condition for multivariate normality is practically performed for the normalized data from Table II and Table III and does not hold for standardized data from Table I by the formula (6).

TABLE III. MSD FOR THE NORMOLIZED DATA AFTER OUTLIER CUTOFF

i	Z_1	Z_2	Z_3	Z_4	d_i^2	TS for MSD
1	0.4880	0.3457	0.5780	0.1988	0.49	0.11
2	0.9629	0.8150	0.8221	1.3986	2.10	0.45
3	0.7517	0.4811	0.7286	1.4633	3.18	0.69
4	-1.4990	0.2334	-1.9193	-0.6276	10.43	2.25
5	0.6135	0.7350	0.6191	0.5219	0.58	0.13
6	-0.3870	0.5051	0.4369	-0.5809	3.23	0.70
7	0.5713	0.4472	-1.5320	-0.1487	11.66	2.52
8	0.7397	0.6531	0.6819	0.4949	0.66	0.14
9	1.1222	1.0433	0.8982	2.3257	7.32	1.58
10	0.0142	-0.0371	0.3221	-0.7184	2.09	0.45
11	-0.4323	-0.3193	0.5379	-0.1621	2.29	0.49
12	0.6883	0.7284	0.6281	0.5667	0.61	0.13
13	0.3886	0.6736	0.6052	0.0808	1.06	0.23
14	0.4470	0.4472	0.5106	-0.1253	1.21	0.26
15	0.4308	0.3381	0.3354	-0.0067	0.59	0.13
16	-2.1285	-2.0400	-2.0162	-2.0706	5.33	1.15
17	-2.0332	-1.8843	-1.6313	-1.4974	4.75	1.03
18	0.3589	0.2720	0.6104	0.1257	0.56	0.12
19	-0.1743	-0.1869	-0.1577	-0.2327	0.05	0.01
20	0.8571	0.8783	0.5657	1.1705	1.53	0.33
21	0.5051	0.1392	0.4073	-0.8330	4.59	0.99
22	-1.5628	-1.6314	-1.3458	-0.6892	4.74	1.02
23	0.7024	0.6967	0.6406	0.9673	0.94	0.20
24	0.4244	-1.6878	0.2276	-0.6276	9.57	2.06
25	-0.1502	-1.8345	-1.6834	-0.8629	8.16	1.76
26	0.7883	0.7201	0.6733	0.3732	0.98	0.21
27	-1.4346	0.4216	0.3920	-0.4009	8.38	1.81
28	0.8847	0.9188	0.8159	1.4942	2.46	0.53
29	-1.9372	-1.8719	-1.7512	-1.5977	4.33	0.93

The F plots are computed for a test statistic for MSD from Table I, Table II and Table III. A F plot is similar to a Q-Q plot. The F plots for the test statistic for MSD (4) from Table I, Table II and Table III are shown on Fig. 1, Fig. 2 and Fig. 3.

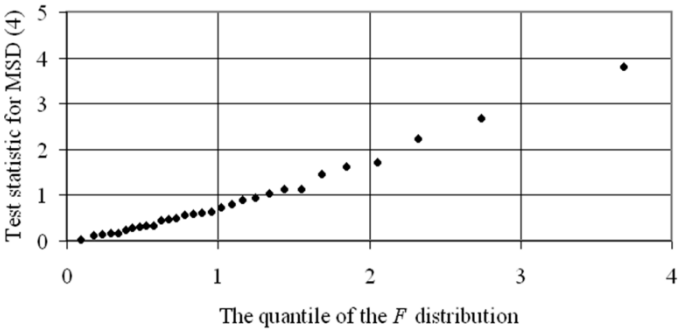


Fig. 1. *F* plot for a test statistic for MSD (4) from Table I

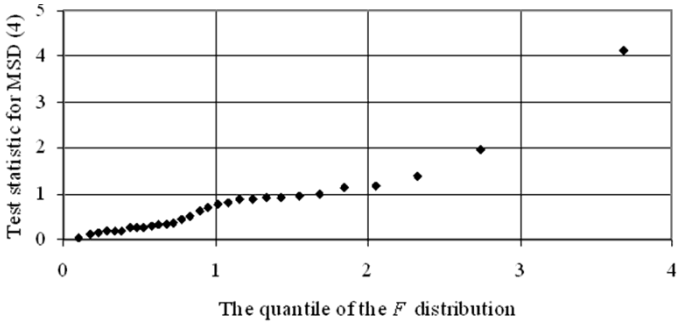


Fig. 2. *F* plot for a test statistic for MSD (4) from Table II

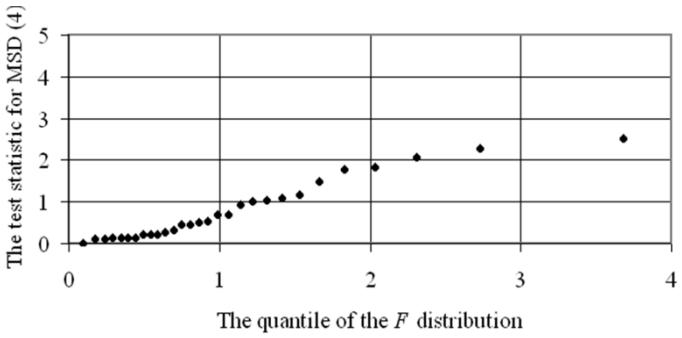


Fig. 3. *F* plot for a test statistic for MSD (4) from Table III

On Fig. 1 and Fig. 2 the quantiles of the *F* distribution are calculated for $m = 4$ and $N - m$ =26 degrees of freedom and the 30 probability levels, which are computed as $\frac{(1-0.5)}{30} = 0.0167$, $\frac{(2-0.5)}{30} = 0.05$,…, $\frac{(30-0.5)}{30} = 0.9833$.

On Fig. 3 the quantiles of the *F* distribution are calculated for $m = 4$ and $N - m$ =25 degrees of freedom and the 29 probability levels, which are computed as $\frac{(1-0.5)}{29} = 0.0172$, $\frac{(2-0.5)}{29} = 0.0517$,…, $\frac{(29-0.5)}{30} = 0.9828$.

It is necessary to check if points of the *F* plot approximately fall on a straight line. On Fig. 1 and Fig. 2 the one specimen with largest MSD is clearly removed from straight line pattern. We note on Fig. 1 and Fig. 2 the point with largest MSD is situated upper from an approximate straight line. On Fig. 3 all points do not appear to be an outlier. The same results were obtained by proposed technique for detecting outliers in multivariate non-Gaussian data.

V. CONCLUSIONS

From the examples we conclude that the proposed technique is promising, since, firstly, for multivariate outlier the value of test statistic for the Mahalanobis squared distance for the normalized data is greater than the same value for the standardized data and, secondly, the multivariate data point is not flagged as an outlier when we only consider the univariate data. However, some refinements may be necessary. We intend to try other normalizing transformations and test statistics for the Mahalanobis squared distance.

REFERENCES ЛІТЕРАТУРА

[1] D.M. Hawkins. Identification of Outliers, London; New York: Chapman and Hall, 1980, 188 p.

[2] R.A. Johnson and D.W. Wichern. Applied Multivariate Statistical Analysis. Pearson Prentice Hall, 2007, 800 p.

[3] S. Prykhodko, N. Prykhodko, L. Makarova and K. Pugachenko, "Detecting outliers in multivariate non-Gaussian data on the basis of normalizing transformations", unpublished.

[4] P.M. Stanfield, J.R. Wilson, G.A. Mirka, N.F. Glasscock, J.P. Psihogios, J.R. Davis. "Multivariate input modeling with Johnson distributions", in *Proceedings of the 28th Winter simulation conference WSC'96*, December 8-11, 1996, Coronado, CA, USA, ed. S.Andradyttir, K.J.Healy, D.H.Withers, and B.L.Nelson, IEEE Computer Society Washington, DC, USA, 1996, pp. 1457-1464.

[5] A.A. Afifi and S.P. Azen. Statistical analysis: a computer oriented approach, New York; London: Academic Press, 1972, 366 p.

[6] S.B. Prykhodko. "Statistical anomaly detection techniques based on normalizing transformations for non-Gaussian data", in *"Computational Intelligence (Results, Problems and Perspectives)", Proceedings of the International Conference*, Kyiv-Cherkasy, Ukraine, May 12-15, 2015, pp. 286-287.

[7] K.V. Mardia. "Measures of multivariate skewness and kurtosis with applications", *Biometrika*, 57, pp. 519–530, 1970.

Апроксимація Параметричними Сплайнами для Ущільнення Інформації Електрокардіограми у Трьох Ортогональних Відведеннях

В.В. Вишневський
К.т.н., провідний науковий співробітник
Інститут проблем математичних машин і систем НАН України
Київ, Україна
vit@immsp.kiev.ua

Т.М. Романенко
Науковий співробітник
Інститут проблем математичних машин і систем НАН України
Київ, Україна
romanenko@immsp.kiev.ua

Approximation by Parametrically Specified Splines for Consolidation of Electrocardiogram in Three Orthogonal Leads Information

V.V. Vishnevsky
Ph.D., Senior researcher
Institute of Mathematical Machines and Systems of NASU
Acad. Glushkova Avenue, 42, Kyiv, Ukraine
vit@immsp.kiev.ua

T.M. Romanenko
Scientific researcher
Institute of Mathematical Machines and Systems of NASU
Acad. Glushkova Avenue, 42, Kyiv, Ukraine
romanenko@immsp.kiev.ua

***Анотація*—обговорюється алгоритм обчислення коефіцієнтів параметричного сплайну, що апроксимує QRS-комплекс електрокардіограми (ЕКГ) у трьох ортогональних відведеннях. Алгоритм використовується для попередньої обробки ЕКГ для подальшої її класифікації за допомогою нейромережі.**

***Abstract*—an algorithm for computing the coefficients of a parametrically specified spline which approximates the QRS-complex of the electrocardiogram (ECG) in three orthogonal leads is discussed. The algorithm is used to pre-process the ECG for further classification using a neural network.**

Ключові слова—електрокардіограма (ЕКГ); QRS-комплекс; апроксимація; параметричний сплайн; класифікація; нейромережа.

Keywords—electrocardiogram (ECG); QRS-complex; approximation; cardinal spline; classification; neural network.

I. Вступ

Електрокардіографія, як відомо, є рутинним методом діагностики стану серцево-судинної системи людини. Останнім часом на ринку медичного обладнання з'явилися мобільні кардіографи, які дозволяють будувати телемедичні системи на базі використання смартфонів та планшетів. Побудова таких систем, у свою чергу, робить актуальними задачі експертної оцінки кардіограми за діагностичними ознаками, а також задачу ідентифікації пацієнта за його ЕКГ. З точки зору інформаційної системи і задача діагностики, і задача ідентифікації можуть бути зведені до задачі класифікації нейромережею. В даній роботі розглядається процедура ущільнення інформації, що міститься в електрокардіограмі, та виділення інформаційних ознак, що можуть бути використані нейромережею безпосередньо для класифікації. Програмна реалізація алгоритмів, що описані в цій роботі, використовуються нами для автоматичної ідентифікації пацієнта за його ЕКГ.

II. Постановка задачі

Стандартна електрокардіограма містить 12 відведень, тобто 12 сигналів. Ці 12 сигналів можуть бути перетворені на три ортогональні відведення та розмічені на цикли та елементи циклів за допомогою спеціальних алгоритмів [1]. Окрім стандартної електрокардіограми на 12 відведень існує також електрокардіограма за Франком, с такою схемою накладення електродів, яка дозволяє одразу отримати три ортогональні відведення. Електрокардіограма в ортогональних відведеннях може

бути представлена в фазовому просторі координат у вигляді вектор-кардіограми [2, 3].

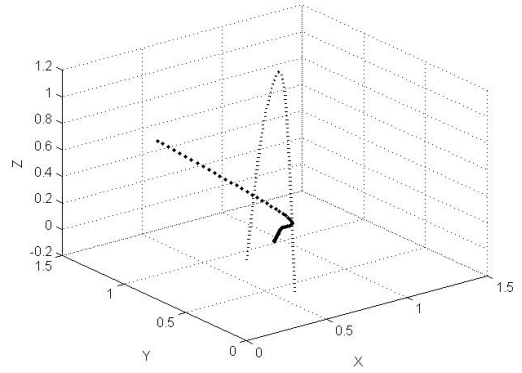


Рис. 1. QRS-комплекси електрокардіограм двох пацієнтів, представлені у вигляді тривимірних дуг

Після розмічення ЕКГ у кожному її циклі візьмемо частину, що відповідає QRS-комплексу. Маючи три відведення, можемо представити її у вигляді тривимірної дуги (рис. 1, 2). Кількість точок у такій дузі достатньо велика, до того ж не однакова для різних ЕКГ. Тому безпосередньо використати координати цих точок як параметри для класифікації ЕКГ за допомогою нейромережі неможливо: по-перше, велика кількість параметрів потребує значного обсягу даних для навчання нейромережі, по-друге, кількість параметрів має бути однаковою для всіх дуг, що класифікуються.

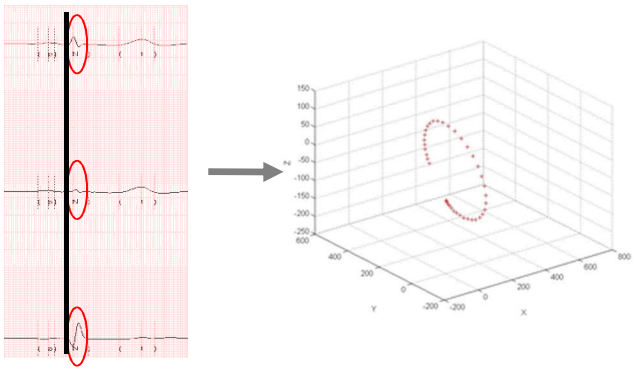


Рис. 2. Три ортогональних відведення ЕКГ та тривимірна дуга, що їм відповідає

Апроксимувавши тривимірну дугу, що відповідає QRS-комплексу, параметричними сплайнами, отримуємо значно меншу кількість точок, які характеризують форму кривої. Для апроксимації можна використати канонічний сплайн (cardinal spline) або криву Безьє [4, 5].

В подальшому класифікація нейромережею виконується для координат контрольних точок параметричного сплайна, кількість яких може бути обмежена, наприклад, чотирма.

III. Опис алгоритму

Тож, вихідними даними маємо дискретні відліки тривимірної дуги кривої, яка є частиною ЕКГ у трьох відведеннях, що відповідає QRS-комплексу: $X = \{x(m)\}$, $Y = \{y(m)\}$, $Z = \{z(m)\}$. Нам необхідно класифікувати форму цієї дуги.

Вихідні дані можна замінити аналітичними кривими, які добре збігаються за формою з тривимірною кривою, що відповідає QRS-комплексу, але водночас є інваріантними щодо кількості відліків та масштабу. Ці аналітичні криві задаються функціями $x(k)$, $y(k)$, $z(k)$, які є поліномами r-ого степеня від параметра k:

$$x(k) = a_0k^r + a_1k^{r-1} + \ldots + a_{r-1}k + a_r,$$

$$y(k) = b_0k^r + b_1k^{r-1} + \ldots + b_{r-1}k + b_r,$$

$$z(k) = c_0k^r + c_1k^{r-1} + \ldots + c_{r-1}k + c_r,$$

Вибір r-ого степеня поліномів залежить від складності дуги, що має бути апроксимованою.

В цій роботі буде описано використання для апроксимації тривимірної дуги, що відповідає QRS-комплексу електрокардіограми, канонічного сплайна (cardinal spline) третього порядку.

Канонічний сплайн (*cardinal spline*) – це поліном третього порядку, який можна записати параметричним

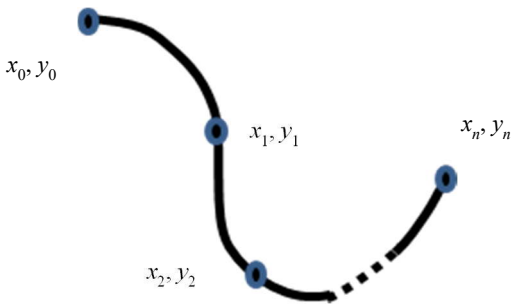


Рис. 3. Канонічний сплайн

рівнянням загального вигляду:

$$x(t) = a_xt^r + b_xt^{r-1} + c_xt + d_x,$$

$$y(t) = a_yt^r + b_yt^{r-1} + c_yt + d_y,$$

$$z(t) = a_zt^r + b_zt^{r-1} + c_zt + d_z,$$

де t – параметр, який змінюється у проміжку $[0, 1]$.

Для випадку канонічного сплайна з чотирма контрольними точками коефіцієнти рівнянь можна обчислити за формулами:

$$a_x = T(x_2 - x_0) + T(x_3 - x_1) + 2x_1 - 2x_2,$$

$$a_y = T(y_2 - y_0) + T(y_3 - y_1) + 2y_1 - 2y_2,$$

$$a_z = T(z_2 - z_0) + T(z_3 - z_1) + 2z_1 - 2z_2,$$

$$b_x = -2T(x_2 - x_0) - T(x_3 - x_1) - 3x_1 + 3x_2,$$

$$b_y = -2T(y_2 - y_0) - T(y_3 - y_1) - 3y_1 + 3y_2,$$

$$b_z = -2T(z_2 - z_0) - T(z_3 - z_1) - 3z_1 + 3z_2,$$

$$c_x = T(x_2 - x_0),$$

$$c_y = T(y_2 - y_0),$$

$$c_z = T(z_2 - z_0),$$

$$d_x = x_1,$$

$$d_y = y_1,$$

$$d_z = z_1,$$

де $x_0, y_0, z_0, x_1, y_1, z_1, x_2, y_2, z_2, x_3, y_3, z_3$ – координати чотирьох контрольних точок $P_0(x_0, y_0, z_0)$, $P_1(x_1, y_1, z_1)$, $P_2(x_2, y_2, z_2)$ і $P_3(x_3, y_3, z_3)$;

T – коефіцієнт натягнення (зазвичай дорівнює 0,5).

Канонічний сплайн плавно проходить через усі свої контрольні точки (рис. 3).

Для тривимірної дуги, що потрібно апроксимувати, знаходимо таку криву, яка щонайкраще передасть її форму. Це означає, що потрібно обчислити координати контрольних точок кривої, отже, коефіцієнти сплайна. Таким чином тривимірна дуга буде представлена вектором ν, компонентами якого є координати контрольних точок канонічного сплайна:

$$\nu = \{x_n, y_n, z_n\},$$

Мірою подібності канонічного сплайна та вихідної тривимірної дуги будемо використовувати оцінку площі фігури, обмеженої контуром, що утворений цим сплайном та ламаною лінією, яка з'єднує точки вихідної тривимірної дуги $S(G, P)$. Тоді пошук сплайна, який щонайкраще апроксимує вихідну тривимірну дугу, полягатиме в обчисленні оптимальних коефіцієнтів ν_{opt}, при яких значення оцінки площі буде мінімальним:

$$\nu_{opt} = \arg\min_{\nu} S(G, P),$$

де G – точки вихідної тривимірної дуги $g_m(x_{gm}, y_{gm}, z_{gm})$;

P – точки канонічного сплайна $p_m(x_{pm}, y_{pm}, z_{pm})$;

$m = 0, M$;

M – кількість точок у вихідній тривимірній дузі.

Мінімізації міри подібності досягаємо методом градієнтного спуску.

Наведемо загальний алгоритм ітераційної процедури апроксимації тривимірної дуги, що відповідає QRS-комплексу електрокардіограми, параметричним канонічним сплайном з використанням методу градієнтного спуску.

На початку алгоритму маємо довільне наближення сплайна до вихідної тривимірної дуги. Оцінюємо міру подібності та виконуємо один крок ітерації: змінюємо координати контрольних точок канонічного сплайна в одному з можливих напрямків. Яку координату змінювати, обираємо, розрахувавши, який з напрямків дасть найкращу зміну міри подібності, тобто найкраще зменшення оцінки площі. Це відбувається у напрямку градієнта. Повторюємо оцінку міри подібності та виконуємо новий крок у напрямку градієнтного спуску. Таку процедуру повторюємо, доки оцінка міри подібності не стане меншою за деяке попередньо задане значення ε. Можлива ситуація, коли відсутня збіжність ітераційного алгоритму. У цьому

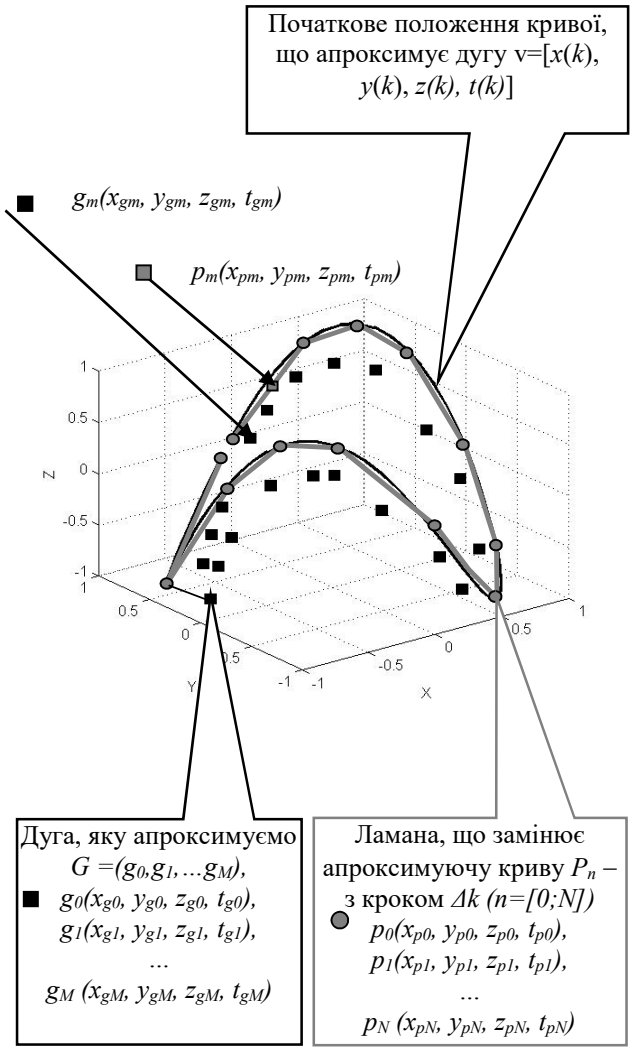


Рис. 4. Апроксимація тривимірної дуги

випадку є блокування: кількість ітерацій обмежена.

Графічна інтерпретація алгоритму наведена на рис. 4.

IV. Перевірка методу

Для перевірки алгоритму був проведений модельний експеримент, у якому виконувалась класифікація ЕКГ для декількох пацієнтів. Для цього ЕКГ були зареєстровані, попередньо оброблені та розділені на елементи кардіоциклів. З цих вихідних даних були сформовані тривимірні дуги, що відповідають QRS-комплексам у трьох відведеннях. Далі ці дуги апроксимовані канонічними сплайнами (рис. 5). Наступним кроком є прийняття рішення про приналежність електрокардіограми до одного з класів, тобто одному з пацієнтів. Воно приймається за результатами порівняння координат точок, що відповідають за форму апроксимуючого сплайна, за допомогою нейромережі.

Для експерименту зареєстрували електрокардіограми різних людей, функціонально здорових, віком від 25 до 55 років, чоловічої і жіночої статі. Дослідження проводились у різний час доби, пацієнти знаходились у різному

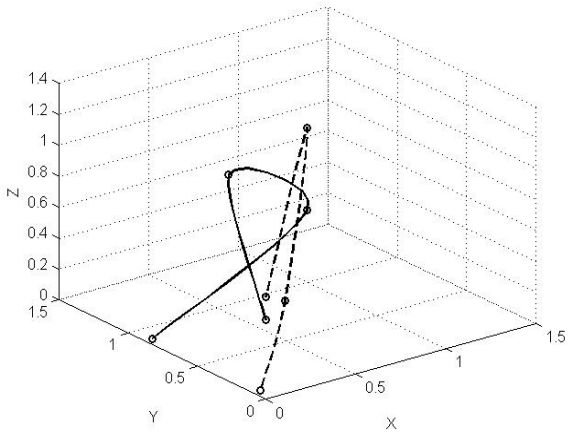


Рис. 5. Результати апроксимації ЕКГ у трьох відведеннях двох людей

емоційному та фізичному стані. Для кожної персони були зареєстровані по 10 електрокардіограм.

Була проведена попередня обробка даних. Для кожної електрокардіограми виконані такі дії: перетворення 12 відведень на три ортогональні відведення, розділення сигналу на кардіоцикли, виділення у кожному з них QRS-комплексу. На наступному етапі визначили репрезентативний цикл електрокардіограми за допомогою метрики Хаусдорфа. Далі всі тривимірні QRS-комплекси електрокардіограми апроксимували канонічними сплайнами. Після цього за форму кожного QRS-комплексу відповідають по чотири контрольні точки канонічного сплайну.

Для класифікації апроксимованих електрокардіограм використовували нейромережу. Вихідні ЕКГ розподіляли на два класи. Провели два варіанти експерименту. Виконували попарну класифікацію, тобто електрокардіограми кожного пацієнта порівнювали з кожним, а також класифікацію «один до всіх» - електрокардіограми кожного пацієнта розрізняли з-поміж електрокардіограм решти пацієнтів. В обох варіантах навчання нейромережі виконувалось на одній з електрокардіограм пацієнта, а решта дев’ять – використовувались для ідентифікації. Навчання виконувалось за всіма циклами кардіограми, а класифікація – за репрезентативним циклом.

У ході експерименту ЕКГ були успішно розділені на класи, що відповідають різним пацієнтам. Правильне рішення про приналежність електрокардіограми було прийняте у приблизно 97% випадків.

V. Висновки

Виконання апроксимації тривимірної дуги параметричними сплайнами для трьох відведень електрокардіограми дозволяє побудувати систему ознак, що будуть чутливі до форми цієї дуги. Таку систему ознак зручно використовувати для подальшої класифікації за допомогою нейромережі, що необхідно як при розв’язанні задачі ідентифікації пацієнта, так і при проведенні діагностики.

Література References

[1] И.А. Чайковский. Анализ электрокардиограммы в одном, шести и двенадцати отведениях с точки зрения информационной ценности: электрокардиографический каскад // Клиническая Информатика и Телемедицина. – 2012. – № 2. – С.102-106.

[2] L.S. Fainzilberg. Computer Analysis and Recognition of Cognitive Phase Space Electrocardiographic Image / Fainzilberg L.S., Potapova T.P. // Proceeding of 6th International Conference on Computer analysis of Images and Patterns (CAIP-95). – Prague, 1995. – P. 668-673.

[3] В.В. Вишневський, Т.М. Романенко, Л.А. Кізуб. Біометрична ідентифікація за допомогою електрокардіограми // П'ята міжнародна науково-практична конференція "Інформаційні технології та комп'ютерна інженерія" ІТКІ 2015. – Івано-Франківськ. – 2015. – С. 130-131

[4] В.Г. Калмыков, В.В. Вишневський, Т.М. Романенко. Аппроксимация плоских и пространственных кривых параметрически заданными сплайнами // Матеріали міжнародної наук.-практ. конф. "Системи підтримки прийняття рішень. Теорія і практика (СППР 2015)". - Київ: ІПММС НАНУ, 2015. – С. 165-169

[5] В.В. Вишневський, В.Г. Калмыков, Т.М. Романенко. Аппроксимация одно-, дву- и трехмерных дуг кривых параметрическими сплайнами // Математичні машини і системи. – 2015. – №4. – С.57-64.

Алгоритми Побудови Ортогональних Базисів на Основі Рекурсивних Функцій Ґалуа

Наталія Превисокова
кафедра інформатики
Прикарпатський національний університет
Івано-Франківськ, Україна
prevysokova.n@gmail.com

Любомир Петришин
кафедра управління
Університет AGH
Краків, Польща
L.b.petryshyn@gmail.com

Algorithms for Constructing Orthogonal Basis From the Recursive Galois Functions

Nataliia Prevysokova
Department of Computer Science
Vasyl Stefanyk Precarpathian National University
Ivano-Frankivsk, Ukraine
prevysokova.n@gmail.com

Lubomyr Petryshyn
dept. of Computer Science
AGH University
Krakow, Poland
L.b.petryshyn@gmail.com

***Анотація*—Розроблено алгоритми побудови ортогональних систем функцій на основі рекурсивної системи функції Ґалуа для виконання дискретних перетворень в цифровій обробці сигналів. Доведено властивості побудованих базисів.**

***Abstract*—It is developed algorithms of orthogonal functions construction on recursive Galois functions system base to perform a discrete transform in digital signal processing. The properties of the built bases are proved.**

Ключові слова—дискретне ортогональне перетворення; системи функцій Ґалуа; базис перетворення; властивості базису

Keywords—discrete orthogonal transform; Galois functions system; transform base; basis properties

I. Вступ

Останнім часом поширеними є методи цифрової обробки інформаційних потоків на основі дискретних перетворень, серед яких найчастіше використовуються дискретні перетворення Фур'є, косинусне, Уолша, Хаара, вейвлет-перетворення та ін. [1-3]. Дискретні ортогональні перетворення використовуються для кодування, обробки, зменшення надлишковості інформації, зокрема, при обробці зображень, у системах зв'язку, спектроскопії, радіолокації.

Підвищення вимог щодо швидкості і якості обробки інформації зумовлює необхідність розробки методів ефективних перетворень [1-3]. Однією із важливих проблем реалізації методів обробки інформації є вибір базису перетворення [1, 3], придатного для подання інформаційного потоку.

Дискретне пряме та обернене перетворення одновимірного інформаційного потоку $X(t)$ зі скінченною енергією, заданого вектором $X = \left[X(0), X(1), ..., X(N-1)\right]$ у базисі ψ_i, визначаються за формулами

$$Y_i = \left\langle X, \psi_i \right\rangle,$$

$$X = \sum_{i=0}^{N-1} \left\langle X, \psi_i \right\rangle \tilde{\psi}_i,$$

де Y_i – координати вектора коефіцієнтів перетворення (спектр) $Y = \left[Y(0), Y(1), ..., Y(N-1)\right]$, $i = 0,1,...,N-1$; $\langle , \rangle$ – позначення скалярного добутку; $\tilde{\psi}_j$ – біортогональний (взаємний) базис до ψ_i, $\left\langle \psi_i, \tilde{\psi}_j \right\rangle = \delta_{i,j}$, $\delta_{i,j}$ – символ Кронекера: $\delta_{i,j} = 1$ для $i = j$, $\delta_{i,j} = 0$ для $i \neq j$; $X(t) = X(j)$, $t \in [j, j+1)$; $\left\langle X(t), \psi_i(t) \right\rangle = \int_a^b X(t)\psi_i(t)dt$ або $\left\langle X, \psi_i \right\rangle = \sum_{j=0}^{N-1} X_N(j)\psi_i(j)$.

Ортогональний базис одночасно біортогональний до себе, тому ортогональне перетворення є частинним випадком біортогонального.

У матричній формі дискретне перетворення вхідного інформаційного потоку X в базисі, заданому матрицею M, подається згідно

$$Y = MX \ .$$

Оптимальним для подання є перетворення Карунена-Лоева, яке потребує виконання обчислень матриць власних значень та власних векторів для кожного вхідного інформаційного потоку, тому на практиці не використовується [1 – 3]. Це зумовлює необхідність дослідження та формування нових базисів, які дозволяють ефективно розв'язувати окремі задачі цифрової обробки інформації.

Автором розвивається напрям синтезу базисів перетворень і розв'язується задача розробки алгоритмів побудови ортогональних базисів на основі рекурсивних систем функцій Ґалуа.

II. Основна Частина

Рекурсивні системи функцій Ґалуа [2] $\{Gal(n,\theta,i)\}$, $i = 0,1,...,2^n - 1$ утворюються відповідно до породжуючого вектора поля Ґалуа $GF(2^n)$.

Зокрема, у полі $GF(2^4)$ існують чотири породжуючі вектори: $(1,0,0,1,1)$, $(1,0,0,1,\bar{1})$, $(1,1,0,0,1)$, $(1,1,0,0,\bar{1})$. На основі даних векторів із початкових векторів $(g_0, g_1, g_2, g_3) = (1,1,1,1)$ і $(g_0, g_1, g_2, g_3) = (0,0,0,0)$ генеруються рекурсивні послідовності $g_0, g_1, g_2, \dots g_{2^n-2}$ за правилами, які відповідають породжуючим векторам [2], наприклад $(1,0,0,1,1) \to g_{j+4} = g_j \oplus g_{j+3}$; $(1,1,0,0,1) \to g_{j+4} = g_j \oplus g_{j+1}$; $j = 0,1,...,2^n - 2$. Фрагмент із $n-1$ нульових елементів послідовності доповнюється нулем.

Рекурсивні функції Ґалуа $\{Gal(n,\theta,i)\}$ в точках $\theta = \frac{j}{N}$ визначаються із послідовностей $\{g_j\}$ та доозначаються до неперервних функцій на інтервалах $\theta \in [\frac{j}{N}; \frac{j+1}{N})$, $N = 2^n$

$$Gal(n,\theta,0) = Gal(n,\frac{j}{N},0) = 1 - 2g_j \ , \qquad (1)$$

$$Gal(n,\theta,i+1) = Gal(n,\theta+\frac{1}{N},i) \ .$$

Відповідно до наведених процедур побудови матриці рекурсивних систем функцій Ґалуа є тоепліцевими циклічними або ганкелевими антициклічними (циркулянтами) [2].

У результаті обчислення рангів матриць рекурсивних систем функцій Ґалуа різних порядків встановлено, що будь-яка підсистема із $2^n - 1$ функцій є лінійно незалежною, оскільки ранг її матриці дорівнює $2^n - 1$, водночас, будь-яка функція системи із 2^n функцій є лінійною комбінацією інших функцій.

З властивості $\sum_{s=0}^{2^n-1} Gal(n,\theta_s,i) = 0$ та властивості симетрії $\sum_{i=0}^{2^n-1} Gal(n,\theta_s,i) = 0$ слідує, що $Gal(n,\theta_s,j) = -\sum_i Gal(n,\theta_s,i)$, де $i \neq j$ і кожна функція є лінійною комбінацією інших функцій системи. Зокрема, остання функція є лінійною комбінацією інших функцій системи.

Отже, рекурсивна система $N = 2^n$ функцій Ґалуа є лінійно залежною і не утворює повної системи в просторі $L_2[0,1)$, що обмежує її застосування для розкладання інформаційних потоків.

Розроблено два алгоритми побудови ортогональних базисів на основі рекурсивних функцій Ґалуа.

A. Алгоритм 1.

На першому етапі алгоритму рекурсивна система функцій Ґалуа [2] модифікується шляхом зміни нумерації перших $2^n - 1$ функцій

$$Gal_m(n,\theta,i) = Gal(n,\theta,i-1) \ ,$$

де $\theta \in [0;1)$, $n = 1,2,\dots$ – порядок функції, $i = 0,1,...,2^n - 1$.

На другому етапі здійснюється доповнення системи одиничною функцією

$$Gal_m(n,\theta,0) = 1 \ .$$

На третьому етапі до функцій одержаної на попередніх двох етапах модифікованої системи застосовується процедура ортогоналізації Грама-Шмідта [3], в результаті якої одержують ортогональні функції $\{G(n,\theta,i)\}$

$$G(n,\theta,0) = 1,$$

$$G(n,\theta,1) = Gal_m(n,\theta,1) - \frac{\langle Gal_m(n,\theta,1), G(n,\theta,0)\rangle}{\|G(n,\theta,0)\|_{L_2}^2} G(n,\theta,0),$$

$$G(n,\theta,2) = Gal_m(n,\theta,2) - \frac{\langle Gal_m(n,\theta,2), G(n,\theta,0)\rangle}{\|G(n,\theta,0)\|^2_{L_2}} G(n,\theta,0) -$$

$$\frac{\langle Gal_m(n,\theta,2), G(n,\theta,1)\rangle}{\|G(n,\theta,1)\|^2_{L_2}} G(n,\theta,1),$$

.....

$$G(n,\theta,k+1) = Gal_m(n,\theta,k+1) -$$

$$-\sum_{i=0}^{k} \frac{\langle Gal_m(n,\theta,k+1), G(n,\theta,i)\rangle}{\|G(n,\theta,i)\|^2_{L_2}} G(n,\theta,i), \qquad (2)$$

де $k = 0,1,\ldots,N-1$, $\|G(n,\theta,i)\|^2_{L_2}$ – норма в просторі інтегровних з квадратом функцій $L_2[0,1)$, $\langle Gal_m(n,\theta,k+1), G(n,\theta,i)\rangle$ – скалярний добуток.

Для прикладу, наведемо матрицю значень ортогональних функцій Ґалуа $G(3,\theta,i)$ в точках $\theta = \frac{j}{8}$

$$\begin{Vmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & 1 & -1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 & -1 & -1 & 1 & 1 \\ -1/2 & 3/2 & -1/2 & -3/2 & -1/2 & 1/2 & 1/2 & 1/2 \\ 3/2 & -1/2 & -3/2 & -1/2 & 1/2 & 1/2 & 1/2 & -1/2 \\ 1/6 & -1/2 & -1/6 & -1/2 & 5/6 & 1/6 & -5/6 & 7/6 \\ -2/5 & 2/15 & -2/5 & 2/15 & 2/3 & -14/15 & 2/3 & 2/15 \end{Vmatrix}.$$

На рис. 1 наведені графіки одержаної в результаті ортогоналізації системи ортогональних функцій Ґалуа $G(3,\theta,i)$.

Рис. 1. Система ортогональних функцій Ґалуа, $n = 3$.

У результаті дослідження залежностей (2) здійснено факторизацію матриці перетворення G і подання її у формі добутку матриць

$$G = C \times Gal_m,$$

де C – трикутна матриця ортогоналізації $N \times N$, Gal_m – матриця значень функцій $\{Gal_m(n,\theta,i)\}$ в точках $\theta = \frac{j}{N}$,

$$C = \begin{bmatrix} c_{00} & 0 & \ldots & 0 \\ c_{10} & c_{11} & .. & 0 \\ \ldots & \ldots & \ldots & \ldots \\ c_{N-1,0} & c_{N-1,1} & \ldots & c_{N-1,N-1} \end{bmatrix}.$$

Елементи матриці C є дійсними числами, які визначаються рекурсивно

$$c_{k+1,j} = Gal_m(n, \frac{j}{N}, k+1) - \sum_{i=0}^{k} \frac{\sum_{l=0}^{N-1} c_{il} \cdot Gal_m(n, \frac{l}{N}, k+1)}{\sum_{l=0}^{N-1} (c_{il})^2} c_{ij}.$$

Встановлено, що для всіх $k \le n$, $j \le n$ елементи матриці C задовольняють умови $c_{kk} = 1$; $c_{kj} = 0, k > j$.

Доведено наступні властивості одержаної на основі даного алгоритму системи функцій [4].

1) Функції системи лінійно-незалежні та ортогональні.

Система, одержана в результаті ортогоналізації, є ортогональною. Ортогональна система функцій є лінійно незалежною.

2) Система функцій є повною в просторі $L_2[0,1)$.

Встановлено, що ортогональні функції Ґалуа утворюють повну систему функцій у просторі $L_2[0,1)$, оскільки система функцій Уолша повна, та досліджено, що оператор лінійного перетворення матриці Уолша розміру $N \times N$ в матрицю Ґалуа такого ж розміру є обмеженим. Властивість повноти дозволяє застосовувати побудовану систему в якості базису для розкладання інформаційних потоків та виконання дискретних ортогональних перетворень.

3) Оскільки система є повною, то може використовуватись в якості базису дискретного ортогонального перетворення.

B. Алгоритм 2.

На першому етапі будується рекурсивна послідовність аналогічно як в першому алгоритмі: вибирається примітивний незвідний поліном певного степеня n і на

основі полінома генеруються періодичні рекурсивні послідовності $g_0, g_1, g_2, \ldots g_{2^n-2}$ з періодом $M = 2^n - 1$.

На другому етапі з рекурсивної послідовності із $2^n - 1$ елементів утворюється перша функція $\{GR(n,\theta,1)\}$ системи $\{GR(n,\theta,i)\}$ та доозначається до неперервної функції на інтервалах $\theta \in [\frac{j}{N}; \frac{j+1}{N})$ за правилом

$$GR(n, \frac{j}{N}, 1) = 1 - 2g_j,$$

$$GR(n,\theta,1) = GR(n, \frac{j}{N}, 1).$$

На третьому етапі здійснюється побудова функцій $\{GR(n,\theta,2)\}, \{GR(n,\theta,3)\}, \ldots, \{GR(n,\theta,2^n-1)\}$ за допомогою циклічного зсуву першої функції та кожної наступної одержаної функції вправо на $\frac{1}{N}$, а також паралельне перенесення всіх функцій на інтервал $\theta \in [\frac{1}{N}; 1)$:

$$Gal_{ort}(n,\theta,i) = GR(n,\theta - \frac{1}{N}, i-1).$$

На четвертому етапі система із $2^n - 1$ функцій доповнюється одиничною функцією

$$G_{ort}(n,\theta,0) = 1.$$

На наступному етапі здійснюється доповнення значень всіх функцій на інтервалі $\theta \in [0; \frac{1}{N})$

$$G_{ort}(n,\theta,i) = 1.$$

Матриця значень ортогональних функцій системи $Gal_{ort}(3, \frac{j}{N}, i)$ на прикладі n=3

$$\begin{Vmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 & -1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & -1 & 1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 & 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & 1 & 1 & -1 & -1 & -1 \\ 1 & -1 & 1 & 1 & -1 & -1 & -1 & 1 \end{Vmatrix}$$

Доведено наступні властивості одержаної на основі другого алгоритму системи функцій.

1) Функції системи лінійно-незалежні та ортогональні;

2) Система функцій є повною в просторі $L_2[0,1)$ і може використовуватись в якості базису дискретного ортогонального перетворення.

Встановлено, що побудовані функції утворюють повну систему функцій у просторі $L_2[0,1)$, оскільки система функцій Уолша повна, та досліджено, що оператор лінійного перетворення матриці Уолша розміру $N \times N$ в матрицю побудованої системи такого ж розміру є обмеженим. Властивість повноти дозволяє застосовувати побудовану систему в якості базису для розкладання інформаційних потоків та виконання дискретних ортогональних перетворень.

3) Матриця оберненого перетворення співпадає з матрицею прямого перетворення, що дозволяє використовувати однакові алгоритми прямого та оберненого перетворень.

III. Висновки

Таким чином, розроблено два алгоритми побудови повних ортогональних систем функцій на основі рекурсивної системи функції Ґалуа. Перший алгоритм реалізується на основі процедури ортогоналізації Грама-Шмідта модифікованої лінійно незалежної рекурсивної системи функцій Ґалуа, доповненої одиничною функцією. Другий алгоритм базується на заміні окремих значень функцій системи. Доведено властивості побудованих базисів.

Запропоновані алгоритми дозволяють будувати повні ортогональні базиси для спектрального аналізу на основі виконання дискретних ортогональних перетворень інформаційних потоків.

Література References

[1] А. И. Солодовников, А. М. Спиваковский, Основы теории и методы спектральной обработки информации, Л.: Изд-во Ленингр. ун-та, 1986.

[2] Л. Б. Петришин, Теоретичні основи перетворення форми та цифрової обробки інформації, К.: ІЗіМН МОУ, 1997.

[3] В. В. Солодовников, А. Н. Дмитриев, Н. Д. Егупов, Спектральные методы расчета и проектирования систем управления, М.: Машиностроение, 1986.

[4] Н. В. Превисокова. Метод обробки інформації на основі дискретного ортогонального перетворення Ґалуа / Н. В. Превисокова // Вісник Хмельницького нац. ун-ту. Технічні науки. – 2010. – № 2 (146). – С.149 –156.

The Dynamic Information in the Images

Vitaliy Boyun, C.-Member of National Academy of Sciences of Ukraine
Department of Intelligent Real Time Video Systems,
V.M.Glushkov Institute of Cybernetics NASU,
Kyiv, Ukraine
vboyun@gmail.com

Динамічна Інформація в Зображеннях

В.П. Боюн, чл.-кор. НАН України
Відділ інтелектуальних відеосистем реального часу
Інститут кібернетики імені В.М.Глушкова НАНУ
Київ, Україна
vboyun@gmail.com

***Abstract*—Taking into consideration the importance of texture features in the problems of image segmentation, texture evaluation and classification, object recognition and so on, the approach to the texture features extraction with the image dynamics is proposed in this work. The conception of the image δ-entropy, which represents mean value of the signal derivative magnitude, is introduced. It allows detecting and evaluating of the amount of useful (dynamic) information in the image. In addition, it allows extracting from the cross section of the image the following texture features: space frequencies over the image rows and columns, granularity of texture, contrast, directivity and regularity. Such approach allows rather simply conducting a segmentation of the images with the areas of texture type, evaluating of texture parameters and classifying of textures, simplifying of the object recognition with the information from the cross section of the image.**

***Анотація*—Враховуючи важливість текстурних ознак в завданнях сегментації зображень, оцінки і класифікації текстур, розпізнавання об'єктів і тому подібне, в роботі пропонується підхід до виділення текстурних ознак по динамічності зображення. Вводиться поняття δ-ентропії зображення, яке є середнім значенням модуля похідної сигналу. Це дозволяє виділяти і оцінювати кількість корисної (динамічної) інформації в зображенні. Крім того, дозволяє виділяти з поперечних зрізів зображення ряд таких текстурних ознак, як: просторові частоти по рядках і стовпцях зображення, зернистість текстури, контраст, спрямованість, регулярність. Такий підхід дозволяє досить просто проводити сегментацію зображень з ділянками типу текстур, оцінювати параметри і проводити класифікацію текстур, спростити розпізнавання об'єктів за інформацією поперечних зрізів зображення.**

Keywords—texture; δ-entropy; dynamic informationм; segmentation; texture characteristics; object recognition

Ключові слова—текстура; δ-ентропія; динамічна інформація; сегментація; текстурні характеристики; розпізнавання об'єктів

I. Introduction

Object search in the image is one of the fundamental problems of the image processing. The quality of the search algorithms depends on the selected set of the informative features, which describe the searched object most adequately. Specific characteristics for the given field are extracted in the concrete subject area. Those can be common characteristics for the object search in the image: color, texture, shape, orientation in the space and the object motion.

Principal consideration will be given to the issues of the segmentation and the extraction of the informative features in the images of texture type in this work.

II. State of Problem

Texture contains the information about the structure of the surfaces, which constituent the image. It has properties of periodicity and a scale, and it can be described by means of the extraction of orientation, contrast features and blurs. Texture complements the color characteristic and goes far during image comparison.

Taking into consideration the variety of texture types, the big arsenal of the methods for extraction of informative features from the textures is developed. The methods for extraction of the texture features are divided into statistical, geometrical, model and spectral ones [1].

Statistical features describe the distribution of the brightness level over the image by the instrumentality of such statistical characteristics:

- simplest statistical parameters, computed over the brightness values of each pixel;
- parameters, computed in accordance with the adjacency matrices;
- texture histograms, constructed in accordance with the Tamura features and others.

The texture determination by the instrumentality of the constituent texture elements and primitives is common for the geometrical features.

Model methods of texture analysis assume construction of the model, which helps to describe the texture and synthesize it.

Spectral features describe the texture in the frequency domain, using the signal decomposition on the certain basis. At that, the decomposition coefficients act as the elements of the feature vector.

Implementation complexity by means of computer engineering is the common drawback of all these methods.

Therefore, the aim of this work is the development of new approaches to the texture evaluation, using as a prototype the mechanisms of the information perception by the human visual analyser.

III. δ-ENTROPY AS THE MEASURE OF THE IMAGE DYNAMICS

On one hand, arbitrary area of field of view with contrast light and dark areas may be analyzed and transformed into the space frequency – number of luminosity variations on the defined space area, or cycles number of alternation of dark and light stripes on the given space area. More stripes are per the unit of the pattern area; space frequency of that pattern is more that is the patterns with the high space frequencies are made of small elements and vice versa. The space frequency per one degree of the angle of view (cycles/degree) is used practically. As it is known, visual pattern represents the pattern of luminosities that is the sum of a series of simple components. The complex space distribution of entire visual field luminance may be transformed into sinusoidal waves of simple components using mathematical apparatus of the Fourier analysis. Inverse action is summing of sinusoidal waves for the complex pattern synthesis. There are the evidences that there are the detectors of space frequency in the visual system – the specialized cells, which are sensual maximally to certain space frequencies. Thus, our perception of any complex pattern is the result of analysis and synthesis of the constituent parts of space frequencies of that pattern by the visual system and the brain reconstructs visual pattern of the object, integrating the information about different space frequencies [2].

On another hand, a human eye responds not to a magnitude of brightness or chromaticity in the image, but on changes of these variables between brightness values of neighbouring receptors or brightness values of a particular receptor in time, i.e. dynamics of this parameter. Taking into consideration the need to determine changes in brightness between adjacent elements in a row and a column of the matrix picture, it can be represented as a matrix of differences between adjacent elements (matrix increments) [2].

Glushkov V.M. gave in [3] such capacious determination of the information concept that it did not lose its actuality to the presence: " Information is most general own understanding presents the measure of ***the heterogeneity*** of the matter and energy distribution in time and in space, the measure of changes, which accompany all running processes in the world". At that two information kinds are marked out – ***static*** information (which characterizes current state of the certain material or energetic system) and ***dynamic*** information (its variability in time and in space). If the hundreds of the works were dedicated to the static information and it became already classic one, no necessary attention was attended to the dynamic information practically till late XX century. However, the determination of the dynamic information concept turned out extremely fruitful under the study of the information properties of the physical systems and processes.

The author laid down the foundations of theory of the dynamic information in the monograph [4]. That allowed extraction and use of the useful (dynamic) information out of the steady stochastic and unsteady signals, images, spatial fields, iteration processes, recurrent procedures and the like with considerable information redundancy reduction.

In contrast to the statistical information theory C. Shannon, which was created for compression of information during transmission and storage, it proposed a dynamic theory of information, which helps to distinguish useful (dynamic) information from signals, images, sequences, iterative processes, and the like, greatly reducing its redundancy, and can be the basis of transformation processes and information processing in real-time systems. This introduced the concept of entropy the value of the parameter random variable and δ-entropy that is a measure of the changes taking place in space or time. ***δ-entropy*** is defined as the average value of the modulus of the derivative of the video signal on rows or columns of the image, and for video stream – between frames [4].

δ-entropy is defined in discrete form for the evaluation of the amount of information in the image and in the video sequence as:

$$H = \sum_{j=1}^{N} \sum_{i=1}^{M} p_{ij} \log_2 \left(\left| \frac{\Delta z_{ij}}{\delta z} \right| + 1 \right), \qquad (1)$$

where: MN – dimensions of the image field,

Δz_{ij} – differences between brightness of the pixels in rows and columns.

δ -entropy is a measure of the dynamic of image and can be used to evaluate the amount of information in the image and video stream. In addition it is proposed to use a dynamic measure for finding highly dynamic areas, the textures in the image, the selection of informative features for their recognition, and the like.

A known method of averaging the cross-sectional profile of the brightness of the image by rows and columns [5] allows you to determine on the image location of the object on the background, if it differs from the background, however, does not allow allocating objects of type texture.

To search the location of such objects we propose to carry out averaging of the cross-section of image on dynamic features, i.e. the brightness difference between neighbouring pixels in rows and columns [6].

$$H\delta_i = \frac{1}{M}\sum_{j=1}^{M}\left|\frac{\Delta Z_{ij}}{\delta z}\right|; \qquad H\delta_j = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\Delta Z_{ij}}{\delta z}\right|. \qquad (2)$$

This dynamic measure of information can be effectively used for:

segmentation of the image into high and low dynamic plot;

segmentation of text information in the image;

search and classification of textures;

barcode search, DMX-codes, fingerprints, numbers, car;

selection of motion, change of brightness or chromaticity, etc;

speed control shooting video camera and the like.

Tools are developed, which enable to extract an arbitrary area of the image and to construct diagrams of the module-averaged cross-sections of the changes of brightness along columns and rows, and also to evaluate the mean value of the module of changes across entire image in accordance with the formula:

$$H\delta = \frac{1}{MN}\sum_{j=1}^{M}\sum_{i=1}^{N}\left|\frac{\Delta Z_{ij}}{\delta z}\right|. \qquad (3)$$

The simplified method with the brightness jumps with the sign in the rows and columns can be used in the several cases for the image dynamics evaluation. In these cases it is reasonable to introduce the verification over the jump value threshold or to conduct preliminary low-frequency filtration for the noised images.

The examples of application of the proposed method are given below (Fig. 1. – Fig. 7.).

Fig. 1. Segmentation of the text information.

The text columns are extracted on the vertical cuts, and on the horizontal cuts – the text rows, even if they are overlapped by the gray scale images.

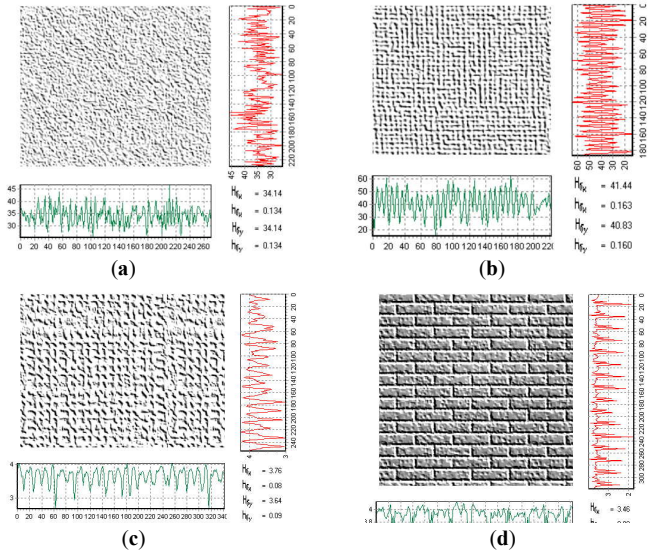


Fig. 2. Evaluation of texture parameters: a) sand, b) linen, c) sacking, d) brickwork.

The step (grain) of texture and mean values of the changes of brightness or chromaticity can be extracted on the figures. The locations of the regularity disturbances (defects) can be detected on the regular textures, for example on the Fig. 2 (c).

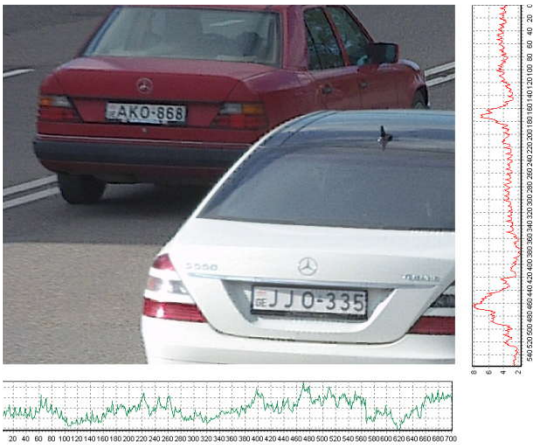


Fig. 3. Search of the car number in the car image.

The car number is clearly extracted on the horizontal cuts and less clearly on the vertical cuts.

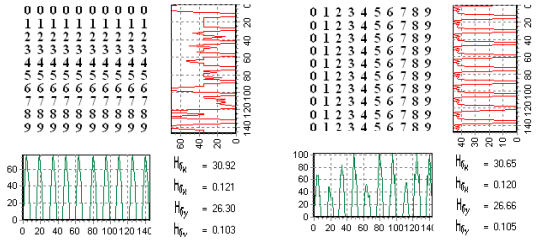


Fig. 4. Dynamical characteristics of the numerals.

All characters on the horizontal cuts a) have the differences in the amplitudes and in the quantity of extreme, which can be used for their recognition. In addition, the amplitude values of the vertical cuts b) may be also used.

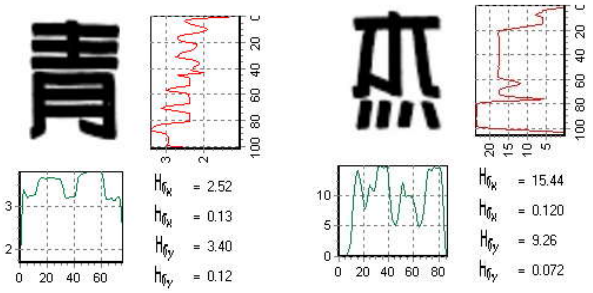


Fig. 5. Dynamical characteristics of Han characters.

Dynamic characteristics allow more simply recognizing more complex figures, for example, Han characters.

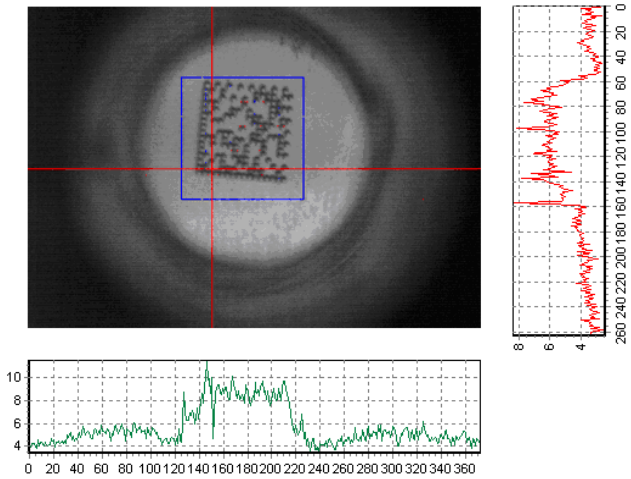


Fig. 6. Searching of the DMX-code in the micro image.

Dynamic characteristics allow determining the location of the bar-codes and DMX-codes on the products.

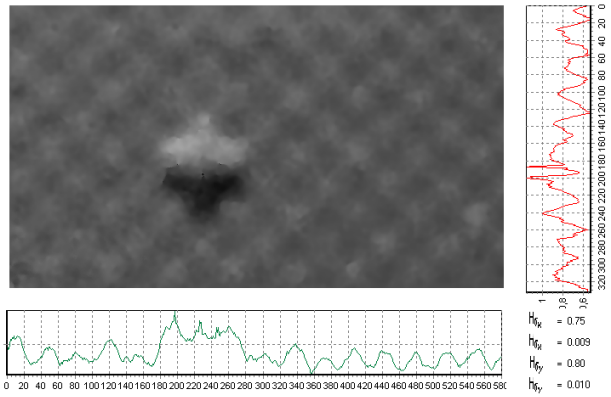


Fig. 7. Searching for defects on the cellular panel.

Dynamic characteristics can be used effectively for the quality control of the different texture surfaces of the products.

IV. Conclusions

The proposed measure of the dynamic information allows evaluating the amount of useful information in the image and in addition allows extracting a series of texture features from the cross section of the image for the image segmentation, for the texture evaluation and classification and for the simplification of recognizing the images of the texture type.

References

[1] R. Gonsales, R. Woods. Digital image processing. Moscow, Russia, Technosphere, 2005. 1072 p. (in Russian)

[2] H.R. Schiffmann. Sensation and perception. 5-th ed. Piter, SPb., Russia, 2003. 928 p. (in Russian)

[3] V. Glushkov. About cybernetics as a science, Cybernetics, thinking, life. Moscow, USSR, 1964, pp. 53-62 (in Russian)

[4] V. Boyun. The dynamic theory of information. Fundamentals and applications. Institute of Cybernetics of NASU, Kyiv, Ukraine, 2001, 326 p. (in Russian)

[5] W. Prett. Digital Image processing. Mir, Moscow, USSR, 1982. vol.2, 560 p. (in Russian)

[6] V. Boyun. Intelligent selective perception of visual information. Informational aspects. Artificial intellect. 2011. №3. pp.16-24. (in Ukrainian)

Застосування Трійкових Симетричних Функцій для Цифрової Обробки Інформації на Основі Ортогональних Перетворень

Артем Ізмайлов
кафедра інформатики
Прикарпатський національний університет
Івано-Франківськ, Україна
aiartefact@gmail.com

Application of Symmetric Ternary Functions for Digital Information Processing Based on Orthogonal Transforms

Artem Izmailov
dept. of Computer Science
Precarpathian National University
Ivano-Frankivsk, Ukraine
aiartefact@gmail.com

***Анотація*—Проаналізовано можливість застосування трійкових симетричних функцій для цифрової обробки інформації на основі ортогональних перетворень. Застосовано логічні операції трійкової симетричної логіки для синтезу нової системи функцій на основі трійкових симетричних, яка може бути використана для цифрової обробки інформації на основі ортогональних перетворень.**

***Abstract*—Application possibility of symmetric ternary functions for digital information processing based on orthogonal transforms was analyzed. Logical operations of symmetric ternary logic were applied for synthesis of the new system of functions based on symmetric ternary functions that can be applied for digital information processing based on orthogonal transforms.**

Ключові слова—цифрова обробка інформації; ортогональне перетворення; трійкові симетричні функції; логічні операції

Keywords—digital information processing; orthogonal transform; symmetric ternary functions; logical operations

I. Вступ

Цифрова обробка інформації є ключовим елементом багатьох технічних систем, які використовуються у різних галузях економіки, управління, виробництва, зв'язку та медицини [1 – 4]. Відповідно, ефективні рішення у галузі цифрової обробки інформації призведуть до підвищення ефективності перебігу процесів, які включають цифрову обробку інформації, у прикладних галузях.

Одним із актуальних завдань цифрової обробки інформації є обробка цифрових сигналів на основі ортогональних перетворень [1 – 3]. У рамках розв'язання даного завдання постає проблема розробки нових ортогональних перетворень, які дозволять з вищою ефективністю проводити обробку конкретних цифрових сигналів, у тому числі сигналів, для яких існуючі методи працюють із недостатнім рівнем ефективності [3].

У більшості сучасних систем цифрової обробки інформації використовується двійкова логіка [1, 4]. Водночас використання трійкових симетричних функцій та відповідної їм логіки є обмеженим, незважаючи на доведену більшу ефективність описаної логіки у порівнянні із двійковою [4, 5].

Метою дослідження є побудова системи функцій на основі трійкових симетричних функцій, яка може бути використана для цифрової обробки інформації на основі ортогональних перетворень.

Наукова новизна отриманих результатів полягає в успішному застосуванні логічних операцій трійкової симетричної логіки для синтезу системи функцій на основі

трійкових симетричних функцій, яка може бути використана для цифрової обробки інформації на основі ортогональних перетворень.

II. Трійкові симетричні функції

Трійкові симетричні функції задаються аналітичним виразом (1) [5].

$$Ter_n(x)=\begin{cases} 0, \text{ якщо } 0 \le modh(x+\dfrac{3^n-1}{2}, 3^{n+1}) < 3^n, \\ 1, \text{ якщо } 3^n \le modh(x+\dfrac{3^n-1}{2}, 3^{n+1}) < 2\cdot 3^n, \\ -1, \text{ якщо } 2\cdot 3^n \le modh(x+\dfrac{3^n-1}{2}, 3^{n+1}) < 3^{n+1}, \end{cases} \quad (1)$$

де n – порядковий номер функції, x – цілочисельний аргумент, $mod(x,p)$ – допоміжна функція, задана аналітичним виразом (2).

$$modh(x,p)=\begin{cases} mod(x,p)+p, \text{ якщо } x<0, \\ mod(x,p), \quad \text{якщо } x \ge 0, \end{cases} \quad (2)$$

де $mod(x,p)$ – функція залишку від ділення числа x на число p.

Однак у вигляді (1) трійкові симетричні функції непридатні для цифрової обробки інформації на основі ортогональних перетворень, у зв'язку з неповнотою даної системи та її неортогональністю [6, 7].

Для подолання проблеми неможливості застосування системи (1) трійкових симетричних функцій для цифрової обробки інформації на основі ортогональних перетворень необхідно побудувати нову систему функцій на основі існуючої системи (1).

Однак, побудувати нову систему функцій, яка може бути застосована для ортогональних перетворень, на основі трійкових симетричних функції у вигляді (1) неможливо, оскільки областю визначення системи (1) є вся числова вісь, а для цифрової обробки інформації у більшості випадків застосовують функції з областю визначення на проміжку [0; 1) [1 – 3]. Це пов'язано зі зручністю проведення необхідних перетворень та обчислень [1 – 3]. Відповідно, виникає необхідність нормалізування функцій (1) до проміжку [0; 1).

Окрім приведення області визначення системи (1) трійкових симетричних функцій до проміжку [0; 1) є ще інші властивості системи функцій (1), які необхідно змінити для того, щоб на основі даної системи можна було побудувати систему функцій, придатну для цифрової обробки інформації на основі ортогональних перетворень. Детальний перелік властивостей системи (1), які підлягають зміні, разом з усіма проміжними обчисленнями можна знайти у [7]. У результаті відповідних перетворень система трійкових симетричних функцій з областю визначення на проміжку [0; 1) може бути представлена у вигляді (3).

$$Ter\,01_{n,i}(\theta)=Ter_{\left|3^n-1-i\right|}(\theta * 3^{3^n}), \quad (3)$$

де $n=\log_3 N$ – порядок набору базисних функцій теоретико-числових перетворень, N – кількість функцій у наборі, $\theta=t/T$ – параметр часу, тобто час, нормований до інтервалу T, t – поточне значення часу, $i=0,1,2,\ldots,N-1$ – порядковий номер функції, $Ter_n(\theta)$ – трійкові симетричні функції, які задані аналітичним виразом (4).

$$Ter_n(x)=\begin{cases} 0, \ 0 \le modh(x-3^n, 3^{n+1}) < 3^n, \\ 1, \ 3^n \le modh(x-3^n, 3^{n+1}) < 2\cdot 3^n, \\ -1, \ 2\cdot 3^n \le modh(x-3^n, 3^{n+1}) < 3^{n+1}. \end{cases} \quad (4)$$

Необхідно відзначити появу наборів трійкових симетричних функцій у аналітичному виразі (3). Уведення наборів пов'язане зі зміною області визначення трійкових симетричних функцій і є необхідним для перетворення трійкових симетричних функцій із вигляду (3) у вигляд (4). Детальний опис принципу роботи уведених наборів можна знайти у [7].

Система трійкових симетричних функцій (4) все ще не є повною і ортогональною, але на її основі може бути побудована система функцій, яка може бути використана для цифрової обробки інформації на основі ортогональних перетворень.

III. Система функцій на основі добутків трійкових симетричних функцій

Одним з найочевидніших варіантів утворення нової системи функцій є утворення системи добутків трійкових симетричних функцій подібно до утворення функціями Радемахера системи функцій Уолша [2]. Однак, у випадку утворення системи добутків на основі хоча б двох трійкових симетричних функцій утвориться лише $2^2=4$ функції і отримана система виявиться неповною. Неповнота даної системи випливає із того факту, що максимальна кількість змін функціями своїх значень у випадку використання трійкових симетричних функцій перших двох порядків з першого набору системи (4) (нульовий набір, як випливає з виразу (4), містить одну функцію, а перший – три) буде рівна 9, тобто для досягнення повноти даної системи необхідно, щоб вона складалась з 9 функцій (на противагу чотирьом наявним).

Для вирішення виявленої проблеми можна скористатись піднесенням до другого степеня (на противагу першому степеню, який використовується при формуванні добутків функцій Радемахера [2]) при формуванні добутків трійкових симетричних функцій. Для спрощення візуального аналізу результатів і без втрати загальності можна розглянути систему добутків трійкових симетричних функцій, утворену трійковими симетричними функціями перших двох порядків першого набору системи (4). Описана система добутків задається аналітичним виразом (5).

$$Ter\,01\,Mult_{num}(\theta)=\prod_{j=1}^{2} Ter\,01_{1,2-j}(\theta)^{A_{num,j}}, \quad (5)$$

де num=0,1,…,3^2-1 – порядок функції, $A_{num,j}$ – елемент рядка num та стовпця j кодової матриці Грея А (Рис. 1).

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 2 \\ 1 & 2 \\ 1 & 1 \\ 1 & 0 \\ 2 & 0 \\ 2 & 1 \\ 2 & 2 \end{pmatrix}$$

Рис. 1. Матриця 9×2 трійкового незваженого коду Грея

Необхідно пам'ятати, що формула (5) може бути узагальнена для довільного порядку. Отримана система функцій є повною, але не є ортогональною. Однак, її можна ортогоналізувати за алгоритмом Грама-Шмідта, опис якого можна знайти, наприклад, у [8]. Доведення факту неортогональності системи (5), як і перебіг процесу ортогоналізації за процедурою Грама-Шмідта, можуть бути знайдені у [7]. Після процедури ортогоналізації Грама-Шмідта система функцій (5) утворить нову систему, яка породжує певну матрицю значень (Рис. 2), яка може бути використана у якості матриці ортогонального перетворення.

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1/3 & 1/3 & 1/3 & -2/3 & -2/3 & -2/3 & 1/3 & 1/3 & 1/3 \\ -1 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & 1 \\ 1 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 \\ 1/3 & -2/3 & 1/3 & 1/3 & -2/3 & 1/3 & 1/3 & -2/3 & 1/3 \\ -1/3 & 2/3 & -1/3 & 0 & 0 & 0 & 1/3 & -2/3 & 1/3 \\ 1/9 & -2/9 & 1/9 & -2/9 & 4/9 & -2/9 & 1/9 & -2/9 & 1/9 \end{pmatrix}$$

Рис. 2. Матриця значень ортогоналізованої системи функцій (5)

Однак, отримана матриця (Рис. 2) у 2-му, 6-му, 7-му та 8-му рядках містить дробові значення, які ускладнюють процедуру перетворення. У зв'язку з цим фактом доцільно перевірити можливість отримання матриці, яка буде містити меншу кількість дробових значень і, відповідно, буде у меншій мірі ускладнювати процедуру перетворення. Одним із способів отримати таку матрицю є заміна добутків трійкових симетричних функцій на результати інших логічних операцій в межах трійкової симетричної логіки.

IV. Системи функцій на основі результатів логічних операцій трійкової симетричної логіки

Для спрощення візуального аналізу результатів і без втрати загальності можна для подальших обчислень модифікувати аналітичний вираз (5) заміною операції множення довільною двомісною логічною операцією трійкової симетричної логіки. Іншими словами, розглядатимуться системи функцій, які породжуються результатами застосування двомісних логічних операцій трійкової симетричної логіки до трійкових симетричних функцій перших двох порядків першого набору системи (4). У результаті отримується аналітичний вираз (6).

$$Ter\,01Op_{num}(\theta) = \overset{2}{\underset{j=1}{*}} Ter\,01_{1,2-j}(\theta)^{A_{num,j}}, \qquad (6)$$

де num=0,1,…,3^2-1 – порядок функції, * – деяка двомісна логічна операція трійкової симетричної логіки, $A_{num,j}$ – елемент рядка num та стовпця j кодової матриці Грея A (Рис. 1)

Необхідно враховувати, що у аналітичному виразі (6) піднесення до степеня реалізує не множення, а певну двомісну логічну операцію трійкової симетричної логіки.

За аналогією з бінарною логікою у трійковій симетричній логіці є 3^{3^n} операцій для n аргументів. Звідси випливає, що у трійковій симетричній логіці є $3^{3^2} = 19683$ двомісні операції. І кожна з них може бути використана у аналітичному виразі (6) для утворення деякої системи функцій на основі трійкових симетричних.

У зв'язку з тим, що переважна більшість операцій трійкової симетричної логіки не має назв (таких, як наприклад, трійкова симетрична диз'юнкція, тощо), доцільно з метою їх розрізнення занумерувати їх за деяким принципом. Зручним способом упорядкування даних операцій є їх упорядкування за таблицями істинності. Так операції із таблицею істинності, яка містить лише значення «*1*» (Табл. 1) доцільно присвоїти номер «1», а операції із таблицею істинності, яка містить лише значення «*–1*» (Табл. 2) – «19683».

ТАБЛИЦЯ I. Таблиця істинності логічної операції з порядковим номером «1»

Значення першого операнду	**Значення другого операнду**	**Результат виконання операції**
1	1	1
1	0	1
1	–1	1
0	1	1
0	0	1
0	–1	1
–1	1	1
–1	0	1
–1	–1	1

ТАБЛИЦЯ II. Таблиця істинності логічної операції з порядковим номером «19683»

Значення першого операнду	**Значення другого операнду**	**Результат виконання операції**
1	1	–1
1	0	–1
1	–1	–1
0	1	–1
0	0	–1

Значення першого операнду	Значення другого операнду	Результат виконання операції
0	–1	–1
–1	1	–1
–1	0	–1
–1	–1	–1

Більшість операцій при підстановці у вираз (6) утворюють у результаті лінійно залежні системи функцій (визначники матриць значень рівні нулю). Серед лінійно незалежних систем, однак, не зустрічаються ортогональні системи, а отже, дані системи теж потребують застосування процедури ортогоналізації. Встановлено, що має місце наступна тенденція: чим більший за модулем визначник матриці значень, тим у ортогоналізованій матриці більшими за модулем є знаменники дробових значень. З цього твердження випливає доцільність аналізу матриць значень з якнайменшими визначниками.

Якщо розглянути матрицю значень системи функцій, одержаної на основі операції з порядковим номером «701» (Рис. 3), то можна дійти висновку, що кількість різних за модулем дробових значень у ній вища, ніж у матриці, отриманій за добутками (Рис. 2) (7 у порівнянні з 5). Тобто за критерієм кількості дробових значень ортогоналізована матриця значень за системою (5) (Рис. 2) є більш придатною для цифрової обробки інформації на основі ортогональних перетворень.

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 2/3 & 2/3 & 2/3 & -1/3 & -1/3 & -1/3 & -1/3 & -1/3 & -1/3 \\ 0 & 0 & 0 & 1/2 & 1/2 & 1/2 & -1/2 & -1/2 & -1/2 \\ -1/3 & -1/3 & 2/3 & 0 & 0 & 0 & -1/3 & -1/3 & 2/3 \\ 1/6 & 1/6 & -1/3 & -1/3 & -1/3 & 2/3 & -1/6 & -1/6 & 1/3 \\ -1/9 & -1/9 & 2/9 & -1/9 & -1/9 & 2/9 & 1/9 & 1/9 & -2/9 \\ 1/2 & -1/2 & 0 & 1/2 & -1/2 & 0 & 1/2 & -1/2 & 0 \\ -1/3 & 1/3 & 0 & 1/6 & -1/6 & 0 & 1/6 & -1/6 & 0 \\ 0 & 0 & 0 & -1/4 & 1/4 & 0 & 1/4 & -1/4 & 0 \end{pmatrix}$$

Рис. 3. Матриця значень ортогоналізованої системи функцій (6) за операцією з порядковим номером «701»

Однак, визначник матриці значень за системою (5) (Рис. 2) рівний 64, у той час, як визначник матриці за системою (6) на основі операції з порядковим номером «701» (Рис. 3) рівний 1. Іншими словами, з точки зору критерію обчислення оберненої матриці друга матриця є більш прийнятним варіантом. Звідси випливає, що відкидати з подальшого аналізу ефективності застосування трійкових симетричних функцій для цифрової обробки інформації на основі ортогональних перетворень матрицю за системою (6) на основі операції з порядковим номером «701» (Рис. 3) не можна, оскільки за частиною критеріїв вона може виявитися більш ефективною, ніж матриця за системою (5) (Рис. 2).

V. Висновки

Трійкові симетричні функції у чистому вигляді (система (1)) не можуть бути використані для цифрової обробки інформації на основі ортогональних перетворень. З метою їх адаптації для даних процесів необхідно будувати на їх основі нові системи функцій такі, як наприклад, системи (5) та (6).

Принципово новим підходом є застосування логічних операцій трійкової симетричної логіки для формування системи функцій на основі трійкових симетричних. У результаті такого застосування отримано ортогоналізовану матрицю значень (Рис. 3), яка може бути використана для цифрової обробки інформації на основі ортогональних перетворень і за рядом критеріїв може перевищити ефективність застосування попередньо отриманої матриці на основі добутків (Рис. 2).

Подальші дослідження полягають у адаптуванні критеріїв ефективності застосування систем функцій для цифрової обробки інформації на основі ортогональних перетворень до роботи зі значеннями трійкової симетричної логіки з метою аналізу ефективності застосування отриманих систем функцій у цільовій області.

Література References

[1] Э. Айфичер, Б. Джервис. Цифровая обработка сигналов: практический подход, 2-е издание: Пер. с англ. – М.: Издательский дом «Вильямс», 2004. – 992 с.

[2] Л.А. Залманзон, Преобразования Фурье, Уолша, Хаара и их применение в управле-нии, связи и других областях. – М.: Наука, 1989. – 496 с.

[3] Н. Ахмед, Ортогональные преобразования при обработке цифровых сигналов / Н. Ахмед, К.Р. Рао; пер. с англ. Т.Э. Кренкеля. – М.: Связь, 1980. – 248 с.

[4] B. Hayes, Computing science. Third base. A reprint from American Scientist, the magazine of Sigma Xi, the Scientific Research Society, vol. 89, Nr. 6. November–December 2001, pp. 490-494

[5] A. Izmailov, Effectiveness analysis of bases and function systems used in digital information processing / A. Izmailov // Збірник тез 52 конференції студентських наукових кіл Гірничо-металургійної академії імені Станіслава Сташіца у Кракові. – Краків, 2015. – С. 281.

[6] А.В. Ізмайлов, Трійкові симетричні функції та їх застосування у цифровій обробці інформації / А. В. Ізмайлов, Л. Б. Петришин // Системи обробки інформації. — 2016. — № 4. — С. 41-44.

[7] A. Izmailov, Symmetric ternary functions and their application in orthogonal transforms / A. Izmailov, L. Petryshyn // IEEE Xplore Digital Library, in press.

[8] L.E. Franks, “Signal Theory”, Prentice Hall, 1969, p. 318.

СЕКЦІЯ 5.

Захист інформації в інформаційно-телекомунікаційних системах

SECTION 5.

Information security in information and telecommunication systems

Математичні Моделі Модифікованої Несиметричної Крипто-Кодової Системи Мак-Еліса на Модифікованих Еліптичних Кодах

Євсеєв Сергій
Кафедра інформаційних систем
Харківський національний економічний університет ім. С. Кузнеця
Харків, Україна
Serhii.Yevseiev@hneu.net

Mathematical Models of the Modified Asymmetric Crypto-Code System of McElice on Modified Elliptic Codes

Serhii Yevseiev
Information Systems Department
Simon Kuznets Kharkiv National University of Economics
Kharkiv, Ukraine
Serhii.Yevseiev@hneu.net

***Анотація*—Математичні моделі модифікованих засобів захисту інформації розроблені на основі теоретико-кодової схеми Мак-Еліса з використанням алгеброгеометричних блокових кодів з укороченням і подовженням інформаційного простору**

***Abstract*—Mathematical models of modified crypto-code information protection tools are developed on the basis of the McElis theoretical-code scheme using algebraic geometric block codes with shortening and extension of the information premise**

Ключові слова— несиметрична крипто-кодова система, теоретико-кодові схеми, модифіковані коди з виправленням помилок

Keywords—asymmetric crypto-code system, theoretical-code system, modified error-correcting codes

I. Введення і аналіз літератури

Розвиток телекомунікаційних систем і технологій, бурхливе зростання обчислювальної техніки висувають нові вимоги до основних критеріїв якості обслуговування клієнтів (уповноважених користувачів). Основними показниками за результатами аналізу стандартів в цій галузі є забезпечення достовірності (надійності) передачі даних і забезпечення безпеки всього циклу обробки і зберігання даних [1, 2, 3]. Для забезпечення достовірності використовуються механізми перешкодостійкого кодування, а для забезпечення безпеки – криптографічні механізми на основі методів симетричної і несиметричної криптографії. Перспективним напрямком, на наш погляд, є використання несиметричних криптосистем на основі теоретико-кодових схем Мак-Еліса, що забезпечують інтегровано (одним механізмом) показники достовірності на рівні $2^9 – 2^{12}$ і криптостійкості – $2^{30} – 2^{35}$ групових операцій при її побудові над полем $GF(2^{10})$. Дана криптосистема отримала широке застосування з розвитком обчислювальних можливостей комунікаційних пристроїв і їх програмного забезпечення.

В роботі [4] автори пропонують використовувати криптосистему Мак-Еліса в програмному забезпеченні Sequitur, яка дозволяє інтегровано вирішувати завдання швидкодії і безпеки при передачі конфіденційної інформації. У роботах [5, 6, 7] криптосистему Мак-Еліса пропонують використовувати для забезпечення основних послуг безпеки: конфіденційності і цілісності в стегасистемі на основі звукових файлів MPEG Layer-III або MP3, для забезпечення доступності та цифрового підпису при передачі конфіденційної медичної інформації. Разом з тим, проведений в роботі [8] аналіз програмної реалізації несиметричної крипто-кодової системи на ТКС Нідеррайтера показав на значні складності реалізації, що істотно ускладнює використання теоретико-кодових схем для побудови криптостійкі несиметричних систем. В роботі [9] розглянуто нові підходи до злому криптосистеми Мак-Еліса на основі

рандомізованих зчеплених кодів.

Для забезпечення необхідних показників криптостійкості і збільшення обсягу переданих даних автором пропонується модифікована несиметрична крипто-кодова система (МНККС) Мак-Еліса на модифікованих (скорочених/подовжених) еліптичних кодах, що є перспективним напрямком у вирішенні даної науково-технічної задачі. Проведений аналіз в роботі [10] швидкості криптоперетворень в модифікованих несиметричних крипто-кодових системах Мак-Еліса порівняння з блоково-симетричними шифрами, але забезпечує математичну доказову криптостійкість на основі теоретико-складної задачі – декодування випадкового коду.

II. Основна частина

В роботі [10] автором розглянуті основні принципи модифікації завадостійких кодів, та запропонований математичний апарат модифікації алгеброгеометричних кодів на еліптичних кривих (модифікованих еліптичних кодів, МЕС).

На основі запропонованих модифікацій МЕС пропонуються математичні моделі модифікованої несиметричної крипто-кодової системи (МНККС) Мак-Еліса на МЕС (подовжених/скорочених кодів).

Математична модель НККС з використанням ТКС Мак-Еліса на основі укорочення (скорочення інформаційних символів) формально задається сукупністю наступних елементів [9]:

– множина відкритих текстів

$$M = \{M_1, M_2, ..., M_{q^k}\},$$

де $M_i = \{I_0, I_{h_1}, .. I_{h_j}, I_{k-1}\}$, $\forall I_j \in GF(q)$ h_j – інформаційні символи рівні нулю, $|h| = \frac{1}{2}k$, т. е. $I_i = 0$, $\forall I_i \in h$;

– множина закритих текстів (кодограм)

$$C = \{C_1, C_2, ..., C_{q^k}\},$$

$$\text{де } C_i = (c^*_{X_0}, c^*_{h_1}, ..., c^*_{h_j}, c^*_{X_{n-1}}),\ \forall c^*_{X_j} \in GF(q);$$

– множина прямих відображень (на основі використання відкритого ключа – породжувальної матриці):

$$\varphi = \{\varphi_1, \varphi_2, ..., \varphi_s\},$$

де $\varphi_i : M \to C_{k-h_j}$, $i = 1,2,...,s$;

– множина зворотних відображень (на основі використання закритого (особистого) ключа – матриць маскування):

$\varphi^{-1} = \{\varphi_1^{-1}, \varphi_2^{-1}, ..., \varphi_s^{-1}\}$, где $\varphi_i^{-1} : C_{k-h_j} \to M$, $i = 1,2,...,s$;

– множина ключів, параметрізірующіх прямі відображення (відкритий ключ уповноваженого користувача):

$$K_{a_i} = \{K_{1_{a_i}}, K_{2_{a_i}}, ..., K_{s_{a_i}}\} = \{G_X^{EC_1}{}_{a_i}, G_X^{EC_2}{}_{a_i}, ..., G_X^{ECs}{}_{a_i}\}$$

де $G_X^{ECi}{}_{a_i}$ – яка породжує $n \times k$ матриця замаскованого під випадковий код алгеброгеометричного блокового (n,k,d) коду з елементами з $GF(q)$ тобто $\varphi_i : M \xrightarrow{K_{1a_i}} C_{k-h_j}$; $i = 1,2,...,s$; a_i – набір коефіцієнтів багаточлена кривої $a_1 ... a_6$, $\forall a_i \in GF(q)$, однозначно ставить конкретний набір точок кривої з простору P^2.

– множина ключів, які параметрізують зворотні відображення (особистий (закритий) ключ уповноваженого користувача):

$$K^* = \{K_1^*, K_2^*, ..., K_s^*\} = \{\{X,P,D\}_1, \{X,P,D\}_2, ..., \{X,PD\}_s\},$$

$$\{X,P,D\}_i = \{X^i, P^i, D^i\},$$

де X^i – маскуюча невироджена випадково рівноймовірно сформована джерелом ключів $k \times k$ матриця з елементами з $GF(q)$; P^i – перестановочна випадково рівноймовірно сформована джерелом ключів $n \times n$ матриця з елементами з $GF(q)$; D^i – діагональна сформована джерелом ключів матриця з елементами з $GF(q)$, тобто

$$\varphi_i^{-1} : C \xrightarrow{\text{K}_\text{i}^*} M,\ i = 1,2,...,s,$$

складність виконання зворотного відображення φ_i^{-1} без знання ключа $K_i^* \in K^*$ пов'язане з рішенням теоретико-складної задачі декодування випадкового коду (коду загального положення).

Вихідними даними при опису розглянутої несиметричної крипто-кодової системи захисту інформації є:

– алгеброгеометричний блоковий (*n, k, d*) код C_{k-h_j} над *GF(q)*, тобто множина кодових слів $C_i \in C_{k-h_j}$ таких, що виконується рівність $C_i H^T = 0$ де H – перевірочна матриця алгеброгеометричного блокового коду;

– a_i – набір коефіцієнтів багаточлена кривої $a_1 ... a_6$, $\forall a_i \in GF(q)$, однозначно ставить конкретний набір точок кривої з простору P^2 для формування матриці, яка породжує;

– h_j – інформаційні символи, рівні нулю, $|h| = 1/2k$, т. е. $I_i = 0$, $\forall I_i \in h$;

– маскуючі матричні відображення, задані безліччю матриць $\{X,P,D\}_i$, де X – невироджена $k \times k$ матриця над *GF(q)*, P – переставна $n \times n$ матриця над *GF(q)* з одним ненульовим елементом в кожному рядку і в кожному стовпці матриці, D – діагональна $n \times n$ матриця над *GF(q)* з ненульовими елементами на головній діагоналі.

У несиметричній крипто-кодовій системі на основі ТКС Мак-Еліса модифікований (укорочений) алгеброгеометричний (*n, k, d*) код C_{k-h_j} з швидким алгоритмом розкодування маскується під випадковий

(n, k, d) код C_{k-h_j}* за допомогою множення породжувальної матриці G^{EC} код C_{k-h_j} на збережені в секреті маскуючі матриці X^u, P^u і D^u, [10], що забезпечує формування відкритого ключа уповноваженого користувача:

$$G_X^{ECu} = X^u \cdot G^{EC} \cdot P^u \cdot D^u, \; u \in \{1,2,...,s\},$$

де G^{EC} – породжувальна $k \times n$ матриця алгеброгеометричного блокового (n,k,d) коду з елементами з *GF(q)*, побудовану на основі використання вибраних користувачем коефіцієнтів багаточлена кривої $a_1 ... a_6$, $\forall a_i \in GF(q)$, однозначно, які задають конкретний набір точок кривої з простору P^2.

Формування закритого тексту $C_j \in C_{k-h_j}$ по введеному відкритому тексту $M_i \in M$ і заданому відкритому ключу $G_X^{ECu}{}_{a_i}$, $u \in \{1,2,...,s\}$ здійснюється шляхом формування кодового слова замаскованого коду з додаванням до нього випадково сформованого вектору $e = (e_0, e_1, ..., e_{n-1})$

$$C_j = \varphi_u\left(M_i, G_X^u\right) = M_i \cdot \left(G_X^u\right)^T + e,$$

причому вага Хемінга (число ненульових елементів) вектору e не перевищує виправляючої здатності використовуваного алгебраїчного блокового коду:

$$0 \le w(e) \le t = \left\lfloor \frac{d-1}{2} \right\rfloor,$$

$\lfloor x \rfloor$ – ціла частина дійсного числа *x*.

Для кожного закритого тексту, який формується $C_j \in C_{k-h_j}$ відповідний вектор $e = (e_0, e_1, ..., e_{n-1})$ виступає в якості одноразового сеансового ключа, тобто для конкретного E_j вектор e формується випадково, рівноймовірно і незалежно від інших закритих текстів.

У канал зв'язку надходить $C_j^* = C_j - C_{k-h_j}$.

На приймальній стороні уповноважений користувач, який знає правило маскування, кількість і місця нульових інформаційних символів може скористатися швидким алгоритмом розкодування алгеброгеометричного коду (поліноміальної складності) для відновлення відкритого тексту [10]:

$$M_i = \varphi_u^{-1}\left(C_j^*, \{X, P, D_u\}\right).$$

Для відновлення відкритого тексту уповноважений користувач додає нульові інформаційні символи $C_j^* = C_j + C_{k-h_j}$ з відновленого закритого тексту C_j знімає дію секретних перестановної і діагональної матриць P^u і D^u:

$$C = C_j^* \cdot \left(D^u\right)^{-1} \cdot \left(P^u\right)^{-1} = \left(M_i \cdot \left(G_X^u\right)^T + e\right) \cdot \left(D^u\right)^{-1} \cdot \left(P^u\right)^{-1} =$$
$$= \left(M_i \cdot \left(X^u \cdot G \cdot P^u \cdot D^u\right)^T + e\right) \cdot \left(D^u\right)^{-1} \cdot \left(P^u\right)^{-1} =$$
$$= M_i \cdot \left(X^u\right)^T \cdot (G)^T \cdot \left(P^u\right)^T \cdot \left(D^u\right)^T \cdot \left(D^u\right)^{-1} \cdot \left(P^u\right)^{-1} + e \cdot \left(D^u\right)^{-1} \cdot \left(P^u\right)^{-1} =$$
$$= M_i \cdot \left(X^u\right)^T \cdot (G)^T + e \cdot \left(D^u\right)^{-1} \cdot \left(P^u\right)^{-1},$$

розкодує отриманий вектор за алгоритмом Берлекемпа-Месси [10]:

$$C = M_i \cdot \left(X^u\right)^T \cdot \left(G^{EC}\right)^T + e \cdot \left(D^u\right)^{-1} \cdot \left(P^u\right)^{-1},$$

тобто позбувається від другого доданку і від співмножника $(G)^{EC^T}$ в першому доданку в правій частині рівності, після чого знімає дію матриці маскування X^u. Для цього отриманий результат розкодування $M_i \cdot \left(X^u\right)^T$ слід помножити на $\left(X^u\right)^{-1}$: $\left(M_i \cdot \left(X^u\right)^T\right) \cdot \left(X^u\right)^{-1} = M_i$. Отримане рішення – суть відкритий текст M_i.

Математична модель модифікованої несиметричної крипто-кодової системи захисту інформації з використанням алгеброгеометричних блокових кодів на основі теоретико-кодової схеми Мак-Еліса на основі подовження (збільшення інформаційних символів) формально задається сукупністю наступних елементів

– множина відкритих текстів

$M = \{M_1, M_2, ..., M_{q^k}\}$, де $M_i = \{I_0, I_{h_1}, ... I_{h_{rj}}, I_{k-1}\}$, $\forall I_j \in GF(q)$, h_j – інформаційні символи рівні нулю, |h| = $\frac{1}{2}k$, тобто $I_i = 0, \forall I_i \in h$; h_r – інформаційні символи подовження k, |h| = $\frac{1}{2}k$;

– множина закритих текстів (кодограм)

$$C = \{C_1, C_2, ..., C_{q^k}\},$$

де $C_i = (c_{X_0}^*, c_{h_{r1}}^*, ..., c_{h_{rj}}^*, c_{X_{n-1}}^*)$, $\forall c_{X_j}^* \in GF(q)$;

– множина прямих відображень (на основі використання відкритого ключа - породжує матриці)

$\varphi = \{\varphi_1, \varphi_2, ..., \varphi_s\}$, де $\varphi_i : M \to C_{h_r}$, $i = 1,2,...,s$;

– множина зворотних відображень (на основі використання закритого (особистого) ключа - матриць маскування)

$\varphi^{-1} = \{\varphi_1^{-1}, \varphi_2^{-1}, ..., \varphi_s^{-1}\}$, де $\varphi_i^{-1} : C_{h_r} \to M$, $i = 1,2,...,s$;

– множина ключів, які параметрізують прямі відображення (відкритий ключ уповноваженого користувача))

$$K_{a_i} = \{K_{1_{a_i}}, K_{2_{a_i}}, ..., K_{s_{a_i}}\} = \{G_X^{EC_1}{}_{a_i}, G_X^{EC_2}{}_{a_i}, ..., G_X^{ECs}{}_{a_i}\},$$

де $G_{X\ a_i}^{ECi}$ – породжувальна $n\times k$ матриця замаскованого під випадковий код алгеброгеометричного блокового (n,k,d) коду з елементами з $GF(q)$, тобто

$$\varphi_i : M \xrightarrow{K_{i a_i}} C_{h_r}; i = 1,2,...,s;$$

a_i – набір коефіцієнтів багаточлена кривої a_1 ... a_6, $\forall a_i \in GF(q)$, однозначно ставить конкретний набір точок кривої з простору P^2.

– множина ключів, які параметрізують зворотні відображення (особистий (закритий) ключ уповноваженого користувача))

$$K^* = \{K_1^*, K_2^*, ..., K_s^*\} = \{\{X,P,D\}_1, \{X,P,D\}_2, ..., \{X,P,D\}_s\},$$

$$\{X,P,D\}_i = \{X^i, P^i, D^i\},$$

де X^i – маскуюча невироджена випадково рівноймовірно сформована джерелом ключів $k\times k$ матриця з елементами з $GF(q)$; P^i – перестановочна випадково рівноймовірно сформована джерелом ключів $n\times n$ матриця з елементами з $GF(q)$; D^i – діагональна сформована джерелом ключів $n\times n$ матриця з елементами з $GF(q)$, тобто

$$\varphi_i^{-1} : C \xrightarrow{\mathrm{K_i^*}} M, i = 1,2,...,s,$$

складність виконання зворотного відображення φ_i^{-1} без знання ключа $K_i^* \in K^*$ пов'язане з рішенням теоретико-складної задачі декодування випадкового коду (коду загального положення).

Вихідними даними при описі розглянутої несиметричної крипто-кодової системи захисту інформації є параметри, описані в попередній моделі.

У несиметричній крипто-кодовій системі на основі ТКС Мак-Еліса модифікований (подовжений) алгеброгеометричний (n, *k, d*) код C_{h_r} з швидким алгоритмом розкодування маскується під випадковий (*n, k, d*) код C_{h_r}* за допомогою множення породжує матрицю G^{EC} коду C_{k-h_j} на маскують матриці, які зберігаються в секреті X^u, P^u і D^u, що забезпечує формування відкритого ключа уповноваженого користувача: $G_X^{ECu} = X^u \cdot G^{EC} \cdot P^u \cdot D^u$, $u \in \{1,2,...,s\}$,

де G^{EC} – яка породжує $n\times k$ матрицю алгеброгеометричного блокового (n,k,d) коду з елементами з $GF(q)$, побудована на основі використання вибраних користувачем коефіцієнтів багаточлена кривої a_1 ... a_6, $\forall a_i \in GF(q)$, однозначно задають конкретний набір точок кривої з простору P^2.

Формування закритого тексту $C_j \in C_{h_r}$ по введеному відкритому тексту $M_i \in M$ і заданому відкритому ключу $G_{X\ a_i}^{ECu}$, $u \in \{1,2,...,s\}$ здійснюється шляхом формування укороченого кодового слова, а потім подовження замаскованого коду з додаванням до нього випадково сформованого вектору $e = (e_0, e_1, ..., e_{n-1})$:

$$C_j = \varphi_u\left(M_i, G_X^u\right) = M_i \cdot \left(G_X^u\right)^T + e$$

Для кожного формованого закритого тексту $C_j \in C_{h_r}$ відповідний вектор $e = (e_0, e_1, ..., e_{n-1})$ виступає в якості одноразового сеансового ключа, тобто формується випадково, рівноймовірно і незалежно від інших закритих текстів.

У канал зв'язку надходить $C_j^* = C_j - C_{k-h_j} + C_{h_r}$

На приймальній стороні, уповноважений користувач, який знає правило маскування, кількість і місця нульових інформаційних символів може скористатися швидким алгоритмом розкодування алгеброгеометричного коду (поліноміальної складності) для відновлення відкритого тексту: $M_i = \varphi_u^{-1}\left(C_j^*, \{\mathrm{X,P,D}\}_\mathrm{u}\right)$.

Для відновлення відкритого тексту уповноважений користувач заміняє символи подовження на нульові інформаційні символи: $C_j^* = C_{h_r} \to C_{k-h_j}$,

з відновленого закритого тексту C_j знімає дію секретних перестановної і діагональної матриць P^u и D^u:

$$C = C_j^* \cdot \left(D^u\right)^{-1} \cdot \left(P^u\right)^{-1} = \left(M_i \cdot \left(G_X^u\right)^T + e\right) \cdot \left(D^u\right)^{-1} \cdot \left(P^u\right)^{-1} =$$

$$= \left(M_i \cdot \left(X^u \cdot G \cdot P^u \cdot D^u\right)^T + e\right) \cdot \left(D^u\right)^{-1} \cdot \left(P^u\right)^{-1} =$$

$$= M_i \cdot \left(X^u\right)^T \cdot (G)^T \cdot \left(P^u\right)^T \cdot \left(D^u\right)^T \cdot \left(D^u\right)^{-1} \cdot \left(P^u\right)^{-1} + e \cdot \left(D^u\right)^{-1} \cdot \left(P^u\right)^{-1} =$$

$$= M_i \cdot \left(X^u\right)^T \cdot (G)^T + e \cdot \left(D^u\right)^{-1} \cdot \left(P^u\right)^{-1},$$

розкодує отриманий вектор за алгоритмом Берлекемпа-Месси [10]:

$$C = M_i \cdot \left(X^u\right)^T \cdot \left(G^{EC}\right)^T + e \cdot \left(D^u\right)^{-1} \cdot \left(P^u\right)^{-1},$$

тобто позбавляється від другого доданку і від співмножника $(G)^{EC^T}$ в першому доданку в правій частині рівності, після чого знімає дію матриці маскування X^u. Для цього отриманий результат розкодування $M^*{}_i$ слід помножити на $\left(X^u\right)^{-1}$:

$$M^*{}_i \cdot \left(X^u\right)^{-1} = M_i.$$

Отримане рішення – відкритий текст M_i, до якого додаються символи подовження: $M_j = M_i + h_r$ – суть переданого повідомлення.

В роботах [8, 10] проведений аналіз оцінки енергетичних витрат на програмну реалізацію і складності кодоперетворень в МНККС Мак-Еліса. Для оцінки часових і швидкісних показників прийнято використовувати одиницю виміру *cpb* (*cycles per byte*) – кількість тактів

процесора, який необхідно витратити для обробки 1 байту інформації. Складність алгоритму обчислимо за виразом:

$$Per = Utl * CPU_clock / Rate ,$$

де Utl – утилізація ядра процесора (%), $Rate$ – пропускна здатність алгоритму (байт/с).

В табл. 1 наведені результати досліджень залежності довжини кодової послідовності алгеброгеометричного коду в МНККС Мак-Еліса від кількості тактів процесора на виконання елементарних операцій в програмній реалізації модифікованих крипто-кодових систем. В табл. 2 результати досліджень оцінки тимчасових і швидкісних показників процедур формування і розкодування інформації в несиметричних крипто-кодових системах на основі ТКС Мак-Еліса.

ТАБЛИЦЯ 1 РЕЗУЛЬТАТИ ДОСЛІДЖЕНЬ ЗАЛЕЖНОСТІ ДОВЖИНИ КОДОВОЇ ПОСЛІДОВНОСТІ КОДУ В НККС МАК-ЕЛІСА І МОДИФІКОВАНОЇ НККС ВІД КІЛЬКОСТІ ТАКТІВ ПРОЦЕСОРА

Довжина кодової послідовності		**MacElis на укорочених кодах**			**MacElis на подовжених кодах**			**MacElis**		
		10	***100***	***1000***	***10***	***100***	***1000***	***10***	***100***	***1000***
Кількість викликів функцій, що реалізовують елементарні операції	Читання символу	10 294 397	28 750 457	76 759 874	11 432 131	33 460 317	82 473 442	11 018 042	30 800 328	80 859 933
	Порівняння рядків	3 406 921	9 246 748	25 478 498	3 673 756	12 119 867	29 469 389	3 663 356	10 199 898	26 364 634
	Конкатенація рядків	1 705 544	5 045 748	12 379 422	1 947 681	6 114 478	14 456 729	1 834 983	5 125 564	13 415 329
Сума		15 406 862	43 042 953	114 617 794	17 053 568	17 053 568	51 694 662	126 399 560	46 125 790	120 639 896
Тривалість виконання функцій * в тактах процесора	Читання символу	295 374	810 478	2 001 167	300 479	843 705	2 745 148	297 487	831 609	2 183 218
	Порівняння рядків	178 814	531 379	1 248 684	213 478	561 754	1 739 170	197 821	550 794	1 423 690
	Конкатенація рядків	544 990	1 328 114	3 586 486	578 174	1 647 638	4 007 883	544 990	1 522 293	3 984 353
Сума		1 006 781	2 749 548	7 247 488	109 157	109 157	1 092 131	3 053 097	2 904 696	7 591 261
Тривалість виконання ** в мс		0,52	1,37	3,4	0,56	0,56	1,55	4,1	1,53	4

ПРИМІТКА: * тривалість 1000 операцій в тактах процесора: читання символу – 27 тактів, порівняння рядків – 54 такту, конкатенація рядків – 297 тактів; ** для розрахунку взято процесор з тактовою частотою 2 ггц з урахуванням завантаження операційної системою 5 %

ТАБЛИЦЯ 2 РЕЗУЛЬТАТИ ДОСЛІДЖЕНЬ ОЦІНКИ ТИМЧАСОВИХ І ШВИДКІСНИХ ПОКАЗНИКІВ ПРОЦЕДУР ФОРМУВАННЯ І РОЗКОДУВАННЯ ІНФОРМАЦІЇ

Показники	***Довжина кодової послідовності***	***Пропускна здатність алгоритму, Rate, (байт / с)***	***Утилізація ядра процесора,%***	***Складність алгоритму, Per (cpb)***
MacElis	100	46 125 790	56	61,5
	1000	120 639 896	56	62,0
MacElis на укорочених кодах	100	52 721 778	56	61,5
	1000	127 389 928	56	62,1
MacElis на подовжених кодах	100	51 694 662	56	61,7
	1000	126 399 560	56	62,2

Проведений аналіз табл. 1, 2 підтверджує необхідність використання модифікованих НККС для зменшення енергетичних витрат.

Таким чином, передача ключовий послідовності при використанні модифікованої несиметричної крипто-кодової системи Мак-Еліса на основі укорочених кодів дозволяє використовувати відкриті канали зв'язку комунікаційних систем і суттєво знизити обсяги ключових послідовностей, що зберігаються у користувачів даної системи. Використання подовжених еліптичних кодів дозволяє збільшити обсяг переданих даних на h_r символів, забезпечуючи при цьому стійкість криптосистеми при її формуванні в поле $GF\ (2^6 – 2^8)$, що істотно знижує енергетичні витрати на її реалізацію.

ЛІТЕРАТУРА REFERENCES

[1] S.G. Semenov, Modeli i metody upravlenija setevymi resursami v informacionno-telekommunikacionnyh sistemah [Tekst] : monografija / S. G. Semenov, A. A. Smirnov, E. V. Meleshko – Har'kov : NTU "HPI", 2011. – 212 s.

[2] H.N. Rzaev, Analiz sostojanija i putej sovershenstvovanija protokolov bezopasnosti sovremennyh telekommunikacionnyh setej [Tekst] : monografija / H. N. Rzaev, O. G. Korol' // Informacionnye tehnologii v upravlenii, obrazovanii, nauke i promyshlennosti: monografija/– H. : Izdatel' Rozhko S. G. 2016. – S.217 – 234.

[3] Telekommunikacionnye uslugi v mirovoj jekonomike [Jelektronnyj resurs] : – Rezhim dostupa : http://www.gumer.info/bibliotek_Buks/Econom/world_econom/30.php.

[4] Transmission of Picturesque content with Code Base Cryptosystem [Elektronnyj resurs] : – Rezhim dostupa : https://doaj.org/article/6714b60516cc4aa79e56d0c421febaf3.

[5] Steganography application program using the ID3v2 in the MP3 audio file on mobile phone [Jelektronnyj resurs] : – Rezhim dostupa : https://doaj.org/article/707a6506be9e49698fd75323fcc1302c.

[6] Space-Age Approach To Transmit Medical Image With Codebase Cryptosystem Over Noisy Channel [Jelektronnyj resurs] : – Rezhim dostupa : https://doaj.org/article/5c7da3a1e3ec4f83b552199034bd3241.

[7] An Authenticated Transmission of Medical Image with Codebase Cryptosystem over Noisy Channel [Jelektronnyj resurs] : – Rezhim dostupa : https://doaj.org/article/39a3ac65d5b24b348f069dfc82eb6248.

[8] S.P. Evseev, Analiz programmnoj realizacii prjamogo i obratnogo preobrazovanija po metodu nedvoichnogo ravnovesnogo kodirovanija / S.P. Evseev, H.N. Rzaev, A.S. Cyganenko // Naukovo-tehnichnij zhurnal «Bezpeka informacii». tom.22. № 2. Kiïv. – 2016. – s. 196 – 203

[9] On the Usage of Chained Codes in Cryptography [Jelektronnyj resurs] : – Rezhim dostupa : https://doaj.org/article/c0f40bdb1f6149f4ac107d44a95c9531.

[10] S.P. Evseev, Razrabotka modificirovannoj nesimmetrichnoj kripto-kodovoj sistemy Mak-Jelisa na ukorochennyh jellipticheskih kodah / S.P. Evseev, H.N. Rzaev, O.G. Korol, Z.R. Imanova // Vostochno-evropejskij zhurnal peredovyh tehnologij. – Har'kov. – 2016. – tom 4. 9(82). – S. 18-26.

Development of Data Exchange Protocol Prototype Between Intrusion Detection and Prevention Systems

Oleksandra Yeremenko, Oleksandr Lemeshko, Anatoliy Persikov
Department of Infocommunication Engineering
Kharkiv National University of Radio Electronics
Kharkiv, Ukraine
oleksandra.yeremenko.ua@ieee.org, oleksandr.lemeshko@nure.ua, persikovanatoliy@gmail.com

Розробка Прототипа Протоколу Обміну Даними між Системами Виявлення та Протидії Атакам

Олександра Єременко, Олександр Лемешко, Анатолій Персіков
Кафедра інфокомунікаційної інженерії
Харківський національний університет радіоелектроніки
м. Харків, Україна
oleksandra.yeremenko.ua@ieee.org, oleksandr.lemeshko@nure.ua, persikovanatoliy@gmail.com

***Abstract*—The data exchange protocol prototype between Intrusion Detection and Prevention Systems is proposed. One of the candidates for the universal protocol for managing and exchanging of information between IDPS systems is IDXP, however, it has a number of limitations. As a result there arises the task of developing a protocol that solves the issues of detecting various types of IDPSs in the network, exchanging attack signatures between IDPSs, coordinating actions to manage the state of the network and efficient information exchange in a potentially insecure network with an irregular structure. The protocol should take into account the differences in IDPS workflows and how to capture information about events in the network, the possibilities of analyzing the network state, and the functional capabilities of devices acting as data accumulating elements.**

***Анотація*—Запропоновано прототип протоколу обміну даними між системами виявлення та протидії атакам. Одним з кандидатів на роль універсального протоколу для управління та обміну інформацією між IDPS системами є IDXP, проте він має ряд обмежень. Отже виникає задача розробки протоколу, який вирішує завдання проведення виявлення різних видів IDPS в мережі, обміну сигнатурами атак між IDPS, узгодження дій з управління станом мережі та ефективного обміну інформацією в потенційно незахищеній мережі з нерегулярною структурою. Протокол повинен враховувати відмінності в схемах роботи IDPS і способах фіксації інформації про події в мережі, можливості проведення аналізу стану мережі, а також функціональних здібностях пристроїв, які виступають в ролі елементів, що акумулюють дані.**

Keywords—prototype; protocol; data exchange; Intrusion Detection and Prevention System; IDXP

Ключові слова—прототип; протокол; обмін даними; система виявлення та протидії атакам; IDXP

I. INTRODUCTION

Public networks, such as the Internet, and networks integrated with it are open infocommunications, services of which are used by millions of users. Every day hundreds of thousands attempts to violate network performance in general are made. These attempts are mostly initiated by professional attackers, who possess special software tools and hardware with enormous computing power [1, 2].

That is why a relevant problem of such a type of networks is to provide information security (IS), which means realization of different procedures and tasks: authentication, identification, authorization, audit, confidentiality and integrity of information, etc. [2, 3]. In general, attack prevention in a single segment of a large or small network has no sense because networks based on packet-switching technology and IP protocol are fault-tolerant due to the possibility of selecting a set of information transmission routes [4-8]. That is why the attacker is able to avoid secure segments and spread his actions at vulnerable ones [1-3].

Implementation of passive protective means able to perform auditing of events and data filtering does not give the necessary effect, because such protective means cannot provide quantitative evaluation of the network state [3]. In modern

security systems the emphasis should be done on using active means of security able to accumulate data on different events in the network, to conduct multi-criteria analysis of these data and influence the state of network elements with the help of common or specific interfaces [3] (i.e. to perform active auditing).

The major means for providing active scalable protection is introduction of distributed Intrusion Detection and Prevention Systems (IDPS) [9]. An IDPS should support real time operation to perform the following activities:

- realization of protection mechanisms corresponding to the network security policy;
- detecting an intrusion and forecasting of attacker`s intensions and actions;
- evaluating potential vulnerabilities, data collection and analysis of the current state of the network and security system;
- implementing contractions, including suppression of the attacker`s actions and redistribution of the load between the critical protection mechanisms;
- reducing consequences of intrusion and identifying vulnerabilities, adapting the IS system to better counteract the already studied attacks in the future.

The drawback of modern IDPSs is the lack of effective exchange protocols that allow data exchange between IDPSs of different manufacturers [9] oriented to different formats of data storage and exchange. The protocols offered by the Internet community [10-12] do not meet all the requirements of modern security systems [1, 2]. In addition, these protocols only consider the task of universal data formatting, but they do not cover the issues of dynamic detection of IDPS elements in networks and data routing in a potentially vulnerable network.

Therefore, a relevant problem and task is the development of a prototype of the universal protocol (protocol stack) for IDPS detection and data exchange between different types of IDPS. The protocol should be able to deliver data with the time and quality characteristics given by each of the types of networks, taking into account the potential danger of data transmission over a certain link or a subnet.

II. Basic Requirements to Modern Secure Data Exchange Protocols between IDPS

Currently, there are published several exchange formats and protocols (for example, the Intrusion Detection Exchange Protocol (IDXP)) and used for the exchange of information between different IDPSs [10, 11]. Nevertheless, the IDXP protocol [12] is oriented only to the problem of universal data formatting within Intrusion Detection Message Exchange Format (IDMEF) [13]. Analysis of the capabilities of IDXP/IDMEF has shown that:

- the protocol does not solve the problems of the phase of negotiations on the selected information exchange technologies (including protocols and cryptographic algorithms);
- the protocol is applied and oriented to point-to-point connection, which involves the use of a service transport protocol that can be compromised before a secure connection is established;
- the protocol declares the possibility of multipath data exchange, but does not describe it;
- there is no implementation of the connection pooling, so there is a need to re-generate the secure connection with the subsequent transmission of messages.

However, despite these shortcomings, IDXP/IDMEF ideas can be used in the protocol prototype to solve the problems of universal data formatting.

The protocol prototype should be of the Network Layer (an approach similar to that defined in IPv6 [14] can be selected) to perform data routing and accelerated device discovery by sending frame signals about the presence of an IDPS object of a certain type [15]. Providing mutual authentication of network elements in this case can be realized by using the Extensible Authentication Protocol (EAP) [16].

To deliver messages between hosts where IDPS works, a “routed protocol – transport protocol” binding can be used. As a routed protocol, IPv4 and IPv6 can be chosen [2] (in the case of cryptographic security IPsec [17]), and as the transport protocol we can choose the BEEP-encapsulated [18] TCP protocol [19]. In principle, considering IDPS as a routing device that can monitor the state of routers in the network, one can create an overlay communication network related to the tasks of detecting and preventing attacks.

Delivery of messages describing various types of vulnerabilities and ways to counter them should be carried out using a universal data format, which is a text format. The markup of text data can be performed using a certain markup language based on the Standard Generalized Markup Language (SGML) specification [20]. Since it is assumed that the protocol prototype is compatible with IDXP, Extensible Markup Language (XML) will be selected as the markup language. For the encoding of information, the Multipurpose Internet Mail Extensions (MIME) standard [21] can be used, and in the case of cryptographic security – Secure/Multipurpose Internet Mail Extensions (S/MIME) [22]. In the case of S/MIME, encryption (optionally with compression) of transmitted messages is implemented, and therefore, Network Layer security services can be limited by mutual authentication of nodes and matching of session keys. In fact, there is no need to create a full cryptographic tunnel between hosts where IDPSs operate, which increases the system operation speed in general [23, 24]. This aspect is important when scaling the integration of IDPS systems and under simultaneous operation of tens of thousands of devices. The data request/data management can be carried out on the basis of the HTTP protocol [25], which will allow signaling in heterogeneous networks, where firewalling is used [2, 3] (since the transmission of HTTP messages is usually not blocked).

The second task (after ensuring compatibility of systems by unifying the format of data representation), which is solved when choosing a data markup language, is the task of ensuring that the network state uncertainty does not increase as new

IDPS data are received from other IDPSs. The level of uncertainty increases in the following cases:

receiving incorrect data imposed by an attacker;

receiving corrupted data that will be incorrectly interpreted;

late reception of data due to delays in the network (including those caused by the impact of attacks), untimely data transmissions, as well as the need to convert data;

obtaining correct data, however, they are presented in a different format (resulting in an increase in the total number of system states).

To resist the imposing of data by an attacker is possible through the use of integrity primitives [23, 24] and special MIME markup (S/MIME). Tagging markup of the XML protocol allows to detect syntax mismatches and, therefore, to identify incorrect message fragments. This will allow to avoid sending incorrect data to the analysis system (including those presented in a different format). Delayed data delivery can be avoided by using the Traffic Engineering [6, 26-28] orientation and resource reservation along the data path [4]. In the case if traffic is not delivered over a certain set of paths during the predetermined number of times, these paths must be excluded from the network topology.

III. Stages of Exchange Protocol Operation

The protocol prototype consists of three stages, each of which in turn consists of a certain number of phases (Fig. 1). The division into phases and stages is done in such a way that the actions within the stage can be performed cyclically, and the transition from one stage to the next one is possible only if all the phases of the stage are successfully completed.

Most of the phases are independent elements of the protocol and can be executed regardless of other phases. This enables parallel execution of actions within the protocol with a constant change in the presentation of the network within the integrated IDPS system. The transfer of information and network management can then be performed in an iterative manner, i.e. management will be improved as more information becomes available.

The emphasis in the exchange protocol within the IDPS system should be made on ensuring that the representation of the network state cannot be deteriorated and that only the relevant information, which is not imposed by the attacker, is received.

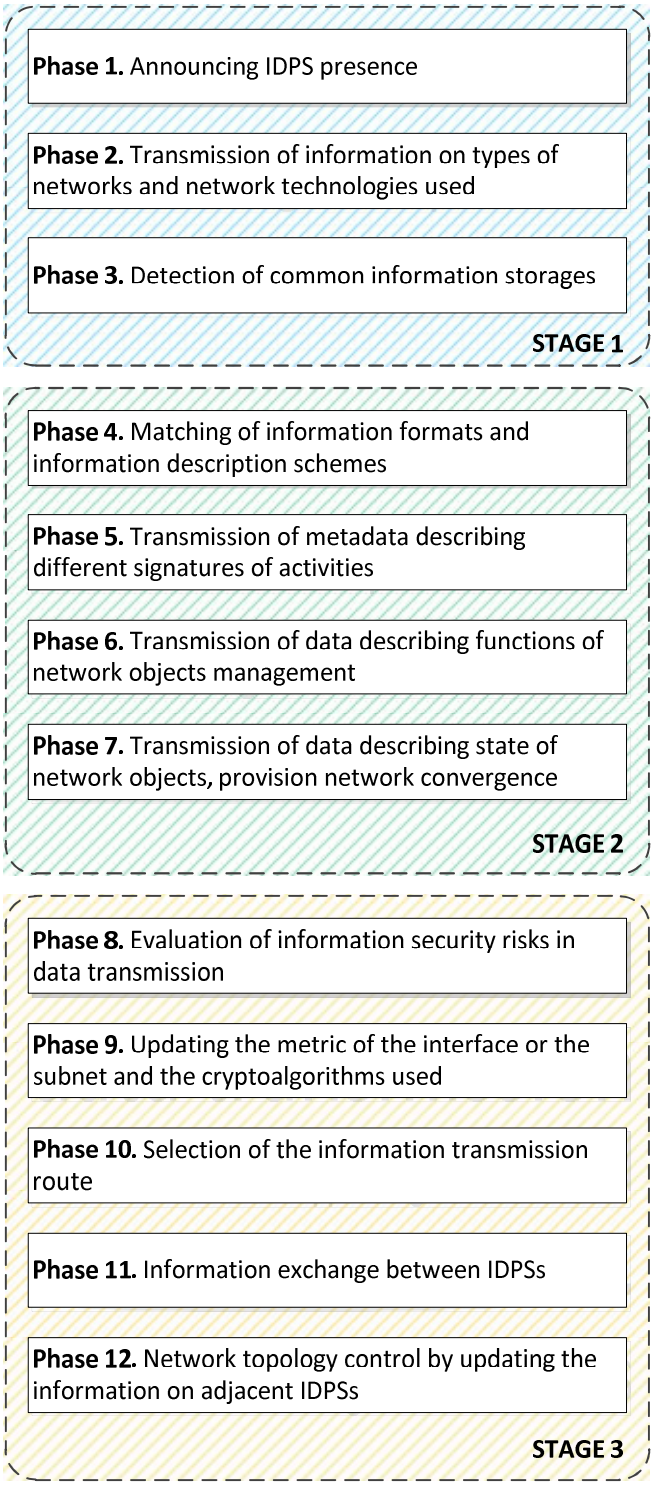


Fig. 1. Tasks supported by the data exchange protocol.

IV. Conclusion

One of the candidates for the universal protocol for managing and exchanging of information between IDPS systems is IDXP, however, it has a number of limitations. That is why there arises the task of developing a protocol (protocol stack) that solves the issues of detecting various types of IDPSs in the network, exchanging attack signatures between IDPSs,

coordinating actions to manage the state of the network and efficient information exchange in a potentially insecure network with an irregular structure. The protocol should take into account the differences in IDPS workflows and how to capture information about events in the network, the possibilities of analyzing the network state, and the functional capabilities of devices acting as data accumulating elements.

When calculating the data route between IDPSs in the network, it is necessary to use a modified metric that takes into account the probability of compromising the information transmission route and the information itself if it is transmitted in the encrypted form. In further development the modified metric is proposed to use, which corresponds to the rule of the Enhanced Interior Gateway Routing Protocol (EIGRP) metric:

as the computing power of supercomputers increases, the metric increases, i.e. the potential security of the information route decreases;

when the probability of compromising the path is increased due to cryptanalysis, the value of the metric increases, i.e. the path will be chosen, where the probability of compromise is minimal;

when the probability of compromising the path is increased due to the vulnerabilities of the communication network, the value of the metric is increased, i.e. the path will be chosen where the number of potential vulnerabilities is minimal.

It should be noted that in this paper we propose only a prototype of the protocol and examine the main provisions that need to be met when developing a real protocol. Therefore, further work in this direction should expand and generalize the provisions of the prototype protocol.

REFERENCES

[1] B. Schneier, "Data and Goliath: The hidden battles to collect your data and control your world," WW Norton & Company, 2015, 398 p.

[2] W. Stallings, "Cryptography and Network Security: Principles and Practice," 7th Edition, Pearson, 2016, 768 p.

[3] V.V. Popovskiy, and A.V. Persikov, "Zaschita informatsii v telekommunikatsionnyih sistemah," OOO "Kompaniya SMIT", 2006, 238 p.

[4] M. Barreiros, and P. Lundqvist, "QOS-Enabled Networks: Tools and Foundations," Wiley Series on Communications Networking & Distributed Systems, 2nd Edition, Wiley, 2016, 254 p.

[5] I. Chen, A. Lindem, and R. Atkinson, "OSPFv3 over IPv4 for IPv6 Transition RFC 7949," IETF, Aug. 2016.

[6] O.V. Lemeshko, O.S. Yeremenko, N. Tariki, and A.M. Hailan, "Fault-Tolerance Improvement for Core and Edge of IP Network," XIth International Scientific and Technical Conference Computer Science and Information Technology (CSIT`2016). Conference Proceedings, Lviv, Ukraine, pp. 161-164, September 6-10, 2016.

[7] T. Szigeti, C. Hattingh, R. Barton, and K. Briley, "End-to-End QoS Network Design: Quality of Service for Rich-Media & Cloud Networks," 2nd Edition, Cisco Press, 2013, 1040 p.

[8] Cisco Networking Academy, ed. Routing Protocols Companion Guide, 1st Edition, Cisco Press, 2014, 792 p.

[9] K. Scarfone, P. Mell, "Guide to intrusion detection and prevention systems (IDPS)," NIST special publication, Feb. 2007, 800(2007):94.

[10] J. Steinberger, A. Sperotto, M. Golling, H. Baier, "How to Exchange Security Events? Overview and Evaluation of Formats and Protocols," 2015 IFIP/IEEE InternationalSymposium on Integrated Network Management (IM 2015), May 2015.

[11] R. Danyliw, J. Meijer, and Y. Demchenko, "The Incident Object Description Exchange Format RFC 5070 (Proposed Standard)," IETF, Dec. 2007.

[12] B. Feinstein, and G. Matthews, "The Intrusion Detection Exchange Protocol (IDXP)," RFC 4767 (Experimental), IETF, Mar. 2007.

[13] H. Debar, D. Curry, and B. Feinstein, "The Intrusion Detection Message Exchange Format (IDMEF) RFC 4765 (Experimental)," IETF, Mar. 2007.

[14] J. Bound, Y. Pouffary, S. Klynsma, T. Chown, and D. Green, "Ipv6 Enterprise Network Analysis-IP Layer 3 Focus RFC 4852", April 2007.

[15] T. Narten, W.A. Simpson, E. Nordmark, and H. Soliman, "Neighbor discovery for IP version 6 (IPv6) RFC 4861," IETF, Sept. 2007.

[16] B. Aboba, L. Blunk, J. Vollbrecht, J. Carlson, and H. Levkowetz, "Extensible authentication protocol (EAP) RFC 3748," IETF, June 2004.

[17] S. Frankel, and S. Krishnan, "IP Security (IPsec) and Internet Key Exchange (IKE) Document Roadmap RFC 6071," IETF, Feb. 2011.

[18] M. Rose, "The Blocks Extensible Exchange Protocol Core," RFC 3080 (Proposed Standard), IETF, Mar. 2001.

[19] M. Rose, "Mapping the BEEP core onto TCP RFC 3081," IETF, Mar. 2001.

[20] Standard, S.G.M.L., Information processing – Text and Office Systems – Standard Generalized Markup Language (SGML). First edition, ISO 8879:1986(E), [Geneva]: International Organization for Standardization, 1986-10-15.

[21] N. Freed and N. Borenstein, "Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies," RFC 2045 (Draft Standard), IETF, Nov. 1996.

[22] B. Ramsdell, and S. Turner, "Secure/Multipurpose Internet Mail Extensions (S/MIME) Version 3.2 Message Specification RFC 5751," IETF, Jan. 2010.

[23] V.V. Popovskiy, and A.V. Persikov, "Osnovy kriptograficheskoy zaschityi informatsii v telekommunikatsionnyih sistemah," Ch. 1, Kharkiv, Kompaniya SMIT, 2010, 350 p.

[24] V.V. Popovskiy, and A.V. Persikov, "Osnovy kriptograficheskoy zaschityi informatsii v telekommunikatsionnyih sistemah," Ch. 2, Kharkiv, Kompaniya SMIT, 2010, 294 p.

[25] R. Fielding, J. Gettys, J. Mogul, H. Frystyk, L. Masinter, P. Leach, and T. Berners-Lee, "Hypertext transfer protocol--HTTP/1.1 RFC 2616," IETF, June 1999.

[26] D. Awduche, A. Chiu, A. Elwalid, I. Widjaja, and X. Xiao, "Overview and Principles of Internet Traffic Engineering RFC 3272, IETF, May 2002.

[27] Y. Lee, Y. Seok, Y. Choi, and C. Kim, "A Constrained Multipath Traffic Engineering Scheme for MPLS Networks," Proc. IEEE ICC'2002, Publisher: IEEE, New York, pp. 2431-2436, May 2002.

[28] Al-shawi, M., and Laurent, A, "Designing for Cisco Network Service Architectures (ARCH): Foundation Learning Guide," 4th Edition, Cisco press, 2017, 944 p.

Огляд Видів Аналізу Соціальних Мереж для Забезпечення Інформаційної Безпеки

Тамара Радівілова
кафедра інфокомунікаційної інженерії
Харківський національний університет радіоелектроніки
Харків, Україна
tamara.radivilova@gmail.com

Людмила Кіріченко
кафедра прикладної математики
Харківський національний університет радіоелектроніки
Харків, Україна
lyudmyla.kirichenko@nure.ua

Даніїл Рудченко
кафедра інфокомунікаційної інженерії
Харківський національний університет радіоелектроніки
Харків, Україна
daniil.rudchenko@nure.ua

Overview Types of Analysis of Social Networks for Information Security

Tamara Radivilova
Department of Infocommunication Engineering
Kharkiv National University of Radioelectronics
Kharkiv, Ukraine
tamara.radivilova@gmail.com

Lyudmyla Kirichenko
Department of Applied Mathematics
Kharkiv National University of Radioelectronics
Kharkiv, Ukraine
lyudmyla.kirichenko@nure.ua

Daniil Rudchenko
Department of Infocommunication Engineering
Kharkiv National University of Radioelectronics
Kharkiv, Ukraine
daniil.rudchenko@nure.ua

***Анотація*—У даній статті представлений огляд основних видів аналізу даних в соціальних мережах для забезпечення інформаційної безпеки. Розглянуто основні математичні методи аналізу даних, які використовуються в кожному з вказаних видів аналізу. Для кожного з видів аналізу вказано дані, які можуть бути проаналізовані, дані, які будуть отримані після аналізу, та їх область та доцільність використання. Представлено типові завдання аналізу соціальних мереж, такі як кластерізація, прогнозування, глибинний аналіз текстів, моніторинг інформаційного простору, визначення взаємозв'язків користувачів та спільнот, класифікація даних, виявлення спільнот та експертів в мережі, виявлення лідерів в спільнотах, аналіз потоків подій, аналіз та контент аналіз та інші.**

***Abstract*—The paper provides an overview of the main types of data analysis in social networks for information security. The basic mathematical methods for analyzing the data used in each of these types of analysis are considered. The main types of social network analysis are hypothesis testing, data operations, sorting, comparison of elements, analysis of photographic and video images, deep analysis of texts, monitoring of the information space, evaluation of user involvement, defining relationships, evaluation of coverage, demographic and geographical classification, analysis tonality, the evaluation of transactions on**

Internet, event stream analysis, prediction. For each type of analysis, the data that can be analyzed, as well as the methods that are most often used to analyze this data, are indicated. Also for each type of analysis, the data to be obtained as a result of the analysis, and their scope and usefulness are indicated. It is shown that main important methods of analysis are graph theory methods and comtlex networks methods, text mining and data mining methods. Typical tasks of analyzing social networks, such as clustering, prediction, determining users relationships and communities, classifying data, identifying communities and experts on the network, identifying leaders in communities, event stream analysis, statistical analysis and content analysis and others are presented.

Ключові слова—аналіз соціальних мереж, методи аналізу даних, передбачення, моніторинг інформаційного простору.

Keywords—social network analysis, data analysis methods, prediction methods, prediction, monitoring of information space.

I. Вступ

У сучасному суспільстві Інтернет і соціальні мережі відіграють дуже важливу роль. Вони дозволяють людям всієї планети спілкуватися з друзями, шукати роботу, займатися просуванням власного бізнесу або організацією різних заходів. Але соціальні медіа - Facebook, Twitter або відеохостінговий сайт YouTube - все в більшій мірі стають інструментами впливу звичайних людей, їх користувачів, на міжнародну політику, соціальні процеси.

Виділяють наступні основні типи кінцевих користувачів, зацікавлених в аналізі, прогнозуванні та управлінні соціальними мережами: 1) органи державної влади та місцевого самоврядування; 2) підприємства державного та приватного сектора економіки, в тому числі комерційні організації; дослідні організації; засоби масової інформації; силові структури; 3) суспільство, в тому числі політичні партії, науково-освітні організації; окремі фізичні особи. Аналіз соціальних мереж, зокрема, використовується в розвідувальних, контррозвідувальних та правоохоронних заходах. Таким чином аналіз інформації та поведінки людей в комп'ютерних соціальних мережах викликає великий інтерес у сучасних дослідників та інших зацікавлених організацій та осіб [1-3].

II. Аналіз інформації соціальних мереж

Тема аналізу інформації дуже велика, і в даній роботі представлена частина видів можливого аналізу інформації соціальних мереж. Мета аналізу - виявлення патернів подій: взаємозв'язків між людьми, закономірностей, за якими відбуваються ці події та передбачення ходу аналогічних подій в подальшому. В ході аналізу соціальних мереж використовують безліч математичних методів. При цьому можуть використовуватися методи багатовимірного статистичного аналізу і штучного інтелекту (в тому числі методи data mining, text mining, image / video mining), а також методи аналізу мережевих структур.

Нижче наведені найпростіші види аналізу даних.

1. Тестування гіпотез. Зіставляючи факти, непрямі ознаки, логічні висновки з фактів можна визначити найбільш ймовірну гіпотезу. Побудова гіпотез (як і будь-якої версії) - це розумовий процес переходу від неповних ймовірних знань до повних і достовірних знань. Висування гіпотези включає всебічне вивчення явищ, що спостерігаються, аналіз і відбір фактів, які перебувають в причинно-наслідковом часовом зв'язку з зазначеними обставинами, аналіз окремих фактів і відносин між ними. Висування гіпотези (версії) складається з наступних кроків: аналіз окремих фактів і відносин між ними; синтез фактів, їх узагальнення; формулювання припущення. Перевірка гіпотези (версії) - цілеспрямований збір доказів, які підтверджують або спростовують ці припущення. Існує кілька способів підтвердження істинності гіпотези (версії):

а) виведення гіпотез, що випливають з наслідків та їх перевірка - встановлення відповідності фактичних даних;

б) безпосереднє виявлення об'єкта, думка про існування якого була основним змістом гіпотези;

в) дедуктивне виведення гіпотези з іншого, але достовірного, знання;

г) підтвердження підстави гіпотези, якщо при побудові гіпотези вона не була достовірною;

д) розширення підстави гіпотези до меж, достатніх для достовірного знання, і т. п.

Спростування гіпотези здійснюється шляхом встановлення фальсифікації наслідків, що випливають з неї, тобто шляхом встановлення їх невідповідності фактичним даним.

2. Операції з даними: сортування, зіставлення елементів і т.п. Можна визначити спільноти, структури, до яких може мати відношення об'єкт, що цікавить. Часто використовують методи кластеризації, основними з яких є наступні [4]:

а) Алгоритми, засновані на поділі даних (Partitioning algorithms), в т.ч. ітеративні: поділ об'єктів на k кластерів; ітеративний перерозподіл об'єктів для поліпшення кластеризації.
б) Ієрархічні алгоритми (Hierarchy algorithms): агломерація: кожен об'єкт спочатку є кластером, кластери, з'єднуючись один з одним, формують більший кластер і т.д.
в) Методи, засновані на концентрації об'єктів (Density-based methods): засновані на можливості з'єднання об'єктів, ігнорують шуми, знаходження кластерів довільної форми.
г) Грід-методи (Grid-based methods): квантування об'єктів в грід-структури.
д) Модельні методи (Model-based): використання моделі для знаходження кластерів, найбільш відповідних даним.
є) Графічні методи (відображення графової структури взаємодії / знайомств об'єкта, груп і т.д).

3. Аналіз фото- і відео-зображень. Використовується для виконання таких завдань: аналіз даних, що випадково потрапили в кадр, як вид з вікна, частина назви географічного пункту або відображення в дзеркалі на задньому плані; виявлення «підозрілих» предметів; ідентифікація осіб, які перебувають «в розшуку»; авторизація доступу за відбитками пальців; записи

відеоспостережень; відбитки пальців, знімки сітківки ока; зображення осіб; невербальні сигнали, які дозволять судити про людину (складки на обличчі, що відкривають переважну емоцію людини; улюблені жести; характер взаємовідносин з іншими людьми і т.п.).

Найбільш використовуваними є наступні методи і алгоритми [5]: методи визначення біометричних характеристик, метод головних компонент (зменшує розмірність даних), інтелектуальні методи розпізнавання (віднесення вихідних даних до певного класу за допомогою виділення істотних ознак, що характеризують ці дані, із загальної маси несуттєвих даних), метод еластичних графів (особа представляється у вигляді графа, вершини якого розташовані на ключових точках особи - вузлах, таких як контури голови, губ, носи та їх крайніх точках, кожна грань позначена відстанями між її вершинами), методи нейронних мереж (система з'єднаних і взаємодіючих між собою простих процесорів), генетичні алгоритми (евристичний алгоритм пошуку, який використовується для вирішення завдань оптимізації та моделювання шляхом випадкового підбору, комбінування і варіації шуканих параметрів з використанням механізмів, аналогічних природному відбору в природі), фрактальний аналіз (обчислення фрактальних характеристик для всього зображення, для окремих фрагментів або об'єктів в полі сканування) та інші.

В даний час існує чотири основні методи розпізнавання особи: eigenfaces (використовує двовимірні зображення в градаціях сірого, які представляють відмінні характеристики зображення особи, часто використовується в якості основи для інших методів розпізнавання особи), аналіз «відмінних рис» (подібна методиці «Eigenface», але більшою мірою адаптована до зміни зовнішності або міміки людини), аналіз на основі нейронних мереж (використовують алгоритм, який встановлює відповідність унікальних властивостей людини, що перевіряється, і параметрів шаблону, що знаходиться в базі даних, при цьому застосовується максимально можливе число параметрів), метод «автоматичної обробки зображення обличчя» (використовує відстані та ставлення відстаней між легко визначними точками особи, такими як очі, кінець носа, куточки рота).

4. Глибинний аналіз текстів. Автоматичне вилучення понять і фактів і формалізованих масивів інформації (баз даних, таблиць) і неструктурованих текстів, представлених в веб-просторі, виявлення глибинних зв'язків між поняттями. Для вирішення цих завдань використовуються технології Knowledge discovery, концепції глибинного аналізу даних і текстів (Data Mining, Text Mining) [1, 3-4]. Завдання Text Mining: витяг з тексту його характерних елементів або властивостей, які можуть використовуватися в якості метаданих документа, ключових слів, анотацій; віднесення документа до деяких категорій з заздалегідь заданої схеми класифікації; семантичний пошук документів. До основних методів Text Mining відносять класифікацію (Classification, завдання розпізнавання, де за деякою контрольною вибіркою система відносить новий об'єкт до якоїсь категорії), кластеризації (Clustering, процес виділення компактних підгруп об'єктів з близькими властивостями), побудову семантичних мереж (аналіз зв'язків між поняттями, що екстрагуються з документів), витяг фактів, понять (Feature Extraction, отримання деяких фактів з тексту з метою поліпшення класифікації, пошуку, кластеризації та побудови семантичних мереж), реферування (Summarization, складання коротких викладів матеріалів), відповіді на запити (Question Answering, формування пошукових образів документів), тематичне індексування (Thematic Indexing), пошук за ключовими словами (Keyword Searching), засоби створення і підтримки таксономії (Taxonomies), тезаурусів (Thesauri) і онтологій (Ontology).

5. Моніторинг інформаційного простору - адаптація концепції глибинного аналізу текстів (Text Mining) і класичних методів контент аналізу до умов формування і розвитку динамічних інформаційних масивів даних. Завдання вирішуються системами моніторингу: створення сучасної ефективної інформаційної технології та заснованого на ній он-лайнового сервісу, що акумулюють різні математичні методи та алгоритми, орієнтовані на моніторинг і багатофакторний аналіз медіа-простору. При цьому повинні виконуватися наступні умови: безперервний автоматизований моніторинг медіа-простору, що включає в себе тисячі найбільш рейтингових веб-ресурсів, основні державні і регіональні інтернет-джерела, ЗМІ та телевізійні канали (відеомоніторинг); акумулювання різноманітних розрізнених баз даних в єдину інформаційну систему; накопичення в структурованому вигляді результатів моніторингу для подальшого аналізу; використання сучасних методів кількісного і якісного аналізу інформації. В цілому існуючі на ринку системи можуть надавати такі можливості: SiteSputnik - програмний комплекс, аналоги якого відсутні, що дозволяє вести пошук і обробку його результатів у видимому і невидимому Інтернеті, використовуючи всі необхідні користувачу пошуковики; WebSite-Watcher – дозволяє проводити моніторинг веб-сторінок, включаючи захищені паролем, моніторинг форумів, RSS каналів, груп новин, локальних файлів. Володіє потужною системою фільтрів. Моніторинг ведеться автоматично і поставляється в зручному для користувача вигляді; http://web-data-extractor.net/ – універсальне рішення для отримання будь-яких даних, доступних в інтернеті; CaptureSaver – професійний інструмент дослідження інтернету. Програма, що дозволяє захоплювати, зберігати і експортувати будь-яку інтернет інформацію, включаючи web сторінки, блоги, RSS новини, електронну пошту, зображення і багато іншого; http://www.orbiscope.net/en/software.html – система веб моніторингу; Maltego – принципово нове програмне забезпечення, що дозволяє встановлювати взаємозв'язок суб'єктів, подій і об'єктів в реалі і в інтернетіта т.п.

6. Оцінка залученості користувачів (аналіз передруків, ретвітів, лайків, коментарів) [4-5]. В результаті аналізу збирається статистика того, хто або що збирає максимум лайків від об'єкта, виявляється оцінка впливовості авторів і кількості ботів серед них. Розглядаючи окремо кількість лайків, репостів, коментарів, як і кількість передплатників, можна

зіткнутися з таким, що вводить в оману, параметром успіху, в якому зростання окремих метрик (від числа лайків до кількості передплатників і реєстрацій) розглядається як безсумнівний показник ефективності в соціальних мережах. Насправді ці показники, відірвані від інших даних, не впливають ні на що. У свою чергу, загальна кількість цих метрик (обсяг) може бути використана для розрахунку важливіших показників: коефіцієнта залучення і (в деяких випадках) вартості залучення. Коефіцієнт залучення відповідає на два головних питання: наскільки актуальний та цікавий контент; наскільки аудиторія хоче отримати інформацію.

Для отримання оцінки використовуються семантичний аналіз (компонентна оцінка кількості слів або фраз, що визначають основний зміст тексту – семантичне ядро, і статистичних показників) [6], дослідження каскадів поширення інформації (необхідність того, щоб інформація, яку потрібно донести до аудиторії, мала властивості інформаційного каскаду. Це показник віральності – цікавості контенту) [7]. Метод океану (методи створення безконкурентних ринків, які дозволяють миттєво отримувати додаткову вигоду компанії, її покупцям, її працівникам за допомогою пошуку нового попиту і роблячи конкуренцію, як таку, непотрібною). Визначення ботів і стеження за їх активністю. Робот пошукової машини (бот від англ. Bot), який ходить по сайтах, індексуючи їх контент. Це автоматичний скрипт, який працює за визначеним розкладом. Пошукові боти мають свої юзер-агенти, через які їх легко визначити і мають певну функцію. Наприклад, існують боти, які індексують текстовий контент, відео та аудіо контент, зображення, боти для блогів, для новин та інші роботи. Методи визначення: за допомогою різних сервісів (FakeOFF, VkFak, Twitter Audit і ін.), за IP-адресою і ін. Для виявлення подібності принципів взаємовідносин учасників соціальних мереж використовується такий напрямок аналізу як рольові алгебри, які фокусуються на виявленні логіки взаємодій учасників мережі в блокових моделях [4, 8]. Аналіз діад і тріад дає важливий показник сили зв'язків між учасниками, яка визначається як лінійна комбінація тривалості, емоційної насиченості, інтимності або конфіденційності та значущості взаємних послуг, які характеризують дані взаємодії.

7. Визначення взаємозв'язків. Виявлення неочевидних закономірностей і зв'язків з текстів веб-сторінок і виявлення їх взаємозв'язків, побудова матриць і графів взаємозв'язків. База даних може розглядатися у вигляді графа, вершинами якого виступають об'єкти - терми, поняття та ін., а ребрами - їх зв'язки. Основа пошуку - це пошук вершин, тобто пошук об'єктів. При проектуванні баз даних зв'язків використовуються перспективні рішення в області створення інформаційно-аналітичних систем – Text Mining, Information Extraction, Big Data (технології баз даних надвеликих обсягів), Complex Network (концепція складних мереж). В рамках теорії складних мереж вивчаються характеристики, пов'язані з топологією мережі, статистичні феномени, розподіл ваг окремих вершин і ребер, ефекти провідності і протікання в мережах.

Пов'язані підгрупи (спільноти) в мережі характеризуються наявністю великої кількості зв'язків між учасниками, що входять до них, та істотно меншим числом зв'язків з іншими учасниками. Аналіз спільнот дозволяє вивчати стійкість соціальних структур [1,3,9]. Найпростіший випадок пов'язаної групи – це спільнота, де кожен учасник пов'язаний з кожним, і до цієї групи не можуть бути включені інші учасники мережі, оскільки вони не мають зв'язків з усіма членами спільноти (кліки). Таким чином, кліка - це максимальний повний підграф даного графа. Якщо аналізувати процеси поширення інформації в графах, то можна дати інше визначення спільноти, як множини учасників, де шлях між двома будь-якими учасниками не містить більше однієї проміжної вершини. В результаті інформація від одного учасника до іншого в пов'язаній групі передається з мінімальними спотвореннями. Пов'язані групи також можуть бути виділені за допомогою багатовимірного шкалювання (аналіз і візуалізація даних за допомогою розташування точок, відповідних досліджуваним – шкаліруемим об'єктам, в просторі меншої розмірності ніж простір ознак об'єктів) або факторного аналізу матриці зв'язків графа (зниження розмірності матриці зв'язків графа, тобто виділення у всій сукупності ознак тих, які дійсно впливають на зміну залежною змінною) [7].

8. Оцінка охоплення. Одна справа – кількість згадок, і зовсім інша справа – охоплення аудиторії, одне без іншого не має практично ніякого сенсу. Охоплення (reach) – це кількість представників цільової аудиторії, які в рамках кампанії мали контакт з поширюваною інформацією (новиною, рекламою, дезінформацією і т.д.) певну кількість разів [3,9]. Чим більше кількість потенційних людей, до яких реклама була донесена, тим більше охоплення. Ідеальне охоплення – отримання повідомлення ста відсотками цільової аудиторії. Однак на практиці повне охоплення залишається, як правило, недосяжною метою. Охоплення аудиторії може бути представлене як характеристика аудиторії, яка бачила (чула) інформацію певну кількість разів, і як характеристика аудиторії, яка бачила (чула) інформацію не менше певної кількості разів. Оцінка охоплення здійснюється по кожному носію інформації. Найбільш важливою є оцінка того, чи вдалося забезпечити досягнення цілей охоплення і частоти. Охоплення може бути виражене як кількісно, тобто в одиницях, тисяча, мільйони людей, так і у відсотках від цільової аудиторії. Для збільшення охоплення використовують різні методи, включаючи використання негативної інформації. Справа в тому, що негативна інформація поширюється набагато швидше і охочіше, а також часто має набагато більшу вагу. Тобто картина співвідношення позитиву і негативу в абсолютній кількості знайдених згадок сильно відрізняється від картини співвідношення при оцінці за охопленням.

9. Демографічна і географічна класифікація. Стать, вік, трудові ресурси, географічний розподіл, расовий, етнічний, релігійний склад населення і інше. Основні методи класифікації: за допомогою дерев рішень (метод автоматичного аналізу даних та подання правил в ієрархічній, послідовній структурі, де кожному об'єкту

відповідає єдиний вузол, що дає рішення) [4]; байєсівська (наївна) класифікація (простий імовірнісний класифікатор, заснований на застосуванні теореми Байєса зі строгими (наївними) припущеннями про незалежність); за допомогою штучних нейронних мереж (система з'єднаних і взаємодіючих між собою простих процесорів (штучних нейронів) [10]; метод опорних векторів (набір схожих алгоритмів навчання з учителем, що використовуються для задач класифікації і регресійного аналізу) [8]; статистичні методи, зокрема, лінійна регресія (регресійна модель залежності однієї змінної від іншої або кількох інших змінних з лінійною функцією залежності) [6]; за допомогою методу найближчого сусіда (метричний алгоритм для автоматичної класифікації об'єктів. Основним принципом методу найближчих сусідів є те, що об'єкт присвоюється тому класу, який є найбільш поширеним серед сусідів даного елемента) [4]; CBR-метод (це метод отримання рішення шляхом пошуку подібних проблемних ситуацій в пам'яті, що зберігає минулий досвід вирішення завдань, і адаптації знайдених рішень до нових умов) [3,4]; за допомогою генетичних алгоритмів (евристичні алгоритми пошуку, які використовуються для вирішення завдань оптимізації та моделювання шляхом випадкового підбору, комбінування і варіації параметрів, які шукаються, з використанням механізмів, аналогічних природному відбору в природі) [10].

10. Аналіз тональності висловлювань по відношенню до інформаційних об'єктів. Проводиться кластеризація повідомлень за сюжетами з можливістю знайти кожне окреме повідомлення і його характеристики (в тому числі тональність); ранжування сюжетів за обговорюванням. Також використовується для попередження про можливі ризики – полягає в спостереженні за кількісною та якісною динамікою обговорення інформаційного об'єкта і прогнозі подальшої динаміки. Отримання інформації з коротких неформальних повідомлень, таких, як твіти та смс, відстеження подій, вилучення проблем, виявлення сарказму та ін. Традиційний підхід до класифікації тональності ґрунтується на присутності негативних і позитивних слів або піктограм, що зображують емоцію (емотикони), які служать індикаторами позитивного або негативного забарвлення. Застосовуються також змішані підходи, де лексичні ресурси (словники тональності) комбінують з методами машинного навчання. Аналіз проводиться з використанням рішення бінарної класифікації [11], завдання класифікації послідовностей (sequential classification) [12], завдання тематичного моделювання або традиційної задачі кластеризації [13]. Мета класифікації – визначити, чи є терміни, іменники та словосполучення шуканим об'єктом (аспектом, щодо якого висловлюється деяка думка). В [14] пропонується підхід, заснований на правилах, в якому використовуються дерева залежностей для пропозицій. Також використовуються методи умовних випадкових полів (Conditional Random Fields, CRF) для класифікації послідовностей в завданні вилучення аспектних термінів.

11. Оцінювання транзакцій в інтернеті (фінансові транзакції через кредитні карти, покупки і використання авіаквитків, інформація про інших людей, з якими об'єкт веде справи, а також про те, як це все вписується в загальну історію транзакцій і т.д.) як частина «загального патерну активності» користувача. Використовуються методи та інструменти оцінки ризиків на базі моніторингу транзакцій, методи машинного навчання, статистичні методи.

12. Аналіз потоків подій (можливість «пов'язувати» і оцінювати вплив зовнішніх подій на ключові події / заходи). Застосовується для своєчасного виявлення шахрайства в сфері фінансових послуг; для оперативного аналізу активностей на біржових торгових майданчиках; для оперативного аналізу відтоку клієнтів і проведення маркетингових кампаній, для оцінки впливу політики на соціальну та економічну області та ін.

Обробка потоків подій (англ. Event Stream Processing, ESP) - набір технологій, призначених для побудови інформаційних систем обробки подій. ESP технологія включає в себе візуалізацію подій, їх зберігання, кероване по подіям сполучне програмне забезпечення та мови програмування обробки подій. Основним завданням для ESP є обробка потоків подій (даних) з метою виявлення в них значущих шаблонів, використовуючи такі методи як пошук взаємозв'язків між подіями, кореляція подій, ієрархії подій, та інші аспекти, такі, як причинність, аналіз складових подій і часових рядів.

Для аналізу використовується концептуальна модель обробки подій. Вона розглядається в термінах мережі обробки подій, яка лежить в її основі (Event Processing Network), та асоційованої з нею концептуальної архітектури для обробки подій (Conceptual Architecture for Event Processing). Також використовується теорія ризиків (ризик - це комбінація ймовірності події та її наслідків), кореляційний аналіз (метод, що дозволяє виявити залежність між декількома випадковими величинами).

13. Прогнозування. Методи прогностичного моделювання, які дозволяють за минулим передбачати майбутнє, застосовуються при вирішенні безлічі завдань, таких як рекомендуючі системи, виявлення шахрайства та зловживань або профілактика захворювань і нещасних випадків.

Прогноз успішності процесу передачі дезінформації, пошук закономірностей в соцмережах для генерації даних, необхідних для прогнозування районів, в яких з найбільшою ймовірністю можуть відбуватися злочини і терористичні акти, рекомендація товарів і послуг, ґрунтуючись на звичках об'єктів, допомога лікарям приймати рятівні профілактичні заходи з урахуванням сприйнятливості людини до конкретних захворювань, для передбачення епідемій, а деякі навіть пророкують стихійні лиха і їх наслідки, аналізуючи контент соціальних медіа [8]. Прогностичні моделі використовуються і для реалізації функції класифікації, коли результатом є клас або категорія. Найбільш популярними є методи: нейронних мереж (NN) [10], кластеризації [4], метод опорних векторів (SVM) , асоціативні правила, дерева рішень, лінійна і логістична регресія і оціночні таблиці. Для передбачення появи нових зв'язків в мережі за проміжок часу в [2] застосовується автоматичне моделювання процесу

розвитку соціальної мережі з залученням деяких характеристик мережі, таких як кількість спільних сусідів, найкоротший шлях, впливовість вершини, момент першого попадання в соціальну мережу. Також застосовуються прогностичні моделі засновані на машинному навчанні, що використовують особисту інформацію про користувачів мережі для підвищення точності передбачення [15]. Іноді застосовують ієрархічні, імовірнісні (марковські) і реляційні моделі для виявлення зв'язків між користувачами [16].

В інших моделях [9, 16] за основу пропонується брати самі властивості користувачів, і, наприклад, наявність великої кількості зв'язків (в блогосфері), що може бути пояснено шляхом зіставлення демографічних груп, спільних інтересів або географічною близькістю.

III. Висновки

У даній статті розглянуто основні види аналізу даних в соціальних мережах. Для кожного виду аналізу наведено перелік необхідних даних, короткий огляд основних математичних методів, які використовуються для проведення аналізу. Також описані показники, які можна отримати після аналізу, їх інформативність та використання в подальшому аналізі або представлені результатів аналізу. Описано взаємозв'зки показників даних та аналітичних розрахунків для кожного виду аналізу соціальних мереж.

Література REFERENCES

[1] Т.В. Батура Модели и методы анализа компьютерных социальных сетей. // Батура Т.В. / Программные продукты и системы, № 3. – 2013. - С. 130-137.

[2] Liben-Nowell D., Kleinberg J. The Link Prediction Problem for Social Networks // Proceedings of the 12th International Conference on Information and Knowledge Management. N. Y .: ACM Press, 2003. P.556-559.

[3] Д.В.Ланде, А.А.Снарский , И.В.Безсуднов Интернетика: навигация в сложных сетях: модели и алгоритмы. М. книжный дом «Либроком», 2009. – 264с.

[4] F.Sebastiani Machine Learning in Automated Text Categorization. URL: http://nmis.isti.cnr.it/sebastiani/Publications/ACMCS02.pdf.3.

[5] Н.Г.Федотов Теория признаков распознавания образов на основе стохастической геометрии и функционального анализа. – М.:ФИЗМАТЛИТ, 2010. – 304с.

[6] G.Sidorov, S.Miranda-Jimenez, F.Viveros-Jimenez, A.Gelbukh, N.CastroSanchez, F.Velasquez, J.Gordon Empirical study of machine learning based approach for opinion mining in tweets // Advances in Artificial Intelligence. Springer Berlin Heidelberg, 2013. P. 1-14.

[7] Как использовать социальный граф для распространения контента URL: http://www.pvsm.ru/sotsial-ny-e-seti/64097)

[8] CJC Burges. A Tutorial on Support Vector Machines for Pattern Recognition. URL: http://www.music.mcgill.ca/_rfergu/adamTex/references/Burges98.pdf

[9] R.Kumar, J.Novak, P.Raghavan, A.Tomkins Structure and Evolution of Blogspace // Commun. ACM. 2004. Vol. 47. No. 12. P. 35-39.

[10] Leonid Lyubchyk, Eugenie Bodyansky, Arkady Rivtis. Adaptive harmonic components detection and forecasting in wave non-periodic time series using neural networks. URL: https://pdfs.semanticscholar.org/c879/8f1ac64fa076c9df437a860645fd5e463040.pdf.

[11] A.M. Popescu, O.Etzioni Extracting product features and opinions from reviews // Natural language processing and text mining. ACL, 2007. P. 9-28.

[12] N.Jakob, I.Gurevych Extracting opinion targets in a single-and crossdomain setting with conditional random fields // Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing. ACL, 2010. P. 1035-1045.

[13] Y.Zhao, B.Qin, T.Liu Clustering product aspects using two effective aspect relations for opinion mining // Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data. Springer International Publishing, 2014. P. 120-130.

[14] S.Poria, E.Cambria, W.Ku, C.Gui, A.Gelbukh A rule-based approach to aspect extraction from product reviews // Proceedings of the Second Workshop on Natural Language Processing for Social Media (SocialNLP). 2014. P. 28-37.

[15] B.Taskar, M.Wong, P.Abbeel, D.Koller Label and Link Prediction in Relational Data. URL: http://kdl.cs.umass.edu/srl2003_upload/files/taskar-paper.pdf.

[16] Saoussen Aouay, Salma Jamoussi, Faiez Gargouri, Ajith Abraham. Modeling Dynamics of Social Networks: A Survey. Sixth International Conference on Computational Aspects of Social Networks. – 2014. P.49-53.

Вдосконалення Методу Моніторингу Цифрових Слідів Інцидентів Інформаційної Безпеки у Сегментах Національної Інформаційної Інфраструктури

Надія Казакова
кафедра комп'ютерних та інформаційно-вимірювальних технологій
Одеська державна академія технічного регулювання і якості
Одеса, Україна
kaz2003@ukr.net

Олексій Фразе-Фразенко
кафедра комп'ютерних та інформаційно-вимірювальних технологій
Одеська державна академія технічного регулювання і якості
Одеса, Україна
frazenko@gmail.com

Monitoring Method Improvement of Information Security Incidents Digital Traces In National Information Infrastructure

Nadiia Kazakova
Department of computer, information and measurement technologies
Odesa State Academy of Technical Regulation and Quality
Odesa, Ukraine
kaz2003@ukr.net

Oleksii Fraze-Frazenko
Department of computer, information and measurement technologies
Odesa State Academy of Technical Regulation and Quality
Odesa, Ukraine
frazenko@gmail.com

***Анотація*—У межах вирішення проблеми виявлення найбільш захищених сегментів національної інформаційної інфраструктури засобами моніторингу сегментів інформаційного простору спеціального призначення, розглядається метод моніторингу стану безпеки заданих сегментів інфраструктури. Показано, що пропоновані удосконалення дозволяють забезпечити підвищення ймовірності оцінювання їх стану по відношенню до захищеності. Враховано, що характеристики сегментів містять велику кількість контрольованих параметрів, які з метою підвищення ефективності роботи систем моніторингу можуть бути об'єднані у відповідності до їх вагових внесків в окремих елементах, що підтримують функціонування різноманітних складових та процесів у системах забезпечення інформаційної безпеки.**

***Abstract*—Within solving the problem of the national information infrastructure most protected segments identifying by the tools of segment information space special purpose monitoring, the method of infrastructure safety monitoring of defined segments is considered. It is shown that the offered improvements allow to provide increase in probability of estimation of their state in relation to security. It is considered that characteristics of segments contain a large number of controlled parameters which for the purpose of increase in overall performance of monitoring systems can be integrated according to their weight contributions in separate elements which support functioning of different components and processes in systems of information security ensuring.**

Ключові слова—інформація; безпека; інформаційна інфраструктура; параметр; моніторинг; інформаційний простір

Keywords—information; security; information infrastructure; parameter; monitoring; information space

I. Вступ

Тривалий час застосування традиційних засобів захисту інформації, таких як системи розмежування прав доступу, використання міжмережевих екранів, антивірусного програмного забезпечення тощо, було стандартним методом забезпечення інформаційної безпеки (ІБ) будь-якої комунікаційної системи чи мережі. Відомості про їх роботу та ефективність можуть бути використані на новому етапі розвитку систем захисту інформації (СЗІ) щодо забезпечення ІБ, а саме – системами моніторингу інформаційного простору з метою виявлення його найбільш безпечних та захищених сегментів для переміщення (міграції) до них обчислювальних ресурсів та даних, що забезпечить підвищення ступеню їх конфіденційності, цілісності та доступності. Такий підхід передбачає постійний динамічний процес моніторингу стану інформаційних процесів, що пов‘язані з забезпеченням ІБ, і з часом може стати невід‘ємною частиною ідеології функціонування національної інформаційної інфраструктури (НІІ) [1, 2].

Застосування систем моніторингу НІІ з метою організації міграції даних суттєво підвищить ефективність наявних в інформаційній системі засобів забезпечення ІБ за рахунок синергетичного ефекту при обробці інформації

II. Метод моніторингу цифрових слідів інцидентів інформаційної безпеки

Методологія забезпечення ІБ державних ресурсів, яка передбачає їх міграцію до найбільш безпечних сегментів інформаційного простору, що може контролюватися єдиним центром обробки даних (ЦОД), вимагає використання різних засобів виявлення у них слідів мережевих атак, наявності та ефективності систем захисту від спаму, дієвості антивірусних засобів, ефективності міжмережевих екранів та сканерів безпеки, доступності, надійності та ефективності обчислювальних та інших технічних ресурсів [3]. При цьому розуміється, що доступ до відомостей, які характеризують зазначене, може бути забезпечений на основі відповідних угод між суб’єктами, які обслуговуються єдиним ЦОД.

Далі під загальним поняттям «ЦОД», який формуватиме управляючі впливи стосовно міграції даних, будемо розуміти комплексне організаційно-технічне рішення, метою функціонування якого є створення та підтримка високопродуктивної та відмовостійкої інформаційно-телекомунікаційної інфраструктури у межах виділеного інформаційного простору. Його загальним завданням буде ефективне консолідоване зберігання та обробка даних користувачів з одночасним наданням їм прикладних сервісів та підтримка функціонування корпоративних додатків. Обробка отриманих даних веде до зростання множини апаратно-програмних засобів забезпечення ІБ та суттєвого росту обсягів інформації, яка може бути необхідною для контролю мережевої безпеки. Відповідно, існує необхідність автоматизації зазначених процесів з метою підвищення продуктивності робіт з обробки даних, та рішення завдань щодо оперативності прийняття управляючих рішень для організації міграції даних та обчислювальних ресурсів. Це дозволить вирішити протиріччя між значним зростанням обсягів інформації, яка обробляється та аналізується для встановлення рівня безпеки визначених мережевих ресурсів, наявності загроз для них та їх ступенем, та оперативністю управління міграцією [4].

У межах вирішення проблеми виявлення безпечних сегментів НІІ на основі застосування систем моніторингу сегментів інформаційного простору спеціального призначення (СМСІПСП) [5, 6], розглянемо метод моніторингу стану безпеки заданих сегментів НІІ, який забезпечує підвищення ймовірності оцінювання їх стану по відношенню до захищеності [2, 7]. Враховуватимемо, що кожна з характеристик сегментів НІІ містить велику кількість контрольованих параметрів, які з метою підвищення ефективності роботи СМСІПСП можуть бути об‘єднані у відповідності до їх вагових внесків у окремих елементах, які забезпечують функціонування різноманітних складових та процесів у системах забезпечення ІБ [8]. Групування може виконуватися у тих випадках, коли є доцільним врахування динаміки керування часовими проміжками при ухваленні управляючих рішень щодо стану захищеності сегментів НІІ та врахування ступеню впливу груп контрольованих параметрів та величин відхилення їх значень від заданих.

Згідно до [7], існує метод, відповідно до якого моніторинг стану систем забезпечення ІБ в деякій системі, базується на попередньо заданій множині з $X > 2$ контрольованих параметрів безпеки. Крім того, метод передбачає, що $Y \geq X$ зразкових значень параметрів безпеки, які підлягають контролю, а також їх вагові коефіцієнти $k^{ваг}$, є заданими.

Маючи зазначені дані у якості вхідних параметрів, може бути виконаний аналіз, суть якого полягає у здійсненні наступних процедур:

- вимірювання значень контрольованих параметрів безпеки;
- порівняння їх зі зразками;
- формування звіту;
- формування управляючого рішення щодо стану систем забезпечення ІБ.

Згідно до [7], можливе додаткове формування $Z \geq 2$ груп параметрів, які підлягають контролю, з числа попередньо заданих контрольованих параметрів. Така процедура є необхідною у зв‘язку з тим, що кожна z-а група контрольованих параметрів, де $z = 1, 2, \ldots, Z$, є окремою характеристикою стану ІБ z-го структурного елемента або функціонального процесу, які відбуваються у системі, що контролюється засобами СМСІПСП. Вважається, що коефіцієнти важливості $k_z^{ваг}$ є незалежними та можуть бути задані для кожної z-ї групи. Для кожної такої групи параметрів повинно бути задано максимальне Δt_z^{max} та мінімальне Δt_z^{min} значення проміжків часу, протягом яких відбувається вимірювання

параметрів, що контролюються, а також момент часу $t_z^{звіт}$, який передбачає формування звіту про стан безпеки досліджуваної системи.

До процедури встановлення попередніх параметрів також може бути віднесено зазначення інтервалу часу вимірювань параметрів, що контролюються, наприклад, для z -ї групи, який дорівнює максимальному, тобто Δt_z^{max}. Після виконання процедури порівняння виміряних значень з заданими зразками, при їх співпадінні, цикл аналізу безпеки системи повинен бути повторений до настання моменту часу $t_z^{звіт}$ формування звіту про безпеку системи – у випадку, який розглядається, це досліджуваний сегмент НІІ. Якщо при порівнянні значення отриманих параметрів не співпадають зі встановленими зразками, то їх запам'ятовують або заносять до бази даних.

Після виконання процедури порівняння, необхідно внести корективи значень часових інтервалів вимірювань, $\Delta t_z^{кор} = \frac{\Delta t_z^{max}}{k_z^{ваг}}$. Далі необхідно отримане значення $\Delta t_z^{кор}$ порівняти з мінімальним Δt_z^{min}. Якщо $\Delta t_z^{кор} = \Delta t_z$, то цикл аналізу безпеки необхідно виконати повторно, а у протилежному випадку, коли $\Delta t_z^{кор} \le \Delta t_z^{\min}$, формується повідомлення про вихід контрольованих параметрів в z -й групі за межі припустимих значень. Як наслідок, СМСІПСП реєструє наявність порушень у системі забезпечення ІБ у контрольованому сегменті.

Аналіз приведеного методу виявив, що для забезпечення достовірного знаходження слідів інцидентів ІБ у спеціальних сегментах НІІ, необхідно достатньо часто та з високою вірогідністю виконувати моніторинг та аналіз параметрів, які характеризують стан СЗІ у них. Окрема актуальна задача – вчасне формування звітів про стан ІБ та передавання його до вищого рівня ієрархії СМСІПСП з метою формування управляючого рішення. Процедура моніторингу параметрів, що визначають стан систем захисту досліджуваної системи, здійснюється на основі локального зчитування їх значень з оперативної пам'яті контрольованих елементів та порівняння зі зразками. Процедура виконується на основі протоколів мережної взаємодії. Таким чином, значення параметрів СМСІПСП отримуються з оперативної пам'яті, що унеможливлює використання процедури для отримання відомостей про ретроспективний стан систем забезпечення ІБ у контрольованому середовищі. Відповідно, необхідно виконати її удосконалення з метою отримання значень параметрів з журналів реєстрації інцидентів ІБ зі збереженням показників економічної ефективності.

III. Вдосконалення методу моніторингу цифрових слідів інцидентів інформаційної безпеки

Вважатимемо, що контрольовані сегменти НІІ є складними інформаційними технічними структурами та містять велику кількість елементів, які підтримують функціонування систем забезпечення ІБ. У загальному випадку для моніторингу ретроспективних станів безпеки інформаційної структури, значна частина ресурсів, наприклад, відомостей про пропускну здатність каналів зв'язку, що змінена внаслідок DDos-атаки, не можуть бути отримані без даних про те, як вона функціонувала у штатному режимі. Збільшення ж множини значень про сукупні контрольовані ресурси шляхом розгортання локальної системи моніторингу веде до значних економічних затрат. З метою вирішення питання використаємо узагальнену схему (рис. 1), яка пояснює групування параметрів структурних елементів досліджуваної системи.

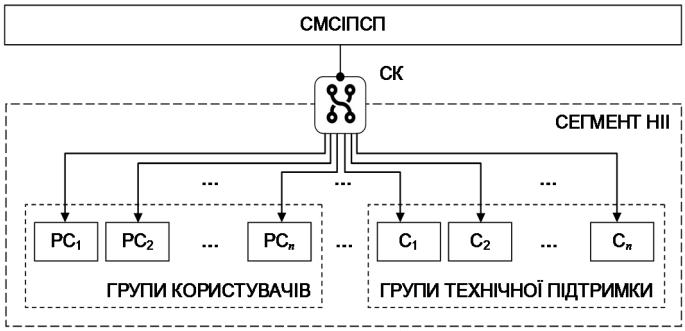


Рис. 1. Принцип групування параметрів структурних елементів досліджуваної системи, де: PCN – робочі станції, CN – сервери, комутатори та маршрутизатори, які програмно керуються; СК – системний комутатор, що управляється СМСІПСП

Організація системи моніторингу, яка приведена на рис. 1, дозволяє СМСІПСП визначати ті об'єкти у структурі сегменту НІІ, які задані завданням щодо визначення станів їх безпеки і, т.ч., управляти отриманням значень параметрів з журналів реєстрації інцидентів ІБ. При цьому підконтрольними СМСІПСП і, відповідно, державному ЦОД, будуть відомості про наступні дані, які мають відношення до інцидентів інформаційної безпеки:

- перелік активних логічних портів, які були задіяні у інцидентах (АЛП);
- про Ір-адреси, які були задіяні у створені інцидентів (ІР-адр);
- про вплив інцидентів на активність диспетчера підключень дистанційного доступу (АДПДД);
- про порушення вимог захисту даних у службах папок обміну даними (ЗДСПОД);
- про порушення локальних налаштувань сервера служби файлового обміну (ЛНССФО).
- про порушення локальних налаштувань захисту служб електронної пошти (ЛНЗСЕП);
- про порушення локальних налаштувань користувацького клієнта служб файлового обміну (ЛНККСФО);

У залежності від коефіцієнтів важливості $k_z^{ваг}$, відомості можуть бути об'єднані у групи, що підвищить ефективність роботи СМСІПСП. Відповідно до [7], доцільним є їх об'єднання у вигляді, наведеному на рис. 2, де також відзначено додаткові характеристики

структурних елементів контрольованого сегменту (відповідно до рис. 1).

Встановимо, що контрольований сегмент НІІ є множиною функціональних вузлів та телекомунікаційного обладнання. На рис. 1 наведено умовний досліджуваний сегмент, який керує з'єднаннями та управляється СМСІПСП. Ним, на основі використання закріплених ідентифікаторів, визначається множина технічних та програмних активів, які підлягають моніторингу. У якості ідентифікаторів найбільш доцільним є застосування мережних адрес з сімейства протоколів TCP/IP [7].

Для перевірки практичного застосування методу та окремих функцій з його удосконалення, встановимо, що множина функціональних вузлів та телекомунікаційного обладнання забезпечує: виконання функціональних процесів з контролю за обміном файлів, виконання функціональних процесів щодо контролю передавання електронної пошти, виконання функціональних процесів у межах мережної взаємодії.

Ці процеси є найбільш характерними при їх дослідженні з метою виявлення наслідків впливу інцидентів ІБ. При практичній перевірці журналів реєстрації інцидентів найбільш просто організувати доступ саме до них. В подальшому розумітимемо, що для СМСІПСП доступ до журналів забезпечується на основі відповідних договорів.

Записи у журналах активності, які підлягають моніторингу та аналізу, базуються на контролюванні функціональних процесів ІБ у точках входу до робочих додатків, які відповідають логічним портам на яких реалізується робочий процес. Кожна множина функціональних вузлів та телекомунікаційного обладнання, а також їх функціональні процеси, характеризуються деякою попередньо встановленою сукупністю параметрів, що описують стан їх ІБ і, т.ч., ці відомості реєструються у журналах та є досяжними для СМСІПСП. До них, як до елементів сегменту, що досліджується засобами СМСІПСП, віднесемо такі ж відомості, які наведено вище та згрупуємо (рис. 2).

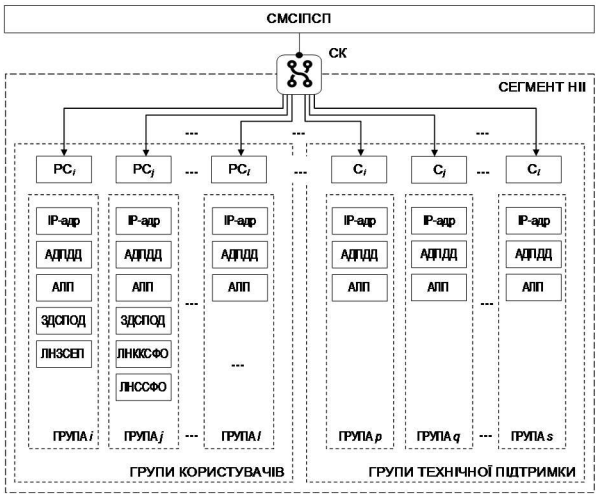
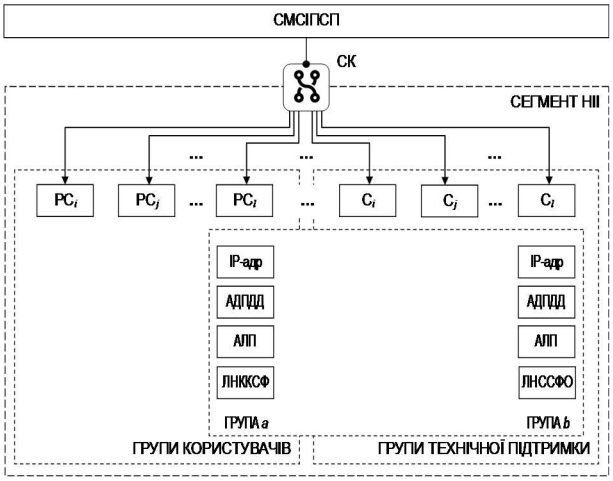
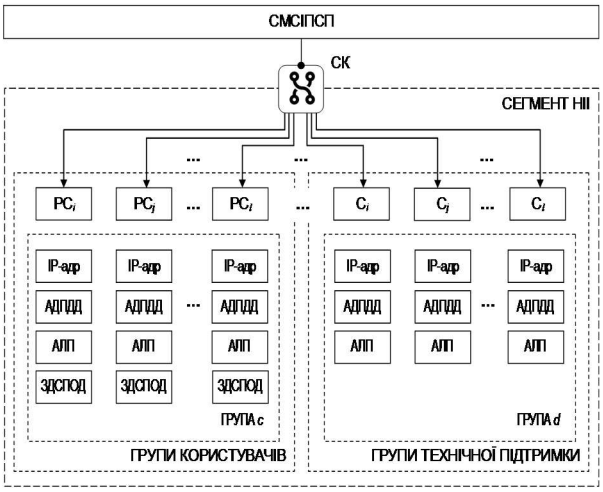


Рис. 2. Узагальнене групування додаткових характеристик структурних елементів контрольованого сегменту

Як видно з нього, взаємозв'язок параметрів контрольованих функціональних процесів ІБ у точках входу до робочих додатків, викликає необхідність їх додаткового групування. Це будуть 4 групи відомостей про інциденти ІБ у службах, що підтримують відповідно: a − обіг файлів; b − функціонування електронної пошти; c − функціонування папок обміну; d − мережевої взаємодії. Зазначені групи відображено на рис. 3.



а)



б)

Рис. 3. Додаткові групи відомостей про інциденти ІБ: a – відомості про загальне використання додатків користувачами та технічною підтримкою; б – відомості про розмежоване використання додатків користувачами та технічною підтримкою

Засоби контролю за інцидентами, які відносяться до приведених груп, внесуть записи до відповідних розділів журналів активності при їх виникненні. Це призведе до значного спрощення та пришвидшення роботи СМСІПСП. При цьому повинно враховуватися, що для кожної групи параметрів, що контролюються, та яка сформована у відповідності до приведеної схеми, попередньо задаються вагові коефіцієнти у відповідності до важливості групи, тобто у відповідності до ступеню порушення інцидентом ІБ гарантованого рівня конфіденційності, цілісності та доступності інформації у контрольованому сегменті.

Крім того, їх значення, а також зразкові величини (чи вимоги), повинні враховувати дані про важливість інформації та можливі потенційні вразливості контрольованих функціональних процесів ІБ. Для цього можуть бути використані методики, що базуються на достатньо відомих методах неформальної логіки та методах залучення експертів. При цьому, найбільш доцільним та ефективним для розрахунку надійності та достовірності отримання даних системами СМСІПСП є завдання значень параметрів у вигляді матриці. Пропонується, для ефективного рішення питання надійного та достовірного отримання даних про інциденти ІБ використання теорії псевдонапівзворотних матриць [14].

Як випливає з вище приведеного, врахування сукупності визначальних ознак веде до реалізації можливості комплексної оцінки стану параметрів у системах забезпечення ІБ з одночасною мінімізацією використовуваних ресурсів, а також використанням можливості адаптивного управління характеристиками процесу моніторингу. Це свідчить про те, що реалізація методу є можливою без доповнення контрольованих сегментів НІІ додатковими активними сенсорами, які виконують моніторинг стану СЗІ. Це, у свою чергу, веде до підвищення ймовірності оцінювання стану захищеності контрольованих сегментів з одночасним підвищенням економічної ефективності СМСІПСП у цілому. В основу роботи СМСІПСП, з незначними доробками, може бути покладено алгоритм, який наведено, наприклад, у [7]. Згідно до зазначеного джерела, алгоритм може забезпечити виявлення цифрових слідів інцидентів ІБ у контрольованих сегментах НІІ.

IV. Висновки

В процесі моніторингу слідів інцидентів ІБ у журналах записів, з високою ймовірністю можна виявити не тільки факт відмінності значень параметрів, що контролюються, від заданих, але й за рахунок їх групування по ваговому внеску підвищити ймовірності оцінювання стану захищеності контрольованого сегменту НІІ. Динамічне управління інтервалом часу прийняття управляючого рішення про ретроспективний стан безпеки сегменту НІІ забезпечує підвищення економічної ефективності СМСІПСП, що є додатковим обґрунтуванням актуальності викладеного.

Наведені дані можуть бути використані для подальшого розвитку теорії забезпечення ІБ у інформаційних структурах, а також для вирішення загальної науково-прикладної проблеми з розробки моделей та комплексних методів проактивного забезпечення ІБ в когнітивних мережах, у яких основою побудови та управління є SDN-технології. При цьому може бути враховано існування некоректних задач, пов’язаних з невизначеністю процесів у СМСІПСП [9], та деструктивних впливів [10], які порушують вимоги щодо конфіденційності, цілісності та доступності. Рішення зазначеної проблеми, в порівнянні з існуючими принципами функціонування комплексних СЗІ, дозволяє розробити концепцію безпечної міграції даних та обчислювальних ресурсів у межах хмарних структур, а також розширити теоретичні та практичні межі загальних принципів функціонування СЗІ у них, що веде до підвищення ступеня захисту інформаційних процесів, які володіють властивостями невизначеності [11]. Доцільними питаннями, які є перспективними щодо подальших досліджень у галузі забезпечення ІБ, є питання надійності та живучості СМСІПСП [12, 13].

Література References

[1] О. О. Скопа, Аналіз розвитку сучасних напрямів інформаційної безпеки автоматизованих систем [Текст] / О. О. Скопа, Н. Ф. Казакова // Системи обробки інформації. — Харків : Харківський ун-т Повітряних Сил ім.І.Кожедуба. — 2009. — № 7(79). — 2009. — С. 48-54.

[2] Н. Ф. Казакова, Моніторинг інформаційних ресурсів в захищених інформаційних мережах [Текст] / Н. Ф. Казакова // Світ інформації та телекомунікацій : VII міжнар. наук.-техн. конф. студентства та молоді, 15-16 квітня 2010 р. — ДУІКТ : Київ. — С. 165-168.

[3] Н. Ф. Казакова, Оцінка живучості систем моніторингу інформаційного простору [Текст] / Н. Ф. Казакова // Восточно-европейский журнал передовых технологий. — Харьков : Технологический центр. — 2012. — № 4/2(58). — С. 12-15.

[4] Н. Ф. Казакова, Визначення показників для вирішення завдань прогностичного контролю мультисервісних телекомунікаційних мереж / Н. Ф. Казакова, О. О. Скопа // Сучасний захист інформації. — К. : ДУІКТ. — 2010. — Спецвипуск (4). — С. 55-61.

[5] Н. Ф. Казакова, Застосування програмно реалізованого прогностичного контролю для вирішення практичних завдань забезпечення якості надання послуг у захищених інформаційних мережах / Н. Ф. Казакова // Сучасна спеціальна техніка. — К. : Державний науково-дослідний інститут МВС України. — 2012. — № 2(29). — С. 86-95.

[6] О. О. Скопа, Проблематика якості послуг інтернет-провайдерів / О. О. Скопа, С. Л. Волков, К. Б. Айвазова // Збірник наукових праць Одеської державної академії технічного регулювання та якості. — 2013. – № 1(2). — С. 27-31.

[7] Спосіб моніторингу безпеки автоматизованих систем : пат. № 2355024 : МПК G06F15/00, G06F17/00 / О. С. Євстигнєєв К. М. , Зорін, М. О. Карпов [та ін.] ; заявник та патентообладач Військова академія зв‘язку ім. С. М. Будьонного ; заявл. 12.02.2007 ; опубл. 10.05.2009.

[8] С. Л. Волков, Оптимізація параметрів телекомунікаційної мережі методом статистичної регуляризації / С. Л. Волков, Н. Ф. Казакова // Сучасна спеціальна техніка. — К. : Державний науково-дослідний інститут МВС України. — 2012. — № 1(28). — С. 54-60.

[9] Н. Ф. Казакова, Некоректні задачі відновлення даних у системах моніторингу інформаційного простору / Н. Ф. Казакова // Вісник Східноукраїнського національного університету імені Володимира Даля. — Луганськ : СНУ ім. В.Даля. — 2012. — № 8(179). — Т. 1. — С. 325-332.

[10] О. В. Грабовський, Регуляризація визначення показників якості функціонування ІВС з врахуванням нечіткості інформації / О. В. Грабовський, С. Л. Волков, О. О. Скопа // Вісник Національного технічного університету «ХПІ» : Нові рішення в сучасних технологіях. — 2013. — №26 (999). — С.169-174.

[11] О. О. Скопа, Інтелектуальні автономні системи: концептуальні положення створення та функціонування / О. О. Скопа, Є. В. Вавілов // Бионика интеллекта. — 2013. — №1(80). — С. 35-40.

[12] О. В. Грабовський, Скорочення випробувань надійності ІВС за рахунок її функціональної надмірності / О. В. Грабовський, Н. Ф. Казакова // Технологічний аудит та резерви виробництва. — 2013. — №2/1(10). — С. 24-27.

[13] О. О. Скопа, Концепція контрольних випробувань резервних систем на основі біноміальної схеми / О. О. Скопа, С. Л. Волков, А. В. Мінін // Інформаційна безпека. — 2011. — №2(6). — С.69-76.

[14] С. А. Петренко, Управление информационными рисками. Экономически оправданная безопасность / С. А. Петренко, С. В. Симонов. — М. : АйТи; ДМК Пресс, 2004. — 384 с.

The Authentication Method in Wireless Sensor Network Based on Neighbor Trust Model

Olesia Voitovych
dept. of Information Protection
Vinnytsia National Technical University
Vinnytsia, Ukraine
voytovych.olesya@vntu.edu.ua

Olga Shulyatitska
dept. of Information Protection
Vinnytsia National Technical University
Vinnytsia, Ukraine
olya_olek@ukr.net

Метод Автентифікації Сусідніх Вузлів в Сенсорній Мережі на Основі Моделі Довіри

Олеся Войтович
кафедра захисту інформації
Вінницький національний технічний університет
Вінниця, Україна
voytovych.olesya@vntu.edu.ua

Ольга Шулятицька
кафедра захисту інформації
Вінницький національний технічний університет
Вінниця, Україна
olya_olek@ukr.net

***Abstract* — Wireless sensor networks are rapidly developed and soon will occupy the dominant place among the other data collection and communication systems. The remote location of sensors and automatic operations within ones increase their vulnerability to external intrusions and attacks. Because of sensor supply necessity the authentication providing is quite challenging whereas the best protocols drain sensor capabilities. The method of sensor authentication within wireless sensor networks, which is based on the trust model, is proposed.**

Keywords — wireless sensor networks, sensor technology, authentication, trust model, energy consuming.

I. INTRODUCTION

Wireless sensor networks (WSN) increasingly penetrate the different fields of human activity. Sensor networks as part of the Internet of Things (IoT) occupy an important place in communication networks. However, research results show safety problems, those arise at the operation of these systems [1] including unsecured communication, weak authentication, lack of authorization, feasibility of used cryptographic algorithms, open libraries and updates exchange etc. Most of these problems are caused by such important feature of WSN as autonomic power supply, and those the duration of such a network performance is paramount.

WSN is a set of sensors, which are able to collect some data, transform it into electromagnetic signals, those are carried out by their broadcast, receiving signals from neighbor sensors and retransmit them on the air [2].

The wireless connection is easy enough for the intruder sensors packet interception. For example, most of the major threat is the denial of service attack, aimed to break the correct functioning of the sensor network. There are various options for anti-theft security systems, but most of them require substantial amount of resources, which is difficult to achieve at the strictly limited power supply conditions. Therefore, WSN requires new solutions for keys generation, their distribution, identification and protection [3].

II. THE WIRELESS SENSOR NETWORK ATTACKS

The problems of WSN information security can be divided into the primary and secondary issue. Primary targets are well known and cover confidentiality, integrity, availability and authentication of data. The secondary security issues cover data freshness, self-organizing, time synchronization, localization data [4].

A. Attacks on the WSN

The known attacks are shown in Fig. 1.

WSN are rapidly developing and soon will occupy a dominant place among the data collection and communication systems [3]. After analyzing the known attacks it can be concluded that to protect from the majority of attacks one needs to use authentication methods.

Classification features of sensor networks are proposed in [5]. One of the main features is power supply. WSN differ by power supply: self-powered (off-line), connected to the power supply (line) and renewable (energy that is collected from resources which are naturally replenished e.g. solar energy). This feature is very important because of WSN nature, which depends on battery or no, and thus can use more or less different processing resources.

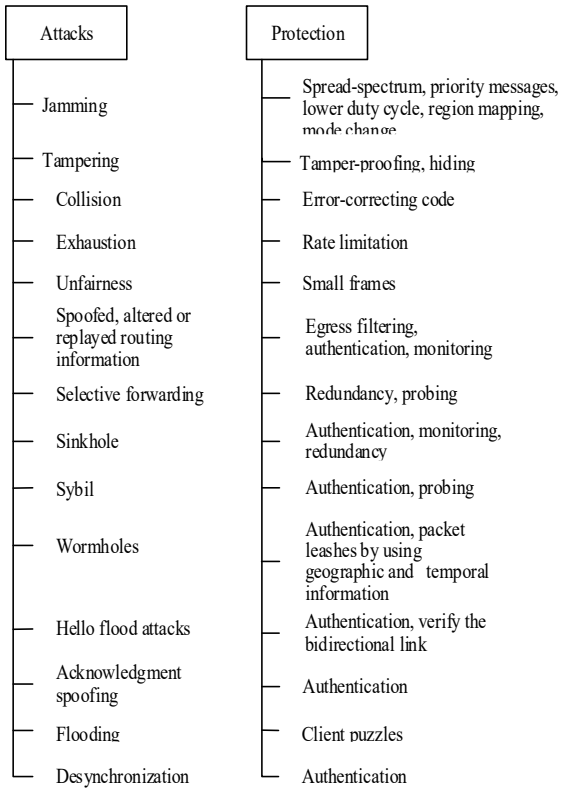


Fig. 1. Attacks on WSN and protection methods

III. THE AUTHENTICATION METHODS ANALIZING

Today there are many authentications schemes in wireless sensor networks, the main ones are:

Based on the cryptography protocols, including lightweight protocols;

Based on hierarchies;

Based on the trust center.

For instance, in paper [6], author proposes a lightweight dynamic user authentication scheme that is built upon Wong et al. There is overall handshake in the scheme. The Wong's scheme consists of four phases: registration, login, authentication and password-changing phases. The first and the last phases are performed via a secure channel. There is a centralized gateway node in Wong's scheme and the performance in authentication process might be improved by designing a decentralized gateway node.

In paper [7], author proposes a broadcast authentication scheme for wireless sensor networks that utilizes a one-time signature scheme and re-keying mechanism. This scheme exhibits individual authentication, robustness to packet loss, and low overhead in computation, communication and storage. Author improves upon the one-time signature scheme, reducing the large key storage by using Merkle hash trees in generating the key pair. Despite increasing computation and communication efficiency, it significantly decreases the storage space usage [8].

In paper [9], the author proposes an enhanced Trust Center mechanism, which improves the performance of current standards in ZigBee Network with mobility node. In Trust Center, node of the initial authentication is achieved through the Trust Center. If the node will be moved to a different subnet, Trust Center shares authentication data with other Trust Centers without a separate authentication process. In order to evaluate the performance of the proposed scheme, the comparing of the energy and memory efficiency was performed.

There is a new authentication protocol in [10]. The trust center is not situated next to the joiner sensor. The master key is pre-shared between a trust center and a joined device, and the network key is distributed to every parent router in the network by the trust center, the same way as at authentication protocol within ZigBee. In addition, the local authentication key, which is used in pre-authentication procedure, is pre-shared between a parent router and the joiner device to be used in pre-authentication procedure.

In [11], author describes communications model and trust model, considering the scenario where a mobile user, associated with its home agent, is visiting a foreign network with a foreign agent. When mobile user is out of its home network, it needs to be authenticated before being allowed to access a visited foreign network. Because mobile user is out of the coverage of its home agent, author assumes that any message between mobile user and home agent has to go through foreign agent. Author further assumes that home agent has a communication link to the foreign agent that is to serve mobile user. A secure channel can be established between home agent and foreign agent, for example, via Kerberos [12].

Fig. 2 shows the trust model, where a dashed line with arrows at both ends indicates that there is mutual trust established between the two end parties, and a dashed line with an arrow at one end only indicates a one-way trust. Following Fig. 2, mobile user cannot trust foreign agent and vice versa, when MU – mobile user, HA – home agent, FA – foreign agent, AS - authentication server.

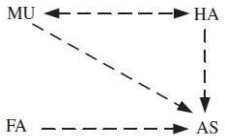


Fig. 2. Trust model accordig to [12]

Typically, there is no bilateral trust between foreign agent and home agent. Although mobile user trusts, authentication server and home agent in case of direct communication absence in foreign network, the proper authentication is to be used. All other trusting pairs connected via dashed lines in Fig. 2 are straightforward to follow.

Thus, the lack of these authentication methods has three problems:

Weak methods are easy to break, as a result of the traffic analyses, data modifying, the sensor substitution etc.

Cryptography methods need too large energy consumption, which is bad from the practical point of view. Using

complex cryptographic authentication scheme consumes a lot of energy for data transmission and processing as a result sensor battery needs charging or replacing, so this issue is relevant for the decision.

Additional trust centers reduce flexibility of WSN and can be victim of attackers.

IV. Proposed Authentication Methods Based on Trust Model

The authentication method based on trust model with a poll of at least three neighbor nodes, which solves the above-mentioned problems, is proposed. At the energy-efficient routing protocols [13-18], the sensor asks neighbors about their readiness for data transferring (level of battery, distance from each other etc) to construct the optimal data transfer route. That is why it is proposed to request additional options (level of trust) during this data exchange. Fig. 3 schematically shows the idea of authentication process performance (ID – Identifier, K – Key, RD - Level of trust (1 or 0), SN - Node (sensor), Bat – Power, Route - Remoteness).

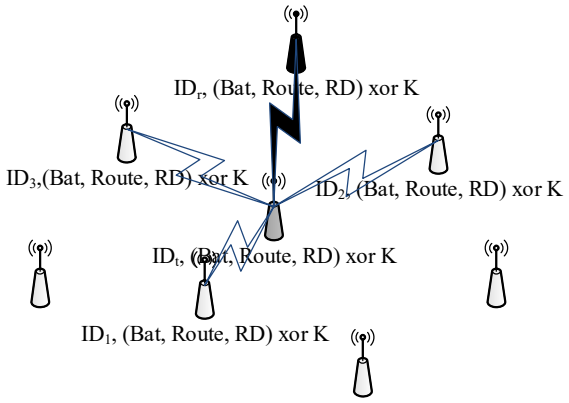


Fig. 3. Transfer data scheme of the WSN (gray - transmitter, black - receiver)

Let's assume there is a sensor network and each sensor (node) has a small database, that contains minimum of three neighbor nodes IDs to further authenticate and determine the level of trust in other nodes. This authentication scheme is a simple request-response protocol, which involves such components: transmitter node (SN_t), receiver node (SN_r) and respond nodes (SN_i). The transmitter node (SN_t), which is to send data firstly requests at least three trusted sensor nodes whether or not the receiver node (SN_r) is trusted by them. Moreover, if the nodes majority answers positively, the receiver node would become trusted. Thus, transmitter node forms an overall picture either transmits data or polls other nodes and searches another route.

1) *Registration phase*

At the registration phase more complex authentication schemes could be used bearing in mind above statements for nodes in WSN. They provide greater stability because of the more complicated authentication usage. Neighbor nodes participating in the scheme and authentication need only to perform authentication three times and new node would be considered as trusted one.

Once the node successfully completes the registration phase for three nearest nodes, it becomes trusted. Since such authentication requires a small number of iterations (only for neighbor nodes) this would save energy supply. As a key for further sharing can be used mentioned battery level. The battery charge value can be used to avoid pseudo-random numbers generation. The value of the battery charge is a true random number after trend elimination (e.g. Bat = 78,854, then ID is 854).

2) *Authentication phase*

Case 1: neighbor sensors are trusted. Transmitter node (SN_t) polls the neighbor sensors about their resources level (Battery) and the routing data (Route); as a result the optimal energy-efficient route is received. As the polled neighbor sensors (SN_1, SN_2, SN_3) are trusted and have the shared key (K) (received at the registration phase) the information is to be transmitted.

Case 2: new neighbor sensor is appeared. When the new sensor (SNr) tries to take part at the communication, it must pass a strong authentication during registration phase to receive a shared key.

Step 1: transmitter node (SN_t) forms the request, which contains the following parameters: identifier ID_t, key (K) (the shared key can be formed from energy values, without trend, and therefore does not require additional resources for its generation), energy level (Battery) and routing data (Route), and requests for the same settings from its neighbor nodes. The request includes the trust level (RD) too. The transmitter node (SN_t) wants to know whether or not receiver node (SN_r) is trusted. Moreover, all data transmitted and accepted to be encrypted using XOR on key (K) that obtained during registration phase.

Step 2: at this step transmitter gets the response from the nodes SN1, SN2, SN3, which contains IDs, energy levels (Battery), routing data (Route) and trust level of the receiver node SN_r and the most importantly his personal ID (the battery without trend obtained at the registration phase).

After these steps transmitter node (SN_t) forms the decision concerning data transfer to receiver node (SN_r). The transmitter node determines the optimal route based on the data, that is received from neighbor nodes. Using the level of supply and optimal route distance the optimal route is formed, and using level of trust transmitter node makes the decision concerning data transferring.

The proposed WSN was simulated using environment Atarraya. Modeling was performed by Atarraya emulator [19] - the simulator control events that could be used for teaching and researching of management protocols and wireless network topology. It is designed to be run on Windows platform. The simulator consists of two topologies: TC - Topology Constructions; TM - Topology Maintenance. Two protocols are selected: DGETRec - each time, that node reaches a critical energy threshold, the topology maintenance algorithm terminates the previous reduced topology and invokes the topology construction algorithm to create a new one [20]; DLEDSR - each time, that node reaches a critical

energy threshold, the local topology maintenance algorithm based on the DSR protocol repairs the topology [20].

The experiment showed (fig.4) the implementation of different energy efficient protocols, but with identical parameters such as the number of nodes (100 nodes), the same amount of transmitted data and the same weight of this information etc. Comparison of results is rather impressive. In the first case of EECDS protocol using WSN was almost not working, nodes lost their charge and couldn't transmit data, in other case of Energy Local Patching DSR protocol using network was fully functioning when nodes have high battery power they were able continue their work even greater amount of time, unlike the first case.

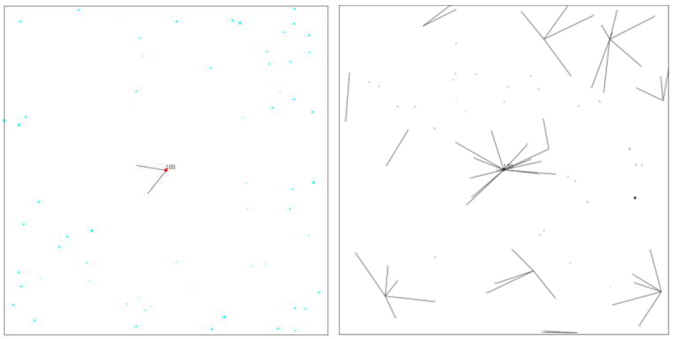


Fig. 4. WSN modeling (1. TC – EECDS, TM – DGETRec time work network = 3000012.345; 2. TC – EECDS, TM – Energy Local Patching DSR time work network = 3000000.3)

So the extension of power supply for each unit was gained, and thus the duration of the service provided by the network was increased for quite much time. It would provide decreasing of network maintenance, consequently the network would be more economical.

V. CONCLUSION

Since WSN found the widespread usage at the Internet of Things, which development is very fast-paced, there are two major problems: security of data transmitted and fast unit (with battery) power reduction. This article was aimed at addressing these two issues. The primary and secondary security issues as well as the real attacks on wireless sensor networks were analyzed. Based on main attacks and protection mechanism the problem of authentication in wireless sensor network is allocated. The existing authentication schemes often drain very much battery energy and cannot be used widely.

The method of authentication in WSN, which provides a stable authentication using the minimum amount of resources, was proposed. The authentication method is based on the trust model, that considers usage of a complex authentication only at the registration phase and provides network protection by doing this. Using neighbor trust model for simplifying authentication during other communications provide power consuming decreasing of the WSN. The experiment proved that the usage of protocols increases efficient supply network lifetime, therefore reducing network maintenance costs.

REFERENCES

[1] Operating problems WSN: http://blog.ioactive.com /2017/02/hacking-robots-before-skynet.html

[2] S. Yinbiao, P. Lanctot, F. Jianbin, “Internet of things: wireless sensor networks,” White Paper, International Electrotechnical Commission, 2014.

[3] A. Roslyakov, S. Vanyashin , A. Scallops, M. Samsonov, “Internet of Things,” Samara, 2014 (in Russion).

[4] G. J. Fan, S. Y. Jin, “Coverage problem in wireless sensor network: A survey,” Journal of Networks, vol. 5, №. 9, pp. 1033-1040, 2010.

[5] O.Voitovych, O.Shulyatitska, V.Malyushytskyy, Simulation and security of sensor networks: Inżynier XXI wieku projectujemy przyszlosc, monografia [pod red: Jacek Rysiński] - Bielsko-Biała: Wydawnictwo Naukowe Akademii Techniczno-Humanistycznej w Bielsku-Białej, 2016 – pp. 367-373. ISBN: 978-83-65182-51-7.

[6] H. R. Tseng, R. H. Jan, W. Yang, “An improved dynamic user authentication scheme for wireless sensor networks,” Global Telecommunications Conference, pp. 986-990, 2007.

[7] S. M. Chang, et al., “An efficient broadcast authentication scheme in wireless sensor networks,” Proceedings of the 2006 ACM Symposium on Information, computer and communications security, ACM, pp. 311-320, 2006.

[8] Volodymyr Luzhetsky, Yurii Baryshev. Methods of Generic Attacks Infeasibility Increasing for Hash Functions // The 7th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS’2013), September 12-14, 2013 Berlin, Germany. Available: idaacs.net/2013/wp-content/uploads/abstracts/i13-071_8d0bb875.rtf

[9] K. Lee, et al., “An enhanced Trust Center based authentication in ZigBee networks,” International Conference on Information Security and Assurance, Springer Berlin Heidelberg, pp. 471-484, 2009.

[10] S. H. Lee, J. H. Kim, “Design of authentication protocol for LR-WPAN using pre-authentication mechanism,” Consumer Communications and Networking Conference, pp. 1-5, 2009.

[11] C. Chen, et al., “Lightweight and provably secure user authentication with anonymity for the global mobility network,” International Journal of Communication Systems, vol. 24, №. 3, pp. 347-362, 2011.

[12] B. C. Neuman, Ts'o T. Kerberos, “An authentication service for computer networks,” IEEE Communications magazine, vol. 32, №. 9, pp. 33-38, 1994.

[13] V. Karthikeyan, A. Vinod, P. Jeyakumar, “An Energy Efficient Neighbour Node Discovery Method for Wireless Sensor Networks,” 2014.

[14] Y. Chung, “An energy-efficient unicast routing protocol for wireless sensor Networks,” Tech. Int. J. Comput. Sci. Emerg. Tech. – 2011. – vol. 2. – pp. 60-64.

[15] R. Tiwari, A. Saxena, “A review on energy efficient routing in wireless sensor networks,” Journal of engineering trends and technology, 2015. Vol 19. pp. 29-34.

[16] A. G. A. Elrahim, et al., “An energy aware WSN geographic routing protocol,” Universal Journal of Computer Science and Engineering Technology, 2010. – vol. 1. pp. 105-111.

[17] G. Krishna Priya, Dr. G. Prakash Babu, “EE-DSR: Energy Efficient Dynamic Source Routing in Wireless Ad Hoc Networks Using Residual Energy,” International Journal of Advanced Research in Computer and Communication Engineering, Vol. 4, 2015 pp. 134-138.

[18] R. Alasem, A. Reda, M. Mansour, “Location based energy-efficient reliable routing protocol for wireless sensor networks,” Recent Researches in Communications, Automation, Signal processing, Nanotechnology, Astronomy and Nuclear Physics, WSEAS Press, Cambridge, UK, 2011 pp. 180-185.

[19] Simulation Atarraya. Available: http://www.csee.usf.edu/~mlabrador/Atarraya/

[20] P. M. Wightman Rojas, Topology control in wireless sensor networks. – 2010.

Методи Байт-Орієнтованого Хешування Даних Низькоресурсної Криптографії

Володимир Лужецький
кафедра захисту інформації
Вінницький національний технічний університет
Вінниця, Україна
lva.zi2002@gmail.com

Дмитро Кисюк
кафедра обчислювальної техніки
Вінницький національний технічний університет
Вінниця, Україна
kneimad@gmail.com

Світлана Слободян
кафедра обчислювальної техніки
Вінницький національний технічний університет
Вінниця, Україна
feride.fe@list.ru

Олександр Ювковецький
кафедра обчислювальної техніки
Вінницький національний технічний університет
Вінниця, Україна
alex.yuvkovetskyi@gmail.com

Methods of byte-oriented hashing of lightweight cryptography

Volodymyr Luzhetskyi
Department of Information Protection
Vinnytsia National Technical University
Vinnytsia, Ukraine
lva.zi2002@gmail.com

Dmytro Kysiuk
Department of Computer Science
Vinnytsia National Technical University
Vinnytsia, Ukraine
kneimad@gmail.com

Svitlana Slobodyan
Department of Information Protection
Vinnytsia National Technical University
Vinnytsia, Ukraine
feride.fe@list.ru

Olexander Yuvkovetskyi
Department of Information Protection
Vinnytsia National Technical University
Vinnytsia, Ukraine
alex.yuvkovetskyi@gmail.com

***Анотація*—Запропоновано принципово новий байт-орієнтований підхід до побудови хеш-функцій для низькоресурсних систем з використанням характеристичних ознак даних. Наведено схеми обчислень і структурні схеми спеціалізованих процесорів для хешування та оцінки їх апаратної складності.**

***Abstract*—A fundamentally new byte-oriented approach to constructing hash functions for lightweight systems using the characteristic features of data is proposed. Schemes of calculations and structural schemes of specialized processors for hashing and estimation of their hardware complexity were presented.**

Ключові слова—хеш-функція, байт-орієнтоване хешування, низькоресурсне хешування, спеціалізований процесор, апаратна складність.

Keywords—hash-function, byte-oriented hashing, lightweight hashing, specialized processor, hardware complexity .

I. Вступ

Основною особливістю сучасних ІТ- та Інтернет-технологій є значне поширення різноманітних інтелектуальних пристроїв, зокрема гаджетів, що мають доступ до Інтернет. Стрімкий розвиток таких технологій формує ряд актуальних задач пов'язаних з їх інформаційною безпекою. Також, разом з широким переліком традиційних інтернет-пристроїв, таких як персональні комп'ютери та ноутбуки, смартфони, розумні годинники і браслети, почали з'являтися побутові пристрої, термінали, безпілотні літаки, транспортні зчитувачі і навіть автомобілі з автопілотом та навіть давачі, що мають доступ до певної мережі (зазвичай Інтернет) [1,2].

Такі технології отримали назву Internet of Things (IoT) або Інтернет речей, що є бездротовою мережею, яка самоконфігурується, з різноманітних засобів, до складу яких входять давачі та сенсори. Також, останнім часом, набули поширення модулі, що використовують RFID-мітки (Radio Frequency Identification) або NFC технологію (Near Field Communication). Крім того, їх регулярно використовують численні програми, пов'язані з обробкою персональних даних, біометричних даних, відомостей фінансового та військового характеру і т. п. [1,2].

Саме тому особливо актуальною постає задача ефективної реалізації алгоритмів захисту інформації, левову частку яких складають криптографічні методи захисту інформації. Беручи до уваги особливості побудови та функціонування таких пристроїв, найкращим рішенням для них є вид криптографії, що носить назву «низькоресурсна криптографія» (LightWeight Cryptography, LWC) [3-7]. Вона потрібна для тих сфер застосування, де є жорсткі вимоги чи обмеження на ціну, габарити, ресурси системи (такі як пам'ять, обчислювальна потужність, джерело живлення і т. п.) при проектуванні та промисловому виготовленні подібних пристроїв. Більшість вимог, які висуваються до алгоритмів, що застосовуються у низькоресурсних засобах, обумовленні міжнародним стандартом ISO/IEC FDIS 29192 – Information technology – Security techniques – Lightweight cryptography.

Одним з декількох основних напрямків розвитку низькоресурсної криптографії є більш ефективна реалізація вже відомих алгоритмів шифрування, у більшості випадків з незначною їх модифікацією. При цьому передбачається, що рівень безпеки такого шифрування знизиться до рівня, який є достатнім для більшості його застосувань.

Усе це обумовлює актуальність проведення досліджень у напрямку створення засобів, які задовольняють LWC. Типовими обмеженнями для низькоресурсної криптографії є складність апаратної реалізації, час, що витрачається на виконання програми, споживана енергія.

Серед криптографічних перетворень особливе місце займають ключові хеш-функції, відомі як коди перевірки справжності повідомлення (Message Authentication Code - MAC) або імітовставка, та безключові хеш-функції або коди виявлення модифікації інформації (Modification Detection Code - MDC).

Відомі стандарти хеш-функцій для своєї реалізації потребують понад 10000 умовних логічних елементів (GE – gate equivalent), тоді як гранична складність не повинна перевищувати 2000 GE [5]. Певною мірою вимоги низькоресурсної криптографії забезпечують хеш-функції сімейства QUARK [5], однак продовжуються пошуки підходів до хешування, які б забезпечили ще кращі показники апаратної складності та часу хешування.

II. Методи хешування

Авторами пропонується новий підхід до побудови LW хеш-функцій з байт-орієнтованою обробкою даних, суть якого полягає в такому. Вхідне повідомлення M розглядається як послідовність байтів $M=\{ m_1, m_2, ..., m_L\}$. Хеш-функції – це функція, яка певним чином пов'язує ASCII–коди байтів m_l та номери позицій l (l = 1, 2, …) цих байтів у повідомленні [8-11].

Передбачається хешування даних довільної довжини та отримання хеш-значення різного розміру (128, 192, 256 біт).

Метод 1.

Хеш-значення формується з ASCII–кодів байтів з урахуванням номерів позицій l. Таку хеш-функцію будемо скорочено позначати SNP.

Виходячи із заданої довжини хеш-значення l_h визначається кількість байт, що будуть використовуватися для обчислень:

$$g = l_h / 8 . \qquad (1)$$

Наприклад, при довжині хеш-значення 128 біт, це хеш-значення представляється у вигляді набору з 16 елементів, кожен з яких є байтом (128=8×16), відповідно для довжини 192 біт набір складається з 24 елементів (192=8×24), а для 256 – з 32 елементів (256=8×32).

При цьому, номер позиції байта, використовуваний у подальшому, визначається за формулою:

$$q = l \bmod 2^g .$$

Початкове хеш-значення h_0 є сукупністю псевдовипадкових чисел і має вигляд:

$$h_0 = \{h_{0,0}, h_{0,1}, ..., h_{0,(g-1)}\} .$$

Номер позиції q байта у повідомленні розглядається як двійковий код:

$$q = \sum_{i=0}^{g-1} a_i \cdot 2^i .$$

Проміжні хеш-значення $h_l = \{h_{l,0}, h_{l,1}, ..., h_{l,(g-1)}\}$ обчислюється на основі попереднього хеш-значення h_{l-1}, двійкового представлення номера позиції q та ASCII–кода байта n_l:

$$h_{l,j} = h_{(l-1),(j-1)} \oplus (n_l \cdot a_{j-1}),$$
$$h_{l,0} = h_{l,(g-1)} \oplus (n_l \cdot a_{g-1}),$$

де $j = 1,2,...,(g-1)$.

Схему обчислень хеш-значеннь за даним методом наведено на рис.1.

Для хешування одного байту даних потрібно виконати g операцій додавання за модулем 2 двох кодів байтів і $(g+1)$ операцій запису коду в регістр.

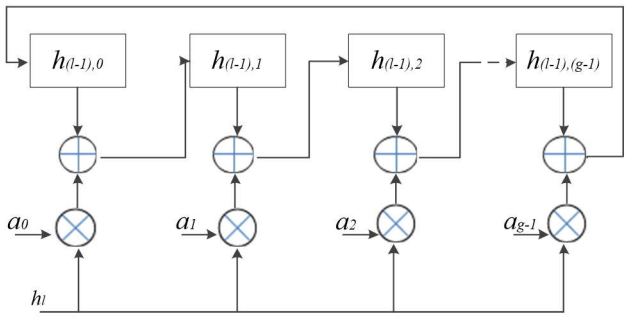


Рис. 1. Схема обчислень хеш-значеннь за методом 1

Метод 2.

Хеш-значення формується з псевдовипадкових чисел, що відповідають номерам позицій l байтів, з урахуванням ASCII–кодів байтів. Таку хеш-функцію будемо скорочено позначати PRNS.

Розрядність двійкового коду псевдовипадкових чисел визначається за формулою (1).

Початкове, проміжні та остаточне хеш-значення складаються з 8 елементів довжиною g біт кожен.

Нехай початкове хеш-значення має вигляд:

$$h_0 = \{h_{0,0}, h_{0,1}, ..., h_{0,7}\}$$

Тут кожна складова може приймати будь-які значення. У разі, коли ці значення відомі, маємо справу з безключовою хеш-функцією, а коли початкове хеш-значення є секретним, то це буде ключова хеш-функція.

Проміжні хеш-значення $h_l = \{h_{l,0}, h_{l,1}, ..., h_{l,7}\}$ обчислюються на основі попереднього хеш-значення, ASCII-коду байту $\{a_{l.0}, a_{l.1} ... a_{l.7}\}$ і чергового псевдовипадкового числа s_l з циклічним зсувом складових хеш-значення:

$$h_{l,j} = h_{(l-1),(j-1)} \oplus (s_l \cdot a_{j-1}),$$
$$h_{l,0} = h_{l,7} \oplus (s_l \cdot a_7),$$

де $j = 1,2,\ldots 7$.

Для генерування послідовності псевдовипадкових чисел пропонується використовувати регістр зсуву з лінійним зворотнім зв'язком (LFSR). Як псевдовипадкове число розглядається стан цього регістру.

Вибір генератора на основі LFSR обґрунтовується такими його властивостями [12].

1. Псевдовипадкові числа належать інтервалу $[1, 2^n - 1]$.
2. В межах періоду генератора числа не повторюються.
3. Регістри зсуву в конфігурації Галуа забезпечують паралельну обробку за один такт кожного з двійкових розрядів, зі швидкістю, що порівняна зі швидкістю перемикання транзисторів типових мікросхем.
4. Складність апаратної реалізації може дорівнювати $(4 \cdot n + 6)$ GE.

Схему обчислень хеш-значеннь за даним методом наведено на рис.2.

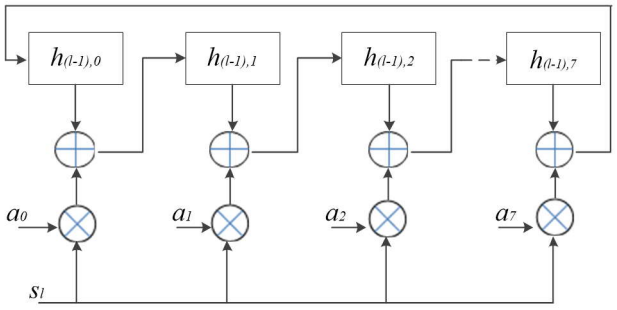


Рис. 2. Схема обчислень хеш-значеннь за методом 2

Для хешування одного байту даних необхідно виконати 8 операцій додавання за модулем 2 двох g-розрядних кодів, 9 операцій запису коду в регістр і одну операцію для формування псевдовипадкового числа.

III. Спеціалізовані процесори для хешування даних

Для отримання хеш-коду використовується спеціалізований процесор МН (message hashing), який певним чином підключається до центрального процесора і обмінюється з ним потоками даних, виконуючи певні функції. Передбачається, що інформація, для якої формується хеш-значення знаходиться в оперативній пам'яті, тому зчитування цих даних покладається на центральний процесор, а обчислення хеш-значення на МН-процесор.

Структурну схему МН-процесора, що реалізує хешування за методом 1, наведено на рис. 3. До складу пристрою входять g однобайтних регістрів для формування хеш-значення, лічильник розрядністю g, для формування коду номера позиції q поточного байту, регістр зсуву для зберігання коду числа q, блок з 8 суматорів за модулем 2 і блок з 8 логічних елементів І.

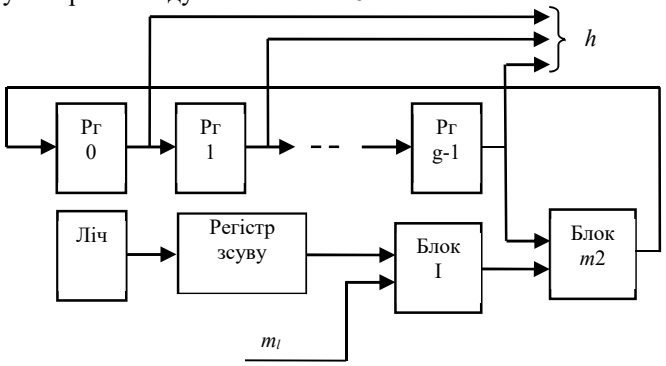


Рис. 3. Узагальнена структурна схема МН-процесора, що реалізує хеш-функцію SNP

Складність пристрою в умовних одиницях GE, при такій розрядності його складових, обчислюється за формулою:

$$S = 60 \cdot g + 24$$

Структурну схему МН-процесора, що реалізує хешування за методом 2, наведено на рис. 4. До складу пристрою входять вісім g-розрядних регістрів Рг0…Рг7, в

яких зберігаються проміжні хеш-значення, *g*-розрядний регістр зсуву з лінійним зворотнім зв'язком РЗЗЛЗЗ, який забезпечує формування послідовності псевдовипадкових чисел, 8-розрядний регістр зсуву, який забезпечує перетворення паралельного 8-розрядного ASCII–коду у послідовність бітів. Блок I складається із *g* логічних елементів I. Блок m2 складається із *g* суматорів за модулем 2.

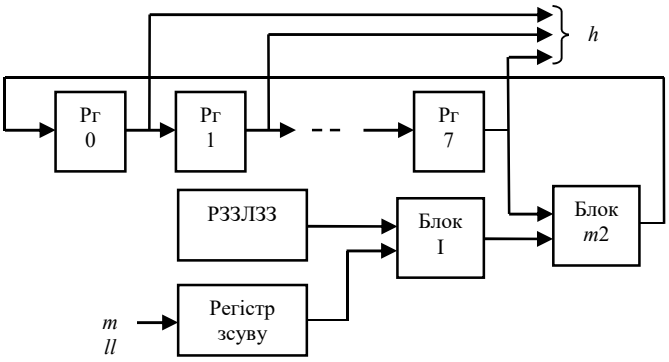


Рис. 4. Узагальнена структурна схема МН-процесора, що реалізує хеш-функцію PRNS

Складність такого спеціалізованого процесора в умовних одиницях GE обчислюється за формулою:

$$S = 57 \cdot g + 48 .$$

У таб. 1 наведено характеристики хеш-функцій сімейства QUARK та запропонованих авторами хеш-функції PRNS та SNP. Тут С/байт – позначає кількість операцій для обчислення хеш-значення для одного байту даних.

ТАБЛИЦЯ I. ХАРАКТЕРИСТИКИ ХЕШ-ФУНКЦІЙ

Хеш-функція	**Довжина хеш-значення, *біт***	**Складність, *GE***	**С/байт**
D-QUARK	160	1702	547
T-QUARK	224	2296	33
U-QUARK	128	1379	33
PRNS	128	960	18
PRNS	192	1416	18
PRNS	256	1872	18
SNP	128	984	33
SNP	192	1464	49
SNP	256	1944	65

IV. Висновки

Запропоновані методи хешування забезпечують апаратну реалізацію, складність якої не перевищує 2000 GE, що відповідає вимогам міжнародного стандарту ISO/IEC FDIS 29192, і ця складність у 1,44 рази менша за складність реалізації найкращих відомих алгоритмів хешування.

Для запропонованої хеш-функції PRNS кількість операцій, що потрібна для хешування одного байту даних, не залежить від довжини хеш-значення і в 1,8 рази менша порівняно з відомими хеш-функціями сімейства QUARK.

Результати дослідження запропонованих методів хешування з використанням тестів NIST показали, що вони забезпечують виконання вимог щодо криптографічної стійкості хеш-функції.

Література References

[1] Bogdanov A., Leander G., Paar C., Posch A., Robshaw M., Seurin Y. Hash Functions and RFID Tags: Mind the Gap -10. In ter na tio nal Work shop on Cryp to gra phic Hard ware and Em bed ded Sys tems - CHES 2008, Wa shington, USA. Au gust 10 - 13, 2008. Pages 283-299.

[2] Yoshida H., Watanabe D., Okeya K., Kitahara J., Wu J., Kucuk O., and Preneel B.. MAME: A Compression Function With Reduced Hardware Requirements. In P. Paillier and I. Verbauwhede, editors, Cryptographic Hardware and Embedded Systems — CHES 2007, volume 4727 of Lecture Notes in Computer Science, pages 148–165. Springer-Verlag, 2007.

[3] Manifavas C., Hatzivasilis G., Fysarakis K., Rantos K. Lightweight Cryptography for Embedded Systems - A Comparative Analysis, SETOP'2013.

[4] Panasenko S., Smagin S. Lightweight Cryptography: Underlying Principles and Approaches. International Journal of Computer Theory and Engineering, Vol. 3, No. 4, August 2011, pp. 516-520.

[5] Жуков А. Е. Легковесная криптография / А. Жуков // – Москва: Вопросы кибербезопасности №1(9), 2015.

[6] Bogdanov A., Knudsen L., Leander G., Paar C., Poschmann A., Robshaw M., Seurin Y., Vikkelsoe C. Present - An Ultra-Lightweight Block Cipher. In Proceedings of Workshop on Cryptographic Hardware and Embedded Systems — CHES 2007, Lecture Notes in Computer Science v. 4727, pp. 450-466, 2007.

[7] Preneel B. Perspectives on Lightweight Cryptography," Inscript 2010, Shanghai, China, 20-24 October 2010.

[8] Лужецький В. А. Новий підхід до побудови криптографічних хеш-функцій / В. А. Лужецький, Д. В. Кисюк // «Інформаційні технології та комп'ютерна інженерія»; матеріали статей п'ятої міжнародної науково-практичної конференції, м. Івано-Франківськ, 27-29 травня 2015 року. – Івано-Франківськ: Супрун В. П., 2015 р. – с. 206-208.

[9] Лужецький В. А. Метод хешування даних на основі їх характеристичних ознак / В. А. Лужецький, Д. В. Кисюк // II міжнародна науково-практична конференція (Закарпатська область, Міжгірський район, село Верхнє Студене, туристичний комплекс «Едельвейс». 24-27 лютого 2016 р.). – К.: Видавництво Європейського університету, 2016. – 97-100 с.

[10] Лужецький В. А. Метод байт-орієнтованого хешування / В. А. Лужецький, Д. В. Кисюк, Л. А. Савицька // V Міжнародна науково-практична конференція «Методи та засоби кодування, захисту й ущільнення інформації». 19-21 квітня 2016 року, Вінниця – Немирів, ВНТУ. – 37-39 с.

[11] Лужецький В. А. Узагальнений метод хешування байтової форми представлення інформації / В. А. Лужецький, Д. В. Кисюк // Тези доповідей Четвертої Міжнародної науково-практичної конференції «Інформаційні технології та комп'ютерна інженерія» м. Вінниця, 28-30 травня 2014 року. – Вінниця: ВНТУ, 2014., – ⁻275 с.

[12] Watanabe D., Okamoto K., Kaneko T. A Hardware-Oriented Light Weight Pseudo-Random Number Generator Enocoro-128v2. The 2010 Symposium on Cryptography and Information Security, SCIS 2010, 3D1-3, 2010.

Концепція Сучасної Архітектури Менеджменту Процесів Інформаційної Безпеки Soapa

Копиченко Іван
начальник відділу технічного захисту інформації
КП «Обласний інформаційно-аналітичній центр»
Одесса, Україна
ikopychenko@odessa.gov.ua

The Concept of Modern Information Security Processes Management Architecture Soapa

Kopychenko Ivan
head of Technical Information protection department
Municipal Enterprise. Regional Information and Analytical Center
Odessa, Ukraine
ikopychenko@odessa.gov.ua

***Анотація*—Безперервне збільшення інформаційних і телекомунікаційних систем мережевої інфраструктури компаній, так саме як і значне збільшення обсягів телеметричної інформації, що генерується такими системами, вимагає впровадження більш розвинених моделей управління подіями інформаційної безпеки. Саме через це фахівцями ESG (Enterprise strategy group) було запропоновано нову програмну архітектуру SOAPA.**

***Abstract*—The continuous increase in information and telecommunications systems of the company's network infrastructure, as well as a significant increase in the volumes of telemetric information generated by such systems, requires the introduction of more advanced models of information security event management. That is why the specialists of the Enterprise strategy group proposed a new software architecture called SOAPA.**

Ключові слова—безпека; SIEM; СУІБ; кібербезпека; архітектура; аналітика; управління; мережа; захист; автоматизація

Keywords—security; SIEM; ISMS; cybersecurity; architecture; analytics; management; network; defense; automation

I. Вступ

У сучасних реаліях забезпечення інформаційної безпеки перестало сприйматися як технічна задача, а стала серйозною бізнес проблемою для багатьох підприємств стурбованих своєю репутацією і місцем на ринку. Для забезпечення безпеки даних, які обробляються за допомогою інформаційно-телекомунікаційних систем підприємства, недостатньо лише побудувати надійну систему захисту інформації, а й вкрай важливо забезпечити її постійну підтримку, розвиток і управління протягом усього життєвого циклу.

Для вирішення саме такого завдання фахівцями ESG була запропонована концепція нової архітектури менеджменту процесів інформаційної безпеки SOAPA (security operations and analytics platform architecture). Дана архітектура стала неминучим кроком у розвитку SIEM (Security information and event management) систем.

Фахівці із інформаційної безпеки та інсайдери стверджують, що 2017 рік стане важким роком, наповненим атаками на державні ресурси, вірусною активністю та витоками даних. В першу чергу це пов’язано із тенденціями розвитку інформаційних технологій, які в значній мірі ускладнюють забезпечення інформаційної безпеки:

Хмарні сервіси. Все більше систем та інформаційних навантажень переносять в приватні та публічні хмарні сервіси.

Інтернет речей. Згідно із прогнозами, до 2020 року очікується більш ніж 20 мільярдів підключених до мережі Інтернет пристроїв. Вже сьогодні сфери діяльності пов’язані із енергетикою, охороною здоров’я та промисловістю активно впроваджують IoT (internet of things) пристрої.

Зростання мережі. Фізичні мережі і мережеві магістралі розширюються від 10Гб до 40 / 100Gb, продовжується перехід від IPv4 до IPv6. Все це призводить до збільшення об’ємів трафіку, сеансів,

пакетів, потоків та протоколів, за якими необхідно стежити.

Перенесення робочих процесів в електронні системи. Велика кількість державних та приватних установ автоматизують робочі процеси та переносять роботу в електронні системи.

II. Призначення архітектури

В процесі управлінні інформаційною безпекою дуже важливо не губитися в великих обсягах інформації, що генерується системами управління інформаційною безпекою (СУІБ). Навіть якщо всередині підприємства існує центр управління інформаційною безпекою (security operations center - SOC), фахівець часто не здатний впоратися з обробкою всіх даних, що надходять.

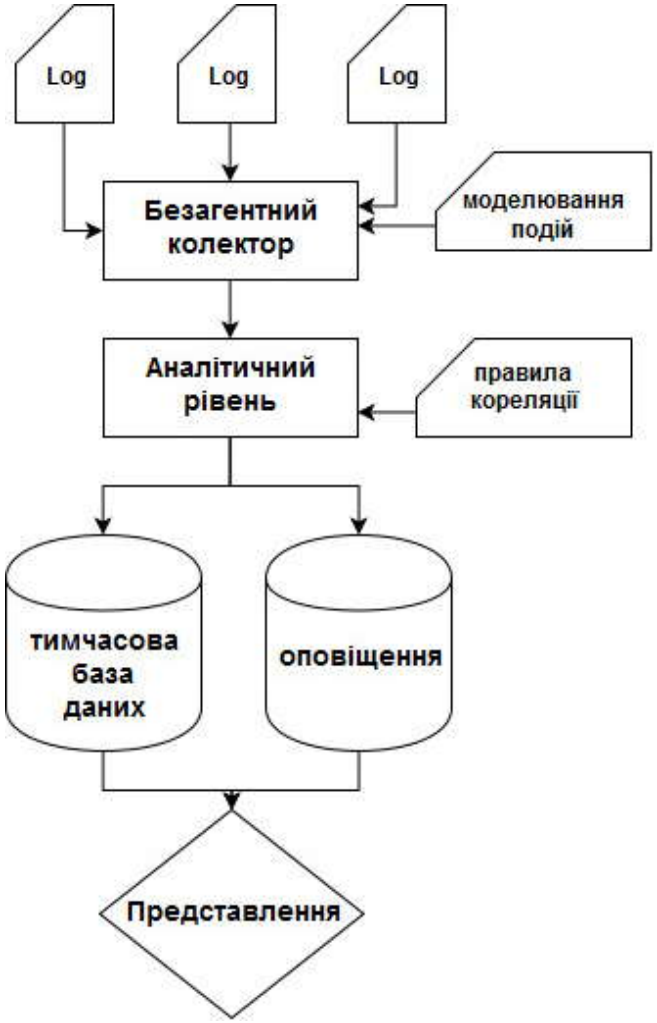


Рис. 1. Типовий схема взаємодії вузлів SIEM системи

Рішенням такої проблеми для підприємства може стати впровадження системи, побудованої на базі архітектури SOAPA, яка дозволяє підвищити ефективність роботи фахівців з інформаційної безпеки, в умовах значного нарощування обсягів внутрішньої інформаційно-телекомунікаційної структури підприємства.

На сьогодні основною категорією споживачів таких систем є підприємства, що працюють в банківській сфері. Адже саме їм необхідно регулярно проводити зовнішні аудиторські перевірки на відповідність. У банківській сфері здійснюється обробка великої кількості чутливої інформації, тому в разі виникнення інцидентів стає критично важливим володіння інформацією про те, хто/коли/звідки допустив витік, чи було це зловмисною дію або випадковою, а також наявність та характеристика супутніх факторів.

Так само потреба в SOAPA виникає у приватних підприємств з великим географічним розподілом. У таких підприємств часто виникають проблеми усереднення даних. Замість того щоб дивитися на проблему комплексно, події розглядаються в потоці загальної інформації, і часто залишаються без належної уваги через низький пріоритету окремих показників.

Нарешті, SOAPA може стати наступним кроком розвитку інформаційної безпеки для підприємств, у яких вже побудована повноцінна система управління подіями інформаційної безпеки, але існує потреба в поліпшенні процесу запобігання та захисту від атак.

Основна перевага SOAPA перед іншими архітектурами полягає в тому, що вона дозволяє не тільки аналізувати події, створювати красиві звіти і управляти ризиками, а й автоматизовано управляти ролями користувачів системи, розмежовувати права доступу, відстежувати аномалії поведінки користувача та трафіку всередині мережі, аналізувати вхідні дані на уразливості та, що найважливіше - об'єднати всі розрізнені системи інформаційної безпеки в один комплексний механізм.

Дана архітектура дозволяє підтримувати взаємодію підсистем за допомогою великого числа протоколів і інтерфейсів:

Syslog

SNMPv2 SNMPv3

HTTP, HTTPS

SQL

WMI

FTP SFTP

SSH

Rsync

SAMBA

Однак, слід також зазначити, що впровадження подібних систем без попереднього аналізу та підготовчих робіт може привести не тільки до нераціонального використання виділених ресурсів і надлишковості в роботі системи, але і до отримання абсолютно зворотного ефекту, очікуваного від системи, а саме значного збільшення ризиків пов'язаних з інформаційною безпекою підприємства.

Перш ніж впроваджувати системи менеджменту інформаційної безпеки, керівництву, спільно з фахівцями з інформаційної безпеки, в обов'язковому порядку необхідно провести ряд підготовчих робіт, а саме:

визначення області дії СУІБ

збір вихідних даних про бізнес-процеси, структурні підрозділи, інфраструктуру, методи і засоби забезпечення інформаційної безпеки.

аналіз діючої організаційно-розпорядчої документації, що регламентує питання забезпечення інформаційної безпеки

проведення оцінки ризиків:

інвентаризація та класифікація активів;

формування карти загроз;

розробка процесів управління інформаційною безпекою;

розробка процесів забезпечення інформаційної безпеки;

розробка комплекту організаційно-розпорядчої документації, що регламентує питання забезпечення інформаційної безпеки;

впровадження процедур і документації СУІБ:

впровадження процесів управління інформаційною безпекою;

впровадження процесів забезпечення інформаційної безпеки;

впровадження системи обробки подій та журналів.

III. Елементи архітектури

Дуже часто, у випадках коли у підприємства виникає питання автоматизації роботи своєї системи управління інформаційною безпекою, стає очевидною недостатність функціоналу звичних рішень. SOAPA розвивається як динамічна архітектура, і це значить що в майбутньому в неї можуть бути інтегровані нові системи. На сьогодні до складу архітектури SOAPA входять:

інструменти виявлення/реагування на кінцевих вузлах (EDR)—часто аналітику інформаційної безпеки підприємства необхідно поглибитися в оповіщення про тривогу спостерігаючи за користувачем системи та досліджуючи його поведінку

платформи реагування на інциденти (IRPs)—крім збору, обробки та аналізу даних, фахівцю з безпеки важливо якомога швидше сповіщати про проблеми безпеки та виправляти їх

системи аналітики мережевої безпеки—доповнює системи виявлення та реагування на загрози аналізом на рівні аналізу потоку пакетів

системи аналітики поведінки користувача/алгоритми машинного навчання (UBA|MLA)—автоматизує обробку великих обсягів інформації

сканери уразливостей та засоби керування безпекою—допомагає спеціалісту із безпеки встановлювати пріоритети аналізу та управління уразливостями

пісочниці для запобігання шкідливому програмному забезпеченню (anti-malware sandboxes)— дозволяє проводити ручний або автоматизований аналіз та зворотній інжиніринг можливих заражених файлів, а також пошук уразливостей нульового дня (0-day)

системи розвідки загроз— для організацій дуже важливо проводити порівняння внутрішніх аномалій із відомими та задокументованими уразливостями. Також необхідно створювати та наповнювати загальнодоступні джерела таких загроз

Керуючись схемою типового життєвого циклу хакерської атаки на інформаційні ресурси підприємства, яка наведена на рис. 1, можна побачити, що такий тип архітектури системи менеджменту процесів інформаційної безпеки дозволяє не тільки детектувати та передбачати атаку, але й автоматично запобігати їй ще на рівні системи управління інформаційною безпекою. Спираючись на метод статистичного модулювання, система має можливість автоматично ідентифікувати незвичайну поведінку, пов’язану із хакерською атакою. Далі, використовуючи процес машинного самонавчання, вона може самостійно відрізняти «нормальне» від «аномального».

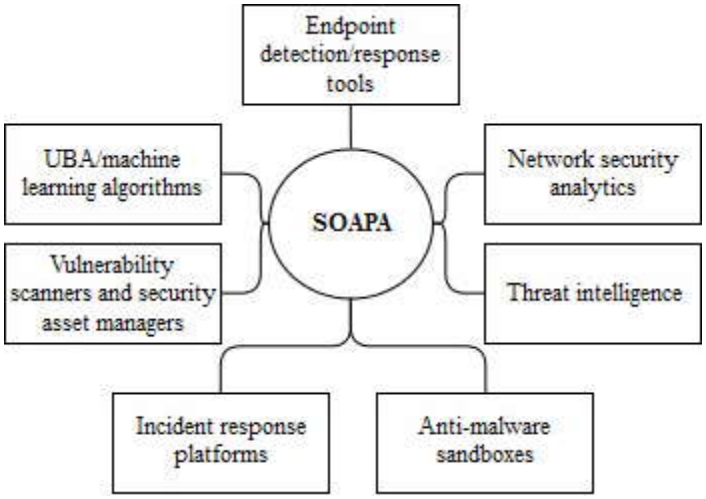


Рис. 2. Елементи, що входять в склад архітектури SOAPA

Збір інформації		
Тип атаки	***Цілі***	***Засоби***
Дослідження відкритої інформації	- веб сервери - зовнішні додатки - соціальні медіа	- пошукові сервіси - публічна інформація - зовнішнє сканування
Первісна експлуатація		

Тип атаки	*Цілі*	*Засоби*
Ініціювання атаки	- керівництво та фахівці - віддаленні співробітники	- вразливості нульового дня - соціальна інженерія - фішинг - створення дір
Створення пладарму	- робочі станції - web-сервери	- встановлене шкідливе ПЗ - вкрадені облікові записи

Управління та контроль		
Тип атаки	*Цілі*	*Засоби*
Отримання доступу	- системи безпеки - операційні системи	- руткіти - трояни - створення облікових записів - створення VPN
Внутрішня розвідка	- спільні каталоги - робочі станції - сервери - маршрутизатори	- встановлене шкідливе ПЗ - вкрадені облікові записи

Підвищення привілеїв		
Тип атаки	*Цілі*	*Засоби*
Горизонтальне переміщення	- спільні каталоги - робочі станції - сервери - маршрутизатори	- віддалене управління - вкрадені облікові записи
Підвищення привілеїв	- облікові записи адміністраторів - сервери - маршрутизатори	- руткіти - трояни - створення облікових записів

Управління даними		
Тип атаки	*Цілі*	*Засоби*
Збір та шифрування даних	- спільні каталоги - робочі станції - сервери - документи	- ftp або email - ZIP & RAR компресія - криптолокери
Викрадення даних	- документи - дослідження і розробки	- ftp або email - веб-розміщення - шифровані тунелі

Рис. 3. Типовий життєвий цикл хакерської атаки на інформаційні ресурси підприємства.

IV. Висновки

Архітектура SOAPA дозволяє більш поглиблено автоматизувати менеджмент процесів інформаційної безпеки підприємства в умовах значного збільшення обсягів даних, що генеруються системою управління інформаційною безпекою. Така архітектура дозволяє значно спростити життя фахівцям, і допомагає розставити пріоритети реагування на інциденти. Але однією з основних причин розробки нової архітектури стала не тільки необхідність автоматизації обробки інформації, а й автоматичне реагування на зовнішні і внутрішні загрози інформаційним ресурсам підприємства.

Слід також окремо відзначити, що доцільність та ефективність можливості впровадження SOAPA в державному секторі доки що не визначено. Але доцільність використання в банківській сфері не викликає сумнівів, так як вона підпорядковується міжнародним стандартам, і не суперечить існуючому законодавству в сфері захисту інформації в Україні. Так само варто пам'ятати що це всього лише концепція, і в найближчому майбутньому в неї можуть бути додані нові підсистеми.

Перевагами даної концепції є динамічність та швидка адаптованість умов інформаційно-комунікаційної архітектури.

Література References

[1] David R. Miller, Shon Harris, Stephen Vandyke, Security Information and Event Management (SIEM) implementation, McGrawHill, 2011.

[2] ISO/IEC27000 seriesFAQ–ISO27kForum[Електронний ресурс]. – Режим доступу: http://www.iso27001security.com/ html/faq.html.

[3] J. Oltsik, Security: Goodbye SIEM, Hello SOAPA?[Електронний ресурс]. – Режим доступу: https://www.channele2e.com/2016/12/28/security-goodbye-siem-hello-soapa

[4] J. Oltsik, Security: Security data growth drives SOAPA [Електронний ресурс]. – Режим доступу: http://www.networkworld.com/article/3154090/security/security-data-growth-drives-security-operations-and-analytics-platform-architecture-soapa.html

[5] O. Salah, Mandatory Information Security Management System Documents Required for ISO/IEC 27001 Certification [Електронний ресурс] / O. Salah, G. Hinson. – Режим доступу: http://www.iso27001security.com/ISO27k_mandatory_ISMS_documents.rtf.

[6] SAS® Security Intelligence [Електронний ресурс]. – Режим доступа: https://www.sas.com/software/security-intelligence

[7] Security Intelligence Operations [Електронний ресурс]. – Режим доступу: http://tools.cisco.com/security/center/home.x

[8] М. В. Гайворонський, О. М. Новіков, Безпека інформаційно-комунікаційних систем. – К.: Видавнича група BHV, 2009. – 608 с.

[9] Інформаційні технології. Методи захисту. Звід правил для управління інформаційною безпекою (ISO/IEC 27002:2005, MOD). ГСТУ СУІБ 2.0/ISO/IEC 27002:2010. –К.: Національний банк України, 2010. –163 с.

[10] Інформаційні технології. Методи захисту. Система управління інформаційною безпекою. Вимоги (ISO/IEC 27001:2005, MOD). ГСТУ СУІБ 1.0/ISO/IEC 27001:2010. –К.: Національний банк України, 2010. –49 с.

Оцінка Ризиків Інформаційної Безпеки в Мережах Wi-Fi на Основі Апарату Нечіткої Логіки

Леонід Куперштейн
Кафедра захисту інформації
Вінницький національний технічний університет
Вінниця, Україна
kupershtein.lm@gmail.com

Олеся Войтович
Кафедра захисту інформації
Вінницький національний технічний університет
Вінниця, Україна
voytovych.op@gmail.com

Татарчук Артем
Кафедра захисту інформації
Вінницький національний технічний університет
Вінниця, Україна
tatarchuka@bk.ru

Ратников Назар
Кафедра захисту інформації
Вінницький національний технічний університет
Вінниця, Україна
nazarratnikov@mail.ru

Information Security Risk Assessment of Wi-Fi Network Based on Fuzzy Logic

Leonid Kupershtein
dept. of Information Protection
Vinnitsia National Technical University
Vinnytsia, Ukraine

Olesya Voytovych
dept. of Information Protection
Vinnitsia National Technical University
Vinnytsia, Ukraine

Tatarchuk Artem
dept. of Information Protection
Vinnytsia National Technical University
Vinnytsia, Ukraine

Ratnikov Nazar
dept. of Information Protection
Vinnytsia National Technical University
Vinnytsia, Ukraine

Анотація—Розглядається модель оцінки рівня ризику загроз інформаційної безпеки в бездротових мережах WiFi за допомогою апарату нечіткої логіки. Використано алгоритм нечіткого виведення Мамдані. Дослідження проведено з використанням інструментарію Fuzzy Logic Toolbox пакету MatLAB.

Abstract—A model of risk assessment of information security threats in WiFi network is considered using fuzzy logic. The algorithm of Mamdani is used. The research was performed using the Fuzzy Logic Toolbox of the MatLAB.

Ключові слова—ризик, інформаційна безпека, загроза, нечітка логіка, лінгвістична змінна.

Keywords—risk, information security, thear, fuzzy logic, linguistic variable.

I. Вступ

Бездротові мережі на сьогоднішній день використовуються практично у дуже багатьох сферах людської діяльності. Широке використання бездротових мереж обумовлено тим, що вони можуть використовуватися не тільки на персональних комп'ютерах, а й телефонах, планшетах, ноутбуках, а а також їх зручністю і порівняно невисокою вартістю.

В наш час методи впливу на конкурентів переходять від фізичного впливу до інтелектуального. При цьому використовуються новітні способи і засоби несанкціонованого отримання інформації. Саме тому актуальним є необхідність оцінки ризиків інформаційної безпеки для побудови ефективних систем захисту інформації (СЗІ).

Існуючі методи оцінки ризиків в більшості засновані на теоріях ймовірності і класичних множин. Ці методи не дозволяють врахувати той факт, що будь-яка складна система є динамічною системою з набором невизначених даних. Системи оцінки ризиків, побудовані на застосуванні нечіткої логіки, можуть

характеризуватися логічністю і високою стійкістю в тому випадку, коли аналіз ризиків здійснюється в умовах нестачі даних і знань [1, 2].

II. Загрози в Wi-Fi мережах

Бездротове середовище передачі радіосигналу створює умови для неконтрольованого підключення до мережі. Класичні способи захисту Wi-Fi мережі, закладені в специфікації її стандартах, включають в себе шифрування і автентифікацію користувачів, і не забезпечують повного захисту мережі. Тому слід визначити ряд загроз, які можна оцінити при побудові чи проведенні аудиту мережі. Загрози Wi-Fi можна поділити на такі групи: «чужаки», нефіксована природа зв'язку, вразливості мереж та пристроїв, нові загрози і атаки, витік інформації з дротової мережі, особливості функціонування бездротових мереж [3 -5].

A. «Чужаки».

Чужаками називаються пристрої, які дають змогу несанкціонованого доступу до корпоративної мережі, часто в обхід механізмів захисту, визначених корпоративною політикою безпеки. Найчастіше це ті самі самовільно встановлені точки доступу. Статистика по всьому світу вказує на чужаків, як на причину більшості зломів мереж організацій. Навіть якщо організація не використовує бездротовий зв'язок і вважає себе захищеною від бездротових атак - впроваджений (навмисне чи ні) чужак з легкістю виправить це положення. Крім точок доступу в ролі чужака можуть виступити домашній роутер з Wi-Fi, програмна точка доступу Soft AP, ноутбук з одночасно включеними дротовим і бездротовим інтерфейсом, сканер, проектор і т.д.

B. Нефіксована природа зв'язку.

Бездротові пристрої не "прив'язані» кабелем до розетки і можуть змінювати точки підключення до мережі прямо в процесі роботи. Наприклад, можуть відбуватися «випадкові асоціації», коли ноутбук з Windows XP (досить довірливо ставиться до всіх бездротових мереж) або просто некоректно налаштований бездротовий клієнт автоматично асоціюється і підключає користувача до найближчої бездротової мережі. Такий механізм дозволяє зловмисникам «перемикати на себе» нічого не підозрюючого користувача для подальшого сканування вразливостей, фішингу або атак типу «людина в середині».

C. Загрози пов'язані із вразливостями мереж та пристроїв.

Деякі мережеві пристрої, можуть бути більш уразливі, ніж інші - можуть бути неправильно налаштовані, використовувати слабкі ключі шифрування або методи аутентифікації з відомими вразливостями. В першу чергу зловмисники атакують саме їх. Понад 70 відсотків успішних зломів бездротових мереж відбулися саме в результаті неправильної конфігурації точок доступу або клієнтського програмного забезпечення. До вразливостей також відносять злам шифрування, не правильні налаштування клієнтів та точок доступу.

D. Нові загрози і атаки.

Бездротові технології породили нові способи реалізації старих загроз, а також деякі нові, досі неможливі в провідних мережах. У всіх випадках, боротися з атакуючим стало набагато важче, тому що неможливо ні відстежити його фізичне місце розташування, ні ізолювати його від мережі. До них відносять: розвідку, імперсонацію, відмову в обслуговуванні.

E. Витік інформації з дротової мережі

Практично всі бездротові мережі в якийсь момент з'єднуються з провідними. Відповідно, будь-яка бездротова точка доступу може бути використана як плацдарм для атаки. Але це ще не все: деякі помилки в конфігурації точок доступу в поєднанні з помилками конфігурації провідної мережі можуть відкривати шляхи для витоків інформації. Найбільш поширений приклад - точки доступу, що працюють в режимі моста (Layer 2 Bridge), підключені в плоску мережу (або мережа з порушеннями сегментації VLAN) і передають в ефір широкомовні пакети з дротового сегмента, запити ARP, DHCP, фрейми STP і т.п . Деякі з цих даних можуть бути корисними для організацій атак Man-in-The-Middle, різних Poisoning і DoS атак, та й просто розвідки.

Інший поширений сценарій ґрунтується на особливостях реалізації протоколів 802.11. У разі, коли на одній точці доступу налаштовані відразу кілька ESSID, широкомовний трафік буде поширяться відразу в усі ESSID. В результаті, якщо на одній точці налаштована захищена мережа і публічний хот-спот, зловмисник, підключений до хот-споту, може, наприклад, порушити роботу протоколів DHCP або ARP в захищеній мережі.

F. Особливості функціонування бездротових мереж.

Деякі особливості функціонування бездротових мереж породжують додаткові проблеми, здатні впливати в цілому на їх доступність, продуктивність, безпеку і вартість експлуатації. Для грамотного вирішення цих проблем потрібен спеціальний інструментарій підтримки і експлуатації, спеціальні механізми адміністрування та моніторингу, не реалізовані в традиційному інструментарії управління бездротовими мережами

III. Формулювання лінгвістичних змінних

У процесі аналізу факторів ризику виявлено показники, які можуть бути джерелами ризику ІБ в WiFi мережах. При завданні лінгвістичних змінних, що характеризують фактори ризику: ймовірність виникнення загрози - загрози це можливість здійснення загрози по відношенню до будь-якої системи або ресурсу; збиток від загрози це втрати які понесе організація від реалізації загрози; ймовірність реалізації атаки – можливість успішного проведення атаки.

Фактори ризику представлено в формі значення тер-множин:

Ймовірність виникнення загрози (U) = {Низька(Н), середня(С), висока(В)}.

Збиток від загрози (Р) = {Дуже низький(ДН), низький(Н), середній(С), високий(В), дуже високий(ДВ)}.

Ймовірність реалізація загрози (І)= {Низька(Н), середня(С), висока(В)}.

Для отримання вихідної змінної ризику загрози визначимо таку лінгвістичну змінну:

Ризик загрози (R)= {Низький(Н), помірний(П), середній(С), високий(В), екстремальний(Е)}.

IV. Нечітка продукційна модель оцінки ризику

На етапі фазифікації здійснюється встановлення співвідношення між конкретним (чисельним) значенням окремої вхідної змінної системи нечіткого виводу і значенням функції приналежності відповідного їй терму вхідної лінгвістичної змінної. Після завершення цього етапу для всіх вхідних змінних повинні бути визначені конкретні значення функцій приналежності по кожному із лінгвістичних термів, які використовуються в підумовах бази правил системи нечіткого виведення [6, 7].

На рис. 1−3 представлено трикутні функції приналежності лінгвістичних змінних. Трикутні функції приналежності задаються аналітичною формулою:

$$\mu = \begin{Bmatrix} 0, x \le a \\ \dfrac{x-a}{b-a}, a \le x \le b \\ \dfrac{c-x}{c-b}, b \le x \le c \\ 0, c \le x \end{Bmatrix},$$

де [a, c] - діапазон зміни змінної;

b - найбільш можливе значення змінної.

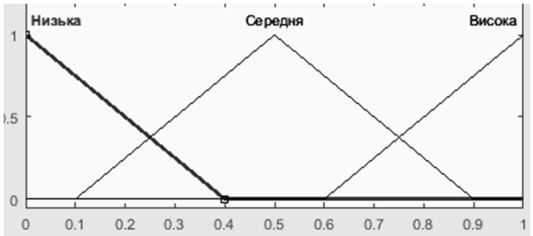


Рис. 1. Функція приналежності ймовірності загрози

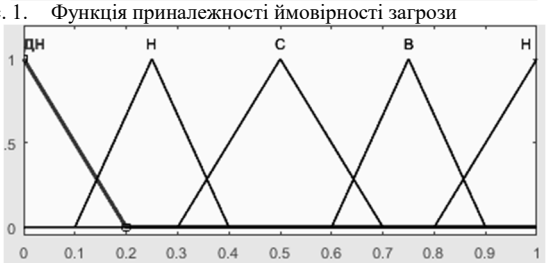


Рис. 2. Функція приналежності збитку від реалізації загрози

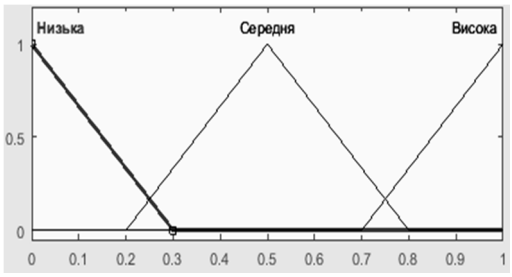


Рис. 3. Функція приналежності ймовірності реалізації загрози

Формування нечітких продукційних правил найчастіше представляється у формі узгоджених щодо використовуваних лінгвістичних змінних структурованого тексту:

ПРАВИЛО_1: ЯКЩО «Умова_1» ТО «Заключення_1» (F 1 т),

...

ПРАВИЛО_n: ЯКЩО «Умова_n» ТО «Заключення_n» (F n т),

де F і ∈ [0; n] є коефіцієнтом визначеності або ваговим коефіцієнтом відповідного правила.

Наприклад:

Правило 1. (U=H and P=ДН and I=H) or (U=C and P=ДН and I=H) or (U=H and P=ДН and I=C) or (U=H and P=H and I =H) or (U=H and P=C and I=H) then R = H.

Правило 2. (U=H and P=C and I=C) or (U=C and P=ДН and I=C) or (U=C and P=H and I=C) or (U=C and P=C and I=H) or (U=B and P=ДН and I=H) or (U=H and P=B and I=H) or (U=H and P=ДН and I=B) or (U=H and P=H and I=B) or (U=H and P=C and I=B) or (U=H and P=B and I=C) or (U=B and P=ДН and I=C) then R = П.

Правило 3. (U=C and P=B and I=H) or (U=C and P=ДН and I=B) or (U=C and P=H and I=B) or (U=B and P=ДН and I=H) or (U=B and P=H and I=H) or (U=B and P=C and I=H) or (U=C and P=C and I=C) or (U=C and P=B and I=C) or (U=B and P=B and I=H) or (U=B and P=ДН and I=C) or (U=C and P=C and I=B) or (U=H and P=B and I=B) or (U=B and P=C and I=H) or (U=B and P=H and I=C) or (U=H and P=C and I=C) or (U=C and P=ДВ and I=C) then R = C.

Правило 4. (U=B and P=C and I=C) or (U=B and P=C and I=B) or (U=B and P=B and I=C) or (U=C and P=B and I=B) or (U=C and P=ДВ and I=B) then R = B.

Правило 5. (U=B and P=B and I=B) or (U=B and P=ДВ and I=B) then R = E.

В алгоритмі Мамдані база правил задається у вигляді структури з трьома входами і одним виходом (рис. 4).

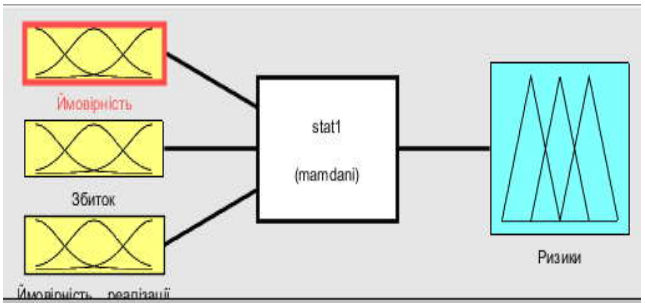


Рис. 4. Структура нечіткої моделі

Етап дефазифікації полягає в тому, щоб, використовуючи результати вихідних лінгвістичних змінних, отримати звичайне кількісне значення кожної із вихідних змінних, яке може бути використано спеціальними пристроями [6].

Перехід від вихідної лінгвістичної змінної до числового значення відбувається методами центру ваги [7]:

$$x = \frac{\int_{\min}^{\max} x\mu(x)dx}{\int_{\min}^{\max} \mu(x)dx},$$

де x − це модальне значення (мода) нечіткої множини, відповідної вихідної змінної після фазифікації.

Реалізуючи систему нечіткого виведення на етапі дефазифікації, отримується оцінка пріоритету ризику. Пріоритезація є основною метою аналізу ризиків і основоположним чинником в процесі прийняття рішень з управління ризиками ІБ організації.

Дослідження було проведено з використанням інструментарію Fuzzy Logic Toolbox пакету MatLAB.

V. Використання моделі оцінки ризиків ІБ Wi-Fi мережі

Припустимо, що на основі попереднього обстеження отримані деякі оцінки рівня захисту мережі, які введемо в вікно механізму виведення графічного інтерфейсу Fuzzy Logic Toolbox.

При значенні ймовірності загрози 0,6, значення лінгвістичної змінної відповідає терму С. При значенні збитку від загрози 0,8 значення лінгвістичної змінної відповідає терму В. При значенні ймовірності реалізації загрози 0,2, значення лінгвістичної змінної відповідає терму Н. За заданими вихідними умовами активізується правило №3. Результуюче значення змінної ризик відповідає значенню 0.5, що визначає значення лінгвістичної змінної С.

На рис. 5 представлено графік «кривої логічного виведення».

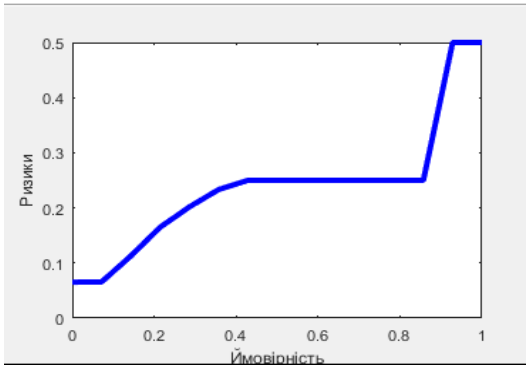


Рис. 5. Залежність змінної R від U

VI. Висновки

Розроблена нечітка продукційна модель оцінки ризиків загроз інформаційної безпеки мереж Wi-Fi дозволяє істотно розширити можливості існуючих методик, зняти обмеження на кількість врахованих вхідних змінних і інтегрувати як якісні, так і кількісні підходи до оцінки ризиків.

Основна складність механізму отримання оцінок ризику на основі нечіткої логіки полягає в побудові моделі для проведення лінгвістичного аналізу ризиків мереж Wi-Fi, однак, даний механізм є ефективним інструментом, коли інші підходи до оцінки ризику неприйнятні. Він володіє широкими можливостями і дозволяє адаптувати його до нових загроз, а також модифікувати з урахуванням реального стану мережі.

Література References

[1] А. Г. Корченко, Построение систем защиты информации на нечетких множествах / А. Г. Корченко - Киев : МК-Пресс, 2006. - 312 с.

[2] А. М. Астахов, Искусство управления информационными рисками / А. М. Астахов – Москва : ДМК Пресс, 2010. - 312 с.

[3] В. Б. Щербаков, Риск-анализ атакуемых беспроводных сетей: Монография/ В.Б. Щербаков, С.А. Ермаков, Н.С. Коленбет; под ред. чл.-корр. РАН Д.А. Новикова. – Воронеж: Издательство «Научная книга», 2013. – 160 с.

[4] В. Б Щербаков, Безопасность беспроводных сетей: стандарт IEEE 802.11. / В. Б. Щербаков, С. А. Ермаков – Москва : Радио Софт, 2010. - 256 с.

[5] В. Ф. Шаньгін, Захист інформації в комп'ютерних системах і мережах / В. Ф. Шаньгін – Київ : МК Прес, 2012. - 592 с.

[6] Д. Рутковская, Нейронный сети, генетические алгоритмы и нечеткие системы / Д. Рутковская, М. Пилиньский, Л. Рутковский - Москва : Горячая линия-Телеком, 2006. - 388 с.

[7] А.П. Ротштейн, Интеллектуальные технологии идентификации: нечеткие множества, нейронные сети, генетические алгоритмы / А.П. Ротштейн. - Винница: Універсум-Вінниця, 1999. - 295 с.

Access Control Method Based on Workstation Binding

Yurii Baryshev
dept. of Information Protection
Vinnytsia National Technical University
Vinnytsia, Ukraine
yuriy.baryshev@gmail.com

Kristina Neuimina
dept. of Information Protection
Vinnytsia National Technical University
Vinnytsia, Ukraine
kris.vladimirovna99@gmail.com

Метод Розмежування Прав Доступу з Прив'язкою до Робочої Станції

Юрій Баришев
кафедра захисту інформації
Вінницький національний технічний університет
Вінниця, Україна
yuriy.baryshev@gmail.com

Крістіна Неуйміна
кафедра захисту інформації
Вінницький національний технічний університет
Вінниця, Україна
kris.vladimirovna99@gmail.com

***Abstract*—This paper presents an analysis of access control models. The proposed method of access control based on the hashing process peculiarities allowing to limit the quantity of workstations from which users are to get remote access to information resources.**

***Анотація*—В даній роботі представлено аналіз моделей розмежування прав доступу. Запропонований метод розмежування доступу, який базуючись на особливостях процесу гешування дозволяє обмежити перелік робочих станції, з яких користувачу дозволяється отримувати віддалений доступ до інформаційних ресурсів.**

Keywords—authentication; hash; access control model; workstation parameters; authentication factors.

Ключові слова—автентифікація; гешування; модель розмежування прав доступу, параметри робочої станції; фактори автентифікації.

I. Introduction

There are a lot of information threats for one processed by computer means, that is caused by the presence of different possible threat sources within hardware and software used for the data processing. In particular the role of such sources might be performed by an enterprise personnel, intruders, malware, technical failures, backdoors etc [1-3]. The latter determines importance of tasks of information protection providing and simultaneously avoiding the substantial worsening computational processes performance quality. One of protection methods, which could be used for this task solving, is user access control. It allows to limit informational resources available to users in the computer system [2, 3]. At the same time mobile devices are spreading among users, and they also could be used to access this informational resources, but ones are quite limited of computing resources and hence information protection providing abilities. The typical workstation might provide high level of informational security and if it is situated at the enterprise, the information protection department of the enterprise could ensure that. All of these are out of reaching in the case, when valid users apply mobile devices or home personal computers for accessing critical informational resources of the enterprise.

The goal of this research is information confidentiality improvement for remotely accessed critical informational resources.

For reaching this goal it is essential to solve the set of different tasks. This work is to solve the following ones of these tasks:

- known access models analyses performing for determining ones, which could be used for reaching goal of the research;
- upgrading the access control models considering research goal;
- the method of user authentication development, that allows to implement selected access control models.

II. Access Control Models Analyses

The access control systems perform access management for subjects using information system to the objects of this system. The certain set of rules lies in the basis of access control model. The well-known models are divided into following groups: discretionary; mandatory; role-based.

During this research the following models were analysed: Harrison-Ruzzo-Ulman, Take-Grant, access matrix, Bell-LaPadula, RBAC [1, 4-6].

The Harrison-Ruzzo-Ulman model (known also as HRU) is the example of discretionary model of access control. The model allows to consider access control system as an automaton, that functions according to the certain rules of transition [1, 4, 5]. This model caused theoretical impact, because it allows research performing based on the automaton theory.

Another analyzed discretionary access control model is Take-Grant [1, 5, 6]. It provides possibility to analyse and check up the state of information system security. The Take-Grant model uses as basic elements type of user access and transformation allowed. It is based on the graph model, which allows to visualise enterprise information protection policy. The main task of the model is determination of user rights to access system as its subject relatively to certain object, which is described as some graph of allowed access types. Using this model, it is possible to study the states of information system depending of access granted to information system users [4, 6]. The access matrix hash the same approach as the latter model, but is presented in the way two-dimensional matrix.

The advantage of discretionary access control models is implementation evidence, independence from the information system type and extreme flexibility. However a key lack is necessity of the manual control over protection systems based on this model, and thus increasing of human factor impact on protection system that uses such model of access control.

The Bell-LaPadula model guarantees that a subject is able to become familiar with information only in case, when he has sufficient authority, and no subject (except the administrator authority to set up the confidentiality levels of objects) in no way will be able to carry out data transfer from an object with the higher level of confidentiality to an object with less level of confidentiality. The Bell-LaPadula model uses classifications of subjects by authority (or mandate) levels and objects by level of confidentiality [1].

If the user of the information system, that owns the high level of authority, would send some data (for instance, that have a level of secrecy, that equals his authority) to an object with less level of confidentiality, then it could become accessible to the subject with lower authority level, than it is settled by information protection policy [1, 5]. This model is also difficult for implementation and requires the considerable amount of computer system resources, thus it limits quantity of enterprises, where it can be used.

The role-based access control (RBAC) model considers additional element comparatively to the above-mentioned models. This model presents the computer system as sets sequence: the set of users, the set roles, the set of rights and privileges, the set of user sessions with the information system [1]. This access control model is appears as compromise between discretionary models manual control management complexity and its flexibility. The former is reduced by sacrificing the latter. However from the practical point of view this model seems to be most useful fro the most types of information systems [2, 3, 5].

The general lack of all these abovementioned models is that they do not provide limitations of workstations, that can be applied by the user to gain an access to informational resources of the system. And according to the previous analyses results, this is the most inappropriate for the remote access distribution systems such as file servers or cloud services providers.

For this research aim reaching the access control model, which is proposed at the article [7], was chosen to be implemented. This model allows access control rules forming depending on authentication data of the user and workstations, those are authorised to be used by him to get access to the certain information resource.

III. AUTHENTICATION METHOD ANALYSES

User gaining access process consists of three related consistently executable procedures: identification, authentication and authorization [3, 5].

There are a few methods of authentication, that differ in complexity, reliability, cost and other parameters. Each of these methods has its pros and cons. The methods of authentication conditionally can be divided into one factor (weak, from the security point of view) and multivariable (strong) [3-5], where features of the subject, who are to be authenticated, are used as factors.

The password authentication is the most widespread, simple and common method, where knowledge of the certain secret are used as an authentication factor. The role of password usually is performed by some word, but recently other types of passwords gain spreading, graphical ones, for instance [5].

The well-known methods of authentication, which are based on the unique objects usage. They could provide more reliable protection, than password authentication methods. These objects usually are divided into two groups [3, 5]:

- passive objects, that contain authentication information (for example, the randomly generated number, which could be interpreted as some kind of a password) and pass it during authentication process after the respective query from authentication module
- active objects, that has sufficient processing resources and actively participate at the authentication process (for example microprocessor-based smart cards and USB-tokens).

The authentication based on unique objects usage has the set of drawbacks concerning their application at the remote access control system [5]:

- the object can be stolen from the user;
- in most cases the special equipment is needed for working with such unique objects;

it is possible to make a copy of the object or emulate its presence.

The biometrical methods of authentication work on the basis of the usage of the equipment for measuring and comparing to the standards of the individual user's body features [5]. It could be implemented by the means of automated methods, that makes it posiible to be used at the informational systems. The authentication of people is performed on the basis of their physiology (static) and behavioural characteristics (dynamic) [5].

As physiological characteristics could be used features of finger-prints, retina and cornea of eyes, geometry of hands or face etc. As behavioural characteristics could be used the style of working tasks solution performing, for example using a keyboard or a mouse. Authentication systems based on voice recognition are also could be selected as the latter type.

Such means allow with high accuracy to recognize a user using certain biometrical characteristics, and they are difficult to being forged. The general lack of biometrical authentication means is a necessity for the addtitional equipment for biometrical characteristics estimation, that can be expensive and inconvenient in the case of remote access control systems. For workstation authentication it is suggested to use combination from a few unique parameters of this station such as [5, 7, 8]: hard drive serial number; date of creation and checksum of BIOS; hardware productivity; MAC-address and others.

In certain cases for granting iniquity to each of authentication sessions it is proposed to add cryptographic salt – a random number.

IV. Access Control Method

The implementation of selected access control model requires development of users authentication method. The method that is executed in accordance with a scheme presented on figure 1 [7]. The database of the authentication contains data presented by the set of vectors of following kind $\{h_i, h_{ij}, Access_i, ID_i, PC_j\}$, where h_i – hash value of *i*th user password; h_{ij} – hash value gotten on side of the client; $Access_i$ – the set of rights and privileges for *i*th user access; ID_i – *i*th user identifier; PC_j – *j*th workstation identifier [8].

It is seen from fig. 1, that the method using the iterative hash construction, which allows to keep at the server the hash value of user authentication factors (such as his password) without storing key used for the hashing process. Further hashing of *j*th workstation parameters guarantees coincidence of hash value yielded by user and server. For reduction of server computational resources usage it is proposed to execute the process of workstations parameters hashing after they are inserted into the database and their storing in the same way as other database content.

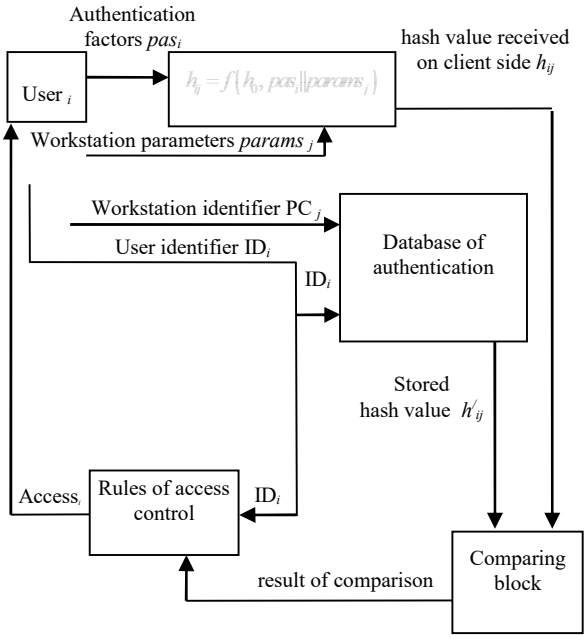


Fig. 1. Scheme of user authorisation

The method is based on the discretionary access control model. But the proposed approach of the workstations involvement into the authentication process could be applied to the all access control models, which were analysed above in this work.

In particular a discretionary model on the basis of access matrix after approach application would change in the following way: instead of two-dimensional matrix, a three-measurable matrix is to be used. Thus there are the dimension for subjects properties, the dimension – for user authentication factors and one – for workstation authentication factors.

For this user authentication method the scheme of the authentication system was developed, which implements access control model for distributed informational resources, those are stored remotely from the used workstation. An user sends his registration data to the hashing block on the client side. Then the hash value yielding is performed on the basis of user authentication factors and parameters of the workstation. The workstation identifier is sent to the server. On the server side it workstation parameters hashing is performed using the user authentication factors hash value as a key for the hashing process. If received from user hash value match the yielded one the authentication is considered to be successful and the user authentication – to be completed. Otherwise parameters of the next workstation allowed to be used be the user for accessing the informational resource are hashed.

On figure 2 the scheme of client side of the authentication system is presented.

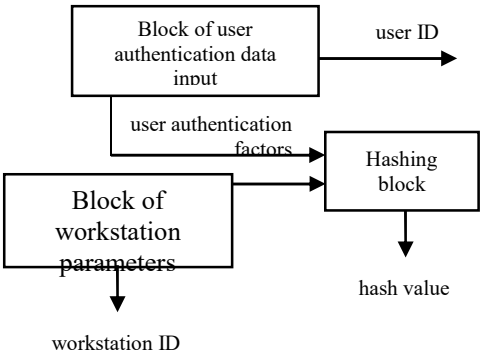


Fig. 2. The scheme of client side of the authentication system

The side of server of the system is shown on figure 3.

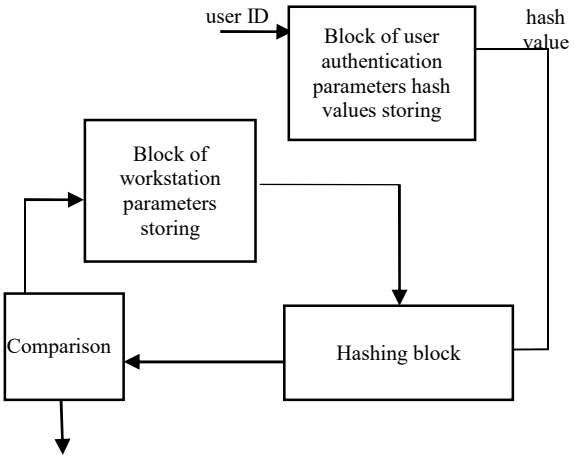


Fig. 3. The scheme of server side of the authentication system

The user authentication is performed in the following way. On side of the client user enters his identifier and authentication factors. The latter is concatenated with workstation parameters, the list of which was presented with the previous analyses results. The concatenation result is hashed using secret key known to user only. The hashing result is sent to the server as well as workstation identifier and user identifier.

The server side using received user identifier the hash value of the user authentication factors is determined by the means of the respective block. Parameters of the first workstation allowed for user to be used are yielded from the block of workstation parameters storing. Using the hashing block the server performs hashing of the parameters using the output of the clock of user authentication parameters hash values storing as a key.

The yielded hash value is compared with the received one. If they match the user is allowed to get access. Otherwise the next workstation parameters would be hashed at the next iteration. If there is no next workstation parameters stored at the respective block the access is denied for the user.

The proposed access control method could be used for other types of access control model (mandatory or role-based ones). The system is also quite flexible to the change of user authentication factors type, because if caused respective change of the user software/hardware platform without altering server. The latter is essential for the remote access granting information systems.

V. CONCLUSIONS

According to the performed analysis of access control models the number of tasks, for those it is important to limit the list of the workstations from which the user can get access to the critical information resources, were determined. As the result of access control system analyses the access control model, which allows to reach the research goal, was found. The method of remote users authentication and structure of the access control system, those due to the iterativeness of hashing process allow to execute users binding to the workstations, were performed for the implementation of the selected access control model. The method is planned to be used for the file server access control system.

REFERENCES

[1] П.Н. Девянин, Модели безопасности компьютерных систем. – М.: Издательский центр «Академия», 2005. – 144 с.

[2] Теория и практика обеспечения безопасного доступа к информационным ресурсам. Учебное пособие для вузов / А. А. Афанасьев, Л. Т. Веденьев, А. А. Воронцов и др.; Под ред. А. А. Шелупанова, С. Л. Груздева. – М.: Горячая линия – Телеком, 2009. – 552 с.

[3] В. А. Лужецький, Основи інформаційної безпеки : навчальний посібник / В. А. Лужецький, А. Д. Кожухівський, О. П. Войтович. – Вінниця: ВНТУ, 2013. – 221 с.

[4] В. В. Жора, Підхід до моделювання рольової політики безпеки *Правове нормативне та метрологічне забезпечення систем захисту інформації в Україн*.– 2003, Вип. 7 – С. 45–49

[5] Аутентификация. Теория и практика обеспечения безопасного доступа к информационным ресурсам. Учебное пособие для вузов. / А. А. Афанасьев, Л. Т. Веденьев, А. А. Воронцов и др.; Под ред. А. А. Шелупанова, С. Л. Груздева, Ю. С. Нахаева. – М. : Горячая линия-Телеком, 2009. – 552 с.

[6] В. Г. Миронова, А. А. Шелупанов, Н. Т. Югов, Реализация модели Take-Grant как представление систем разграничения прав доступа в помещениях *Доклады ТУСУРа.* – 2011. – № 2 (24). – С. 206– 210.

[7] Ю. В. Баришев, В. А. Каплун, Метод автентифікації віддалених користувачів для мережевих сервісів. *Інформаційні технологій та комп'ютерна інженерія*. – 2014. – №2. – С. 13-17.

[8] Ю. В. Баришев, К. В. Неуйміна, Метод розмежування прав доступу з прив'язкою до робочої станції. *Науково-технічна конференція факультету інформаційних технологій та комп'ютерної інженерії,* 2017 – https://conferences.vntu.edu.ua/index.php/all-fitki/all-fitki-2017/paper/view/2937/2501

Перевірка Модельних Припущень у Криптоаналізі Arx-Шифрів

Олександр Деркач
кафедра математичних методів захисту інформації
Фізико-технічний інститут
Національний технічний університет України «Київський політехнічний інститут ім. Ігоря Сікорського»
Київ, Україна
alex7derkach@gmail.com

Verification of Model Assumptions in Cryptanalysis of Arx-Ciphers

Oleksandr Derkach
Department of Mathematical Methods of Information Security
Institute of Physics and Technology
National Technical University of Ukraine “Igor Sikorsy Kyiv Polytechnic Institute”
Kyiv, Ukraine
alex7derkach@gmail.com

***Анотація*—На прикладі шифру ARX-ГОСТ експериментально перевірені модельні припущення двох найпоширеніших методів аналізу ARX-криптосистем – диференціального та обертального криптоаналізу.**

***Abstract*—We present experimental verification of model assumptions for two the most popular methods of ARX-ciphers cryptanalysis – differential and rotational cryptanalysis, applied to ARX-GOST cipher.**

Ключові слова—ARX-криптосистеми, диференціальний криптоаналіз, обертальний криптоаналіз, шифр ГОСТ

Keywords—ARX-cryptosystems, differential cryptanalysis, rotational cryptanalysis, GOST cipher

I. Вступ

ARX-криптосистеми (від англ. Add-Rotation-Xor) – окремий клас криптопримітивів, побудованих лише із застосуванням операцій модульного та побітового додавання, а також циклічного зсуву бітових векторів. Такі операції підтримуються усіма сучасними обчислювальними архітектурами на рівні базових, що дозволяє будувати надшвидкі криптопримітиви. Протягом останніх років було запропоновано багато криптографічних алгоритмів на основі ARX, зокрема, потокових шифрів (Salsa, ChaCha), блокових шифрів (Simon, Speck), геш-функцій (BLAKE, Skein) тощо. Для потокових шифрів швидкість роботи має важливе значення, тому дослідження ARX-примітивів в цьому напрямку є актуальною задачею.

У той же час методи аналізу стійкості ARX-криптосистем ґрунтуються здебільшого на експериментальних обчисленнях та модельних припущеннях, справжність яких іноді виглядає сумнівною. У даній роботі на прикладі конструкції ARX-ГОСТ буде розглянуто адекватність таких припущень для двох найпоширеніших в наш час методів аналізу ARX-криптосистем – диференціального та обертального криптоаналізу. Відштовхуючись від одержаних даних, будуть визначені параметри конструкції ARX-ГОСТ, за яких вона досягає максимальної стійкості до диференціального криптоаналізу.

II. Теоретичні відомості

Нехай V_n – множина всіх n-бітових векторів, на якій визначено три операції: операція $\lll$ циклічного зсуву вектора на один біт, операція $\oplus$ побітового додавання та операція $+$ додавання за модулем 2^n; при цьому вважається, що бітові вектори природним чином ототожнені із лишками кільця $\mathbf{Z}_{2^n}$: кожне число ототожнюється із вектором власного представлення у двійковій системі числення. *ARX-криптосистемою* будемо називати криптопримітив, побудований лише з використанням введених трьох операцій.

Стійкість ARX-криптосистем ґрунтується на тому, що операція додавання за модулем є складним нелінійним перетворенням над полем $\mathbf{F}_2$. Будемо називати *диференціалом* трійку векторів $(\alpha, \beta \to \gamma)$, які задовольняють рівнянню

$$(x \oplus \alpha) + (y \oplus \beta) = (x + y) \oplus \gamma,$$

а відповідну імовірність такої події (*xor differential probability*) позначимо через

$$xdp^{+}(\alpha, \beta \to \gamma) = Pr_{x,y}\{(x \oplus \alpha) + (y \oplus \beta) = (x + y) \oplus \gamma\}.$$

Аналітичні умови для перевірки існування диференціалів такого типу із ненульовими імовірностями, обчислення значення імовірностей, оцінки кількості та розподіли нетривіальних диференціалів, а також алгоритми пошуку диференціалів та їх окремих векторів із максимально можливими значеннями імовірностей наведені у роботі [1].

Багато існуючих досліджень стійкості ARX-криптосистем ґрунтується на експериментальному обчисленні максимумів імовірності диференціалів. При цьому зазвичай висувається *гіпотеза про марковську поведінку імовірностей*: імовірності проходження диференціалів через кожну операцію + вважаються незалежними, що дозволяє обчислювати імовірності диференціалів складних ARX-конструкцій шляхом звичайного множення імовірностей на окремих їх вузлах. Однак справедливість даної гіпотези не підтверджена, а для інших конструкцій симетричної криптографії (наприклад, SP-мереж із модульним додаванням у ключовому суматорі) ця гіпотеза не виконується. Тому в даному розділі буде проведена експериментальна перевірка гіпотези про марковість окремих ARX-примітивів.

Іншим методом криптоаналізу ARX-конструкцій є *обертальний криптоаналіз* (*rotational cryptanalysis*) [2]. Цей метод споріднений до диференціального криптоаналізу, але замість диференціалів в якості досліджуваної статистики розглядаються так звані *пари обертання* – пари векторів виду (x, x'), де $x' = x >>> r$ для певного фіксованого значення r. Пари обертання однозначно проходять через операцію $\oplus$ та циклічні зсуви. Імовірність проходження пар обертання через операцію + була встановлена у [3] (див. також [4]):

$$p(r) = Pr_{x,y}\{(x + y)' = x' + y'\} = \frac{1}{4}\left(1 + \frac{1}{2^r} + \frac{1}{2^{n-r}} + \frac{1}{2^n}\right).$$

При обертальному криптоаналізі ARX-криптосистем також часто висувалось припущення про маркову поведінку імовірностей проходження пар обертання через операції додавання; в цьому випадку імовірність проходження пари обертання через всю систему обчислювалась як $(p(r))^q$, де q – кількість операцій додавання за модулем у системі. Втім, вже самими розробниками даного методу встановлено, що за наявності в системі декількох операцій додавання підряд така оцінка є неадекватно заниженою [5]. Далі буде показано, що таке модельне припущення не буде виконуватись навіть якщо всі операції додавання у ARX-криптосистемі розмежовані побітовими додаваннями та циклічними зсувами.

Основною конструкцією, яка досліджується в даній роботі, є блоковий шифр ARX-ГОСТ, одержаний із блокового шифру ДСТУ ГОСТ 28147:2009 [6] (далі – просто ГОСТ) шляхом усунення S-блоків із раундової функції. Важливість такої конструкції зумовлена легкістю її реалізації на існуючих апаратних та апаратно-програмних рішеннях (оскільки вона, власне, є модифікованим шифром ГОСТ). Формалізуємо шифр ARX-ГОСТ у вигляді такого сімейства відображень.

Раундове перетворення F є параметризованим за двома параметрами n, r ключезалежним перетворенням виду:

$$F: \{0,1\}^n \times \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n \times \{0,1\}^n,$$

$$F_k(x, y) = (y, x \oplus ((y + k) <<< r)).$$

Тут n – довжина вхідного напівблоку, r – величина циклічного зсуву. Раундові ключі k також розглядаються як n-бітові вектори.

Шифр ARX-ГОСТ є ітеративним блоковим шифром, який складається із послідовного застосування раундових перетворень F із незалежними рівноімовірними раундовими ключами. Шифр ARX-ГОСТ, який складається з m раундів, будемо позначати через F^m; раундові ключі k_1, …, k_m для зручності запису будемо опускати.

III. Перевірка модельних припущень диференціального та обертального криптоаналізу для шифру ARX-ГОСТ

Блокові шифри, які використовують у ключовому суматорі додавання за модулем замість побітового додавання, зазвичай втрачають властивість марковості: імовірності раундових диференціалів перестають бути незалежними, оскільки вони будуть напряму залежати від значення точки входу. Продемонструємо це на такому простому прикладі.

Нехай S – 8-бітовий S-блок шифру AES (можна використати довільний криптографічний S-блок) і $g_k(x) = S[x + k]$. Двохраундовий шифр E визначимо як

$$E_{k_1,k_2}(x) = g_{k_2}(g_{k_1}(x)),$$

де раундові ключі k_1 та k_2 є незалежними та рівноімовірними.

Визначимо імовірності диференціалів шифру E двома способами: шляхом безпосереднього обчислення за визначенням:

$$P_1\{\alpha \to \beta\} = \frac{1}{\left(2^8\right)^3}\sum_x \sum_{k_1,k_2}\left[E_{k_1,k_2}(x \oplus \alpha) = E_{k_1,k_2}(x) \oplus \beta\right],$$

(тут квадратними дужками позначено індикаторну функцію – дужки Айверсона), а також через імовірності диференціальних характеристик, припускаючи, що такий шифр є марковським:

$$d^g(\alpha,\beta)=\frac{1}{2^{16}}\sum_x\sum_k\left[g_k(x\oplus\alpha)=g_k(x)\oplus\beta\right],$$

$$P_2\{\alpha\to\beta\}=\sum_\gamma d^g(\alpha,\gamma)d^g(\gamma,\beta).$$

Обчислення по всіх диференціалах (α, β) показали, що:

1) для 65005-ти диференціалів з 65536-ти можливих (тобто у 99.19%) значення імовірностей P_1 та P_2 відрізняється, причому у 32510-ти диференціалах (49.61% від загальної кількості розбіжних) істинна імовірність більша за модельну, а у 32495-ти (50.39%) – менша;

2) загальна середня квадратична відстань між істинними розподілами імовірностей диференціалів та модельними розподілами складає ≈ 0.027, тобто доволі суттєву величину з точки зору статистичної розрізнюваності.

Таким чином, експериментально встановлено, що при використанні різних алгебраїчних операцій блокові шифри втрачають властивість марковості. Тобто при аналізі блокових шифрів довільної структури необхідно завжди доводити чи спростовувати гіпотезу про марковість перетворень.

Далі експериментальним шляхом було перевірено, чи виконується гіпотеза про марковість для шифру ARX-ГОСТ. Через брак обчислювальних ресурсів була перевірена зменшена версія шифру із розміром напівблоку $n = 4$ та величиною внутрішнього циклічного зсуву $r = 1$.

Неважко показати, що імовірності раундових диференціалів шифру ARX-ГОСТ зводяться до імовірностей xdp^+ таким чином:

$$d^F(\alpha,\beta)=[\alpha_2=\beta_1]\,xdp^+(\alpha_2,0\to(\alpha_1\oplus\beta_2)>>>r),$$

де вхідна та вихідна різниці розбиваються на дві n-бітові частини: $\alpha=(\alpha_1,\alpha_2)$, $\beta=(\beta_1,\beta_2)$.

Відповідно, для m-раундового шифру ARX-ГОСТ необхідно порівняти значення істинних імовірностей диференціалів

$$P_1\{\alpha\to\beta\}=$$

$$=\frac{1}{2^{2n}}\frac{1}{\left(2^n\right)^m}\sum_{(x,y)}\sum_{k_1,\ldots,k_m}\left[F^m(x\oplus\alpha_1,y\oplus\alpha_2)=F^m(x,y)\oplus\beta\right],$$

із розрахунковою модельною імовірністю

$$P_2\{\alpha\to\beta\}=\sum_{\gamma_1,\ldots,\gamma_{m-1}}d^g(\alpha,\gamma_1)d^g(\gamma_1,\gamma_2)\ldots d^g(\gamma_{m-1},\beta).$$

Обчислення імовірностей всіх диференціалів для $m = 2$, 3 та 4 раундів шифрування показало, що для шифру ARX-ГОСТ гіпотеза про марковську поведінку імовірностей диференціалів повністю виконується: істинні значення імовірностей всіх диференціалів співпали із модельними. Таким чином, для оцінювання стійкості шифру ARX-ГОСТ до диференціального криптоаналізу можна використовувати припущення про марковість. Це дуже важливий результат з огляду на суттєвішу спрощенність аналітичного дослідження марковських шифрів у порівнянні із немарковськими. Втім, необхідно пам'ятати, що результати експериментальних обчислень в даному випадку потребують аналітичного підтвердження перед застосуванням під час оцінювання стійкості конкретних шифрів.

На наступному етапі дослідження була перевірена гіпотеза про марковість імовірностей пар обертання у шифрі ARX-ГОСТ. Для цього для різних значень зсуву s між векторами x та $x' = x >>> s$ порівнювалась істинна імовірність

$$P_1=Pr_{x,y,k_1,\ldots,k_m}\left\{F^m(x',y')=F^m(x,y)'\right\}$$

(тут обертання як на вході, так і на виході виконується над кожним напівблоком окремо, що виправдане структурою шифру) із модельною імовірністю

$$P_2=\left(p(s)\right)^m.$$

Результати обчислень наведені у таблиці 1.

ТАБЛИЦЯ I. РІЗНИЦІ ІМОВІРНОСТЕЙ ПАР ОБЕРТАННЯ

m	**Зсув на 1**		
	Модельне	***Істинне***	***Різниця***
2	0.178	0.178	0
3	0.075085	0.073261	0.001823
m	**Зсув на 2**		
	Модельне	***Істинне***	***Різниця***
2	0.153	0.153	0
3	0.059605	0.053963	0.005642
m	**Зсув на 3**		
	Модельне	***Істинне***	***Різниця***
2	0.178	0.178	0
3	0.075085	0.073261	0.001823

З табл. 1 видно, що припущення про марковість для імовірностей пар обертання у шифрі ARX-ГОСТ перестає виконуватись вже починаючи з третього раунду. Для шифру ARX-ГОСТ бачимо, що теоретична (модельна) імовірність має більше значення, аніж істинне, тому загальна оцінка стійкості до обертального криптоаналізу не повинна порушитись; однак невідомо, чи буде така залежність справджуватись для більшої кількості раундів та для інших ARX-примітивів.

IV. Пошук параметрів шифру ARX-ГОСТ, які максимізують стійкість до диференціального криптоаналізу

Наступною задачею даного дослідження є визначення параметрів шифру ARX-ГОСТ, за яких даний шифр є найбільш захищеним від диференціального криптоаналізу. Знову розглядалась зменшена восьмибітова версія шифру (n = 4), для якої визначалась максимальна імовірність диференціалів в залежності від величини внутрішнього зсуву r.

Оскільки в попередньому розділі було експериментально встановлено, що для шифру ARX-ГОСТ виконується припущення про марковість, для обчислення імовірностей диференціалів була обрана більш проста формула $P_2\{\alpha \rightarrow \beta\}$, яка використовує імовірності диференціальних характеристик. Стійкість шифру до диференціального криптоаналізу визначалась за двома параметрами:

- максимальним значенням імовірності диференціалу (повинна бути мінімально можливою для посилення стійкості шифру);
- кількістю диференціалів, які мають максимальну імовірність (повинно бути якомога менше для ускладнення пошуку таких диференціалів).

Результати експериментальних обчислень наведені у таблиці 2.

ТАБЛИЦЯ II. Максимуми імовірностей диференціалів та їх кількість (у дужках) в залежності від зсуву та кількості раундів

Кількість раундів	***r = 1***	***r = 2***	***r = 3***
2	0.5 (32)	0.5 (32)	0.5 (32)
3	0.5 (8)	0.5 (6)	0.5 (10)
4	0.5 (2)	0.5 (4)	0.5 (4)
5	0.25 (5)	0.5 (2)	0.25 (7)
6	0.125 (2)	0.25 (6)	0.125 (6)

З табл. 2 випливає, що найкращі показники стійкості шифру ARX-ГОСТ за двома параметрами досягаються при значенні внутрішнього зсуву r = 1. Таким чином, можна припустити, що для загальної конструкції шифру ARX-ГОСТ максимальна стійкість до диференціального криптоаналізу також буде досягатись при невеликих значеннях внутрішнього зсуву, зокрема, можливо, також на один біт.

V. Висновки

У даній роботі був проведений аналіз модельних припущень у диференціальному та обертальному криптоаналізі ARX-криптосистем на прикладі зменшеного восьмибітового шифру ARX-ГОСТ. Експериментально показано на прикладі S-блоку шифру AES, що не для будь-яких конструкцій виконується гіпотеза марковості імовірностей диференціалів. У той же час показано, що для конструкції ARX-ГОСТ це припущення виконується, що дозволяє значно спростити аналіз стійкості такого шифру. Але аналогічне припущення про незалежність імовірностей пар обертання для шифру ARX-ГОСТ порушується.

Ґрунтуючись на одержаних результатах, було знайдено величину внутрішнього зсуву для шифру ARX-ГОСТ, за якої такий шифр сягає максимальної стійкості до диференціального криптоаналізу. Виявилось, що внутрішній зсув повинен бути невеликий: для зменшеної версії шифру найкращі результати показав шифр із внутрішнім зсувом у один біт. Такий зсув зменшить максимум імовірності диференціала і зменшить загальну кількість таких диференціалів.

Результати даної статті можуть використовуватись для побудови нових стійких криптопримітивів, зокрема, потокових шифрів та блокових шифрів у режимах вироблення гами, на основі ARX-конструкцій, а також для аналізу існуючих алгоритмів легкої криптографії.

Подальші дослідження даної теми передбачають перевірку модельних припущень для збільшених версій шифру ARX-ГОСТ – 16 та 32 бітних і аналітичне обґрунтування побудованих моделей. Також для покращення параметрів шифру варто врахувати оцінки стійкості до обертального криптоаналізу.

Література References

[1] H. Lipmaa, S. Moriai, “Efficient algorithms for computing differential properties of addition” in *Lecture Notes in Computer Science*, Springer, 2001, vol. 2355, pp. 1-17.

[2] D. Khovratovich and I. Nikolic, “Rotational cryptanalysis of ARX” in *Fast Software Encryption*: 17th International Workshop (FSE 2010), Seoul, Korea, February 7-10 2010. Lecture Notes in Computer Science, vol. 6147, Springer, 2010, pp. 333–346.

[3] Th. A. Berson, “Differential cryptanalysis mod n with applications to MD5” in *Advances in Cryptology*: CRYPTO ’92. Lecture Notes in Computer Science, vol. 740, Springer-Verlag, 1993, pp. 71–80.

[4] M. Daum, “Cryptanalysis of hash functions of the MD4-Family,” Ph.D. thesis , Ruhr-Universität Bochum, May 2005.

[5] D. Khovratovich, I. Nikolic, J. Pieprzyk, Prz. Sokolowski and R. Steinfeld, “Rotational cryptanalysis of ARX revisited” [Online]. Available: http://eprint.iacr.org/2015/095

[6] Система обробки інформації. Захист криптографічний. Алгоритм криптографічного перетворення (ГОСТ 28147-89), ДСТУ ГОСТ 28147:2009.

Асимтотичні Розподіли Імовірностей Змішаних Диференціалів Випадкових S-Блоків

Всеволод Бахтігозін, Сергій Яковлєв
кафедра математичних методів захисту інформації
Фізико-технічний інститут
Національній технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»
Київ, Україна
seba.bakh@gmail.com, yasv@rl.kiev.ua

Asymptotic Distributions of Mixed-Type Differential Probabilities of Random S-Boxes

Vsevolod Bakhtigozin, Serhii Yakovliev
Department of Mathematical Methods of Information Security
Institute of Physics and Technology
National Technical University of Ukraine “Igor Sikorsy Kyiv Polytechnic Institute”
Kyiv, Ukraine
seba.bakh@gmail.com, yasv@rl.kiev.ua

***Анотація*—У роботі наводяться асимптотичні розподіли ймовірностей диференціалів випадкових перестановок за різними груповим операціями на вході та на виході (побітове та модульне додавання); також знайдено довірчий інтервал для ймовірностей диференціалів такого типу.**

***Abstract*—We present asymptotic distributions of probabilities of random permutation differentials with respect to different group operations for inputs and outputs (bitwise and modular additions are considered). Also we provide confidence interval for such mixed-type differential probabilities.**

Ключові слова—диференціальний криптоаналіз, розподіли диференціалів

Keywords—differential cryptoanalysis, distributions of differentials

I. Вступ

Диференціальний криптоаналіз є одним з найпотужніших методів криптоаналізу симетричних блокових шифрів. Сучасні методи доведення стійкості алгоритмів шифрування до диференціального криптоаналізу фактично зводять відповідні оцінки стійкості до певних обчислювальних параметрів раундових перетворень та їх окремих компонент, таких як S-блоки (див., наприклад, [1-4]). Зазвичай для підвищення стійкості до даного методу аналізу у блокових шифрах обираються бієктивні S-блоки (тобто підстановки на бітових векторах) із мінімально можливими імовірностями диференціалів.

У роботах [5,6] було досліджено асимптотичні розподіли ймовірностей диференціалів випадкових перестановок за різними груповими операціями, а саме побітового додавання (XOR), додавання за модулем та множення за модулем. Зокрема, було показано, що імовірності таких диференціалів підкорюються розподілам Пуассона із різними параметрами (1/2 для побітового додавання та 1 для інших операцій, які розглядались). Також були побудовані довірчі інтервали для диференціалів за кожною операцією та показано, що диференціали за побітовим додаванням мають в цілому суттєво більші імовірності.

Однак для деяких класів блокових шифрів оцінки доказової стійкості до диференціального аналізу будуються через значення імовірностей змішаних диференціалів, в яких різниці на вході та на виході S-блоку обчислюються за різними операціями (див., зокрема, [7-9]). У даній роботі ми розглянемо асимптотичні розподіли імовірностей диференціалів найпоширенішого змішаного типу, в яких різниці обчислюються за побітовим та модульним додаванням. Буде доведено, що імовірності змішаних диференціалів асимптотично підкорюються розподілу Пуассона із параметрами 1 або ½. Також для довільного розміру S-блоку нами буде одержано довірчій інтервал для диференціальних імовірностей змішаного типу.

II. Необхідні терміни та позначення

Нехай V_n – простір n-бітових векторів, і бінарна операція $\otimes$ визначає на V_n структуру абелевої групи. В якості $\otimes$ ми будемо розглядати, зокрема, операції $\oplus$ (побітове додавання) та + (додавання за модулем 2^n); в останньому випадку бітові вектори трактуються як цілі невід'ємні числа у двійковому записі.

Позначимо через π деяку підстановку на V_n. *Диференціалом* π називається довільна пара бітових векторів (α, β), які трактуються як різниці на вході та на виході π за операцією $\otimes$:

$$u \otimes v^{-1} = \alpha \Rightarrow \pi(u) \otimes (\pi(v))^{-1} = \beta,$$

де $\alpha, \beta, u, v \in V_n$, $\otimes$ – групова операція, обернені елементи обчислюються за цією операцією. *Імовірністю диференціала* (α, β) називається величина

$$DP^{\pi}_{\otimes}(\alpha,\beta) = Pr_{x \in V_n}\{\pi(x \otimes \alpha) = \pi(x) \otimes \beta\}.$$

Змішаним, або $(+,\oplus)$-диференціалом π, будемо називати пара бітових векторів (α, β), які трактуються як різниці на вході та виході за різними операціями: для двох вхідних значень u, v

$$(u - v) \bmod 2^n = \alpha,\ \pi(u) \oplus \pi(v) = \beta.$$

Імовірністю $(+,\oplus)$-диференціалу називається величина

$$DP^{\pi}_{+,\oplus}(\alpha,\beta) = Pr_{x \in V_n}\{\pi(x + \alpha) = \pi(x) \oplus \beta\}.$$

Аналогічним чином визначаються $(\oplus,+)$-диференціали підстановки π та їх імовірності $DP^{\pi}_{\oplus,+}$.

Нехай $a = ord\alpha$, $b = ord\beta$, де (α, β) – довільний диференціал. Ймовірність такого диференціалу залежить лише від порядків елементів у групах, які визначаються відповідними операціями [5]:

$$p_t(a,b) = Pr_{\pi}\{2^n \cdot DP^{\pi}(\alpha,\beta) = t\}.$$

Для абелевої групи $\langle V_n, \otimes \rangle$ визначимо такі множини:

$$E_{\delta} = \{(u,v) \mid u \otimes v^{-1} = \delta\}$$

– множина пар елементів, які мають різницю δ за відповідною операцією, та

$$A_{uv} = \{\pi \mid (\pi(u), \pi(v)) \in E_{\delta}\}$$

– множина перестановок, що відображає пару елементів (u,v) у пару елементів, яка належить множини E_{δ}.

III. Асимптотичні розподіли імовірностей змішаних диференціалів

Розглянемо довільний диференціал (α, β), де $\alpha, \beta \neq 0$, $a = ord\alpha$, $b = ord\beta$. Імовірність $p_t(a,b)$ визначається як $p_t(a,b) = P_t/2^n!$, де P_t – кількість перестановок, які відображають рівно t елементів з множини E_α у множину E_β.

У роботі [5] було показано, що

$$P_t = \sum_{i=0}^{2^n - t} (-1)^i C^i_{t+i} S_{t+i},$$

де $S_k = \sum_{Y \subseteq E_\alpha, |Y| = k} \left| \bigcap_{uv \in Y} A_{uv} \right|$.

Наведений вище вираз для S_k стає надзвичайно складним для a і b більше 2. Розглянемо інший вираз, визначений не в термінах пар елементів, а в термінах окремих елементів. Нехай $Y \subseteq E_\alpha$, $|Y| = k$, тоді кількість окремих елементів $p(Y)$ у множині Y буде знаходитись у межах від k (всі елементи утворюють єдиний цикл) до $2k$ (жоден окремий елемент не утворює двох диференціалів). Розглянемо функцію:

$$\varphi(k,j) = \sum_{Y \subseteq E_\alpha, |Y| = k, p(Y) = j} \left| \{\pi \mid \pi(Y) \subseteq E_\beta\} \right|.$$

Тоді вираз для S_k приймає наступний вигляд:

$$S_k = \sum_{j=k}^{2k} \varphi(k,j)$$

В цій сумі домінуючим додатком є $\varphi(k,2k)$. Якщо $j \neq 2k$, маємо два випадки:

1) $ord\alpha = 2$ – таких елементів всього два: 0 та 2^{n-1}, причому перше значення відповідає тривіальному диференціалу, який має константний розподіл. Оскільки $b = 2$ для будь-якого диференціалу, то одержуємо ситуацію, яка була повністю досліджена у [5, 6]. Зокрема, було показано, що величина $2^n DP(\alpha,\beta)$ має розподіл Пуассона з параметром 1/2

2) $ord\alpha \neq 2$ – в цьому випадку маємо дві пари елементів (x, y), $(y, z) \subseteq E_\alpha$, причому $x \neq z$ (інакше це попередній випадок, оскільки тоді $ord\alpha = 2$) Але тоді маємо

$$\pi(x) \oplus \pi(y) = \pi(y) \oplus \pi(z) = \beta.$$

Отже, $\pi(x) \oplus \pi(z) = 0$, що можливо тільки при $x = z$; маємо протиріччя, а тому одна з цих пар не належить E_β. Отже, доданків такого виду у S_k не міститься взагалі.

Перейдемо до оцінки величини $\varphi(k,2k)$. Побудова перестановок $\pi(Y) \subseteq E_\beta$ відбувається у такий комбінаторний спосіб. Існує 2^n способів відобразити перший елемент однієї пари із різницею α; другий елемент

цієї пари після цього відображається однозначно. Для наступної пари маємо вже $2^n - 2$ способів побудови образу, і так далі, доки не отримаємо k пар для диференціалу (α, β). Елементи, які залишились, можна довільно відобразити $(2^n - 2k)!$ способами. Таким чином, кількість перестановок π: $\pi(Y) \subseteq E_\beta$ дорівнює

$$\prod_{i=0}^{k-1}(2^n - 2i)(2^n - 2k)! = \left(\frac{2^n}{2}\right)^k 2^k (2^n - 2k)! \approx 2^{nk}(2^n - 2k)!$$

Відповідно, кількість множин $Y \subseteq E_\alpha$, $|Y| = k$ оцінюється у такий спосіб: першу пару можна вибрати 2^n способами, другу – $(2^n - 3)$ способами (оскільки не можна вибирати пари, суміжні з вже обраними) і т.д. Кількість множин $Y \subseteq E_\alpha$, $|Y| = k$, дорівнює

$$\frac{1}{k!}\prod_{i=0}^{k-1}(2^n - 3i) = \frac{1}{k!}\left(\frac{2^n}{3}\right)^k 3^k \approx \frac{2^{nk}}{k!}$$

(Відповідні апроксимації наведено у [6].)

Тоді вираз для $\varphi(k,2k)$ приймає вигляд:

$$\varphi(k,2k) = 2^{nk}(2^n - 2k)!\frac{2^{nk}}{k!} \approx \frac{2^n!\,2^{2nk}}{k!\,2^{2nk}}.$$

Підставляючи це в формулу для S_k, а потім і для P_t, остаточно одержуємо:

$$P_t = \sum_{i=0}^{2^n - t}(-1)^i C_{t+i}^i \frac{2^n!}{(t+i)!} = \sum_{i=0}^{2^n - t}(-1)^i \frac{(t+i)!\,2^n!}{i!\,t!\,(t+i)!} = \frac{2^n!}{t!}\sum_{i=0}^{2^n - t}\frac{(-1)^i}{i!} \approx \frac{2^n!}{t!}e^{-1}$$

Звідси $p_t(a,b) = e^{-1}/t$, тобто імовірності $(+,\oplus)$-диференціалів для випадкової перестановки підкорюються розподілу Пуассона з параметром 1 – так само, як і імовірності $(+,+)$-диференціалів.

Зауважимо, що одержані асимптотичні оцінки залишаються вірними і для $(\oplus,+)$-диференціалів через природну відповідність: будь-якому $(\oplus,+)$-диференціалу (α, β) перестановки $\pi(x)$ відповідає $(+,\oplus)$-диференціал (β, α) зворотної перестановки $\pi^{-1}(x)$; оскільки розподіли диференціалів обчислюються на множині всіх перестановок, то асимптотичні розподіли означених диференціалів будуть співпадати. Оскільки ж асимптотичні розподіли всіх $(+,\oplus)$-диференціалів (β, α) при $\mathrm{ord}\beta \neq 2$ однакові, то вони будуть однаковими й для всіх $(\oplus,+)$-диференціалів (α, β).

IV. Довірчі інтервали для імовірностей змішаних диференціалів

Побудова довірчого інтервалу для імовірностей $(+,\oplus)$-диференціалів (а також для $(\oplus,+)$-диференціалів) відбувається в майже такий само спосіб, як і побудова довірчого інтервалу для $(+,+)$-диференціалів, описаного у [5].

Позначимо $DP(\pi) = max_{\alpha,\beta\neq 0} DP^{\pi}_{+,\oplus}(\alpha,\beta)$. Знайдемо обмеження для $DP(\pi)$ з асимптотичною ймовірністю 1. Визначимо величину

$$\theta_t(\pi) = \frac{1}{(2^n - 1)^2}\sum_{\alpha\neq 0}\sum_{\beta\neq 0} I\left(2^n \cdot DP(\alpha,\beta) = t\right),$$

що є часткою вхідних/вихідних різниць, які справедливі рівно для t пар входів (виходів), $0 \le t \le 2^n$. Маємо, що $M\theta_t(\pi) \sim e^{-1}/t!$

Покладемо $\Omega^{(t)} = (2^n - 1)^2\theta_t(\pi)$, а також $\Omega = \Omega^{(2B_n)}$, де

$$B_n = \frac{\ln N^2}{\ln\ln N^2},\ N = 2^n - 1$$

Тоді $M\Omega = (2^n - 1)^2 M\theta_{2B_n}(\pi) \sim N^2 e^{-1}/(2B_n)!$

Використаємо формулу Стірлінга для асимптотичної оцінки факторіалу:

$$(2B_n)! \sim \left(\frac{2B_n}{e}\right)^{2B_n}\cdot\sqrt{2\pi\cdot 2B_n} = \frac{\left(\ln N^2\right)^{2\ln N^2/\ln\ln N^2}}{\left((e/2)\ln\ln N^2\right)^{2B_n}} 2\sqrt{\pi B_n},$$

де $\left(\ln N^2\right)^{2\ln N^2/\ln\ln N^2} = \left(e^{\ln\ln N^2}\right)^{2\ln N^2/\ln\ln N^2} = N^2$

Звідси

$$M\Omega \sim N^2\frac{e^{-1}}{N^4}\left((e/2)\ln\ln N^2\right)^{2\ln N^2/\ln\ln N^2}\frac{1}{2\sqrt{\pi B_n}} = \frac{e^{-1}}{2\sqrt{\pi B_n}}\left(\frac{\left((e/2)\ln\ln N^2\right)^{2/\ln\ln N^2}}{(N^2)^{1/\ln N^2}}\right)^{\ln N^2} = \frac{e^{-1}}{2\sqrt{\pi B_n}}\left(\frac{\left((e/2)\ln\ln N^2\right)^{2/\ln\ln N^2}}{e}\right)^{\ln N^2}.$$

Вираз в дужках позначимо через $y(N)$. Можна показати, що $y(N) \le 1$. Отже,

$$M\Omega \sim \frac{e^{-1}}{2\sqrt{\pi B_n}} y(N)^{\ln N^2} = o(1),$$

при $n \to \infty$. Відповідно, при $2B_n = o(2^{n/2})$ маємо

$$M\Omega^{(t)} = M\Omega/(2B_n)^{t-2B_n}.$$

Таким чином, середня кількість нетривіальних диференціалів, імовірність яких більша або рівна за $2B_n/2^n$, дорівнює

$$\sum_{t \geq B_n} M\Omega^{(t)} \leq \sum_{t \geq B_n} \frac{M\Omega}{(2B_n)^{t-2B_n}} = M\Omega \sum_{i \geq 0} \frac{1}{(2B_n)^i} =$$

$$= \frac{M\Omega}{1 - 1/(2B_n)} \sim M\Omega = o(1)$$

Ймовірність $DP(\pi) \geq B_n/2^{n-1}$ менше за середню кількість диференціалів із імовірністю $t/2^n \geq 2B_n/2^n$. Звідси випливає, що

$$Pr\left\{DP(\pi) \geq \frac{B_n}{2^{n-1}}\right\} = o(1), \text{ ν.}$$

Таким чином, величина $DP(\pi)$ для змішаних диференціалів майже напевно лежить у межах від 0 до $2B_n/2^n$.

V. Перевірка швидкості збіжності

Швидкість збіжності імовірностей диференціалів до їх асимптотичних розподілів була оцінена експериментально для перестановок розміру від 4 до 10 бітів.

В ході експерименту було згенеровано 100 000 випадкових перестановок відповідного розміру, на основі яких обчислено вибіркові розподіли кожного диференціала. Для перевірки збіжності обчислювалось середньоквадратична відстань (СКВ) між вибірковим розподілом та розподілом Пуассона із параметром 1. Для оцінки швидкості розглядались максимальне та середнє по всіх диференціалах значення середньоквадратичної відстані. Результати обчислень наведені на рис 1,2.

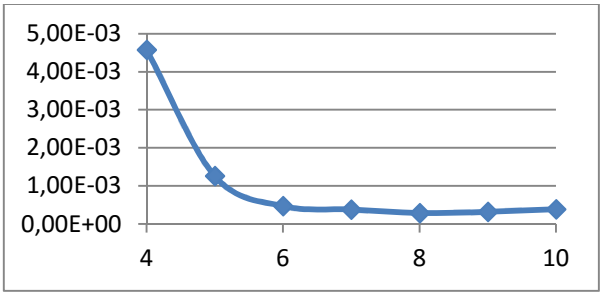


Рис. 1. Максимальне значення СВК імовірностей по всіх диференціалах

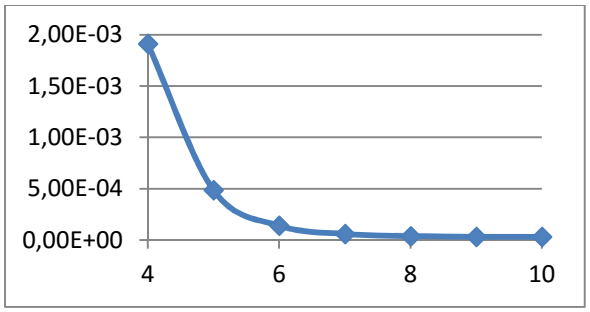


Рис. 2. Середнє значення СВК імовірностей по всіх диференціалах

VI. Висновки

У даній роботі було розглянуто асимптотичні розподіли імовірностей диференціалів найпоширенішого змішаного типу, в яких різниці обчислюються за операцією модульного додавання на вході та операцією побітового додавання на виході або навпаки. Було встановлено, що ймовірності диференціалів такого типу асимптотично підкорюються розподілу Пуассона з параметром 1, окрім класу диференціалів із вхідною різницею $\alpha = 2^{n-1}$, для яких відповідний параметр розподілу Пуассона становить ½. Також було знайдено довірчій інтервал для диференціалів змішаного типу і показано, що майже завжди значення імовірностей таких диференціалів не дуже великі. Одержані аналітичні твердження були перевірені експериментально; результати експериментів показали, що вже для 6-бітових S-блоків середньоквадратичне відхилення розподілів від асимптотичних не перевищує 10^{-4}.

Одержані результати можуть бути використані для побудови адекватних моделей та оцінювання стійкості до диференціального криптоаналізу немарковських блокових шифрів із різними алгебраїчними операціями у раундових перетвореннях та ключову суматорі.

Література References

[1] K. Nyberg, Provable Security Against a Differential Attack / K. Nyberg, L.R. Knudsen // Journal of Cryptology. – Vol.8. – No.1. – 1995.

[2] C. Adams, Designing S-boxes resistant to differential cryptanalysis [електронний ресурс] / C. Adams, St. Tavares. – Режим доступу : [14] http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.40.2536

[3] S. Park, On the security of Rijndael-like structures against differential and linear cryptanalysis / S. Park, S.H. Sung, S. Chee, E.-J. Yoon, J. Lim // Advances in Cryptology, ASIACRYPT 2002. – LNCS, vol. 2501. – Berlin: Springer, 2002. – pp. 176-191.

[4] S. Park, Improving the upper bound on the maximum differential and the maximum linear hull probability for the SPN structures and AES / S. Park, J. Sung, S. Lee, J. Lim // Fast Software Encryption. – FSE'03, Proceedings. – Springer Verlag, 2003. – P. 247 – 260.

[5] P. M. Hawkes and L. J. O'Connor, "XOR and NON-XOR Differential Probabilities" in Lecture Notes in Computer Science, May 1999, 10.1007/3-540-48910-X_19

[6] J.L. Massey, Nomination of SAFER++ as candidate algorithm for the New European Schemes for Signatures, Integrity, and Encryption (NESSIE). / J.L. Massey, G.H. Khachatrian, M.K. Kuregian. – Primitive submitted to NESSIE by Cylink Corp. – 2000.

Реалізація Методу Виділення Зон Управління в Захищеній Комп'ютерній Мережі

Любомир Пархуць
кафедра захисту інформації
Національний університет "Львівська політехніка"
Львів, Україна
par7@i.ua

Марина Костяк
кафедра захисту інформації
Національний університет "Львівська політехніка"
Львів, Україна
markost@i.ua

Implementation of The Method of Allocation Zones Management in Secure Computer Network

Lubomyr Parkhuts
dept. of Information Protection
Polytechnic National University
Lviv, Ukraine
par7@i.ua

Maryna Kostiak
dept. of Information Protection
Polytechnic National University
Lviv, Ukraine
markost@i.ua

***Анотація*—Робота присвячена вирішенню проблеми що стосується підвищення ефективності функціонування захищених інформаційних мереж спеціального призначення при їх створенні чи модернізації. Запропоновано евристичні алгоритми для реалізації методу виділення зон управління в захищеній комп'ютерній мережі та приведено основні результати досліджень.**

***Abstract*—The work deals with the problem relating to improving the functioning of secure information networks for special purposes in their creation or modernization. A heuristic algorithms to implement the method of allocation zones management in a secure computer network and shows the main results of the research..**

Ключові слова—алгоритм, інформаційна мережа, захист інформації, управління трафіком, оптимізація мережі.

Keywords—algorithm, information network, information security, traffic management, network optimization

I. Вступ

Подання структури захищеної інформаційної мережі у вигляді сукупності і способів взаємодії елементів, що входять до її складу, передбачає задання конфігурації мережі зв'язку, через яку проходить обмін інформацією між елементами мережі. При цьому, конфігурація мережі в основному визначається її топологією, просторовим розташуванням джерел та користувачів інформації, а конфігурація і вид мережі зв'язку є важливою структурною характеристикою її [1].

Для оптимізації структури захищеної інформаційної мережі та мінімізації трафіків зручно використовувати апарат теорії графів [2], який дозволяє реалізувати алгоритм впорядкованого перебору при пошуку оптимальної структури та мінімального трафіку [3, 4].

Із збільшенням розмірності інформаційної мережі при реалізації алгоритмів маршрутизації і управління інтенсивністю потоків виникає ряд труднощів, обумовлених багатьма причинами та специфікою функціонування захищеної інформаційної мережі спеціального призначення. Так, зокрема, при збільшенні розмірності мережі збільшується частка службового трафіку в загальному обсязі мережного трафіку і, отже зменшується продуктивність мережі, оскільки частина запитів користувачів не може бути обслужена через наявність службової інформації.

Маршрутні таблиці, якщо вони містять повну інформацію про те, як досягти будь-якого адресата мережі, можуть виявитися дуже великими, унаслідок чого значно ускладнюється реалізація вузлів комутації. Збільшується час доставки службової інформації і, отже, при виборі маршрутів використовується інформація, яка може в значній мірі не відповідати реальній ситуації, що має місце в мережі в даний момент часу.

Можливим варіантом вирішенням проблеми маршрутизації і управління обсягом потоків в мережі є застосування принципу ієрархічної зонової адресації і маршрутизації [6].

В роботах [5-7] розглянуто питання оптимізації потоків трафіку в інформаційних мережах, запропоновано алгоритми адаптивної маршрутизації в захищеній інформаційній мережі з обмеженим вибором витікаючих каналів зв'язку, ієрархічну зонова адресацію і маршрутизацію, а також розглянуто алгоритми адаптивної внутрішньозонової маршрутизації в захищеній локальній мережі.

Дана робота є логічним продовженням вказаних робіт. В ній розглянуто евристичні алгоритми для реалізації методу виділення зон управління в захищеній інформаційній мережі.

II. Основна частина

Знайти математично оптимальне теоретичне рішення щодо виділення зон управління в захищеній інформаційній мережі при достатньо великій кількості вузлів комутації є дуже непростою задачею, оскільки практично неможливо врахувати всі особливості її функціонування в екстремальних умовах, а також той факт, що в таких умовах необхідно впаховувати особливу пріоритетність окремих видів інформації.

В такому випадку запропоновано застосувати евристичні алгоритими для реалізації методу виділення зон управління в захищеній інформаційній мережі спеціального призначення.

Розглянемо два евристичні алгоритми, що реалізовують запропонований нами метод виділення зон управління [7].

При описі алгоритмів вважатимемо, що орграф, який відображає інформаційну мережу, перетворений в неорієнтований граф, причому довжина кожного з ребер рівна сумі довжин ребер, що сполучають відповідні вершини орграфа, тобто довжина ребра одержуваного графа відображатиме інтенсивність сумарного (в двох напрямах) транзитного трафіку між суміжними вузлами інформаційної мережі.

Введемо такі позначення: N_{∂}^{∂} – максимально допустиме число вузлів комутації (ВК) в зоні при використовуванні децентралізованого методу управління; $N_{ц}^{\partial}$ – максимально допустиме число в зоні при використовуванні централізованого методу управління. Нехай для визначеності $N_{\partial}^{\partial} > N_{ц}^{\partial}$.

Розглянемо послідовний алгоритм виділення зон управління, суть якого полягає в тому, що куски $G_i = (X_i, Y_i)$ графа G_i формуються так, щоб кожний з них в множині Y_{ii} містив ребра з можливо більшою сумарною довжиною.

При цьому отримаємо локальний мінімум сумарної довжини V сполучних ребер.

Послідовний алгоритм складається з таких кроків:

1. Визначити початкову вершину x_i. Як початкова вибирається вершина з найменшим локальним ступенем $p(x_i) = \min\limits_{x_i \in X} p(x_i)$, тобто вершина, що має мінімальну сумарну довжину суміжних з нею ребер. Якщо таких вершин декілька, то вибираємо ту, у якої одне з ребер має найбільшу довжину.

При такому виборі початкової вершини формування зони починається з периферійних ВК, тобто ВК, які найменше пов'язані з іншими.

2. Сформувати черговий кусок розбиття. Включаємо в X_j вершину x_i і всі вершини, суміжні з нею. Позначимо множину вершин, включених в X_j, через Γx_i.

3. Якщо

$$\left(|X_j| = N_{ц}^{\partial} \;\&\; k_{ц} \geq k_{ц}^{(2)}\right) \vee \left(|X_j| = N_{\partial}^{\partial} \;\&\; k_{ц} < k_{ц}^{(2)}\right),$$

то перейти до п.6.

Якщо

$$\left(|X_j| = N_{ц}^{\partial} \;\&\; k_{ц} \geq k_{ц}^{(2)}\right) \vee \left(|X_j| = N_{\partial}^{\partial}\right)$$

то перейти до п.4.

Якщо

$$\left(|X_j| = N_{ц}^{\partial}\right) \vee \left(|X_j| < N_{o}^{\partial} \;\&\; k_{ц} < k_{ц}^{(2)}\right),$$

то перейти до п.5.

4. Видалити "зайві вершини", тобто вершини, сумарна довжина зв'язків яких з рештою вершин в куску G_j мінімальна. Перейти до п.6.

5. Вибрати в G вершину, що задовольняє умові

$$\sigma(x_k) = p(x_k) - a_k = \max\limits_{x_k \in X} G(x_k),$$

де a_k – сумарна довжина ребер, що з'єднують вершину x_k зі всіма вершинами, що не належать G_j. Побудувати множину Γx_k і перейти до п.3.

6. Вилучити кусок G_j, з графа G^*. Після вилучення отримаємо граф

$$G^* = (X^*, Y^*),\ X^* = X \setminus X_1,\ Y^* = Y \setminus Y_1.$$

Якщо для графа G^* :

$$\left(k_{ц} \geq k_{ц}^{(2)} \;\&\; |X^*| \leq N_{ц}^{\partial} + \varepsilon^{\partial}\right) \vee \left(k_{ц} < k_{ц}^{(2)} \;\&\; |X^*| \leq N_{\partial}^{\partial} + \varepsilon^{\partial}\right),$$

то закінчити.

(Тут ε^{∂} позначено допустиме перевищення розміру куска, яке задається для того, щоб не вводити зайву зону управління, коли число вершин, що залишилося, не перевищує ε^{∂}).

В протилежному випадку слід перейти до п.2, використовуючи як початкове граф G^*.

До переваг послідовного алгоритму відноситься його простота, результат можна отримати при невеликих витратах машинного часу. Проте в загальному випадку даний алгоритм може привести до неоптимальних результатів.

Кращі результати можуть бути отримані при розбитті мережі на зони з використанням ітераційних алгоритмів.

Розглянемо ітераційний алгоритм виділення зон управління, суть якого полягає в тому, що спочатку знаходиться (наприклад, за допомогою описаного вище алгоритму) деяке початкове розбиття графа $G = \{G_1^н, \ldots G_i^н\}$.

Потім обчислюється величина, на яку зміниться вага розбиття V при обміні вершинами (або групами вершин) між двома кусками $G_i^н$ і $G_j^н$. Якщо вага розбиття зменшиться, то вершини (групи вершин) міняються місцями.

Процес продовжується до тих пір, поки не будуть виконані всі можливі перестановки. В алгоритмі, розробленому в даному параграфі, проводиться обмін одиничними вершинами.

Визначимо умови, при яких проводиться обмін вершинами між кусками. Розглянемо два куски з розбиття – A і B.

Нехай зовнішня і внутрішня ваги розбиття відповідно рівні

$$E_a = \sum_{y \in B} r_{ay}; \quad I_a = \sum_{y \in B} r_{ax}$$

Позначимо $D_z = E_z - I_r$. При обміні вершинами a і b вага розбиття V змінюється на величину $\Delta V = D_a - 2r_{ab}$. Отже, мінятися місцями повинні лише ті вершини, для яких $\Delta V > 0$.

Крім того, оскільки при обміні вершинами може змінитися значення коефіцієнта централізації зони, необхідно ввести поняття допустимого обміну.

Обмін вершинами a і b назвемо допустимим, якщо при обміні значення коефіцієнта централізації кожного з кусків розбиття не переходить через граничне значення $k_ц^{(2)}$, тобто якщо до обміну виконувалася нерівність $k_ц \le k_ц^{(2)}$, то вона повинна виконуватися і після обміну; аналогічно, якщо до обміну виконувалася нерівність $k_ц > k_ц^{(2)}$, то вона повинна виконуватися і після обміну.

Ітераційний алгоритм складається з таких кроків:

1. Визначити початкове розбиття графа мережі G з використанням послідовного алгоритму.

2. Зафіксувати пару кусків розбиття G_i і G_j. Обчислити D_{x_k} для всіх $x_k \in X_i$ для всіх $x_i \in X_j$.

3. Вибрати пару вершин $x_k \in X_i$, $x_i \in X_j$ таких, що

$$\begin{cases} \Delta V = \max\limits_{(x_k, x_i)} \Delta V, \\ \Delta V > 0. \end{cases}$$

Якщо така пара знайдена і обмін допустимий, то поміняти вершини місцями. В протилежному випадку перейти до п. 5.

4. Сформувати множини $X_i = X_i \setminus x_k$ і $X_j = X_j \setminus x_i$, перейти до п. 3.

5. Якщо є не розглянуті пари G_i і G_j, то перейти до п. 2, якщо таких пар немає, то закінчити.

Даний ітераційний алгоритм дозволяє значно поліпшити результати розбиття графа інформаційної в порівнянні з послідовним алгоритмом. Подальше зменшення ваги розбиття може бути досягнуто ціною значного ускладнення даного алгоритму, що полягає в перестановці груп вершин.

III. Результати дослідженнь

Дослідження запропонованих евристичних алгоритмів для реалізації методу виділення зон управління в захищеній інформаційній мережі було проведено на ряді топологічних структур захищених інформаційних мереж, які генерувалися випадковим чином, з використанням комп'ютерного моделювання.

Результати отриманих результатів дослідження залежно від числа вузлів в мережі N і середньої зв'язаності мережі $\nu = M/N$, де M – число каналів зв'язку (КЗв), приведено на рис.1.

Крива *1* характеризує залежність ваги розбиття від числа вузлів в мережі при розподілі вузлів по зонах випадковим чином; крива *2* – залежність, відповідна послідовному алгоритмі розбиття; крива *3* – залежність, відповідна ітераційному алгоритму розбиття.

На рис.2 приведено графіки залежності відносного зменшення ваги розбиття при використанні розроблених алгоритмів від числа вузлів мережі.

Крива *1* – відповідає ітераційному алгоритму; крива *2* – послідовному алгоритму. Значення Δ обчислюється за формулою $\Delta = (V_c - V_1)/V_c$, де V_c – вага розбиття при випадковому розподілі вузлів інформаційної мережі по зонах; V_i – вага розбиття при використовуванні

одного з алгоритмів розбиття (послідовного при $i = 1$ або ітераційного при $i = 2$).

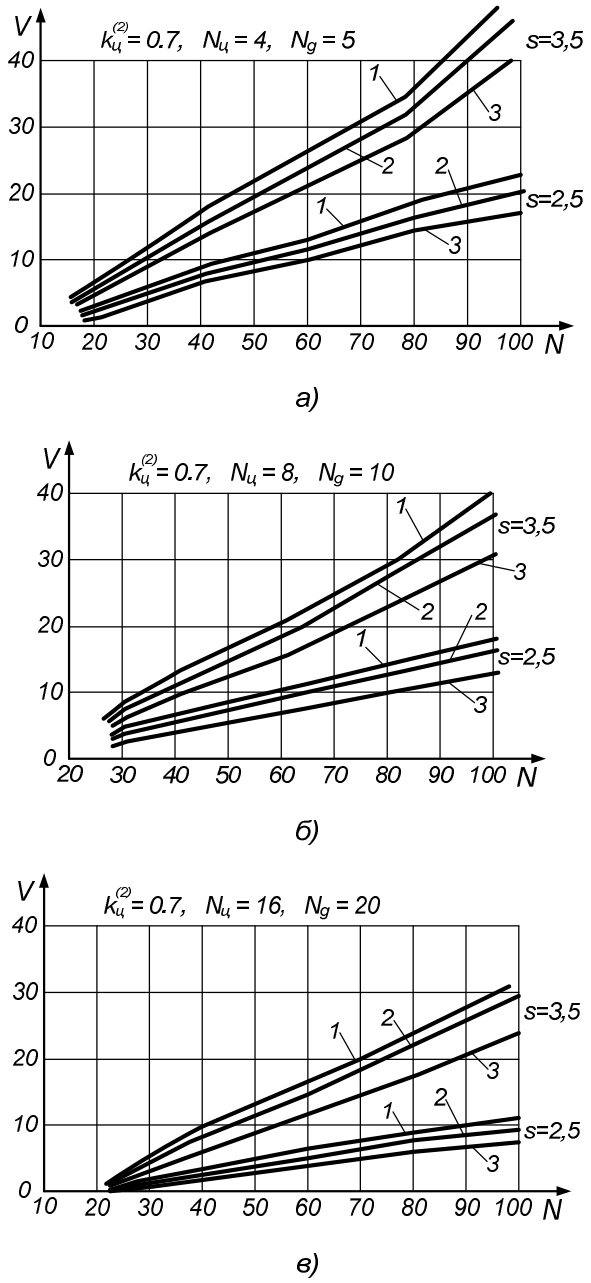


Рис. 1. Результати дослідження розроблених алгоритмів на ряді топологічних структур захищених інформаційних мереж

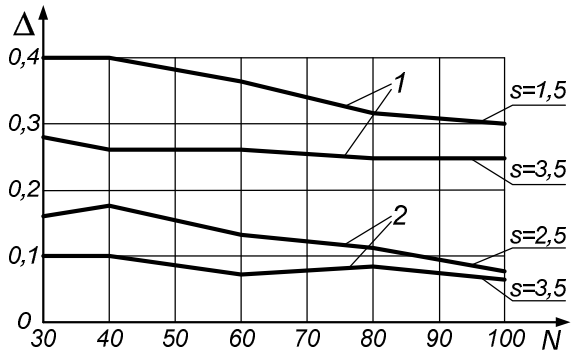


Рис. 2. Залежність відносного зменшення ваги розбиття при використанні розроблених алгоритмів від числа вузлів мережі

IV. Висновки

Запропоновані евристичні алгоритми для реалізації методу виділення зон управління в інформаційній мережі дозволяють оптимізувати обсяг трафіків в захищеній інформаційній мережі, а також збільшити пропускну здатність та швидкодію інформаційної мережі.

Це є важливим фактором для роботи захищених інформаційних мереж спеціального призначення, особливо в екстремальних умовах та при надзвичайних ситуаціях, коли частина структури інформаційної мережі може бути пошкоджена чи зруйнована вніслідок агресивних дій супротивника.

Проведені дослідження та результати комп'ютерного моделювання підтверджують достатню ефективність запропонованих алгоритмів.

Література References

[1] В.А. Кудинов, Л.Т. Пархуць, В.А. Хорошко, Оптимизация структуры информационной сети. Науково-технічний журнал «Захист інформації», № 3, 2004. – с.44-49.

[2] К. Берж, Теория графов и ее применения. М.: Изд-во иностр.лит., 1962. -264с.

[3] И.Р. Конеев, А.В. Беляева, *Информационная безопасность предприятия.* – СПб.: БХВ-Петербург, 2003. –752с.

[4] Дж. Ирвин, Д. Харль, *Передача данных в сетях: инженерный подход.* – СПб.: БХВ-Петербург, 2003. –448с.

[5] М.Ю. Костяк, Л.Т.Пархуць, Задача управління обмеженням інтенсивності потоків в захищеній інформаційній мережі. *Системи обробки інформації.* – Харків: ХУПС. – 2013. – Вып. 3(110), Том 2. – С. 105-107.

[6] В.Б. Дудикевич, М.Ю. Костяк, Л.Т. Пархуць, В.О. Хорошко, Ієрархічна зонова адресація і маршрутизація в захищеній інформаційній мережі. *Вісник Східноукраїнського національного університету імені Володимира Даля*. – 2009. – № 6 (136), Ч.1., – С.143-149.

[7] М.Ю. Костяк, Алгоритми адаптивної внутрішньозонової маршрутизації в захищеній локальній мережі / М.Ю.Костяк, Л.Т.Пархуць // *Інформаційна безпека Східноукраїнського національного університету імені Володимира Даля*. – 2009. – № 1, – С. 38-43.

Оцінки Імовірностей Диференціалів Безключової R-Схеми Блокового Шифрування

Євсюкова Яна, Яковлєв Сергій
кафедра математичних методів захисту інформації
Фізико-технічний інститут
НТУУ "Київський політехнічний інститут ім. Ігоря Сікорського"
Київ, Україна
yanayevsyukova@mail.ru, yasv@rl.kiev.ua

Estimations of Differential Probabilities of Unkeyed R-Scheme of Block Encryption

Yana Yevsyukova, Serhii Yakovliev
Dept. of Mathematical Methods of Information Security
Institute of Physics and Technology
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"
Kyiv, Ukraine
yanayevsyukova@mail.ru, yasv@rl.kiev.ua

***Анотація*—R-схема блокового шифрування є одним з аналогів популярної схеми Фейстеля. У даній роботі одержано аналітичні оцінки для імовірностей диференціалів трираундової безключової R-схеми через відповідні параметри її раундових функцій.**

***Abstract*—Block encryption R-scheme is one of the analog of wide known Feistel scheme. We present analytic bounds for differential probabilities of three-round keyless R-scheme, expressed with corresponding parameters of its round mappings.**

Ключові слова—блокові шифри; R-схема; диференціальний криптоаналіз; легка криптографія.

Keywords—block ciphers; R-scheme; differential cryptanalysis; lightweight cryptography.

I. Вступ

Диференціальний [3] та лінійний [4] криптоаналіз є двома потужними методами аналізу симетричних блокових шифрів. Стійкість до даних методів є обов'язковою вимогою для усіх сучасних алгоритмів шифрування.

Природний спосіб оцінювання стійкості шифрів до диференціального та лінійного криптоаналізу полягає у дослідженні максимальних ймовірностей диференціалів (потенціалів лінійних наближень) шифруючих перетворень, усереднених по усіх можливих ключах. Однак цей підхід не застосовний для випадку ітеративних безключових перетворень, які останнім часом широко використовуються у легкій криптографії для побудови надійних та ефективно обчислюваних нелінійних відображень (наприклад, S-блоків). У цьому випадку для забезпечення стійкості необхідно гарантувати невеликі значення диференціальних імовірностей та лінійних потенціалів перетворення в цілому. Встановлювати ці параметри шляхом безпосередньої перевірки можна лише для перетворень із невеликим розміром блоку. Тому дуже слушними стають аналітичні методи оцінювання криптографічних параметрів ітеративних безключових перетворень через відповідні параметри їх складових елементів.

У роботі [1] Лі та Ванг одержали аналітичні оцінки для диференціальних імовірностей та лінійних потенціалів для трираундової безключової схеми Фейстеля; ці оцінки побудовані на основі значень диференціальних імовірностей та лінійних потенціалів раундових перетворень (S-блоків) схеми Фейстеля. А.Канто та ін. [2] покращили ці оцінки та поширили їх на трираундову схему MISTY.

У даній роботі буде розглянуто ще одну модифікацію схеми Фейстеля – так звану R-схему. Для трираундової безключової R-схеми із певними додатковими умовами будуть одержані оцінки диференціальної імовірності через відповідні параметри її раундових функцій.

II. Необхідні терміни та позначення

У роботі розглядаються S-блоки, що мають однакову кількість вхідних та вихідних бітів. Стійкість до диференціального та лінійного криптоаналізу визначається максимальним значенням у таблиці розподілів диференціалів (таблиці лінійних апроксимацій відповідно) [7]. Визначимо ці два параметри формально.

Нехай V_n – множина всіх n-бітових векторів і F – це відображення з V_n на V_n. Для будь-якої пари різниць (a,b) з V_n^2 визначимо множину

$$D_F(a \to b) = \left\{ x \in F_n^2 \mid F(x \oplus a) \oplus F(x) = b \right\}.$$

Комірка з індексом (a,b) в таблиці розподілів диференціалів F тоді відповідає потужності множини $D_F(a \to b)$; позначимо її як $\delta_F(a,b)$.

Диференціальна рівномірність F – це величина

$$\delta(F) = max_{a \neq 0, b}\, \delta F(a,b).$$

Максимальна імовірність диференціалу *MDP* пов'язана із диференціальною рівномірністю очевидним чином: $MDP(F) = \delta(F)/2^n$. Для будь-яких перетворень F справедлива оцінка $\delta(F) \geq 2$. Функції F, для яких виконується рівність, називаються майже досконалими нелінійними функціями (almost perfect nonlinear mappings, APN). [8]

Перетворення Уолша відображення F – це функція

$$\lambda : V_n^2 \times V_n^2 \to Z$$

$$(a,b) \mapsto \lambda_F(a,b) = \sum_{x \in V} (-1)^{b \cdot F(x) \oplus a \cdot x},$$

де крапкою позначено скалярний добуток бітових векторів. *Нелінійність* F – це величина

$$\alpha(F) = \max_{a,b \in V_n, b \neq 0} \left| \lambda_F(a,b) \right|$$

Дійсно, з точністю до множника 2^n нелінійність відповідає імовірності неспівпадіння значення функції F та її найкращої лінійної апроксимації:

$$Pr_X\left[b \cdot F(x) + a \cdot X = 1\right] = \frac{1}{2^n}\left(2^{n-1} - \frac{1}{2}\sum_{x \in F_n^2} (-1)^{b \cdot F(x) \oplus a \cdot x}\right) =$$

$$= \frac{1}{2}\left(1 - \frac{\lambda_F(a,b)}{2^n}\right)$$

Варто зауважити, що для будь-якої фіксованої вихідної маски $b \in V_n$ функція $a \mapsto \lambda_F(a,b)$ відповідає перетворенню Уолша n-змінної булевої функції $b \cdot F(x)$, що є лінійною комбінацією координатних функцій F.

R-схема блокового шифрування є одним з аналогів широко розповсюдженої схеми Фейстеля. Діаграму обчислення R-схеми представлено на рис. 1. Властивості R-схеми як шифруючого перетворення, кожен раунд якого параметризовано ключем, разом із властивостями деяких інших фейстель-подібних схем були досліджені у [5, 6] – зокрема, у [5] були одержані аналітичні оцінки для імовірностей диференціалів та лінійних потенціалів. У даній роботі розглядаються три раундові R-схеми без ключів.

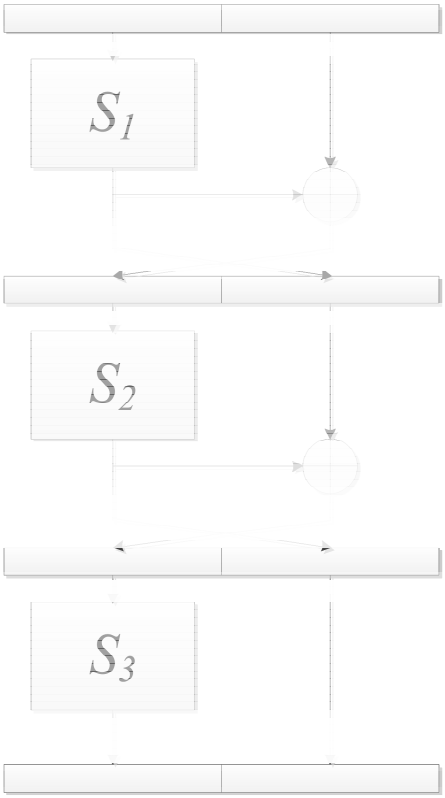


Рис. 1. Діаграма трираундової R-схеми блокового шифрування. В останньому раунді не використовується додавання частин блоку між собою.

III. Диференціальна рівномірність для трираундової R-схеми

Одержані оцінки диференціальної рівномірності для трираундової R-схеми базуються на розгляданні окремих різниць, для яких вхідна різниця одного з раундових S-блоків дорівнює нулю.

Теорема 1. Нехай S_1, S_2 та S_3 – це три n-бітні S-блоки (не обов'язково різні), F – це $2n$-бітова функція, побудована за структурою трираундової R-схеми із відображеннями S_1, S_2 та S_3 в якості раундових перетворень. Тоді для будь-яких a, b та c з V_n маємо:

1) якщо S_1 – бієктивний, то

$$\delta_F(0 \,||\, a,b \,||\, c) = \delta_{S_2}(a, b \oplus c) \times \delta_{S_3}(b \oplus c, c)$$

2) якщо S_2 – бієктивний, то

$$\delta_F(b \,||\, a,c \,||\, c) = \delta_{S_1}(b,a) \times \delta_{S_3}(a,c)$$

3) якщо S_3 – бієктивний, то

$$\delta_F(a \,||\, b,c \,||\, 0) = \delta_{S_1}(a,c) \times \delta_{S_2}(b \oplus c, c).$$

Доведення: Для вектору $x \in V_n^2$ позначимо через x_L та x_R його ліву та праву частини відповідно. Розглянемо проходження різниці входів x та $x \oplus (0 \,||\, a)$ через функцію F таким чином, щоб одержати на виході різницю $(b \,||\, c)$.

Випадок 1:

Перший раунд:

$$(x_R \oplus S_1(x_L), S_1(x_L)) \oplus (x_R \oplus a \oplus S_1(x_L), S_1(x_L)) = [a,0].$$

Другий раунд:

$$\begin{aligned}
&(S_1(x_L) \oplus S_2(x_R \oplus S_1(x_L)), S_2(x_R \oplus S_1(x_L))) \oplus \\
&\oplus (S_1(x_L) \oplus S_2(x_R \oplus a \oplus S_1(x_L)), S_2(x_R \oplus a \oplus \\
&\oplus S_1(x_L))) = \\
&= (S_2(x_R \oplus S_1(x_L)) \oplus S_2(x_R \oplus a \oplus S_1(x_L)), \\
&S_2(x_R \oplus S_1(x_L)) \oplus S_2(x_R \oplus a \oplus S_1(x_L))) = \\
&= [b \oplus c, b \oplus c]
\end{aligned}$$

Третій раунд:

$$\begin{aligned}
&(S_2(x_R \oplus S_1(x_L)) \oplus S_3(S_1(x_L) \oplus S_2(x_R \oplus \\
&\oplus S_1(x_L))), S_3(S_1(x_L) \oplus S_2(x_R \oplus S_1(x_L)))) \oplus \\
&\oplus (S_2(x_R \oplus a \oplus S_1(x_L)) \oplus S_3(S_1(x_L) \oplus S_2(x_R \oplus \\
&\oplus a \oplus S_1(x_L))), S_3(S_1(x_L) \oplus S_2(x_R \oplus a \oplus S_1(x_L)))) = \\
&= (S_2(x_R \oplus S_1(x_L)) \oplus S_2(x_R \oplus a \oplus S_1(x_L)) \oplus \\
&\oplus S_3(S_1(x_L) \oplus S_2(x_R \oplus S_1(x_L))) \oplus S_3(S_1(x_L) \oplus \\
&\oplus S_2(x_R \oplus a \oplus S_1(x_L))), S_3(S_1(x_L) \oplus S_2(x_R \oplus \\
&\oplus S_1(x_L))) \oplus S_3(S_1(x_L) \oplus S_2(x_R \oplus a \oplus S_1(x_L)))) = [b,c]
\end{aligned}$$

Таким чином, вектор $x = (x_L, x_R)$ задовольняє рівнянню

$$F(x_L \,||\, x_R) \oplus F(x_L \,||\, (x_R \oplus a)) = b \,||\, c$$

тоді і тільки тоді, коли:

$$\begin{aligned}
&S_2(x_R \oplus S_1(x_L)) \oplus S_2(x_R \oplus a \oplus S_1(x_L)) \oplus S_3(S_1(x_L) \oplus \\
&\oplus S_2(x_R \oplus S_1(x_L))) \oplus S_3(S_1(x_L) \oplus S_2(x_R \oplus a \oplus S_1(x_L))) = \\
&= b, \\
&S_3(S_1(x_L) \oplus S_2(x_R \oplus S_1(x_L))) \oplus S_3(S_1(x_L) \oplus \\
&\oplus S_2(x_R \oplus a \oplus S_1(x_L))) = c,
\end{aligned}$$

тобто

$$\begin{aligned}
&S_2(x_R \oplus S_1(x_L)) \oplus S_2(x_R \oplus a \oplus S_1(x_L)) = b \oplus c, \\
&S_3(S_1(x_L) \oplus S_2(x_R \oplus S_1(x_L))) \oplus S_3(S_1(x_L) \oplus \\
&\oplus S_2(x_R \oplus a \oplus S_1(x_L))) = c,
\end{aligned}$$

що рівносильне тому, що

$$y_L \in D_{S_2}(a \to b \oplus c),$$

$$y_R \in S_2(y_L) \oplus D_{S_3}(b \oplus c \to c),$$

якщо S_1 – бієктивний.

Таким чином доведено, що існує $\delta_{S_2}(a, b \oplus c)$ значень x_R та для кожного з них $\delta_{S_3}(b \oplus c, c)$ значень x_L, таких що x досягає різниці.

Випадок 2:

Розглянемо проходження різниці входів x та $x \oplus (b \,||\, a)$ через функцію F таким чином, щоб одержати на виході різницю $(c \,||\, c)$.

Аналогічно, бачимо, що вектор $x = (x_L, x_R)$ задовольняє рівнянню

$$F(x_L \,||\, x_R) \oplus F((x_L \oplus b) \,||\, (x_R \oplus a)) = c \,||\, c$$

тоді і тільки тоді, коли:

$$\begin{aligned}
&S_1(x_L) \oplus S_1(x_L \oplus b) = a, \\
&S_3(S_1(x_L) \oplus S_2(x_R \oplus S_1(x_L))) \oplus S_3(S_1(x_L \oplus b) \oplus \\
&\oplus S_2(x_R \oplus a \oplus S_1(x_L \oplus b))) = c,
\end{aligned}$$

що рівносильно системі

$$\begin{aligned}
&x_L \in D_{S_1}(b \to a), \\
&x_R \in S_1(x_L) \oplus D_{S_3}(a \to c),
\end{aligned}$$

якщо S_2 – бієктивний.

Випадок 3:

Розглянемо проходження різниці входів x та $x \oplus (a \,||\, b)$ через функцію F таким чином, щоб одержати на виході різницю $(c \,||\, 0)$

Аналогічно, бачимо, що вектор $x = (x_L, x_R)$ задовольняє рівнянню

$$F(x_L \,||\, x_R) \oplus F((x_L \oplus a) \,||\, (x_R \oplus b)) = c \,||\, 0$$

тоді і тільки тоді, коли:

$$S_1(x_L) \oplus S_1(x_L \oplus a) = c,$$
$$S_2(x_R \oplus S_1(x_L)) \oplus S_2(x_R \oplus b \oplus S_1(x_L \oplus a)) = c,$$

що рівносильно

$$x_L \in D_{S_1}(a \to c),$$
$$x_R \in S_1(x_L) \oplus D_{S_2}(b \oplus c \to c),$$

якщо S_3 – бієктивний.

Основний результат щодо оцінки диференціальної рівномірності без ключової R-схеми подамо у вигляді такої теореми.

Теорема 2. Нехай S_1, S_2 та S_3 – це три n-бітні S-блоки (не обов'язково різні), F – це $2n$-бітова функція, побудована за структурою трираундової R-схеми із відображеннями S_1, S_2 та S_3 в якості раундових перетворень. Тоді,

$$\delta(F) \geq \delta(S_2) \max(\delta_{\min}(S_3), \delta_{\min}(S_1)),$$

де $\delta_{\min}(S) = \min_{a \neq 0} \max_b \delta_S(a,b)$. Зокрема, якщо S_2 є перестановкою, то

$$\delta(F) \geq \max_{i \neq 2, j \neq 2, i} \max(\delta(S_i)\delta_{\min}(S_j), \delta(S_i)\delta_{\min}(S_2^{-1})) .$$

Якщо S_2 не є перестановкою, то $\delta(F) \geq 2^{n+1}$.

Доведення. Даний результат є прямим наслідком Теореми 1. Доведемо наведену границю для першого випадку Теореми 1; інші випадки доводяться аналогічно.

Розглянемо диференціал (α, β), на якому S_2 досягає диференціальної рівномірності: $\delta(S_2) = \delta_{S_2}(\alpha, \beta)$. Оберемо $a = \alpha$ та якщо $b = \beta \oplus c$; тоді для будь-яких $c \in V_n$

$$\delta_F(0 \,||\, \alpha, (\beta \oplus c) \,||\, (\beta \oplus b)) = \delta(S_2) \times \delta_{S_3}(\beta, \beta \oplus b)$$

Якщо $b \neq \beta \oplus c$, то добуток може бути більше або рівний за $\delta(F)$, або менший. Якщо ж добуток менший, то можна знайти таке значення b, щоб максимізувати його.

Тоді можемо вибрати для b значення, яке максимізує $\delta_{S_3}(\beta, \beta \oplus b)$. Це значення завжди більше або дорівнює $\delta_{\min}(S_3)$.

Так само розглянемо диференціал (α, β), на якому S_3 досягає диференціальної рівномірності: $\delta(S_3) = \delta_{S_3}(\alpha, \beta)$. Оберемо $\beta = c$ та $b = \alpha \oplus \beta$; тоді для довільного $a \in V_n$:

$$\delta_F(0 \,||\, \alpha, (\beta \oplus \alpha) \,||\, \beta) = \delta_{S_2}(a, \alpha) \times \delta(S_3) .$$

Можемо вибрати для a значення, яке максимізує $\delta_{S_2}(a, \alpha)$. Це значення завжди більше або дорівнює $\delta_{\min}(S_2^{-1})$, коли S_2 бієктивний.

Припустимо тепер, що S_2 не бієктивний. Це означає, що існує деякий ненульовий $a \in V_n$ такий, що $\delta_{S_2}(a,0) \geq 0$. Тоді з першого пункту Твердження 1 випливає, що при $b = c = 0$ співвідношення

$$F(x_L \,||\, x_R) \oplus F(x_L \,||\, (x_R \oplus a)) = (0,0)$$

має $\delta_{S_2}(a,0) \times \delta_{S_3}(0,0) \geq 2 \times 2^n = 2^{n+1}$ розв'язків у V_n^2.

Результати, аналогічні твердженням Теорем 1 та 2, також одержуються для лінійних потенціалів безключової трираундової R-схеми, що дозволяє виводити безпосередні оцінки стійкості до лінійного криптоаналізу.

IV. Висновки

У даній роботі було проведено аналіз безключової R-схеми блокового шифрування. Одержано аналітичні оцінки для імовірностей диференціалів R-схеми, виражені через відповідні параметри її раундових перетворень (S-блоків).

Дані результати можуть бути використані для побудови надійних алгоритмів легкої криптографії.

Література References

[1] Li, Y., Wang, M.: Constructing S-boxes for Lightweight Cryptography with Feistel Structure. In: Cryptographic Hardware and Embedded Systems – CHES 2014, LNCS, vol. 8731, pp. 127–146. Springer (2014).

[2] Anne Canteaut, Sebastien Duval, and Gaetan Leurent "Construction of Lightweight S-Boxes using Feistel and MISTY structures (Full Version)", Inria, project-team SECRET, France [Online.] – http://eprint.iacr.org/2015/711.pdf

[3] Matsui, M.: Linear cryptanalysis method for DES cipher. In: Advances in Cryptology—EUROCRYPT'93. LNCS, vol. 765, pp. 386–397. Springer (1994)

[4] Biham, E., Shamir, A., "Differential Cryptanalysis of DES-like Cryptosystems", in. Journal of Cryptology. – 1991. – V. 4. – № 1. – P. 3 – 72.

[5] Y. Kaneko, F. Sano, K. Sakurai, "On provable security against differential and linear cryptanalysis in generalized Feistel ciphers with multiple random functions" // Proc. of SAC'97. – Springer, 1997.

[6] Henri Gilbert and Marine Minier, "New Results on the Pseudorandomness of Some Blockcipher Constructions" // FSE 2001. – LNCS vol. 2355. – Berlin: Springer, 2002. – pp. 248-266.

[7] Heys Howard M. "A Tutorial on Linear and Differential Cryptanalysis" [Online.] – Available at http://www.engr.mun.ca/~howard/PAPERS/ldc_tutorial.pdf

Ланцюги Підпросторів Калина-Подібних Шифрів

Коляда Марія, Яковлєв Сергій
кафедра математичних методів захисту інформації
Фізико-технічний інститут
НТУУ «Київський політехнічний інститут імені І.Сікорського»
Київ, Україна
kolyadamariya1710@gmail.com, yasv@rl.kiev.ua

Subspace Trails of Kalyna-Like Block Ciphers

Mariya Kolyada, Serhii Yakovliev
dept. of Mathematical Methods of Information Security
Institute of Physics and Technology
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"
Kyiv, Ukraine
kolyadamariya1710@gmail.com, yasv@rl.kiev.ua

***Анотація*—У даній роботі наведено типи та приклади ланцюгів підпросторів для одного, двох, трьох та чотирьох раундів «Калина»-подібних шифрів.**

***Abstract*—We present types and examples of subspace trails for one, two, three and four rounds of "Kalyna"-type block ciphers.**

Ключові слова—симетрична криптографія; блокові шифри; «Калина»; ланцюги підпросторів.

Keywords—symmetric cryptography; block ciphers; Kalyna cipher; subspace trails

I. Вступ

Наявність у блоковому шифрі характеристик із нерівноімовірним розподілом дозволяє будувати ефективні стастистичні атаки відновлення раундових ключів або ключа шифрування загалом. Пошук та аналіз поведінки таких нерівномірних статистик є постійно актуальною криптографічною задачею.

Дослідження ланцюгів підпросторів було вперше запропоноване у 2011 році для криптоаналізу шифру PRINTcipher [3], після чого ідея такого аналізу була ефективно застосована для шифру AES [1]. Для проведення такого аналізу у [1] запропонована техніка використання спеціальним чином підібраних вхідних даних, поведінку яких можна спрогнозувати після декількох раундів шифрування; при цьому від шифру не вимагається наявність спеціальних симетрій або констант.

У даній роботі ми розглянемо ланцюги підпросторів, які можна побудувати для блокових шифрів із структурою алгоритму шифрування «Калина» (ДСТУ 7624:2014 [2]), та покажемо, як можна застосувати техніку аналізу ланцюгів підпросторів для «Калина»-подібних шифрів.

II. Короткий опис блокового шифру «Калина»

Симетричний блоковий шифр «Калина» [3] був стандартизований як національний стандарт ДСТУ 7624 у 2014 році та введений у дію з 2015 року [4] (зауважимо, що попередня опублікована версія шифру «Калина» під такою ж назвою має суттєві структурні відмінності [5]). Структура шифру «Калина» подібна до структури шифру *Rijndael*, але орієнтована на 64-бітні обчислювальні архітектури. Кількість раундів залежить від довжини відкритого тексту та довжини ключа. Відкритий текст подається у вигляді матриці розміром a×8, де a∈{2,4,8}. Базові перетворення для шифрування:

$$E_{l,k}^{K} = AddKey^{K'} \circ MixColumns \circ ShiftRows \circ SubBytes \circ$$
$$\circ \prod_{i=1}^{t-1} (AddKey^{Ki} \circ MixColumns \circ ShiftRows \circ SubBytes) \circ$$
$$\circ AddKey^{K0'}$$

де l – розмір внутрішнього стану блокового шифру (у бітах), K – ключ шифрування, k – довжина ключа шифрування (у бітах), $AddKey^{K'}$ – функція додавання раундового ключа K за модулем 2^{64}, *MixColumns* – лінійне перетворення (множення матриці лінійного перетворення на матрицю внутрішнього стану над скінченним полем), *ShiftRows* – перестановка елементів $g_{i,j} \in GF(2^8)$ внутрішнього стану (циклічний зсув рядків вправо при матричному поданні), *SubBytes* (*SB*) – шар нелінійного бієктивного відображення, який виконує обробку векторів,

заданих над V_8 (байтова підстановка), $AddKey^K$ функція додавання циклового ключа K за модулем 2.

Позначимо через R одне раундове перетворення, тобто послідовне виконання процедур *SubBytes*, *ShiftRows* та *MixColumns*, а також додавання із ключем. Через $R^{(i)}$ будемо позначати процедуру, яка складається з виконання послідовних i раундів (включно із додаванням проміжних раундових ключів).

Процедура *ShiftRows* (*SR*) виконує циклічний зсув вправо рядків матриці стану $g_{i,j} \in GF(2^8)$. Кількість елементів зсуву залежить від номеру рядку $i \in \{0,1,\ldots,7\}$, розміру блоку $l \in \{128\ 256,\ 512\}$, та обчислюється за формулою $\delta_i = \left|\frac{i*l}{512}\right|$.

Процедура *MixColumns* (*MC*) виконує множення кожного стовпчику матриці стану на спеціально підібрану матрицю. Кожен елемент $g_{i,j}$, матриці внутрішнього стану $G=(g_{i,j})$ розглядається як елемент скінченного поля $GF(2^8)$, яке утворене незвідним поліномом $\vartheta(x) = x^8 + x^4 + x^3 + x^2 + 1$. Відповідно, кожен елемент результуючої матриці стану $W=(w_{i,j})$ одержується як результат множення векторів довжини 8 над скінченним полем $GF(2^8)$ за формулою $w_{i,j} = v >>> i \otimes G_j$, де υ = (0x01, 0x01, 0x05, 0x01, 0x08, 0x06, 0x07, 0x04) – вектор, що утворює циркулянтну матрицю МДР-коду і складається з послідовності байтових констант у шістнадцятковому поданні, які інтерпретуються як елементи поля $GF(2^8)$, при цьому циклічний зсув виконується відносно елементів вектора над скінченним полем.

Більш детальну інформацію про структуру та особливості блокового шифру ДСТУ 7624:2014 «Калина» можна одержати у [3, 4].

III. Ланцюги підпросторів ітеративних блокових шифрів

Нехай F – раундова функція в ітеративному блочному шифрі:

$$E_K(m) = k_n \oplus F(\ldots F(k_1 \oplus F(k_0 \oplus m))),$$

де k_i – раундові ключі, отримані з основного ключа K за допомогою певного ключового розкладу. Вхідні повідомлення m розглядаються як бітові вектори із лінійного простору всіх бітових векторів відповідної довжини. Підпростір V лінійного простору вхідних повідомлень є інваріантним відносно F, якщо $F(V) = V$; однак наявність таких підпросторів для сучасних блокових шифрів є малоімовірною.

Нехай для підпростору V існують класи суміжності $V \oplus a$ та $V \oplus a'$ такі, що $F(V \oplus a) = V \oplus a'$; тоді, якщо раундовий ключ K міститься в $V \oplus (a \oplus a')$, то $F(V \oplus a) \oplus K = V \oplus a$ і клас суміжності $V \oplus a$ є інваріантним відносно F. Якщо для довільного a, існує унікальне b таке, що $F(V \oplus a) \oplus K = V \oplus b$, то будемо казати, що підпростір V *зберігає інваріантність відносно* F. Це значить, що для довільного початкового підпростору $V \oplus a$ ми можемо поставити в відповідність інший підпростір $V \oplus b$, де b залежить від a та від раундового ключа. У більш загальному випадку ми розглядаємо пару підпросторів V_1 та V_2 таких, що для довільного вектору a існує унікальний вектор b (який залежить від a та ключа) повинне виконуватись співвідношення $F(V_1 \oplus a) \oplus K \subseteq V_2 \oplus b$, тобто F переводить кожен клас суміжності у якийсь інший клас суміжності.

Ланцюгом підпросторів довжини r назвемо простий кортеж з $r+1$ підпросторів $(V_1, V_2, \ldots, V_{r+1})$, для яких виконуються співвідношення

$$F(V_i \oplus a_i) \oplus K \subseteq V_{i+1} \oplus a_{i+1}.$$

Позначимо через $E=\{e_{0,0},\ldots,e_{a,8}\}$, $a \in \{2,4,8\}$, простір початкових станів шифру «Калина», де $e_{i,j}$ – окремі байти (8-бітові рядки), a – кількість стовпчиків. Визначимо чотири сімейства підпросторів E:

діагональний простір,

інверсно-діагональний простір,

стовпчиковий простір,

змішаний простір.

У шифрі «Калина» передбачено 3 різні форми вхідних даних, в залежності від їх довжини; опис підпросторів буде наведено для всіх можливих довжин блоків.

Стовпчиковий простір C_i визначимо як $C_i = \langle e_{0,i}, e_{1,i}, e_{2,i}, e_{3,i}, e_{4,i}, e_{5,i}, e_{6,i}, e_{7,i} \rangle$. Наприклад, C_0 для випадку 8*8 буде мати вид, наведений на рис. 1:

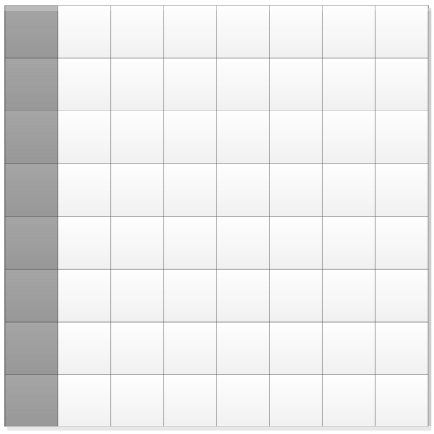


Рис. 1. Схематичний вид простору C_0 розміром 8*8. Сірим позначено ненульові координати елементів простору.

Діагональний простір D_i визначимо таким чином:

Для випадку 8x8: $D_i = SR^{-1}(C_i) = \langle e_{7,i}, e_{6,i+1}, e_{5,i+2}, e_{4,i+3}, e_{3,i+4}, e_{2,i+5}, e_{1,i+6}, e_{0,i+7} \rangle$, де індекс $i+j$ обчислюється за модулем 8, $i \in \{0,\ldots,7\}$. Наприклад, D_0 буде мати вид, наведений на рис. 2:

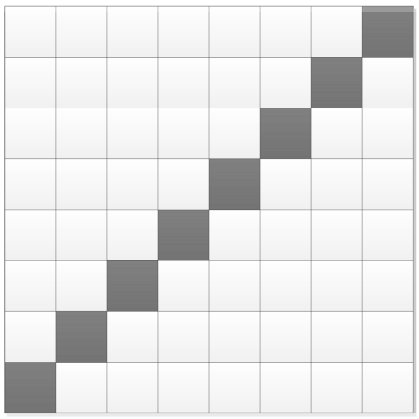


Рис. 2. Схематичний вид простору D_0 розміром 8*8. Сірим позначено ненульові координати елементів простору.

Для випадку 4x8: *:* $D_i = SR^{-1}(C_i) = \langle e_{7,i}, e_{6,i}, e_{5,i+1}, e_{4,i+1}, e_{3,i+2}, e_{2,i+2}, e_{1,i+3}, e_{0,i+3}\rangle$, де індекс $i+j$ обчислюється за модулем 4, $i \in \{0,1,2,3\}$.

Для випадку 2x8: $D_i = SR^{-1}(C_i) = \langle e_{7,i}, e_{6,i}, e_{5,i}, e_{4,i}, e_{3,i+1}, e_{2,i+1}, e_{1,i+1}, e_{0,i+1}\rangle$, де індекс $i+j$ обчислюється за модулем 2, $i \in \{0,1\}$.

Інверсно-діагональний простір ID_i визначимо таким чином:

Для випадку 8*8*: ID_i = SR(C_i) = <e_{0,i}, e_{1,i+1}, e_{2,i+2}, e_{3,i+3}, e_{4,i+4}, e_{5,i+5}, e_{6,i+6}, e_{7,i+7}>* ,де індекс $i+j$ обчислюється за модулем 8, $i \in \{0,\ldots,7\}$. Наприклад, ID_0 буде мати вид, наведений на рис. 3:

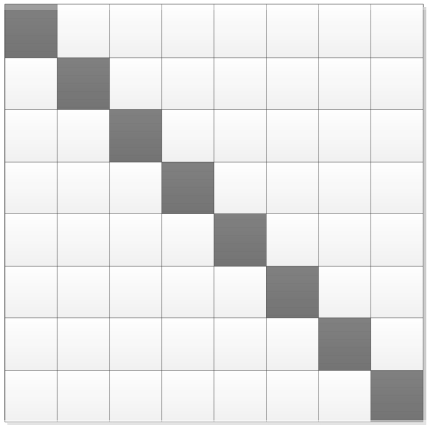


Рис. 3. Схематичний вид простору ID_0 розміром 8*8. Сірим позначено ненульові координати елементів простору.

Для випадку 4*8*: ID_i = SR(C_i) = <e_{0,i}, e_{1,i}, e_{2,i+1}, e_{3,i+1}, e_{4,i+2}, e_{5,i+2}, e_{6,i+3}, e_{7,i+3}>* ,де індекс $i+j$ розраховується за модулем 4, $i \in \{0,1,2,3\}$.

Для випадку 2*8*: ID_i = SR(C_i) = <e_{0,i}, e_{1,i}, e_{2,i}, e_{3,i}, e_{4,i+1}, e_{5,i+1}, e_{6,i+1}, e_{7,i+1}>* ,де індекс $i+j$ розраховується за модулем 2, $i \in \{0,1\}$.

Змішаний простір M_i визначимо таким чином:

$$M_i = MC(ID_i).$$

IV. Ланцюги підпросторів для одного та двох раундів шифру «Калина»

Побудуємо ланцюги підпросторів для двох, трьох та чотирьох раундів шифру «Калина-512» (вхідні дані трактуються як матриця 8x8). Для шифрів «Калина-256» та «Калина-128» (вхідні дані трактуються як матриці 4x8 та 2x8 відповідно) існують аналогічні ланцюги, які описуються таким само чином.

Опишемо деякі ланцюги підпросторів для одного раунду шифрування.

1) Нехай $I \subseteq \{0,1,2,3,4,5,6,7\}$, де $0 < |I| < 8$ та $a \in D_I^{\perp}$, тоді існує унікальне $b \in C_I^{\perp}$, таке що

$$R_K(D_I \oplus a) = C_I \oplus b .$$

2) Нехай $I \subseteq \{0,1,2,3,4,5,6,7\}$, де $0<|I|<8$, та $a \in C_I^{\perp}$, тоді існує унікальне $b \in M_I^{\perp}$, таке, що

$$R_K(C_I \oplus a) = M_I \oplus b .$$

Процес проходження підпросторів, які розглядались у випадках 1) та 2), наведено на рис. 4.

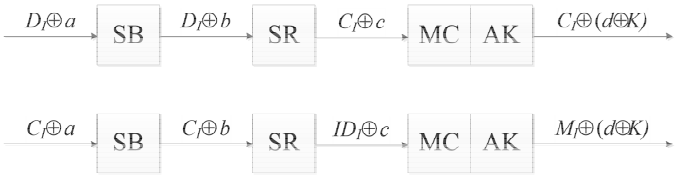


Рис. 4. Візуальне представлення ланцюгів підпросторів для одного раунду.

Нехай $R^{(2)}$ визначає процедуру зашифрування двох раундів «Калини» з раундовими ключами K_1, K_2. Опишемо ланцюги підпросторів для них.

3) Якщо відкритий текст належав класу суміжності діагонального підпростору, то результат його шифрування буде належати класу суміжності змішаного підпростору. Зокрема, після 2-х раундів шифрування з фіксованим ключем отримаємо:

$$Pr(R^{(2)}(u) \oplus R^{(2)}(v) \in M_I \mid u \oplus v \in D_I) = 1 \ \text{™}\sum u \neq v .$$

4) Якщо два відкритих тексти належать різним класам суміжності діагонального простору, то результат їх шифрування буде належати різним класам суміжності змішаного простору. Інакше

$$Pr(R^{(2)}(u) \oplus R^{(2)}(v) \in M_I \mid u \oplus v \notin D_I) = 0 , \ \text{™}\sum u \neq v .$$

5) З відкритого тексту, який належить класу суміжності діагонального простору, неможливо за два раунди

шифрування одержати шифртекст із класу суміжності діагонального простору:

$$Pr(R^{(2)}(u) \oplus R^{(2)}(v) \in D_J \mid u \oplus v \in D_I) = 0 , \ ™\sum u \neq v .$$

6) З відкритого тексту, який належить класу суміжності змішаного простору, неможливо за два раунди шифрування одержати шифртекст із класу суміжності змішаного простору:

$$Pr(R^{(2)}(u) \oplus R^{(2)}(v) \in M_J \mid u \oplus v \in M_I) = 0 , \text{ де } u \neq v .$$

Ці властивості використовуються як відмінна риса для криптоаналізу двох раундів шифрування. Візьмемо два відкритих тексти з D_I, з $I \subseteq \{0,1,2,3,4,5,6,7\}$, де $0 < |I| < 8$; з імовірністю 1 після двох раундів вони потраплять у підпростір M_I. Якщо замість «Калини» застосувати до цих текстів випадкову перестановку (ідеальний шифр), то ймовірність того, що вони потраплять в однаковий підпростір M_I дорівнює $(2^8)^{-64+8*|I|}$. Таким чином, достатньо однієї пари, щоб відрізнити випадкову перестановку від двох раундів шифрування «Калиною».

V. Ланцюги підпросторів для трьох та чотирьох раундів шифру «Калина»

Опишемо ланцюги підпросторів для трьох раундів шифрування «Калини».

Для довільних M_I та C_J маємо:

$$Pr(x \in C_j | x \in M_I) = (2^8)^{-8|I|+|I||J|}.$$

Відповідно, клас суміжності простору M_I можна представити як об'єднання класів суміжності простору C_J:

$$M_I \oplus a = \bigcup_{x \in M_I \oplus a \setminus C_J} C_J \oplus x .$$

Відмітимо, що кількість векторів $x \in M_I \oplus a \setminus C_J$ точно дорівнює $(2^8)^{8|I|-|I||J|}$.

Розглянемо два елементи з одного класу суміжності D_I. Після двох раундів они будуть належати одному класу суміжності простору M_I. Клас суміжності M_I може бути представлений у вигляді об'єднання $N = (2^8)^{8|I|-|I||J|}$ класів суміжності $A_1, A_2, \ldots, A_N$ простору C_J. Імовірність, що ці два елементи належать одному класу суміжності C_J після двох раундів шифрування, дорівнює $(2^8)^{-8|I|+|I||J|}$. Відповідно, після третього раунду шифрування ці два елементи перейдуть у один клас суміжності змішаного простору (див. рис. 5).

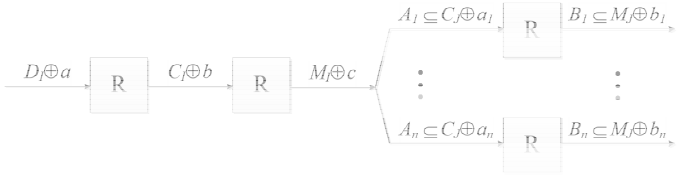


Рис. 5. Візуальне представлення ланцюгів підпросторів для трьох раундів шифрування.

Позначимо через p_1 ймовірність того, що випадкових відкритих тексти після трьох раундів шифрування належатимуть одному класу суміжності простору M_J, а через p_2 – ймовірність того, що у один клас суміжності M_J після трьох раундів потраплять два відкритих тексти, які належать одному класу суміжності простору D_I. Дані ймовірності обчислюються за формулами

$$p_1 = C_8^{|J|} (2^8)^{-64+8*|J|}, \ p_2 = C_8^{|J|} (2^8)^{-8*|I|+|I||J|}.$$

Легко помітити, що ймовірність одержати колізію у другому варианті вища, ніж при випадковому виборі. Зокрема, для $|J|=7$ та $|I|=1$ маємо $p_2 = 2^{-5}$, в той час коли $p_1 = 2^{-61}$. Звернемо увагу, що I та J підкоряються умові: $0<|I|+|J|\leq 7$.

Для чотирьох раундів можна побудувати ланцюги підпросторів, використовуючи просту композицію властивостей ланцюгів підпросторів для двох раундів, описану у пунктах 3) та 6) попереднього розділу. Маємо:

$$Pr(R^{(4)}(u) \oplus R^{(4)}(v) \in M_J \mid u \oplus v \in D_I) = 0 , \text{ де } u \neq v .$$

Іншими словами, за чотири раунди шифрування тексти, які належать одному класу суміжності діагонального простору, не можуть потрапити у один клас суміжності змішаного простору. Це також дає змогу побудувати розпізнавач для чотирираундової «Калини». Інший шлях для побудови чотирираундових ланцюгів – розбиття класу суміжності змішаного простору на сукупність класів суміжності діагональних просторів та застосування техніки, аналогічної для побудови трираундових ланцюгів.

VI. Висновки

У даній роботі було побудовано декілька ланцюгів підпросторів для перших чотирьох раундів шифрування шифру «Калина» та проведено їх попередній аналіз. Дані результати залежать лише від структури матриці стану шифру та його лінійних перетворень, а тому переносяться на довільні «Калина»-подібні шифри.

В подальшому на основі проведеного аналізу планується побудувати атаки розпізнавання на «Калина»-подібні шифри із зменшеною кількістю раундів.

Література References

[1] Gregor Leander, Mohamed Ahmed Abdelraheem, Hoda AlKhzaimi, and Erik Zenner. A Cryptanalysis of PRINTcipher: The Invariant Subspace Attack. In *Advances in Cryptology* – CRYPTO 2011: 31st Annual Cryptology Conference,Santa Barbara, CA, USA, 2011. Proceedings, pages 206–221, 2011.

[2] Lorenzo Grassi, Christian Rechberger and Sondre Ronjom, "Subspace Trail Cryptanalysis and its Applications to AES" [Online] – http://eprint.iacr.org/2016/592

[3] Roman Oliynykov, Ivan Gorbenko, Oleksandr Kazymyrov et al. "A New Encryption Standard of Ukraine: The Kalyna Block Cipher" [Online]. – http://eprint.iacr.org/2015/650

[4] Інформаційні технології. Криптографічний захист інфомації. Алгоритм симетричного блокового перетворення: ДСТУ 7624:2014. – К.: Держспоживстандарт України, 2015. – 238 с.

[5] І.Д. Горбенко, Перспективний блоковий шифр "Калина" – основні положення та специфікація / І.Д. Горбенко, О.С. Тоцький, С.В. Казьміна та ін. // Прикладна радіоелектроніка. – 2007. – Т.6, №2. – С.195-208

Формування Сеансовиих Ключів на Базі Концепції Прийняття Рішень

Глущенко В.Є.
кафедра кібернетики та комп'ютерних систем
Східноукраїнський національний університет
імені В.Даля
Сєвєродонецьк, Україна
2847@i.ua

Петришин М.Л.
кафедра інформатики
Прикарпатський національний університет
імені Василя Стефаника
Івано-Франківськ, Україна
M.L.Petryshyn@gmail.com

Session Keys Forming on the Base of Decision-Making Concept

V. Glushchenko
Department of Cybernetics and Computer Systems
Volodymyr Dahl East Ukrainian National University
Severodonetsk, Ukraine
2847@i.ua

M. Petryshyn
Department of Computer Science
Vasyl Stefanyk Precarpathian National University
Ivano-Frankivsk, Ukraine
M.L.Petryshyn@gmail.com

***Анотація*—У статті проаналізовано застосування математичного апарату формування сеансових ключів, заснованого на концепції прийняття рішень в структурованому просторі.**

***Abstract*—The article deals with the mathematical apparatus of session keys formation based on the concept of decision-making in a structured space.**

Ключові слова—інформаційна безпека, сеансові ключі, завадостійкість, кодування, інформація

Keywords—information security, session key, noise resistance, encoding, information

I. Вступ

Як показують останні дослідження вітчизняних і зарубіжних авторів, одним з найбільш ефективних методів шифрування повідомлень в процесі інфообміну є методи, засновані на концепції сеансових ключів. До основних переваг алгоритмів, заснованих на даній методології, є можливість використання різних ключів для шифрування різних обсягів даних в інфопотоках повідомлень.

В ідеальному випадку необхідно забезпечити множинність та різноманітність ключів. Це зумовлює актуальність вирішення проблеми створення та застосування математичного апарату, що дозволяє формувати різноманітні сеансові ключі і забезпечує базу для створення ефективних алгоритмів роботи з такими ключами.

Мета дослідження полягає в здійсненні аналізу математичного апарату формування сеансових ключів, який базується на концепції прийняття рішень в структурованому просторі.

II. Концепція формування сеансових ключів в структурованому просторі

Концепція формування сеансових ключів в структурованому просторі полягає на наступних основних положеннях:

1. Під структурованим простором розуміється простір, що має метрику, яка дозволяє визначити віддаль між будь-якою парою точок цього простору.

2. Основним поняттям є поняття "образ ключа", під яким розуміється опис ключа за допомогою заданої системи ознак. Нехай QL^M - структурований простір з носієм A, $||A|| = M$, де M - кількість об'єктів множини A. Множина QL - вихідна множина описів сеансових ключів $QL \in QL^M$. Описи ключів є точками простору QL^M.

3. Взаємозв'язки між точками сусідніх гіперповерхностей QA визначено відображенням $V^S \to V^{S-1}$, яке бієктивно відображає точку $y \in V^S$ на підмножину точок X_y^{S-1}, де $V^{S-1} = \bigcup_{y \in V^S} X_y^{S-1}$,

4. На множині QL визначено кластери $Q = \{Q_1, Q_2. ... Q_m\}$ множини D - допустимих описів сеансових ключів, які представлено як центри відповідних кластерів. Описи кластерів включають: описи центру, максимальної і мінімальної точок кожного кластера.

Максимальною точкою кластера Q_i, $Q_i = \{Q_1, Q_2, ..., Q_t\}$, визначено точку $R_i^M \in Q_i$, для якої справедливо $Pot\ R_i^M \geq Pot\ R_i$, $i = \overline{1, t}$, а мінімальною - точка $R_i^m \in Q_i$, $Pot\ R_i^m \leq Pot\ R_i$, $i = \overline{1, t}$. Віддаль від центру класу R_i до R_i^M визначено віддаллю до верхньої межі кластера Q_i, віддаль від R_i до R_i^m - віддаллю до нижньої межі кластера Q_i. Сукупність кластерів формують модель простору описів сеансових ключів.

5. Процес розпізнавання ключа трактується як знаходження точки погодження в моделі простору описів за множиною представлених описів. Функцією вибору точки погодження в просторі опису визначено триплет:

$$(QL,\ D \subseteq QD,\ C{:}D \rightarrow QL),$$

де QL - вихідна множина допустимих описів розпізнавання відображень $QL \in QA$, D - деяка частина множини QD всіх підмножин множини QL, $C{:}D \rightarrow QL$ - відношення, в якому для будь-якої підмножини G з QL, $D(G) \rightarrow T$, T – точка погодження. В якості точки погодження множини приймається точка T, для якої справедлива умова:

$$\bigcap_{i=1}^{n} T_i \subset T \subset \bigcup_{i=1}^{n} T_i\ .$$

III. Використання простору лінійних квазіпорядків для формування опису сеансових ключів

Розглянемо реалізацію вище наведеної концепції для формування описів сеансових ключів в просторі лінійного квазіпорядка.

Опис образів здійснюється за допомогою множини ознак A. Зв'язок між парою ознак $a, b \in A$ визначено за допомогою наступних бінарних відношень:

а) відношення еквівалентності (a–b):

$$\rho = \{(\alpha, b) : z(\alpha) = z(b)\}; \quad (1)$$

б) відношення строгого пріоритету (a, b):

$$\rho = \{(\alpha, b) : z(\alpha) > z(b)\}; \quad (2)$$

де $z\ (i)$ - значення ознаки i в заданій шкалі.

Отриманий таким чином опис способу представляє лінійний квазіпорядок у вигляді строки, що відображає впорядкування примітивів A.

Відношення еквівалентності породжує декомпозицію множини A наступним чином: α і $b \in A$ відносяться до одного класу, якщо $(\alpha, b) \subset \rho$. Такий опис способу можна подати як

$$R = \{k_1, k_2, ..., k_m\}, \quad (3)$$

$$\bigcup_{i=1}^{m} k_i = \mathrm{A}, \quad (4)$$

де k_i- клас декомпозиції A, $k_i \cap k_j = \varnothing$, $i \neq j$.

Класи декомпозиції R можна пронумерувати таким чином, що відношення ρ збігатиметься з відношенням слідування. Отримуємо, таким чином, впорядковану декомпозицію.

У випадку, коли в декомпозиції R еквівалентні елементи відсутні, тобто всі класи декомпозиції одноелементні, то відношення ρ утворює лінійний порядок.

Множини всіх відношень лінійного квазіпорядка на A з геометричної точки зору утворюють простір лінійних квазіпорядків (ПЛК), множини лінійних порядків - простір строгого порядку (ПСП).

Ранжування елементів множини A трактовано як точки цих просторів.

Для характеристики ранжування, а, отже і точок, якими вони відображаються в ПЛК, використано поняття потенціалу.

Нехай об'єкт $b \in B$ описано ранжуванням $R = \{K_1, K_2, ..., K_m\}$. Тоді для R справедливо:

$$K_S \in R \Rightarrow \left\{\forall j \neq s,\ j = \overline{1, m}, s \in [1, m] : |K_S| \geq |K_i|\right\}. \quad (5)$$

де $|K_i|$ - число елементів в класі K_i. Клас $K_S \in R$ називається класом, що визначає потенціал R, $Pot\ R = |Ki|$. Точка x, що відображає R в ПЛК, має той же потенціал, що і R. Точки потенціалу s утворюють множину V^S, розташовану на гіперповерхні U^S. ПЛК представляється у вигляді множини вкладених гіперповерхонь $QA = \bigcup_{i=2}^{N-1} U^i$, для яких:

$$\forall U^i \in QA,\ \forall i \neq j,\ i,\ i = \overline{2, N-1} \Rightarrow U^i \cap U^j = \varnothing.$$

Потенціали сусідніх гіперповерхонь відрізняються на одиницю, тобто

$$\forall U^i \in QA,\ \forall i,\ i = \overline{2, N-1} \Rightarrow Pot\ U^i - Pot\ U^{i+1}.$$

Потенціали гіперповерхонь ПЛК змінюються від 2 до $N–1$, $N = |A|$. В якості функції, що характеризує взаємне розташування точок ПЛК, використано метрику Хеммінга [3].

Позначимо через QL^M простір лінійних квазіпорядків з носієм A, $||A|| = M$, де M - кількість об'єктів множини A.

Нехай $R = \{R_1, R_2, ..., R_m\}$ і $P = \{P_1, P_2, ..., P_m\}$ – дві декомпозиції множини A. Матриці $\boldsymbol{r} = \left\| r_{i,j} \right\|^{M_{i,j}=1}$,

$\rho = \left\| \rho_{i,j} \right\|^{M_{i,j}=1}$ - матриці зв'язку елементів в декомпозиціях R і P відповідно.

Нехай

$$e_{i,j} = \begin{cases} 1, \text{якщо } (i,j) \in \rho \\ 1, \text{якщо } (i,j) \in \tilde{\rho} \\ 0, \text{якщо } (i,j) \notin \rho \end{cases} \quad (6)$$

Тоді віддаль d_x (R,P) між елементами R і P визначається за формулою:

$$d_x(R,P) = \sum_{i,j=1}^{M} \left| r_{i,j} - p_{i,j} \right|. \quad (7)$$

Отже, d_x (R,P) - кількість декомпонованих розбіжностей елементів матриць $\boldsymbol{r}$ і ρ.

Нехай сеансовий ключ b_i описується n різними образами, які належать кластеру ключа b_i. Отримані при цьому описи ключа b_i утворюють множину $W = \{R_1, R_2, ..., R_m\}$. Множина W розглядається як представлення описів зображень ключа b_i, за якими здійснюється його розпізнавання, що представляють собою процес знаходження погодженого рішення за елементами W. В якості такого рішення приймається опис, розташований між усіма елементами W. Вважається, що ранжування R знаходиться між елементами W, якщо для нього справедлива умова:

$$\bigcap_{n=1}^{n} R_i \subset R \subset \bigcup_{i=1}^{n} R_i \,. \quad (8)$$

Для знаходження погодженого рішення згідно заданого представлення W використовується апарат геометричного підходу до знаходження групового рішення [3].

Використовуючи апарат геометричного підходу до знаходження групових рішень, розроблені стандартні процедури знаходження точки погодження в ПЛК [1, 2].

При цьому досліджено властивості різних функцій вибору з урахуванням допустимих видів представлення. Показано, що функції вибору для точок $W \subseteq X_y^{S-1}$ задовольняють умову слідування, тобто $|W| > 1$, $W' \subseteq W$, $|W'| > 1$, $R \in C(W) \Rightarrow R \in C(W')$, та умову сепарування: $|W| > 1$, $W' \subseteq W$, $|W'| > 1$, $\forall W' : C(W) \Rightarrow R$, $C(W') \in C(W)$. Якщо $LW \subseteq X_{R_1}^S \cap X_{R_2}^S$, $|LW| > 1$, то клас функції вибору для LW має властивість узгодженості, тобто

$$VW, Z \subseteq LW, C(W) \cap C(Z) \subseteq C(W \cup Z) \,.$$

Отже, завдання пошуку узгоджувального рішення в ПЛК по заданому пред'явленню G в ПЛК зводиться до вибору і використання процедур знаходження точки, що задовольняє умові (8), на основі процедурних знань системи.

Таким чином, відмінна особливість запропонованого методу полягає у введенні "універсального", в певному сенсі, простору описів зображень для класифікації образів за мінімумом відстані і розробці стандартних процедур прийняття рішень на основі вхідних даних і бази знань системи.

IV. Алгоритм застосування одноразового сеансового ключа

На рис. 1 представлена схема алгоритму застосування сеансового одноразового ключа.

Першим кроком вибирається модель опису сеансових ключів, яка буде використана для вибору образу ключа, використовуваного для шифрування переданого блоку даних. Обраній моделі опису відповідає конкретне значення контрольного маркера.

На наступному кроці проводиться шифрування переданого блоку даних.

На підставі образу сеансового ключа формуються його описи, які утворюють множину представлень. Описи способу ключа генеруються випадковим чином так, щоб образ ключа був точкою погодження для точок, що утворюють множину представлень. Кількість точок множини представлень визначається рівнем секретності, який необхідно забезпечити, і якістю передачі, яку забезпечує використовувана мережа. Як показує практика, кількість точок у множині представлень не повинна перевищувати 3-4 описів [3].

Контрольний маркер і елементи множини представлень шифруються за допомогою відкритого ключа користувача. З урахуванням специфіки використовуваних алгоритмів і протоколів обміну інформацією, контрольний маркер і множина представлень можуть передаватися окремо від основних даних або бути поміщеними всередину переданого блоку інформації (пакета).

Отримувач після одержання інформації проводить дешифрування контрольного маркера і множини представлень за допомогою відкритого ключа користувача. За значенням контрольного маркера визначається модель опису образів сеансових ключів, яка використовується відправником. По точках множини представлення відновлюється образ сеансового ключа.

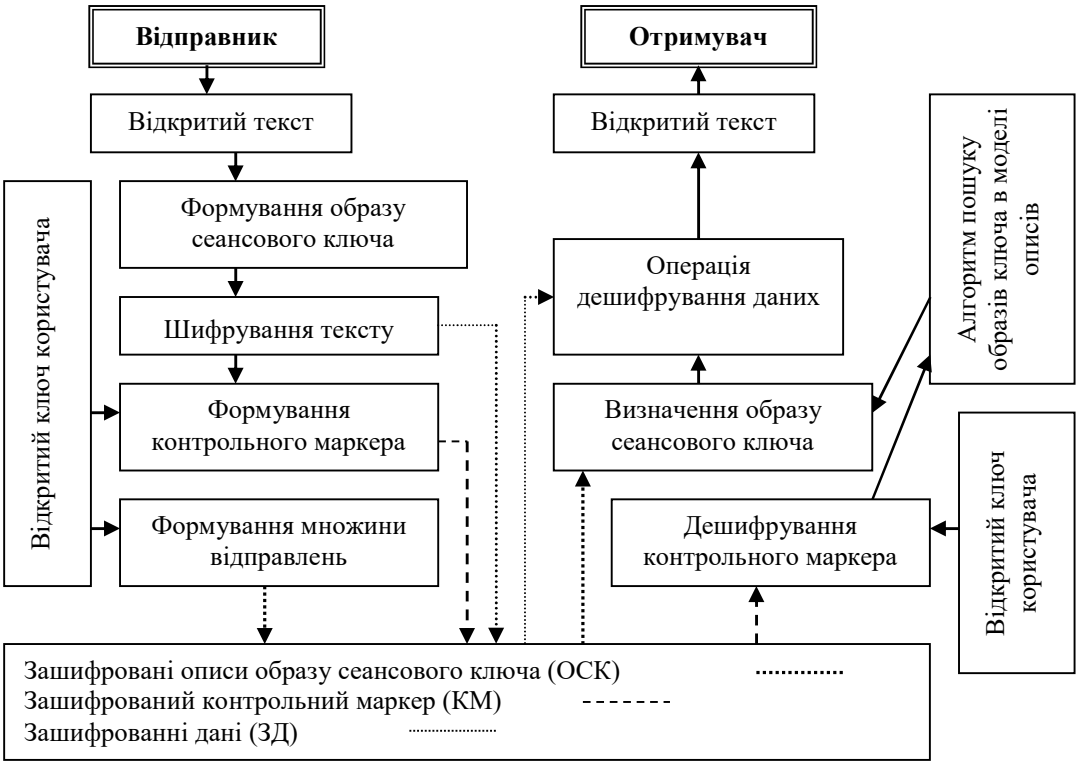


Рис. 1. Алгоритм застосування одноразового сеансового ключа.

На даному етапі автоматично вирішується завдання ідентифікація відправника.

Після визначення способу сеансового ключа проводиться дешифрування отриманих даних.

Перевагою алгоритму є те, що ключ, який використовується для шифрування повідомлень, заснований на інформації, яка є у обох користувачів, а той факт, що ключ не залежать від передачі їх один одному, забезпечує повну таємність. Сторони можуть погодити секретне значення без шифрування, і ця загальна величина може бути відразу використовуватися для шифрування даних і/або засвідчити автентичність.

V. Висновки

1. Запропонований підхід дозволяє розширити множину можливих описів сеансових ключів, і працювати з ключами нефіксованої довжини, що значно розширює продуктивність застосованого методу.

2. При формуванні сеансового ключа обидві сторони мають у своєму розпорядженні одну і ту ж інформацію про простір описів, проте операція вибору конкретного сеансового ключа базується на даних тільки однієї сторони.

3. Створення стандартних процедури опису образів сеансових ключів за моделлю описів дозволило уніфікувати операції розпізнавання образів і значно скоротити необхідний обсяг обчислень.

Література References

[1] В.Е. Глущенко, Ю.В. Глущенко, “Концептуальные вопросы построения интеллектуальных систем защиты от несанкционированного доступа,” *Вістник Східноукраїнського національного університету ім. Володимира Даля*. – 2006. – № 5 [111] – С. 48-53.

[2] V. Glushchenko, M. Petryshyn, “Investigation of the space structure of session keys patterns description” “Дослідження структури простору опису образів сеансових ключів” Матеріали Міжнародної науково-практичної конференції "*Інформаційні технології та комп’ютерне моделювання*", - Івано-Франківськ, 2016. – С. 113–116.

[3] В.Є. Глущенко, М.Л. Петришин, “Формування завадостійкого коду сеансових ключів.” Матеріали п’ятої Міжнародної науковопрактичної конференції "*Інформаційні технології та комп’ютерна інженерія*", - Івано-Франківськ. 2015. – С. 171–174.

Семантичний Аналіз Великих Даних в Задачах Кібербезпеки

Вероніка Островська
кафедра захисту інформації
Вінницький національний технічний університет
Вінниця, Україна
vera_nika.ostrovskaya@mail.ru

Олеся Войтович
кафедра захисту інформації
Вінницький національний технічний університет
Вінниця, Україна
voytovych.op@gmail.com

Big Data Semantic Analysis in the Cybersecurity Tasks

Veronika Ostrovska
dept. of Information Protection
Vinnytsia National Technical University
Vinnytsia, Ukraine
vera_nika.ostrovskaya@mail.ru

Olesia Voitovych
dept. of Information Protection
Vinnytsia National Technical University
Vinnytsia, Ukraine
voytovych.op@gmail.com

Анотація—В роботі надана загальна постанова процедури класифікації текстів, описані основні підходи, що використовуються в задачі класифікації семантичних даних, розглянуті найбільш поширені математичні методи класифікації та регресійний аналіз текстових документів. Розкрито особливості використання, переваги та недоліки зазначених методів. Зроблено висновок щодо необхідності подальшого розроблення алгоритмів класифікації на основі зазначених методів, що були б простими в реалізації, ефективними, мали низькі обчислювальні витрати при навчанні та високу якість класифікації в реальних завданнях.

Abstract—The general terms in classification of texts, main approaches, which are used in classification of semantic data, are presented at the article. Both mathematical classification and text documents regression analysis methods, which are the most common ones, were researched. The features, advantages and disadvantages of these methods were determined. The deduction of the further development necessarily of classification algorithms based on these methods which would be simple for implementation, efficient, low computational costing in training and high classification quality of real data, was yielded.

Ключові слова—Великі Дані, кібербезпека, аналіз тональності, семантичні дані, машинне навчання, словник тональності, класифікація, опорний вектор, Байєсівський класифікатор, k-найближчий об'єкт, регресія

Keywords—Big Data, cybersecurity, sentiment analysis, semantic data, machine learning, dictionary of tonality, classification, support vector, Bayesian classifier, k-nearest object, regression

I. Вступ

У сфері захисту від сучасних кібератак однією із новітніх тенденцій є використання аналітичних систем кібербезпеки із застосуванням систем машинного навчання і штучного інтелекту на великих обсягах даних (Machine Learning-based Security Analytics using Big Data). Ці системи дають змогу помічати відхилення у поводженні систем чи користувачів від норми і таким чином виявляти більшість небезпечних кібератак [1].

У джерелі [2] наведено спосіб збирання даних із соціальної мережі Facebook з допомогою функцій мови програмування R. За основу дослідником було взято подію – другий тур дебатів кандидатів у президенти США Дональда Трампа та Гілларі Клінтон, що відбулися 10 жовтня 2016 року. Це були гострі політичні протистояння, що спричинили інформаційну війну між таборами прихильників кандидатів. Результати дослідження довели, що реакція людей та підтримка Гілларі Клінтон і Дональда Трампа після дебатів змінилася.

Аналітична обробка Великих Даних дає змогу накопичувати знання, виявляти закономірності і виробляти оптимальні методи. Після терористичного акту на Бостонському марафоні у 2013 році великі набори повідомлень, знімків і відеозаписів із соціальних мереж класифіковано та проаналізовано з допомогою високопродуктивних систем, що в кінцевому підсумку допомогло виявити організаторів теракту [3]. Таким чином, хмари надали обчислювальні потужності для вирішення завдання, а результативність роботи

автоматизованих засобів аналітики було покращено завдяки інформації від учасників соціальних мереж [3].

II. Результати дослідження

Важливою проблемою аналізу текстів у глобальній мережі є забезпечення безпеки і виявлення потенційно небезпечних або небажаних повідомлень (поширення спаму і вірусів, терористичні загрози тощо) [4; 5].

Віртуальні спільноти все активніше і масштабніше використовують у власних інтересах засоби інформаційно-психологічного впливу. Вони надають широкі можливості в плані впливу на формування громадської думки, прийняття політичних, економічних і військових рішень, впливу на інформаційні ресурси противника і поширення спеціально підготовленої інформації (дезінформації) [6]. Процеси в соціальних мережах викликають підвищений інтерес в науці, однак темпи теоретичних досліджень істотно відстають від темпів розвитку соціальних мереж [7].

Є два основних підходи до проблеми аналізу тональності семантичних даних: підхід, що ґрунтується на методах машинного навчання, та підхід, що ґрунтується на використанні словників тональної лексики. У першому підході завдання аналізу тональності зводиться до класифікації текстів, яка може бути вирішена шляхом навчання класифікатора на заздалегідь розміченій колекції текстів. Під класифікацією текстів (Text Categorization) розуміється розподіл текстових документів по заздалегідь визначених категоріях.

Загальна методика класифікації полягає в тому, щоб за допомогою набору прикладів із кожного класу відшукати правила, які можуть бути застосовані до нових прикладів. Це один з найважливіших режимів машинного навчання [8].

Основою другого підходу є аналіз тональності окремих слів (термів) у тексті і подальше визначення тональності тексту згідно з оцінками окремих слів, що входять до даного тексту. Для цього в основному використовуються словники тональності, в яких кожному слову відповідає величина, яка відображає «вагу» слова в тональності всього тексту. У подальшому згідно із запропонованим методом будується функція, яка на вхід приймає кількість входжень у текст кожного слова й обчислює агреговану величину тональності всього тексту.

Однак проблемою такого аналізу є те, що не завжди можна просто визначити точне емоційне забарвлення тексту, опираючись тільки на окреме слово. Поширення набули слова, які в сукупності можуть мати зовсім інший емоційний зміст, ніж поодинці. Або ж текст може містити велику кількість негативних або позитивних слів і все одно виражати протилежну думку. Тому одним із напрямів аналізу тональності тексту є вибір методів таким чином, щоб здійснювати класифікацію максимально точно, враховуючи різні можливі комбінації.

Завдання методів аналізу тональності текстів полягає в тому, щоб якнайкраще обрати такі ознаки і сформулювати правила, відповідно до яких прийматиметься рішення про віднесення документа до певного класу. До найвідоміших методів аналізу тональності текстів на основі машинного навчання відносяться метод Байєсової (наївної) класифікації, метод опорних векторів, метод k-найближчого сусіда та регресія.

Метод Байєсової (наївної) класифікації використовує ймовірнісну модель, в якій класифікація та включення у відповідну категорію документів реалізується шляхом оцінювання ймовірності появи слів у документі. Ймовірності можуть бути використані для оцінювання найбільш близьких категорій тестового документа. Основні переваги байєсівського класифікатора – простота реалізації і низькі обчислювальні витрати при навчанні та класифікації [9]. Особливо добре цей класифікатор справляється з задачею узагальнення, коли навчальний набір містить недостатньо даних, щоб покрити весь простір ймовірностей. У тих рідкісних випадках, коли ознаки дійсно незалежні (або майже незалежні), байєсівський класифікатор (майже) оптимальний [8]. Основним недоліком методу є відносно невисока якість класифікації в більшості реальних завдань. Зазначений метод часто використовується як базовий при порівнянні різних методів машинного навчання.

Метод опорних векторів (Support Vector Machine, SVM) використовує процес пошуку площини вирішення, яка може розділити позитивні і негативні приклади в багатовимірному просторі функції, в якому навчальні документи представлені як вектори. Метод набув величезної популярності завдяки своїй високій ефективності. Результати класифікації текстів з допомогою методу опорних векторів є одними з найкращих у порівнянні з іншими методами машинного навчання [9]. Точність даного методу залежить від добору оптимальної площини вирішення; області розв’язання є ефективними за умови збільшення вхідних змінних [10]. Однак, швидкість навчання даного алгоритму є однією з найнижчих. Метод опорних векторів вимагає великого обсягу пам’яті і значних витрат машинного часу на навчання, що знижує його масштабованість. Проте даний алгоритм можна використовувати як еталон з точки зору якості класифікації. Так, метод працюватиме ефективно, якщо опорних векторів буде порівняно небагато, якщо ж їхня кількість зростатиме, то метод стає малоефективним через значно збільшену складність.

Метод k-найближчого сусіда (k-nearest neighbor) [8] як навчальну вибірку використовує набір об'єктів, кожен з яких належить до одного з двох або більше класів. Невідомий об'єкт відноситься до одного з класів за таким принципом: знаходяться k-найближчі об'єкти з навчальної вибірки в просторі образів (зазвичай використовується міра відстані Евкліда). Потім визначається, до якого класу належить більшість найближчих об'єктів навчальної вибірки – до цього класу належить і невідомий об'єкт.

Даний метод є найпоширенішим алгоритмом плоскої кластеризації, однак його недоліком є необхідність зберігання в оперативній пам’яті комп’ютера всіх об’єктів для порівняння кожного із них з невідомим об’єктом [8]. У дослідженнях, присвячених аналізу роботи різних

алгоритмів машинного навчання для задачі класифікації текстів, цей метод демонструє одні з найкращих результатів [11].

Головною особливістю, що виділяє зазначений метод серед інших, є відсутність у нього стадії навчання. Належність документа відповідному класу визначається без побудови функції класифікації. Основною перевагою такого підходу є можливість оновлювати навчальну вибірку без перенавчання класифікатора [12].

У великих даних задачу класифікації розглядають як визначення значення одного з параметрів об'єкту на основі значення інших параметрів [13]. Задача регресії подібна до задачі класифікації і дозволяє визначити за відомими характеристиками об'єкту значення деякого його параметру. Тут значенням параметру є не кінцева множина класів, а множина дійсних чисел. Перевагою цього алгоритму є те, що на виході для кожного об'єкта отримується ймовірність приналежності до класу. До найбільш дієвих видів регресії в задачах, що стосуються Big Data, відносяться логістична, лінійна та гребнева регресії.

До переваг лінійної регресії можна віднести швидкість і простоту отримання моделі, а також модель можна інтерпретувати. Лінійна модель є прозорою і зрозумілою для аналітика. За отриманими коефіцієнтами регресії можна судити про те, як той чи інший фактор впливає на результат, зробити на цій основі додаткові корисні висновки [14].

Логістична регресія має деякі недоліки, що також властиві лінійній регресії – низька стійкість до помилок, залежність від набору даних, але в загальному працює краще, і може бути приведена до вигляду лінійної регресії для спрощення обчислень [15].

Гребнева регресія – удосконалення лінійної регресії з підвищеною стійкістю до помилок, що накладає обмеження на коефіцієнти регресії для отримання більш наближеного до реальності результату [15]. До того ж, цей результат набагато простіше інтерпретувати.

Класифікація та регресія передбачають здійснення двох обов'язкових етапів. Перший етап – виділення набору об'єктів, для яких відомі значення залежних і незалежних змінних. На основі отриманого набору будується модель визначення значення залежної змінної (функція класифікації або регресії). На другому етапі побудовану модель застосовують до об'єктів, які аналізуються. Недоліком класифікації та регресії є те, що розробник системи повинен фіксувати кількість класів та характеристик, за якими буде проводитись дослідження. Це означає, що якщо система не виявить ознаки або класу, він не буде коректно оброблений.

Кожен із проаналізованих методів має свої переваги і недоліки залежно від конкретного вирішуваного завдання. Ефективність розв'язання завдання значною мірою залежить саме від особливостей засобу. У табл. 1 наведена порівняльна характеристика основних методів аналізу тональності семантичних даних, що застосовуються при обробці природної мови. Тут знак «+» означає високий, знак «–» – низький та «+/−» – нейтральний показник характеристики методу.

ТАБЛИЦЯ І. ПОРІВНЯЛЬНА ХАРАКТЕРИСТИКА МЕТОДІВ АНАЛІЗУ ТОНАЛЬНОСТІ СЕМАНТИЧНИХ ДАНИХ

Метод	**Характеристика**						
	Трудомісткість	***Швидкість***	***Масштабованість***	***Сфера використання***	***Алгоритми***	***Відстань між словами***	***Поєднання з іншими методами***
Метод Байєсової (наївної) класифікації	+\-	+\-	-	Аналіз переходів по web-сторінках	Класичні методи статисти–ки	+	-
Метод опорних векторів	+	+	-	Класифіка ція та регресій–ний аналіз даних	Методи аналізу	+	+
Метод k-середніх	+\-	+	-	Машинне навчання	Методи ієрархіч–ного аналізу	+	-
Регресія	+\-	+	+	Оцінка допусти–мості прийняття рішення	Класичні методи статисти–ки	-	-

Для підвищення оцінок ефективності алгоритмів, ймовірно, слід доповнити їх елементами лінгвістичного аналізу. Таким чином, необхідною умовою для аналізу тональності є складання словарного списку тональної лексики. Для цього в тексті виділяються оціночні слова, для них обчислюється емоційна вага, потім ці ваги об'єднуються за допомогою деякої функції (наприклад, середнє арифметичне або сума). Існує кілька підходів до вилучення оцінних слів і обчисленню їх емоційної ваги.

В джерелі [16] спочатку вибираються дві еталонних множини оціночних слів: позитивна і негативна. Далі з відгуків вибираються набори, що складаються з прикметників в поєднанні з іменниками і прислівники в поєднанні з дієсловами. Автор Turney використовує набори, вважаючи, що, хоча ізольоване слово може вказувати на суб'єктивність, його може виявитися недостатньо для визначення контексту емоційної оцінки. Тональність відгуку розраховується як середнє арифметичне емоційних оцінок наборів, взятих з цього відгуку. Для розрахунку емоційної оцінки для набору Turney використовував пошукову систему Altavista.

Множини оціночних слів також створюються вручну експертами. Для поповнення даних множин можуть використовуватися словники. В джерелі [17] запропонований метод, який використовує тезаурус для поповнення заданої вручну множини оціночних слів. Ідея методу полягає в наступному: якщо слово оціночне, то його синоніми також будуть оціночними і відносяться до однієї тональності, а антоніми – до протилежної

тональності. Ще один підхід представлений в роботі [18], де за допомогою тлумачень слів в словнику визначається їх орієнтація. Даний метод ґрунтується на ідеї, що слова з однаковою емоційною оцінкою мають схожі тлумачення.

В джерелі [19] описується підхід до автоматичного створення словника оціночної лексики в області товарів і послуг для російської мови ProductSentiRus. Словник ProductSentiRus був отриманий застосуванням навченої моделі до наборів відгуків в декількох предметних областях. Словник представлений як список 5 тисяч слів, упорядкованих у міру зниження обчисленої ймовірності їх оціночності без вказівки позитивної або негативної тональності. Для покращення якості списку оціночних слів використовується тезаурус російської мови РуТез [20].

На вхід алгоритму надходить список слів, що впорядковані за ймовірністю їх оціночності. Основна ідея уточнення оціночного лексикону – це автоматичний розподіл понять тезауруса на оціночні та нейтральні, і подальше використання отриманої розмітки для переранжування списку оціночних слів. Цей процес регулюється за допомогою раніше отриманих ваг оціночності слів [19].

Чим більше ресурсів, таких як семантично розмічені корпуси та тезауруси, розроблено для природної мови, тим вище вдається дослідникам підняти якість вирішення цього завдання. На даний момент найбільший синтаксично розмічений корпус російською мовою – SynTagRus [21], який входить до складу Національного корпусу російської мови. На сайті Національного корпусу російської мови (www.ruscorpora.ru) також містяться корпуси для інших мов, в тому числі корпус текстів української мови, однак він доступний лише для онлайн-пошуку.

Переважна більшість кращих підходів у задачах класифікації семантичних даних ґрунтуються на застосуванні методу опорних векторів SVM, який може комбінуватися з додатковими ресурсами на зразок словників або правил. Даний результат узгоджується з результатами, отриманими для завдання аналізу тональності англомовних текстів, де також було доведено, що метод опорних векторів зазвичай породжує кращі за якістю результати в цьому завданні.

III. Висновки

В процесі дослідження методів аналізу тональності текстових даних надано загальну процедуру класифікації текстів, наведено огляд підходів до вирішення задачі класифікації, описані основні підходи, що використовуються в задачі класифікації текстів, визначено етапи процесу класифікації та розглянуті найбільш поширені математичні методи класифікації текстових документів. Розкриті особливості використання, переваги та недоліки зазначених методів дають змогу зробити висновок про необхідність подальшого вдосконалення алгоритмів класифікації на основі зазначених методів, що були б простими в реалізації, ефективними, мали низькі обчислювальні витрати при навчанні та високу якість класифікації в реальних завданнях.

Література References

[1] Експерт: кіберзахист – це не параноя [Електронний ресурс]. – Режим доступу : http://www.bbc.com/ukrainian/features-39364360. – Назва з екрану.

[2] Voitovych O. Badania sieci społecznych jako źródła informacji w czasie wojny / Voitovych O., Holovenko V. // Inżynier XXI wieku projectujemy przyszlosc : monografia / [pod red. : Jacek Rysiński]. – Bielsko-Biała : Wydawnictwo Naukowe Akademii Techniczno-Humanistycznej w Bielsku-Białej, 2016. – С. 111–119.

[3] Analytics of Big Data and social networks [Електронний ресурс]. – Режим доступу : http://www.osp.ru/os/2013/08/13037856. – Назва з екрану.

[4] Kontostathis A. Text mining and cybercrime / Kontostathis A., Edwards L., Leatherman A. // Text Mining. Applications and Theory / ed. by Berry M. W., Kogan J. – Chichester : Wiley. – 2010. – P. 149–164.

[5] Dua S. Data Mining and Machine Learning in Cybersecurity / Dua S., Du X. – New York. – 2011. – 224 p.

[6] Гриненко І. Вплив віртуальних спільнот на інформаційну безпеку: сучасний стан та тенденції розвитку / І. Гриненко, Д. Прокофьєва-Янчиленко // Правове, нормативне та метрологічне забезпечення систем захисту інформації в Україні. – 2012. – № 1 (23). – С. 18–23.

[7] Ефимов Е. Г. Социальные интернет-сети (методология и практика исследования : монография / Е. Г. Ефимов / Волгоградский гос. тех. ун-т. – Волгоград. – 2015. – 169 с.

[8] Коэльо Л. П., Ричарт В. Построение систем машинного обучения на языке Python. 2-е издание / пер. с англ. Слинкин А. А. – М.: ДМК Пресс. – 2016. – 48 – 167 с.

[9] Text Mining: Aplications and theory / ed. by Berry M. W., Kogan J. [Електронний ресурс]. – Режим доступу : http://www.mohamedrabeea.com/books/book1_1165.pdf. – Назва з екрану.

[10] Thorsten J. Text Categorization with Support Vector Machines: Learning with Many Relevant Features [Електронний ресурс]. – Режим доступу : https://www.cs.cornell.edu/people/tj/publications/joachims_98a.pdf. – Назва з екрану.

[11] Sebastiani F. Machine learning in automated text categorization / F. Sebastiani // ACM Comput. Surv. – 2010. – P. 1-47.

[12] Yang Y. A re-examination of text categorization methods / Y. Yang, X. Liu // Proc. of Int. ACM Conference on Research and Development in Information Retrieval (SIGIR-99). – 2007. – P. 42-49.

[13] Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И.И. Технологии анализа данных: Data Mining, Viual Mining, Text Mining, OLAP. – Спб.: БХВ-Петербург, 2007. – 384 с.

[14] Линейная регрессия [Електронний ресурс]. – Режим доступу : https://basegroup.ru/deductor/function/algorithm/linear-regression. – Назва з екрану.

[15] 10 типов регрессии – какой выбрать? [Електронний ресурс]. – Режим доступу : http://datareview.info/article/10-tipov-regressii-kakoy-vyibrat/. – Назва з екрану.

[16] Turney P. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews // Proceedings of ACL-02, 40th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, 2002, pp. 417–424.

[17] Hu M., Liu B. Mining and Summarizing Customer Reviews // KDD, Seattle, 2004, pp. 168–177.

[18] Esuli A., Sebastiani F. Determining the Semantic Orientation of Terms through Gloss Classification // Conference of Information and Knowledge Management (Bremen). ACM, New York, NY, 2005, pp. 617–624.

[19] Лукашевич Н. В., Четверкин И. И. Комбинирование тезаурусных и корпусных знаний для извлечения оценочных слов, Системы и средства информ., 2015, том 25, выпуск 1, С. 20–33.

[20] О лингвистической онтологии «Тезаурус РуТез» [Електронний ресурс]. Режим доступу: http://www. labinform. ru/pub/ruthes/index.htm. – Назва з екрану.

[21] Синтаксически размеченный корпус русского языка: информация для пользователей [Електронний ресурс]. Режим доступу: http://www.ruscorpora.ru/instruction-syntax.html. – Назва з екрану.

СЕКЦІЯ 6.

Архітектоніка та компоненти комп'ютерних систем та мереж

SECTION 6.

Computer networks, systems and components architectonics

Комп'ютерна Система Визначення Реологічних Характеристик Суднових Технічних Рідин

Віталій Нікольський
кафедра теорії автоматичного управлення і обчислювальної техніки, НУ “ОМА”
Одеса, Україна
prof.Nikolskyi@ukr.net

Марк Нікольський
кафедра автоматизації суднових газотурбінних і дизельних установок, НУ “ОМА”
Одеса, Україна
markdezert@ukr.net

Кирило Бережний
кафедра автоматизації суднових паросилових установок, НУ “ОМА”
Одеса, Україна
90demon777@ukr.net

Computer System for Determination of Rheological Characteristics Ships Technical Fluids

Vitalii Nikolskyi
Department of Automatic Control and Computing Technologies, NU “OMA”
Odesa, Ukraine
prof.Nikolskyi@ukr.net

Mark Nikolskyi
Department of Automation of Vessel Gas Turbines and Diesel Installations, NU “OMA”
markdezert@ukr.net
Odesa, Ukraine

Kyrylo Berezhnyi
Department of Automation of Ships Steam Power Units, NU “OMA”
Odesa, Ukraine
90demon777@ukr.net

***Анотація*—Суднові технічні рідини при експлуатації розглядають як "ньютонівські", а в реальних умовах при знаходженні в зазорах триботехнічних вузлів дизелів проявляють себе, як "нен’ютонівські" з ефектом тиксотропії. При бункеровці судна на них надається технічний паспорт з одним значенням кінематичної в'язкості та щільності. Не врахування цих властивостей збільшує знос елементів паливної апаратури та підшипників. А це призводить до збільшення аварійності, експлуатаційних витрат. Негативний ефект посилюється в маневровому режимі дизеля при різкій варіації навантажувальних характеристик і, відповідно, різкій зміні коефіцієнтів в'язкості.**

Нами розроблено комп'ютерну систему, що включає зонд-реометр, що занурюється, з п'єзоелектричним приводом, і новизною якої є можливість використання при комплексному аналізі "нен’ютонівської" рідини та вивченні ефекту тиксотропії в режимі перманентного моніторингу. Існуючі суднові стаціонарні лабораторії не дозволяють провести детальні аналізи.

***Abstract*—Ship technical fluids in operation are considered as "Newtonian", and in real conditions, while in the gaps of the tribotechnical units of the diesel engines, they manifest themselves as "non-Newtonian" with the effect of thixotropy. When bunkering the vessel, they are provided with a technical passport with one value of kinematic viscosity and density. Not taking into account these properties increases the wear of fuel equipment and bearings. And this leads to an increase in accidents, operating costs. The negative effect is amplified in the maneuvering mode of the diesel engine with a sharp variation in load characteristics and, accordingly, a sharp change in the viscosity coefficients.**

We have developed a computer system that includes a submerged type rheometer probe with a piezoelectric drive, and the novelty of which is the possibility of using a non-Newtonian fluid for complex analysis and studying the effect of thixotropy in

the permanent monitoring mode. Existing ship's stationary laboratories do not allow for detailed analyzes.

Ключові слова—"нен'ютонівська" тиксотропна рідина, триботехнічний вузол, комп'ютерна система, реологічні характеристики, віскозиметр з п'єзоелектричним приводом.

Keywords—"non-Newtonian" thixotropic fluid, tribotechnical unit, computer system, rheological characteristics, viscometer with piezoelectric drive.

I. ВСТУП

Зменшення експлуатаційних витрат і впливу на екосистему суднових двигунів внутрішнього згоряння (ДВЗ) і, зокрема, дизелів, які працюють на "важкому" паливі, можливо за допомогою конструктивних змін (змінна ступені стиснення за рахунок пастелі, що рухається), а також застосування електронного управління упорскуванням палива. Поряд з цими заходами постійно йде вдосконалення систем регулювання в'язкості палива, яке відноситься до суднових технічних рідин.

Загальновідомо, що суднові технічні рідини ідентичних марок можуть мати різні коефіцієнти динамічної в'язкості при однаковій температурі. Збільшення в'язкості при інших рівних умовах призводить до погіршення якості розпилювання палива і умов роботи прецизійних пар паливних насосів, більш інтенсивному утворенню нагару на елементах циліндропоршневої групи (ЦПГ) та газо-повітряного тракту, зміни раніше встановлених моментів подачі палива.

При малої в'язкості збільшується кут розпилювання палива і зменшується довжина факела, що при даній конструкції форсунки і камери згоряння призводить до погіршення утворення суміші (нестачі повітря в глибині факела) і, як наслідок, до збільшення витрати палива [1].

Поряд з в'язкістю палива існує проблема регулювання в'язкості мастильних циркуляційних масел, що володіють властивостями "нен'ютонівських" рідин з ефектом тиксотропії в триботехнічних вузлах ДВЗ. "Нен'ютонівські" властивості різних видів палив та мастил полягають в зменшенні їх в'язкості при збільшенні швидкості зсуву, а тиксотропні - зменшення їх в'язкості згодом при постійній швидкості деформації [2].

Специфічні особливості "нен'ютонівських" тиксотропних рідин в вузлах тертя тягнуть за собою недостатній рівень мастила вузлів ЦПГ, збільшення коефіцієнта тертя і, в наслідок чого, зменшення моторесурсу і виникнення аварійної ситуації в роботі механізмів.

Але якщо автоматичні системи регулювання в'язкості палива набули широкого поширення на судах, то автоматичні системи регулювання в'язкості мастильних циркуляційних мастил належного розвитку не отримали. В даний час головним робочим параметром регулювання циркуляційного мастила є його температура. Точна відповідність параметрів температури і в'язкості мастила визначається за допомогою номограм, що містять фізичні характеристики даного сорту. В даний час використання мастильного циркуляційного мастила в суднових ДВЗ відбувається протягом декількох років і тому через його "вироблення" визначити точну в'язкість по його температурі неможливо навіть в умовах спокою рідини, не кажучи вже про його поведінку при зміні швидкостей зсуву і часу. Тому, дослідження специфічних властивостей суднових технічних рідин і розробка нових засобів вимірювання є актуальними.

II. ОСНОВНА ЧАСТИНА

В даний час при створенні віскозиметрів вживаються такі методи вимірювання, як: капілярний; ротаційний; співвісних циліндрів; паралельних нерухомих або рухомих площин, що переміщують; падаючої кульки; крутильних коливань; вібраційний [2].

У роботах [3-7] вивчалися реологічні характеристики суднових технічних рідин на основі ротаційного віскозиметра з метою дослідження їх "нен'ютонівських" і тиксотропних властивостей. Основною невирішеною проблемою проведених досліджень була відсутність на той момент промислових контролерів, які працюють в «жорсткому» реальному часі, що дозволило б не тільки проводити вимірювання в'язкості, але і визначати реологічні характеристики суднових технічних рідин.

В [8] основною метою був аналіз можливостей, достоїнств і недоліків віскозиметрів і реометрів, що використовують п'єзоелементи у вигляді датчика, або приводу вимірювального зонду. В [9-13] запропоновано використання п'єзоелектричних двигунів та актуаторів з наборних шайб при побудові віскозиметрів ротаційних і на співвісних циліндрах. Встановлено, що характер реологічних характеристик, за якими визначається коефіцієнт динамічної в'язкості однієї і тієї ж рідини, істотно відрізняється при дослідженнях на ротаційному віскозиметрі та віскозиметрі на співвісних циліндрах. Процеси в вимірювальному зонді співвісного віскозиметра ідентичні процесам, що протікають в вузлах паливної апаратури [14]. Однак, в дослідженнях використовувався тільки ротаційний п'єзоелектричний двигун, тому що ціна лінійного двигуна становила понад 700 євро. З впровадженням 3D принтерів вдалося спростити виробництво таких двигунів, що призвело до вирівнювання цін лінійних і ротаційних двигунів.

Нами розроблена комп'ютерна система визначення реологічних характеристик суднових технічних рідин в реальному часі, в якій використовується обладнання Phoenix Contact та зонд-реометр, що занурюється, з лінійним п'єзоелектричним приводом. Робота зонду заснована на методі співвісних циліндрів. Ця система є складовою частиною автоматичної системи регулювання в'язкості суднових технічних рідин. На рис. 1 наведена її структурна схема, де ОУ - об'єкт управління, ЗР - зонд-реометр, ПЗР - привід зонду реометра, ПЛК - контролер, МРЧ – монітор реального часу, ВМ - виконавчий механізм.

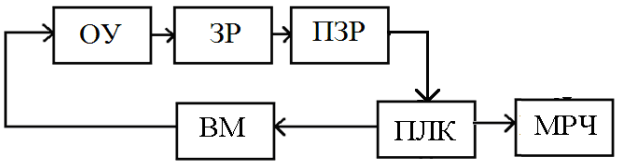


Рис. 1. Структурна схема комп'ютерної системи визначення реологічних характеристик суднових технічних рідин

На рис. 2 наведено функціональну схема комп'ютерної системи визначення реологічних характеристик суднових технічних рідин, де: 1 - електричний підігрівач; 2 - стакан, що занурюється; 3 - досліджувана рідина; 4 - вимірювальний зонд; 5 - реверсивний лінійний п'єзоелектричний двигун LPM-5 [15]; 6 - датчики фіксації переміщення направляючої двигуна; 7 - контролер (ПЛК) ILC 150 GSM / GPRS а також сенсорна WEB-панель WP 04T [16] з призначеним для користувача інтерфейсом, яка використовується як монітор реального часу.

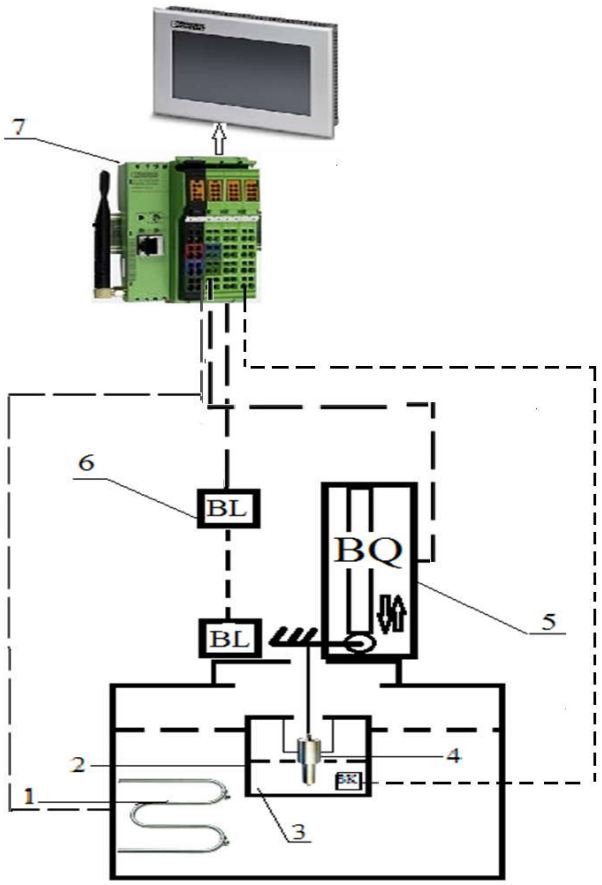


Рис. 2. Функціональна схема комп'ютерної системи

В якості вимірювального пристрою використовується зонд-реометр, що занурюється, та складається з прецизійних "підігнаних" один до одного втулки і плунжера з зазором близько 10 мікрон. Зонд фіксується в отворі майданчику, що регулює, та кріпиться до допоміжної кришки склянки, яка занурюється.

Плунжер реометру через металевий шток з'єднаний з п'єзоелектричним двигуном, що здійснює зворотно - поступальні рухи. В табл. 1 наведено технічні характеристики лінійного п'єзоелектричного двигуна LPM-5-SP.

ТАБЛИЦЯ I. ТЕХНІЧНИЙ ПАСПОРТ LPM-5-SP

Parameters		Measured value	Specifications
Max. speed	Forward (Left)	0.2 m/sec	≥ 0.2 m/sec
	Backward (Right)	0.2 m/sec	
Max. force	Forward (Left)	1.2 N	≥ 4 N
	Backward (Right)	1.2 N	
Self-braking force		1.5 N	≥ 5 N
Resolution (min step)		0.04 μm	≤ 0.1 μm
Excitation Frequency	Forward (Left)	149.3 kHz	140÷160 kHz
	Backward (Right)	149.0 kHz	
Supply voltage (transformer 16:80)			12 V(U_A~60V)
Operating current	Forward (Left)	400 mA	≤ 500 mA
	Backward (Right)	310 mA	
Motor weight			35 g
Size			59x47x10 mm

Положення двигуна в "мертвих" нижньої і верхньої точках і, відповідно, час між ними фіксується світлодіодними оптичними датчиками. Дана конструкція за допомогою головної кришки вміщується всередині підігрівача, що представляє собою виконавчий механізм, і складається з нагрівального елементу, який обігріває рідину.

Робота системи здійснюється наступним чином. Досліджувана рідина поміщається в стакан (2), що занурюється (рис. 2). При попаданні молекул досліджуваної рідини в зазор між плунжером і втулкою (4) виникає опір руху плунжера залежно від сил в'язкого тертя рідини. Це призводить до зміни лінійної швидкості п’єзодвигуна (5). Отримане значення швидкості руху порівнюється з еталонною швидкістю холостого ходу двигуна ПЛК (7), який обробляє отримані сигнали та обчислює в'язкість. Далі відбувається зміна швидкості двигуна у всьому діапазоні швидкостей елементів паливної апаратури малообертових суднових дизелів, а потім будується реологічна характеристика для даного значення температури.

На наступному етапі відбувається зміна температури і проводиться вимір в'язкості з наступною побудовою реологічної характеристики. Таким чином, зробивши дослідження "проби" рідини для всього температурного діапазону і швидкостей переміщення, отримуємо еталонну характеристику. А в процесі експлуатації ДВЗ можна синхронізувати режим роботи вимірювального зонду з режимами роботи вузлів паливної апаратури і отримати реологічні характеристики рідин в «жорсткому» реальному

часі. Це і є основною перевагою і новизною розробленої комп'ютерної системи.

На рис. 3 пояснено, як здійснюється програмування контролеру ILC 150 GSM/GPRS за допомого інтегрованих середовищ від компанії Phoenix Contact.

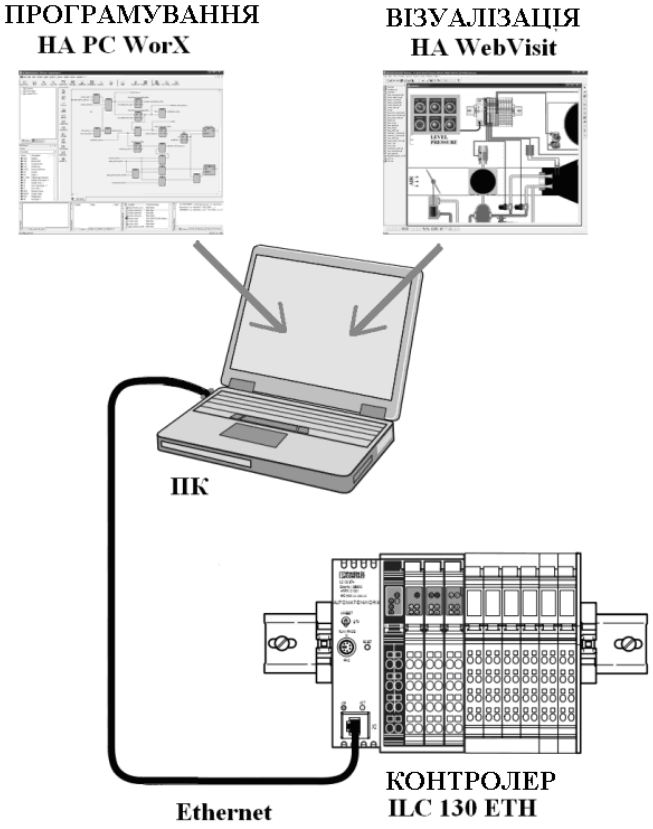


Рис. 3. Програмне забезпечення Phoenix Contact, яке застосоване при створенні системи.

В інструментальному середовищі PC Worx здійснюється програмування логічної частини на мові FBD, а в середовищі WebVisit здійснюється програмування Web серверу.

III. Висновки

Розроблена комп'ютерна система управління дозволяє досліджувати властивості "нен'ютонівської" псевдопластичної рідини з ефектом тиксотропії, що знаходиться в проміжку між втулкою і плунжеру зонду, подібно тому, як вона знаходиться в зазорі підшипників ковзання триботехнічних вузлів ДВЗ, так і плунжерних пар паливних насосів високого тиску дизелів. У перспективі дані дослідження дозволять виробити алгоритм регулювання в'язкості палива і мастила з "попередженням", що дозволяє мінімізувати ефект "нен'ютонівської" тиксотропної рідини у вузлах тертя, що в свою чергу збільшить моторесурс і знизить експлуатаційні витрати суднової енергетичної установки.

Література References

[1] В.И. Ланчуковский, Автоматизированные системы управления судовых дизельных и газотурбинных установок. / В.И. Ланчуковский, А.В. Козьминых // Учебник. - М.: Транспорт, 1983. - 320с.

[2] П.Ф. Овчинников, Виброреология // Киев: наук. думка, 1983. – 272 с.

[3] Б.А. Алтоиз, Трибологические особенности граничных смазочных слоев судовых гоплив и масел / Б.А. Алтоиз, С.А. Ханмамедов // Судовые энергетические установки: научн.-техн. сб. - 2003. - Вып. 9. - С. 80 - 86.

[4] С.А. Ханмамедов, Совершенствование функциональных свойств систем смазывания судовых энергетических установок: Дис. …докт. техн. наук: 05.08.05. – Николаев, 1990. – 382 с.

[5] Б.А. Алтоиз, Ротационный вискозиметр для исследования микронных прослоек / Б.А. Алтоиз, С.К. Асланов, А.Ф. Бутенко // [Електроний ресурс]. - Режим доступу: http://phys.onu.edu.ua/files/journals/fas/articles/42/fas42_altoiz.pdf (Дата звернення 09.03.2017).

[6] Е.Б. Плавинский, И.В. Никитенко, В.В. Никольский, А.Ю. Поповский, Д.Ю. Крохмаль, О.В. Глазева, Измерения вязкости жидких сред пьезоэлектрическим вискозиметром // Судовые энергетические установки: науч. - техн. сб. - Одесса: ОГМА, 2001. - Вып. 6. - С. 61 - 62.

[7] В.В. Никольский, Л.Б. Багдасарян, Определение реологических характеристик топлив и масел, используемых в СЭУ // Судовые энергетические установки: научн.-техн. сб. – Одесса: ОНМА, 2005. – Вып. 14. – С. 31-35.

[8] М. В. Никольский, Пьезоэлектрический Вискозиметр / М.В. Никольский, К.Ю. Бережной // Наукові праці ЧДУ ім. Петра Могили. – Серія: "Комп'ютерні технології". – Вип. 275. – Т. 287. – Миколаїв, 2016. – С. 60–63.

[9] В.В. Нікольський, Деклараційний патент України, МКІ 7 G01N11/10 Віскозиметр. - № 2003054350; Заявл. 15.05.2003; Опубл. 15.01.2004, Бюл. № 1.

[10] В.В. Нікольський, Деклараційний патент України, МКІ 7 G01N11/10 П'єзоелектричний віскозиметр. - № 2003109671; Заявл. 28.10.2003; Опубл. 15.09.2004, Бюл. № 9.

[11] В.В. Нікольський, Деклараційний патент України, МКІ 7 G01N11/10 Віскозиметр. - № 20031211819; Заявл. 18.12.2003; Опубл. 15.11.2004, Бюл. № 11.

[12] В.В. Нікольський, Деклараційний патент на корисну модель України, МКІ 7 G01N11/10 Віскозиметр. - №u200500629; Заявл. 24.01.2005; Опубл. 15.07.2005, Бюл. №7.

[13] В.В. Никольский, Оценка реологических свойств тиксотропных жидкостей реометрами с пьезоприводом // Автоматизация судовых технических средств: науч. - техн. сб. - Одесса: ОНМА, 2005. - Вып. 10. - С. 61 - 64.

[14] В.В. Никольский, Моделирование процессов в вискозиметрах с пьезоэлектрическим приводом // Автоматика, автоматизация, электротехнические комплексы и системы: науч. - техн. журнал - Херсон: ХГТУ. - 2004. - №2 (14) - С. 60 - 67.

[15] В.В. Лавриненко, В.С. Коваль, С.Ф. Петренко, В.В. Лукін, Р.В. Франченко, Патент на винахід № 76759 України, МПК(2006) H02N 2/00 Пєзоелектричний двигун. - № 20040110441; Заявл. 18.12.2003; Опубл. 15.09.2006, Бюл. № 9.

[16] Сенсорная панель - WP 04T - 2913632 [Електроний ресурс]. - Режим доступу: https://www.phoenixcontact.com/online/portal/ua?uri=pxc-oc-itemdetail:pid=2913632&library=uauk&pcck=P-14-15-06-09&tab=1 (Дата звернення 09.03.2017).

Система контролю завантаження контейнеровозів “ShipLoad”

Юрій Накул
кафедра теорії автоматичного управлення і обчислювальної техніки, НУ “ОМА”
Одеса, Україна
nak01@mail.ru

Марк Нікольський
кафедра автоматизації суднових газотурбінних і дизельних установок, НУ “ОМА”
Одеса, Україна
markdezert@ukr.net

Владислав Стовманенко
кафедра інтелектуальних та інформаційних систем, ЧНУ ім. Петра Могили
Миколаїв, Україна
flaviusglamfenix@gmail.com

Loading System Control Container Ships “ShipLoad”

Yuriy Nakul
Department of Automatic Control and Computing Technologies, NU “OMA”
Odesa, Ukraine
nak01@mail.ru

Mark Nikolskyi
Department of Automation of Vessel Gas Turbines and Diesel Installations, NU “OMA”
Odesa, Ukraine
markdezert@ukr.net

Vladyslav Stovmanenko
Department of Intellectual and Information systems, Black Sea State University of Petro Mohyla
Mikolaiv, Ukraine
flaviusglamfenix@gmail.com

***Анотація*—Завантаження вантажів на судно є складним завданням, що вимагає координації багатьох складових. Недбалість під час цього процесу може привести до дуже серйозних наслідків. Зокрема показовий випадок судна “CSCL Uranus”. Потрапивши в шторм контейнеровоз нахилився і за борт потрапило 80 контейнерів, деякі контейнери було пошкоджено. Це відбулося через неправильне закріплення та завантаження контейнерів на борт. Це призвело до фінансових втрат. Все могло обійтися ще серйозніше, якби судно нахилилося на більший кут внаслідок недотримання правил техніки безпеки. Судно могло досягнути значного крену й перевернутися, тоді мова йшла б про значно більші витрати, а також про людські життя.**

Нами розроблено метод контролю завантаження та на його основі спроектовано систему, що включає програму для керуванням процесом завантаження та має можливості індикації, структуру файлів, що міститимуть необхідну для роботи програми інформацію, модулі відслідковування координат контейнера та систему бездротового зв’язку.

***Abstract*—Ship loading is a difficult task that requires the coordination of many components. Negligence in the process might cause very serious consequences. Particularly illustrative case of “CSCL Uranus” ship. Being in storm boat leaned and 80 containers fell overboard, some of others was corrupted. This happened because of wrong containers loading and mounting. This led to financial losses. Everything might be much more seriously if ship bends at greater angle. Ship might achieve significant roll and overturn, than it was on to more financial losses and even to human’s lives. All this because of violations of some load rules.**

We developed load control method and designed system that includes program for loading process control with some

possibilities for indication, file structure with all necessary information needed for right program's work, container's coordinate tracking module and system of wireless communications based on it.

Ключові слова—судно, завантаження, контроль, бездротовий зв'язок, мікроконтролер, інерційні навігаційні системи, декартовий простір.

Keywords—ships, downloading, control, wireless connection, controller, inertial navigation systems, cartesian space.

I. ВСТУП

Морський та річковий транспорт існує з давніх-давен, але з розвитком комп'ютерних технологій з'явилася можливість контролювати завантаження вантажів на судно задля прискорення процесу та підвищення надійності.

Як тільки контейнеровози увійшли в широкий ужиток, з'явилися проблеми (зокрема проблеми із завантаженням), які не вдалося вирішити і до теперішнього часу. Не завжди можливо оперативно встановити, яка кількість контейнерів знаходиться на борту; чи правильно вони розміщені. Для отримання актуальної та точної інформації потрібно посилювати контроль над процесом. Це не завжди вдається, так як в гонитві за зниженням собівартості перевезень судновласники скорочують екіпаж при одночасному збільшенні тоннажу. Все це призводить до затримок лоцмана в процесі обслуговування судна і збільшення часу використання буксирів. У підсумку отримуємо додаткові витрати на перевезення.

Окрім того, на багатьох суднах капітан або його помічник мають безпосередньо спостерігати за процесом задля того, щоб пересвідчитися в тому, що відповідний контейнер було встановлено саме на попередньо вказане місце, а для перевірки закріплення треба було направити декілька людей, які б змогли все оглянути в безпосередній близькості до контейнера.

Зараз існують програми, які дозволяють графічно відображати розміщення контейнерів в зонах зберігання на судні, але заносити дані про перебування цих вантажів треба вручну, що приховує у собі загрозу помилок з неуважності. Також програми можуть пропонувати оптимізацію розміщення контейнерів, але вже після того, як дані про нього було внесено. А якщо контейнер встановлено і лише потім внесено в програму, то доведеться переставляти його, тобто витрачати додатковий час та ресурси.

Зазначені системи мають такі недоліки:

– необхідність купувати вантажні крани, які реалізують уже готові алгоритми автоматизованого завантаження, що вимагає вливання значних коштів за відносно короткий проміжок часу;

– закритість коду програм, що не дозволяє гнучко налаштовувати програму для потреб того чи іншого судна/порту/транспортної компанії;

– висока ціна на продукти фірм, які пропонують аналогічні системи(наприклад DELFTload коштує 150 євро за одну ліцензію);

– націленість на певний, конкретний тип суден, неуніверсальність;

– недостатній рівень автоматизації у програмних продуктів(необхідність вручну вносити дані про перебування чи відсутність певного контейнера на судні);

Тому постає актуальне питання щодо встановлення контролю над процесом завантаження в портах для забезпечення правильності розміщення, дотримання показань ваги на різних зонах корабля, дотримання правил техніки безпеки при розміщенні контейнерів з вантажами різних класів "небезпечності".

Інформація про судно та його параметри має зберігатися у доступному для різних програмних засобів вигляді, що дозволить відносно легко створювати нові інструменти для роботи з цими даними. Проте в той же час має існувати можливість забезпечувати певну конфіденційність за необхідності. Певні види вантажів мають певні правила розміщення. Відповідно до цих правил має підбиратися правильна позиція для контейнера. Також ця система має допомогти отримати доступ до контейнерів, які мають бути вивантажені в порту, тобто потрібна оптимізація розміщення відповідно до маршруту для найкращої позиції кожного контейнера при максимальній завантаженості корабля.

II. ОСНОВНА ЧАСТИНА

На всіх GPS-системах у порту судна є точкою, що дуже ускладнює контроль за процесом завантаження, так як дослідити положення тіла на основі даних GPS не можна. Використання цієї технології перше, що спало на думку, коли мова зайшла про контроль за місцезнаходженням об'єктів. Так як перевізники не можуть ніяк модифікувати контейнер, то варіант із зовнішнім маркуванням також відпадає. Треба знайти спосіб контролювати положення тіла, який би не передбачав ніяких додаткових об'єктів-орієнтирів та дозволяв залишити контейнер саме у тому вигляді, в якому його надав клієнт. Відповідно до цих вимог було обрано координатний метод контролю.

Кожне тіло має координати в просторі. Увівши певну точку відліку можна точно сказати про місцезнаходження об'єкта у просторі. Звісно, одного набору координат буде замало інакше ми стикатимемося з ігноруванням фактичних розмірів тіла. Якщо ми розглядаємо контейнер (рис. 1), то говоримо про прямокутний паралелепіпед із заданими розмірами. Ідеально описати його положення у просторі можна зафіксувавши координати (x, y, z) усіх його вершин відносно певної базової точки (початку координат), але для спрощення, маючи розміри того чи іншого контейнера можна обмежитися координатами чотирьох точок, координати нижньої основи можна розрахувати, якщо вона буде паралельна до земної поверхні. Паралельність можна оцінити розглянувши координати z точок верхньої основи.

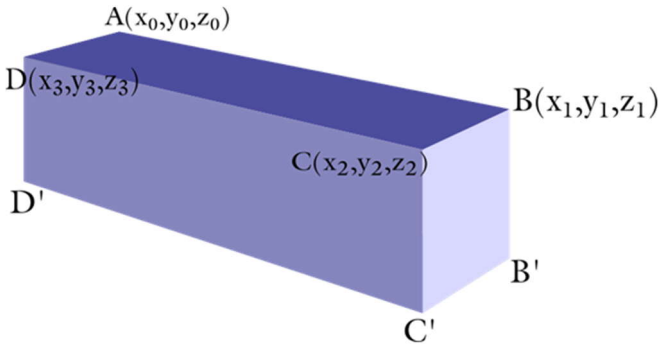


Рис. 1. Фіксовані точки для визначення координат (A' прихована)

Система працюватиме наступним чином:

1. Крановий, отримавши координати потрібного контейнера підіймає його краном. В момент зчеплення в базі виконується пошук інформації про вантаж, який буде перенесено. Отримавши розміри контейнера розраховуються координати всіх інших вершин.

2. Виконується переміщення контейнера на судно відповідно до координат, які були задані капітаном для даного контейнера.

3. Після того, як контейнер було встановлено на потрібне місце, суднова частина системи оповіщається про це. В портову базу відправляється команда, які підтверджує успішне перенесення вантажу.

Для того, щоб програма на борту корабля могла перевірити правильність розміщення треба мати карту зон завантаження з усіма потрібними параметрами. Під час старту система зчитує ці дані зі спеціального файлу і працює з ними не змінюючи оригінальний файл. Для можливих змін та корекцій можна застосовувати спеціальну утиліту, яка поставлятиметься разом з основною програмою.

В якості основи для файлу з параметрами судна було взято формат XML. Це дозволить використовувати карту з багатьма програмними інтерфейсами так як механізм запису/зчитування в XML безпосередньо чи через сторонні бібліотеки підтримують всі основні мови програмування.

На рис. 2 наведено карту збереження вантажів, інтерфейс якої дозволяє налаштувати специфічні параметри та оновити, чи відкоригувати карту, та на рис. 3 наведено приклад файлу в якому описані зони завантаження корабля.

Якщо ж власники корабля чи певні вимоги з безпеки обмежуватимуть відкритість даних, всю цю інформацію можна зберігати у бінарному файлі, тоді ніхто сторонній, не знаючи алгоритму заповнення цього файлу не зможе зчитати звідти корисну інформацію. На рис. 4 показано спробу зчитати бінарний файл за допомогою текстового редактора.

Спеціальна утиліта знатиме точно структуру такого файлу щоб, за необхідності, зчитати його чи змінити вносячи корективи у відповідні місця файлу. Для підтримки найвищого рівня захисту, якщо в цьому постане необхідність, можна додатково скористатися синхронними механізмами шифрування такими як AES чи RSA.

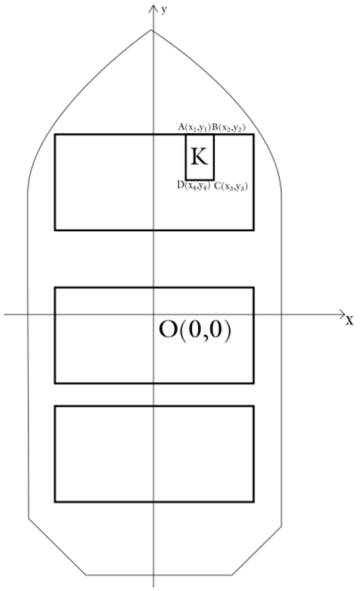


Рис. 2. Карта збереження вантажів, K - контейнер для прикладу, який характеризує розміщення

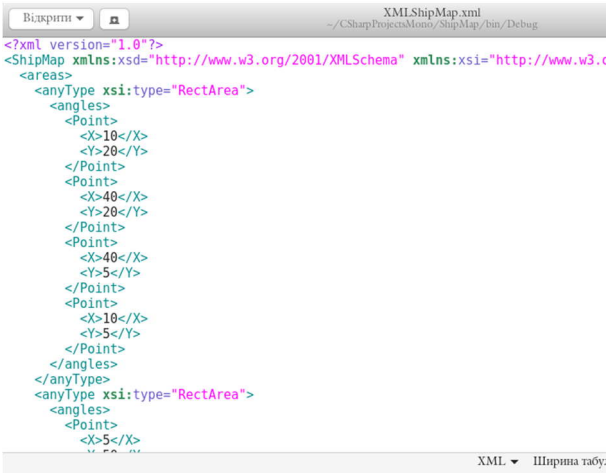


Рис. 3. Приклад фрагменту файлу з характеристиками корабля

Всі модулі, встановлені на кранах відправлятимуть інформацію на суднову частину системи. Вона являє собою комп'ютер приєднаний до роутера, який створює мережу через яку здійснюються комунікації. На комп'ютері встановлено програмне забезпечення, яке виконує обробку всіх даних, які надходять від кранів. Заповнюється судновий реєстр вантажів, який являє собою надбудову над картою завантаження і, по суті, поєднує координати інформацію про місцезнаходження об'єкта і дані про цей об'єкт. Напряму з даними, які надходять з мережі працюватиме лише один програмний модуль. Буде відбуватися логування та занесення контейнера до суднового реєстру вантажів. Всі інші програмні модулі працюватимуть з ним. В програмі передбачено графічний інтерфейс користувача, який дозволяє відобразити розміщення контейнерів на кораблі, модуль, який дозволяє

переглянути історію перевезень та модуль аналізу розміщення, який підбиратиме найбільш оптимальне положення для контейнера, а також модуль зв'язку через який здійснюватиметься в'язок між кораблем та крановими.

Рис. 4. Спроба зчитати бінарний файл (фрагмент)

Кранова ж частина системи складається з двох модулів Інерційної системи навігації (ІНС) та мікроконтролеру з вбудованим Web - сервером, через який здійснюватиметься передача інформації. ІНС визначає зміну положення тіла через вимірювання прискорень, які діють на тіло та його кутових швидкостей. На основі змін цих параметрів відбувається розрахунок координат, які показують відхилення тіла від точки початку руху (місце, в якому ініціалізовано ІНС). Датчики розташовано на захваті контейнеру. Таким чином в систему заводитимуться координати не контейнеру, а захвату крану.

Для розробки було обрано мікроконтролер Phoenix Contact, а саме модель ILC 150 GSM/GPRS 2916545. Зараз на його основі створюється модуль для закріплення на крані. Після його закінчення буде проведено серію повномасштабних тестів. Перевагою нашого модулю є те, що він може встановлюватися на майже будь-який кран, який уже встановлено в порту, що знижує витрати на її запровадження. Мікроконтролер під'єднається через інтерфейс RJ-45 до більш потужного роутера. Можлива пряма передача з мікроконтролера, якщо до його складу входить вбудований модуль Wi-fi. Таким чином маємо мережу, що зображена на рис. 5.

Нами також вивчається можливість використання додаткового каналу передачі даних з контролеру до комп'ютерної системи за допомогою GSM/GPRS.

III. Висновки

Розроблено комп'ютерну систему управління процесом завантаження транспортних суден.

Проаналізовано переваги та недоліки системи-конкурента з метою виявлення можливостей для покращення запропонованої нами системи.

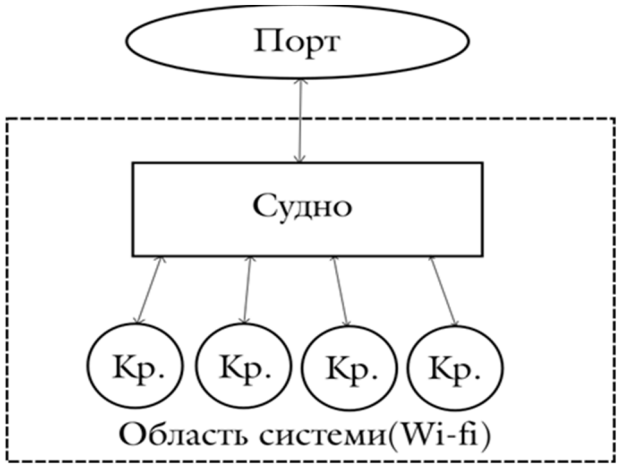


Рис. 5. Ієрархія системи

Створено модель роботи комп'ютерної системи завантаження контейнеровозу та план її реалізації із зазначенням алгоритму роботи.

Реалізовано структуру файлів для збереження інформації про судно.

Бездротові мережі та інерційні навігаційні системи покладено в основу реалізації обраного нами координатного методу. Запропоновано встановлення датчиків ІНС на захваті крану.

Було проведено експеримент, який демонструє приклад використання системи з використанням малопотужних аналогів у зменшеному масштабі, а також проаналізовано результати експерименту.

Література References

[1] Нікольський В. В. Система підтримки прийняття рішення по навантаженню великотоннажного контейнеровоза / В.В. Нікольський, М.В. Нікольський, Ю.А. Накул // Наукові праці ЧДУ ім. Петра Могили. – Серія: "Комп'ютерні технології". – Вип. 271. – Т. 283. – Миколаїв, 2016. – С. 60–63.

[2] Delta Marine | Ship Loading Software [Електорнний ресурс] – Режим

[3] доступа. – URL: http://www.deltamarine.com.tr/en/services/ship-loading-software.php (дата звернення 23.02.2017);

[4] DELFTload [Електронний ресурс] – Режим доступа. – URL:

[5] http://www.delftship.net/DELFTship/index.php/delftload/delftload (дата звернення 24.02.2017);

[6] MEECOS Suite: INTERSCHALT maritime systems [Електронний ресурс] – Режим доступа. – URL: http://www.interschalt.com/software/meecos-suite/ (дата звернення 20.02.2017);

[7] What is Inertial Navigation Systems Guide | OxTS [Електронний ресурс] – Режим доступа. – URL: http://www.oxts.com/what-is-inertial-navigation-systemsguide/ (дата звернення 02.03.2017);

[8] Роберт Тейбор. Реалізація XML Web-служб на платформі Microsoft .NET = Реализация XML Web-служб на платформе Microsoft .NET. — М. : «Вільямс», 2002. — С. 464. — ISBN 0-6723-2088-6.

[9] Баричев С. Г., Гончаров В. В., Серов Р. Е. 2.4.2. Стандарт AES. Алгоритм Rijdael // Основы современной криптографии — 3-е изд. — М.: Диалог-МИФИ, 2011. — С. 30–35. — 176 с. — ISBN 978-5-9912-0182-7.

[10] Тилборг ван Х. К. А. Основы криптологии = Fundamentals of Cryptology. — М. : Мир, 2006. — 472 с.

Скорочення Необхідного Ресурсу ПЛІС Для Обчислювальних Систем на Кристалі

Валерій Жабін
каф. обчислювальної техніки
НТУУ «КПІ ім. Ігоря Сікорського»
Київ, Україна
vizhabin@comsys.kpi.ua

Валентина Жабіна
каф. програмного забезпечення комп'ютерних систем
НТУУ «КПІ ім. Ігоря Сікорського»
Київ, Україна
val.zhabina2@gmail.com

Reduction of Necessary FPGA Source for Computing System on Chip

Valerii Zhabin
dept. of Computer Engineering
NTUU "Igor Sikorsky Kyiv Polytechnic Institute"
Kyiv, Ukraine
vizhabin@comsys.kpi.ua

Valentina Zhabina
dept. of Computer Systems Software
NTUU "Igor Sikorsky Kyiv Polytechnic Institute"
Kyiv, Ukraine
val.zhabina2@gmail.com

***Анотація*—Показана можливість зменшення необхідного апаратного ресурсу ПЛІС за рахунок порозрядної обробки даних в системах з безпосередніми зв'язками між обчислювальними модулями. Наведено приклад реалізації такої системи для обчислення степеневих функцій з порозрядною обробкою операндів зі старших розрядів в симетричних надлишкових системах числення. Запропонована реалізація обчислень дозволяє при порівняно високій швидкодії скоротити необхідний ресурс ПЛІС завдяки малому числу зв'язків між модулями**

***Abstract*—In the article there is shown a possibility to reduce necessary FPGA hardware sources by means of digit-by-digit data processing in data-flow systems with direct couplings between processing units. One of the approaches to solving the problem of reduction of number of connections between processing units is to use quasi-parallel processors, which enable to realize concurrently the processes of digit-by-digit operands input, as well as the process of digit-by-digit result output starting from higher positions. The digit-by-digit data interchange starting from higher positions enables not only to reduce the number of connections between processing units, but also to provide realization of data-flow operations in concurrency mode. In a simple case at each step of calculations one operand position is put into processing unit starting from the higher one, as well as one result position is formed. In this calculation mode, the next operation execution will not be started after previous operation is performed, but right after first result position is received. Samples of algorithms of calculation of power functions in on-line mode with the use of symmetrical redundant numerical system with digits {-1, 0, 1} for operands representation are considered. To raise numbers to powers there is used the binary tree method. Elements of internal sources (execution and communication units) as well as FPGA output pins are saved. As a result of the above-mentioned, a possibility to place the system on one microchip is increased. It enables to increase system reliability and reduce energy usage and dimensions.**

Ключові слова—неавтономний режим; степеневі функції; порозрядна обробка; системи на кристалі; ПЛІС

Keywords—on-line mode; power functions; digit-by-digit processing; SoC; FPGA

I. Вступ

Тривалість обробки інформації у паралельних системах залежить не тільки від часу виконання операцій, але і від затрат часу на обмін інформацією між гілками алгоритмів, тобто між обчислювальними модулями (ОМ) паралельної системи. Зменшити затрати часу на обмін даними дозволяє використання потокових систем з безпосередніми зв'язками (ПСБЗ) між ОМ [1, 2].

У ПСБЗ виходи одних ОМ підключаються до входів інших ОМ (можливо через елементи комутації) відповідно до графа потоку даних (ГПД). ОМ працюють у неавтономному (on-line) режимі. У процесі обчислень дані пересилаються безпосередньо від одних ОМ до інших, перетворюючись на кожному кроці відповідно до операцій, що задані вершинами ГПД. У такому випадку відсутні складні процедури пересилання даних між ОМ через загальну пам'ять, тобто створюються передумови до

зменшення непродуктивних витрат часу в процесі обміну інформацією.

Досягнення в області інтегральної технології дозволяють створювати паралельні системи за технологією SoC (System on Chip), які можна реалізувати на ПЛІС з використанням інтерактивних засобів проектування, що суттєво прискорюють термін розробки систем. Однак з боку елементної бази накладається ряд обмежень, пов'язаних з числом виводів (pins) мікросхем, наявністю вбудованих функціональних вузлів і пристроїв. Наприклад, ресурсів однієї мікросхеми може не вистачити для реалізації всієї системи, тобто може виникнути необхідність застосування кількох мікросхем, пов'язаних між собою зовнішніми лініями зв'язку. Це визначається як розмірністю задач, так і складністю обчислювальних засобів, що залежить від алгоритмів виконання операцій.

Витрата необхідних ресурсів ПЛІС багато в чому визначається способом передачі інформації між компонентами системи. При передачі даних паралельним кодом витрачається великий обсяг комунікаційних засобів мікросхеми, а також виникають проблеми, пов'язані з можливою нестачею виводів мікросхеми для введення та виведення даних. При нестачі виводів можлива ситуація, коли частина функціональних ресурси мікросхеми може виявитися незадіяною.

З огляду на важливість цієї проблеми, компанія Virtual Machine Works розробила технологію під назвою VirtualWire (віртуальні з'єднання), представлену як технологія виробництва великих систем, для реалізації яких доводиться використовувати кілька ПЛІС [3]. Ідея, закладена в основі технології, полягає в тому, що частина ресурсів мікросхем використовується для реалізації спеціальних схем, які дозволяють підключати виводи мікросхем поперемінно до різних джерел сигналів усередині ПЛІС. Дана технологія, хоча і допомагає вирішити проблему нестачі виводів, не дозволяє в повній мірі використовувати незадіяні ресурси мікросхем для вирішення безпосередньо основної задачі. Крім того, мультиплексування створює додаткові часові затримки просування потоків даних, що суперечить самій ідеї потокової моделі обчислень.

Таким чином, існує важлива проблема зменшення апаратних ресурсів ПЛІС при побудові на їх основі обчислювальних систем класу ПСБЗ. В роботі досліджується можливість вирішення даної проблеми шляхом скорочення кількості зв'язків між компонентами системи за рахунок порозрядної передачі даних.

II. Постановка задачі

Одним з підходів до вирішення проблеми зменшення кількості зв’язків між компонентами системи є використання квазіпаралельних ОМ, що обмінюються між собою даними послідовним кодом. Хоча на входах і виходах таких ОМ числа представлені послідовним кодом, вони по внутрішній організації ближчі до паралельних пристроїв. У зв’язку з цим вони отримали назву квазіпаралельні [1, 4]. За швидкістю виконання арифметичних операцій (множення, ділення тощо), такі пристрої ближче до паралельних пристроїв.

Порозрядний обмін даними дозволяє зменшити необхідну кількість зовнішніх виводів і необхідний ресурс комутаційної системи ПЛІС. Підвищується ймовірність реалізації системи на одній або на меншій кількості мікросхем. У свою чергу, це дає можливість підвищення надійності систем, зменшення енергоспоживання та габаритів.

Нижче на прикладі обчислення степеневих функцій розглядаються структури на базі квазіпаралельних ОМ, що працюють в неавтономному режимі з використанням надлишкової системи числення для подання операндів та результату. Обґрунтовано вибір надлишкової системи числення. Запропоновано структуру ОМ і алгоритми виконання операцій. Показана можливість виконання залежних операцій в режимі часткового суміщення на рівні обробки розрядів операндів. Порівнюються системи на базі ОМ з послідовним та паралельним обміном інформацією. Одержано формули для оцінки часу виконання обчислень в різних системах.

III. Організація апаратних засобів

Узагальнену функціональну модель систем класу ПСБЗ можна подати у вигляді

$$SYS = \langle DI, DO, M, F, C \rangle,$$

де DI – множина вхідних даних; DO – множина результатів; M – множина ОМ; F – система функцій перетворення даних; C – характеристика засобів комутації.

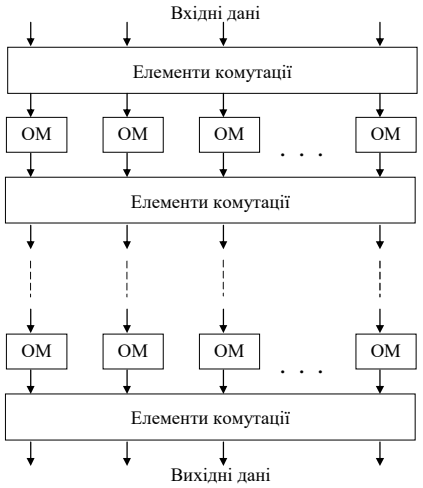


Рис. 1. Структура системи класу ПСБЗ, що реконфігурується

Система функцій F подається в вигляді алгоритмів виконання операцій в ОМ. Характеристика C визначає можливості взаємодії ОМ у просторі.

На рис. 1 показано приклад організації ПСБЗ для реалізації ярусно-паралельних форм алгоритмів. ОМ налагоджуються на виконання певних операцій, а

елементи комутації забезпечують зв'язок між ОМ. Формалізовані методики переходу від ГПД до структури ПСБЗ відомі [2, 5, 6].

В квазіпаралельних ОМ відбувається суміщення процесів порозрядного введення операндів, їх обробки і виведення результату. Це дає потенційну можливість виконувати залежні операції (результат однієї операції є операндом для іншої операції) у режимі часткового суміщення в часі.

На кожному кроці обчислень в ОМ вводиться по одному розряду операндів, починаючи зі старших, і формується один розряд результату. При цьому розряд проміжного результату, отриманий на i-му кроці в одному ОМ при виконанні j-ї операції, може бути використаний на $(i+1)$-му кроці в іншому ОМ при виконанні $(j+1)$-ї операції. При такому режимі обчислень виконання наступної операції буде починатися не після завершення виконання попередньої операції, а відразу ж після отримання першого розряду результату цієї операції. Для виконання послідовності операцій необхідна сумісна робота кількох ОМ, що обмінюються інформацією у процесі роботи.

ПЛІС дозволяють конфігурувати логічні блоки для роботи у якості ОМ, а також за допомогою спеціальних засобів можна сформувати необхідний набір команд [3]. Цей факт є дуже важливим для побудови ОМ зі спеціальними алгоритмами виконання операцій, в тому числі, з використанням надлишкової системи числення.

ПСБЗ містить однотипні ОМ. Для обчислення степені чисел використовується метод бінарного дерева [7].

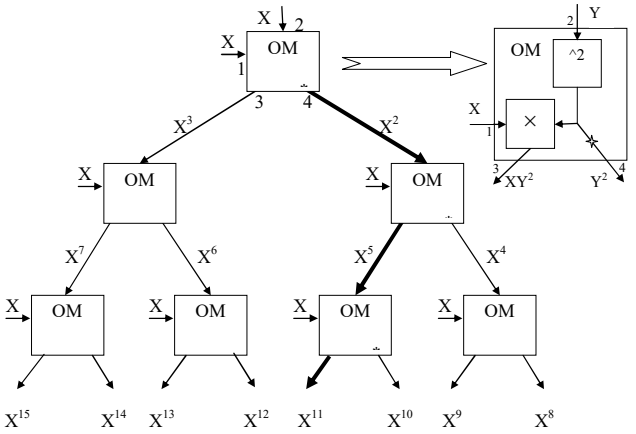


Рис. 2. Дерево однотипних модулів для обчислення степеневих функцій

На рис. 2 показано схему з'єднання ОМ, що дозволяє отримувати всі степені числа від 2-ї до 15-ї. У загальному випадку дерево, що містить m ярусів, дозволяє обчислювати функції $f(X)=X^j$, де $j=\overline{2,M}$. Максимальний показник степені визначається формулою $M=2^{m+1}-1$.

На рис. 2 окремо показана внутрішня організація ОМ. Модуль включає пристрій множення і пристрій для обчислення квадрата числа (позначені відповідно через «×» і «^2»). Для вирівнювання затримок формування розрядів результатів на обох виходах на відміченому зірочкою виході ОМ включені елементи затримки.

Якщо в один момент часу необхідно одержати значення будь якої із зазначених функцій, що формуються на старшому ярусі, то кількість ОМ може бути зменшена до m практично без зниження швидкодії. При обчисленні функції X^Q шлях проходження інформації крізь дерево ОМ визначається m молодшими розрядами двійкового коду показника степеня $Q=q_1q_2q_3q_4$. Розряд q_i керує виходами ОМ на $(i-1)$-му ярусі. При рівності розряду q_i одиниці в ОМ $(i-1)$-го ярусу формується функція XY^2 (див. рис. 2), а при рівності нулю – функція Y^2. В якості прикладу на рис. 2 потовщеною лінією показаний шлях проходження інформації через дерево ОМ при обчисленні 11-го степеня числа (Q = 1011).

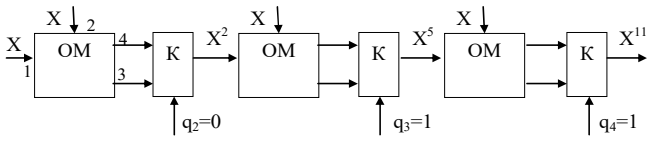


Рис. 3. Лінійне з'єднання модулів для обчислення степеневих функцій

Таким чином, для обчислення функції X^Q, де Q – ціле $\left(2^m \le Q \le 2^{m+1}-1\right)$, достатньо мати m ОМ, з'єднаних між собою, як вказано на рис. 3, де К – комутатор.

IV. Алгоритми виконання операцій

Неавтономній режим виконання операцій зі старших розрядів можливий як у зміщених, так і симетричних надлишкових системах числення [4, 8]. Наприклад, з основою $k=2$ це відповідно системи з цифрами $g=\{0,1,2\}$ та $g=\{-1,0,1\}$. В зміщених системах для визначення знаку числа необхідно мати спеціальний знаковий розряд, який при виконанні операцій треба обробляти окремо, що ускладнює обробку чисел зі знаками. В симетричних системах знак числа визначається знаком старшого основного розряду числа, тобто знаковий розряд не потрібен. Всі розряди обробляються однаково, що дає можливість прискорити виконання операції. Якщо операнди можуть мати різні знаки, то доцільно виконувати операції в симетричних системах числення.

В ОМ виконуються дві операції: множення ($Z=2^{-p}XY$) та піднесення до квадрату ($Z=2^{-p}X^2$). В неавтономному режимі обчислень кожний операційний пристрій вносить затримку формування старшого розряду на p кроків.

Будемо вважати, що операнди є нормалізованими дробовими числами (мантисами) і представлені у формі:

$$X = \sum_{i=1}^{n} x_i 2^{-i},\ Y = \sum_{i=1}^{n} y_i 2^{-i}, \qquad (1)$$

де $x_i, y_i \in \{-1, 0, 1\}$ – цифри операндів, n – розрядність операндів.

В даній надлишковій системі одне і те ж саме число може мати кілька подань, наприклад, $0{,}1011 = 0{,}11\bar{1}1 = 0{,}1\bar{1}01$ (символ $\bar{1}$ є більш зручної формою запису цифри -1). Це дозволяє в ОМ формувати результат порозрядно без переносів у старші розряди.

Операнди, що містять тільки i розрядів праворуч від коми, позначимо відповідно через X_i та Y_i. Наприклад, $X_3 = x_1 x_2 x_3 0...0$, $Y_2 = y_1 y_2 0...0$.

Цифри z_i результату Z формуються зі старших розрядів в такий самій системі числення порозрядно. Розряди операндів і результату передаються між ОМ за допомогою двох провідників. Цифри з множини $\{$-1, 0, 1$\}$ можуть кодуватись відповідно двома сигналами $\{$10, 00, 01$\}$.

A. Алгоритм множення.

Вимагатимемо, щоб похибка результату була знакозмінною і за абсолютною величиною не перевищувала половини ваги n -го розряду після коми. Ця вимога буде виконуватися, якщо на i -му кроці цифру z_i результату вибирати таким чином, щоб мало місце співвідношення

$$Z_i - 2^{-i-1} \le 2^{-p} X_i Y_i < Z_i + 2^{-i-1}, \qquad (2)$$

де p – число кроків латентної затримки початку формування цифр результату множення в ОМ.

Використовуючи методику, описану в [4], і формули (1), (2), можна отримати алгоритм обчислення Z у неавтономному режимі для обраної форми подання чисел.

В даному випадку $p = 2$. Для отримання n розрядів результату після коми необхідно виконати $n+2$ кроки обчислення.

Алгоритм обробки операндів має наступній вигляд.

1. X_0, Y_0, R_0 надати значення 0.
2. Для $i = \overline{1, n+2}$ виконувати пункти 3-7.
3. $H_i = 2R_{i-1} + 2^{-2} X_{i-1} y_i + 2^{-2} Y_{i-1} x_i + 2^{-2-i} x_i y_i$.
4. $X_i = X_{i-1} + x_i 2^{-i}$.
5. $Y_i = Y_{i-1} + y_i 2^{-i}$.
6. $z_i = \begin{cases} -1, \text{ якщо } H_i < -2^{-1}; \\ 0, \text{ якщо } -2^{-1} \le H_i < 2^{-1}; \\ 1, \text{ якщо } 2^{-1} \le H_i. \end{cases}$
7. $R_i = H_i - z_i$.

В алгоритмі H_i і R_i – допоміжні змінні.

B. Алгоритм обчислення квадрату числа.

Аналогічно операції множення висунемо вимогу, щоб похибка результату була знакозмінною і за абсолютною величиною не перевищувала половини ваги n -го розряду після коми. Ця вимога буде виконуватися, якщо на i -му кроці цифру z_i результату вибирати таким чином, щоб мало місце співвідношення

$$Z_i - 2^{-i-1} \le 2^{-p} X_i^2 < Z_i + 2^{-i-1}. \qquad (3)$$

Використовуючи методику, описану в [4], і формули (1), (3), можна отримати алгоритм обчислення Z у неавтономному режимі.

1. X_0, R_0 надати значення 0.
2. Для $i = \overline{1, n+2}$ виконувати пункти 3-6.
3. $H_i = 2R_{i-1} + 2^{-1} X_{i-1} x_i + 2^{-2-i} x_i^2$
4. $X_i = X_{i-1} + x_i 2^{-i}$.
5. $z_i = \begin{cases} -1, \text{ якщо } H_i < -2^{-1}; \\ 0, \text{ якщо } -2^{-1} \le H_i < 2^{-1}; \\ 1, \text{ якщо } 2^{-1} \le H_i. \end{cases}$
6. $R_i = H_i - z_i$.

V. Порівняльна оцінка систем

Послідовність виконання операцій у неавтономному режимі ілюструється діаграмою на рис. 4.

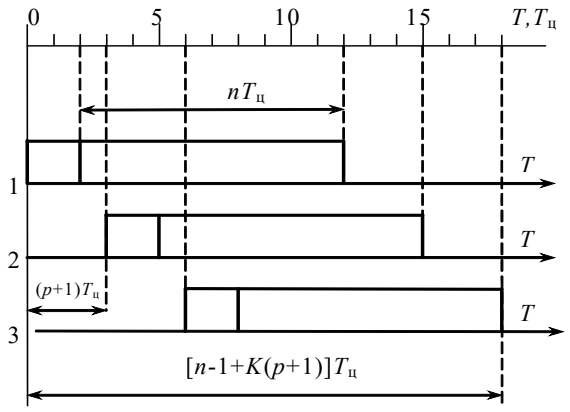


Рис. 4. Часова діаграма виконання залежних операцій в 3-х ОМ

Операції виконуються в неавтономному режимі за допомогою ланцюга ОМ, що формують цифри проміжних результатів із затримкою на p циклів. Наступний у ланцюгу ОМ починає свою операцію через ($p+1$) циклів відносно попереднього ОМ. Це означає, що час обчислень, необхідний для отримання остаточного результату, складає

$$T = \left[n - 1 + \sum_{j=1}^{K} \left(p_j + 1 \right) \right] t_0 , \qquad (4)$$

де n – розрядність результату; K – число операцій у ланцюгу; p_j – затримка формування розрядів результату j -ї операції; $T_{\text{ц}}$ – тривалість циклу формування одного розряду результату в ОМ.

У випадку, що розглядається, для кожної операції затримка складає $p = 2$. Наприклад, для обчислення 15-го степеня числа необхідний ланцюг з 3-х ОМ, в кожному з яких зв'язані між собою 2 операційних пристроя, тобто в неавтономному режимі працюють 6 пристроїв. Наприклад, для 64-розрядних операндів знадобиться виконати 64-1+18=81 цикл обчислень. Тривалість одного циклу визначається в основному обчисленням змінної H_i у приведених вище алгоритмах. Вказаний час визначається додаванням 3-х чисел для операції обчислення квадрату числа і 4-х чисел для операції множення за допомогою дерева комбінаційних суматорів і складає одиниці наносекунд при реалізації на ПЛІС.

При використанні для виконання послідовності залежних за даними операцій паралельних ОМ, що працюють в канонічній двійковій системі числення, час виконання ланцюга операцій буде визначатися формулою

$$T = T_0 + \sum_{i=1}^{K} t_i , \qquad (5)$$

де t_i – час виконання i -ї операції; T_0 – додаткові витрати часу. До останніх витрат можна віднести, наприклад, витрати на вибір і розшифрування команд при програмній реалізації або на процедури обміну даними між ОМ (можливо, через загальну пам'ять), прийом операндів до пристрою, тощо.

Для кількісної порівняльної оцінки часу обчислень за формулами (4) та (5) необхідний більш детальний розгляд конкретної реалізації ОМ паралельного типу. Враховуючи, що при одночасному паралельному вводі операндів до ПЛІС втрачається багато зовнішніх виводів, навряд чи такий режим вводу є ефективним. Більш реальним є режим попереднього вводу 64-розрядних операндів у певному порядку з наступною їх обробкою у паралельному пристрої, що додатково затримує обчислення.

Крім того, під час передачі інформації між ОМ паралельним кодом використовується багато комунікаційного ресурсу ПЛІС, тобто апаратний ресурс ПЛІС витрачається неекономно.

VI. Висновки

У роботі досліджена можливість підвищення ефективності систем з безпосередніми зв'язками між ОМ за рахунок реалізації неавтономних обчислень при порозрядній передачі даних між модулями системи.

У порівнянні з системами, в яких передача даних між ОМ проводиться паралельним кодом, для реалізації систем з порозрядною обробкою даних на базі квазіпаралельних ОМ витрачається менше ресурсів ПЛІС. Економляться як елементи внутрішнього ресурсу (функціональні і комунікаційні модулі, вбудовані пристрої), так і виводи ПЛІС.

Завдяки цьому збільшується ймовірність розміщення системи на одній мікросхемі. Реалізація системи на одній мікросхемі забезпечує підвищення надійності систем, зменшення енергоспоживання та габаритів.

Обмін даними між ОМ всередині мікросхеми виконується швидше, ніж між компонентами системи, що реалізовані на різних мікросхемах. Це дає потенційну можливість підвищити частоту тактування, що, у свою чергу, прискорює обробку інформації.

Таким чином, отримані результати підтверджують ефективність застосування неавтономних методів порозрядної обробки інформації в обчислювальних системах на кристалі.

Література References

[1] В.И.Жабин, «Выполнение последовательностей зависимых операций в режиме совмещения», *Вісник НТУУ «КПІ». Інформатика, управління та обчислювальна техніка*, №46, с. 226-233, 2007.

[2] В.И.Жабин, В.И.Корнейчук, В.П.Тарасенко, «Построение быстродействующих специализированных вычислителей для реализации многоместных выражений», *Автоматика и вычислительная техника*, №6, с. 18-22, 1981.

[3] К. Максфилд, *Проектирование на ПЛИС. Архитектура, средства и методы*, М: Издательский дом «Додэка-XXI», 2007, 408 с.

[4] V.I.Zhabin, V.I.Korneĭchuk, V.P.Tarasenko, "Computation of rational functions for many arguments", *Automation and Remote Control*, vol. 38, no 12, pp. 1864–1871, 1978. [В.И.Жабин, В.И.Корнейчук, В.П.Тарасенко, «Некоторые машинные методы вычисления рациональных функций многих аргументов» *Автомататика и телемеханика*, № 12, с. 145–154, 1977].

[5] А.В.Палагин, В.Н.Опанасенко, *Реконфигурируемые вычислительные системы*: *Основы и приложения*, К: Просвіта, 2006, 280 с.

[6] И.А.Каляев, И.И.Левин., Е.А.Семерников, «Ахитектура семейства реконфигурируемых вычислительных систем на основе ПЛИС», *Искусственный интеллект*, № 3, с. 663-674, 2008.

[7] Д.Кнут, *Искусство программирования для ЭВМ. Т. 2. Получисленные алгоритмы*, М: «Мир», 1977, 724 с.

[8] И.А.Дичка, В.В.Жабина, «Совмещение зависимых операций на уровне обработки разрядов операндов», *Искусственный интеллект*, №3, с. 649-654, 2008.

Глибокі Нейронні Мережі для Вирішення Завдань Розпізнавання і Класифікації Зображення

Віктор Синєглазов
Кафедра авіаційних комп'ютерно-інтегрованих комплексів,
Навчально-науковий інститут інформаційно-діагностичних систем,
Національний авіаційний університет, Київ, Україна
svm@nau.edu.ua

Олена Чумаченко
Кафедра технічної кібернетики,
Національний технічний університет України «Київський політехнічний інститут ім. Ігоря Сікорського», Київ, Україна
chumachenko@tk.kpi.ua

Deep Neural Networks for Solution Problems of Recognition and Classification of Images

Victor Sineglazov
Computer-Integrated Complexes Department,
Educational and Scientific Institute of Information and Diagnostic Systems, National Aviation University
Kyiv, Ukraine
svm@nau.edu.ua

Olena Chumachenko
Technical Cybernetic Department
National Technical University of Ukraine
"Ihor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine
chumachenko@tk.kpi.ua

***Анотація*—Розглянуто вирішення завдання обробки зображень використовуючи згорткові нейронні мережі. Показано, що незважаючи на свою ефективність вони мають недоліки, які полягають у відсутності оптимального вибору кількості карт ознак, розмірів вікна, площі перекриття, чергування шарів різних типів. У результаті досліджень поставлено і розв'язано задачу пошуку оптимальних значень вищевказаних параметрів на основі використання гібридного генетичного алгоритму. Синергетичний ефект запропонованого алгоритму реалізується завдяки використанню в ньому елементів алгоритмів найбільш поширених модифікацій генетичних алгоритмів, що реалізують різні схеми призначення придатності і селекції. Розглянуто приклад побудови оптимальної структури загорткової мережі.**

***Abstract*— It is considered the problem of image processing using convolutional neural network. It is shown that despite their effectiveness they have disadvantages that are the lack of optimal choice of number of features maps, window size, the area of overlap, alternating layers of different types. Based on the research results it is determined and solved the problem of the above parameters optimal values search, based on hybrid genetic algorithm. It is considered the design of convolutional neural network.**

Ключові слова—обробка зображень; генетичний алгоритм; синергетичний ефект; згорткові нейронні мережі.

Keywords—image processing; genetic algorithm; synergistic effect; convolutional neural network.

I. Вступ та постановка завдання

Постановка задачі класифікації. Нехай X – множина описань об'єктів, Y – множина номерів (чи найменувань) класів. Існує невідома цільова залежність – відображення y^*: $X \to Y$, значення якої відомі тільки на об'єктах кінцевої навчальної вибірки:

$$X^m = \{(x_1, y_1), \ldots, (x_m, y_m)\},$$

де X^m множина елементів навчальної вибірки розмірністю m.

Потрібно побудувати алгоритм здатний визначити належність довільного об'єкта $x \in X$ до класу $y \in Y$.

II. Використання глибоких нейронних мереж для вирішення задачі розпізнавання образів неформалізованих елементів

Застосування багатошарового персептрона з традиційною структурою при вирішенні реальних задач розпізнавання і класифікації зображень викликає певні труднощі. По-перше, зображення, як правило, мають велику розмірність, внаслідок чого зростає число нейронів

і синаптичних зв’язків у мережі. У свою чергу, це вимагає збільшення навчальної вибірки, внаслідок чого збільшується час і обчислювальна складність процесу навчання. По-друге, ігнорується топологія вхідних даних. Компоненти вхідного шару можуть бути представлені в будь-якому порядку, без урахування мети навчання. Однак зображення мають строгу двовимірну структуру, в якій існує залежність між просторово-сусідніми пікселями.

Від цих недоліків вільні так звані згорткові нейронні мережі, які являють собою особливий клас багатошарових персептронів, спеціально створених для розпізнавання двовимірних поверхонь з високим ступенем інваріантності до масштабування, зсуву, повороту, зміни ракурсу та інших просторових спотворень, і глибокі НМ побудовані на базі автоенкодерів, попереднє навчання яких здійснюється на базі обмежених машин Больцмана.

III. Згорткові нейронні мережі підвищеної ефективності

Згорткова нейронна мережа (ЗНМ) – тип багатошарової нейронної мережі, яка свою назву «згорткова мережа» отримала за назвою операції – згортка, вона часто використовується для обробки зображень і може бути описана наступною формулою

$$(\mathbf{f} \times \mathbf{g})[m,n] = \sum_{k,l} \mathbf{f}[m-k, n-l] \cdot \mathbf{g}[k,l],$$

де **f** – вихідна матриця зображення; **g** – ядро (матриця) згортки.

Неформально цю операцію можна описати наступним чином – вікном розміру ядра **g** проходимо з заданим кроком (зазвичай 1) все зображення **f** на кожному кроці поелементно множимо вміст вікна на ядро **g**, результат сумується і записується в таблицю результату.

Ідея згорткових нейронних мереж полягає в чергуванні згорткових шарів (англ. convolution layers) і субдискретизуючих шарів (англ. subsampling layers, верств підвибірки). Структура мережі – односпрямована (без зворотних зв’язків), багатошарова (рис. 1).

Модель згорткової мережі складається з трьох типів шарів: згорткові (convolutional) шари, субдискретизуючі (subsampling, підвибірка) верстви і прошарки «звичайної» нейронної мережі – персептрона.

Архітектура згорткових нейронних мереж реалізує три ідеї, які забезпечують інваріантність мережі до невеликих зрушень, змін масштабу і спотворень:

- кожен нейрон отримує вхідний сигнал від локального рецептивного поля (local receptive fields) у попередньому шарі, що забезпечує локальну двовимірну зв’язність нейронів;
- кожен прихований шар мережі складається з безлічі карт ознак, на яких всі нейрони мають загальні ваги (shared weights), що забезпечує інваріантність до зміщення і скорочення загальної кількості вагових коефіцієнтів мережі;
- за кожним шаром згортки слідує обчислювальний шар, який здійснює локальне усереднення та підвибірку, що забезпечує зменшення розширення для карт ознак.

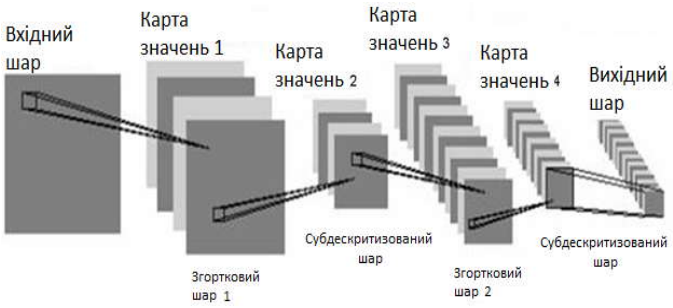


Рис. 1. Структура згорткової мережі.

Робота згорткової нейронної мережі забезпечується двома основними елементами.

1) Фільтри (filters) (визначники ознак).

2) Карти ознак (feature maps).

Фільтр – це невелика матриця, що представляє ознаку, яку необхідно знайти на вихідному зображенні. За допомогою верхнього фільтра визначаються частини вихідного зображення з вертикальними лініями, нижній фільтр служить для визначення частин зображення з горизонтальними лініями.

Безпосередньо процес визначення заснований на операції згортки фільтром оригінального зображення. Результати згортки, які визначають місце розташування ознак вихідного зображення, називаються картами ознак.

Мета процесу згортки – зменшити розмірність карти ознак до такої міри, щоб з повним набором ознак могла працювати мережа прямого поширення (в більшості випадків багатошаровий персептрон).

Згортковий шар реалізує ідею локальних рецептивних полів, тобто кожен вихідний нейрон з’єднаний тільки з певною (невеликою) областю вхідної матриці і таким чином моделює деякі особливості людського зору.

Недоліками згорткових нейронних мереж (ЗНМ) є:

- висока складність архітектури;
- повнозв’язаність;
- фіксована площа вікна шару згортки.

З метою підвищення ефективності роботи ЗНМ необхідно знайти оптимальні значення наступних параметрів:

- кількість карт ознак;
- щільність зв'язків між картами ознак;
- розмір вікна;
- площа перекриття;
- початкова ініціалізація ваг.

IV. Генетичні алгоритми

Генетичні алгоритми належать до класу еволюційних алгоритмів і мають ряд характеристик, які роблять їх кращими, ніж класичні методи оптимізації:

- для пошуку ефективних рішень з використанням генетичних алгоритмів не потрібні специфічні знання про саму задачу і параметри, що входять до неї;
- в генетичних алгоритмах замість детермінованих використовуються стохастичні оператори, які показали себе досить стійкими в умовах зашумленості зовнішнього середовища;
- властивий генетичним алгоритмам паралелізм – одночасне врахування великої кількості індивідів популяції – робить їх менш чутливими до локальних оптимумам і впливу шумів.

V. Загальна конструкція генетичного алгоритму

Використання генетичних алгоритмів для розв'язання задач багатокритеріальної оптимізації, дозволяє позбутися від основних недоліків класичних методів, так як генетичні алгоритми підходять для задач великої розмірності і здатні захопити Парето-оптимальні точки навіть при одноразовому запуску алгоритму. За допомогою підтримки популяції рішень і застосування концепції Парето-оптимальності, генетичні алгоритми можуть знаходити різні Парето-оптимальні рішення паралельно.

Таким чином, на відміну від більшості класичних підходів до розв'язання задач багатокритеріальної оптимізації, коли для отримання кожної окремої точки необхідно проводити окремий запуск алгоритму пошуку Парето-оптимальних рішень, застосовуючи еволюційний підхід до векторної оптимізації, завдяки закладеному в генетичних алгоритмах паралелізму, можливо отримання різних точок множини Парето при одному прогоні алгоритму. Дана обставина є очевидною перевагою еволюційного підходу до вирішення задач багатокритеріальної оптимізації перед традиційними методами їх вирішення [1] – [4].

За аналогією з природною еволюцією, в генетичних алгоритмах кандидати-рішення називаються ***індивідами*** або ***особинами***, а безліч кандидатів-рішень – ***популяцією***. Кожен індивід визначає можливе рішення задачі, при цьому, однак, сам по собі він не є вектором рішень, а скоріше кодує його, ґрунтуючись на відповідній структурі кодування рішення. У генетичних алгоритмах ця структура визначається вектором – вектором бітів або вектором дійсних чисел – набором генів, що утворюють ***хромосоми***. Безліч всіх можливих векторів утворює простір індивідів або популяцію.

Якість індивіда під час розв'язання задачі оптимізації визначається скалярним значенням, так званою ***функцією пристосованості (fitness function)***.

В процесі ***селекції,*** який може бути, як стохастичним, так і детермінованим, гірші рішення – непристосовані індивіди – видаляються з популяції, в той час як індивіди з більшою пристосованістю – найбільш пристосовані – піддаються репродукції. Мета полягає в тому, щоб посилити пошук в певних областях пошукового простору і збільшити середню «якість» всередині популяції.

Проблемою навчання глибоких нейронних мереж обробки зображень, що використовуються в системі медичної діагностики, є складність отримання навчальної вибірки. Навчальна вибірка виходить на підставі наявної невеликої кількості прикладів шляхом повороту їх на невеликі кути в різних площинах з подальшою обробкою зображень іншим способом.

Задача розпізнавання зображень в такому формулюванні представляє собою досить складну комплексну задачу з низкою обмежень. В роботі пропонується системний підхід до обробки зображень, заснований на структурній декомпозиції (рис. 2).

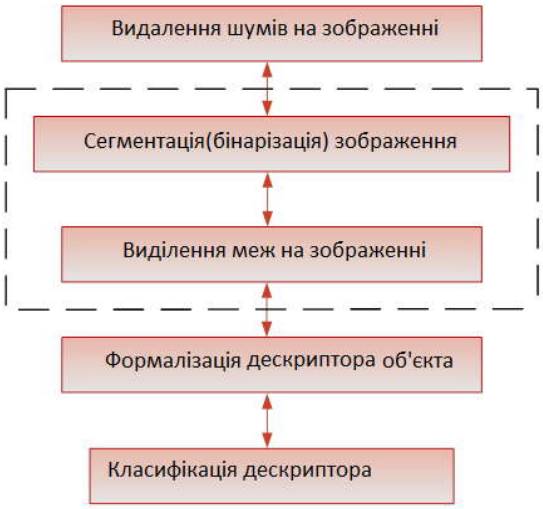


Рис. 2. Алгоритм обробки зображень створення навчальної виборки для згорткової нейронної мережі.

VI. Приклад розрахунку оптимальної структури згорткової мережі

Спробуємо підібрати оптимальну структуру згорткової мережі для розв'язання аналогічної задачі (прогнозування позиції у просторі маніпулятора робота). Розрахунок відрізнятиметься від попереднього прикладу лише способом кодування (і тлумачення) структури згорткової мережі.

Загальну структуру гена для кодування одного шару згорткової мережі показано на рис. 3.

Тип шару (CONV, POOL, FC, ABS)	Параметр 1	Параметр 2	Параметр 3

Рис. 3. Загальна структура гена для кодування одного шару згорткової мережі.

Тоді хромосома, що кодує топологію згорткової мережі буде складатися з N подібних генів, де N – задана максимальна кількість шарів. Для демонстрації розглянемо випадок $N = 4$.

Оскільки кожен тип шару має свої специфічні (часто несумісні) параметри, то тлумачення кожного параметрів гену залежатиме від типу шару. Призначення кожного з параметрів залежно від типу шару наведено в табл. I.

ТАБЛИЦЯ I. ПРИЗНАЧЕННЯ КОЖНОГО З ПАРАМЕТРІВ ЗГОРТКОВОЇ НЕЙРОННОЇ МЕРЕЖІ ЗАЛЕЖНО ВІД ТИПУ ШАРУ

Тип шару (CONV, POOL, FC, ABS)	**Параметр 1**	**Параметр 2**	**Параметр 3**
CONV (00)	Кількість фільтрів (K)	Розмір фільтра (*F*)	Розмір перекриття (зсув) (*S*)
POOL (01)	Ігнорується	Розмір фільтра (*F*)	Розмір перекриття (зсув) (*S*)
FC (10)	Кількість нейронів у шарі		
ABS (11)	Ігнорується		

Розглянемо приклад кодування детальніше.

Крок 0. Ініціалізація: Згенеруємо випадковим чином набір хромосом у вигляді бітових ланцюжків:

- Індивід 1 — 00 11 10 10 00 10 11 01 01 00 10 10 10 101101;
- Індивід 2 — 00 11 10 10 00 10 01 11 01 00 10 11 11 101101;
- Індивід 3 — 00 11 10 10 00 10 01 11 01 00 10 11 11 101101;
- Індивід 4 — 00 11 10 10 00 10 01 11 01 00 10 11 11 101101;
- Індивід 5 — 00 11 10 10 00 10 01 11 01 00 10 11 11 101101.

Кожна згенерована хромосома описує певну структуру згорткової нейромережі. Спробуємо «декодувати» хромосому Індивід 1 та побудувати згорткову мережу, що їй відповідає.

1) Розіб'ємо обрану хромосому на N (N = 4 для даного прикладу) рівних частин:

Індивід 1 — 00 11 10 10 | **00 10 1 01** | 01 00 10 10 | **10 101101.**

2) Кожна з отриманих частин кодує один шар згорткової мережі, а саме: тип шару та значення його параметрів (табл. II–V).

ТАБЛИЦЯ II. КОДУВАННЯ ПАРАМЕТРІВ ЗГОРТКОВОГО ШАРУ CONV

	Тип шару	**Параметр 1**	**Параметр 2**	**Параметр 3**
Шар 1	$00_2 = 0_{10}$	$11_2 = 3_{10}$	$10_2 = 2_{10}$	$10_2 = 2_{10}$
Тлумачення	CONV	Кількість фільтрів (K) = 3	Розмір фільтра (F) = 2	Розмір перекриття (S) = 2

ТАБЛИЦЯ III. КОДУВАННЯ ПАРАМЕТРІВ ЗГОРТКОВОГО ШАРУ POOL

Шар 2	$00_2 = 0_{10}$	$10_2 = 2_{10}$	$11_2 = 3_{10}$	$_{012 = 110}$
Тлумачення	CONV	Кількість фільтрів (K) = 2	Розмір фільтра (F) = 3	Розмір перекриття (S) = 1

ТАБЛИЦЯ IV. КОДУВАННЯ ПАРАМЕТРІВ ЗГОРТКОВОГО ШАРУ ABS

Шар 3	$01_2=1_{10}$	$00_2=0_{10}$	$10_2=2_{10}$	$10_2–2_{10}$
Тлумачення	POOL	Ігнорується	Розмір фільтра (F) = 2	Розмір перекриття (S) = 2

ТАБЛИЦЯ V. КОДУВАННЯ ПАРАМЕТРІВ ЗГОРТКОВОГО ШАРУ CONV

Шар 4	$10_2=2_{10}$	$101101_2=45$
Тлумачення	FC	Кількість нейронів у шарі = 45

Таким чином, даний індивід кодує згорткову мережу наступного вигляду:

- вхідний шар – кількість нейронів визначається параметрами задачі;
- згортковий шар CONV – 3 фільтри 2×2 з розміром перекриття 2;
- згортковий шар CONV – 2 фільтри 3×3 з розміром перекриття 1;
- субдискретизуючий шар POOL – розмір фільтра 2×2 з розміром перекриття 2;
- повнозв'язний шар FC – 45 нейронів у шарі.

Крок 1-8. Аналогічні прикладу розрахунку оптимальної структури мережі глибокої довіри.

VII. ВИСНОВКИ

Проаналізовано недоліки використання згорткової нейронної мережі при обробці зображень. Обгрунтовано необхідність розв'язання задачі оптимізації для визначення оптимальних параметрів налаштування нейронної мережі. Запропоновано гібридний генетичний алгоритм для її розв'язання. Наведено приклад, який пояснює отримання оптимальної структури згорткової мережі.

ЛІТЕРАТУРА REFERENCES

[1] J. Horn, N. Nafpliotis, and D. E. Goldberg, "A niched Pareto genetic algorithm for multiobjective optimization," *In Proceedings of the First IEEE Conference on Evolutionary Computation*, Piscataway, vol. 1, pp. 82-87, 1994.

[2] S. Watanabe, T. Hiroyasu, and M. Miki, "NCGA: Neighborhood Cultivation Genetic Algorithm for Multi-Objective Optimization Problems," *Proceedings of the Genetic and Evolutionary Computation Conference*, 2002, pp. 458-465.

[3] E. Zitzler, K. Deb, and L. Thiele, "Comparison of multiobjective evolutionary algorithms: Empirical results," *Evolutionary Computation*, vol. 8, pp. 173-195, 2000.

[4] E. Zitzler and L. Thiele, "Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach," *IEEE Transactions on Evolutionary Computation*, vol. 3, no. 4, pp. 257-271, 1999.

Deep Learning Classifier Based on NEFCLASS and NEFPROX Neural Networks

Olena Chumachenko
Technical Cybernetic Department, National Technical University of Ukraine
"Ihor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine
E-mail: chumachenko@tk.kpi.ua

Класифікатор Глибокого Навчання на Основі Нейронних Мереж NEFCLASS та NEFPROX

Олена Чумаченко
Кафедра технічної кібернетики, Національний технічний університет України
"Київський політехнічний інститут ім. Ігоря Сікорського ", Київ, Україна
E-mail: chumachenko@tk.kpi.ua

***Abstract*—It is proposed a new class of fuzzy classifiers. It is a deep learning classifier based on NEFCLASS and NEFPROX neural networks. The pre-learning is supplied with help of Restricted Boltzman Machine. The adjustable parameters of NEFCLASS and NEFPROX neural networks are parameters of membership functions, the type of which coincides with membership functions type of Restricted Boltzman Machine.**

***Анотація*—Запропоновано новий клас нечітких класифікаторів. Це новий клас класифікаторів глибокого навчання на основі нейронних мереж NEFCLASS та NEFPROX. Попереднє навчання здійснюється за допомогою обмеженої машини Больцмана. Налаштовуваними параметрами нейронних мереж NEFCLASS та NEFPROX є параметри функцій членства, тип яких співпадає з типом функцій членства обмеженої машини Больцмана.**

Keywords—Fuzzy classifiers; deep learning; NEFCLASS and NEFPROX neural network Restricted Boltzman Machine.

Ключові слова—Нечіткі класифікатори; глибоке навчання; нейронні мережі NEFCLASS та NEFPROX; обмежена машина Больцмана.

I. INTRODUCTION

The neural networks are currently used for the solution of various kinds of problems, such as approximation, classification, pattern recognition, prediction. Along with the classical approach, there are a number of methods for these problems solution, to simplify and speed up the computations. One of such method is the transition to the fuzzy logic. Fuzzy logic systems transform the numerical inputs of neural networks into data of fuzzy nature using the fuzzification process. The main element of fuzzification is a membership function μ. Its value determines whether the numerical value of the input "low", "normal" or "great." The most common types of membership functions are triangular, trapezoidal, bell-shaped and Gaussian function. Further there is the aggregation of obtained results using the fuzzy rules, logical conclusion and defuzzification – transform of fuzzy values into numeric format [1], [5].

NEFCLASS and NEFPROX networks belong to a class of three-layer fuzzy perceptrons [9].

1) The first layer is the input data and does not change their values.

2) The neurons of the hidden layer (Layer 2) contain fuzzy rules, for example, fuzzy rule R_j takes the form: if x_1 belongs $\mu_k^{(1)}$, x_2 belongs $\mu_l^{(2)}$, … , x_n belongs $\mu_p^{(n)}$, the output R_j will be equal to 1, where $\mu^{(1)} \ldots \mu^{(n)}$ are membership functions.

In general, the membership function for the network NEFCLASS and NEFPROX is bell-shaped function of the following form.

$$\mu(x) = \frac{1}{1+\left|\frac{x-c}{a}\right|^{2b}}, \text{ for NEFCLASS,} \qquad (1)$$

$$\mu(x) = \begin{cases} \dfrac{x-a}{b-a}, & a < x < b, \\ \dfrac{c-x}{c-b}, & b < x < c, \\ 0, \end{cases} \quad \text{for NEFPROX,} \qquad (2)$$

where c, a, b are adjustable parameters during training process.

3) The third layer consists of output neurons, each of which corresponds to one of the classes. The output value is calculated by the equation

$$Y = \sum R_j k, \quad \text{for NEFCLASS,} \qquad (3)$$

$$Y = \sum R_j, \quad \text{for NEFPROX,} \qquad (4)$$

where R_j is the output of the second layer equal to 0 or 1; k is the weight coefficient equal to 1.

II. Problem Statement

The standard approach to learning is based on the method of back propagation [1], however, a large number of adjustable parameters lead to a halt in the algorithm stop in local extremums, that influences badly which adversely affects the accuracy of the network [1].

To improve the accuracy of the network uses the concept pre-learning [2] – [4], [6] used in training deep neural networks, for which the basic learning algorithm is a method of backpropagation. Pre-learning is done by constructing autoassociator – neural network, the output of which is to be closest to the value of the input data. Autoassociator trained teachers without the inputs identified in the training set, then the weighting coefficients autoassociator transferred to the main network and continue learning by back propagation. Thus, there is an initial setup that allows you to be as close as possible to the global extremes, which increases the accuracy of the network. The most common autoassociator are Restricted Boltzman Machine (RBM) [7] and autoencoders [3]. In this paper we used the RBM is limited.

III. Problem Solution

Since pre-learning assumes autoassociator structure (the number of neurons and interneuronal connections, as well as activation of neuronal function) coincides with the structure of the network that pre-learning. For NEFCLASS and NEFPROX pre-learning network uses an approach similar [9]. It includes calculating the membership functions in a single layer and neurons of the initializing layer coefficients using Restricted Boltzman Machine (RBM). Process pre-learning NEFCLASS and NEFPROX network is shown in Figs 1 and 2.

IV. Network Training Algorithm

Step 1. Structural adjustment network

1) For all possible combinations of values $\mu_j(x_i)$ it is created the rules R_k. During training for sigmoidal functions of the form (1) it is necessary to adjust two parameters: slope w and the bias c.

2) For all $\mu_j(x_i)$ it is set the initial values. The initial values of membership functions must correspond to the rule

$$\forall x_k \in O(x_i) : \exists \mu_j(x_i) \neq \pm 1,$$

where $O(x_i)$ is the domain of input variable definition. That is, for each value of the input variable domain of definition, there is a corresponding value of the membership functions other than ± 1. An example of the initial distribution of membership functions for the domain [-1,1] is shown in Fig. 3.

Step 2. Parametric adjustment of the RBM is made by the algorithm independent differences

1) The initial values v_0 set equal to the input data.

2) Calculate the probability that the state of the neuron of the hidden layer is equal to 1:

$$p(h_j = 1 \mid v) = \frac{1}{1 + \exp\left(\sum_{i=1}^{m} w_{ij} v_i + b_i\right)}, \qquad (4)$$

where c_i is the bias of the hidden layer.

3) Assign the value of the hidden layer neuron based on the probability (if the probability of 0.9, then with probability 90% of neuron will be 1, and with a probability of 10% – 0).

4) Calculate the probability that the state of the neuron of the visible layers will be equal to 1:

$$p(v_i = 1 \mid h) = \sigma\left(\sum_{j=1}^{m} w_{ij} h_j + c_j\right), \qquad (5)$$

where c_j is the bias of the visible layer; σ is the function

$$\sigma = \frac{1}{1 + e^{-x}}. \qquad (7)$$

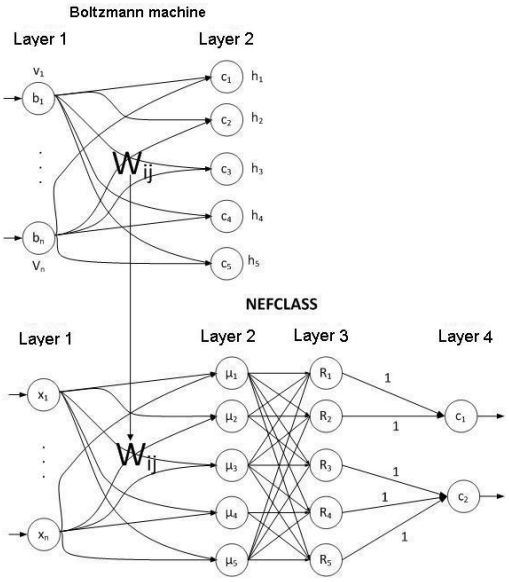


Fig. 1 Pre-learning NEFCLASS neural network.

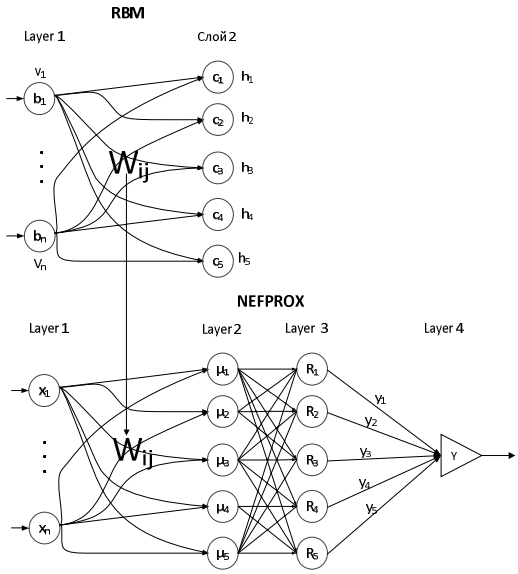


Fig. 2 Pre-learning NEFPROX neural network.

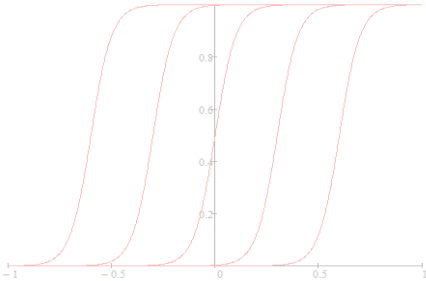


Fig. 3. The initial distribution of the membership functions.

1) Assign neuron layer visibility based on the probability value.
2) Repeat *step 2* if the iteration number is less than *k*.
3) Calculate the probability of the hidden layer.

A. Initial parameter values

1) The states of the hidden layer neurons are given by input data.
2) The initial value of neural connections weights are given as a small value of about 0 to Gause distribution.
3) The biases of the visible layer are given.

B. Parameters modification

1) Synaptic connections

$$w_{ij}(t+1) = w_{ij}(t) + \alpha\left[v_i(0)h_j(0) - v_i(k)h_j(k)\right].$$

2) Biases of the visible layer

$$b(t+1) = b_i(t) + \alpha\left[v_i(0) - v_i(k)\right].$$

3) Biases of the hidden layer

$$c(t+1) = c_j(t) + \alpha\left[h_j(0) - h_j(k)\right],$$

where α is the training speed.

Step 3. Parametric fuzzy network

The parameters obtained by learning MBP transferred to the network as follows:

1) The weighting coefficients $w_{ij}^{\text{NEFPROX}} = w_{ij}^{\text{RBM}}$
2) The bias of the second layer neurons $\beta_j^{\text{NEFPROX}} = b_j^{\text{RBM}}$.

For training example, we find a difference between the reference value and the value obtained as a result of the network functioning

$$\Delta_c = y_e - y_r.$$

3) We calculate Δ_R for the rule.

$$\Delta_R = U_R(1-U_R)\sum_{c\in U_3} W(R,c)\Delta_c,$$

where U_R is the output of the rules layer; $W(R,c)$ is the connection weight of rule neuron with class neuron.

4) We find x' which is the minimum of function $W(x',R)U_{x'}$

$$W(x',R)U_{x'} = \min\{W(x,R)U_x\}$$

where $W(x',R)U_{x'}$ is the membership functions.

5) For the membership function it is determined the values on which it is necessary to change the parameters values

$$\Delta_w = -\sigma\Delta_R \frac{-xe^{wx'+\beta}}{(e^{wx}+e^{\beta})}, \qquad \Delta_\beta = -\sigma\Delta_R \frac{e^{wx'+\beta}}{(e^{wx}+e^{\beta})^2},$$

where σ is the training speed.

6) Change the parameters, according to step 5 and calculate the rule error

$$E = U_R(1-U_R)\sum_{c\in U_3}(2W(R,c)-1)\Delta_c.$$

The complex criterion of stop is:

1) $\sum_{i=1}^{n}\Delta_{c_i} \le \Delta_c^*$, i the number of examined examples; Δ_c^* is the threshold value; n is the size of examined sample.

2) The Δ_{c_i} doesn't change during some iterations

$$\left|\Delta_{c_i} - \Delta_{c_j}\right| \le \varepsilon,$$

where i, j is the number of iterations $i \neq j$.

V. RESULTS

Sample format (Table I):
number of examples in the training sample – 150;

number of examples in the test sample – 30;

number of classification parameters – 4;

number of classes – 3.

The given results show that the quality of the problem solution is improved. It is especially important when it is necessary to solve the problem of image recognition [8].

TABLE I. Results of Problem Solution by Iris Database Samples

No	Neural Networks type	Method of teaching	Correct results	Incorrect results
1	Classifier based on Softmax function	Reverse propagation errors	22	8
2	NEFCLASS, bell-shaped membership function	Reverse propagation errors	20	10
3	NEFCLASS, sigmoidal membership function	Reverse propagation errors	19	11
4	NEFCLASS, sigmoidal membership function	Preliminary Training Limited Boltzmann machine (CD-1) + reverse propagation errors	26	4
5	NEFPROX, bell-shaped membership function	Reverse propagation errors	21	9
6	NEFPROX, sigmoidal membership function	Reverse propagation errors	19	11
7	NEFPROX, sigmoidal membership function	Preliminary Training Limited Boltzmann machine (CD-1) + reverse propagation errors	25	5
8	ANFIS, bell-shaped membership function	Reverse propagation errors	25	5
9	NEFPROX, sigmoidal membership function	Reverse propagation errors	26	5
10	NEFPROX, sigmoidal membership function	Preliminary Training Limited Boltzmann machine (CD-1) + reverse propagation errors	27	3

VI. CONCLUSIONS

The opportunity to train NEFCLASS and NEFPROX neural network according to the paradigm of pre-learning autoassociator is proved. These results demonstrate the improved accuracy on the test sample. It permits to use these classifiers in convolution neural networks for image processing quality improving.

REFERENCES

[1] D. Nauck, U. Nauck, and R. Kruse, "Generating classification rules with the neurofuzzy system NEFCLASS." *Biennial Conference of the North American, Fuzzy Information Processing Society*, 1996, pp. 466-470.

[2] Y. Bengio, "Learning deep architectures for AI." *Foundations and trends® in Machine Learning*, 2(1), pp. 1-127, 2009.

[3] Y. Bengio, *Deep learning of representations: Looking forward, Statistical Language and Speech Processing*. Springer Berlin Heidelberg, 2013, pp. 1-37.

[4] J. Schmidhuber, "Deep learning in neural networks: An overview." *Neural Networks*, 61, pp. 85-117, 2015.

[5] L. A. Zadeh, "Fuzzy algorithms." *Information and Control*. 12 (2), pp. 94–102, 1968.

[6] G. Hinton, *A Practical Guide to Training Restricted Boltzmann Machines*, 2010, pp. 3-17.

[7] A. Fischer, and C. Igel, *An Introduction to Restricted Boltzmann Machines*, 2011, pp. 1-23.

[8] M. Nielsen, (viewed on September 20, 2016). Using neural nets to recognize handwritten digits [Electronic resourse] – Electronic data.– Mode of access: http://neuralnetworksanddeeplearning.com/chap1.html

[9] O. I. Chumachenko. "Deep Learning Classifier Based on NEFCLASS Neural Network." *Electronics and Control Systems*, no.3(49), Kyiv: NAU, 2016, pp. 79-83.

Ітераційний Метод Самокалібрування Передатних Характеристик Вимірювальних Каналів Біомедичної Аналогово-Цифрової Системи

Леонід Крупельницький
кафедра обчислювальної техніки,
Вінницький національний технічний університет,
м. Вінниця, Україна,
krupost@gmail.com

Дмитро Куций
кафедра обчислювальної техніки,
Вінницький національний технічний університет
м. Вінниця, Україна,
dimakutsiy@gmail.com

Iterative Method of Self Calibrating Transfer Characteristic Measurement Channels of Biomedical Analog to Digital System

Leonid Krupelnytskyi
department of computer technology,
Vinnytsia National Technical Univercity,
Vinnytsia, Ukraine,
krupost@gmail.com

Dmytro Kutsyi
department of computer technology,
Vinnytsia National Technical Univercity,
Vinnytsia,Ukraine,
dimakutsiy@gmail.com

***Анотація*—Запропоновано ітераційний метод самокалібрування вимірювальних каналів багатоканальних аналого-цифрових систем опрацювання сигналів біомедичних сенсорів. Розглянуто структурну й алгоритмічну реалізацію процедур калібрування і корекції, що дозволяє зменшити похибки та підвищити ідентичність параметрів вимірювальних каналів.**

***Abstract*—Iterative method of self calibrating measuring channel of multichannel analog to digital system for biomedical sensors is preposed. Structure and algorithm realization procedure for calibrating and correction, that can reduce errors and increase identity of measure channels are considered.**

Ключові слова—аналого-цифрова система; вимірювальний канал, передатна характеристика; самокалібрування, коригування похибок.

Keywords—analog to digital system; measure channel; transfer characteristics,self calibrating; errors correction.

I. Вступ

Актуальність дослідження обумовлена підвищеними вимогами щодо точності і ідентичності вимірювальних каналів біомедичних аналого-цифрових систем (АЦ-систем) для узгодженого, одночасного багатоканального опрацювання сигналів акустичних і оптичних сенсорів. Вимірювальні канали цих АЦ-систем містять вхідні аналогові комутатори, програмовані підсилювачі струму і напруги, аналого-цифрові перетворювачі, схеми живлення резистивних і мостових сенсорів постійною чи імпульсною напругою та струмом [1-3]. Пропонується ітераційний метод та алгоритм самокалібрування статичних (а відтак – і динамічних), передатних характеристик послідовності каналів з кроками «зверху-вниз» для різних коефіцієнтів підсилення.

II. Функціональна схема самокаліброваної АЦ-системи

Розглянемо пропонований ітераційний метод самокалібрування на прикладі розробленої авторами самокаліброваної 4-канальної АЦ-системи для опрацювання сигналів біомедичних акустичних і оптичних сенсорів.

На рисунку 1 наведено функціональну схему системи. Вхідні вимірювальні канали складаються з аналогових комутаторів АК1-АК4, програмованих підсилювачів ПП1-ПП4, антиаліасінгових фільтрів низьких частот ФНЧ1 – ФНЧ4, аналого-цифрових перетворювачів АЦП1 – АЦП4. Вказані аналогові і аналого-цифрові компоненти мають реальні (дійсні) статичні та динамічні передатні характеристики, які можуть бути суттєво покращеними за рахунок калібрування та корекції.

Також функціональна схема містить джерела живлення активних акустичних та оптичних сенсорів постійною чи змінною напругою, формувачем якої є окремий 4-канальний ЦАП , характеристики якого також можна

калібрувати, оскільки його виходи підключені на окремі входи аналогових комутаторів. Для калібрування в функціональній схемі наявні додаткові компоненти: джерело опорної напруги, калібрувальний перемножувальний ЦАП, шина «нуля».

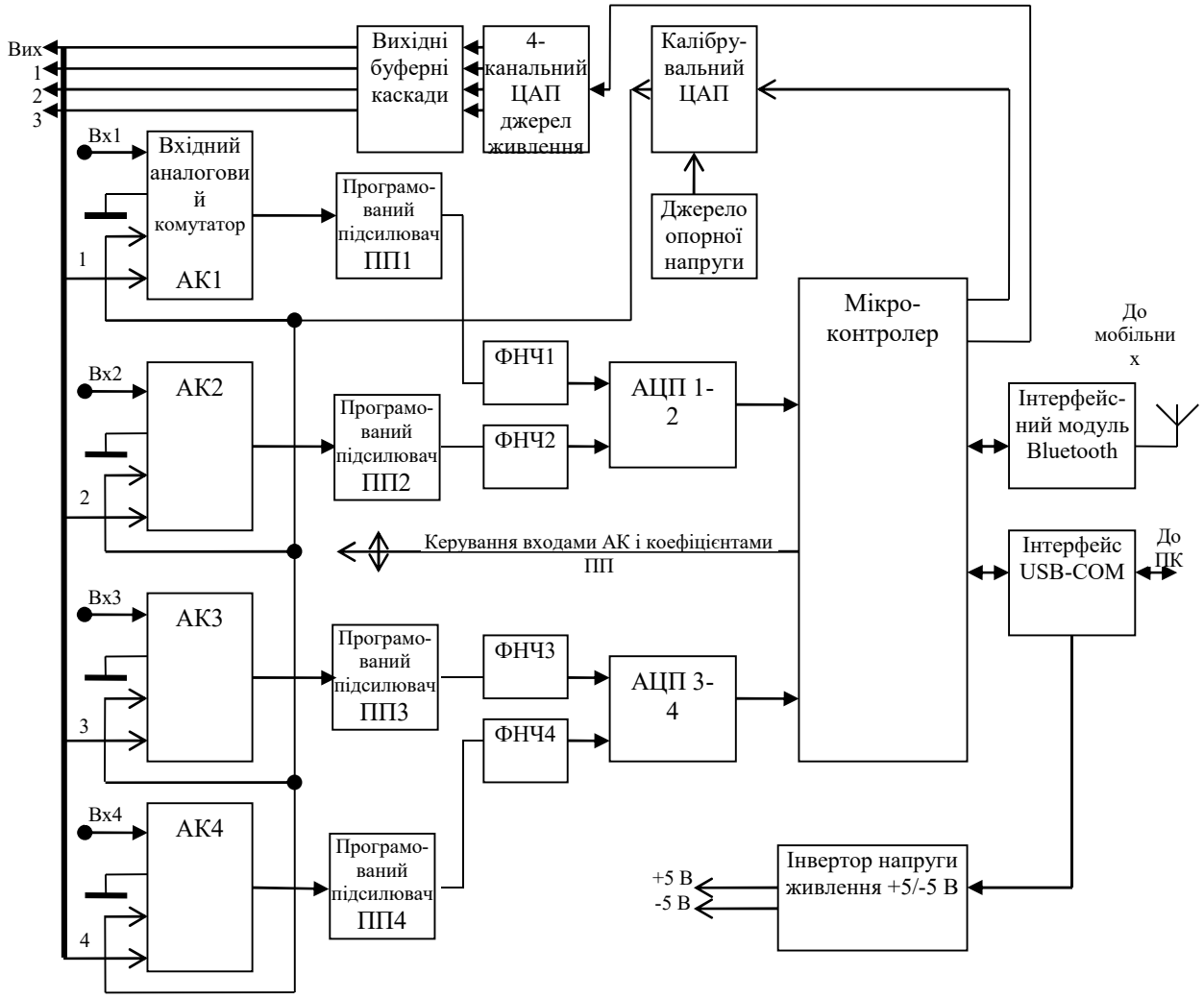


Рис. 1. Функціональна схема 4-канальної біомедичної самокаліброваної АЦ-системи.

Використання перемножувального ЦАП в якості джерела калібрувальних сигналів обумовлено тим, що особливість його схемотехніки на основі R – 2R матриці та КМОН-ключів забезпечує точне відтворення значення напруги зовнішнього опорного джерела. Отже, в верхній шкалі діапазону (код 111…1) напруга на виході ЦАП дорівнюватиме напрузі опорного джерела Uоп, а це дає можливість сформувати калібрований за амплітудою опорний сигнал. Цей сигнал може бути постійним, імпульсним чи моногармонійним з заданою опорною частотою fo. Саме моногармонійний сигнал найкраще сумісний зі спектральними методами опрацювання, в яких просто визначається його амплітуда на заданій опорній частоті. Нелінійні спотворення калібрувального ЦАП через його обмежену розрядність (в експериментальному взірці - 12 двійкових розрядів) не впливають на амплітудне значення сигналу. Зсув «нуля» ЦАП також не впливатиме на результат калібрування, оскільки визначення амплітуди результату АЦ-перетворення сигналу здійснюватиметься тільки за його спектральним значенням на опорній частоті fo. Крім того, загальний зсув «нуля» вимірювального каналу буде визначено та скориговано після підключення шини «нуля» до вхідного аналогового комутатора.

III. Ітераційний алгоритм самокалібрування статичних характеристик

Розглянемо узагальнену блок-схему алгоритму калібрування статичних передатних характеристик вимірювальних каналів, наведену на рис. 2.

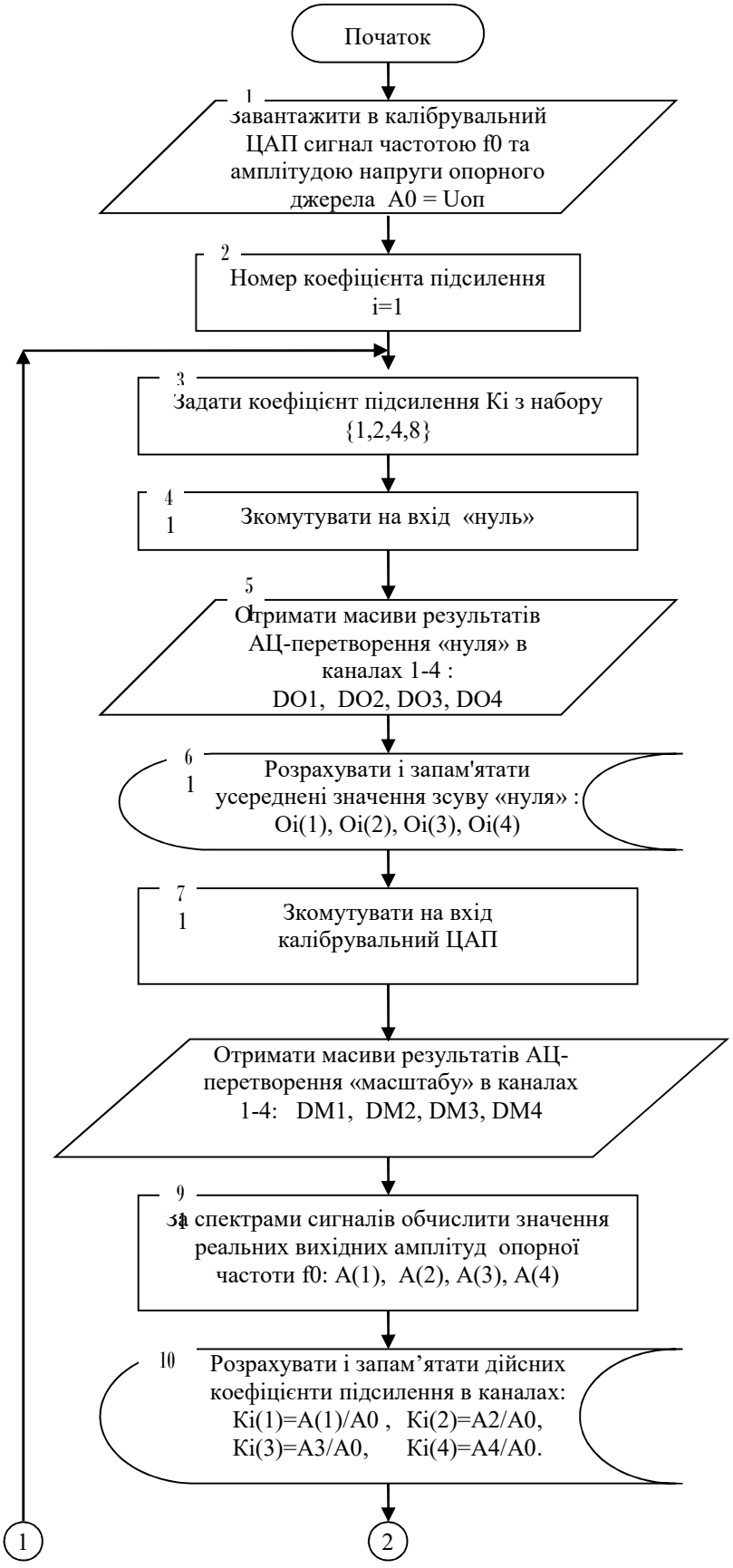
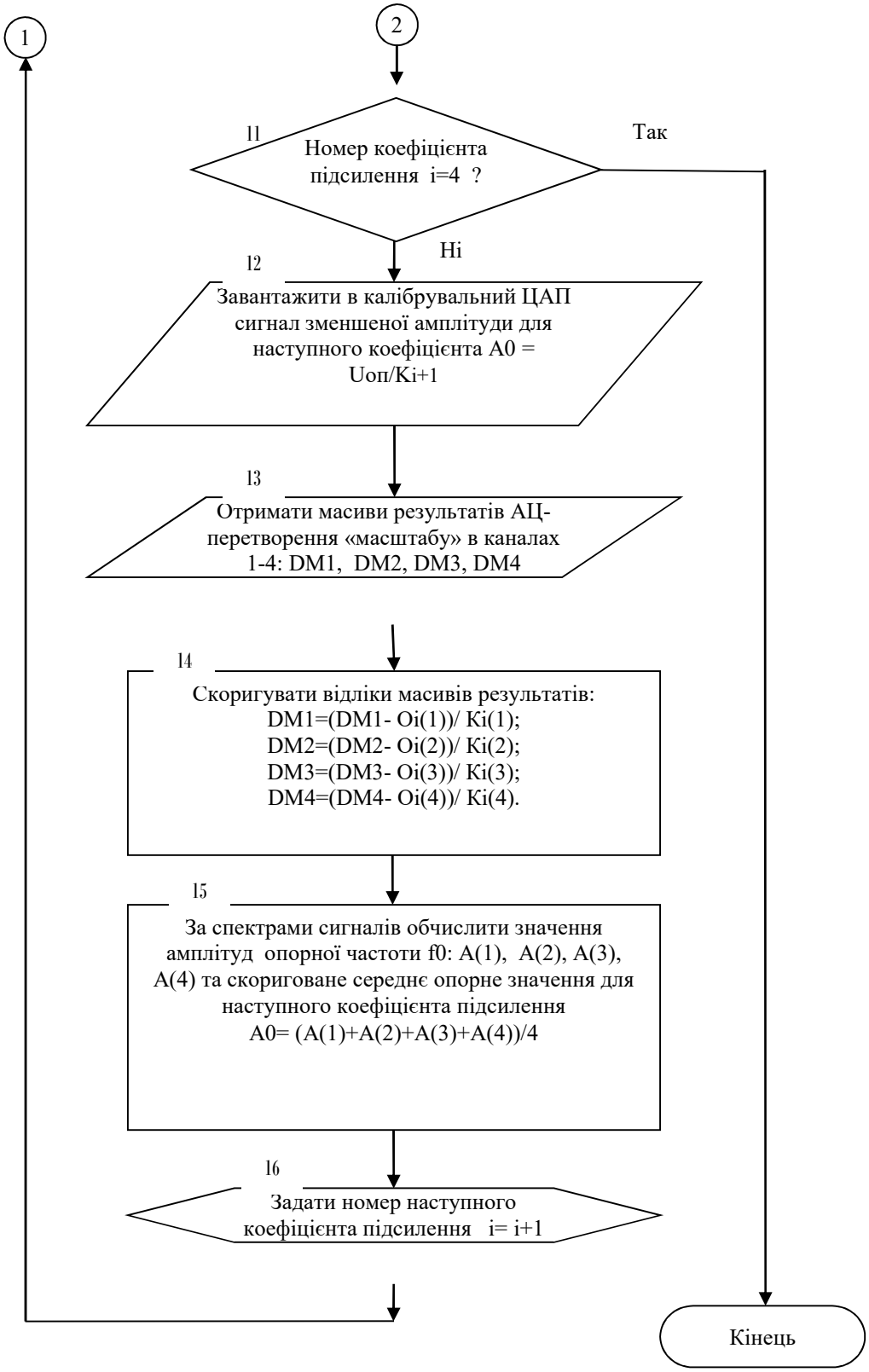


Рис. 2. Блок-схема алгоритму калібрування зсувів «нуля» і коефіцієнтів підсилення в каналах.



Продовження рис. 2.

Передбачається визначення зсувів «нуля» та дійсних коефіцієнтів підсилення в кожному з чотирьох каналів та на кожному з чотирьох коефіцієнтів підсилення Ki з набору {1,2,4,8}.

Алгоритм калібрування коефіцієнтів підсилення реалізовано за процедурою «зверху вниз». Спочатку на калібрувальний ЦАП подається гармонійний сигнал амплітуда повної шкали А0, яка дорівнює напрузі опорного джерела. Для коефіцієнта підсилення Кі=1 визначаються поточні цифрові значення результатів АЦ-перетворення «нуля» О1 та опорного сигналу А.

Дійсний коефіцієнт підсилення в режимі калібрування коригується в кожному з каналів за виразом: К1 = А/А0. В режимі безпосереднього АЦ-перетворення вхідних відліків сигналів S на першому коефіцієнті підсилення відбувається наступне коригування: S* = (S – О1) *К1.

Після закінчення калібрування на першому коефіцієнті підсилення, отримані значення О1 і К1 використовуються для визначення точного значення напруги опорного сигналу для наступного коефіцієнта К2. Для цього опорний сигнал калібрувального ЦАП зменшується в К2/К1 раз та подається на вимірювальні канали, в яких встановлено попередній коефіцієнт підсилення К1. Після застосування відкаліброваних значень зсуву «нуля» О1 і коефіцієнта підсилення К1 визначається дійсне значення опорного сигналу А0, яке і буде використовуватись під час калібрування О2 і К2 на другому коефіцієнті підсилення. Описана процедура повторюється також для коефіцієнтів К3 і К4..

На рисунку 2 наведено основні операційні блоки алгоритму запропонованого методу калібрування. На початку із зовнішнього комп'ютера здійснюється завантаження в калібрувальний ЦАП максимального за амплітудою опорного моногармонійного сигналу частотою f0 (блок 1). При цьому амплітуда циклічно відтворюваного сигналу буде дорівнювати напрузі опорного джерела живлення А0=U0. Для використаної мікросхеми ADR03 ця напруга складає (2,5 ± 0,005) В, а її значення залишається стабільним в широкому діапазоні температур експлуатації.

Далі виконується описаний вище ітераційний алгоритм «зверху вниз» калібрування «нуля» і коефіцієнтів підсилення в кожному з чотирьох вимірювальних каналів. Номер діапазону підсилення і = {1,2,3,4} визначає одне із чотирьох значень коефіцієнта підсилення Кі з набору і = {1,2,3,4} (блоки 2,3). Після комутації «нуля» на вхід аналогових комутаторів проводиться отримання проміжних масивів цифрових даних результатів АЦ-перетворення в каналах 1 - 4: D01, D02, D03, D04 (блоки 4, 5). Кожний із масивів розмірністю 256 – 1024 відліків усереднюється для знаходження значення зсувів «нуля» для кожному і-ого коефіцієнта підсилення (блок 6): Оі (1), Оі (2), Оі (3), Оі (4).

Визначення дійсних коефіцієнтів підсилення в каналах здійснюється після подачі сигналу калібрувального ЦАП на входи вимірювальних каналів (блок 7) та отримання проміжних масивів результатів АЦ-перетворення в каналах 1 - 4 (блок 8): DM1, DM2, DM3, DM4. Кожний з масивів даних, розмірністю 256 – 1024 відліків передається на зовнішній комп'ютер, де відбувається спектральний аналіз сигналів методом швидкого перетворення Фур'є. Здійснюється виділення спектральних складових А (1), А (2), А (3), А (4) на опорній частоті f0 (блок 9). Відношення цих амплітуд до амплітуди опорного сигналу визначає дійсні коефіцієнти підсилення в каналах (блок 10):

Кі (1) = А1 /А0, Кі (2) = А2 /А0,

Кі (3) = А3 /А0, Кі (4) = А4 /А0.

Наступним кроком калібрування є завантаження і точне вимірювання сигналу калібрувального ЦАП для наступного (і + 1)-го діапазону підсилення, але з використанням відкоригованих значень на і-ому діапазоні (блоки 11 -15). Ця процедура також передбачає отримання проміжних масивів даних DM1, DM2, DM3, DM4 (блок 13), їх попереднє коригування з урахуванням раніше отриманих зсуву «нуля» та дійсних значень коефіцієнтів підсилення Кі (блок 14).

Безпосереднє визначення амплітуди сигналу калібрувального ЦАП для (і +1) діапазону підсилення також здійснюється за спектрами сигналів в кожному із чотирьох каналів та з наступним отриманням скоригованого середнього значення А0 (блок 15). Описані операції калібрування і корекції виконуються послідовно для всіх чотирьох діапазонів підсилення в циклі, організованому по номеру діапазону і (блоки 1,11,16).

Описаний алгоритм калібрування реалізовано в розробленій авторами 4-канальній самокаліброваній АЦ-системі опрацювання біомедичних сигналів SADS-4. Початкові значення приведених до діапазонів вимірювань похибки та неідентичності статичних передатних характеристик вимірювальних каналів зменшується в результаті коригування від 3 до 10 раз, до рівня 0.03% в діапазоні робочих температур 0...+50 ºС..

IV. Висновки

Запропоновано ітераційний метод самокалібрування передатних характеристик багатоканальних АЦ-систем з програмованими діапазонами підсилення. Наведено структурну і алгоритмічну реалізацію методу калібрування і корекції. Застосування запропонованого методу калібрування дозволяє суттєво зменшити похибки та підвищити ідентичність вимірювальних каналів розробленої АЦ-системи для опрацювання акустичних і оптичних біомедичних сигналів.

Література References

[1] Л.В. Крупельницький, О. Д. Азаров. Аналого-цифрові пристрої систем, що самокоригуються, для вимірювань і обробляння низькочастотних сигналів : монографія / Під заг. ред. О.Д. Азарова, - УНІВЕРСУМ-Вінниця, 2005.- 167 с.

[2] О. Д. Азаров, Л. В. Крупельницький, А. В. Снігур, О. О. Решетнік, В. А. Гарнага. Коригування статичних похибок вимірювального каналу ІВС, який містить АЦП із ваговою надлишковістю // Проблеми інформатизації та управління. – 2007. – N2. – С. 5 – 9.

[3] Л.В. Крупельницький. Характеристики і структури багатоканальних АЦ-систем, що самокорегуються, для аналізу аудіо сигналів / Л.В.Крупельницький // Тези доповідей П'ятої Міжнародної науково-практичної конференції "Методи та засоби кодування, захисту й ущільнення інформації". Україна, Вінниця, 19-21 квітня 2016 р. – Вінниця: ВНТУ, 2016. – С.129-133.

Апаратна Реалізація Алгоритму Сортування на Основі Просторово-Часових Графів

В.М. Грига, І.Т. Когут, В.І. Голота
кафедра комп'ютерної інженерії та електроніки
Прикарпатський національний університет
імені Василя Стефаника
Івано-Франківськ, Україна
volodymyr.gryga@pnu.edu.ua

Л.В. Николайчук
кафедра теорії оптимальних процесів
Львівський національний університет
імені Івана Франка
Львів, Україна
Lesia_Nykolaichuk@gmail.com

Hardware Implementation of Sorting Algorithm Based on Spatio-Temporal Graph

V. Gryga, I. Kogut, V. Golota
Department of Computer Engineering and Electronics
Vasyl Stefanyk Precarpathian
National University
Ivano-Frankivsk, Ukraine
volodymyr.gryga@pnu.edu.ua

L. Nykolaichuk
Department of Optimal Process Theory
Ivan Franko National University
of Lviv
Lviv, Ukraine
Lesia_Nykolaichuk@gmail.com

***Анотація*—Обґрунтовано застосування алгоритму сортування методом "парно-непарної" перестановки, який дозволяє виконати сортування даних апаратним способом за короткий час. Запропоновано нові структури багато-тактових операційних пристроїв для заданого алгоритму сортування побудовані на основі просторово-часових графів, наведено аналітичні вирази для розрахунку затрат обладнання. Проведено порівняльний аналіз апаратної складності розроблених структур сортування даних з відомими реалізаціями.**

***Abstract*—The application of the algorithm sorting permutation method "even-odd", which allows you to sort the data hardware way in a short time. The new structure multiple operating devices for a given sorting algorithm are based on the spatio-temporal graph provides analytical expressions for calculating the cost of equipment. A comparative analysis of hardware complexity developed sorting data structures with known implementations.**

Ключові слова—алгоритм сортування методом "парно-непарної" перестановки; багатотактовий операційний пристрій; потоковий, просторово-часовий граф алгоритму

Keywords—sorting algorithm by "even-odd" permulation method; multiple operating device; data flow, spatio-temporal graph of algorithm

I. Вступ

Сортування є одною з типових проблем обробки даних і зазвичай розуміється, як задача розміщення елементів невпорядкованого набору значень масивів даних, в порядку монотонного зростання або спадання [1]. Операція сортування в середньому займає 25% машинного часу [1] і найбільш часто використовується в задачах цифрової обробки сигналів та зображень. Паралельне виконання операцій алгоритму сортування декількома операційними пристроями одночасно значно прискорює час виконання алгоритму. Серед алгоритмів сортування паралельне виконання операцій можна виконувати для алгоритмів, в яких послідовність виконуваних операцій залежить тільки від числа вхідних даних і не залежить від значень їхніх ключів (неадаптивні або інваріантні до зсуву алгоритми).

Забезпечити виконання такої операції у реальному часі є можливим на спеціалізованих засобах, архітектура яких повністю відображає структуру алгоритму сортування і орієнтована на реалізацію у вигляді ПЛІС чи НВІС.

Проте при сортуванні масивів даних, структура пристрою містить велику кількість базових операцій та зв'язків між ними, які у випадку однократного виконання іноді є не задіяними та простоюють. Тому постає задача розробки та дослідження нових видів пристроїв

сортування, які б дали можливість при не критичних часових параметрах зменшити апаратну складність пристрою та площу кристалу на ПЛІС чи НВІС яку буде займати спроектована система.

Поставлене завдання вирішується шляхом представлення заданого алгоритму у вигляді ярусно-паралельної форми потокового графу (ЯПФ ПГ) та його просторово-часового перетворення [2], з використанням різних типів просторово-часових графів (ПЧГ), що дозволить вибрати оптимальну структуру відносно заданих технічних параметрів для вирішення потрібної задачі.

II. Структура та апаратна реалізація алгоритму сортування методом “парно-непарної” перестановки

На рис. 1 зображено граф алгоритму сортування “парно-непарної” перестановки для 6-ти вхідних значень.

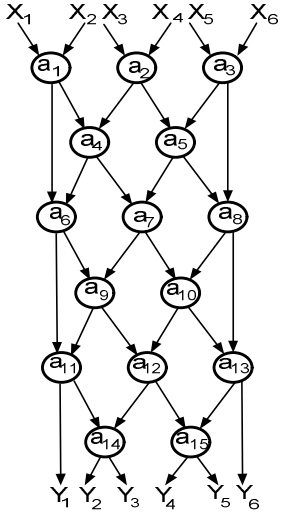


Рис. 1. Граф алгоритму сортування 6-ти значень методом “парно-непарної” перестановки.

Алгоритм сортування методом “парно-непарної” перестановки потребує для N вхідних значень - $N(N–1)/2$ операцій “порівняти і переставити” а часова складність даного алгоритму складає N операцій “порівняти і переставити” [1,3].

Для виявлення паралелізму графу алгоритму сортування методом “парно-непарної” перестановки, його потрібно подати у вигляді ярусно-паралельної форми (ЯПФ), в якій всі вершини графу розділені на яруси таким чином, що в межах одного ярусу між вершинами не має зв’язків. ЯПФ визначає степінь паралелізму графу (максимальна кількість вершин на одному ярусі або ширина графу), а також мінімально-можливий час обчислення даного алгоритму (кількість ярусів або висота графу)

Ширина ПГ алгоритму сортування методом “парно-непарної” перестановки складає - $N/2$, а висота - N операцій “порівняти і переставити”.

Апаратна реалізація описаного вище алгоритму сортування передбачає повне відображення його ПГ у структуру операційного пристрою, в якому вершинам графу (функціональним операторам) буде відповідати апаратний блок (операція) а дугам – лінії для передаі вхідних даних та проміжних і кінцевих результатів

На рис. 2 наведено структуру пристрою сортування 6-ти 8-ми розрядних чисел методом “парно-непарної” перестановки.

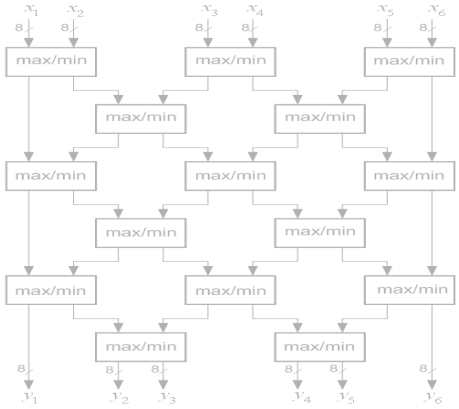


Рис. 2. Структура операційного пристрою сортування методом “парно-непарної” перестановки для 6-ти вхідних значень

Основою для вище поданої структури пристрою сортування є базова операція “порівняти і переставити” (рис. 3), яка дозволяє виконати порівняння двох чисел за допомогою схеми порівняння “на більше” та видати на один з виходів більше число (max) а на другий – менше число (min) [4].

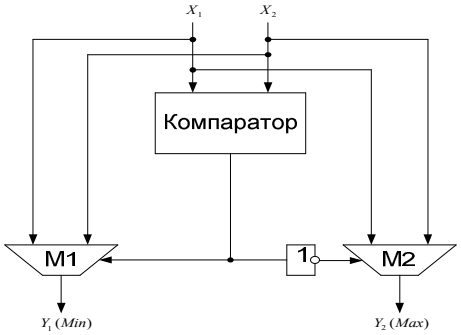


Рис. 3. Структурна схема операції “порівняти і переставити”

Структурна схема операції “порівняти і переставити” складається з компаратора та двох мультиплексорів (М1, М2). В основі компаратора є схема порівняння, яка порівнює два числа (X_1, X_2) та формує істинну ознаку, якщо ($X_1 > X_2$) в протилежному випадку формується хибне значення. Мультиплексор (М1) видає на вихід (Y_1) менше з двох чисел, а мультиплексор (М2) видає на вихід (Y_2) більше з двох чисел, в залежності від сформованого значення на їх керуючих входах, яке формується вихідним сигналом компаратора.

Структура схеми порівняння “на більше” для двох чотирьохрозрядних чисел [4,5] зображена на рис. 4.

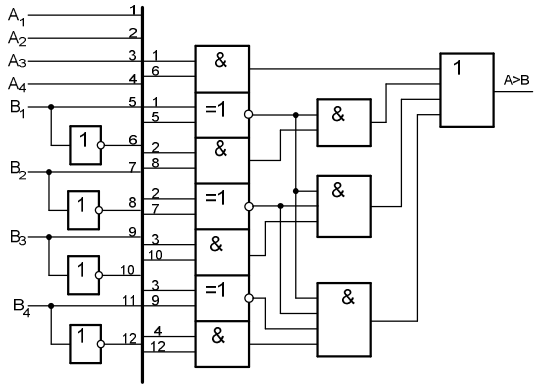


Рис. 4. Структура комбінаційної схеми порівняння двох чотирьох-розрядних чисел “на більше”

Дана комбінаційна схема складається з логічних елементів “НЕ”, “І”, “АБО” та “Виключаюче АБО”. Логічний елемент “Виключаюче АБО” містить 4 вентилі та має часову затримку 3 мікротакти [6].

Сумарна кількість вентилів даної схеми буде становити – 24 вентилі, а сумарна часова затримка – 6 мікротактів.

III. Просторово-часове перетворення алгоритму сортування методом “парно-непарної” перестановки

Для реалізації різних типів багатотактових операційних пристроїв (БОП) сортування методом “парно-непарної” перестановки потрібно здійснити перетворення ПГА у відповідні їм просторово-часові графи, які відповідають різним типам таких БОП [7].

На рис. 5 зображено структуру БОП комбінованого типу алгоритму сортування методом “парно-непарної”перестановки.

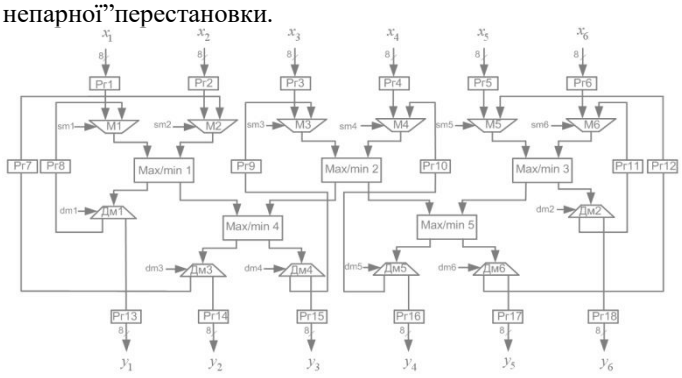


Рис. 5. Структура БОП комбінованого типу алгоритму сортування методом “парно-непарної” перестановки.

Дана структура БОП складається з 6-ти вхідних регістрів (Рг1,…, Рг6) на які подаються вхідні дані (x_1,…, x_6), 5-ти операційних блоків, які послідовно в часі виконують однотипні базові операції алгоритму, 6-ти проміжних регістрів (Рг7,…,Рг12), які затримують проміжні результати на відповідну кількість тактів, 6-ти мультиплексорів (М1,…,М6), які за допомогою значень керуючих сигналів (sm1,…,sm6) згенерованих пристроєм керування послідовно пропускають вхідні дані та проміжні результати на входи операційних блоків, 6-ти демультиплексорів (Дм1,…,Дм6), які за допомогою значень керуючих сигналів (dm1,…,dm6) розподіляють у потрібні моменти часу проміжні результати на входи операційних блоків а кінцеві результати записуються у вихідні регістри (Рг13,…,Рг18).

Затрати обладнання на реалізацію БОП комбінованого типу для заданого алгоритму сортування будуть дорівнювати:

$$W_{БОП} = (N-1)W_{ОБ} + NW_{Rвх} + NW_{МП(2\to1)} + NW_{ДМП(1\to2)} + NW_{ЕЗПР} + NW_{Rвих}. \quad (1)$$

Згідно формули (1) комбінований БОП складається з $(N-1)$ обчислювальних блоків, які виконують операцію “порівняти та переставити”, N вхідних, проміжних і вихідних n-розрядних регістрів та N мультиплексорів на 2 входи і демультиплексорів на 2 виходи.

На рис. 6 зображено структуру БОП ітераційного типу алгоритму сортування методом “парно-непарної”перестановки.

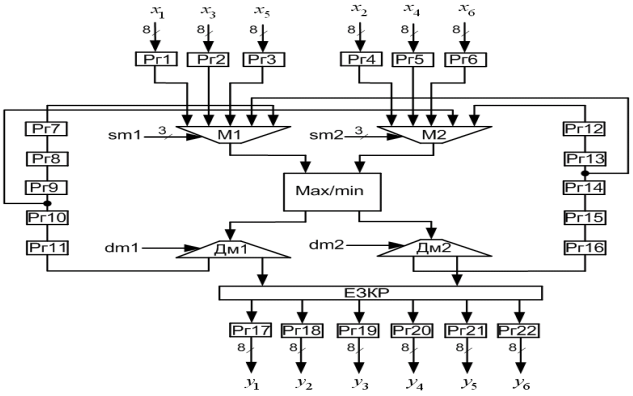


Рис. 6. Структура БОП ітераційного типу алгоритму сортування методом “парно-непарної” перестановки.

Дана структура БОП складається з 6-ти вхідних регістрів (Рг1,…, Рг6) на які подаються вхідні дані (x_1,…, x_6), одного операційного блоку, який послідовно в часі виконує всі однотипні операції алгоритму, 10-ти проміжних регістрів (Рг7,…,Рг16), які затримують проміжні результати на відповідну кількість тактів, 2-ох мультиплексорів (М1, М2), які за допомогою значень керуючих сигналів (sm1, sm2) згенерованих пристроєм керування послідовно пропускають вхідні дані та проміжні результати на входи операційного блоку, двох демультиплексорів (Дм1, Дм2), які за допомогою значень керуючих сигналів (dm1, dm2) розподіляють у потрібні моменти часу проміжні результати на входи операційного блоку а кінцеві результати на елементи затримки кінцевих результатів (ЕЗКР) після яких результати записуються у вихідні регістри (Рг17,…,Рг22).

Затрати обладнання на реалізацію БОП ітераційного типу для заданого алгоритму сортування будуть дорівнювати:

$$W_{БОП} = W_{ОБ} + NW_{Rвх} + 2W_{МП((N/2+2)\to 1))} + 2W_{ДМП(1\to 2)} + \\ +(2N-2)W_{ЕЗПР} + ((N/2-1)+(N-2))W_{ЕЗКР} + NW_{Rвих}. \quad (2)$$

Згідно формули (2) БОП ітераційного типу складається з одного обчислювального блоку, який виконує операцію ”порівняти і переставити” (max/min), N вхідних та вихідних n-розрядних регістрів, $(2N-2)$ елементів затримки проміжних результатів (ЕЗПР), двох $(N/2+2)$-входових мультиплексорів, двох 2-виходових демультиплексорів та $(N/2-1)+(N\text{-}2)$ елементів затримки кінцевих результатів (ЕЗКР).

Підставивши у формули розрахунку затрат обладнання (1) і (2) для різних типів багатотактових операційних пристроїв зачення кількості вентилів, з яких складаються схема порівняння, мультиплексори, демультиплексори та регістри, отримано числові значення, на основі яких можна побудувати графіки залежності кількості логічних елементів (вентилів) від кількості вхідних даних для алгоритму сортування двійкових чисел методом “парно-непарної перестановки”.

На рис. 7 подано графік зележності загальної кількості вентилів від величини вхідних даних для різних типів БОП розраховані на основі отриманих аналітичних виразів.

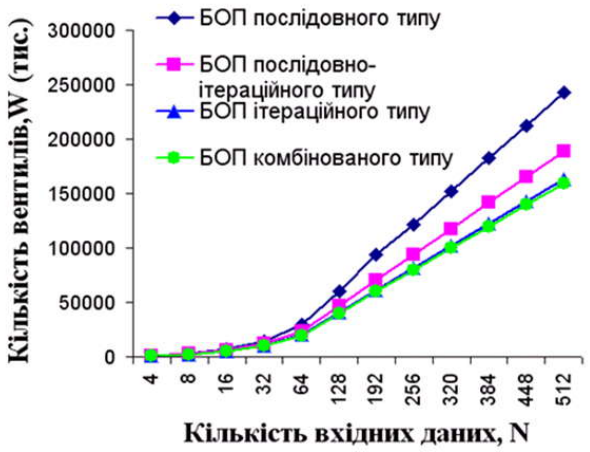


Рис. 7. Графіки залежності загальної кількості вентилів від величини вхідних даних для різних типів БОП

На рис. 8 подано графік зележності загальної кількості вентилів від величини вхідних даних для відомих одно-тактового та конвеєрного операційних пристроїв (ОП) розраховані на основі відомих аналітичних виразів.

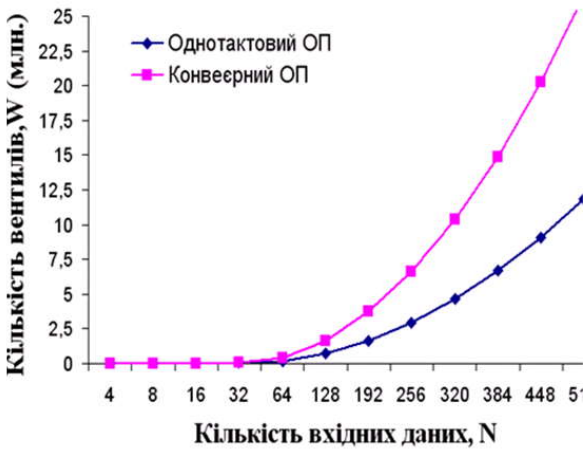


Рис. 8. Графіки залежності загальної кількості вентилів від величини вхідних даних для однотактового і конвеєрного ОП

Як видно з графічних залежностей кількість логічних елементів, які необхідно для реалізації БОП алгоритму сортування методом “парно-непарної” перестановки при ($N \geq 128$) зменшується в десятки разів у порівнянні з відомими реалізаціями однотактового та конвеєрного ОП.

IV. Висновки

При дослідженні різних способів апаратного представлення алгоритму сортування методом “парно-непарної” перестановки на основі просторово-часових графів, побудовано різні типи базових структур БОП та наведено формули розрахунку затрат обладнання на їхню реалізацію.

Отримані результати дослідження апаратної складності показали, що комбінований БОП має найменше логічних вентилів серед різних типів БОП. При порівнянні з відомими реалізаціями однотактового та конвеєрного ОП для сортування великої кількості вхідних даних БОП сортування містять в десятки разів менше обладнання.

Література References

[1] Д. Кнут, Искусство программирования для ЭВМ. Т.3: Пер. с англ. – М.: Мир, 1978. -841 с.

[2] Д.А. Поспелов, Введение в вычислительные системы. Москва: “Советское радио”, 1972. – 323 с.

[3] С. Кун, Матричные процессоры на СБИС: Пер. с англ. – М.: Мир, 1991. – 672 с.

[4] А.О. Мельник, Пам’ять із впорядкованим доступом: монографія – Львів: Видавництво Львівської політехніки, 2014. – 296 с.

[5] Б.Я. Цилькер, С.А. Орлов, Организация ЭВМ и систем: Учебник для вузов. – спб:. Питер, 2006. – 668 с.

[6] В.Л. Шило, Популярные цифровые микросхемы: Справочник. – М.: Радио и связь, 1987. – 352 с.

[7] В. М. Грига, Особливості побудови багатотактових операційних пристроїв. // Матеріали 5-ї міжнародної конференції “Сучасні комп’ютерні системи та мережі: Розробка та використання” – Львів: НУ “Львівська політехніка”, 2011. – С. 243-244.

СЕКЦІЯ 7.

Математичне та комп'ютерне моделювання складних систем

SECTION 7.

Mathematical and computer modelling of complex systems

Застосування Алгоритму Імітації Відпалу для Розподілення Даних при Паралельному Моделюванні Неізотермічного Напружено-Деформованого Стану з Урахуванням Фільтрації

В.О. Богаєнко
Інститут кібернетики ім. В.М. Глушкова НАН України
Київ, Україна
sevab@ukr.net

Application of Simulated Annealing Partitioning for Parallel Non-Isothermal Filtration-Affected Stress-Strain State Modelling

V.O. Bohaienko
V.M. Glushkov Institute of cybernetics of NAS of Ukraine
Kyiv, Ukraine
sevab@ukr.net

***Анотація*—У роботі розглядаються схеми розподілу даних в межах паралельного алгоритму скінчено-елементного моделювання напружено-деформованого стану неукріплених стінок дренажних каналів у неізотермічних умовах з урахуванням фільтраційних процесів. Для статичного та динамічного розроділу елементів сітки були застосовані алгоритм імітації відпалу та алгоритм декомпозиції на основі зменшення півширини стрічки. Алгоритм імітації відпалу був модифікований за допомогої евристики, аналогічної застосовуваній у алгоритмах мурашиних колоній. Приведені результати тестування ефективності цих алгоритмів.**

***Abstract*—Data partitioning schemes within parallel implementation of finite element modelling of non-isothermal filtration-affected stress-strain state of drainage channel walls are considered in the paper. Simulated annealing and reduced bandwidth decomposition algorithms are used for static and dynamic partitioning of a mesh. Heuristic, similar to the one used in ant colony algorithms, is proposed for simulated annealing algorithm. Algorithms efficiency testing results are presented.**

Ключові слова—напружено-деформований стан; неізотермічність; фільтрація; метод скінчених елементів; паралельні алгоритми; розбиття сітки; алгоритм імітації відпалу

Keywords—stress-strain state; non-isothermal; filtration; finite elements method; parallel algorithms; mesh partitioning; simulated annealing

I. Introduction

High computation complexity of numerical methods for modelling coupled processes in soils makes urgent a development of parallel algorithms for distributed memory systems. When differential equations system is discretized using finite element method, partitioning of the mesh used to map its elements onto the network of processes highly influence parallel computations performance.

Several approaches to solve optimal mesh partitioning problem are known. The first is recursive bisection which focuses on splitting nodes graph (e.g., spectral graph bisection [1, 2]) or mesh geometry (e.g., recursive coordinate bisection [3]). Such methods are successfully used for initial static mesh partitioning. Diffusive algorithms [4, 5] transfer mesh nodes from one process to another using information about their performance and are widely used for dynamic repartitioning. Another efficient approach for this purpose is an usage of simulated annealing as a solver of combinatorial optimization problem [6, 7]. Greedy [8] and heuristic [9] algorithms are also developed and used. Reduced bandwidth decomposition algorithm [10, 11] can be used for both initial partitioning and repartitioning. This algorithm renumbers mesh nodes with one of bandwidth minimization algorithms and assigns equally

sized blocks of nodes with sequential numbers to each process. Some of researches [12-14] focus on using swarm intelligence algorithms for solving problems of task allocation in general and finite element mesh partitioning particularly.

Usage of simulated annealing partitioning with reduced bandwidth decomposition is studied in the paper to accelerate parallel finite element analysis of stress-strain state dynamic of non-reinforced walls of drainage channels. Ant colony-like heuristics is proposed to increase efficiency of simulated annealing.

II. Problem statement

Modelling of non-isothermal stress-strain state of a soil is considered in the paper taking filtration into account. Mathematical model of the process consists of dynamic filtration, heat transfer, and stress-strain state equations as follows [15]:

$$\tilde{\mu}\frac{\partial h}{\partial t}-\frac{\partial}{\partial x}\left(K_f(T,P)\frac{\partial h}{\partial x}\right)-\frac{\partial}{\partial y}\left(K_f(T,P)\frac{\partial h}{\partial y}\right)=0,$$

$$c_T\frac{\partial T}{\partial t}-\frac{\partial}{\partial x}\left(\lambda_T\frac{\partial T}{\partial x}-c_w v_x T\right)-\frac{\partial}{\partial y}\left(\lambda_T\frac{\partial T}{\partial y}-c_w v_y T\right)=0,$$

$$\rho_s\frac{\partial^2 u}{\partial t^2}-\mu\Delta u-(\lambda+\mu)\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right)-\frac{\partial P}{\partial x}=0, \qquad (1)$$

$$\rho_s\frac{\partial^2 v}{\partial t^2}-\mu\Delta v-(\lambda+\mu)\frac{\partial}{\partial y}\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right)-\frac{\partial P}{\partial y}=-\rho_s g,$$

$$(x,y)\in\Omega,$$

where h is a water head, T – temperature, $w=(u,v)^T$ – soil displacement vector, $P=\rho_w g(h-y)$ – pressure, $K_f(T,P)$ – filtration coefficient, $(v_x,v_y)^T$ – filtration velocity, λ, μ – Lamé parameters, $\tilde{\mu}$ – soils water capacity, c_T, c_w – volumetric heat capacities of soil and water, λ_T – thermal conductivity coefficient, ρ_s, ρ_w – densities of soil and water.

Equation system (1) is considered in the domain Ω depicted in fig.1 that represents soil massif with two drainage canals.

The following boundary and initial conditions were set:

$$h(x,y,t)=y,\ (x,y,t)\in\Gamma^1,$$

$$h(x,y,t)=y_0+\tilde{h}_1(t),\ (x,y,t)\in CD,$$

$$\tilde{h}_1(t)=\frac{1}{|CD|}\int_{CD}\int_0^t v_y(x,y,\tau)d\tau d\Gamma,$$

$$h(x,y,t)=y_0+\tilde{h}_2(t),\ (x,y,t)\in GH, \qquad (2)$$

$$v_y(x,y,t)=-K_f(T,P)\frac{\partial h}{\partial y},\ \frac{\partial h}{\partial n}=0,\ \frac{\partial T}{\partial n}=0,\ (x,y)\in\Gamma^2,$$

$$T(x,y,t)=T_1,\ (x,y)\in\Gamma^1,\ T(x,y,t)=T_2,\ (x,y)\in OK,$$

$$u=0,\ (x,y)\in OA\cup JK,\ u=v=0,\ (x,y)\in OK,$$

$$\tau_s=0,\ \sigma_n=0,\quad (x,y)\in\Gamma^1,$$

$$\tilde{h}_2(t)=\frac{1}{|GH|}\int_{GH}\int_0^t v_y(x,y,\tau)d\tau d\Gamma,$$

$$h(x,y,0)=H_0(x,y),\quad T(x,y,0)=T_0(x,y),$$
$$u(x,y,0)=U_0(x,y),$$
$$v(x,y,0)=V_0(x,y),\quad \frac{\partial u}{\partial t}(x,y,0)=\frac{\partial v}{\partial t}(x,y,0)=0, \qquad (3)$$
$$(x,y)\in\Omega.$$

where

$$\Gamma^1=AB\cup BB'\cup E'E\cup EF\cup FF'\cup I'I\cup IJ,$$

$\Gamma^2=OA\cup OK\cup JK$, σ_n, τ_s are normal and tangent components of stress vector.

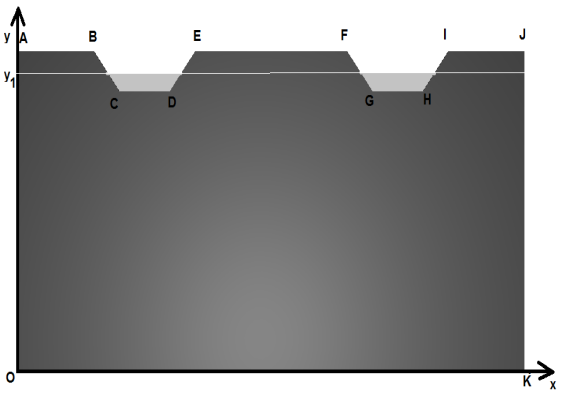


Fig.1. Solution domain

Problem (1)-(3) was numerically solved using finite element discretization on space variables and Crank-Nicolson discretization on time variable by the scheme similar to described in [16].

III. Parallel algorithms features

Solution of the discrete problem on distributed memory systems is considered on the base of the approach described in [16, 17]. Linear systems matrix is stored in compressed row storage format. Parallel version of BiCGStab [18] was used to solve the linear system with multithreaded execution within each process. Complete finite element mesh was stored in the memory of each process.

Performance of the parallel algorithm depends on uniformity of nodes distribution over the network of processors and time spent on data exchanges. The latter itself depends on the volume of data and the number of exchange operations. Partitioning of finite element mesh has to be done statically, without knowledge about computational system performance, at the start of solution process while dynamic repartitioning can be done further.

Considered problem is solved over simple domains that describe soil massifs. On the other hand, it takes into consideration several coupled physical processes that leads to linear systems with ill-conditioned matrices of high density. In such situation, different uniform data distributions can be simply reached and minimization of data exchange time along with dynamic adaptation of mesh distribution on the performance of the system plays an essential role.

We consider two variants of linear system matrix organization:

– with rows and columns grouped by functions having a matrix of 6x6 blocks;

– with rows and columns grouped by mesh nodes;

In the first case, rows of the matrix are initially distributed among processes the way to make each processes working time equal. Each row of the matrix is given a weight equal to average number of operations needed to compute one non-zero entry in this row multiplied by the number of non-zero elements. Sequential blocks of rows are assigned to processes such that a sum of row weights in the block of each process is equal. In the process of computations, as each processes performance can be measures, dynamic matrix repartitioning is done once per a given number of time steps. While doing repartitioning, weights of rows are multiplied by average time spent by owning process to make computations associated with this row.

In the second case, as the problem is solved on relatively simple meshes and for long time intervals, usage of the simplest algorithm for initial partitioning with probabilistic dynamic repartitioning can be efficient. We consider the partitioning scheme where initial partitioning is done by reduced bandwidth decomposition that uses Cuthill-McKee algorithm [19] and dynamic repartitioning done by modified simulated annealing method.

IV. Simulated annealing partitioning

Simulated annealing is a general probabilistic optimization method that simulates physical system cooling. It minimizes goal function F associated with systems state T called 'temperature' by proposing and then accepting or rejecting changes to systems state. The change is accepted unconditionally if it lowers goal function value. Otherwise, it is accepted with probability $e^{-k(F_1-F_0)/T}$, where F_1 - goal function values in changed state, F_0 - goal function value in initial state, k - given constant. Algorithm ends search when a given number of sequentially rejected changes or a given total number of changes are reached.

For mesh partitioning, system state change is done by random selection of a node which is then reassigned from current owner to random process. Several heuristics are used to make algorithm more efficient:

[1] only node on boundaries of node blocks owned by different processes can be selected for changing;

[2] node can be transferred only to the process that owns one of the neighbouring nodes;

[3] as it can be more efficient to change assignment of a cluster of connected nodes owned by the same process than of a single node, such clusters are generated with given probability of node addition;

We propose a modification of the algorithm to make use of information about nodes where changes were successful. For this purpose, approach used in ant colony algorithm [20] can be applied. Ant colony algorithm simulates ants' behaviour: an ant searching for food spreads pheromone that is used by other ants to do further search in 'efficient' areas. Pheromone is evaporating over time. For the considered algorithm such heuristic can be used the following way:

[4] pheromone value is associated with every node;

[5] nodes, neighbouring to nodes of changed cluster, have pheromone value of $m_f(F1-F0)$ added, where m_f - given multiplier;

[6] pheromone values for every node are divided by given constant d_f after each successful change;

[7] node selection is done weighted-randomly with pheromone values as weights;

We use the following form of goal function:

$$F = k_1 \max_p N_p + k_2 \max_p C_p + k_2 \max_p E_p, \qquad (4)$$

where N_p - number of nodes owned by process p, C_p - number of nodes, own by process p, that have neighbours owned by other processes, E_p - number of processes with which process p have communications, k_1, k_2, k_3 - given coefficients that depend on computational system and network performance. Based on the form of goal function, the following additional heuristics were used:

[8] weight used for selecting node for changing is multiplied by $k_1 N_p + k_2 C_p + k_2 E_p$, where p is a process that owns the node. Such additional weight multiplier makes more probable selection of nodes owned by processes that highly influence goal function value;

[9] if a change of nodes' ownership do not change a value of goal function, additional check is done: change is accepted if $\max_p E_p$ lowers, where E_p - number of edges which connect a node owned by process p and a node owned by other process.

V. Parallel algorithms testing

Several numerical experiments were carried out to test an efficiency of considered algorithms and their proposed modifications on SKIT-3 cluster of Institute of cybernetics of NAS of Ukraine on 4-cored nodes.

For testing purposes, problem (1)-(3) was being solved with the following values of parameters:

[10] $K_f(T,P) = \overline{K}_f(10)(0.7 + 0.03T)$;

[11] $\overline{\overline{K}}_f(10) = \begin{cases} \overline{\overline{K}}f, P \geq 0, \\ \overline{\overline{K}}_f e^{2.48P}, P < 0, \end{cases} \quad \overline{\overline{K}}_f = 0.3\,m/day$;

[12] $\rho_s = 1940\,kg/m^3$, $\tilde{\mu} = 10^{-3}$;

[13] Young modulus of soil - $E = 5 \cdot 10^6\ kg/m^3$;

[14] Poisson coefficient of soil - $\nu = 0.3$.

In the first series of experiments, parallel algorithms that uses per function and per node row blocking in linear systems matrices were tested for their performance. Two variants of mapping onto computational system were considered: the first with one thread per process and four processes per node, the second – with one process per node and four threads per process. Initial mesh nodes distribution over processes was computed using reduced bandwidth decomposition approach with Cuthill-McKee numbering. Above described weight coefficients were used in the case of per function row blocking. Obtained algorithms speed-up depending on total number of threads is given in fig.2.

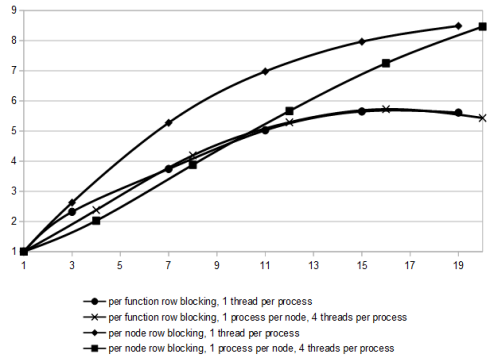


Fig. 2. Parallel algorithms speedup

The results show that per node row blocking leads to better performance because it significantly (up to 5 times) lowers a volume of data exchanges comparing with per function blocking. Use of multithreading also lowers time spent on exchanges but makes linear system matrix filling slower. The latter can be explained by less efficient memory cache utilization. In the conducted experiments, scheme without use of multithreading was faster in most cases.

In the second series of experiments simulated annealing partitioning response on parameter values was studied comparing obtained values of goal function (4) with $k_1 = k_1 = 1$, $k_3 = 0$. Number of iterations done by the algorithm was limited by 10000. As the algorithm is probabilistic, series of 20 runs were done and minimal values of goal function were averaged for further analysis. Standard deviation in all series of experiments was not more than 10% of average minimal goal function value.

Highest efficiency in the case of partitioning the mesh to map onto 15 processes and initial distribution given by reduced bandwidth decomposition algorithm was obtained for $m_f = 10$, $d_f = 1.01$, $k = 100$. Such initial distribution is close to the minimal possible, so efficiency of partitioning optimization in the terms of goal function value minimization was relatively small and was equal to 15%. When ant colony-like heuristics was not used, only 4% efficiency was obtained.

Efficiency of the algorithm in the case of obtained optimal parameter set for random and predefined initial distributions depending on number of processes to map onto is presented on fig.3.

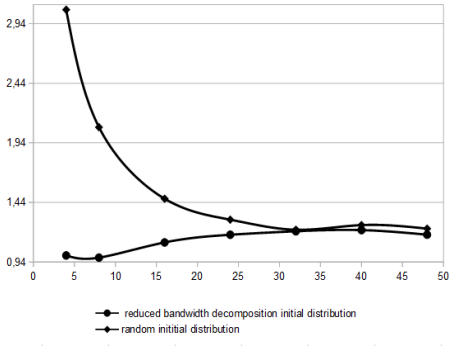


Fig. 3. Efficiency of simulated annealing partitioning

The results show that efficiency lowers when number of processes becomes higher for random initial distribution case. When initial distribution is generated by reduced bandwidth decomposition algorithm, efficiency, in contrary, grows. However, in this case, algorithm lowers initial goal function value only for number of processes more than 15.

When applying considered mesh partitioning scheme for solving problem (1)-(3) on specific computational system, initial guess of its performance parameters is crucial. Series of experiments for testing simulated annealing partitioning as static method to partition a grid were conducted and times spent on calculations and data exchanges per one iteration of BiCGstab were compared for the case of initial distribution generated by reduced bandwidth decomposition algorithm and the case when it is optimized by proposed modification of simulated annealing with $m_f = 10$, $d_f = 1.01$, $k = 250$, $k_1 = 1$, $k_2 = 0.175$, $k_3 = 10$ and maximal number of iteration equal to 30000. In situations when up to 56 cores were used, only in several cases up to 6% speed-up was obtained.

Regarding low efficiency of considered algorithms as static partitioners, the last series of numerical experiments were carried out to study computational systems performance coefficients assessment and further repartitioning influence on time spent on problem solution. Repartitioning was conducted on each 10-th iteration. Computational systems performance coefficients were calculated using averaged measures done while finding solution on each time step. For simulated annealing algorithm, maximal number of iterations was set to 1/10 of initial when doing repartitioning. Speed-up was calculated as time spent per one BiCGstab iteration in the case of non-optimized partitioning divided by the same time spent for solution on optimized mesh partition. Obtained results for speed-up depending on number of repartitions for different number of processes are depicted on fig.4 for per function row blocking and on fig.5 for per node row blocking and

application of the proposed modification of simulated annealing algorithm.

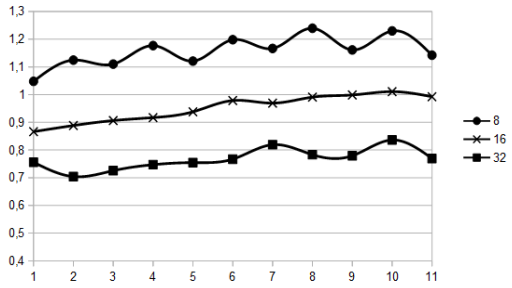


Fig. 4. Speed-up while using dynamic repartitioning with per function matrix rows blocking

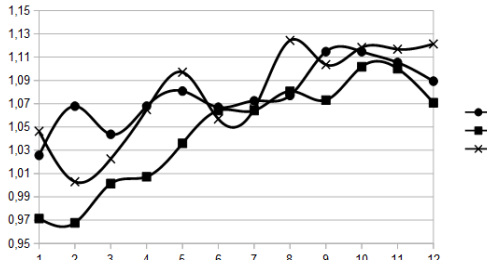


Fig. 5. Speed-up while using dynamic repartitioning with per node matrix rows blocking

Algorithm for dynamic repartitioning in the case of per function row blocking implicitly takes time spent on data exchanges into account. It increases the volume of data to be exchanged on increase of the number of processes becoming inefficient when it grows more than 16. However it uniforms time spent on computations by different processes giving ~20% speedup in the experiment with 8 processes.

In the case of per node row blocking, efficiency of dynamic repartitioning slightly depend on number of processes. Simulated annealing algorithm has given performance speed-up up to ~13% in all experiments focusing on minimization of volume and especially number of data exchange.

VI. Conclusions

Considered problem is solved in rather simple domains making application of even the simplest mesh partitioning algorithms efficient while doing parallel computations. However, optimization of such partitioning is possible both statically and dynamically. It was shown that per node row blocking in linear systems matrix with reduced bandwidth decomposition partitioning is more efficient for the considered problem comparing with the similar decomposition and per function row blocking. Simulated annealing partitioning modified by ant colony-like heuristic was used to further optimize the partition after reduced bandwidth decomposition algorithm. While giving no or negligible performance speed-up when used as static partitioner, simulated annealing produce ~13% speed-up when used for dynamic repartitioning taking computations systems performance coefficients estimations into account.

Considering a significant influence of the simulated annealing algorithm parameters on its performance, further research can be done to study efficiency of heuristic algorithms, e.g. neural networks, usage for finding their optimal values.

References

[1] H. Simon, “Partitioning of un-structured problems for parallel processing” in Computing Systems in Engineering, vol. 2, no. 2/3, 1991, pp. 135-148.

[2] A. Pothen et al., “Partitioning sparse matrices with eigenvectors of graphs” in SIAM J. Matrix Anal., 11, 1990, pp. 430–452.

[3] M.J. Berger and, S.H. Bokhari, “A partitioning strategy for nonuniform problems on multiprocessors” in IEEE Trans. Computers, 36, 1987, pp. 570–580.

[4] G. Cybenko, “Dynamic load balancing for distributed memory multiprocessors” in J. Parallel Distrib. Comput., 7, 1989, pp. 279–301.

[5] H. Meyerhenke at al, "A new diffusion-based multilevel algorithm for computing graph partitions of very high quality" in IEEE International Symposium on Parallel and Distributed Processing 2008, 2008, pp.1-13.

[6] J. Chen and V.E. Taylor, “Mesh partitioning for efficient use of distributed systems” in IEEE Transactions on Parallel and Distributed Systems, 1, 2002, pp. 67–79.

[7] R.D. Williams, “Performance of dynamic load balancing algorithms for unstructured mesh calculations” in Concurrency: Pract. Exper., 3, 1991, pp. 457–481.

[8] C. Farhat, “A simple and efficient automatic FEM domain decomposer” in Computer & Structures, vol. 28, 5, 1988, pp. 579-602.

[9] B.W. Kernighan and S. Lin, “An efficient heuristic procedure for partitioning graphs” in The Bell System Technical Journal, 49, 1970, pp. 291-307.

[10] J.G. Malone, “Automatic mesh decomposition and concurrent finite element analysis for hypercube multiprocessor computers” in Comput. Methods Appl. Mech. Eng., 70, 1988, pp.27-58.

[11] R.J. Collins, “Bandwidth reduction by automatic renumbering” in Int. j. numer. method eng., 6, 1973, pp. 345-356.

[12] A.I. Khan and B.H.V. Topping, “Subdomain generation for parallel finite element analysis” in Computing Systems in Engineering, vol. 4, issues 4–6, 1993, pp. 473-488.

[13] N. Chmait and K. Challita, “Using Simulated Annealing and Ant-Colony Optimization Algorithms to Solve the Scheduling Problem” in Computer Science and Information Technology, 1(3), 2013, pp.208-224.

[14] A.J. Page et al., “Multi-heuristic dynamic task allocation using genetic algorithms in a heterogeneous distributed system” in J. Parallel Distrib. Comput. 70, 2010, pp. 758-766.

[15] V.A. Bohaienko et al., “Analysis of numerical modelling of non-isothermal processes in soil massif (in Russian)” in Kompiuternaia matematica, №2, 2016, pp.2-13.

[16] V.A. Bohaienko et al., “Analysis of numerical modelling of soil massif dynamic in the case of non-steady state pressure filtration (in Russian)” in USiM, №4, 2014, pp.33-40.

[17] V.A. Bohaienko, “Parallel algorithm for modelling dynamic consolidation process in soil media of layered structure (in Russian)” in International Conference “Parallel and Distributed Computing Systems PDCS 2013” Collection of scientific papers, 2013, pp. 58-62.

[18] Y. Saad, “Iterative methods for sparse linear systems, 2 edition” - Society for Industrial and Applied Mathematics, 2003.

[19] E. Cuthill and J. McKee, “Reducing the bandwidth of sparse symmetric matrices” in Proc. 24th Nat. Conf. ACM, 1969, pp.157–172.

[20] M. Dorigo, “Optimization, Learning and Natural Algorithms”, PhD dissertation, Politecnico di Milano, Italy, 1992

Про Природу Вхідних Даних в Деяких Задачах Штучного Інтелекту

Н.Тимофієва
відділ комплексних досліджень інформаційних технологій
МННЦІТіС НАН та МОН України
Київ, Україна
TymNad@gmail.com

About Nature of Input Date Into Some Problems of Artificial Intelligence

N. K. Tymofijeva
integrated research department of information technology
ISTCITS of NAS and MES of Ukraine
Kiev, Ukraine
TymNad@gmail.com

***Анотація*—Аналізуються способи представлення вхідних даних в задачах штучного інтелекту. Показано, що вони задаються матрицями та скінченними послідовностями, а в задачах розпізнавання природних сигналів описуються розміщенням з повтореннями. Ця комбінаторна конфігурація задає нечіткість у вхідних даних.**

***Abstract*—The methods of presentation of input data are analysed in the problems of artificial intelligence. It is shown that they are given matrices and finite sequences, and in the problems of recognition of natural signals described placement of repetitions. This combinatorial configuration sets vagueness in input data.**

Ключові слова—вхідні дані; комбінаторна конфігурація; задача класифікації; розпізнавання; штучний інтелект

Keywords—input data; *combinatorial* configuration *problem of classification; recognition; artificial intelligence (key words)*

I. ВСТУП

До штучного інтелекту, як правило, відносять задачі, пов'язані з розпізнаванням образів [1], звукових (мовленнєвих) сигналів [2], розпізнаванням текстів, клінічної діагностики тощо. Як правило, в задачах цього класу вхідні дані мають нечітку структуру. В процесі їхнього розв'язання проводиться порівняння вхідної інформації та еталону. З цією метою використовують міри подібності. Знаходження оптимального розв'язку в них потребує перебору варіантів. Перебірним же задачам властива комбінаторна природа. Цю властивість можна дослідити, змоделювавши їх в рамках теорії комбінаторної оптимізації. Деякі задачі із штучного інтелекту зводять до задач класифікації чи кластеризації. Далі наведемо загальну постановку задачі комбінаторної оптимізації.

II. ЗАГАЛЬНА ПОСТАНОВКА ЗАДАЧІ КОМБІНАТОРНОЇ ОПТИМІЗАЦІЇ

Задачі цього класу, як правило, задаються однією або кількома множинами, наприклад A та B, елементи яких мають будь-яку природу [3]. Назвемо ці множини *базовими*. Наявні два типи задач. В *першому* типі кожну з цих множин подамо у вигляді графа, вершинами якого є її елементи, а кожному ребру поставлено у відповідність число $c_{lt} \in R$, яке називають вагою ребра (R – множина дійсних чисел); $l \in \{1,...,n\}$, $t \in \{1,..., \tilde{n}\}$, n – кількість елементів множини A, $\tilde{n}$ – кількість елементів множини B. Покладемо, що $n = \tilde{n}$. Між елементами цих множин існують зв'язки, числове значення яких назвемо вагами. Величини c_{lt} назвемо *вхідними* даними та задамо їх матрицями. В *другому* типі задач між елементами заданої множини зв'язків не існує, а вагами є числа $v_j \in R$, $j \in \{1,...,n\}$, яким у відповідність поставлено деякі властивості цих елементів, числові значення яких задаються скінченними послідовностями, що також є вхідними даними. Ці величини визначають значення цільової функції.

Для обох типів задач із елементів однієї або кількох із заданих множин, наприклад $a_l \in A$, $l \in \{1,...,n\}$, утворюється комбінаторна множина W – сукупність комбінаторних конфігурацій певного типу (перестановки, вибірки різних типів, розбиття тощо). На елементах w комбінаторної множини W вводиться цільова функція $F(w)$. Необхідно знайти елемент w^* множини W, для якого $F(w)$ набуває екстремального значення при

виконанні заданих обмежень, тобто функціонал $F(w^*) = glob \underset{w \in W^0 \subset W}{extr} F(w)$, де $extr = \{\min, \max\}$, W^0 – підмножина, яка визначається обмеженнями задачі.

Для моделювання прикладних задач в рамках теорії комбінаторної оптимізації необхідно:

- визначити вид задачі (статична або динамічна) за способом обчислення цільової функції;
- визначити базові множини, якими задається певна задача;
- визначити її тип за вхідними даними;
- визначити аргумент цільової функції (комбінаторну конфігурацію);
- змоделювати цільову функцію.

Аналіз задач із штучного інтелекту показує, що вони відносяться як до першого так і до другого типів. Аргументом цільової функції в них виступають комбінаторні конфігурації різних типів, зокрема вибірки (сполучення та розміщення як з повтореннями так і без повторень, розбиття n-елементної множини на підмножини тощо). Задачі цього класу можуть бути як статичними так і динамічними.

III. Комбінаторні конфігурації

Комбінаторною конфігурацією назвемо будь-яку сукупність елементів, яка утворюється з усіх або з деяких елементів заданої множини $A = \{a_1, ..., a_n\}$ [3]. Позначимо її впорядкованою множиною $w^k = (w_1^k, ..., w_{\eta^k}^k)$. Верхній індекс k ($k \in \{1, ..., q\}$) в w^k – порядковий номер w^k в W, q – їхня кількість. Множину $A = \{a_1, ..., a_n\}$ назвемо базовою. Під символом $w_l^k \in A$ розуміємо як окремі елементи, так і підмножини (блоки), $\eta^k \in \{1, ..., n\}$ – кількість елементів у $w^k \in W$. Залежно від умови задачі η позначатимемо без індексу або з верхнім індексом η^k. Дві нетотожні комбінаторні конфігурації w^k та w^i назвемо ізоморфними, якщо $\eta^k = \eta^i$.

В прикладних задачах обчислювального інтелекту аргументом цільової функції виступають різні типи вибірок. З поняттям вибірки пов'язують як саму операцію виділення підмножин заданої множини, так і її результат: вибрану підмножину. В подальшому маємо на увазі друге поняття.

Нехай задано базову множину $A = \{a_1, ..., a_n\}$. З неї одержимо η-вибірку. Число η називають об'ємом вибірки. В η-вибірках в залежності від умови задачі або ураховується порядок розташування в них елементів (тоді їх називають η-перестановками або η-розміщеннями) або не враховують. У цьому випадку вони називаються η-сполученнями.

Отже існують такі типи вибірок: упорядковані та неупорядковані. Неупорядковані це – сполучення без повторень і сполучення з повтореннями. Упорядковані це – розміщення з повтореннями та розміщення без повторень. Множина будь-якого типу вибірок складається з підмножин ізоморфних вибірок $W_{\eta^k} \subset W$. В задачі розпізнавання та синтезу мовленнєвих сигналів та задачі клінічної діагностики аргументом цільової функції є сполучення без повторень та розміщення з повтореннями.

IV. Моделювання вхідних даних функціями натурального аргументу

Розглянемо задачі першого типу, вхідні дані в яких задаються матрицями, одна з яких – комбінаторна. Подамо елементи h наддіагоналей симетричної комбінаторної матриці $Q(w^k)$ комбінаторною функцією $\beta(f(j), w^k)|_1^m = (\beta_1(f(1), w^k), ..., \beta_m(f(m), w^k))$, а елементи h наддіагоналей симетричної матриці C – функцією натурального аргументу $\varphi(j)|_1^m = (\varphi(1), ..., \varphi(m))$, де $m = \frac{n(n-1)}{2}$ – кількість елементів h наддіагоналей матриць C та $Q(w^k)$, $h = \overline{1, n-1}$. Якщо матриці $Q(w^k)$ та C – несиметричні, то $\beta(f(j), w^k)|_1^m$ та $\varphi(j)|_1^m$ містять усі їхні елементи, а $m=n^2$ (або $m = n\,\tilde{n}$). Функція цілі $F(w^k)$ набуде вигляду

$$F(w^k) = \sum_{j=1}^{m} \beta_j(f(j), w^k)\,\varphi(j).$$

Змоделюємо вхідні дані для задач другого типу на прикладі розбиття заданої множини на підмножини. Маємо множину елементів $A = \{a_1, ..., a_n\}$, кожному з яких присвоєно додатне число – його вагу V_j. Необхідно розділити цю множину на η частин за виконання певних умов. Для різних класів задач ці умови відрізняються.

Отже, маємо одну множину $A = \{a_1, ..., a_n\}$, між елементами $a_j \in A$ якої відсутні зв'язки. Натомість вхідні дані задано скінченною послідовністю (функцією натурального аргументу) $\varphi(j)|_1^n$, значення якої $\varphi(j)$ визначають вагу j-го елемента.

Для визначення розподілення елементів множини A по підмножинах w_s^k, $s \in \{1, ..., \eta^k\}$ для k-го варіанту розв'язку задачі уведемо комбінаторну функцію $\beta(f(j), w^k)|_1^n = (\beta_1(f(1), w^k), ..., \beta_n(f(n), w^k))$, де $\beta_j(f(j), w^k) \in \{0,1\}$ і $\beta_j(f(j), w^k) = 1$, якщо a_j-й елемент входить у підмножину $w_s^k \subset w^k$, а $\beta_j(f(j), w^k) = 0$ в іншому разі. Для підмножини w_s^k

значення цільової функції обчислюється за виразом (1), а для w^k за таким виразом:

$$F(w^k) = \max_{\substack{w_s^k \subset w^k \\ s=1,\eta^k}} \left(\sum_{j=1}^{n} \beta_j (f(j), w_1^k) \varphi(j), \ldots . \right.$$

$$\left. \ldots, \sum_{j=1}^{n} \beta_j (f(j), w_\eta^k) \varphi(j) \right) \qquad (2)$$

В цьому разі задача полягає в знаходженні такого w^{k*}, для якого цільова функція (2) на підмножині $W_{\eta^k} \subset W$ набуває найменшого значення, тобто $F(w^{k*}) = \min_{w^k \in W_{\eta^k} \subset W} F(w^k)$. Таким чином в задачах, вхідні дані в яких відносяться до другого типу, на цільову функцію накладається певна умова.

V. Вхідні дані в задачах розпізнавання різної природи

Розглянемо вхідні дані в задачах розпізнавання мовленнєвих сигналів, текстів, клінічної діагностики. Деякі з цих задач за певними ознаками розділяються на підзадачі, які потребують для свого розв'язання розроблення гібридних алгоритмів. Серед них виділимо кластеризацію та класифікацію.

Задача розпізнавання мовленнєвих сигналів полягає у знаходженні для вхідного сигналу найбільш правдоподібного еталонного з усіх можливих еталонних сигналів [2]. Для розв'язання цієї задачі необхідно провести пошук певного еталону в бібліотеці та порівняти його із вхідним сигналом. Аргументом цільової функції в ній вважають вхідний сигнал [2]. Моделювання задачі розпізнавання мовленнєвих сигналів з використанням комбінаторної оптимізації показує, що аргументом цільової функції в ній є комбінаторні конфігурації різних типів, а вхідними даними є розміщення з повтореннями.

Побудуємо математичну модель задачі клінічної діагностики як задачу комбінаторної оптимізації. Позначимо $A = \{a_1, \ldots, a_n\}$ множину захворювань, описання яких знаходиться в бібліотеці (множина еталонів), де елемент $a_s \in A$, $s \in \{1, \ldots, n\}$, відповідає певному захворюванню, якому поставлено у відповідність характерні ознаки $V^{(t)} = (v_1^{(t)}, v_2^{(t)}, \ldots, v_{q_t'}^{(t)})$, q_t' – кількість ознак t-го захворювання. Вхідною інформацією в задачі клінічної діагностики є множина ознак $\widetilde{V} = (\widetilde{v}_1, \widetilde{v}_2, \ldots, \widetilde{v}_{\widetilde{q}})$, що описує одне або кілька захворювань. Позначимо їх $B = \{b_1, \ldots, b_{n^*}\}$, де $b_\sigma \in B$ – захворювання, яке потрібно визначити, n^* – кількість можливих захворювань, а $q_t' \neq \widetilde{q}$ або $q_t' = \widetilde{q}$. Ознаки $\widetilde{v}_r \in \widetilde{V}$ вхідної інформації мають той же зміст, що і описані в еталоні ознаки $v_l^{(t)} \in V^{(t)}$, $r \in \{1, \ldots, \widetilde{q}\}$, $l \in \{1, \ldots, q_t'\}$, $\sigma \in \{1, \ldots, n^*\}$.

Задача полягає у знаходженні для $\widetilde{B}$ із множиною ознак $\widetilde{V}$ найбільш правдоподібного одного або кількох еталонів із множини $A = \{a_1, \ldots, a_n\}$, тобто за вхідними ознаками встановлюється одне або кілька захворювань $b_\sigma \in B$. Ознаки в цій задачі відіграють роль критеріїв, за якими оцінюється її розв'язок. Як і в розпізнаванні мовленнєвих сигналів, для розв'язання цієї задачі необхідно провести пошук певного еталону в бібліотеці та порівняти його із вхідними ознаками.

Отже, в оговорених задачах вхідними даними є інформація, яка поступає на вхід, та еталонна інформація, яка знаходиться в бібліотеці. Для встановлення їхньої подібності уводяться міри подібності. В процесі розв'язання задачі значення обчислених мір подібності є вхідними даними. Вони задаються матрицями, а обидві задачі зводяться до першого типу.

Оскільки розпізнавання мовленнєвих сигналів та клінічна діагностика задачі перебірні, то в реальному часі знайти оптимальний розв'язок досить складно. Тому для зведення цих задач до розв'язних проводимо за певними ознаками струкутризацію бібліотеки еталонів. На цьому етапі виникає задача кластеризації.

Кластеризація – спосіб групування однорідних об'єктів з метою виділення кластерів або "згустків" цих об'єктів, тобто необхідно виділити такі однорідні кластери, щоб об'єкти всередині них були схожі один на одного, а об'єкти різних кластерів – несхожі. Аргументом цільової функції в ній є розбиття n-елементної базової множини A на η підмножин. Ця задача відноситься до першого типу.

Для задачі кластеризації сформулюємо такі умови. Назвемо множину підмножин $\rho = (\rho_1, \ldots, \rho_\eta)$ таку, що $\rho_1 \cup \ldots \cup \rho_\eta = A$, $\rho_p \cap \rho_l = \varnothing$, $p \neq l$, $\rho_p \neq \varnothing$, $p, l \in \{1, \ldots, \eta\}$. Непуста підмножина $\rho_p = \{a_1, \ldots, a_{\xi_p}\}$, $a_s \in A$, $s \in \{1, \ldots, n\}$, може мати від 1 до n елементів. Кількість підмножин ρ_p у розбитті ρ також може бути від 1 до n ($\eta \in \{1, \ldots, n\}$). Їхню множину позначимо Θ.

Задача кластеризації полягає в розбитті заданої множини A на кластери так, щоб змодельована цільова функція набувала оптимального значення. Вхідні дані в ній є числове значення подібності між певними ознаками, якими є елементи заданої базової множини.

В задачах розпізнавання (самонавчання, розпізнавання текстів) виникає задача, яку називають таксономією. Її відносять до задач класифікації та систематизації складно організованих областей діяльності, які мають ієрархічну побудову. Математично таксономією є деревовидна структура класифікацій певного набору об'єктів. Деякі автори зводять її до задачі про купу каміння.

Як правило, при моделюванні задачі класифікації аргументом цільової функції вважають вхідні дані. Якщо змоделювати цю задачу в рамках теорії комбінаторної оптимізації, можна побачити, що вона відноситься до класу задач розбиття, аргументом цільової функції в яких є розбиття n-елементної множини A на підмножини з повтореннями. В задачі класифікації виділимо такі підзадачі:

- задано скінченну базову множину A. Класи можуть бути як задано так і не задано. Необхідно розподілити елементи базової множини по класах так, щоб останні не перетиналися. Ця задача зводиться до задачі кластеризації;
- задано скінченну базову множину A. Класи можуть бути як задано так і не задано. Елементи множини A розподіляються так, що один елемент може належати різним класам. В даному разі аргументом цільової функції є розбиття n-елементної множини A на підмножини з повтореннями;
- задано нескінченну базову множину A, частина елементів якої відома, а частина визначається в процесі розв'язання задачі, тобто інформація поступає в процесі розв'язання задачі та змінюється в часі. Аргументом цільової функції в ній є часткове розбиття нескінченної множини A на підмножини з повтореннями. В цьому разі уводиться часткова цільова функція та часткове розбиття.

Оскільки для перших двох задач розбиття утворюється з елементів скінченної множини, яке характерне для задачі кластеризації, розглянемо аргумент цільової функції для третьої задачі. Уведемо базову нескінченну множину $\widetilde{A}$, в якій елементи $\widetilde{a}_s$ для $s=\overline{1,n}$ задано, а для $s>n$ визначаються в процесі розв'язання задачі. З відомих елементів $\widetilde{a}_r \in \widetilde{A}$, $r=\overline{1,q}$, утворюємо часткове розбиття множини $\widetilde{A}$ на η підмножин (блоків) $\rho=(\rho_1,...,\rho_\eta)$, $q>n$ – кількість відомих елементів. Тоді множина підмножин $\rho=(\rho_1,...,\rho_\eta)$ має такі характеристики: $\rho_1 \cup ... \cup \rho_\eta = \widetilde{A}$, $\rho_p \cap \rho_l = \varnothing$ або $\rho_p \cap \rho_l \neq \varnothing$, $p \neq l$, $\rho_p \neq \varnothing$, $p,l \in \{1,...,\eta\}$. Непуста підмножина $\rho_p=\{a_1,...,a_{\xi_p}\}$ може мати від 1 до q' елементів ($\xi_p \in \{1,...,q'\}$), $\eta \in \{1,...,q\}$, $q'>q$, $\widetilde{a}_r=\widetilde{a}_s$ або $\widetilde{a}_r \neq \widetilde{a}_s$, $\widetilde{a}_r,\widetilde{a}_s \in \rho_p$, $r,s \in \{1,...,\xi_p\}$. Їхню множину позначимо $\widetilde{\Theta}$.

В цій задачі оцінка результату проводиться за частковими цільовими функціями, аргументом якої є часткове розбиття нескінченної множини на підмножини з повтореннями ρ^k, тобто $F(\rho^{k*})=\underset{\rho^k \in \widetilde{\Theta}}{extr} F(\rho^k)$. Вхідними даними в ній є числове значення подібності між певними ознаками, якими є елементи заданої базової множини. Тобто, описана задача класифікації відноситься до першого типу.

VI. Вхідні дані в задачах розпізнавання природних сигналів

В задачах розпізнавання мовленнєвих сигналів, електрокардіограм, електроенцефалограм тощо вхідними даними є розміщення з повтореннями. Розглянемо це детальніше.

Мовленнєвий сигнал під дією певних чинників утворюється різноманітними комбінаціями активних та пасивних органів творення мови. Позначимо A – базову множину, елементам $a_1 \in A$ якої відповідають органи мовленнєвого тракту.

Мовленнєві сигнали, що відповідають одному і тому ж слову, але вимовлені різними дикторами, відрізняються як частотою так і величиною амплітуди. Його можна описати упорядкованою послідовністю чисел, що вибрані з базової множини A. Тому отриманий сигнал є вибірка – розміщення з повтореннями з n елементів $a_j \in A$ по η, в якій ураховується порядок елементів. Отримані розміщення з повтореннями для різних сигналів одного слова містять різну кількість елементів. Ця комбінаторна конфігурація визначає нечіткість вхідної інформації.

Процес розгортання сигналу, що характерний для роботи серця або мозку (електрокардіограма, енцефалограма) також описується як розміщення з повтореннями.

VII. Висновки

Отже, задачі штучного інтелекту зводяться до задач комбінаторної оптимізації. Аргументом цільової функції в них є комбінаторні конфігурації різних типів. Вхідні дані в них задаються як матрицями так і скінченними послідовностями. У задачах з розпізнавання природних сигналів вхідними даними є ці сигнали, які описуються розміщенням з повтореннями (мовленнєві сигнали, електрокардіограми, електроенцефалограми). Оскільки ця комбінаторна конфігурація на підмножині ізоморфних розміщень є скінченною, а на усій множині – нескінченною, то вона визначає нечіткість вхідної інформації.

Література References

[1] М. Шлезингер, В. Главач, "Десять лекций по статистическому и структурному распознаванию", Киев: Наук. думка, 2004, 546 с.

[2] Т. К. Винцюк, "Анализ, распознавание и интерпретация речевых сигналов", К.: Наукова думка, 1987, 262 с.

[3] Н. К. Тимофієва, "Теоретико-числові методи розв'язання задач комбінаторної оптимізації". Автореф. дис... докт. техн. Наук, Ін-т кібернетики ім. В. М. Глушкова НАН України, Київ, 2007, 32 с.

Застосування Віртуалізованої IT-Інфраструктури для Збільшення Рівня Екологічної Безпеки Технологічних Систем

Вавулін Петро
кафедра кібернетики хіміко-технологічних процесів
Національний технічний університет України «Київський політехнічний інститут» ім. І. Сікорського
Київ, Україна
lestatxa81@gmail.com

Using of Virtualized IT-Infrastructure for Ecological Risk Decrease of Automation Systems of Technological Objects

Vavulin Petro
Department of Cybernetics of Chemical Technology Processes
National Technical University of Ukraine
«Igor Sikorsky Kyiv Polytechnic Institute»
Kyiv, Ukraine
lestatxa81@gmail.com

***Анотація*—Зважаючи на глобальні перспективи та досвід впровадження сучасних інформаційних технологій у виробництво, було вирішено розглянути можливі шляхи зменшення екологічного ризику українських виробництв. Було розглянуто проблеми і методи використання програмно – апаратних рішень в області віртуалізації для експлуатованих виробничих комплексів. Головним методом дослідження є комп'ютерне моделювання реальних систем автоматизації та керування технологічними процесами з використанням серверів та засобів віртуалізації, в тому числі гіпервізорів (MS Hyper V). Метод базується на використанні повністю віртуалізованої IT-інфраструктури, у тому числі систем керування та робочих станцій.**

***Abstract*—Taking into account the global trends and experience of implementation of modern information technologies in production processes, with the aim of updating and increasing the ecological safety of Ukrainian industrial complexes, the issues and methods for using of hardware and software and technological solutions in the field of virtualization are considered. The main method of research is computer simulation of real automation systems (including server hardware) using the tools of virtualization (Microsoft Hyper V). The method based on using of a virtual IT infrastructure, including primary and backup servers with process control system and workstations.**

The ways of using of traditional automation systems, which are deployed on the basis of virtualization platform Hyper V, are considered. Also considered deployment of these systems under server OS with support for one of the many virtualization technologies, such as: MS Hyper V, Citrix Xen Server, and others.

The possible positive effect of using modern IT infrastructure for technological objects under operation conditions is also analyzed. It lays in the fact of theoretically increase of fault tolerance level, practical simplification of system administration, and creation of backup storage of virtual machines.

Ключові слова—IT-інфраструктура; гіпервізор; тонкий клієнт; прикладні IT-технології

Keywords—IT infrastructure virtualization; hypervisor; thin client; applied IT

I. Вступ

Системи керування і захисту технологічних процесів стають все більш складними і виникає проблема інтегрованості цих систем з сучасними технологіями для їх використання в виробничих галузях, які пов'язані з екологічним ризиком.

Кожний технологічний процес, пов'язаний з виробництвом, повинен використовувати сучасні технології, які дозволяють максимально знизити ймовірність аварії і зменшити надходження небезпечних речовин у навколишнє середовище. Також, на даному етапі розвитку технологій, забезпечити повністю безаварійну роботу технічних систем не можливо. Саме тому, побудова гіперконвергентних рішень з використанням

віртуалізованих середовищ для побудови систем автоматизації надзвичайно важливе, особливо у випадку наявності вимог відповідності виробництв стандартам ISO 14000 та ISO/IEC 17799 [1, 2].

Використання віртуалізації (та її складових як програмних, так і апаратних) передбачає підвищення рівня надійності та відмовостійкості, оскільки воно передбачає відновлення працездатності системи шляхом запуску резервних віртуальних машин чи застосування резервних апаратних елементів (тонкі клієнти). Слід відмітити, що окрім ймовірного підвищення рівня надійності, використання сучасних гіпервізорів також спрощує адміністрування та керування системами автоматизації [3].

Слід зазначити, що питання оцінки і попередження виникнення аварійних ситуацій і катастроф, а також їх економічної та екологічної складової, носить надзвичайно важливий характер [4]. В контексті актуальності питань аварійної безпеки, виявляється необхідним розвиток та впровадження сучасних IT-технологій для зменшення екологічних ризиків. Також, підвищення надійності і рівня відмовостійкості технологічних об'єктів, включаючи елементи чи навіть повноцінні системи автоматизації, є невід'ємною частиною еволюційного розвитку IT-інфраструктури промислових об'єктів та систем.

II. ОБ'ЄКТ І МЕТА ДОСЛІДЖЕННЯ

Об'єктом дослідження даної роботи являється комп'ютерно-інтегроване управління технічними системами з використанням засобів віртуалізації.

Розглядається система автоматизації та керування технологічним об'єктом з використанням програмно-апаратних рішень компанії Honeywell, рис. 1. Дана система включає – контролер C200, ТОУ (технологічний об'єкт управління) та програмне забезпечення Experion PKS [5].

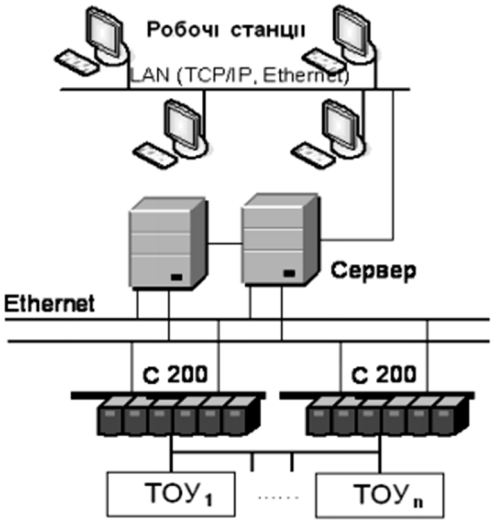


Рис. 1. Узагальнена принципова схема системи керування Honeywell

Дана система використовує сучасні сервери на компонентній базі корпорації Intel та програмне забезпечення Microsoft Windows Server та Honeywell Experion PKS. Також допускається використання серверів на базі AMD. В якості сенсорів для отримання необхідних даних стосовно стану об'єкта управління найбільш поширеним є використання сенсорів виробництва компанії Honeywell.

Одним з проблемних та вузьких місць даного рішення є відсутність автоматизованого резервного копіювання сервера та робочих станцій. Також до наявних недоліків можна віднести відсутність централізованої консолі управління та моніторингу стану робочих станцій і серверів та їх резервних копій. У контексті розвитку сучасних технологій та для зменшення рівня екологічної безпеки, вивчення можливостей побудови системи комп'ютерно-інтегрованого управління з використанням можливостей та інструментів віртуалізації є надзвичайно важливим та необхідним.

Метою даної роботи є дослідження можливості впровадження віртуалізованих середовищ при експлуатації систем автоматизації технологічних процесів для пілвищення рівня екологічної безпеки. Впровадження даних технологій дозволить спростити інтеграцію елементів технічних систем та агрегування генерованих ними даних. Це розширить можливості застосування методів визначення функції розподілу випадкової величини, яка характеризує рівень техногенного і екологічного ризику відмови складної системи у часі [6, 7]. Сучасна концепція безпеки полягає у тому, що міжнародні стандарти розглядають систему безпеки комплексно, в цілому із урахуванням апаратного та програмного резервування всіх компонентів. До компонентів системи захисту включають вимірювальні і виконавчі пристрої, конфігурації приладів та операційних систем і т. д.

III. АНАЛІЗ ЛІТЕРАТУРНИХ ДАНИХ

Проблеми застосування математичних методів до оцінки екологічних та економічних ризиків постійно перебувають у полі зору науковців. Слід зазначити, що за сучасних умов розвитку технологій для забезпечення високго рівня надійності та відмовостійкості технічних систем та об'єктів необхідно мати надзвичайно ефективну та сучасну IT-інфраструктуру [8].

Відмова від використання сучасних рішень та управління технологічними об'єктами з використанням неефективних або застарілих технологій може обернутися негативним результатом в діяльності підприємства. Як наслідок, проблеми адаптації сучасних технологій до існуючих технологічних систем (об'єктів) з урахуванням методології оцінювання ризику актуальні у будь-який час [8]. На даному етапі питання інтеграції традиційних систем автоматизації із сучасними IT-технологіями недостатньо досліджене.

Розглянемо задачу проектування та впровадження IT-інфраструктури та системи автоматизації (керування) ним з використанням сучасних програмно-апаратних рішень. Використання віртуалізованих систем (середовищ) є поширеним, і що більш важливо – ефективним [9].

На сьогодні, технології віртуалізації набувають досить широкого розповсюдження у різноманітних сферах виробничої діяльності. У харчовій промисловості віртуалізація розглядається як елемент системи автоматизації та збільшення економічної ефективності виробництва [10]. У промисловості та при використані розподілених обчислювальних середовищ, технології віртуалізації застосовуються для забезпечення гарантованого рівня надійності (Service Level Agreements) [11].

Для виробничих процесів та систем управління набуває все більшого розповсюдження використання технологій та засобів віртуалізації і обробки великих об'ємів даних [8]. Слід відмітити, що питання інформаційної безпеки мережевої, апаратної та програмної складових віртуалізованої IT-системи вирішується застосуванням спеціалізованих засобів та інструментів [12–14]. Дані інструменти включають як адміністративні засоби (посадові інструкції з розмежуванням прав доступу) так і програмні (антивірусне програмне забезпечення, апаратні засоби від взлому та DDos attack).

Слід відмітити, що до традиційних систем автоматизації, зазвичай майже не включають, або ігнорують можливості використання засобів віртуалізації для підвищення відмовостійкості та оптимізації керування [12].

Оптимальна стратегія керування технологічним процесом базується на традиційних підходах до побудови автоматизованих систем керування технологічним процесом (АСУ, процес обробки та відображення інформації та ін.). Також при побудові використовується сукупність критеріїв та параметрів (в тому числі керуючих впливів) для прийняття оптимальних рішень.

З проведеного літературного аналізу можна зробити висновок, що проблеми адаптації сучасних технологій до існуючих технологічних систем (об'єктів) актуальні у будь-який час.

IV. Матеріали та методи дослідження

Розглянемо задачу проектування та впровадження IT-інфраструктури технологічних процесів та системи автоматизації (керування) з використанням сучасних програмно-апаратних рішень у складі:

– спеціалізоване програмне забезпечення (MS Hyper V, VMWare VSphere, Citrix Xen Server та ін.);

– тонкі клієнти (Напр. Raspberry PI чи HP, Dell);

– сервери керування та збереження даних (HP, Dell, Supermicro та ін.).

При віртуалізації програмне забезпечення використовується для імітації наявності обладнання і створення віртуальної комп'ютерної системи, що в свою чергу може бути використано для оптимізації роботи систем керування технологічними процесами [10]. Слід відмітити, що питання безпечності використання даних технологій вирішуються у відповідності до вимог виробництва та сучасних понять кібербезпеки [11].

Розглянемо апаратну віртуалізацію, як інструмент збільшення ефективності утилізації обчислювальних ресурсів. Апаратна віртуалізація забезпечує продуктивність, порівнянну з продуктивністю невіртуалізованної машини. Це в сою чергу дозволяє використовувати віртуальні машини у продуктивному середовищі. Найбільш поширені технології віртуалізації – Intel-VT і AMD-V [13].

Розглянемо сценарій практичного застосування запропонованого підходу, з використанням гіпервізора Hyper V для системи автоматизації, побудованої на базі технологічних рішень американської компанії Honeywell.

Для віртуалізації використаємо Microsoft Hyper-V – систему апаратної віртуалізації для x64-систем, на базі гіпервізора. Microsoft Hyper-V дозволяє створювати віртуальну IT-інфраструктуру на базі існуючої фізичної.

Архітектурно система включає в себе глобальний сервер (або сервери), на яких розміщується базове програмне забезпечення та віртуальні машини. Глобальні сервери об'єднуються між собою та з тонкими клієнтами локальною обчислювальною мережею (також допускається віртуалізації мережі засобами Hyper V чи будь-якого іншого гіпервізора) [13]. Тонкі клієнти під'єднуються до виділених їм віртуальних машин. Варіантами вирішення питань інформаційної безпеки, в залежності від потреб можуть бути: засоби ОС, використання Microsoft Active Directory, засоби LDAP та інші [13]. Керування технологічним процесом здійснюється через віртуальні машини, які під'єднані до обладнання та програмного забезпечення Honeywell. На рис. 2 зображено апаратну складову, на рис. 3 – програмну.

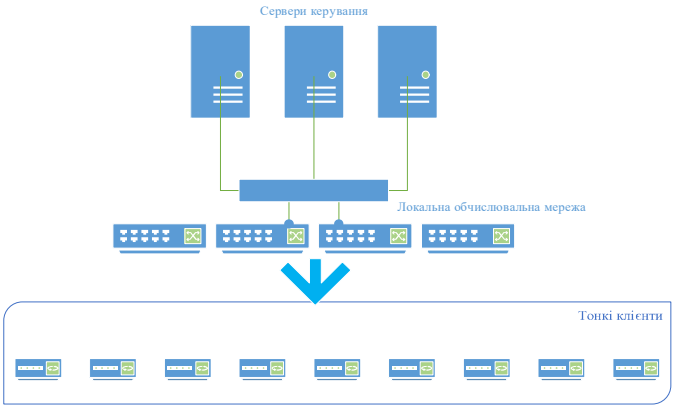


Рис. 2. Графічне зображення базової апаратної інфраструктури системи автоматизації з застосуванням засобів віртуалізації.

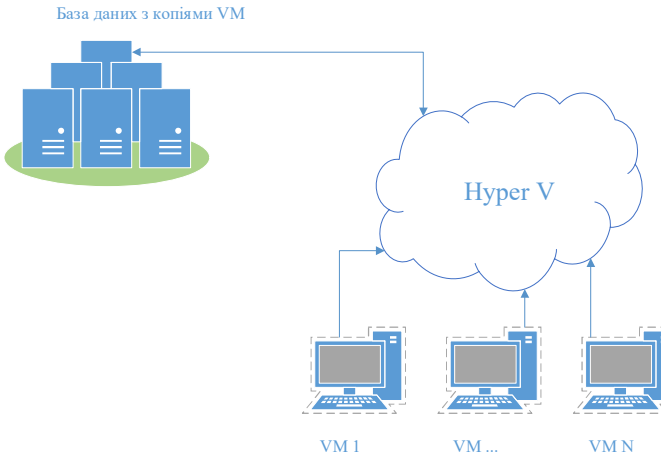


Рис. 3. Графічне зображення віртуальної інфраструктури системи автоматизації.

Програмна складова включає в себе гіпервізор Hyper V, за допомогою якого відбувається моніторинг та керування віртуальним середовищем. Адміністратор мережі задає налаштування політик відновлення та створення копі1 активних віртуальних машин, для можливості швидкого відновлення роботи системи автоматизації у разі відмови елементів системи.

Схему процедури відновлення працездатності елементів системи автоматизації наведено на рис. 4. У випадку відмови віртуальної машини сервер в автоматичному режимі запускає повністю ідентичну їй копію, попередньо збережену у базі даних.

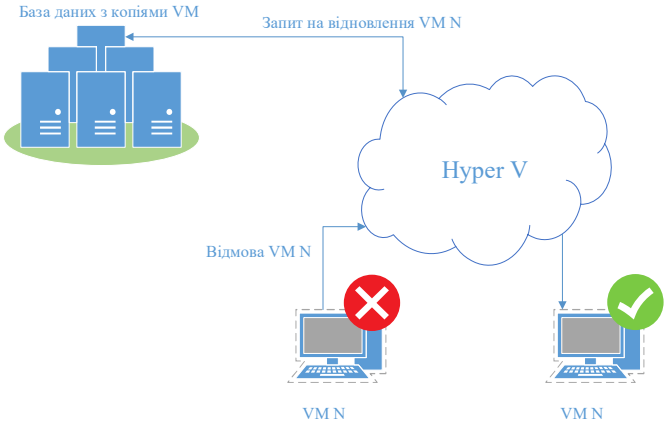


Рис. 4. Схема процедури відновлення працездатності віртуальних машин.

Для цього піднімемо роль Hyper V Server, а на ньому створимо віртуальну машину з сервером Experion PKS. На рис. 5 відображено диспетчер Hyper V, який надає засоби керування віртуальною інфраструктурою.

Для запуску віртуального сервера достатньо натиснути кнопку «пуск» у відповідному меню доступних дій. Як видно з рис. 6–7, створений віртуальний сервер за всіма атрибутами відповідає класичному серверу. Слід зазначити, що використання засобів віртуалізації спрощує інтеграцію між об'єктами керування та засобами прогнозування техногенного ризику [8]. При віртуалізації серверів та робочих станцій Honeywell було використано звичайний персональний комп'ютер (об'єм оперативної пам'яті 8 Гб, двох-ядерний процесор Intel Core i5-3230M з тактовою частотою 2.6 GHz).

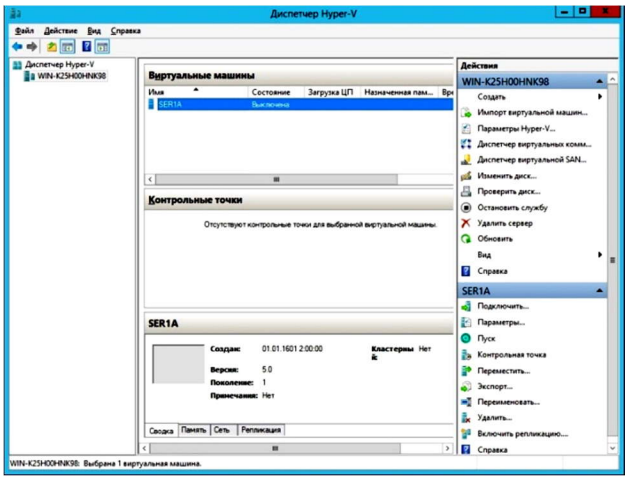


Рис. 5. Середовище керування віртуальною інфраструктурою MS Hyper V.

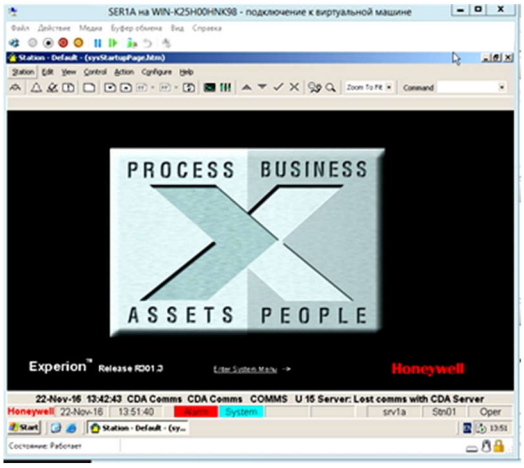


Рис. 6. Запуск віртуального сервера керування з спеціалізованим програмним забезпеченням Experion PKS.

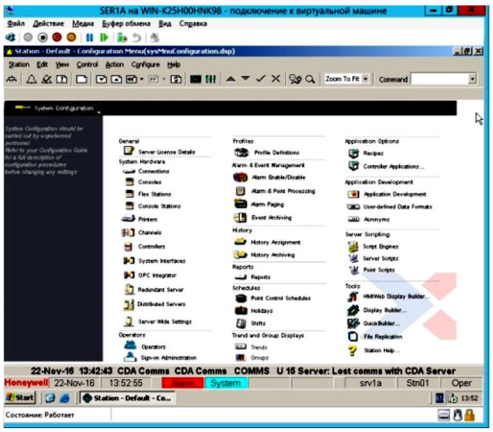


Рис. 7. Налаштування системи у віртуалізованому середовищі.

Для оцінки економічного ефекту розглянемо систему, що складається з 30 робочих станцій, та її аналог з використанням засобів віртуалізації. Результати розрахунків наведено в табл. 1.

ТАБЛИЦЯ I. ПОРІВНЯННЯ ВАРТОСТІ ТРАДИЦІЙНОГО І ВІРТУАЛІЗОВАНОГО РІШЕННЯ

Найменування	Вартість за одиницю, грн.	Кількість	Сума, грн.
Робоча станція (ПК чи ноутбук середнього цінового діапазону)	13 000	30	390 000
Тонкий клієнт RASPBERRY PI 3 MODEL B	1 400	30	42 000
Сервер віртуалізації Supermicro SYS-6028R-T	90 000	2	180 000

Ціни приведені станом на березень 2017, для ринку України. Програмне резервування (точні копії віртуальних машин, які використовуються тонкими клієнтами для операційної діяльності) обмежене лише вільним місцем на сервері збереження даних, апаратне резервування в даній конфігурації не розглядається.

V. РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Розрахуємо сумарну економію при використанні віртуалізованого середовища, у порівнянні з традиційним рішенням:

$$E = ((Ц_{трад} - Ц_{вірт})/ Ц_{трад})\cdot 100\% = ((390\,000 - 222\,000) / 390\,000)\cdot 100\% \approx 43\ \%$$

Розглянемо використання серверу Supermicro SYS-6028R-T (США) у конфігурації з 8 дисками 1000 GB HDD Enterprise SATA 7200 rpm у випадку побудови RAID 1 (дзеркалювання інформації, що зберігається, тобто з дублюванням збережених даних). Даний сервер було обрано виходячи з цінових та якісних характеристик обладнання. Кількість – 2 сервери обумовлена гранично допустимою кількістю дисків, та обчислювальними можливостями центральних процесорів кожного з них. В даній конфігурації для побудови віртуалізованої інфраструктури буде доступно 2 ТБ (терабайти) дискового середовища на кожному з трьох серверів (загальна кількість – 4 ТБ). З урахуванням реальних даних стосовно програмного середовища (операційна система та спеціалізоване програмне забезпечення Experion PKS) серверів керування технологічними процесами Honeywell, для кожної віртуальної машини буде достатньо 15 Гб дискового місця.

Проведемо дослідження впливу застосування віртуалізації на рівень надійності та екологічної безпеки автоматизованої системи з 30 робочими місцями. Результати дослідження наведено у табл. 2. На рис. 8 відображено загальну допустиму кількість копій відповідних вузлів необхідних для нормальної експлуатації системи автоматизації.

ТАБЛИЦЯ II. ПОРІВНЯННЯ НАДІЙНОСТІ ТРАДИЦІЙНОГО І ВІРТУАЛІЗОВАНОГО РІШЕННЯ

Найменування	Традиційні рішення	Віртуалізована IT-інфраструктура
Кількість фізичних або відповідних їм віртуальних машин, шт.	30	260
Гарантована мінімальна кількість унікальних копій фізичних або відповідних їм віртуальних машин, шт.	0	2
Загальна кількість резервних копій окремих віртуальних машин, шт.	1	8

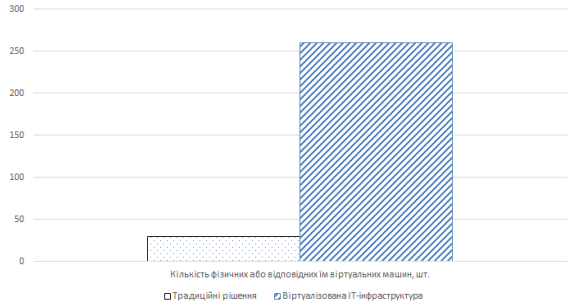


Рис. 8. Максимально можлива кількість екземплярів робочих місць чи серверів, або їх копій у відповідності до місця на дисках.

VI. ВИСНОВКИ

Проаналізувавши одержані результати, можна сказати, що використання засобів віртуалізації значно підвищує рівень екологічної безпеки та відмовостійкості систем автоматизації. Також до побічних ефектів можна віднести значний економічний ефект від впровадження більш надійних систем. Слід зазначити, що на рис. 8 мається на увазі не кількість одночасно запущених віртуальних машин, а лише кількість копій (знімків) робочих станцій у базі даних трьох серверів (кількість одночасно запущених копій обмежена об'ємом оперативної пам'яті та можливостями серверних процесорів). Також, у табл. 2 мається на увазі максимально можлива кількість копій однієї робочої станції у базі даних для швидкого відновлення працездатності системи. Стосовно гарантованої мінімальної кількості унікальних копій фізичних або відповідних їм віртуальних машин – мається на увазі випадок, коли в традиційних системах використовуються звичайні робочі станції (без RAID-контролерів), тому відповідна гарантована кількість фізичних копій дорівнює нулю, на відміну від віртуалізованого рішення в якому побудовано RAID 0 з дублюванням збережених даних (в тому числі й віртуальних машин).

Проаналізувавши результати досліджень можна зробити наступні висновки:

– програмно-апаратна реалізація віртуалізованої інфраструктури має значний потенціал збільшення надійності систем автоматизації, оскільки дані рішення дозволяють швидко відновлювати програмну та апаратну працездатність елементів систем автоматизації. До переваг можна віднести

простоту практичного застосування даних систем у поєднанні з традиційними системами автоматизації;

– застосування віртуалізованої IT-інфраструктури економічно доцільне (табл. 1) і є більш вигідним ніж традиційні рішення у зв'язку з повною ідентичністю систем які розгортаються у віртуалізованому середовищі традиційним.

– застосування віртуалізованої IT-інфраструктури підвищує рівень екологічної безпеки (табл. 2) завдяки значно вищому рівню резервування експлуатованих систем та їх програмних складових.

Також слід зазначити, що практичне застосування запропонованих технологій може призвести до різкого збільшення рівня надійності, керованості та простоти адміністрування сучасних систем автоматизації. Також слід відзначити еволюційність даних підходів, адже вони ніяк не заперечують і мають зворотну сумісність з майже усіма наявними програмно-апаратними системами автоматизації технологічних процесів.

Головними перевагами впровадження платформ віртуалізації таких, як MS Hyper V, Citrix Xen Server, VMWare та ін. є:

– відносна простота використання;

– практично необмежені можливості по створенню клонів фізичних систем автоматичне їх відновлення (обмежене лише доступним місцем для збереження даних);

– «легка міграція» або переїзд системи автоматизації з одного географічного регіону до іншого та ін.

REFERENCES

[1] T. Haeberlen, L. Dupre, (2012, December), Cloud Computing Benefits, risks and recommendations for information security. European Network and Information Security Agency (ENISA). Available: https://resilience.enisa.europa.eu/cloud-security-and-resilience/publications/cloud-computing-benefits-risks-and-recommendations-for-information-security/at_download/file

[2] M. M. Bychenok, S. P. Ivaniuta, Ye. O. Yakovliev, Institute for National Security of the National Security and Defense of Ukraine. (2008). Ryzyky zhyttiediialnosti u pryrodno-tekhnohennomu seredovyshchi. Kyiv, 160.

[3] , S. H. Bondarenko, D. O.Skoretskyi, (2016), Computer-microprocessor system of technological processes control. Kompiuterne modeliuvannia v khimii i tekhnolohiiakh ta systemakh staloho rozvytku – KMKhT-2016: zbirnyk naukovykh statei Piatoi mizhnarodnoi naukovo-praktychnoi konferentsii, 18-20 travnia 2016 roku, m. Kyiv. Kyiv: National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute», 50–56.

[4] P. Vavulin, T. Boyko, (2016), Analysis of algorithm for estimating distribution functions of random variables for the prediction of technogenic risk. Technology Audit And Production Reserves, 2(3(28)), 17–23. doi:10.15587/2312-8372.2016.66754

[5] R. M. Sinnamon, J. D. Andrews, (1996), Quantitative Fault Tree Analysis Using Binary Decision Diagrams. Journal Europeen des Systemes Automatises, 30 (8), 1051–1071.

[6] E. Galante, D. Bordalo, M. Nobrega, (2014), Risk Assessment Methodology: Quantitative HazOp. Journal of Safety Engineering, 3 (2), 31–36. doi:10.5923/j.safety.20140302.01

[7] I. S. Saguy, (2016), Challenges and opportunities in food engineering: Modeling, virtualization, open innovation and social responsibility. Journal of Food Engineering, 176, 2–8. doi:10.1016/j.jfoodeng.2015.07.012

[8] , A. Kertesz, G. Kecskemeti, I. Brandic, (2014), An interoperable and self-adaptive approach for SLA-based service virtualization in heterogeneous Cloud environments. Future Generation Computer Systems, 32, 54–68. doi:10.1016/j.future.2012.05.016

[9] R. F. Babiceanu, R. Seker, (2016), Big Data and virtualization for manufacturing cyber-physical systems: A survey of the current status and future outlook. Computers in Industry, 81, 128–137. doi:10.1016/j.compind.2016.02.004

[10] V. A. Tiurin, (2006), Avtomatizirovannye sistemy upravleniia tehnologicheskimi protsessami. St. Petersburg: St. Petersburg State Academy of Forestry Engineering, 153.

[11] T. Hegazy, M. Hefeeda, (2015), Industrial Automation as a Cloud Service. IEEE Transactions on Parallel and Distributed Systems, 26 (10), 2750–2763. doi:10.1109/tpds.2014.2359894

[12] R. Chandramouli, (2014), Analysis of Protection Options for Virtualized Infrastructures in Infrastructure as a Service Cloud. Fifth International Conference on Cloud Computing, GRIDs, and Virtualization. Venice, Italy, 37–43.

[13] A. Menon, A. L. Cox, W. Zwaenepoel, (2006), Optimizing Network Virtualization in Xen. In Proceedings of the annual conference on USENIX'06 Annual Technical Conference. USENIX Association. Available:

[14] K. Hashizume, D. G. Rosado, E. Fernandez-Medina, E. B. Fernandez, (2013), An analysis of security issues for cloud computing. Journal of Internet Services and Applications, 4 (5), 15–28. doi:10.1186/1869-0238-4-5

Комп’ютерне Моделювання Дифузійних Потоків Водню у Шаруватому Композиті Мідь-Залізо за Арксинус-Розподілу Включень

А.Є. Чучвара
Відділ математичного моделювання нерівноважних процесів
Центр математичного моделювання
Інституту прикладних проблем механіки і математики ім. Я.С. Підстригача НАН України
Львів, Україна
davydoka@gmail.com

Simulation of Hydrogen Diffusion Flow in a Layered Copper-iron Composite under Arcsine Distribution of Inclusions

A. Chuchvara
Department of mathematical modeling nonequilibrium processes
Centre of Mathematical Modeling of Pidstryhach Institute for Applied Problems of Mechanics and Mathematics,
The National Academy of Sciences of Ukraine
Lviv, Ukraine
davydoka@gmail.com

***Анотація*—Одержано формулу для усередненого потоку маси у двофазній випадково неоднорідній смузі з арксинус-розподілом шаруватих включень за ненульової сталої початкової концентрації домішки у тілі. На прикладі дифузії водню у шаруватому композиті мідь-залізо проведено комп’ютерне моделювання та встановлено залежність усередненого дифузійного потоку від характеристик середовища та параметрів структури. Показано, що збільшення товщини прошарків заліза у композиті збільшує потоки водню.**

***Abstract*—The formula for the averaged diffusion mass flow in a two-phase random nonhomogeneous strip with the arcsine distribution of layered inclusions is obtained under the condition of nonzero constant initial concentration of admixture in the body. As an example of hydrogen diffusion in a copper-iron composite simulation is carried out and dependence of the averaged diffusion flow on medium characteristics and structure parameters is established. It is shown that increase of the thickness of iron inclusions in the composite leads to increase of hydrogen flows.**

Ключові слова—комп’ютерне моделювання; дифузійний потік; випадково неоднорідна структура; арксинус-розподіл; композит мідь-залізо

Keywords—simulation; diffusion flow; random nonhomogeneous structure; arcsine distribution; copper-iron composite

I. Вступ

У сучасному автомобілебудуванні, металургійній та гірничовидобувній промисловості, сільському господарстві, приладобудуванні, медицині, будівельній галузі тощо одним із найбільш поширених класів композитних матеріалів є шаруваті композити з основою із металічної матриці (рис. 1) [1, 2].

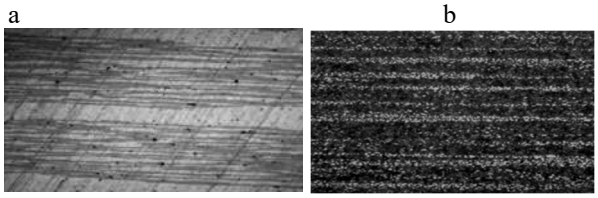


Рис. 1. Мікрофотографії шаруватого композиту мідь-залізо [3] (a) та булату із включеннями вторинного цементиту [4] (b)

Використання полі- та біметалів дозволяє істотно скоротити витрату високолегованих сталей, дефіцитних і дорогих кольорових металів. Комбінуючи об’ємний вміст компонентів, можна отримувати композити з необхідними значеннями міцності, модуля пружності, термо-, зносо- і корозостійкості, а також створювати матеріали з необхідними магнітними, діелектричними, радіопоглинаючими та іншими спеціальними властивостями.

За умов неповної інформації про геометричні параметри тіл зі складною внутрішньою структурою виникає необхідність опису фізичних процесів, які протікають у таких середовищах, випадковими функціями або полями. Тоді для встановлення основних закономірностей процесів дифузії використовують процедуру усереднення за ансамблем конфігурацій фаз. Однак при моделюванні такої характеристики процесу масоперенесення як дифузійний потік виконання процедури усереднення викликає значні труднощі, оскільки невідомою є функція кореляції градієнта випадкового поля концентрації домішкової речовини та стохастичного коефіцієнта дифузії. У працях [5, 6] для вирішення цієї проблеми запропоновано складати балансові рівняння для вже гомогенізованих середовищ з фізичними характеристиками, що є усередненими величинами і враховують відмінності між фазами, проте нехтується взаємодією між ними. Й. Келлер визначає випадковий потік за законом Дарсі, коефіцієнт фільтрації якого є випадковою функцією просторової координати і для побудови розв'язку задачі використовує методи малих збурень та згладжування з відповідними обмеженнями [7].

У праці [8] запропоновано підхід до математичного опису дифузійних потоків у випадково неоднорідних шаруватих тілах, який базується на отриманні рівняння дифузії для функції потоку, обґрунтуванні коректного формулювання початкових і граничних умов та знаходженні розв'язку крайової задачі у вигляді ряду Неймана, зручного для проведення процедури усереднення за ансамблем конфігурацій фаз. За таким підходом у даній роботі досліджено потоки домішкової речовини у двофазній випадково неоднорідній шаруватій смузі на прикладі дифузії водню у композитному матеріалі мідь-залізо (рис. 1а) за арксинус-розподілу залізних включень.

II. Математична модель

Нехай у двофазній смузі товщини z_0, яка складається з n_0 підшарів фази $j=0$ (матриця) та n_1 підшарів фази $j=1$ (включення), відбувається процес дифузії домішкової речовини. Припускаємо, що точні координати розташування включень є невідомими, але приймаємо, що прошаки у тілі найбільш ймовірно розташовані біля границь тіла (рис. 2). При цьому коефіцієнти дифузії вважаємо сталими в межах кожної з фаз і рівні відповідно D_0 для області матриці та D_1 для області включення.

Для опису випадково неоднорідної шаруватої структури типу поданої на рис.1а та рис. 2 використано ймовірнісний арксинус-розподіл, що є частковим випадком бета-розподілу з функцією густини вигляду [9]

$$f(z)=\begin{cases}\dfrac{\Gamma(2\alpha)}{\Gamma^2(\alpha)}\left(\dfrac{z_0}{z}\right)^{1-\alpha}\left(\dfrac{z_0}{z_0-z}\right)^{1-\alpha}, & z\in[0;z_0];\\ 0, & z\notin[0;z_0],\end{cases} \qquad (1)$$

де $\Gamma(x)$ – гама-функція, α – ступінь вільності розподілу ($0<\alpha<1$).

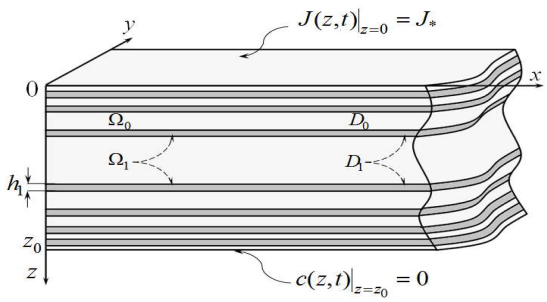


Рис. 2. Схематична реалізація двофазної багатошарової смуги з найбільш ймовірним розташуванням включень біля границь тіла

Рівняння дифузії для випадкової функції потоку $J(z,t)$ в одновимірному за просторовою координатою випадку має вигляд [10]

$$\frac{\partial J(z,t)}{\partial t}=D(z)\frac{\partial^2 J(z,t)}{\partial z^2}, \qquad (2)$$

де $D(z)=\begin{cases}D_0, z\in\Omega_0;\\ D_1, z\in\Omega_1,\end{cases}$ – випадковий коефіцієнт дифузії, Ω_j – багатозв'язна область j-ої фази ($j=0;1$).

Прийнято, що у початковий момент часу відсутній дифузійний потік у тілі, на поверхні $z=0$ підтримується стале значення потоку $J_*\equiv const$, а на границі $z=z_0$ концентрація мігруючої речовини $c(z,t)$ рівна нулю, тобто задано наступні крайові умови:

$$J(z,t)|_{t=0}=0;\quad J(z,t)|_{z=0}=J_*,\quad c(z,t)|_{z=z_0}=0. \qquad (3)$$

Крім того, прийнято, що в початковий момент часу концентрація домішки є сталою і рівна c_*. Гранична умова на межі $z=z_0$ задається для функції концентрації, оскільки визначення дифузійного потоку на цій границі потребує додаткових досліджень.

Приймаючи неоднорідність структури тіла як внутрішні джерела, вихідну задачу (2)-(3) зведено до еквівалентного інтегро-диференціального рівняння, що є сумою розв'язку однорідної крайової задачі $J_0(z,t)$ та згортки детермінованої функції Гріна $G(z,z',t,t')$ з джерелом

$$J(z,t)=J_0(z,t)+\int_0^t\int_0^{z_0}G(z,z',t,t')L_s(z')J(z',t')dz'dt'. \qquad (4)$$

Тут $L_s(z)=\sum_{i=1}^{n_1}(D_1-D_0)\eta_{i1}(z)\dfrac{\partial^2}{\partial z^2}$, $\eta_{i1}(z)=\begin{cases}1, & z\in\Omega_{i1};\\ 0, & z\notin\Omega_{i1},\end{cases}$ – випадкова «функція структури», $\sum_{i=1}^{n_1}\Omega_{i1}=\Omega_1$.

Розв'язок однорідної крайової задачі за ненульової початкової концентрації має вигляд [8]

$$J_0(z,t) = J_* - \frac{2}{z_0}\sum_{n=1}^{\infty} e^{-D_0\xi_n^2 t}\left(\frac{J_*}{\xi_n} + (-1)^n c_* D_0\right)\sin(\xi_n z)\,. \quad (5)$$

Тут $\xi_n = \pi(2n-1)/2z_0$.

Функція Гріна є розв'язком однорідної крайової задачі з точковим джерелом [8]

$$G(z,z',t,t') = \frac{\theta(t-t')}{\pi}\sum_{k=1}^{\infty} e^{-D_0 y_k^2 (t-t')} \times$$

$$\times\left[\cos(y_k(z-z')) - \cos(y_k(z+z'))\right], \quad (6)$$

де $y_k = k\pi/z_0$, $\theta(t-t')$ – одинична сходинкова функція Хевісайда.

Розв'язок рівняння (4) знайдено методом послідовних наближень у вигляді абсолютно і рівномірно збіжного ряду Неймана [11, 12]. Для виконання процедури усереднення за ансамблем конфігурацій фаз обмежимось двома першими членами цього ряду

$$J(z,t) \approx J_0(z,t) + \int_0^t\int_0^{z_0} G(z,z',t,t')L_s(z')J_0(z',t')dz'dt'\,.$$

Приймаючи, що включення мають характерну (середню) товщину h_1, знайдено формулу для дифузійного потоку, усередненого за ансамблем конфігурацій фаз із функцією розподілу включень $f(z)$ (1)

$$\langle J(z,t)\rangle = J_0(z,t) + \frac{(D_1-D_0)Bv_1}{h_1}\int_0^t\left[\int_0^{h_1} z'^{\alpha} z_0^{\alpha}\,{}_2\overline{F}_1\left(\frac{z'}{z_0}\right)\times\right.$$

$$\times G(z,z',t,t')\frac{\partial^2 J_0(z',t')}{\partial z'^2}dz' + \int_{h_1}^{z_0} z_0 h_1^{\alpha}(z_0+h_1-z')^{\alpha-1}\times$$

$$\left.\times\,\overline{F}\left(\frac{h_1}{z_0+(z'-h_1)}\right)G(z,z',t,t')\frac{\partial^2 J_0(z',t')}{\partial z'^2}dz'\right]dt', \quad (7)$$

де $B = \Gamma(2\alpha)\big/\alpha\Gamma^2(\alpha)$, $\overline{F}(z) = {}_2F_1(\alpha,1-\alpha;1+\alpha,z)$,

${}_2F_1(a,b;c;z) = \sum_{m=0}^{\infty}\frac{(a)_m(b)_m}{(c)_m}\frac{z^m}{m!}$ – гіпергеометрична функція, $(a)_m = a(a+1)\ldots(a+m-1)$, $(a)_0 = 1$.

Підставивши у (7) вирази для функції Гріна (5) та дифузійного потоку в однорідному тілі (6), отримаємо

$$\frac{\langle J(z,t)\rangle}{J_*} = 1 - \frac{2}{z_0}\sum_{n=1}^{\infty} e^{-D_0\xi_n^2 t}\left(\frac{1}{\xi_n} + (-1)^n\frac{D_0 c_*}{J_*}\right)\sin(\xi_n z) +$$

$$+\frac{2(D_1-D_0)v_1 B}{D_0 z_0^{\alpha} h_1}\sum_{k=1}^{\infty}\sum_{n=1}^{\infty}\frac{\xi_n}{y_k^2-\xi_n^2}\left(e^{-D_0\xi_n^2 t} - e^{-D_0 y_k^2 t}\right)\times$$

$$\times\left(1+(-1)^n\frac{D_0 c_*}{J_*}\xi_n\right)\sin(y_k z)\left[z_0^{\alpha-1}\times\right.$$

$$\times\left(f^+(0,x_{kn}^-) - f^+(0,x_{kn}^+) + \sum_{m=1}^{\infty}\frac{\alpha(1-\alpha)\ldots(m-\alpha)}{(m+\alpha)m!z_0^m}\times\right.$$

$$\left.\times\left[f^+(m,x_{kn}^-) - f^+(m,x_{kn}^+)\right]\right) + h_1^{\alpha}\left(f^-(0,x_{kn}^-) -\right.$$

$$-f^-(0,x_{kn}^+) + \sum_{m=1}^{\infty}\frac{\alpha(1-\alpha)\ldots(m-\alpha)h_1^m}{(m+\alpha)m!}\times$$

$$\left.\left.\times\left[f^-(m,x_{kn}^-) - f^-(m,x_{kn}^+)\right]\right)\right], \quad (8)$$

де $x_{kn}^{\pm} = y_k \pm \xi_n$, $f^+(a,p) = \int_0^{h_1}(z)^{\alpha+a}\cos(pz)dz$,

$f^-(a,p) = \int_{h_1}^{z_0}(z_0+h_1-z')^{\alpha-(a+1)}\cos(pz)dz$.

Вираз (8) є розрахунковою формулою для усередненого дифузійного потоку у двофазній смузі з найбільш ймовірним розташуванням включень біля границь тіла за умови ненульової початкової концентрації частинок домішки.

III. Потік водню в композиті купрум-ферум

Проаналізовано кількісну та якісну залежність усередненого дифузійного потоку у двофазній багатошаровій смузі з арксинус-розподілом включень на прикладі міграції водню у шаруватому композиті $Cu-Fe$ (рис. 1а), де в якості базової фази прийнято мідь. Оскільки дифузійні процеси належать до повільних фізичних процесів – часи їх протікання (в секундах) є великими в числовому вимірі, а характерні довжини при цьому вимірюються в метрах, – тому при комп'ютерному моделюванні здійснено перехід до безрозмірних змінних $\varsigma = z/z_0$, $\tau = D_0 t/z_0^2$. Коефіцієнт дифузії водню у міді складає $D_{Cu} = 4{,}34\cdot 10^{-10}$ м^2/с, у залізі – $D_{Fe} = 1{,}8\times 10^{-11}$ м^2/с [13]. Прийнято, що характерна товщина прошарків заліза $h_{Fe} = 0{,}01$, а об'ємна частка заліза у міді становить $v_{Fe} = 0{,}2$, ступінь вільності арксинус-розподілу $\alpha = 0{,}5$, а $c_*/J_* = 0{,}1$.

На рис. 3 проілюстровано розподіли усереднених потоків водню $\langle J_H(\varsigma,\tau)\rangle/J_*$ у композиті мідь-залізо залежно від відношення початкової концентрації водню в тілі до його потоку на нижній границі в початковий момент часу $c_*/J_* = 0{,}01;\ 0{,}1;\ 0{,}2;\ 0{,}4$ (криві 1-4) для $\tau = 0{,}1$. На рис. 4 наведено залежність функції потоку від різних значень товщини прошарків заліза $h_{Fe} =$ 0,002; 0,005; 0,01; 0,1 (криві 1-4) в момент часу $\tau = 0{,}1$. Штриховою

лінією позначено відповідні потоки в однорідному шарі міді.

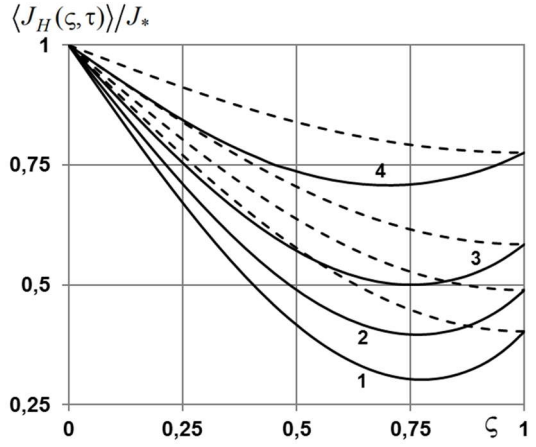


Рис. 3. Розподіли потоків водню в композиті $Cu-Fe$ для різних значень c_*/J_*

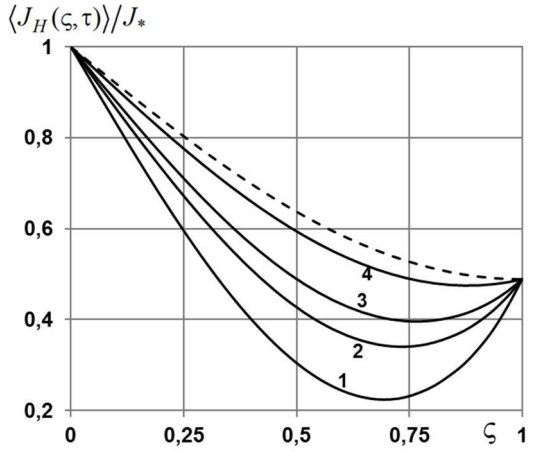


Рис. 4. Розподіли потоків водню в композиті $Cu-Fe$ для різних значень товщини залізних прошарків

Показано, що зі зростанням початкової концентрації водню у композиті c_* усереднені потоки H збільшуються в усьому тілі (рис. 3), при цьому чим більше значення відношення c_*/J_*, тим менше потоки у шаруватому композиті відрізняються від потоків у однорідному тілі. При збільшенні характерної товщини прошарків заліза усереднені потоки домішкової речовини в композиті збільшуються (рис. 4), тоді як зі зростанням об'ємної частки включень Fe за сталого значення h_{Fe} потоки водню зменшуються.

IV. Висновки

На основі підходу, відповідно до якого крайові задачі масоперенесення формулюються безпосередньо для функції дифузійного потоку, досліджено потоки домішкової речовини у двофазній випадково неоднорідній смузі з найімовірнішим розташуванням включень біля границь тіла. Побудовано інтегро-диференціальне рівняння, еквівалентне вихідній крайовій задачі. Його розв'язок знайдено методом послідовних наближень у вигляді абсолютно і рівномірно збіжного ряду Неймана. Проведено усереднення одержаного розв'язку за ансамблем конфігурацій фаз з функцією розподілу арксинуса за ненульової сталої концентрації частинок домішки в початковий момент часу.

Розглянуто модельний випадок міграції атомів водню у шаруватому композитному матеріалі мідь-залізо, з областями найбільш ймовірного розташування включень заліза біля границь тіла, та встановлено закономірності розподілів усередненого потоку домішкової речовини.

Одержана розрахункова формула для усередненого за ансамблем конфігурацій фаз дифузійного потоку за ненульової сталої початкової концентрації домішки в тілі може бути використана при моделюванні дифузійних процесів у двофазних тілах випадкової структури, розташування неоднорідностей яких можна описати за допомогою ймовірнісного арксинус-розподілу. Наприклад при комп'ютерній симуляції шаруватих покриттів з метою підбору оптимальних технологічних параметрів, прогнозування їх надійності, стійкості та інших властивостей.

Література References

[1] В. Копань, *Композиційні матеріали*, Київ, Унів. в-во «Пульсари», 2004.

[2] M. Balzannikov, A. Mikhasek, "The Use of Modified Composite Materials in Building Hydraulic Engineering Structures", *Procedia Engineering,* vol. 91, pp. 183-187, 2014.

[3] X. Sauvage, F. Wetscher, P. Pareige, "Mechanical alloying of Cu and Fe induced by severe plastic deformation of a Cu–Fe composit", *Acta Materialia*, vol. 53, no. 7, pp. 2127-2135, 2005.

[4] Д. Шевченко. (2011) *Дамаск из вторичных материалов* [Online]. Доступ: http://www.knives.com.ua

[5] C. Bergins, S. Crone, K. Strauss, "Multiphase Flow in Porous Media with Phase Change. Part II: Analytical Solutions and Experimental Verification for Constant Pressure Stream Injection", *Transport in Porous Media*, vol. 60, pp. 275-300, 2005.

[6] T. Shulenberg, U. Muller, "An improved model for two-phase flow through beds of coarse particles", *Int. J. Multiphase Flow,* vol. 13, no. 1, pp. 87–97, 1987.

[7] J. Keller, "Flow in random porous media", *Transport in Porous Media*, vol. 43, pp. 395-406, 2001.

[8] О. Ю. Чернуха, Ю. І. Білущак, А. Є. Чучвара, *Моделювання дифузійних процесів у стохастично неоднорідних шаруватих структурах*, Львів, Растр-7, 2016.

[9] В. С. Королюк, Н. И. Портенко, А. В. Скороход, А. Ф. Турбин, *Справочник по теории вероятности и математической статистике,* Москва, Наука, 1985.

[10] Y. Chaplya, O. Chernukha, A. Davydok, "Mathematical modeling of random diffusion flows in two-phase multilayered stochastically nonhomogeneous bodies", *Task Quarterly,* vol. 19, no. 3, pp. 297-320, 2015.

[11] С. Рытов, Ю. Кравцов, В. Татарский, *Введение в статистическую радиофизику. Ч. II. Случайные поля,* Москва, Наука, 1978.

[12] O. Yu. Chernukha, V. E. Goncharuk, A. E. Davydok, "Mathematical modeling of the processes of thermodiffusion of the decaying substance in a stochastically inhomogeneous layered strip", *J. Math. Sci.*, vol. 217, no. 3, pp. 312–329, 2016.

[13] Л. Лариков, В. Исачев, *Диффузия в металлах и сплавах. Структура и свойства металлов и сплавов,* Київ, Наук. думка, 1990.

Parametric Analysis of Input Circles of Measuring Transducers Impedance of Potentiometric Type

G. Barylo, R. Holyaka, V. Virt, F. Vezyr
Department of Electronic Devices
Lviv Polytechnic National University
Lviv, Ukraine
skb_mp@ukr.net

Параметричний Аналіз Вхідних Кіл Вимірювальних Перетворювачів Імпедансу Потенціометричного Типу

Г.І. Барило, Р.Л. Голяка, В.В. Вірт, Ф.Ф. Везир
кафедра електронних приладів
Національний університет «Львівська політехніка»
м. Львів, Україна
skb_mp@ukr.net

***Анотація*—В роботі розглянуто особливості параметричного аналізу вхідних кіл вимірювальних перетворювачів імпедансних потенціометричного типу, на основі SPICE – моделювання. В процесі моделювання використано потенціометричного метод вимірювань який грунтується на основі визначення часової залежності миттєвого значення струму при заданій напрузі модуляції. При низьких значеннях ширини смуги частот операційних підсилювачів частотні характеристики вимірювального перетворювача імпедансу проявляють резонансні процеси, які, в свою чергу, призводить до певної втрати стабільності кола від'ємного зворотного зв'язку. Ці виявлені закономірності необхідно враховуються при виборі елементної бази та схемотехнічній оптимізації розроблюваних вимірювальних перетворювачів імпедансу.**

***Abstract*—In this paper we examine the peculiarities of parametric analysis of input circles of measuring transducers impedance of potentiometric type based on SPICE - simulation. This simulation uses potentiometric measurement method which is based on the determination of time depending on the instantaneous value of the current at a given voltage modulation. The frequency response of measuring transducers impedance has resonance process, which in turn leads to the loss of stability of a circle negative feedback, at low value of bandwidth of operational amplifiers. These identified rules should be taken into account at a choice of components and further optimize the circuit of measuring transducers impedance.**

***Ключові слова*—параметричний аналіз, вхідні кола, вимірювальний перетворювач імпедансу, потенціометричного питу, імпедансана спектроскопія, SPICE модель, діаграма Найквіста.**

***Keywords*—parametric analysis, input circles, measuring transducers impedance, potentiometric type, impedance spectroscopy, SPICE model, Nyquist plot.**

I. INTRODUCTION

The field of information and computer technology, especially with the development of a new direction - the Internet of Things (IoT) or Physical World Internet [1], based on a new generation of microelectronic sensor devices is in constant progress. Sensor devices provide various measurement parameters - temperature, humidity, lighting substances, environmental pollution, these parameters are performing as the interface between the physical world and the Internet [2, 3].

A significant segment of modern sensory devices are based on impedance spectroscopy methods, informative parameters which are frequency dependent of active and reactive components of electrical or electro-chemical impedance [4, 5]. Specialized transducers provide formation of informative parameters of impedance spectroscopy sensor devices [6, 7]. These measuring transducers are based on two methods - galvanostatic and potentiostatic. In the galvanostatic method, informative value of measured impedance is time-dependent instantaneous voltage $V_Z(t)$ at the investigated two-terminal test at predetermined modulation current through it. The informative value of the potentiostatic method is time-dependent instantaneous value $I_Z(t)$ current modulation specified voltage. During the result of transformation, both

methods of measuring have two components of the impedance of the research object: - its active resistance (Re$\dot{Z}$) and its reactive resistance (Im$\dot{Z}$).Using components of impedance an impedance hodograph in the complex area was built, it is also called Nyquist plot. A form of the Nyquist plot allows you to define the structure and parameters of the two-terminal [8].

The actual problem that arises in the development process of sensor devices based on methods of impedance spectroscopy is analyzing the impact of the characteristics components for precision measuring transducers impedance operation. This work is directed to the problem of parametric analysis input circles of impedance measuring transducers of potentiometric type. The basis of these research is schematic simulation SPICE [9] using real mathematical functions Re and Im imaginary signal components. Such specialized functions of impedance analysis can be used in modern versions of software packages circuit simulation, in particular - package MicroCAP [10].

II. Block Diagram of Potentiostatic Type Impedance Measuring

In potentiostatic method of impedance measuring the main value is time-dependent instantaneous value of the current at a given modulation voltage. The main units of measurement transducers are (Fig. 1): VS – Voltage Source, with controlled frequency; CTA – Current Transducer and Amplifier; SD – Synchronous Detector; INT – Integrator; LFA – Low Frequency Amplifier; ADC – Analog-to-Digit Converter; MC – Microcontroller, which manages the device and the subsequent conversion of digital signals – CT and DS – Control and Digit Signal.

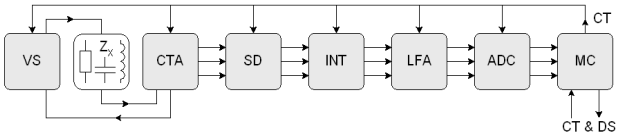


Fig.1. Generalized block diagram of potentiostatic type impedance measuring

The main criteria for the choice of components of impedance measuring is the compliance with the requirements of microelectronic devices to minimize the size while still providing high efficiency, functioning with galvanic low-voltage power sources and the implementation of low cost modes of operation by rail-to-rail.

The main approaches to parametric analysis of input circles impedance measuring transducers potentiostatic type based on operational amplifiers are described in this work.

III. Methods of Parametric Analysis

Precision of impedance conversion measuring will be measured by qualitative and quantifiable value distortion Nyquist plot. Such research is advisable to carry out on an example elementary RC two-terminal with a characteristic frequency f_0, which is on the verge of frequency measurement range. Active Z_{RE} and reactive Z_{IM} impedance components are equated $Z_{RE}(f_0) = Z_{IM}(f_0)$ on the characteristic frequency. The value f0 is expedient to set tenfold smaller than the gain bandwidth (GBW) of the operational amplifier.

For example, consider a modern CMOS (Complementary Metal-Oxide-Semiconductor) an operational amplifier of general purpose AD8541/2/4 (Analog Devices), which meets the criteria of Rail-to-Rail (low-voltage devices with full-scale signal - from the negative to the positive supply voltage) and minimum power consumption.

Thus, when GBW = 1MHz (such as is the case in the AD8541/2/4) and research is carried out on RC link with the characteristic frequency $f_0 = 1/(2\pi RC) = 0.1$ MHz, we obtain the frequency dependence of active Re and reactive Im impedance components and the Nyquist plot shown in Fig. 2.

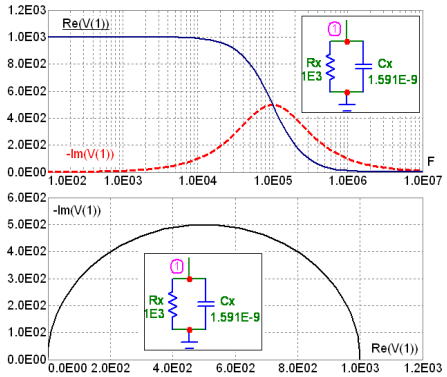


Fig.2. Frequency dependence of active Re and reactive Im impedance components and Nyquist plot of the RC circle with the characteristic frequency = 0.1 MHz

The research influence of amplifier cascades in characteristic impedance synthesizing SPICE equivalent circuit (Fig. 3) involve a source of sinusoidal voltage Vi, voltage controlled current source G1 with the transmission coefficient K_I=1, RxCx link with the characteristic frequency f0=0,1 MHz (Rx = 1 KΩ, Cx = 1,591 nF), isolating amplifier X0 with idealized parameters and two researched cascades at the operational amplifier X1, X2. The first amplifier is non-inverting with a voltage gain of $K_V = 1+R2/R1$, and the second is inverting with K_V = -R4/R3. Isolating amplifier X0 is a supplementary and has two functions: first, to prevent parasitic shunting RxCx link researched amplifier cascades, and secondly, provides a nominal (idealized) input voltage values. The input voltage is inverse to values the gain of K_V. The value of output voltage amplifiers is constant at changing K_V that is providing correct methods of research.

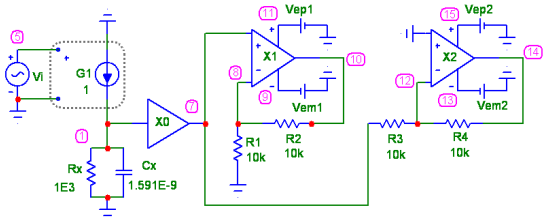


Fig.3. Scheme model research of amplifier cascades

An example result of research of parasitic influence of the amplifier cascade (in this case non-inverting amplifier with an output voltage V(10)) is shown in Fig. 4. The research was conducted at five values of gain K_V = 1, 2, 3, 4, 5, which set the

appropriate value of resistor R_2 circle non-inverting amplifier feedback: R_2 = 0, 10, 20, 30, 40 KΩ. Simultaneously with the change of values of the resistor by transmission gain of isolating amplifier X0 was simultaneously changed - X0.E1 = 1, 1/2, 1/3, 1/4, 1/5. Thus, the nominal value of informative signals remained constant.

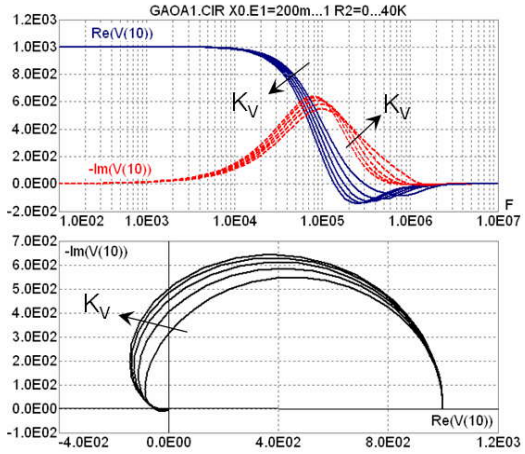


Fig.4. The drift impedance component and Nyquist plot

at K_V = 1,2,3,4,5

The results show that the above values of GBW = 1 MHz, f_0 = 0,1 MHz present a fairly significant drift of the active Z_{RE} (informative signal - Re(V(10)) and the reactive Z_{IM} (informative signal - Im(V(10)) components of impedance. Nyquist plot undergoes substantial changes deformation at high frequencies in the area of negative values of the active component of impedance. Such drift and deformation have a place even at unity gain K_V = 1 (voltage follower mode). The increase KV leading to a corresponding increase in error informative voltage amplification of the measured impedance. Accordingly, in the design of impedance measuring transducers it is necessary, first, to consider these errors and, secondly, to set the gain to the lowest possible value.

Similar research was conducted for inverting amplifier. Inasmuch as qualitatively results of this research identical to the above research non-inverting amplifier and for reducing the volume of the presented material no expediency in their detailed consideration. Comparison of the inverting amplifier and the non-inverting amplifier are reasonable. Such comparison shown in Fig. 5 and Fig. 6, where, the following symbols: A_0 - signal at the input amplifiers, A+ signal on the output of non-inverting amplifier with K_V = 1, A- signal at the output of inverting amplifier with K_V = -1. The transmission gain-controlled current source G1 is $K_I = 10^{-3}$ at the amplitude voltage of controlled source V(Vi) = 1 V and the amplitude of the current I (G1) = 1 mA.

The obtained results allow the formulation of the following important rules and conclusions:

- first of all, the gain signals in the input circles of broadband transducers should be the minimum possible, and it should only a single signal.
- secondly, frequency distortion characteristics of the impedance inverting amplifier with KV = -1 compared to non-inverting amplifier at KV = 1 (a voltage) is somewhat larger.

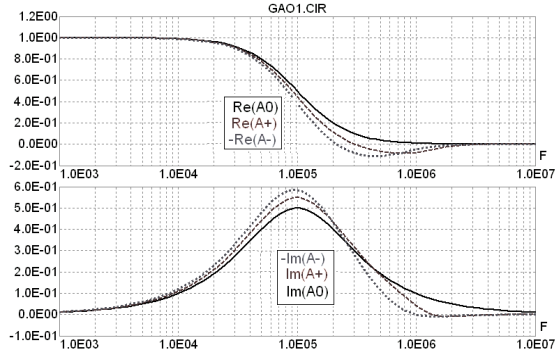


Fig.5. Frequency dependence of impedance components in circles A0, A1, A2

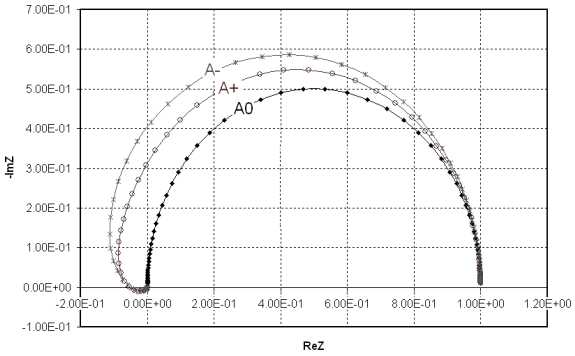
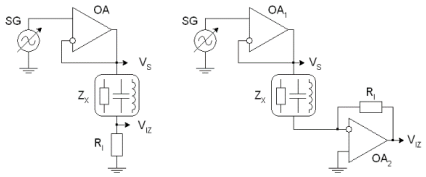


Fig.6. Nyquist plot in circles A0, A1, A2

IV. Parametric Analysis of Input Circles of Potentiometric Type

Functional scheme of input circle of measuring transducers of potentiometric type generates harmonic voltage fluctuations on the investigated two-terminal ZX two-terminal Z_X, where the current is informative on the value of its conductivity. As presented earlier, the general block diagram of potentiostatic type (Fig. 1) perform the function of the source voltage VS and manageable broadband transducer input CTA.

The elementary variant of the input circle (Fig. 7, a) contains only: controlled voltage source SG, the investigated two-terminal ZX and supplementary resistor RI, voltage drop V_{IZ} which is the informative value of measuring conversion.



a) b)

Fig.7. Schemes of input circles of potentiometric type

However, one should note the following features. Firstly, it is necessary to use a voltage follower (OA) to minimize the shunting influence of the investigated two-terminal on the controlled voltage source SG. Secondly, voltage drop on the resistor R_I is backwards function to the measuring impedance (conductive two-terminal) and this must be considered in the model researches (as shown in Fig. 8, informative signals

components of impedance is the value of Re(1/V) and Im(1/V). Thirdly, voltage drop should be minimized on the resistor R_I. The transimpedance transducer at the operational amplifier OA2 (Fig. 7, b) performs these tasks, that just also is aforementioned broadband transducer CTA of the input current to the output voltage (Fig. 1). The operational amplifier inverts the phase of the input current and informative signals change their sign: Re (-1/V) and Im (-1/V).

Scheme model research of the input circles of the potentiometric type shown in Fig. 9.A family of impedance characteristics shown in Fig. 10.

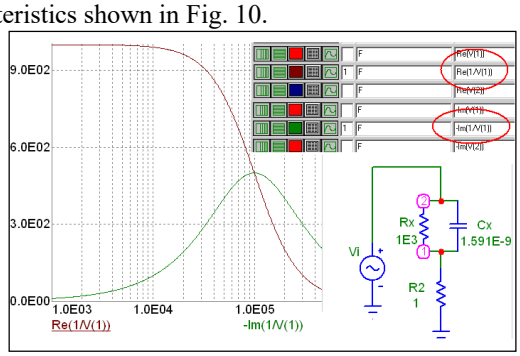


Fig.8. Demonstration process of forming informative signal

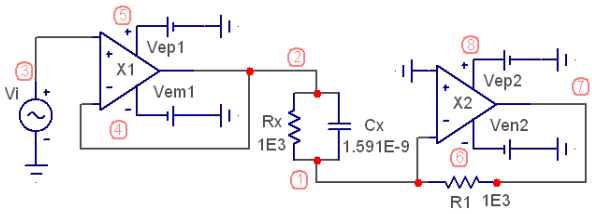


Fig.9. Scheme model research of input circles of potentiometric type

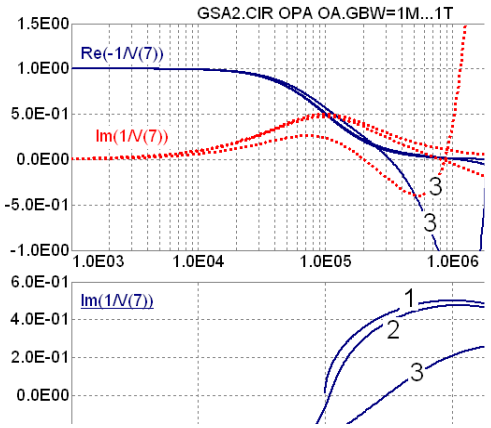


Fig.10. The family of impedance characteristics of input circles of potentiometric type at GBW = 1E12 (1), *1E7 (2). 1E6 (3)*

The frequency response of measuring transducers impedance has resonance process, which in turn leads to the loss of stability of a circle negative feedback, at low value of bandwidth of operational amplifiers. Thus, Nyquist plots not only deformed but also move into the region inductive reactance. These identified rules should be taken into account at a choice of components and further optimize the circuit of measuring transducers impedance.

V. CONCLUSION

The proposed new approaches of parametric analysis of measuring transducers impedance are using SPICE models. The precision of measuring conversion impedance is assessed by qualitative and quantifiable value distortion Nyquist plot. It is shown that parametric analysis is appropriate to use as the research object RC two-terminal with a characteristic frequency f_0, which is on the verge of frequency measurement range. Active Z_{RE} and reactive Z_{IM} impedance components are equated $Z_{RE}(f_0) = Z_{IM}(f_0)$) on the characteristic frequency.

The obtained results allow the formulation of the following important rules and conclusions: first of all, the gain signals in the input circles of broadband transducers should be the minimum possible, and the best single; secondly, frequency distortion characteristics of the impedance inverting amplifier with $K_V = -1$ compared to non-inverting amplifier at $K_V = 1$ (a voltage) is somewhat larger. These identified rules should be taken into account at a choice of components and further optimize the circuit of measuring transducers impedance.

REFERENCES

[1] Alessandro Bassi, Martin Bauer, Martin Fiedler, Thorsten Kramp, Rob van Kranenburg, Sebastian Lange, Stefan Meissner, Enabling Things to Talk: Designing IoT solutions with the IoT Architectural Reference Model // Springer, 2013. 352 P.

[2] Arne Bröring, Stefan Schmid, Corina-Kim Schindhelm, Abdelmajid Khelil, Sebastian Käbisch, Denis Kramer, Danh Le Phuoc, Jelena Mitic, Darko Anicic, Ernest Teniente, Enabling IoT Ecosystems through Platform Interoperability // IEEE Software Year: 2017, Volume: 34, Issue: 1 Pages: 54 - 61,

[3] Z. Hotra, R. Holyaka, T. Marusenkova, J. Potencki, Signal transducers of capacitive microelectronic sensors // Electronika. Rzeszow. Poland, 2010, № 8, P.129 – 132.

[4] An Audio Jack-Based Electrochemical Impedance Spectroscopy Sensor for Point-of-Care Diagnostics / Haowei Jiang; Alexander Sun; Alagarswamy G. Venkatesh; Drew A. Hall / IEEE Sensors Journal /Year: 2017, Volume: 17, Issue: 3 / Pages: 589 - 597, DOI: 10.1109/JSEN.2016.2634530

[5] Measurement of complex dielectric material properties of ice using electrical impedance spectroscopy / M. Flatscher; M. Neumayer; T. Bretterklieber; B. Schweighofer / 2016 IEEE SENSORS / Year: 2016 Pages: 1 - 3, DOI: 10.1109/ICSENS.2016.7808533

[6] Mohammad Takhti; Yueh-Ching Teng; Kofi Odame. A high frequency read-out channel for bio-impedance measurement // IEEE International Symposium on Circuits and Systems (ISCAS) 2016 Pages: 1514 - 1517,

[7] Seongheon Shin; Soon-Jae Kweon; Jeong-Ho Park; Yong-Chang Choi; Hyung-Joun Yoo. An efficient, wide range time-to-digital converter using cascaded time-interpolation stages for electrical impedance spectroscopy // IEEE Asia Pacific Conference on Circuits and Systems (APCCAS) 2016 Pages: 425 - 428

[8] І. І. Григорчак, Г. В. Понеділок, Імпедансна спектроскопія: навч. посібник. – Львів. Видавництво Львівської політехніки, 2011. – 352 с.

[9] Steven Sandler, Charles Hymowitz, SPICE Circuit Handbook // The McGraw Hill. 2006. 326 p.

[10] MICRO-CAP. Electronic Circuit Analysis Program. Spectrum Software. – 2014. – 8 p. [Online]. Available: http://www.spectrum-soft.com.

Математичне Моделювання Процесу Графітування Вуглецевих Виробів

Олексій Жученко, Марія Волощук
кафедра «Автоматизації хімічних виробництв»
НТУУ «КПІ імені Ігоря Сікорського»
Київ, Україна
azhuch@ukr.net, masha-voloshchuk@i.ua

Mathematical Modeling of Graphitization in Carbon Electrodes Production

Oleksii Zhuchenko, Mariia Voloshchuk
dept. of Automatization of chemical production
NTUU “Igor Sicorsky Kyiv Politechnic Institute”
Kyiv, Ukraine
azhuch@ukr.net, masha-voloshchuk@i.ua

***Анотація*—В роботі проведений аналіз існуючих математичних моделей процесу і економічний вплив, оптимізацію технологічного процесу, систему керування процесом графітації в печах Ачесона. Обране програмне забезпечення для вирішення поставленої задачі. Обрана найкраща структура спрощеної математичної моделі процесу графітації, яка може бути використана у подальших дослідженнях, що дозволить в найкоротші терміни і з мінімальними матеріальними витратами, виконати оцінку механічного і енергетичного стану печей графітації при різноманітних наперед завданих характеристиках технологічного регламенту, та забезпечить у синтезованій на її основі системі керування ресурсоенергозбереження і зменшення техногенного впливу на оточуюче середовище. Основні результати роботи можуть бути використані у подальших дослідженнях для оцінки механічного і енергетичного стану печей графітації, а також для подальшої розробки на їх основі системи керування процесом графітації необхідного для ресурсоенергозбереження і зменшення техногенного впливу на оточуюче середовище.**

***Abstract*—We analyzed the existing mathematical models and economic side, optimization of technological modes, system of management of process furnaces graphitization in Acheson. Favorites software to solve the problem of the listing application. Selected mathematical model of the process graphitization that can be used in further research, which will in the shortest possible time and with minimal material costs, its estimates of mechanical and power status furnaces graphitization at various pre caused by the characteristics of technological requirements and provide the synthesized its resource conservation and energy based management system and reduction of anthropogenic impact on the environment. The main results can be used in further research for further development on their basis of process control required to graphitization resource conservation and energy and reducing anthropogenic impact on the environment.**

Ключові слова—графіт, графітація, вуглецеві вироби, графітувальна піч Ачесона, математична модель.

Keywords—graphite, graphitization, carbon products, abundant graphite furnace Acheson furnace, mathematical model.

I. Вступ

У сучасній промисловості виробництво вуглецевих виробів є важливим процесом, оскільки продукція даного виробництва широко використовується в різноманітних галузях промисловості, які нерозривно пов'язані з необхідністю використання електротермічних процесів. Зокрема, до таких виробництв відносяться підприємства чорної та кольорової металургії, машинобудування, хімічної промисловості та інші.

Виробництво вуглецевих виробів є досить ресурсо- та енергозатрат ним. Окрім того, існує необхідність суворого дотримання підтримки певного рівня великої кількості технологічних змінних. Саме тому актуальною є задача введення оптимальних режимів роботи на ключових етапах виробництва з метою зменшення затрат та забезпечення високої якості продукції.

Одним з визначальних технологічних процесів виробництва вуглецевої продукції є процес графітування. Інтегральним показником якості графітованої продукції є ступінь графітації [1]. Тому задача системи керування є

забезпечення максимального ступеню графітації. При цьому на заключну стадію – процес графітування – припадає до 80 % загальних енерговитрат [2], що становить 8…10 МВт-годин на 1 тонну продукції [3]. Тому задача підвищення ефективності процесу графітування у виробництві вуглецевих виробів є актуальною науково-технічною задачею особливо в сучасних умовах постійного зростання вартості енергоносіїв.

Ще одним чинником перевитрат енергоресурсів є неточне визначення тривалості кампанії графітування, яка для гарантованого забезпечення нормативної якості готової продукції триває довше, ніж потрібно.

II. Аналіз існуючих математичних моделей

Знамеровський В. Ю. у своїй роботі [3] для розрахунку розподілу температурних полів в печі Ачесона запропонував використовувати методику чисельного рішення, засновану на методі елементарних теплових балансів з використанням явної і неявної схем апроксимації. В цих роботах піч графітації розглядається як єдина цілісна система, що містить: джерело електроживлення, коротку мережу, вироби, що підлягають графітації, допоміжні шихтові матеріали і футеровку печі. В запропонованій чисельній моделі розрахунковим елементом є вертикальний напівряд, що складається із заготовок, кернової пересипки, теплоізоляції і кладки печі. З метою зменшити похибку розрахунків теплових і електричних параметрів процесу графітації, автором була запропонована оригінальна методика розрахунків, що базувалася на тому, що процес енергообміну між елементарними об'ємами розглядалася з врахуванням одночасної зміни їх температури від дії внутрішнього джерела теплоти. Тепло- і електрофізичні характеристики матеріалу обирались в залежності від середнього значення температури від дії внутрішнього джерела теплоти, що дозволило врахувати вплив останнього на кожному часовому кроці [2].

В чисельній моделі, запропонованій Знамеровським В. Ю., внутрішнє джерело теплоти в елементах печі враховувало лише джоулеві теплоту. З метою прискорення розрахунку застосовувалося спрощення геометрії керну, круглий переріз заготовок замінювався на прямокутні рівновеликі площі.

Аналізуючи існуючі промислові системи керування процесом графітування у печах Ачесона, можна зробити висновок про те, що усі вони здебільшого належать до категорії людино-машинних (автоматизованих) і найчастіше реалізують методи керування за електротехнічними параметрами [2], за температурою футеровки, за досягнутою температурою керна, за градієнтом температур у поперечному перерізі заготовок тощо. А також, можна сказати, що керування даним процесом ведеться за наперед заданою програмою зміни потужності (напруги) електричного струму засобів електроживлення. Така програма формується як усереднена функція на основі аналізу результатів багаторазового проведення процесу графітування. Такий підхід безумовно має недоліки, основним з яких є та обставина, що кожна окремо взята кампанія графітування у загальному випадку відрізняється від узагальненого варіанту, наприклад, складом вуглецевих виробів, що обробляються, або складом та формою пересипки тощо. Крім того, у процесі графітування трапляються (особливо останнім часом) раптові відключення електроживлення, які ніяким чином не можуть бути передбачені наперед сформованою програмою.

Коржик М. В. [2] для оцінки теплоелектричного стану печі графітації запропонував спрощену модель електричного поля керна і адаптивну модель температурного поля печі графітації, яка враховує витрати енергії на випаровування вологи і функціонує в умовах неповної інформації про вологість і фізичні властивості матеріалів завантаження.

До загальних недоліків розглянутих вище чисельних моделей можна віднести наступні:

- не враховується залежність фізичних властивостей сипких матеріалів від тиску;
- немає врахування масо- і теплообміну при випаровуванні і конденсації вологи сипких матеріалів, а також в деяких, вище розглянутих, чисельних моделях не враховується навіть вологовміст;
- не враховується теплота хімічних реакцій, що відбуваються в керні, за винятком врахування теплоти, пов'язаної зі звільненням сірки від зв'язків з вуглецем всередині коксових частинок;
- в багатьох чисельних моделях застосовується спрощення геометрії керна, а врахування торцевих втрат виконується шляхом задання усереднених коефіцієнтів форми, що зменшують точність розрахунку;
- одномірні і двовимірні моделі, не дозволяються оцінити радіальні перепади температур по горизонтальному перерізу заготовок, а усереднення властивостей рядів керну (заготовки і пересипка) збільшують неточність результатів розрахунку [2].

З вищесказаного слідує наступне, що в розглянутих чисельних моделях враховуються досить важливі особливості процесу графітації в печах опосередкованого нагріву, а їх застосування обмежено областю, в якій виконувалася перевірка їх адекватності.

До основних показників процесу графітування відносять дві групи показників: технологічні й режимні. Технологічні показники включають схему укладання керна, типи та розміри печей, перетин та довжину виробів, що графітуються, пересип очні та теплоізоляційні матеріали. До режимних параметрів відносяться температурні та енергетичні, які є тісно пов’язані між собою.

Основні характеристики температурного режиму процесу графітування включають: максимальну температуру у керні; час витримки керна при максимальній температурі; темп нагрівання заготовок; перепад температури у заготовках та ін. Енергетичний режим є визначальним фактором температурного режиму процесу графітування, який значно залежить від

властивостей кернової пересипки і геометрії керна та характеризується регламентом підводу потужності у піч, терміном кампанії графітування питомими витратами електроенергії.

Процес графітування виробів є надзвичайно енергоємним. Теоретичні витрати електроенергії на графітування електродних заготовок, тобто на їх нагрів до 2750°С, дорівнюють 1800 кВт·год/т, а фактичні – на печах Ачесона становлять 4000-5000 кВт·год/т (до 6500 кВт·год/т при графітуванні особливо якісних виробів)[17]. Таким чином понад 60% становлять теплові втрати у довкілля[18]. На печах Кастнера ПВЕ значно менші і становлять 3000 – 3500 кВт·год/т. Тривалість процесу графітування в печах Ачесона – від 36 до 120 годин, а в печах Кастнера – від 10 до 30 годин.

Таким чином, виходячи з ПВЕ печі Ачесона, витікає необхідність зменшення тепловитрат при графітування, наприклад, за рахунок зменшення тривалості процесу. Проте на шляху до прискорення процесу можуть стояти такі дві суттєві перешкоди: 1) збільшення ймовірності розтріскування вуглецевих заготовок при прискореному нагріванні; 2) недостатня потужність пічних трансформаторів.

III. Постановка завдання дослідження

Одним із важливих чинників, які впливають на показники графітування є операція завантаження. Робочий об’єм печі повинен бути заповнений керном таким чином, щоб його електричний опір був однаковим по усьому перетину. Це забезпечить рівномірність розподілу струму по всьому перетину і, відповідне, рівномірне нагрівання керна. Нерівномірна укладка керна призводить до значних перекосів у розподілі струму особливо при високих температурах.

Суттєвий вплив на показники графітування надають допоміжні матеріали – пересипка і теплоізоляційна шихта. Пересипка визначає теплоелектричний режим печі графітування. Дослідження температурного поля математичної моделі показали, що розподіл джоулевої теплоти по перетину керна визначається характером кривої температурної залежності ПВО пересипки та якістю теплоізоляції. Наприклад, для матеріалів зі значною залежністю ПВО від температури (антрацит, сирий кокс) перепад температури між центром і периферією керна збільшується.

Застосування кернової пересипки із графітованого коксу дозволило отримати доволі рівномірне температурне поле у період усієї кампанії графітування. Вплив зазору між заготовками та ПЕО пересипки і на якість кінцевих виробів досліджувався у. Було встановлено, що зменшення ПВО пересипки призводить до більш рівномірного розподілу джоулевої теплоти між заготовками і пересипкою та дозволяє більш рівномірно нагрівати заготовки [8].

Наразі у промисловості відсутні універсальні показники якості графітованих виробів. Тому вимоги до якості визначаються для кожного типу продукції окремо, виходячи з умов її експлуатації безпосередньо у промислових агрегатах. В залежності від вимог проводиться оцінка за такими показниками:

фізичні властивості: істинна і позірна або об’ємна густина та поруватість, питомий електричний опір, теплопровідність, масова теплоємність, термічний коефіцієнт лінійного розширення;

механічні властивості: міцність, модуль пружності та зсуву, коефіцієнт тертя;

хімічні властивості: окислення, участь у реакціях утворення карбідів;

вміст домішок: вміст заліза, сірки та ін.

Названі вище обставини викликають необхідність з метою підвищення ефективності процесу графітування вуглецевих виробів впровадження сучасної системи оптимального керування ним на основі аналізу поточного стану технологічного процесу.

IV. Математичне моделювання процесу графітування

Як вже зазначалося, вуглеграфітове виробництво є дуже енергоємним. Тому для оптимізації його роботи доцільно використати рентабельність даного виробництва як показник ефективності:

$$R = \frac{Q}{Cek} \qquad (1)$$

де $C_{ЕК}$ – собівартість експлуатаційних витрат, Q – прибуток від реалізації продукції.

Використання даного критерію (1) дозволяє підвищити чутливість критерію оптимальності по відношенню до технологічних змінних, реалізація яких на об’єкті керування забезпечить функціонування останнього в оптимальному технологічному режимі [2].

При розв’язанні задачі оптимізації технологічних режимів виробництва потрібно враховувати обмеження на змінні, що оптимізуються. На виробництві вуглеграфітових виробів обмеження можуть бути сформульовані по різному.

Так, обмеження можуть бути виражені щодо кількості кожного виду продукції, що виробляється:

$$F_i \le F_i^{\max}, \qquad (2)$$

де F_i^{max} – максимальна продуктивність виробництва на i-му виду продукції.

Обмеження, що інтегрує у собі всі види продукції, може бути сформульоване таким чином:

$$\sum_{i=1}^{N} g_i F_i \le G_{\max}, \qquad (3)$$

де g_i – вага i-го виду продукції; G_{max} – максимальна потужність виробництва, виражена в одиницях ваги.

Таким чином, задача оптимізації технологічних режимів виробництва математично включає у себе критерій оптимальності (1) разом з обмеженнями (2) і (3). Дана задача являє собою задачу цілочисельного нелінійного програмування, яку можна розв'язати відомими методами, використовуючи стандартне програмне забезпечення [2] та може бути розв'язана у 2 етапи з використанням методів та програмних засобів лінійного програмування.

На процес графітування впливає цілий ряд факторів, зокрема, індивідуальні властивості вуглецевих матеріалів, газове середовище, передісторія термічної обробки, тривалість процесу, температура тощо. Вплив декотрих із них на процес графітування наразі до кінця не досліджений. Однак, як показують численні дослідження [2, 3], головним чинником, який визначає якість графітування вуглецевої речовини є кінцева температура. Причому для отримання якісного штучного графіту кінцева температура обробки повинна бути не нижче 2200 – 2800°*C*. Вимірювати такі температури в автоматичному режимі сучасними технічними засобами неможливо, а це означає, що нема перспективи побудувати класичну систему автоматичного керування на основі, скажімо, ПІД-регуляторів.

Таким чином, метою даної статті є створення нової системи керування процесом графітування вуглецевих виробів, яка забезпечить заощадження енергоресурсів при збереженні потрібної якості готової продукції, для чого функціонування системи, на відміну від існуючих, повинно відбуватися не за наперед заданою програмою, а з урахуванням поточного стану об'єкта керування.

Задача системи керування процесом графітування вуглецевих виробів формулюється таким чином: забезпечити мінімізацію витрат енергоресурсів при збереженні потрібної якості готової продукції. Енерговитрати формуються двома складовими – потужністю електричного струму, що подається на процес, та часом кампанії графітування.

Як відомо [1 – 3], графітування вуглецевих виробів відбувається при температурах понад 2000°*C*. Тому процес графітування умовно можна розділити на дві стадії – розігрівання та безпосередньо графітування.

Для мінімізаціїї енерговитрат у режимі розігрівання треба забезпечити максимально допустиму швидкість підвищення температури до досягнення в усіх вуглецевих заготовках температури графітування [1, 2]. При цьому швидкість підвищення температури обмежується механічною міцністю продукції, яка при значних градієнтах температур може бути пошкоджена у результаті виникнення тріщин. Отже, задачею системи керування на стадії розігрівання є визначення та реалізація оптимальної функції зміни напруги електричного струму у часі.

Після досягнення потрібної температури графітування задачею системи керування є підтримання температури на заданому рівні за рахунок відповідної зміни напруги електричного струму до досягнення заданого значення ступеня графітування, який виступає як узагальнений показник готової вуглецевої продукції [2]. При цьому вирішальним моментом є тривалість процесу графітування, бо, враховуючи питомі витрати електроенергії (до 10 МВт-год. на 1 т продукції), визначення оптимального моменту зупинки подачі електроживлення є дуже важливим.

V. Висновки

Наприкінці, можна сказати, що в результаті літературного огляду за станом проблем математичного моделювання процесу графітування вуглецевих виробів в печах Ачесона виявлено, що в існуючих математичних моделях не враховуються такі важливі особливості процесу як залежність фізичних властивостей сипких матеріалів від тиску і їх вологовміст, масо- і теплообмін при випаровування і конденсації вологи, теплота хімічних реакцій, що відбуваються у керні тощо, а їх застосування обмежено областю. В якій виконувалася оцінка їх адекватності.

Таким чином, у відповідності до розглянутих стадій технологічного процесу графітування потрібно розробити системи керування двома стадіями - розігріву та графітування, задача керування процесом графітування вуглецевих виробів включає у себе критерій оптимальності та обмеження. І для розв'язання поставленої задачі керування у подальших дослідженнях потрібно розробити математичну модель процесу та метод врахування взаємозв'язаних параметричних обмежень, що і є завданням подальших досліджень.

Література References

[1] D. M. Kuznetsov, V.P. Fokyn "The process of grafitization of carbon materials. Modern research methods": Novocherkassk: YuRHTU, 2001, 131 p., ISBN 5-88998-233-8, pp. 126-131

[2] M. V. Korzhyk "Mathematical modeling and automated process control of grafitization in Acheson furnaces" : Avtoref. dys. PhD: 05.13.07; The Ministry of Education and Science of Ukraine, NTUU «KPI». – Kyiv, 2010.

[3] V. Yu. Znamerovskyy, V. V. Yashkyna " The study of energy input modes in a grafite furnace ": Promushlennaya enerhetyka, 1985, # 11, pp. 40–42.

Метод Змінних Радіусів в Задачі Розміщення Нерівних Кіл

С.В. Яковлев, О.В. Карташов, К.П. Коробчинський
кафедра інформатики
Національний аерокосмічний університет ім. М.Є. Жуковського «ХАІ»
Харків, Україна
svsyak@mail.ru

The Method of Variable Radii in the Placement Problem of Unequal Circles

S. Yakovlev, O. Kartashov, K. Korobchinskiy
Department of Computer science
National Aerospace University of N.E. Zhukovsky "HAI"
Kharkiv, Ukraine
svsyak@mail.ru

***Анотація*—Розглядається задача оптимального розміщення кіл у колі мінімального радіуса. Здійснено її формалізацію як задачі математичного програмування. Виділено комбінаторну структуру задачі за допомогою формування множини переставлень радіусів кіл, що розміщуються. На основі функціонального представлення множини переставлень формулюється еквівалентна постановка задачі, в якій радіуси кіл є незалежними змінними. Такий підхід дозволяє підвищити ефективність і поліпшити значення локальних екстремумів розв'язання вихідної задачі.**

***Abstract*—The problem of the optimal placement of circles in a circle of minimum radius is considered. Its formalization as a problem of mathematical programming is given. The problem combinatorial structure is obtained by forming a set of permutations of the radii of the placed circles. The functional representation of the set of permutations allows to get an equivalent formulation of the problem, in which the radii of the circles are independent variables. Such an approach makes it possible to increase the efficiency and improve the values of local extrema for the solution of the original problem.**

Ключові слова—задача розміщення; коло; оптимізація; множина переставлень

Keywords—packing problem; circle; optimization; permutation set

I. Вступ

Задачі оптимального розміщення кіл викликають постійний інтерес дослідників [1-5]. Це стосується як виділення спеціальних класів задач, для яких можна запропонувати як нові ефективні методи розв'язання, так і застосувати сучасні методи теорії оптимізації для розв'язання задач в досить загальній постановці. Поштовхом до розвитку цього напрямку стало створення теорії Ф-функцій Ю.Г. Стояна [6,7]. У даній роботі пропонується новий погляд на формалізацію і методи розв'язання завдань розміщення як задач математичного програмування шляхом виділення їх комбінаторної структури.

II. Постановка задачі та її математична модель

Розглянемо задачу розміщення набору кіл S_i, $i \in J_n$, заданих радіусів, де $J_n=\{1,2,\ldots,n\}$, в колі S_0 мінімального радіуса. Позначимо радіуси кругів через r_i, а координати центрів кіл – $p^i=(\mathrm{u}_i,\mathrm{v}_i)$, $i \in \{0\} \bigcup J_n$. Зафіксуємо $p^0=(0,0)$. Тоді математична постановка задачі набуде виду

$$r_0 \to \min \tag{1}$$

при обмеженнях

$$u_i^2+v_i^2 \le (r_0-r_i)^2, i \in J_n \tag{2}$$

$$(\mathrm{u}_i-u_j)^2+(\mathrm{v}_i-v_j)^2 \ge (r_i-r_j)^2, i \in J_n, j \in J_n, i<j. \tag{3}$$

Таким чином, наведена завдача є задачою математичного програмування з $2n + 1$ змінними r_0, v_i, u_i, $i \in J_n$.

III. Метод змінних радіусів

Здійснимо наступні еквівалентні перетворення задачі (1) - (3). З одного боку, будемо вважати, що радіуси r_i, $i \in J_n$, є незалежними змінними. З іншого боку, сформуємо таку систему обмежень задачі, щоб в результаті її розв'язання допустимими стали ті і тільки ті значення змінних r_i, $i \in J_n$, які збігаються з вихідними фіксованими значеннями. Такий підхід назвемо методом змінних радіусів. Зауважемо, що розгляд змінних радіусів кіл, але в іншій концептуальній постановці, розглядались в роботах [4,5].

З метою реалізації метода змінних радіусів виділимо таку комбінаторну структуру задачі. Оскільки в наведеній задачі (1)-(3) радіуси є константами, зафіксуємо $r_1^0 = r_i$, $i \in J_n$, і будемо вважати, що $r_1^0 \le r_2^0 \le ... \le r_n^0$. У відповідності з наведеними вище міркуваннями будемо розглядати r_i, $i \in J_n$, як незалежні змінні та сформуємо систему рівнянь

$$\sum_{i=1}^{n}\left(r_i - \tau\right)^k = \sum_{i=1}^{n}\left(r_i^0 - \tau\right)^k, \; k \in J_n \tag{4}$$

$$\tau = \frac{1}{n}\sum_{i=1}^{n} r_i^0 . \tag{5}$$

Система рівнянь (4) має таку властивість, що множина її розв'язків збігається з множиною всіляких переставлень з чисел $\left\{r_1^0, r_2^0, ..., r_n^0\right\}$ [6]. Таким чином, сформована задача (1) - (5) еквівалентна вихідній (при фіксованих радіусах) і є задачею умовної оптимізації з 3n + 1 змінними r_0, r_i, v_i, u_i, $i \in J_n$.

Зауважимо, що при розв'язанні задачі (1) - (5) можуть виникати обчислювальні складності, пов'язані з високими ступінями в системі рівнянь (4), що призводить до накопичення похибок обчислень в задачах підвищеного розміру. Тому становить інтерес формування функціональних обмежень в поліедрально-сферичному вигляді [8]

$$\begin{gathered}\sum_{i=1}^{n} r_i = \sum_{i=1}^{n} r_i^0 \\ \sum_{i\in W} r_i \ge \sum_{i=1}^{|W|} r_i^0, \; \forall W \subseteq J_n, \; |W| < n\end{gathered} \tag{6}$$

$$\sum_{i=1}^{n}\left(r_i - \tau\right)^2 = \sum_{i=1}^{n}\left(r_i^0 - \tau\right)^2, \tag{7}$$

де $|W|$ - потужність множини W, а τ задається виразом (5).

Зауважимо, що множина, що описується системою (6), (7), збігається з множиною всіляких переставлень з чисел $\left\{r_1^0, r_2^0, ..., r_n^0\right\}$ [8].

Важливою властивістю задачі розміщення кіл у вигляді (1) - (3), (6), (7) є той факт, що ця задача є квадратичною. Однак кількість лінійних обмежень в системі (6) велика і оцінюється порядком 2^n. Тому реалізація класичних методів нелінійної оптимізації обмежується розміром задачі. Разом з тим врахування властивостей лінійних і квадратичних функцій на комбінаторних багатогранниках дозволяють обійти труднощі, що виникають.

У задачах локальної оптимізації великої розмірності пропонується здійснювати їх декомпозицію з використанням такого підходу. Розглянемо множину. $\Pi = \left\{r_1^0, r_2^0, ..., r_n^0\right\}$. Здійснимо його розбиття на L попарно неперетинних підмножин і введемо позначення

$$\Pi = \bigcup_{k=1}^{L} \Pi^k, \quad \Pi^k = \left\{r_{m_1}, r_{m_2}, ..., r_{m_{l_k}}\right\}, \; M^k = \left\{m_1, m_{2,...,} m_{l_k}\right\},$$

$$m_1 < m_2 < ... < m_{l_k}, \; \sum_{k=1}^{L} l_k = n .$$

Для кожної множини Π^k, $k \in J_L$, будемо вважати $\lambda = l_k$ і запишемо

$$\begin{gathered}\sum_{i=1}^{\lambda} r_{m_i} = \sum_{i=1}^{\lambda} r_{m_i}^0 \\ \sum_{i\in W} r_i \ge \sum_{i=1}^{|W|} r_{m_i}^0, \; \forall W \subseteq M^k\end{gathered} \tag{8}$$

$$\sum_{i=1}^{\lambda}\left(r_{m_i} - \nu\right)^2 = \sum_{i=1}^{\lambda}\left(r_{m_i}^0 - \nu\right)^2, \tag{9}$$

$$\nu = \frac{1}{\lambda}\sum_{i=1}^{\lambda} r_{m_i}^0 .$$

Нехай отримано локальний розв'язок задачі (1) - (3) або його наближення при заданих r_i, $i \in J_n$. Цей розв'язок можна поліпшити, вибираючи його в якості початкової точки і розглядаючи r_i, $i \in J_n$, як незалежні змінні. При цьому вибір способу розбиття з подальшим формуванням обмежень виду (8), (9) задає відповідну модифікацію запропонованого підходу. Більш того, отриманий новий локальний розв'язок можна знову спробувати поліпшити, вибираючи його в якості початкової точки і формуючи нове розбиття множини $\Pi = \left\{r_1, r_2, ..., r_n\right\}$. Таким чином, перспективним видається вибір і обгрунтування способів розбиття, що забезпечують підвищення ефективності методу.

IV. Приклад застосування методу змінних радіусів

Розглянемо задачу розміщення 10 кіл, радіуси яких були вибрані за допомогою генератора випадкових чисел. Їх значення наведено в таблиці 1.

ТАБЛИЦЯ I. Радіуси кіл, що розміщуються

r_1	r_2	r_3	r_4	r_5	r_6	r_7	r_8	r_9	r_{10}
3	10	1	4	4	1	3	5	9	4

Початкові координати розміщення кіл та початковий радіус r_0, v_i, u_i, $i \in J_n$ теж були згенеровані випадково та наведені в таблиці 2.

ТАБЛИЦЯ II. Початкові координати кіл

i	u_i	v_i
1	12,30	8,96
2	-1,75	9,09
3	0,90	-17,86
4	-15,26	0,17
5	-14,61	7,14
6	17,28	-3,71
7	15,57	2,32
8	-11,72	-8,11
9	6,22	-8,16
10	-5,22	-14,34

На Рис. 1 наведено початкове розміщення кіл, що не задовольняє умовам взаємного неперетину кіл і їх розміщення всередині області.

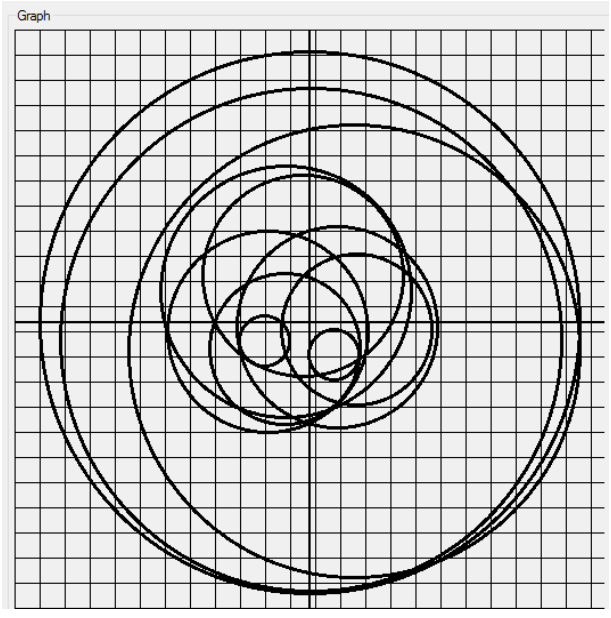


Рис. 1. Початкове розміщення кіл.

Пошук локального розв'язку задачі (1)-(3) при заданих початкових параметрах розміщення здійснювався з використанням пакета програм IPOPT (Interior Point OPTimizer, https://projects.coin-or.org/Ipopt) - бібліотеки для пошуку локальних екстремумів безперервних задач нелінійного програмування з використанням методів внутрішньої точки.Після оптимізації отримуємо точку локального мінімуму для задачі (1)-(3). Координати кругів для неї наведено в таблиці 2, а відповідне розміщення кіл навено на Рис. 2.

ТАБЛИЦЯ III. Координати кіл локального мінімуму

i	u_i	v_i
1	12,30	8,96
2	-1,75	9,09
3	0,90	-17,86
4	-15,26	0,17
5	-14,61	7,14
6	17,28	-3,71
7	15,57	2,32
8	-11,72	-8,11
9	6,22	-8,16
10	-5,22	-14,34

Оптимальний радіус зовнішнього кола – 19,54.

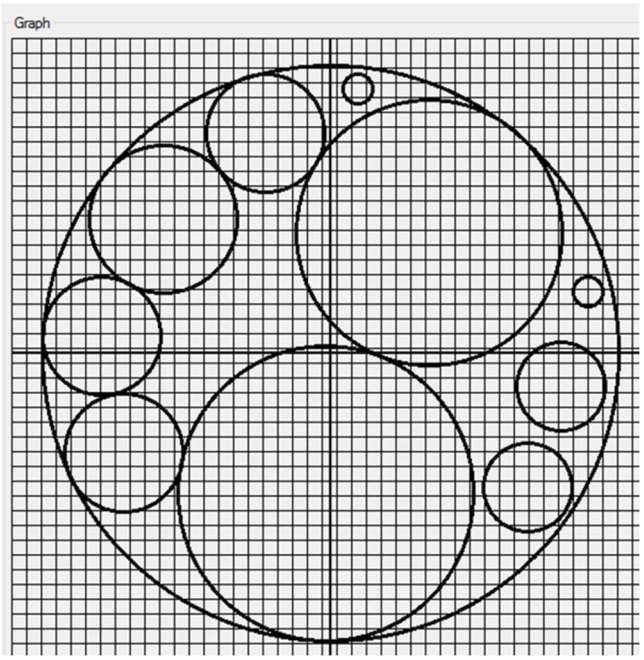


Рис. 2. Локальний мінімум

Вважаємо це розміщення початковою точкою для задачі зі змінними радіусами (1) - (3), (6), (7). Для початкових значень радіусів обираємо випадкові значення в діапазоні між найменшим і найбільшим із вихідних значень кругів, що наведено в таблиці 1. Знаходимо локальний розв'язок цієї задачі. Оптимальне розміщення наведено на Рис. 3, а відповідні оптимальні значення параметрів в таблиці 4.

Нове значення локального оптимуму для радіуса зовнішнього кола – 19,26.

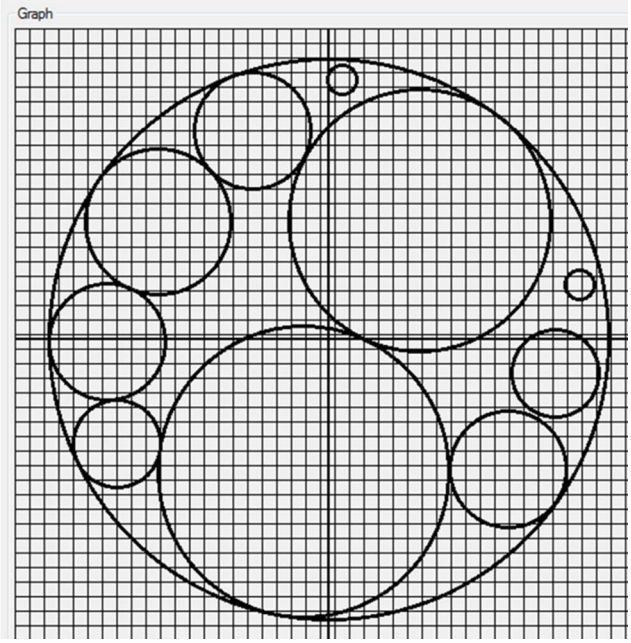


Рис. 3. Локальний мінімум для задачі (1) - (3), (6), (7).

ТАБЛИЦЯ IV. ПАРАМЕТРИ КІЛ ЛОКАЛЬНОГО МІНІМУМУ ДЛЯ ЗАДАЧІ (1) - (3), (6), (7).

i	u_i	v_i	r_i
1	12,30	8,96	4,00
2	-1,75	9,09	10,00
3	0,90	-17,86	1,00
4	-15,26	0,17	4,00
5	-14,61	7,14	3,00
6	17,28	-3,71	1,00
7	15,57	2,32	3,00
8	-11,72	-8,11	5,00
9	6,22	-8,16	9,00
10	-5,22	-14,34	4,00

Тепер в отриманій новій точці знову обираємо випадкові значення для кіл і знову шукаємо розв'язок для задачі (1) - (3), (6), (7). І так декілька раз. Результат кращого з отриманих локальних мінімумів наведено на рис. 4, а відповідні оптимальні значення параметрів в табл. 5. Значення радіуса зовнішнього кола для цієї точки – 19,01.

ТАБЛИЦЯ V. ПАРАМЕТРИ КІЛ ДЛЯ НАЙКРАЩОГО З ОТРИМАНИХ РОЗВ'ЯЗКІВ

i	u_i	v_i	r_i
1	11,83	-13,24	1,00
2	-6,02	-7,99	9,00
3	-12,21	10,36	3,00
4	-17,49	-4,31	1,00
5	5,94	-13,78	4,00
6	12,04	-7,17	5,00
7	-6,07	13,73	4,00
8	-15,94	1,02	3,00
9	6,03	6,69	10,00
10	-7,85	4,88	4,00

Зазначимо, що отриманий розв'язок є глобальним розв'язоком задачі (1)-(3).

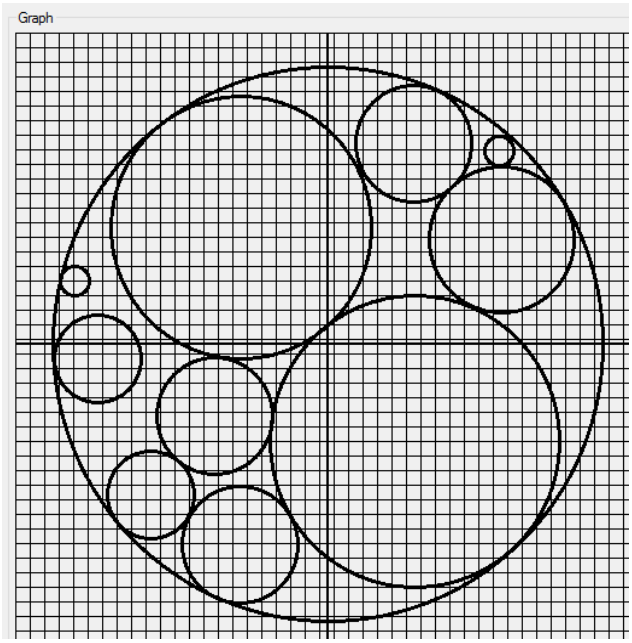


Рис. 4. Найкращий з отриманих розв'язків.

V. ВИСНОВКИ

В статті запропоновано новий погляд на формалізацію задач розміщнння нерівних кіл в колі як задачі математичного програмування. В результаті виділення комбінаторной структури задачі за допомогою додаткових змінних вдалося сформувати додаткову систему обмежень, що описують множину переставлень радіусів кіл. Це дозволяє за рахунок змінних радіусів подолати область тяжіння локальних екстремумів задачі. Таким чином, метод змінних радіусів насамперед слід розглядати як спосіб поліпшення локальних розв'язків задачі.

ЛІТЕРАТУРА REFERENCES

[1] C. Che, Y. Wang, H. Teng, Test problems for quasi-satellite packing: cylinders packing with behavior constraints and all the optimal solutions known// Optimization Online (2008).

[2] G Fasano, J. D. Pinte'r (eds.), Modeling and Optimization in Space Engineering. Series: Springer Optimization and Its Applications, Vol.73, XII, p.404 (2013)

[3] P. I. Stetsyuk, T. E. Romanova, G. Scheithauer, *On the global minimum in a balanced circular packing problem* in Optimization Letters, 2015, 10 (6), P. 1347–1360.

[4] Yu. Stoya, G. Yaskov, Packing unequal circles into a strip of minimal length with a jump algorithm,Optimization Letters, 2014, Vol.8(3), p. 949-970. DOI:10.1007/s11590-013-0646-1.

[5] G. N. Yaskov, *Packing non-equal hyperspheres into a hypersphere of minimal radius* in Пробл. Машиностроения, 2014, Т. 17, № 2, P. 48-53.

[6] Ю. Г. Стоян, Ф-функция и ее основные свойства // Докл. НАН Украины, 2001, №8, С.112-117.

[7] Ю. Г. Стоян, С. В. Яковлев, Математические модели и оптимизационные методы геометрического проектирования. — К. : Наук. думка, 1986. 268 с.

[8] О. С. Пичугина, С. В. Яковлев, Функционально-аналитические представления общего перестановочного множества // Eastern-European Journal of Enterprise Technologie, 2016, № 1, С. 101–126.

Двовимірна Математична Модель Реакції Окислення Чадного Газу на Поверхні Платинового Каталізатора

Петро Костробій, Ірина Рижа
кафедра прикладної математики
Національний університет "Львівська політехніка"
Львів, Україна
ira.saj@gmail.com

Two-Dimensional Mathematical Model for Reaction of Carbon Monoxide Oxidation on the Platinum Catalyst Surface

Petro Kostrobij, Iryna Ryzha
department of Applied Mathematics
Lviv Polytechnic National University
Lviv, Ukraine
ira.saj@gmail.com

***Анотація*—Досліджено двовимірну математичну модель оксидації чадного газу (CO) на поверхні платинового каталізатора з урахуванням процесів перебудови поверхні каталізатора та впливу температури підложки. Показано, що в двовимірному випадку область стійкості реакції окислення CO змінюється.**

***Abstract*—The two-dimensional mathematical model for carbon monoxide (CO) oxidation on the surface of Platinum (Pt) catalyst is investigated accounting for the processes of the catalyst surface reconstruction and the effect of the substrate temperature. It is shown that the stability region for reaction of CO oxidation changes in two-dimensional case.**

Ключові слова—каталітична реакція окислення; реакційно-дифузійна модель; математичне моделювання реакційно-дифузійних процесів.

Keywords—reaction of catalytic oxidation; reaction-diffusion model; mathematical modeling of reaction-diffusion processes

I. Вступ

Дослідження проблем каталітичного окислення чадного газу стало однією з найважливіших задач в сучасній науці [1]. З технологічної сторони, цей процес має життєво важливе значення для багатьох областей застосування: контроль за викидами вихлопних газів автотранспорту, очищення повітря, вуглекислотні лазери, сенсори тощо. З теоретичної сторони, це одна з найпростіших каталітичних реакцій, яка широко використовується в якості модельної системи для розуміння процесів гетерогенного каталізу. Проблема ефективного протікання таких хімічних реакцій на поверхнях металевих каталізаторів охоплює доволі широкий спектр наукових задач. Нажаль досі немає фізичних моделей, які дозволяли би водночас точно та адекватно пов'язати результати теоретичних і практичних досліджень.

Очевидно, без розуміння механізмів каталітичного окиснення чадного газу (CO) побудувати модель для кількісного опису цього процесу неможливо. Одним з механізмів каталітичного окислення CO на поверхні платинового каталізатора (Pt) є механізм Лангмюра-Гіншелвуда (LH), дослідження якого проведено Бакстером та ін. [2].

За механізмом (LH) для того, щоби реакція відбулася, необхідно, щоб обидві реагуючі речовини CO та кисень (O) спочатку адсорбувалися на поверхні каталізатора (в місцях, які називають активними щодо адсорбції), а далі між адсорбованими CO_{ads} та O_{ads} (індекс "ads" позначає адсорбовану фазу) за термодинамічно вигідних умов проходить реакція синтезу із швидким процесом десорбції продукту реакції вуглекислого газу (CO_2) з поверхні каталізатора.

Рівняння реакцій, які можуть проходити при окисленні CO на поверхні каталізатора, згідно механізму (LH) записуються у вигляді:

$$O_2 + 2* \rightarrow 2O_{ads}$$

$$CO + * \leftrightarrow CO_{ads}$$

$$CO_{ads} + O_{ads} \rightarrow CO_2 + 2*$$

Тут символ "*" позначає вільне активне місце на каталітичній поверхні.

При моделюванні реакції окислення CO в рамках рівнянь хімічної кінетики вважається, що реакція проходить лише на поверхні каталізатора, тобто взаємодії можливі лише між частинками, які адсорбовані на поверхні каталізатора. При цьому адсорбція молекули кисню відбувається в результаті її розпаду біля поверхні каталізатора на два атоми, кожен з яких самостійно адсорбується на окреме вільне активне місце [3]. Молекули CO адсорбуються на поверхню і перебувають на ній, не розпадаючись на атоми. Перетворення частинок можливе лише в адсорбованому шарі. При адсорбції молекули кисню швидко дисоціюють на атоми, а десорбція O_2 не враховується через її малу імовірність у діапазоні температур, при яких проводяться експерименти. Так як коефіцієнт дифузії адсорбованого кисню на три-чотири порядки менший, ніж коефіцієнт дифузії CO [4], то адсорбований кисень вважається нерухомим.

Важливу роль в гетерогенному каталізі відіграють процеси поверхневої перебудови атомів каталізатора. Чиста поверхня Pt(110) є реконструйованою і являє собою структуру (1×2) [5]. Під час реакції кисень і CO адсорбуються на поверхню Pt. Якщо величина поверхневого покриття CO перевищує певне критичне значення, поверхня перебудовується до об'ємної конфігурації (1×1) [6]. Коли кисень і CO прореагують, виділяється вуглекислий газ, і поверхня повертається до своєї початкової конфігурації. Такі структурні зміни впливають на швидкості інших елементарних процесів, тому їх необхідно враховувати при побудові математичної моделі реакції.

Для тонкоплівкових каталізаторів тепло, яке генерується при хімічних процесах, може динамічно змінювати температуру T каталізатора. Навіть незначні зміни можуть суттєво вплинути на динаміку окиснення CO. Тому доречно проаналізувати також умови теплової рівноваги, що дають можливість зрозуміти характерні часові і просторові масштаби температурного поля, яке може впливати на перебіг реакції.

II. Опис математичної моделі

Розглянуто модель реакції каталітичного окислення CO, що враховує процес дифузії молекул CO на поверхні платини Pt(110). Поверхня каталізатора вважається плоскою з заданою декартовою системою координат XOY. Часова еволюція покриттів CO (u) та O (v) на поверхні каталізатора визначається наступними кінетичними рівняннями [7-8]:

$$\frac{\partial u}{\partial t} = p_u \kappa_u s_u \left(1 - \left(u/u_{sat}\right)^q\right) + k_{des} u + k_r uv + D_x \frac{\partial^2 u}{\partial x^2} + D_y \frac{\partial^2 u}{\partial y^2}, \quad (1)$$

$$\frac{\partial v}{\partial t} = p_v \kappa_v s_v \left(1 - u/u_{sat} - v/v_{sat}\right)^2 + k_r uv. \quad (2)$$

Тут p_u, p_v – парціальні тиски речовин; κ_u, κ_v – частоти зіткнень молекул з поверхнею; s_u, s_v – коефіцієнти налипання; u_{sat}, v_{sat} – відповідають максимальному покриттю речовин, тобто покриттю насичення; k_r, k_{des} – коефіцієнти, які характеризують швидкості реакції та десорбції CO; D_x, D_y – коефіцієнти дифузії CO в напрямку координат x та y, відповідно. Показник степеня $q = 3$ враховує механізм прекурсора [9] для кінетики адсорбції CO, оскільки пригнічення адсорбції для CO та O_2 є асиметричним – адсорбований CO сильніше блокує адсорбцію кисню.

Структурний фазовий перехід (1×2)↔(1×1) на поверхні Pt(110) моделюємо таким законом [10-11]:

$$\frac{\partial w}{\partial t} = k_{ph}\left(\left(1 + exp\left(\frac{u_0 - u/u_{sat}}{\delta u}\right)\right)^{-1} - w\right). \quad (3)$$

Змінна w позначає частку поверхні неперебудованої структури (1×1), коефіцієнт k_{ph} – швидкість структурного фазового переходу. Параметр u_0 визначає значення порогу, вище якого адсорбовані молекули CO істотно впливають на структуру поверхні, а δu – крутизну порогу.

Коефіцієнт налипання кисню s_v у рівнянні (2) запишеться як лінійна комбінація значень для структур (1×2)–(1×1):

$$s_v = s_v^{1\times1} w + s_v^{1\times2}\left(1 - w\right).$$

Тут $s_v^{1\times1}$, $s_v^{1\times2}$ – коефіцієнти налипання кисню у фазах (1×1) та (1×2), відповідно.

Постійні швидкостей реакції, десорбції та фазового переходу залежать від температури і підпорядковуються рівнянням Арреніуса [12]:

$$k_r[T] = k_r^0 \, exp\left(-E_r/RT\right),$$

$$k_{des}[T] = k_{des}^0 \, exp\left(-E_{des}/RT\right),$$

$$k_{ph}[T] = k_{ph}^0 \exp\left(-E_{ph}/RT\right).$$

Тут k_r^0, k_{des}^0, k_{ph}^0 – коефіцієнти, які не залежать від температури, E_r, E_{des}, E_{ph} – енергії активації, R – універсальна газова стала.

Для повноти моделі враховано генеровані при хімічних реакціях теплові процеси. У феноменологічному формулюванні [13] рівняння теплового балансу включає в себе теплопровідність та тепловиділення при реакції окислення і є таким:

$$\begin{aligned} c_p \rho \frac{\partial T}{\partial t} &= \kappa_{cond}\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2}\right) + \\ &+ \frac{\sqrt{2}}{d^2 L N_A}\left[\Delta H_{u,ads} p_u \kappa_u s_u \left(1 - \left(u/u_{sat}\right)^3\right) + \right. \\ &+ \Delta H_{v,ads} p_v \kappa_v s_v \left(1 - u/u_{sat} - v/v_{sat}\right)^2 + \\ &+ \left. \Delta H_{des} k_{des} u + \Delta H_r k_r u v\right]. \end{aligned} \quad (4)$$

Тут c_p – теплоємність, ρ – густина, κ_{cond} – теплопровідність каталізатора, d – довжина ребра елементарної комірки кристалічної ґратки Pt, L – товщина плівки, N_A – число Авогадро.

Рівняння (1)–(4) складають двовимірну математичну модель реакції окислення CO на поверхні Pt.

III. Числовий аналіз математичної моделі

Результати числового аналізу моделі (1)–(4) подано на рис.1–4. Величини параметрів, що використовувалися в числових розрахунках, наведені в [14].

Динаміку залежності поверхневого покриття CO $u(x,y,t)$ розраховано при заданому співвідношенні компонент тензора дифузійних коефіцієнтів:

$$\hat{D} = \begin{Vmatrix} D_x & 0 \\ 0 & D_y \end{Vmatrix},$$

а саме $D_x/D_y = 10$ (D_x – коефіцієнт дифузії CO вздовж осі OX, а D_y – коефіцієнт дифузії CO вздовж осі OY, які вважаються постійними). Динаміка поверхневого покриття CO $u(x,y,t)$ має яскраву осциляційну поведінку з періодом осциляції $\tau_0 = 5$, який слабо залежить від координат. В той же час амплітуда осциляцій $u(x,y,t)$ залежить від координати y і змінюється від 0,05 (при $y = 0{,}1$) до 0,1 (при $y = 1$), тобто зростає в два рази. Це говорить про вплив дифузії вздовж осі OY, незважаючи на те, що коефіцієнт дифузії вздовж осі OY: $D_y = 0{,}1D_x$. Осциляційна поведінка $u(x,y,t)$ сильно залежить як від співвідношення D_x/D_y, так і від початкових значень парціальних тисків p_u та p_v. Область осциляційної поведінки (область стійкості) реакції окислення CO, на відміну від одновимірної моделі [15], звужується і реальніше відповідає експериментально отриманій області стійкості реакції окислення CO.

Результати числового аналізу для поверхневого покриття кисню $v(x,y,t)$ при заданому співвідношенні коефіцієнтів дифузії $D_y = 0{,}1D_x$ показують, що динаміка залежності $v(x,y,t)$ має яскраво виражену осциляційну поведінку з періодом осциляції $\tau_0 = 5$ (як і для $u(x,y,t)$), який мало змінюється зі зміною y. Осциляції покриття $v(x,y,t)$ є в протифазі до осциляцій $u(x,y,t)$, а саме: максимум u_{max} покриття CO відповідає мінімуму v_{min} покриття O. Це фізично зрозуміло, оскільки окислення CO потребує атомарного кисню. Як і для залежності $u(x,y,t)$ поведінка $v(x,y,t)$ значно змінюється зі зміною співвідношення D_x/D_y, а саме: при незмінному періоді коливань амплітудні значення значно змінюються, що пов'язано з можливістю дифундування адсорбованого CO вздовж обох осей координат.

Динаміка частки поверхні неперебудованої структури (1×1) $w(x,y,t)$ теж носить осциляційний характер з періодом коливань $\tau_1 \approx 7 \div 8$, який слабо залежить від y-координати, але відрізняється від періоду зміни $u(x,y,t)$ та $v(x,y,t)$. Це зрозуміло, оскільки релаксація структури поверхні каталізатора повинна йти із запізненням щодо зміни $u(x,y,t)$, так як зміна поверхневого покриття CO є тим чинником, що приводить до структурної перебудови поверхні каталізатора. Як і в попередніх випадках період слабо залежить від y-координати, однак амплітудні значення залежать від y.

Поведінка $T(x,y,t)$ має осциляційний характер з періодом, що співпадає з періодом зміни покриттів $u(x,y,t)$ та $v(x,y,t)$. Залежність $T(x,y,t)$ від координати y є слабкою і незначно змінюється в залежності від співвідношення D_x/D_y.

IV. Висновки

У роботі побудовано та досліджено двовимірну математичну модель оксидації чадного газу для механізму Ленгмюра-Гіншелвуда. Показано, що врахування двовимірності приводить до зміни області стійкості реакції окислення CO. Як і в одновимірній моделі [15], в області стійкості реакція окислення CO на поверхні Pt-каталізатора є періодичною з періодом $\tau_0 = 5$ (що незначно відрізняється від одновимірного випадку) і відповідає експериментально спостережуваній. Амплітудні значення парціальних тисків p_u та p_v залежать від вимірності моделі та від співвідношення D_x/D_y (D_x – складова тензора дифузії вздовж осі OX, D_y – складова тензора дифузії вздовж осі OY).

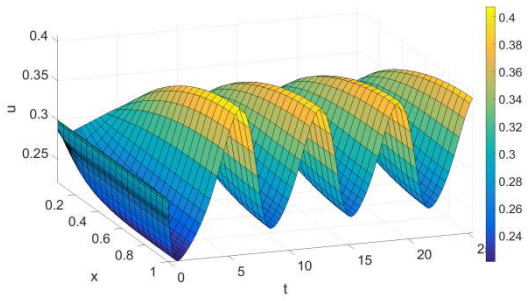


Рис.1. Залежність поверхневого покриття CO *u(x,y,t)* при фіксованому значенні координати *y = 1*.

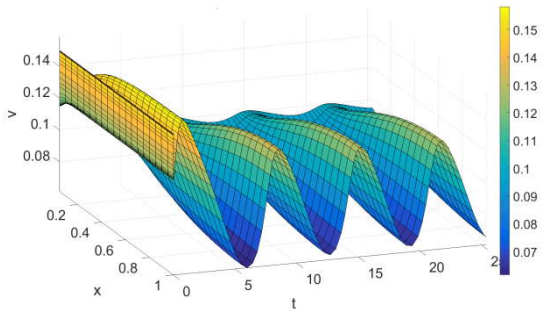


Рис.2. Залежність поверхневого покриття O *v(x,y,t)* при фіксованому значенні координати *y=1*.

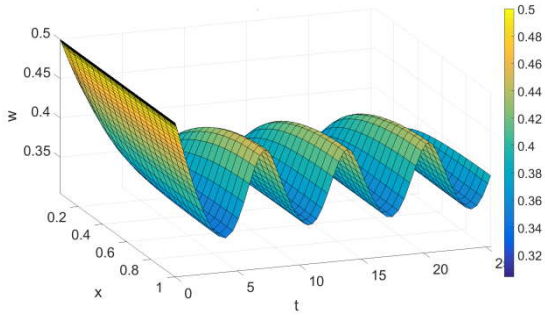


Рис.3. Залежність частки поверхні неперебудованої структури (1×1) *w(x,y,t)* при фіксованому значенні координати *y=1*.

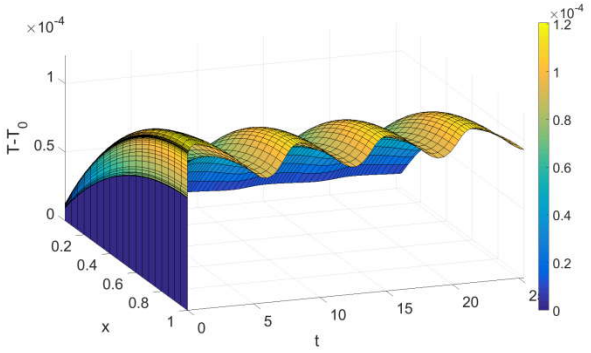


Рис.4. Залежність температури $T(x,y,t)–T_0$ при фіксованому значенні координати *y=1*.

Література References

[1] M. M. Slinko, N. I. Jaeger, *Oscillating Heterogeneous Catalytic Systems*, vol. 86, 1st ed., Eds. Amsterdam: Elsevier, 1994.

[2] R. J. Baxter, P. Hu, "Insight into why the Langmuir-Hinshelwood mechanism is generally preferred," *J. Chem. Phys*., vol. 116, no. 11, pp. 4379-4381, 2002.

[3] M. Wilf, P. T. Dawson, "Adsorption and desorption of oxygen on the Pt(110) surface; A thermal desorption and LEED/AES study," *Surf. Sci*., vol. 65, pp. 399-418, 1977.

[4] R. Gomer, "Diffusion of adsorbates on metal surfaces," *Rep. Prog. Phys*., vol. 53, no. 7, pp. 917-1002, 1990.

[5] G. L. Kellogg, "Direct observations of the (1×2) surface reconstruction on the Pt(110) plane," *Phys. Rev. Lett*., vol. 55, no. 20. pp. 2168-2171, 1985.

[6] T. Gritsch, D. Coulman, R. J. Behm, G. Ertl, "Mechanism of the CO-induced (1×2)-(1×1) structural transformation of Pt(110)," *Phys. Rev. Lett*., vol. 63, no. 10, pp. 1086-1089, 1989.

[7] K. Krischer, M. Eiswirth, G. Ertl, "Oscillatory CO oxidation on Pt(110): Modeling of temporal self-organization," *J. Chem. Phys*., vol. 96, pp. 9161-9172, 1992.

[8] M. Baer, M. Eiswirth, H. H. Rotermund, G. Ertl, "Solitary-wave phenomena in an excitable surface-reaction," *Phys. Rev. Lett*., vol. 69, no. 6, pp. 945-948, 1992.

[9] R. P. H. Gasser, E. B. Smith, "A surface mobility parameter for chemisorption," *Chem. Phys. Lett*., vol. 1, no. 10, pp. 457-458, 1967.

[10] M. Bertram, A. S. Mikhailov, "Pattern formation on the edge of chaos: Mathematical modeling of CO oxidation on a Pt(110) surface under global delayed feedback," *Rhys. Rev. E*., vol. 67, no. 3, pp. 036207-1-11, 2003.

[11] I. S. Bzovska, I. M. Mryglod, "Chemical oscillations in catalytic CO oxidation reaction," *Condens. Matter Phys*., vol. 13, no. 3, pp. 34801:1-5, 2010.

[12] В. Штиллер, *Уравнение Аррениуса и неравновесная кинетика*, М.: Мир, 2000.

[13] Y. Cisternas, P. Holmes, I. G.Kevrekidis, X. Li, "CO oxidation on thin Pt crystals: Temperature slaving and the derivation of lumped models," *J. Chem. Phys*., vol. 118, no. 7, pp. 3312-3328, 2003.

[14] P. P. Kostrobij, I. A. Ryzha, "Modeling of carbon monoxide oxidation process on the two-dimensional catalyst surface," *Math. Model. Comput.*, vol. 3, no. 2, pp. 146-162, 2016.

[15] І. С. Бзовська, І. М. Мриглод, "Поверхневі структури в каталітичній реакції монооксиду вуглецю," *Укр. фіз. журн*., Т. 61, № 2, С. 140-148, 2016.

The Movable Cellular Automata as a Tool to Simulation Biological Process and Systems

Kateryna Gazdiuk, Volodymyr Zhikharevich, Sergiy Ostapov
Department of computer technologies,
Yuriy Fedkovych Chernivtsi National University
Chernivtsi, Ukraine
kateryna.gazdyik@gmail.com, vzhikhar@mail.ru, sergey.ostapov@gmail.com

Рухомі Клітинні Автомати як Інструмент Моделювання Біологічних Процесів та Систем

Катерина Газдюк, Володимир Жихаревич, Сергій Остапов
Відділ комп'ютерних технологій
Чернівецький національний університет імені Юрія Федьковича
Чернівці, Україна
kateryna.gazdyik@gmail.com, vzhikhar@mail.ru, sergey.ostapov@gmail.com

***Abstract*—There is a simulation of the biological process and systems using the movable cellular automata**

***Анотація*—У даній роботі розглянуто моделювання біологічних процесів і систем методом рухомих клітинних автоматів**

Ключові слова—рухомі клітинні автомати, моделювання біологічних процесів, штучні нейронні мережі

Keywords—movable cellular automata; simulation biological proces; artificial neural network

I. INTRODUCTION

Recently the trend of computer simulation related to various aspects of the organization and functioning of biological systems is actively developing. The range of objects for modeling is quite wide - from elementary processes of self-replication [1] to the dynamics of multicellular organisms and nervous subsystem [2]. Typically, the implementation of these models is very complicated task. To realize the full potential of this field, it is essential to use the heritage and make connections with such sciences as biology, the theory of algorithms, cognitive systems, adaptive systems, robotics and others. An important component in modeling of biological processes and systems is artificial neural networks (ANN) or their equivalents, since it enables learning and adapting components of artificial organism to changing environmental conditions.

II. PREVIOSLY INVESTIGATIONS

The interesting approach to the problem of artificial life simulations is a 'Gene Pool' project created by Jeffrey Ventrella [4]. Gene Pool is a virtual Darwinian Aquarium where the user witness evolution unfolding before users eyes. Mate selection criteria can be tweaked to see what happens when "fat" or "immobile" or "long" are considered beautiful - even if these attributes run counter to having the anatomy and motor control for efficient swimming. This simulation has no goal or win and lose but the user can make up goals for instance to breed the fastest swimmer or try and breed two types of swimmers which can coexist in harmony.

Other concept of multicellular development [5] to evolve bodies and controllers of soft-bodied animats, evaluated within simulated physics environment. Growth of an embryo is driven by a neural network-like controller in every cell. The authors shows how, through an evolutionary process, this simple approach produces rich and complex morphologies of animats made of hundreds of cells and exhibiting life-like gaits without centralized control. They then show how the system allows us to evolve animats that optimize both for distance and the use of actuating material, and demonstrate how it can be used to evolve developmentally plastic individuals, whose genomes encode more than one phenotype, each adapted to a different type of environment.

In the paper [6] authors have investigated how the concept of metamorphosis can enhance artificial developmental processes used to evolve designs of robotic morphologies and controllers. Metamorphosis allowed to evolve solutions in

which a single genome encodes two related phenotypes, each crafted for a different problem and/or environment, together with a method of self-restructuring one phenotype into the other. Given a biologically inspired model of development, this required simply employing a fitness function that rewards each developmental stage for a different objective. It was observed how evolution and searching for morphologies de novo, without any prior assumptions about what they should be, was able to discover designs that are well adapted to different tasks at each stage of development. Reversing the order of environments in which growing individuals were evaluated allowed us to confirm that observed morphological changes were clearly a result of evolution shaping growing morphologies in a convergent way: legged locomotion for a terrestrial environment and undulatory swimming for an aquatic one, regardless of the order of evaluation. A further comparison with the evolution allowed finding that metamorphic individuals indeed produced useful specialization, outperforming environmentally robust solutions.

Other interesting approach for the evolution of multicellular animats can be observed in the [7]. In this article was presented a model of parallel co-evolution of development and motion control in soft-bodied – multicellular animats without neural networks. Development is guided by an artificial gene regulatory network (GRN) with real-valued expression levels, contained in every cell. Evolution was successful for generating moving animats and discovering several functional locomotion strategies. Motion was controlled through coordinated cell actions, where individual cells displayed emergent periodic patterns of expansion and contraction. The animat model used in this paper, a collection of springs modifying their resting length, is similar to a model of a soft-bodied robot.

A very interesting approach is morphogenetic robotics by Jin and Meng [8] examine methodologies for designing self-organizing, self-recongurable and self-adaptive robots inspired by biological morphogenesis. They categorized these systems into three main areas: swarm robotics, modular robots and body-brain co-design, and discuss their relationships with epigenetic robotics, which deals with cognitive development, and evolutionary robotics, which is concerned with the automated design of controllers. In their investigation they follow broad distinction between “swarm” and “modular” robots, keeping focus on morphogenetic abilities, and with the additional limitation to small-scale and exact-t construction systems. In a third subgroup, they described nonrobotic block structures, whether made by external “worker” robots or arising spontaneously from physical interactions.

Complex self-reconguration and locomotion tasks are shown M-TRAN system [9] is a well-known self-recongurable modular robot. Their robotic unit mechanical design was simpler and yet capable of more complex movements than other models such as crystalline atoms. The shape of this unit is anisotropic with two rotational degrees of freedom. The authors fully implemented the physical robots and successfully demonstrated intriguing behaviors, including snake-like locomotion, caterpillar track, quadruped walk, obstacle avoidance, and dynamic self-transformation among those shapes or behaviors.

Modular robots that can demonstrate complex behaviors including self-reconguration and self-reproduction were developed by Lipson with his lab. [10]. Their robotic unit, called “molecube”, is a cubic module with only one rotational degree of freedom that swivels two halves of the cube around a diagonal axis. The architecture of the system becomes even simpler and more programmable than earlier models. Using this framework, scientists of this lab demonstrated self-reproduction and self-repair of simple morphologies, as well as evolutionary acquisition of novel forms in a simulated universe.

Based on the theoretical knowledge, a new class of systems collectively called artificial embryogeny (AE, sometimes “artificial ontogeny” or simply “artificial development”) was appeared. This direction of science develops well among others through the works and reviews of Miller and Banzhaf [11], or Stanley and Miikkulainen [12]. Inspired by the biological development of multicellular organisms, AE systems realize the indirect mapping from genotype to phenotype via more or less complex developmental stages. In this investigations proposed codding the parameters of the genotype code for microscopic features of the cells (the agents), i.e. their abilities to communicate, their propensity for motion, and their affinities for assembly with other cells, instead of coding directly for macroscopic features phenotype (the system). Like biological cells, the agents of an AE system generally share the same genotype, i.e. the same set of developmental and behavioral rules. Imitating cell division, differentiation and self-positioning, an agent spawns other agents, follows its own execution path within the common program which may diverge from its neighbors' path depending on its position and creates specific links with other nodes according to its fate.

Particular attention should be given to the shaping of artificial neural networks in the literature. Structured ANNs can provide highly interesting auto-teaching structures as mentioned by Nolfi in [13]. “Auto-teaching networks” consist of two subnets, a “teaching network” and a “controlling network”. The authors describe, that such a network, if it is shaped by an artificial evolutionary process, has “genetically inherited predispositions to learn”. The virtual embryogenetic approach described in this section aims for the evolution of neural networks with substructures with this ability.

Another work that deals with the problem of differing functions within an ANN is [14], in which different predispositions for learning are implemented by “virtual adaptive neuro-endocrinology”. Within such a network, different types of cells exist: gland-cells, which influence the learning of the network, and regular cells, which are influenced by the gland-cells.

A different approach to shape the structure of an ANN has been described in [15]. In the system described, the position of the cell is fixed from the beginning, the growth of axons is controlled by morphogenes diffusing throughout the embryo. The advantages of this system seem to be the higher tendency for self-similarity of network structures during evolutionary processes, but on the other hand it seems that it has a low probability to develop differentiating cell-types with different functions.

III. THE SIMULATION OF LOCOMOTION

In this paper it is suggested to use moveable cellular automata (MCA) as a tool for modeling various biological processes and systems. As it is known, these computing structures are characterized by its simplicity, versatility and possess natural parallelism. Important advantage of the MCA method is a possibility of direct simulation of materials fracture including damage generation, crack propagation, fragmentation and mass mixing. In the case of modeling biological processes and systems, complex metabolic processes that occur in living systems can be presented by the appropriate exchange of contents of the MCA nodes, the synthesis of tissue – by division of the nodes, dying – by removing of the nodes, muscle contraction – by reducing the distance between the nodes etc.

In particular the simulation of a process of locomotion of life-like organisms [16], the simulation object is divided into two subsystems – mechanical and neural.

Mechanical subsystem represent the corresponding nodes of MCA of organism and simulates different body tissue i.e. muscle, connective, nervous. The scheme of cellular automata algorithm is asynchronous. It provides a random convenience sampling of one CA from the set and the appropriate modification of its state and the state of its nearest neighbors according to the rules of interaction. After interaction finishes another MCA is chosen randomly and the process is repeated. A signal for cellular automata interactions is the state of the corresponding "nerve ending" of the neuron subsystem that is associated with corresponding CA.

During the simulation of a neuron subsystem the elementary analogues of artificial neurons (perceptrons) are implemented. Their basic task, similar to that of biological neurons, is to form output signals guided evolutionally algorithm, according to the set of input signals, their state, and the function of transformation. For every separate CA the coordinates of remote fragments of the designed organism are pointed out, the states of which are input signals for the corresponding neuron. Thus, input connections appear chaotically. The amount of connections is fixed for all CA.

Arbitrary evolutionary algorithm shows the random changes (mutations) of parameters and estimation of influence of these changes on the dynamics of the system according to a corresponding criterion. If a criterion is satisfied, changes are fixed, otherwise they are rejected, and that is the complete analogue of natural selection. The criterion can constantly change during the existence of the body, from reducing energy consumption during the locomotion, achieving the maximum speed, the struggle for survival up to adaptation to new environmental conditions.

This approach works well in the absence of growth of the organism and "reversible deformation", since neighboring cells and the connections between them remain unchanged.

IV. THE GROWS PROCESS

Much more complex modeling process is associated with the ability of the organism to grow or change shape in the process of adaptation. In this case, cell shift occurs, or connections between them change within the body.

In concept of virtual growth, which is considered here, we are mimicking processes observable in biology during the developmental phase of most multicellular life-forms. We also included general concepts of biological embryogenesis and artificial embryogenesis [17]. The fast calculation of single embryological processes is necessary, due to the requirement for fast simulation of embryological processes on systems with limited hardware resources, especially for projects dealing with evolution in autonomous robotic systems

During the process of artificial growth of the organism, the embryo is considered as the cellular automata field of separate cells linked in some way. The cells of embryo can reproduce itself, die, acquire a certain type, build connections with other cells of the embryo. An important aspect of this model is that the cells do not have the ability to move actively, but can be "pushed out" from their spatial position during the growth (duplication of other cells). The growth of the embryo (and thus built artificial neural network) is subject to the outcome of the process of self-organization. The process of growth of the embryo is considered complete when there is no more cell duplication or cell death over a long time span. It should be noted also that the cases of infinite growth of the embryo or its complete disappearance are possible. In connection with these "pathological" forms of embryo growth, it is necessary to stop the process if the number of cells exceeds (in the case of infinite growth) or is less than (in the case of death) certain prescribed limits.

After the process of growth stops, the cells and relevant links between them undergo a detailed analysis. It is clear that the cells, which were neighbors at an early stage, can

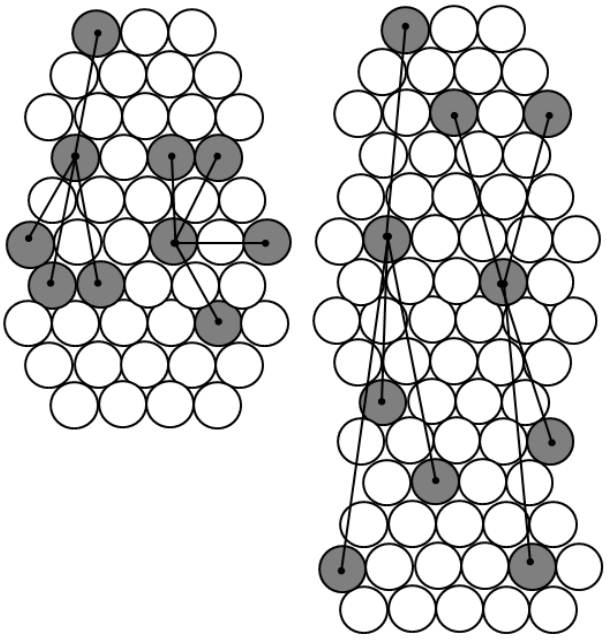


Fig. 1. An example of connection between cells before (in the left fig.) and after (in the right fig.) some period of process of organism's grows.

significantly move away from one another. In this case, there is

a problem of constant search for the nearest neighbors, and establishing ties between them.

The method of search of the nearest neighbors under the conditions of their fixed quantity is described in the papers [18]. In this work, the criterion of the neighboring cells search is the minimum distance between them, and the search is carried out by introducing auxiliary index array. The suggested algorithm showed good results for both uniform and random distribution of cells in space. A significant advantage is speed and low need for computing resources that is important in the development and operation of processes connected with the evolution of autonomous systems, as mentioned above.

V. CONCLUSIONS

This paper provides an overview of works in the field of modeling biological processes and systems as well as fundamental approaches to solve problems of interaction between cells in multicellular organisms.

Study of the works of the famous scientists in this field enabled conceptual and technological breakthrough for achieving extended reliability, advanced adaptability and artificial evolution. It is a main tool for sustainable development for the next generation of collective robotics and adaptive autonomous systems. Based on this investigation, artificial organisms reflect many similarities with biological organisms. This, in turn, opens a way for many great ideas, like artificial sexuality or inspiration from “natural chemistry”, which may seem speculative at this moment, but are very promising in terms of their biological analogies. Since the research should be focused on the specific goals, many of these ideas will be not followed further till their development into the real systems.

As a method for modeling the dynamics and development of bio-organisms, the method of movable cellular automata was chosen. Significant advantages of this method are increased speed and decreased capacity of computing resources.

As a result it is possible to simulate both simple and more complex multicellular organisms, collect the data and study their behavior, embryogenesis, self-organization, self-replication, their nervous system, muscular system etc. Of course this will be possible only with the development in this area and a more detailed analysis of the subject area. But it gives significant impetus to future research.

Development in this branch can give important results, which require a lot of time for development and research in both biology and programming. This study presents rather primitive algorithms of the nervous system’s simulation, but further development of hardware enables simulation of much more complex examples and gaining the speed of signal transmission at the level of human neural network.

REFERENCES

[1] H. Fellermann, S. Rasmussen, Physically Grounded “Simulations of a Self-Replicating Chemical Aggregate”, Proc. of the Alife XII Conference, Odense, Denmark, 2010, pp. 145-146.

[2] B. Wrobel,“Brainless Bodies: Controlling the Development and Behavior of Multicellular Animats by Gene Regulation and Diffusive Signals”, Artificial Life, vol. 13, 2012, pp 349-356.

[3] A. M. Turing,: “The Chemical Basis of Morphogenesis. Philosophical Transactions of the Royal Society of London”, series B, Biological Sciences 237(641), 1952, pp 37–72.

[4] J. Ventrella, Sexual Swimmers, Emergent Morphology and Locomotion Without a Fitness Function. From Animals to Animats, MIT Press, 1996 pp. 484-493.

[5] M. Joachimczak, R. Suzuki, and T. Arita, "Fine grained artificial development for Body-Controller coevolution of Soft-Bodied animats," in Artificial Life 14: Proceedings of the Fourteenth International Conference on the Synthesis and Simulation of Living Systems, the MIT Press, Jul. 2014, pp. 239-246. Available: https://mitpress.mit.edu/sites/default/files/titles/content/alife14/978-0-262-32621-6-ch040.pdf

[6] M. Joachimczak, R. Suzuki, and T. Arita, "Artificial metamorphosis: Evolutionary design of transforming, Soft-Bodied robots," Artificial Life, May 2016. Available: http://www.mitpressjournals.org/doi/pdf/10.1162/ARTL_a_00207

[7] M. Joachimczak, T. Kowaliw, R Doursat., and B. Wróbel, “Brainless bodies: Controlling the development and behavior of multicellular animats by gene regulation and diffusive signals” in Adami, C., Bryson, D. M., Ofria, C., and Pennock, R. T., editors, Artificial Life XIII: Proceedings of the Thirteenth International Conference on the Simulation and Synthesis of Living Systems, Cambridge, MA. MIT Press, 2012 pp. 349-356.

[8] Y. Jin, Y. Meng, (2011), Morphogenetic robotics: An emerging new field in developmental robotics. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on 41(2), 145-160.

[9] S. Murata, E. Yoshida, A. Kamimura, H. Kurokawa, K. Tomita, and S. Kokaji, (2002), M-tran: Self-recongurable modular robotic system. Mechatronics, IEEE/ASME Transactions on 7(4), 431-441.

[10] P. Grouchy, and H. Lipson, (2012), Evolution of self-replicating cube conglomerations in a simulated 3d environment. In Artificial Life, vol. 13, pp. 59-66.

[11] J. F. Miller, and W. Banzhaf, (2003), Evolving the program for a cell: from french ags to boolean circuits. On Growth, Form and Computers pp. 278-301.

[12] K. O. Stanley, and R. Miikkulainen, (2003), A taxonomy for artificial embryogeny. Artificial Life 9(2), 93-130.

[13] S. Nolfi, D. Parisi, Learning to adapt to changing environments in evolving neural networks.Adaptive Behavior 5(1), (1997), pp. 75–98

[14] J. Timmis, M. Nealz, J. Thorniley, An Adaptive Neuro-Endocrine System for Robotic Systems. In: Proc. of the IEEE Workshop on Robotic Intelligence in Informationally Structured Space, RiiSS 2009

[15] A. N. Hampton, C. Adami, Evolution of Robust Developmental Neural Networks. ArXiv Nonlinear Sciences e-prints (May 2004)

[16] V. Zhikharevich, K. Gazdiuk, S. Ostapov, “Simulation of the dynamics of worm-like organisms using the movable cellular automata”, VI-th International Conference of Students, PhD-Students and Young Scientists “Engineer of XXI Century”, 02 Dec. 2016, pp. 469-475.

[17] D. Basanta, M. Miodownik, B. Baum, “The Evolution of Robust Development and Homeostasis in Artificial Organisms”. PLoS Computational Biology 4(3), 2008.

[18] V. V. Zhykharevych, K. P. Hazdiuk, Alhorytm vyznachennia susidnikh elementiv mnozhyny rukhomykh klitynnykh avtomativ za umov fiksovanoi kilkosti susidiv. Visnyk Natsionalnoho tekhnichnoho universytetu "Kharkivskyi politekhnichnyi instytut". Zbirnyk naukovykh prats. Seriia: Informatyka ta modeliuvannia, Kharkiv: NTU "KhPI", № 33 (1142), 2015, s.75-82.

Поверхнева Енергія Деформації і Поляризації у Межах Градієнтної Теорії Пружних Діелектриків з Квадрупольними Електричними Моментами

Ольга Грицина
відділ математичних методів
обчислювального експерименту,
Центр математичного моделювання
Інституту прикладних проблем механіки і математики
ім. Я. С. Підстригача НАН України
Львів, Україна
gryt@cmm.lviv.ua, gryt045@gmail.com

Василь Кондрат
кафедра інженерної механіки,
Національна академія сухопутних військ
імені гетьмана Петра Сагайдачного
Львів, Україна
vasyl.kondrat@gmail.com

Surface Energy of Deformation and Polarization in the Framework of Gradient Theory of Elastic Dielectrics with Quadruple Polarization

Olha Hrytsyna
department for mathematical methods
of computing experiment,
Center of Mathematical Modeling of Pidstryhach Institute
for Applied Problems of Mechanics and Mathematics,
National Academy of Sciences of Ukraine
Lviv, Ukraine
gryt@cmm.lviv.ua, gryt045@gmail.com

Vasyl Kondrat
department of engineering mechanics,
Hetman Petro Sahaydachnyi Academy
of Army Ground Forces
Lviv, Ukraine
vasyl.kondrat@gmail.com

***Анотація*—У роботі сформульовано замкнену систему рівнянь нелінійної градієнтного типу теорії термоелектропружності неферомагнітних діелектриків. Розроблена модель враховує поряд із електричним дипольним моментом також квадрупольний момент. З умови інваріантності рівняння балансу енергії відносно просторових трансляцій, за використання рівнянь балансу маси й ентропії, а також рівняння балансу енергії електромагнітного поля, отримано рівняння балансу механічного імпульсу і нелокальні визначальні співвідношення. Одержані рівняння описують п'єзоелектричні й флексоелектричні явища, а також термоізоляційний ефект. В ізотермічному наближенні співвідношення розробленої теорії використані для розрахунку поверхневої енергії деформації і поляризації твердих тіл. Показано, що зазначена енергія визначається поверхневим значенням скалярного добутку вектора електричного поля і квадрупольного моменту.**

***Abstract*—In the present work, a complete nonlinear gradient-type theory of the thermoelectroelasticity of non-ferromagnetic dielectrics is proposed. In addition to the electric dipole moment, this model includes the quadruple one as well. If we subject the energy balance equation to invariance under rigid body translation and make use of the mass and entropy balance equations as well as of the balance equation for the energy of an electromagnetic field, then the moment equation and the nonlocal constitutive relations are obtained in a systematic manner. The resulting equations also encompass such phenomena as piezoelectricity, flexoelectricity, and the thermopolarization effect. For an isothermal approximation, the proposed theory is used to calculate the surface energy of deformation and polarization. It is shown that the mentioned energy is defined by the scalar product of the electric field vector and the quadruple moment.**

Ключові слова—градієнтна електротермопружність; неферомагнітні діелектрики; нелокальні визначальні співвідношення; дипольний і квадрупольний електричні моменти; поверхнева енергія деформації і поляризації.

Keywords—gradient electrothermoelasticity; non-ferromagnetic dielectrics; nonlocal constitutive equations; dipole and quadruple electric moments; surface energy of deformation and polarization.

I. Вступ

Останні кілька десятиліть у науковій літературі значна увага надається розвитку нелокальних теорій діелектриків. Побудова нових математичних моделей діелектриків була стимульована, з одного боку, необхідністю подолати низку невідповідностей, які виникають у класичних (локальних) теоріях (наприклад, уникнути сингулярностей розв'язків у задачах із зосередженими чинниками), а з іншого — потребою обґрунтувати ряд спостережуваних ефектів, які не охоплюють класичні моделі. Один із шляхів побудови узагальнених моделей діелектриків – розвиток теорій поляризовних середовищ із внутрішніми ступенями свободи (урахування мікроповоротів, мікродеформацій тощо). До таких належать, наприклад, мікроморфні і мікрополярні теорії та похідні від них [1]. Інший шлях розбудови узагальнених моделей діелектриків полягав у розвитку теорій градієнтного типу [2, 3]. Їх будували, постулюючи залежність внутрішньої енергії від градієнтів деяких фізичних величин, наприклад, від градієнтів тензора деформації [4], вектора поляризації [5] чи вектора напруженості електричного поля [6]. При цьому, зазвичай, відповідні системи рівнянь моделей формували з використанням варіаційних методів. Метою пропонованого дослідження є розвиток континуально-термодинамічного підходу до побудови градієнтного типу моделей діелектриків. Покажемо, що для градієнтної теорії діелектриків «природним» фізичним параметром термодинамічного стану є градієнт вектора напруженості електричного поля. При цьому для одержання нелокальних рівнянь стану достатньо у поляризаційному струмі врахувати вищі електричні моменти, наприклад, квадрупольний момент. Співвідношення розробленої теорії застосуємо для визначення у твердих тілах поверхневої енергії деформації і поляризації

II. Система рівнянь моделі

Розглядаємо ізотропне термопружне поляризовне неферомагнітне тверде тіло, яке займає область (V) евклідового простору та обмежене гладкою поверхнею (Σ). У тілі протікають механічні, теплові й електромагнітні процеси, спричинені зовнішніми діями. Усі поля, які характеризують процеси, що протікають у тілі, повинні задовольняти фундаментальні закони фізики – балансу маси, імпульсу, моменту імпульсу, ентропії й енергії, а також рівняння Максвелла.

Запишемо рівняння балансу маси в інтегральній формі

$$\frac{d}{dt}\int\limits_{(V)} \rho\, dV = -\int\limits_{(\Sigma)} \rho \mathbf{v}\cdot\mathbf{n}\, d\Sigma\,. \tag{1}$$

У локальній формі рівняння (1) запишемо так [7]:

$$\frac{\partial \rho}{\partial t} + \nabla\cdot(\rho\mathbf{v}) = 0\,. \tag{2}$$

Тут ρ – густина маси, $\mathbf{v}$ – вектор швидкості конвективного руху, t – час, $\mathbf{n}$ – зовнішня нормаль до поверхні (Σ), ∇ – оператор Гамільтона, «$\cdot$» – скалярний добуток.

Швидкість зміни ентропії елемента тіла визначається притоком ентропії ззовні, конвективним складником потоку ентропії через поверхню тіла, виникненням ентропії σ_s за одиницю часу та джерелами тепла $\Re$. В інтегральній формі рівняння балансу ентропії має вигляд [8]

$$\frac{d}{dt}\int\limits_{(V)} \rho s\, dV = -\iint\limits_{(\Sigma)} \mathbf{J}_s\cdot\mathbf{n}\, d\Sigma - \iint\limits_{(\Sigma)} \rho s\mathbf{v}\cdot\mathbf{n}\, d\Sigma + \int\limits_{(V)} \sigma_s dV + \int\limits_{(V)} \rho\frac{\Re}{T} dV\,, \tag{3}$$

де s – питома ентропія, $\mathbf{J}_s$ – вектор густини потоку ентропії, T – абсолютна температура.

Рівняння (3) у локальній формі запишемо так

$$\rho\frac{ds}{dt} = -\nabla\cdot\mathbf{J}_s + \sigma_s + \rho\frac{\Re}{T}\,.$$

Враховуючи співвідношення $\mathbf{J}_s = \mathbf{J}_q/T$, яке пов'язує вектори густин потоків ентропії $\mathbf{J}_s$ й тепла $\mathbf{J}_q$, маємо

$$\rho T\frac{ds}{dt} = -\nabla\cdot\mathbf{J}_q + \frac{1}{T}\mathbf{J}_q\cdot\nabla T + T\sigma_s + \rho\Re\,. \tag{4}$$

Рівняння електродинаміки мають вигляд [9, 10]

$$\nabla\times\mathbf{E} = -\frac{\partial\mathbf{B}}{\partial t},\ \nabla\times\mathbf{H} = \mathbf{J}_{ef}\,, \tag{5}$$

$$\nabla\cdot\mathbf{B} = 0\,,\ \nabla\cdot\mathbf{D} = \rho_e \tag{6}$$

Тут $\mathbf{B}$, $\mathbf{H}$ – вектори індукції та напруженості магнітного поля; $\mathbf{E}$ і $\mathbf{D}$ – вектори напруженості та індукції електричного поля; $\mathbf{J}_{ef} = \mathbf{J}_e + \mathbf{J}_{ed} + \mathbf{J}_{es}$ – вектор густини повного електричного струму; $\mathbf{J}_e$ – вектор густини електричного струму, пов'язаного з переміщенням вільних зарядів; $\mathbf{J}_{ed} = \varepsilon_0\left(\partial\mathbf{E}/\partial t\right)$ – струм зміщення; $\mathbf{J}_{es}$ – поляризаційний струм, пов'язаний зі зміною не лише дипольного, а й квадрупольного електричних моментів; ρ_e – густина вільного електричного заряду; ε_0 – електрична стала; «$\times$» – векторний добуток.

Для неферомагнітного середовища, яке розглядаємо тут, можна записати, що

$$\mathbf{B} = \mu_0\mathbf{H}\,, \tag{7}$$

де μ_0 – магнітна стала.

Введемо вектор Π_e електричної поляризації, який пов'язаний з вектором $\mathbf{J}_{es}$ співвідношенням [10]

$$\Pi_e = \int_0^t \mathbf{J}_{es}\, dt \;\Rightarrow\; \mathbf{J}_{es} = \frac{\partial \Pi_e}{\partial t}\,.$$

Відтак, для вектора $\mathbf{J}_{ef}$ густини повного електричного струму маємо співвідношення

$$\mathbf{J}_{ef} = \mathbf{J}_e + \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \frac{\partial \Pi_e}{\partial t}\,. \tag{8}$$

Формулою

$$\rho_{e\pi} = -\nabla \cdot \Pi_e \tag{9}$$

введемо у розгляд густину наведеного електричного заряду [9]. Введена величина справджує закон збереження наведеного електричного заряду [9]

$$\frac{\partial \rho_{e\pi}}{\partial t} + \nabla \cdot \mathbf{J}_{es} = 0\,. \tag{10}$$

Вектор поляризації Π_e пов'язаний з дипольним моментом $\mathbf{P}$ та тензором електричного квадрупольного моменту $\hat{\mathbf{Q}}$ формулою [10]

$$\Pi_e = \mathbf{P} - \frac{1}{6}\nabla \cdot \hat{\mathbf{Q}}\,. \tag{11}$$

Тоді для вектора $\mathbf{D}$ електричної індукції маємо [10]

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \Pi_e = \varepsilon_0 \mathbf{E} + \mathbf{P} - \frac{1}{6}\nabla \cdot \hat{\mathbf{Q}}\,. \tag{12}$$

Із рівнянь Максвелла (5), (6) випливає співвідношення, яке трактують як рівняння балансу енергії електромагнітного поля [7]

$$\frac{\partial U_e}{\partial t} + \nabla \cdot \mathbf{S}_\mathbf{e} + \left(\mathbf{J}_e + \frac{\partial \Pi_e}{\partial t}\right) \cdot \mathbf{E} = 0\,. \tag{13}$$

Тут $U_e = \frac{1}{2}\left(\varepsilon_0 \mathbf{E}^2 + \mu_0^{-1}\mathbf{B}^2\right)$ – густина енергії електромагнітного поля, $\mathbf{S}_e = \mu_0^{-1}\mathbf{E}\times\mathbf{B}$ – густина потоку енергії електромагнітного поля [7].

Зазначимо, що останній доданок у формулі (13) відображає вплив електромагнітного поля на речовину, яка разом з полем складає єдину матеріальну систему. Запишемо цей доданок таким чином, щоб у нього входили тензор квадрупольного моменту $\hat{\mathbf{Q}}_*$ та вектори $\mathbf{E}_*, \mathbf{P}_*, \mathbf{J}_{e*}$ напруженості електричного поля, дипольного моменту й густини електричного струму в системі відліку центрів мас, яка рухається зі швидкістю $\mathbf{v}$ відносно лабораторної системи відліку. Для неферомагнітного тіла у нерелятивістському наближенні маємо: $\mathbf{E}_* = \mathbf{E} + \mathbf{v}\times\mathbf{B}$, $\hat{\mathbf{Q}}_* = \hat{\mathbf{Q}}$, $\mathbf{P}_* = \mathbf{P}$, $\mathbf{J}_{e*} = \mathbf{J}_e - \rho_e \mathbf{v}$. Тут вектор $\mathbf{J}_{e*}$ трактуємо як вектор густини струму провідності. Враховуючи це, а також рівняння балансу маси (2), після деяких перетворень надамо рівнянню (13) балансу енергії електромагнітного поля вигляду

$$\frac{\partial U_e}{\partial t} + \nabla \cdot \mathbf{S}_e + \mathbf{J}_{e*} \cdot \mathbf{E}_* + \rho \frac{d\mathbf{p}}{dt} \cdot \mathbf{E}_* + \rho \frac{d\hat{\mathbf{q}}}{dt} : \left(\nabla \otimes \mathbf{E}_*\right) +$$
$$+ \mathbf{v} \cdot \left[\rho_e \mathbf{E}_* + \left(\mathbf{J}_{e*} + \frac{\partial \Pi_e}{\partial t}\right) \times \mathbf{B} + \rho\left(\nabla \otimes \mathbf{E}_*\right) \cdot \mathbf{p} - \right.$$
$$\left. - \rho\left(\nabla \otimes \nabla \otimes \mathbf{E}_*\right) : \hat{\mathbf{q}}\right] - \nabla \cdot \left\{\left[\mathbf{p} \cdot \mathbf{E}_* + \hat{\mathbf{q}} : \left(\nabla \otimes \mathbf{E}_*\right)\right]\rho\mathbf{v}\right\} = 0\,. \tag{14}$$

Тут « $\otimes$ » – діадний добуток.

Замкнену систему рівнянь моделі одержимо, виходячи з рівняння балансу повної енергії, яке для системи «тіло-електромагнітне поле» запишемо так:

$$\frac{d}{dt}\int_{(V)} \mathscr{E}\, dV = -\oint_{(\Sigma)} \left[\rho\left(u + \frac{1}{2}\mathbf{v}^2\right)\mathbf{v} - \bar{\sigma} \cdot \mathbf{v} + \right.$$
$$\left. + \mathbf{S}_e + \mathbf{J}_q\right] \cdot \mathbf{n}\, d\Sigma + \int_{(V)} \left(\rho\mathbf{F} \cdot \mathbf{v} + \rho\Re\right) dV\,. \tag{15}$$

Тут $\mathscr{E} = \rho u + \frac{1}{2}\rho\mathbf{v}^2 + U_e$ – повна енергія системи, u – питома внутрішня енергія, $\rho\mathbf{v}^2/2$ – кінетична енергія, $\bar{\sigma}$ – тензор напружень Коші, $\mathbf{F}$ – вектор масових сил.

Трансформуючи співвідношення (15) із урахуванням рівнянь балансу маси (2), ентропії (4) та енергії електромагнітного поля (14), одержимо

$$\rho\frac{du}{dt} = \rho T\frac{ds}{dt} + \hat{\sigma}_* : \frac{d\bar{\boldsymbol{e}}}{dt} + \rho\mathbf{E}_* \cdot \frac{d\mathbf{p}}{dt} + \rho\nabla \otimes \mathbf{E}_* : \frac{d\hat{\mathbf{q}}}{dt} +$$
$$+ \mathbf{J}_{e*} \cdot \mathbf{E}_* - \mathbf{J}_q \cdot \frac{\nabla T}{T} - T\sigma_s +$$
$$+ \mathbf{v} \cdot \left(-\rho\frac{d\mathbf{v}}{dt} + \nabla \cdot \hat{\sigma}_* + \mathbf{F}_e + \rho\mathbf{F}\right). \tag{16}$$

Тут $\bar{\boldsymbol{e}}$ – тензор деформації Гріна, який пов'язаний з вектором переміщень $\mathbf{u}$ співвідношенням Коші

$$\bar{\boldsymbol{e}} = \left[\nabla \otimes \mathbf{u} + \left(\nabla \otimes \mathbf{u}\right)^{\mathrm{T}}\right]/2\,. \tag{17}$$

З умови інваріантності співвідношення (16) відносно просторових трансляцій отримаємо рівняння балансу механічного імпульсу

$$\nabla \cdot \hat{\sigma}_* + \mathbf{F}_e + \rho\mathbf{F} = \rho\frac{d^2\mathbf{u}}{dt^2}\,, \tag{18}$$

узагальнене рівняння Гіббса

$$\rho\frac{du}{dt} = \rho T\frac{ds}{dt} + \hat{\sigma}_* : \frac{d\bar{\boldsymbol{e}}}{dt} + \rho\mathbf{E}_* \cdot \frac{d\mathbf{p}}{dt} + \rho\nabla \otimes \mathbf{E}_* : \frac{d\hat{\mathbf{q}}}{dt} \tag{19}$$

та вираз для виробництва ентропії

$$\sigma_s = \mathbf{J}_{e*} \cdot \mathbf{E}_* \frac{1}{T} - \mathbf{J}_q \cdot \frac{\nabla T}{T^2} . \qquad (20)$$

Тут

$$\hat{\sigma}_* = \hat{\sigma} - \rho\left[\left(\mathbf{p} \cdot \mathbf{E}_*\right) + \rho\hat{\mathbf{q}} : \left(\nabla \otimes \mathbf{E}_*\right)\right] \hat{\mathbf{I}} ,$$

$$\mathbf{F}_e = \rho_e \mathbf{E}_* + \rho\left(\nabla \otimes \mathbf{E}_*\right) \cdot \mathbf{p} + \left(\mathbf{J}_{e*} + \frac{\partial \Pi_e}{\partial t}\right) \times \mathbf{B} -$$

$$-\rho\left(\nabla \otimes \nabla \otimes \mathbf{E}_*\right) : \hat{\mathbf{q}} ,$$

$\hat{\mathbf{I}}$ – одиничний тензор.

Наслідком урахування квадрупольного електричного моменту є модифікація тензора напружень $\hat{\sigma}_*$ і виникнення у пондеромоторній силі $\mathbf{F}_e$ додаткових складників.

Формули (19), (20) є основою для формування визначальних співвідношень моделі, а саме, нелокальних рівнянь стану та кінетичних співвідношень.

Якщо у формулі (19) за допомогою перетворення Лежандра

$$f = u - Ts - \mathbf{E}_* \cdot \mathbf{p} - \hat{\mathbf{q}} : \left(\nabla \otimes \mathbf{E}_*\right) , \qquad (21)$$

перейти до вільної енергії Гельмгольца, то отримаємо такі нелокальні рівняння стану:

$$s = -\frac{\partial f}{\partial T} , \ \hat{\sigma}_* = \rho \frac{\partial f}{\partial \hat{\boldsymbol{e}}} , \ \mathbf{p} = -\frac{\partial f}{\partial \mathbf{E}_*} , \ \hat{\mathbf{q}} = -\frac{\partial f}{\partial \nabla \otimes \mathbf{E}_*} . \qquad (22)$$

Нелокальність визначальних співвідношень розробленої математичної моделі зумовлена урахуванням у поляризаційному струмі Π_e поряд із дипольним $\mathbf{P}$ також квадрупольного $\hat{\mathbf{Q}}$ електричного моменту.

Щоб записати рівняння стану в явному вигляді, конкретизуємо вигляд вільної енергії. Розвинемо вільну енергію f в ряд за збуреннями параметрів стану і для малих збурень обмежимося в цьому розвиненні квадратичними членами. Приймемо, що у вихідному стані $\hat{\boldsymbol{e}} = 0$, $\hat{\sigma}_* = 0$, $T = T_0$, $s = s_0$, $\mathbf{E}_* = 0$, $\mathbf{p} = 0$, $\nabla \otimes \mathbf{E}_* = 0$, $\hat{\mathbf{q}} = 0$. Тоді для ізотропного початково однорідного тіла запишемо

$$f = f_0 - s_0\theta + \frac{1}{2\rho_0}\left(K - \frac{2}{3}G\right) I_{e1}^2 + \frac{G}{\rho_0} I_{e2} - \frac{C_V}{2T_0}\theta^2 -$$

$$-\frac{\chi_E}{2}\mathbf{E}_*^2 + \frac{\chi_{q1}}{2} I_{E1}^2 - \chi_{q2} I_{E2} - \frac{K\alpha_T}{\rho_0} I_{e1}\theta +$$

$$+\beta_{TE} I_{E1}\theta - \frac{K\alpha_{E1}}{\rho_0} I_{e1} I_{E1} + 2G\frac{\alpha_{E2}}{\rho_0}\hat{\boldsymbol{e}} : \left(\nabla \otimes \mathbf{E}_*\right) . \qquad (23)$$

Тут $I_{e1} = \hat{\boldsymbol{e}} : \hat{\mathbf{I}} = \boldsymbol{e}$, $I_{E1} = \left(\nabla \otimes \mathbf{E}_*\right) : \hat{\mathbf{I}} = \nabla \cdot \mathbf{E}_*$, $I_{e2} = \hat{\boldsymbol{e}} : \hat{\boldsymbol{e}}$, $I_{E2} = \left(\nabla \otimes \mathbf{E}_*\right) : \left(\nabla \otimes \mathbf{E}_*\right)$; $\theta = T - T_0$ – збурення температури; K, G, C_V, χ_E, χ_{q1}, χ_{q2}, α_T, β_{TE}, α_{E1}, α_{E2} – характеристики матеріалу.

Для ізотропних матеріалів на основі (22) і (23) маємо

$$\hat{\sigma} = 2G\hat{\boldsymbol{e}} + 2G\alpha_{E2}\nabla \otimes \mathbf{E} +$$

$$+\left[\left(K - \frac{2}{3}G\right)e - K\alpha_T\theta - K\alpha_{E1}\nabla \cdot \mathbf{E}\right]\hat{\mathbf{I}} , \qquad (24)$$

$$s = s_0 + \frac{C_V}{T_0}\theta + \frac{K\alpha_T}{\rho_0}e - \beta_{TE}\nabla \cdot \mathbf{E} , \qquad (25)$$

$$\mathbf{p} = \chi_E \mathbf{E} , \qquad (26)$$

$$\hat{\mathbf{q}} = 2\chi_{q2}\nabla \otimes \mathbf{E} - 2G\alpha_{E2}\hat{\boldsymbol{e}} - \left(\chi_{q1}\nabla \cdot \mathbf{E} - \frac{K\alpha_{E1}}{\rho_0}e + \beta_{TE}\theta\right)\hat{\mathbf{I}} . \qquad (27)$$

Аналіз одержаних визначальних співвідношень засвідчив, що розроблена модель враховує електромеханічну взаємодію у матеріалах високої симетрії (ізотропних матеріалах) і описує флексоелектричний та термополяризаційний ефекти у них.

Подамо рівняння (20) для виробництва ентропії у вигляді: $\sigma_s = \sum_{k=1}^{2} \mathbf{j}_k \cdot \mathbf{X}_k$. Тут $\mathbf{j}_k$ та $\mathbf{X}_k$ – термодинамічні потоки та сили: $\mathbf{j}_1 = \mathbf{J}_{e*}$, $\mathbf{X}_1 = \mathbf{E}_*/T$, $\mathbf{j}_2 = \mathbf{J}_q$, $\mathbf{X}_2 = -\nabla T/T^2$. Тоді на основі цього виразу можемо записати такі кінетичні рівняння

$$\mathbf{j}_1 = \mathbf{j}_1(\mathbf{X}_1, \mathbf{X}_2) , \ \mathbf{j}_2 = \mathbf{j}_2(\mathbf{X}_1, \mathbf{X}_2) . \qquad (28)$$

У лінійному наближенні кінетичні рівняння (28) набудуть вигляду

$$\mathbf{J}_{e*} = \sigma_e \mathbf{E}_* + \sigma_e \eta \nabla T , \ \mathbf{J}_q = -\lambda \nabla T + \pi_t \mathbf{J}_{e*} , \qquad (29)$$

де σ_e, λ – коефіцієнти електро- й теплопровідності, коефіцієнти η, π_t характеризують термоелектричні явища. Зазначимо, що $\pi_t = -\eta T_0$.

Замкнена нелінійна система рівнянь градієнтної електромагнітотермомеханіки поляризовних неферомагнітних тіл охоплює: рівняння балансу маси (2), ентропії (4) та імпульсу (18), співвідношення (5), (6), які відповідають законам Ампера, Фарадея, Гауса-Фарадея та Гауса-Кулона, формулу (8), визначальні співвідношення (7), (12), (22) і (28), формулу Коші (17), а також закон збереження наведеного електричного заряду (10).

III. Визначення поверхневої енергії деформації та поляризації

Відомо, що для поділу довільного твердого деформівного тіла на дві частини вздовж деякої поверхні (Σ) необхідно подолати енергію зв'язку між частинками, розташованими на та в околі цієї поверхні. Під такою розуміють енергію, яку необхідно затратити, щоб, залишивши незмінною деформацію тіла, зруйнувати атомні зв'язки вздовж

поверхні (Σ) [11]. Цього можна досягти, приклавши зовнішнє поле. Якщо ж це поле забрати, то тіло в околі поверхні здеформується. Енергію, пов'язану з цією деформацією, називають поверхневою енергією деформації. Отже, поверхнева енергія деформації — це та енергія, яку необхідно відняти від енергії зв'язку, щоб визначити поверхневу енергію. У працях [4, 11] показано, що поверхнева енергія деформації завжди від'ємна і складає до 30% від повної енергії зв'язування, а тому нехтувати нею не можна.

Класичні теорії пружності не пропонують шляху обчислення поверхневої енергії деформації. Використаємо для визначення цієї енергії записані вище співвідношення градієнтної теорії електропружності. Для спрощення викладок приймемо ізотермічне наближення.

Розглянемо рівноважний стан ізотропного тіла, що займає область (V) евклідового простору й обмежене гладкою поверхнею (Σ) із зовнішньою нормаллю $\mathbf{n}$. Нехай тіло є ідеальним діелектриком ($\rho_e = 0$) і контактує з вакуумом (область (V_v)).

У поданні (23) для вільної енергії Гельмгольца f врахуємо лінійні рівняння стану (24)-(27). Відтак, надамо потенціалу f вигляду

$$f - f_0 = \frac{1}{2\rho_0}\hat{\sigma} : \hat{\boldsymbol{e}} - \frac{1}{2}\mathbf{E}\cdot\mathbf{p} - \frac{1}{2}\hat{\mathbf{q}} : (\nabla \otimes \mathbf{E}). \qquad (30)$$

Якщо у співвідношенні (30) врахувати формулу $\mathbf{E} = -\nabla\varphi_e$, де φ_e — електричний потенціал [11], а також подання (21) для вільної енергії, формулу Коші (17), закон Гауса (6_2) і рівняння рівноваги

$$\nabla\cdot\hat{\sigma} + \rho_0\mathbf{F} = 0,$$

то після низки перетворень надамо йому вигляду

$$\mathscr{E} = \frac{1}{2}\rho_0\mathbf{F}\cdot\mathbf{u} + \frac{1}{2}\nabla\cdot(\hat{\sigma}\cdot\mathbf{u}) - \frac{1}{2}\nabla\cdot(\varphi_e\mathbf{D}) + \frac{1}{12}\nabla\cdot(\mathbf{E}\cdot\hat{\mathbf{Q}}).$$

Інтегруємо обидві частини записаної вище рівності по області $(V') = (V)\cup(V_\text{v})$, яку займає тіло з вакуумом. Перейшовши в останніх трьох доданках справа від знаку рівності до поверхневих інтегралів, запишемо

$$\int_{(V')}\mathscr{E}\,dV = \frac{1}{2}\rho_0\int_{(V)}\mathbf{F}\cdot\mathbf{u}dV + \frac{1}{2}\int_{(\Sigma)}\left(\hat{\sigma}\cdot\mathbf{u} - \varphi_e[\mathbf{D}] + \frac{1}{6}\mathbf{E}\cdot\hat{\mathbf{Q}}\right)\cdot\mathbf{n}d\Sigma. \qquad (31)$$

Тут $[\mathbf{D}]$ – стрибок вектора електричної індукції на поверхні (Σ).

Розглядаємо рівноважний стан тіла, поверхні якого вільні від дії зовнішніх зусиль ($\forall\mathbf{r}\in(\Sigma)$: $\hat{\sigma}\cdot\mathbf{n} = 0, [\mathbf{D}] = 0$), масові сили відсутні $(\forall\mathbf{r}\in(V)$: $\mathbf{F} = 0)$. Тоді з виразу (31) за відсутності зовнішньої дії на тіло отримаємо таке співвідношення

$$\int_{(V)}\mathscr{E}\,dV = \frac{1}{2}\rho_0\int_{(\Sigma)}(\mathbf{E}\cdot\hat{\mathbf{q}})\cdot\mathbf{n}d\Sigma.$$

Права частина цієї рівності визначає поверхневу енергію деформації і поляризації U_Σ, для визначення якої у межах розробленої теорії маємо формулу

$$U_\Sigma = \frac{1}{2}\rho_0\mathbf{E}\cdot\hat{\mathbf{q}}\cdot\mathbf{n}\Big|_{\mathbf{r}\in\Sigma}.$$

Отже, питома поверхнева енергія деформації і поляризації визначається поверхневим значенням проекції на нормаль до поверхні тіла скалярного добутку вектора напруженості електричного поля $\mathbf{E}$ і квадрупольного електричного моменту $\hat{\mathbf{q}}$.

IV. Висновки

Наслідком урахування у поляризаційному струмі електричних моментів вищого порядку є нелокальні рівняння стану узагальненої моделі електротермопружних неферомагнітних діелектриків. Показано, що з фізичної точки зору природним параметром для градієнтного типу теорій діелектриків є градієнт вектора напруженості електричного поля, спряженим до якого є питомий квадрупольний електричний момент. Розроблена теорія на рівні рівнянь стану враховує електромеханічну взаємодію в ізотропних матеріалах, а також описує флексоелектричний і термополяризаційний ефекти. Співвідношення цієї теорії дозволяють визначити поверхневу енергію деформації і поляризації твердих діелектриків. Ця енергія визначається поверхневим значенням проекції на нормаль до поверхні тіла згортки вектора напруженості електричного поля та квадрупольного електричного моменту.

Література References

[1] J A. C. Eringen, Microcontinuum field theories: foundations and solids. NewYork: Springer-Verlag, 1999.

[2] J. Yang, "Review of a few topics in piezoelectricity", *Appl Mech Rev.*, vol. 59, pp. 335–345, 2006.

[3] В. Кондрат, О. Грицина, «Лінійні теорії електромагнітомеханіки діелектриків», *Фіз.-мат. моделювання та інформ. технології*, вип. 9, с. 7-46, 2009.

[4] R. D. Mindlin, "Elasticity, piezoelectricity and crystal lattice dynamics", *J. Elasticity*, vol. 2, pp. 217–282, 1972.

[5] R. D. Mindlin, "Polarization gradient in elastic dielectrics", *Int. J. Solids Structures*, vol. 4, pp. 637–642, 1968.

[6] J. S. Yang, X. M. Yang, "Electric field gradient effect and thin film capacitance", *World. J. Eng.*, vol. 2, pp. 41–45, 2004.

[7] Я. Бурак, В. Кондрат, О. Грицина, Основи локально градієнтної теорії діелектриків. Ужгород: Поліграфцентр «Ліра», 2011.

[8] С. Де Гроот, П. Мазур, Неравновесная термодинаміка. Москва: Мир, 1964.

[9] М. М. Бредов, В. В. Румянцев, И. Н. Топтыгин, Классическая электродинамика. Москва: Наука, 1985.

[10] А. М. Федорченко, Теоретическая физика. Классическая электродинамика: Учеб. пособие. Киев: Вища школа, 1988.

[11] Ж. Можен, Механика электромагнитных сплошных сред. Москва: Мир, 1991.

Математичне Моделювання Дифузії у Середовищі з Пастками за Каскадного Розпаду Домішок

Ю.І Білущак
Відділ математичного моделювання нерівноважних процесів
Центр математичного моделювання
Інституту прикладних проблем механіки і математики ім. Я.С. Підстригача НАН України
Львів, Україна
byixx13@gmail.com

Mathematical Modelling Diffusion in a Medium with Traps under Cascade Decay of Admixture

Y. Bilushchak
Department of mathematical modeling nonequilibrium processes
Centre of Mathematical Modeling of Pidstryhach Institute for Applied Problems of Mechanics and Mathematics,
Ukrainian Academy of Sciences
Lviv, Ukraine
byixx13@gmail.com

***Анотація*—Запропонована математична модель масоперенесення у середовищі з пастками за каскадного розпаду домішок, в рамках якої концентрація частинок на певному кроці розпаду є джерелом маси розпадної речовини, яка дифундує і сорбується, на наступному кроці. Розв'язки крайових задач дифузії каскадного типу побудовані за ітераційною процедурою з використанням функцій Гріна. Розроблений пакет програм для комп'ютерного моделювання дифузійних процесів у тілі з пастками за каскадного розпаду домішкових речовин.**

***Abstract*—It is proposed the mathematical model of mass transfer in a medium with traps under cascade decay of admixture with in the scope of which the concentration of particles at certain step of decay is the source of the mass of decaying substance, diffusing and being sorbed at the next step. Solutions of initial-boundary value problems of diffusion of cascade type constructed by the iterative procedure with using Green's functions. The software package is developed for simulation of diffusion processes in the body with traps under cascade decay of admixture substance.**

Ключові слова—математичне моделювання; дифузія; сорбція-десорбція; середовище з пастками; каскадний розпад; функція Гріна; комп'ютерне моделювання

Keywords—mathematical modelling; diffusion; sorption-desorption; medium with traps; cascade decay; Green function; simulation

I. Вступ

Важливою проблемою сьогодення є прогнозування поширення розпадних токсичних забруднень у природному середовищі дифузійним шляхом. Прикладом таких агресивних субстанцій можуть бути пестициди, гербіциди, інсектициди, а також радіонукліди. Так, наприклад, схема розпаду дихлордифенілтрихлорметилметану (ДДТ) є

$$ДДТ \xrightarrow{250^{o}C} ДДД \xrightarrow{250^{o}C} ДДЕ \xrightarrow{250^{o}C} \begin{matrix} нетоксичні \\ продукти \end{matrix}$$

тут ДДД – дихлордифенілдихлоретан, ДДЕ – дихлордифенілдихлоретилен. Період піврозпаду ДДТ – 12–16 років [1, 2]. У випадку радіонуклідів відомі схеми розпаду [3, 4], наприклад:

$$^{97}_{36}Kr \to {}^{97}_{37}Rb \to {}^{97}_{38}Sr \to {}^{97}_{39}Y \to {}^{97}_{40}Zr \to {}^{97}_{41}Nb \to {}^{97}_{42}Mo\,,$$

$$^{113}_{44}Ru \to {}^{113}_{45}Rh \to {}^{113}_{46}Pd \to {}^{113}_{47}Ag \to {}^{113}_{48}Cd \to {}^{113}_{48}In \to {}^{114}_{50}Sn\,.$$

Причому у результаті розпаду радіоактивних речовин можуть виникати як більш, так і менш токсичні речовини. Так, у результаті розпаду плутонію-241 з періодом піврозпаду 14,4 років утворюється дуже токсичний америцій-241 з періодом піврозпаду 432,6 років. В той же час ре-

зультатом розпаду дуже шкідливого стронцію-90 з періодом піврозпаду 28,8 років є радіоактивний ітрій-90 з періодом піврозпаду 64 години. Зазначимо, що в літературі такі розпади субстанцій часто називають каскадними [3].

Актуальність розгляду явища дифузії в середовищах пастками обумовлена його застосуванням в таких галузях як фізика кристалів та напівпровідників, біофізика, наноматеріали, економіка, екологія тощо [5, 6]. Особливо важливе значення дослідження моделей дифузії у середовищах з пастками набуває при розв'язанні задач, пов'язаних з міграцією домішкових речовин у геологічних формаціях [7], оскільки на сьогодні ізоляцію найтоксичніших відходів вирішують шляхом глибинного підземного захоронення, а за великих часів можливий вихід радіоактивного забруднення за межі інженерних бар'єрів. При цьому основною моделлю для прогнозування перенесення домішок є модель дифузії у середовищах з пастками [7].

У даній роботі запропонована математична модель масоперенесення у багатокомпонентному середовищі з пастками, в рамках якої концентрація частинок на певному кроці розпаду є джерелом маси розпадної речовини, яка дифундує, на наступному кроці.

II. Математична модель

Приймаємо, що довільно вибрана фізично мала область середовища складається з твердої фази, в об'єм якої можуть проникати домішкові частинки, та порового простору, в якому дифундує домішкова речовина. Причому коефіцієнт дифузії домішки в об'ємі скелету настільки малий у порівнянні з коефіцієнтом дифузії у поровому просторі, що тверду фазу можна розглядати як пастки для домішкових частинок. Приймаємо, що в процесі такого масоперенесення скелет не деформується, і пористість залишається сталою (не враховуються зміни, пов'язані із сорбцією домішкової речовини).

Домішкова речовина $\mathbf{K}^{(0)}$ з поверхні поступає у тіло $\mathbf{K}^{*}$ і дифундує у поровому просторі, супроводжуючись процесами сорбції-десорбції. Частина $\mathbf{K}^{(0)}$ розпадається і утворюються речовини $\mathbf{K}^{(1)}$ і $\mathbf{K}^{(N)}$. Нехай частинки речовини $\mathbf{K}^{(N)}$ є нерозпадними (нешкідливими). Субстанція $\mathbf{K}^{(1)}$ теж розпадається і утворює речовину $\mathbf{K}^{(2)}$ (яка у свою чергу розпадається) та нерозпадні частинки, які відносимо до $\mathbf{K}^{(N)}$, і т.п. Причому нова речовина на кожному етапі утворюється в обох станах – у поровому просторі і пастках з однаковим коефіцієнтом інтенсивності розпаду, вона дифундує, може сорбуватися і десорбуватися.

Вихідні співвідношення математичної моделі багатокомпонентної дифузії за каскадного розпаду мігруючих частинок у середовищі з пастками сформульовані з використанням континуально-термодинамічного підходу механіки твердих розчинів [8]. Кожній компоненті ставиться у відповідність континуум, з допомогою якого описується кінематичні та деформаційні властивості компонент, а також формулюються балансові співвідношення, які відображають закони збереження маси, імпульсу, енергії та ентропії. З використанням концепції локальної термодинамічної рівноваги будуються рівняння стану та кінетичні рівняння. Вибравши в якості розв'язуючих функцій концентрації домішкових компонент, вектор переміщення, густину та температуру тіла, сформульовано ключову систему рівнянь, яка в подальшому лінеаризується.

У даному випадку деформація середовища і теплові процеси не розглядаються, і за ізотермічних умов, нехтуючи конвективними складовими масоперенесення та деформацією [8], отримаємо наступні рівняння дифузії у середовищі з пастками за каскадного розпаду мігруючих частинок на i -му кроці розпаду:

для $i = 0$

$$\frac{\partial c_1^{(0)}}{\partial t} = D^{(0)}\Delta c_1^{(0)} - k_1^{(0)}c_1^{(0)} + k_2^{(0)}c_2^{(0)} - \lambda^{(1)}c_1^{(0)} - \lambda^{(0N)}c_1^{(0)},$$

$$\frac{\partial c_2^{(0)}}{\partial t} = k_1^{(0)}c_1^{(0)} - k_2^{(0)}c_2^{(0)} - \lambda^{(1)}c_2^{(0)} - \lambda^{(0N)}c_2^{(0)};$$

для $i = \overline{1, N-1}$

$$\frac{\partial c_1^{(i)}}{\partial t} = D^{(i)}\Delta c_1^{(i)} - k_1^{(i)}c_1^{(i)} + k_2^{(i)}c_2^{(i)} + \lambda^{(i-1)}c_1^{(i-1)} - \lambda^{(i+1)}c_1^{(i)} - \lambda^{(iN)}c_1^{(i)},$$

$$\frac{\partial c_2^{(i)}}{\partial t} = k_1^{(i)}c_1^{(i)} - k_2^{(i)}c_2^{(i)} + \lambda^{(i-1)}c_2^{(i-1)} - \lambda^{(i+1)}c_2^{(i)} - \lambda^{(iN)}c_2^{(i)};$$

для $i = N$

$$\frac{\partial c_1^{(N)}}{\partial t} = D^{(N)}\Delta c_1^{(N)} - k_1^{(N)}c_1^{(N)} + k_2^{(N)}c_2^{(N)} + \sum_{i=0}^{N-1}\lambda^{(iN)}c_1^{(i)},$$

$$\frac{\partial c_2^{(N)}}{\partial t} = k_1^{(N)}c_1^{(N)} - k_2^{(N)}c_2^{(N)} + \sum_{i=0}^{N-1}\lambda^{(iN)}c_2^{(i)}. \qquad (1)$$

Тут $c_j^{(i)}$ - концентрація домішкової речовини, яка утворилася на i -му етапі розпаду, у стані j ($i = \overline{0, N}$; $j = 1, 2$); $D^{(i)}$ - коефіцієнт дифузії частинок для етапу $i = \overline{0, N}$; $k_1^{(i)}$ і $k_2^{(i)}$ - коефіцієнти інтенсивності сорбції і десорбції, $\lambda^{(i+1)}$ і $\lambda^{(iN)}$ - коефіцієнти інтенсивності розпаду речовини на i -му кроці, що відповідають утворенню нової речовини на $(i+1)$-му етапі та нерозпадної речовини для $i = N$. Зазначимо, що одиниці вимірювання коефіцієнтів інтенсивності розпаду та сорбції- десорбції співпадають $[c^{-1}]$.

Приведемо системи рівнянь (1) до безрозмірного вигляду. З цією метою введемо нові змінні [9]

$$t = k_2^{(0)}\tau \ ; \quad \xi^{(\alpha)} = \left(k_2^{(0)} / D^{(0)}\right)^{1/2} x^{(\alpha)} , \quad \alpha = \overline{1,3} . \qquad (2)$$

Тоді отримаємо
для $i = 0$

$$\frac{\partial c_1^{(0)}}{\partial \tau} = d^{(0)} \Delta_\xi c_1^{(0)} - a_1^{(0)} c_1^{(0)} + a_2^{(0)} c_2^{(0)} ,$$

$$\frac{\partial c_2^{(0)}}{\partial \tau} = a_3^{(0)} c_1^{(0)} - a_4^{(0)} c_3^{(0)} ; \qquad (3а)$$

для $i = \overline{1, N-1}$

$$\frac{\partial c_1^{(i)}}{\partial \tau} = d^{(i)} \Delta_\xi c_1^{(i)} - a_1^{(i)} c_1^{(i)} + a_2^{(i)} c_2^{(i)} + a_{\lambda 1}^{(i-1)} c_1^{e(i-1)} ,$$

$$\frac{\partial c_2^{(i)}}{\partial \tau} = a_3^{(i)} c_1^{(i)} - a_4^{(i)} c_2^{(i)} + a_{\lambda 2}^{(i-1)} c_2^{(i-1)} ; \qquad (3б)$$

для $i = N$

$$\frac{\partial c_1^{(N)}}{\partial \tau} = d^{(N)} \Delta_\xi c_1^{(N)} - a_1^{(N)} c_1^{(N)} + a_2^{(N)} c_2^{(N)} + \sum_{i=0}^{N-1} a_{\lambda 1}^{(iN)} c_1^{(i)} ,$$

$$\frac{\partial c_2^{(N)}}{\partial \tau} = a_3^{(N)} c_1^{(N)} - a_4^{(N)} c_2^{(N)} + \sum_{i=0}^{N-1} a_{\lambda 2}^{(iN)} c_2^{(i)} . \qquad (3в)$$

де

$$\Delta_\xi = \vec{\nabla}_\xi \cdot \vec{\nabla}_\xi , \ \vec{\nabla}_\xi = \vec{i}^{(\alpha)} \partial / \partial \xi^{(\alpha)} ; \ d^{(i)} = D^{(i)} / D^{(0)} ,$$
$$a_1^{(i)} = \left(k_1^{(i)} + \lambda_1^{(i+1)} + \lambda_1^{(iN)}\right) / k_2^{(0)} , \ a_3^{(i)} = k_1^{(i)} / k_2^{(0)} ,$$
$$a_{\lambda j}^{(i-1)} = \lambda_j^{(i-1)} / k_2^{(0)} , \ a_{\lambda j}^{(iN)} = \lambda_j^{(iN)} / k_2^{(0)} , \ i = \overline{0, N} , \ j = 1,2 .$$

Зазначимо, що безрозмірні змінні (2) стискають часову вісь і розтягують просторові. При цьому не залежать від геометричних розмірів тіла або характерних розмірів структурних неоднорідностей.

III. Крайова задача дифузії у середовищі з пастками каскадного типу

Нехай домішкова речовина мігрує в шарі безрозмірної товщини ξ_0 (рис. 1), причому в тілі наявні пастки для домішкових частинок. Домішка може розпадатися, утворюючи нову речовину. Такі процеси описують рівняння (3) для одновимірного за просторовою координатою випадку.

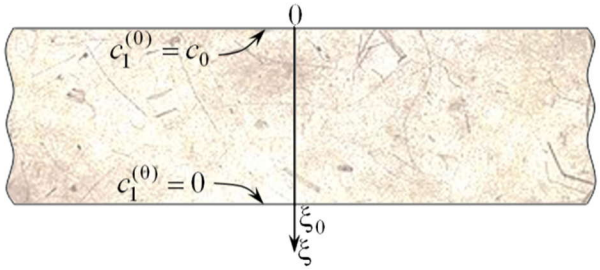


Рис. 1. Шар товщини ξ_0, в якому мігрує домішкова речовина

Приймаємо, що в початковий момент часу в тілі відсутня домішкова речовина:

$$c_1^{(i)}(\xi,\tau)\Big|_{\tau=0} = c_2^{(i)}(\xi,\tau)\Big|_{\tau=0} = 0 , \quad i = \overline{0; N} . \qquad (4)$$

Для часів $\tau > 0$ на межі шару $\xi = 0$ діє стале джерело маси для речовини $\mathbf{K}^{(0)}$ і частинки з поверхні можуть попадати лише у поровий простір. На границі $\xi = \xi_0$ підтримується нульове значення концентрації $c_1^{(0)}$. Також приймаємо, що джерелом маси для речовини $\mathbf{K}^{(i)}$ ($i = \overline{1, N}$) є розпад речовини $\mathbf{K}^{(i-1)}$. Тоді маємо такі граничні умови:

$$c_1^{(0)}(\xi,\tau)\Big|_{\xi=0} = c_0 \equiv const , \quad c_1^{(0)}(\xi,\tau)\Big|_{\xi=\xi_0} = 0 ; \qquad (5а)$$

$$c_1^{(i)}(\xi,\tau)\Big|_{\xi=0} = 0 , \ c_1^{(i)}(\xi,\tau)\Big|_{\xi=\xi_0} = 0 , \ i = \overline{1; N} . \qquad (5б)$$

Таким чином ми отримала крайову задачу (3)-(5) каскадного типу.

Розв'язок крайової задачі для кроку $i = 0$ розпаду дифундуючої речовини, тобто задачі (3а), (4), (5а), знаходимо з допомогою інтегральних перетворень Лапласа за часом і Фур'є [9]. У результаті одержимо

концентрація домішок у поровому просторі

$$\frac{1}{c_0} c_1^{(0)}(\xi,\tau) = 1 - \frac{\xi}{\xi_0} - \frac{2\tilde{B}_1^2}{\xi_0 a_4^{(0)} d^{(0)}} \left[1 - \frac{\xi}{\xi_0} - \frac{\mathrm{sh}\eta^{(0)}(\xi_0 - \xi)}{\mathrm{sh}\eta^{(0)}\xi_0}\right] + \frac{2}{\xi_0} \times$$
$$\times \sum_{k=1}^{\infty} \frac{\sin y_k \xi}{y_k (s_1 - s_2)} \left\{ \left(s_1 + \tilde{A}_1 + \frac{\tilde{B}_1}{s_1}\right) e^{s_1 \tau} - \left(s_2 + \tilde{A}_1 + \frac{\tilde{B}_1}{s_2}\right) e^{s_2 \tau} \right\} ; \qquad (6а)$$

концентрація частинок у пастках

$$\frac{1}{c_0} c_2^{(0)}(\xi,\tau) = \frac{2 a_3^{(0)} \tilde{B}_1}{\xi_0 a_4^{(0)2} d^{(0)2}} \frac{\mathrm{sh}\eta^{(0)}(\xi_0 - \xi)}{\mathrm{sh}\eta^{(0)}\xi_0} + \frac{2}{\xi_0} a_3^{(0)} \times$$
$$\times \sum_{k=1}^{\infty} \frac{\sin y_k \xi}{y_k (s_1 - s_2)} \left\{ \left(\frac{\tilde{B}_2}{s_1}\right) e^{s_1 \tau} - \left(\frac{\tilde{B}_2}{s_2}\right) e^{s_2 \tau} \right\} ; \qquad (6б)$$

сумарна концентрація дифундуючих частинок $\tilde{c}^{(0)}(\xi,\tau)=c_1^{e(0)}(\xi,\tau)+c_2^{(0)}(\xi,\tau)$ *у тілі з пастками*

$$\frac{1}{c_0}\tilde{c}^{(0)}(\xi,\tau)=1-\frac{\xi}{\xi_0}-\frac{2\tilde{B}_1^2}{\xi_0 a_4^{(0)}d_0^{(0)}}\times$$

$$\times\left[1-\frac{\xi}{\xi_0}-\left(1-\frac{a_3^{(0)}}{d^{(0)}a_4^{(0)}\tilde{B}_1}\right)\frac{\mathrm{sh}\eta^{(0)}(\xi_0-\xi)}{\mathrm{sh}\eta^{(0)}\xi_0}\right]+\frac{2}{\xi_0}\sum_{k=1}^{\infty}\frac{\sin y_k\xi}{y_k(s_1-s_2)}\times$$

$$\times\left\{\left(s_1+\tilde{A}_1+\frac{\tilde{B}^{(0)}}{s_1}\right)e^{s_1\tau}-\left(s_2+\tilde{A}_1+\frac{\tilde{B}^{(0)}}{s_2}\right)e^{s_2\tau}\right\},\qquad (6в)$$

™∑ $\tilde{B}^{(0)}=a_3^{(0)}\tilde{B}_2-\tilde{B}_1$, $\tilde{A}_1=a_1^{(0)}+a_4^{(0)}$, $\tilde{B}_1=a_1^{(0)}a_4^{(0)}-a_3^{(0)}a_2^{(0)}$,

$\tilde{B}_2=d_0^{(0)}y_k^2$, $s_{1,2}=\left[-\zeta_1\pm\sqrt{\zeta_1^{\,2}-4\zeta_2}\right]\Big/2$,

$\zeta_1=d_0^{(0)}y_k^2+a_1^{(0)}+\ +a_4^{(0)}$, $\zeta_2=a_1^{(0)}a_4^{(0)}-a_3^{(0)}a_2^{(0)}+a_4^{(0)}d_0^{(0)}y_k^2$

, $\eta^{(0)}=\sqrt{\tilde{B}_1/a_4^{(0)}d_0^{(0)}}$

Зазначимо, що концентрація в пастках є пропорційною не тільки значенню концентрації частинок на границі шару c_0, але й коефіцієнту інтенсивності сорбції $a_3^{(0)}$. Асимптотичні складові концентрацій $c_1^{(0)}(\xi,\tau)$, $c_2^{(0)}(\xi,\tau)$, $\tilde{c}^{(0)}(\xi,\tau)$ є суттєво нелінійними.

Зауважимо, що в рамках моделі, яка розглядається, концентрація частинок у пастках на границі шару $\xi=0$ не дорівнює нулю, і, відповідно, сумарна концентрація дифундуючої речовини на цій межі не дорівнює c_0:

$$\frac{1}{c_0}c_2^{(0)}(\xi,\tau)\Big|_{\xi=0}=\frac{2a_3^{(0)}\tilde{B}_1}{\xi_0 a_4^{(0)2}d^{(0)2}};\ \frac{1}{c_0}\tilde{c}^{(0)}(\xi,\tau)\Big|_{\xi=0}=1+\frac{2\tilde{B}_1 a_3^{(0)}}{\xi_0 a_4^{(0)2}d^{(0)2}}.$$

Це означає, що домішкова речовина може накопичуватись у пастках і на границі шару $\xi=0$. У той же час на границі $\xi=\xi_0$ концентрації розпадних частинок (6а)-(6в) дорівнюють нулю.

Для інших кроків каскадного розпаду мігруючих речовин ($i=\overline{1,N-1}$) розв'язок крайових задач (3б), (4), (5б) подамо через відповідні функції Гріна:

$$c_1^{(i)}(\xi,\tau)=a_{\lambda 1}^{(i-1)}\int_0^{\tau}\int_0^{\xi_0}G_1^{(i)}\left(\xi,\xi';\tau,\tau'\right)c_1^{(i-1)}\left(\xi',\tau'\right)d\xi'd\tau',$$

$$c_2^{(i)}(\xi,\tau)=a_{\lambda 2}^{(i-1)}\int_0^{\tau}\int_0^{\xi_0}G_2^{(i)}\left(\xi,\xi';\tau,\tau'\right)c_2^{(i-1)}\left(\xi',\tau'\right)d\xi'd\tau',\ (7)$$

де $G_1^{(i)}\left(\xi,\xi';\tau,\tau'\right)$, $G_2^{(i)}\left(\xi,\xi';\tau,\tau'\right)$ - функції Гріна задачі (3б), (4), (5б) для $i=\overline{1,N-1}$, тобто є розв'язками наступних крайових задач

$$\frac{\partial G_1^{(i)}}{\partial\tau}-d^{(i)}\frac{\partial^2 G_1^{(i)}}{\partial\xi^2}+a_1^{(i)}G_1^{(i)}-a_2^{(i)}G_2^{(i)}=\delta\left(\tau-\tau'\right)\delta\left(\xi-\xi'\right),$$

$$\frac{\partial G_2^{(i)}}{\partial\tau}-a_3^{(i)}G_1^{(i)}+a_4^{(i)}G_2^{(i)}=\delta\left(\tau-\tau'\right)\delta\left(\xi-\xi'\right)\qquad (8)$$

за нульових крайових умов

$$G_1^{(i)}(\xi,\xi';\tau,\tau')\Big|_{\tau=0}=G_2^{(i)}(\xi,\xi';\tau,\tau')\Big|_{\tau=0}=0,$$

$$G_1^{(i)}(\xi,\xi';\tau,\tau')\Big|_{\xi=0}=G_1^{(i)}(\xi,\xi';\tau,\tau')\Big|_{\xi=\xi_0}=0,\ i=\overline{1,N-1}.(9)$$

Розв'язок крайової задачі (8), (9) також шукаємо з допомогою інтегральних перетворень Лапласа і Фур'є [10, 11]. Формули для визначення функцій Гріна $G_j^{(i)}$ на i-му кроці розпаду ($i=\overline{1,N-1}$) одержано у вигляді:

$$G_1^{(i)}=\theta\left(\tau-\tau'\right)\frac{2}{\xi_0}\sum_{n=1}^{\infty}\frac{\sin\left(y_n\xi\right)\sin\left(y_n\xi'\right)}{s_{1n}-s_{2n}}\times$$

$$\times\left[\left(s_{1n}+\overline{A}_1^{(i)}\right)e^{s_{1n}(\tau-\tau')}-\left(s_{2n}+\overline{A}_1^{(i)}\right)e^{s_{2n}(\tau-\tau')}\right].$$

$$G_2^{(i)}=\theta\left(\tau-\tau'\right)\frac{2}{\xi_0}\sum_{n=1}^{\infty}\frac{\sin\left(y_n\xi\right)\sin\left(y_n\xi'\right)}{s_{1n}-s_{2n}}\times$$

$$\times\left[\left(s_{1n}+\overline{A}_2^{(i)}\right)e^{s_{1n}(\tau-\tau')}-\left(s_{2n}+\overline{A}_2^{(i)}\right)e^{s_{2n}(\tau-\tau')}\right].\qquad (10)$$

де $\overline{A}_1^{(i)}=a_{33}^{(i)}+a_{23}^{(i)}$, $\overline{A}_2^{(i)}=d_0^{e(i)}y_n^2+a_{11}^{e(i)}+a_{31}^{e(i)}$.

Отже, знайшовши функцію Гріна за формулою (10) і визначивши концентрацію розпадних домішкових частинок на кроці $i-1$, за співвідношеннями (7) ми знайдемо концентрацію розпадної речовини для $i=\overline{1,N-1}$.

Для випадку $i=N$ (нерозпадні домішки) процес масоперенесення описується задачею (3в), (4), (5б). Її розв'язок також подамо через відповідні функції Гріна

$$c_1^{(N)}(\xi,\tau)=\int_0^{\tau}\int_0^{\xi_0}G_1^{(N)}\left(\xi,\xi';\tau,\tau'\right)\sum_{i=0}^{N-1}a_{\lambda 1}^{(iN)}c_1^{(i)}\left(\xi',\tau'\right)d\xi'd\tau',$$

$$c_2^{(N)}(\xi,\tau)=\int_0^{\tau}\int_0^{\xi_0}G_2^{(N)}\left(\xi,\xi';\tau,\tau'\right)\sum_{i=0}^{N-1}a_{\lambda 2}^{(iN)}c_2^{(i)}\left(\xi',\tau'\right)d\xi'd\tau',(11)$$

де $G_j^{(N)}(\xi,\xi';\tau,\tau')$ ($j=1,2$) - функції Гріна задачі (3в), (4), (5б) і задовольняють крайову задачу з точковим джерелом

$$\frac{\partial G_1^{(N)}}{\partial\tau} - d^{(N)}\frac{\partial^2 G_1^{(N)}}{\partial\xi^2} + a_1^{(N)}G_1^{(N)} - a_2^{(N)}G_2^{(N)} = \delta(\tau-\tau')\delta(\xi-\xi'),$$

$$\frac{\partial G_2^{(N)}}{\partial\tau} - a_3^{(N)}G_1^{(N)} + a_4^{(N)}G_2^{(N)} = \delta(\tau-\tau')\delta(\xi-\xi'); \quad (12)$$

$$G_1^{(N)}\Big|_{\tau=0} = G_2^{(N)}\Big|_{\tau=0} = G_1^{(N)}\Big|_{\xi=0} = G_1^{(N)}\Big|_{\xi=\xi_0} = 0. \quad (13)$$

Розв'язок крайової задачі (12), (13) має вигляд

$$G_1^{(N)} = \theta(\tau-\tau')\frac{2}{\xi_0}\sum_{n=1}^{\infty}\frac{\sin(y_n\xi)\sin(y_n\xi')}{s_{1n}-s_{2n}}\times$$

$$\times\left[(s_{1n}+\overline{A}_1^{(N)})e^{s_{1n}(\tau-\tau')} - (s_{2n}+\overline{A}_1^{(N)})e^{s_{2n}(\tau-\tau')}\right],$$

$$G_2^{(N)} = \theta(\tau-\tau')\frac{2}{\xi_0}\sum_{n=1}^{\infty}\frac{\sin(y_n\xi)\sin(y_n\xi')}{s_{1n}-s_{2n}}\times$$

$$\times\left[(s_{1n}+\overline{A}_2^{(N)})e^{s_{1n}(\tau-\tau')} - (s_{2n}+\overline{A}_2^{(N)})e^{s_{2n}(\tau-\tau')}\right]. \quad (14)$$

Тоді, послідовно визначивши концентрації на кожному кроці $i=\overline{0,N-1}$ за формулами (11) з урахуванням виразів (14) знайдемо концентрації нерозпадних частинок у тілі.

IV. Схема модуля пакету програм. Числовий аналіз

На основі формул (6), (7), (10), а також розв'язків задач (8), (9) і (12), (13) розроблений пакет програм для комп'ютерного моделювання процесів дифузії у тілі з пастками за каскадного розпаду мігруючих частинок. Схему алгоритму модуля пакету для розрахунку концентрацій розпадних мігруючих частинок для всіх $i=\overline{0,N}$ наведено на рис. 2.

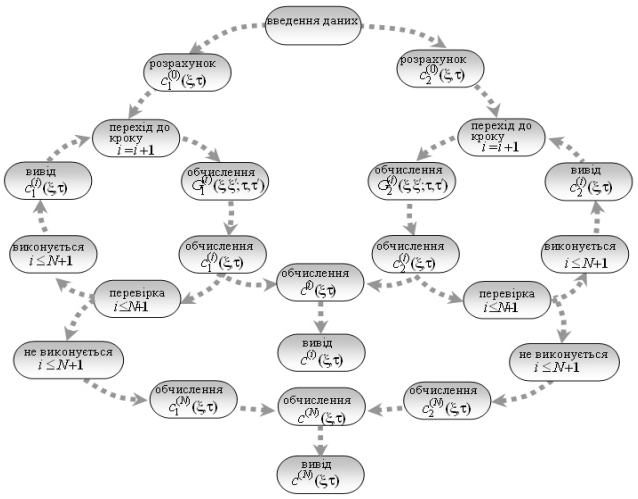


Рис. 2. Схема алгоритму модуля пакету для розрахунку концентрацій домішкових речовин за каскадного розпаду частинок у тілі з пастками

На рис. 3 показано розподіли концентрації домішок на етапі $i=0$ у поровому просторі (криві - 1a-3a), у пастках (криві – 1b-3b) та сумарна концентрація частинок (криві – 1-3) в різні моменти часу $\tau=0{,}01$; 2; 20 для коефіцієнта інтенсивності сорбції $a_3^{(0)}=5$. Рис. 4. ілюструє розподіл сумарної концентрації утвореної після розпаду домішки ($i=1$) при різних значеннях коефіцієнта інтенсивності сорбції $a_3^{(1)}=2$; 3; 4 для коефіцєнта дифузії утвореної речовини $d_0^{(1)}=0{,}05$ (криві a) та $d_0^{(1)}=20$ (криві b).

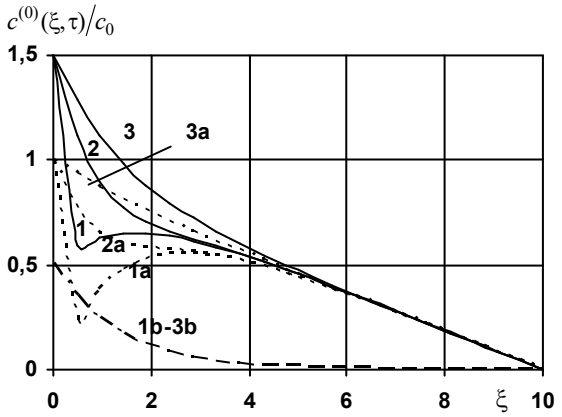


Рис. 3. Розподіли концентрацій $c_1^{(0)}$, $c_2^{(0)}$ та $\tilde{c}^{(0)}$ в різні моменти часу

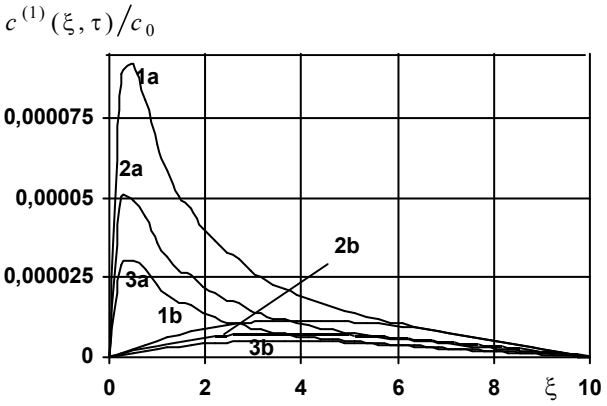


Рис. 4. Розподіли концентрації $\tilde{c}^{(1)}$ при різних значеннях $a_3^{(1)}$

Зазначимо, що для етапу $i = 0$ характерне суттєве при поверхневе зменшення сумарної концентрації $\tilde{c}^{(0)}$ для малих часів (крива 1, рис. 3) проте з часом сумарна концентрація зростає в цій області (крива 2, рис. 3) і виходить на усталений режим (крива 3, рис. 3), набуваючи монотонно спадного характеру.

Зауважимо, що поведінка функції сумарної концентрації на етапі розпаду $i = 1$ суттєво залежить від значення відносного коефіцієнта дифузії $d_0^{(1)} = D_1 / D_0^{(0)}$ (рис. 4) при цьому для великих $d_0^{(1)}$ характерно повільне зростання концентрації $\tilde{c}^{(1)}$ від поверхні з виходом на максимальне значення в середині тіла (криві b, рис. 4). У той же час для $D_1 < D_0$ спостерігається значне накопичення домішкової речовини $\mathbf{K}^{(1)}$ в околі поверхні тіла $\xi = 0$, де діє джерело маси для речовини $\mathbf{K}^{(0)}$ (криві a, рис. 4).

V. Висновки

Таким чином для опису процесів масоперенесення домішкових речовин за їх каскадного розпаду у середовищі з пастками сформована математична модель за континуально-термодинамічним підходои. Сформульовані зв'язані крайові задачі дифузії для тіла з пастками, коли концентрація частинок на певному кроці розпаду є джерелом маси розпадної речовини на наступному кроці, яка теж дифундує, сорбується, десорбується і розпадається.

Розв'язки відповідних крайових задач каскадного типу побудовані за ітераційною процедурою з використанням функцій Гріна. Таке подання розв'язків дозволяє не тільки проводити кількісний та якісний аналіз концентрацій розпадних домішок, а й знайти їхні потоки маси, як і визначити кількість відповідних речовин, що за певний проміжок часу пройшли через одиницю площі деякої поверхні, наприклад, через нижню границю шару. На основі отриманих формул розроблений пакет програм для комп'ютерного моделювання процесів дифузії у тілі з пастками за каскадного розпаду домішкових речовин. Встановлено вплив вхідних даних на поведінку і значення сумарних концентрацій частинок на всіх етапах каскадного розпаду.

Література References

[1] *Н. Н. Мельников,* Пестициды. Химия, технология и применение. – М., Химия, 1987. – 712 с.

[2] *Ю. Білущак, В. Гончарук, Є. Чапля, О. Чернуха,* Математичне моделювання дифузії домішкових компонент за їх каскадного розпаду // Математичні машини і системи. – 2015. – № 1. – С. 146-155.

[3] *В. Колобашкин, П. Рубцов, П. Ружанский, В. Сидоренко* Радиационные характеристики облученного ядерного топлива. – М., Энергоатомиздат, 1983. - 384 с.

[4] *Ю. Білущак, В. Гончарук, О. Чернуха,* Математична модель невзаємодіючих потоків для опису процесів масопереносу двома шляхами за каскадного розпаду частинок // Прикладні проблеми механіки і математики. – 2014. - Вип. 12. – С. 137-145.

[5] *А. Большов, В. Головизнин, А. Дыхне, В. Киселев, П. Кондратенко, В. Семенов,* Новые подходы к оценке безопасности захоронений радиоактивных отходов // Извест. РАН. Энергетика. -2004. - №4. - С. 99-108.

[6] *В. Головизнин, В. Киселев, И. Короткин, Ю. Юрков,* Прямые задачи неклассического переноса радионуклидов в геологических формациях // Известия РАН. Энергетика. - 2004. - №4. - С. 121-130.

[7] *М. Моисеев, И. Завершинский,* Диффузия в среде со случайно распределенными ловушками // Матем. моделирование и краев. задачи. - 2005, часть 2. – С. 185-187.

[8] *Я. Бурак, Є. Чапля, О. Чернуха,.* Континуально-термодинамічні моделі механіки твердих розчинів.– Київ, Наукова думка. - 2006. – 272 с.

[9] *Є. Я. Чапля, О. Ю. Чернуха,* Фізико-математичне моделювання гетеродифузного масопереносу. - Львів: СПОЛОМ. - 2003. – 128 с.

[10] *Y. Chaplya, O. Chernukha, Y. Bilushchak,* Contact initial boundary-value problem of the diffusion of admixture particles in a two-phase stochastically inhomogeneous stratified strip // Journal of Mathematical Sciences. – 2012. – Vol. 183, No 1. – P. 83-99

[11] *И. Снеддон* Преобразования Фурье. - М: Изд-во иностр. лит-ры, 1955. - 667 с.

Математичне та Комп'ютерне Моделювання Процесів Конвективної Дифузії у Двошарових Засипних Фільтрах Води

О. Ю. Чернуха, Ю. І Білущак
Відділ математичного моделювання нерівноважних процесів
Центр математичного моделювання
Інституту прикладних проблем механіки і математики ім. Я.С. Підстригача НАН України
Львів, Україна
zaliznuchna6@gmail.com, byixx13@gmail.com

Mathematical Modelling and Simulation of Convective Diffusion Processes in Two-layered Filled-up Filters of Water

O. Chernukha, Y. Bilushchak
Department of mathematical modeling nonequilibrium processes
Centre of Mathematical Modeling of Pidstryhach Institute for Applied Problems of Mechanics and Mathematics,
Ukrainian Academy of Sciences
Lviv, Ukraine
zaliznuchna6@gmail.com, byixx13@gmail.com

***Анотація*—Побудовано математичну модель для опису конвективної дифузії у двошаровому фільтрі. Розроблено метод знаходження точних розв'язків контактно-крайових задач масоперенесення з урахуванням дифузійного і конвективного механізмів переносу та сорбційних процесів, який базується на застосуванні інтегральних перетворень за просторовими змінними окремо в контактуючих областях. Побудовано аналітичні розв'язки задачі конвективної дифузії для двошарового фільтра. Розвинено метод числового інтегрування подвійного інтеграла зі змінними верхніми межами. Розроблено комплекс програм та встановлено закономірності процесу конвективної дифузії у двошаровій області.**

***Abstract*—The mathematical model is constructed for description of convective diffusion in a two-layered filter. It is developed the method for determining the exact solutions of contact initial-boundary value problems of mass transfer with provision for both diffusive and convective mechanisms of transfer as well as sorption processes, which is based on integral transformations over space variables in the contacting regions. The analytical solutions are constructed for the problem of convective diffusion for the two-layered filter. The method of numerical calculation of the double integral with variable top limits is developed. Is it designed software and established regularities of convective diffusion process in the two-layered region.**

Ключові слова—математичне і комп'ютерне моделювання;конвективна дифузія;двофазний фільтр води;інтегральне перетворення;чисельне інтегрування

Keywords—simulation: mathematical modelling; simulation; convective diffusion; two-phase water filter; integral transformation; numerical integration

I. Вступ

На сьогодні серед актуальних завдань важливе місце займає проблема очистки води до стану питної, а також очистка побутових та промислових каналізаційних стоків для вторинного використання. Одним з різновидів механічних фільтрів є насипні фільтри [1, 2], що складаються із зернистих засипок різних фільтраційних властивостей. У таких конструкціях використовують: кварцовий пісок різних розмірів фракцій, гідроантрацит, алюмосилікат, сульфовугілля тощо (рис. 1). Неперемішуваність фільтруючих прошарків забезпечується значною різницею їх густин.

Для очищення води у промислових масштабах зазвичай застосовують каркасно-насипні фільтри, які є різновидом багатошарових фільтрів [1, 3], і слугують для доочищення стічних вод та затримання грубих, колоїдних та розчинених часток [4], зокрема очистки від таких забруднювачів як нафтопродукти, мідь, залізо, цинк тощо.

Рис. 1. Види засипок: алюмосилікат (a), кварцовий пісок (b), сульфовугілля (c)

Дана робота присвячена математичному моделюванню функціонування двошарових фільтрів води на основі математичної моделі міграції домішкової речовини, коли водний розчин є двокомпонентним, а тіло – двошаровим. При цьому запропоновано підхід до математичного опису процесів конвективної дифузії домішкових частинок, які супроводжуються їхньою сорбцію на скелет, у двофазному багатокомпонентному тілі, що складається з етапів:

- побудова математичної моделі у лінеаризованому варіанті; формулювання на цій основі контактно-крайової задачі з урахуванням умов неідеального контакту на концентрацію забруднення, яке переноситься з конвективно рухомим розчином;
- знаходження точного аналітичного розв'язку сформульованої задачі з допомогою методу інтегральних перетворень окремо в різних пористих шарах фільтра;
- комп'ютерне моделювання отриманих розв'язків, яке містить такі підетапи:
  - побудова алгоритмів для кожної структурної частини розв'язку контактно-крайової задачі для чисельного аналізу:
    - концентрацій забруднюючих частинок у водному поровому розчині в обох шарах фільтра та відповідних потоків маси з використанням відомого методу чисельного інтегрування, а саме Ньютона-Котеса замкненого типу за 7 і 10 вузлами;
    - концентрації частинок забруднення сорбованих на скелеті фільтра, для чого розробляємо новий метод чисельного інтегрування подвійного інтеграла зі змінними верхніми межами;
    - знаходження часу насичення фільтра на основі запропонованого нами алгоритму розв'язання нелінійного функціонального рівняння на відрізку невідомої довжини;
  - розробка програмного забезпечення та використання його для аналізу отриманих результатів.

II. Математична модель

При формулюванні вихідних співвідношень моделі переносу забруднення в двошаровому фільтрі вважаємо, що довільна область кожного шару складається зі скелету та водного розчину, який заповнює поровий простір. Приймемо, що в процесі фільтрації скелет не деформується, і пористість залишається сталою (не враховуються зміни, пов'язані із сорбцією домішкової речовини). Водний розчин є двокомпонентним і складається з частинок води та забруднюючої субстанції. Частинки забруднення знаходяться у двох станах – конвективно рухомому розчині і поверхні скелету.

Основними процесами, що розглядаються, є конвективна дифузія домішок та їхня сорбція скелетом. Ці процеси описуються з використанням наближення континуму центрів мас [5] для рідкої фази, приймаючи швидкість конвективного руху частинок $\vec{v}_\xi$ наближено рівною істиній швидкості порового розчину при фільтрації. Середовище вважається ізотропним, і поверхнева та об'ємна пористості покладаються рівними.

Рівняння балансу для порового розчину записано у вигляді [5]

$$\rho \frac{dc_i(t,\vec{\xi})}{dt} = -\vec{\nabla}\cdot\vec{J}_i + \sigma_{mi}, \quad i=\overline{0;2},$$

де $\rho = \sum_i \rho_i$ - сумарна густина розчину, ρ_i - густина компоненти i, $c_i = \rho_i/\rho$ — масова концентрація компоненти i, $\vec{J}_i$ - дифузійний потік i-ї складової розчину, σ_{mi} - потужність виробництва маси i-ї компоненти; $d/dt = \partial/\partial t + \vec{v}_\xi\cdot\vec{\nabla}$ - повна похідна за часом.

Водній компоненті розчину відповідає значення індексу $i=0$, забруднюючій речовині у розчині - $i=1$, а у зв'язаному стані - $i=2$. При цьому $\sigma_{m0}=0$.

Вважаємо, що дифузійний потік визначається градієнтами хімічних потенціалів μ_i, які лінійно залежать від концентрацій, тобто

$$\vec{J}_i = -L\vec{\nabla}\mu_i = -L^*\vec{\nabla}c_i, \quad i=\overline{0;2},$$

де L, L^* - кінетичні коефіцієнти переносу.

Потужність виробництва маси компоненти i є пропорційна до локальної різниці хімічних потенціалів μ_i ($i=1;2$) і, як наслідок, різниці концентрацій c_i. При лінійній концентраційній залежності хімічних потенціалів, якщо знехтувати зворотним переходом частинок забруднюючої речовини зі зв'язаного стану в розчин, то $\sigma_{m1} = -k^*c_1$, $\sigma_{m2}=0$, $\sigma_{m0}=0$, де k^* - кінетичний коефіцієнт процесу сорбції.

Концентрації c_i та потоки $\vec{J}_i$ ($i=\overline{0;2}$) задовольняють умови нормування $\sum_i c_i = 1$ і $\sum_i \vec{J}_i = 0$. Перенос у зв'язаному стані відсутній. Густина ρ і коефіцієнти L^* та k^* є сталими. Тоді концентрації c_1 і c_2 визначаємо з рівнянь

$$\frac{\partial c_1(t,\vec{\xi})}{\partial t} = D\Delta c_1(t,\vec{\xi}) - \vec{v}_\xi\cdot\vec{\nabla}c_1(t,\vec{\xi}) - kc_1(t,\vec{\xi}), \qquad (1)$$

$$\frac{\partial c_2(t,\vec{\xi})}{\partial t} = kc_1(t,\vec{\xi}), \qquad (2)$$

де $D = L^*/\rho_0$ - коефіцієнт дифузії домішки, $k = k^*/\rho_0$ - приведений кінетичний коефіцієнт процесу сорбції, ρ_0 - стале значення густини розчину.

Приведемо систему рівнянь (1), (2) до безрозмірного вигляду введенням нових змінних [5]

$$\tau = k_* t\,, \quad x = (k_*/D_*)^{1/2}\xi\,,$$

де $k_* = \min\{k_1; k_2\}$, $D_* = \max\{D_1; D_2\}$ для розглядуваних шарів фільтра. Тоді отримаємо

$$\frac{\partial c_1(\tau,\vec{r})}{\partial\tau} = d\Delta_r c_1(\tau,\vec{r}) - \vec{v}\cdot\vec{\nabla}_r c_1(\tau,\vec{r}) - a c_1(\tau,\vec{r})\,,$$

$$\frac{\partial c_2(\tau,\vec{r})}{\partial\tau} = a c_1(\tau,\vec{r})\,,$$

де $d = D/D_*$, $\vec{v} = (k_* D_*)^{1/2}\vec{v}_\xi$, $a = k/k_*$ - безрозмірні коефіцієнти. Причому потік у безрозмірній формі буде

$$\vec{J} = -d\vec{\nabla}_r c_1 + \vec{v} c_1\,. \qquad (3)$$

III. Постановка контактно-крайової задачі

Розглянемо шар безрозмірної товщини x_0, що складається з двох підшарів товщин x' та δx ($\delta\xi = x_0 - x'$) відповідно. Система декартових координат вибрана так, щоб вісь Ox була перпендикулярна до поверхонь шару з початком на верхній границі і спрямована вглиб тіла. Нехай на верхній і нижній поверхнях тіла відомі сталі значення концентрації домішки $c_1^{(1)}(t,x)$ і $c_1^{(2)}(t,x)$ (нижній індекс позначає стан частинок - $i=1$ відповідає розчину, $i=2$ - поверхні скелету; верхній індекс позначає номер підшару фільтра: $j=1$ - для $\Omega_1 =]0; x'[$, $j=2$ - для $\Omega_2 =]x'; x_0[$:

$$c_1^{(1)}(t,x)\Big|_{x=0} = c_0 \equiv \text{const}\,,\ c_1^{(2)}(t,x)\Big|_{x=x_0} = c_* \equiv \text{const}\,.$$

Також приймаємо, що в початковий момент часу

$$c_1^{(1)}(\tau,\xi)\Big|_{\tau=0} = c_2^{(1)}(\tau,\xi)\Big|_{\tau=0} = 0 \text{ (область } \Omega_1\text{)}; \qquad (4)$$

$$c_1^{(2)}(\tau,\xi)\Big|_{\tau=0} = c_2^{(2)}(\tau,\xi)\Big|_{\tau=0} = 0 \text{ (область } \Omega_2\text{)}. \qquad (5)$$

У випадку вертикальної конвективної дифузії масоперенесення домішок описують такі системи рівнянь [5]
у першому шарі (область Ω_1):

$$\frac{\partial c_1^{(1)}(t,x)}{\partial t} = d_1\frac{\partial^2 c_1^{(1)}(t,x)}{\partial x^2} - v_1\frac{\partial c_1^{(1)}(t,x)}{\partial x} - a_1 c_1^{(1)}(t,x)\,, \quad (6)$$

$$\frac{\partial c_2^{(1)}(t,x)}{\partial t} = a_1 c_1^{(1)}(t,x)\,, \quad x\in\Omega_1\,, \qquad (7)$$

у другому шарі (область Ω_2):

$$\frac{\partial c_1^{(2)}(t,x)}{\partial t} = d_2\frac{\partial^2 c_1^{(2)}(t,x)}{\partial x^2} - v_2\frac{\partial c_1^{(2)}(t,x)}{\partial x} - a_2 c_1^{(2)}(t,x)\,, \quad (8)$$

$$\frac{\partial c_2^{(2)}(t,x)}{\partial t} = a_2 c_1^{(2)}(t,x)\,, \quad x\in\Omega_2\,, \qquad (9)$$

де d_j і v_j - коефіцієнти дифузії домішки і конвективного перенесення в Ω_j, a_j - коефіцієнт інтенсивності сорбції.

На границі контакту шарів $x = x'$ виконується умова рівностей хімічних потенціалів і сумарних масових потоків, які запишемо вигляді

$$\lambda c_1^{(1)}(t,x)\Big|_{x=x'} = c_1^{(2)}(t,x)\Big|_{x=x'}\,; \qquad (10)$$

$$d_1\frac{\partial c_1^{(1)}}{\partial x} - v_1 c_1^{(1)}\Big|_{x=x'} = d_2\frac{\partial c_1^{(2)}}{\partial x} - v_2 c_1^{(2)}\Big|_{x=x'}\,, \qquad (11)$$

де $\lambda = \lambda_1/\lambda_2$ - відношення коефіцієнтів концентраційної залежності хімічних потенціалів в станах 1 і 2, які визначають стрибок концентрації на границі контакту.

IV. Метод побудови розв'язку інтегральними перетвореннями

Другі рівняння систем (7) і (9) можна проінтегрувати. З урахуванням початкових умов (4), (5) маємо

$$c_2^{(j)}(t,x) = a_j\int_0^t c_1^{(j)}(t',x)dt'\,.$$

Для побудови точного розв'язку задачі (4)-(11) розвинуто метод на основі інтегральних перетворень за просторовою змінною окремо в контактуючих областях Ω_1 і Ω_2 [6]. Для цього на поверхні $x = x'$ доозначимо шукані функції за допомогою першої контактної умови (10), тобто приймемо, що функції $\lambda c_1^{(1)}$ та $c_1^{(2)}$ є рівні між собою і дорівнюють деякій функції часу $g(\tau)$:

$$\lambda c_1^{(1)}(\tau,x)\Big|_{x=x'} = c_1^{(2)}(\tau,x)\Big|_{x=x'} = g(\tau,x') \equiv g(\tau)\,.$$

Враховуючи вигляд операторів рівнянь (6)-(9), а також те, що задано граничні умови 1-го роду, в області Ω_1 застосуємо таке скінченне інтегральне перетворення [7]:

$$\bar{c}_1(n,\tau) = \int_0^{x'} c_1^{(1)}(x,\tau)\exp\left\{-\frac{v_1 x}{2d_1}\right\}\sin(y_n x)dx\,; \qquad (12)$$

$$c_1^{(1)}(x,\tau)=\frac{2}{x'}\exp\left\{\frac{v_1 x}{2d_1}\right\}\sum_{n=1}^{\infty}\bar{c}_1(n,\tau)\,\sin\left(y_n x\right),\ n=1,2,\ldots. \quad (13)$$

де $y_n=n\pi/x'$ □

Після застосування перетворення (12) крайова задача в зображеннях в області Ω_1 набуде вигляду

$$\frac{d\bar{c}_1}{dt}=-\left(d_1 y_n^2+a_1\right)\bar{c}_1+d_1 y_n c_0-d_1 y_n\frac{(-1)^n}{\lambda}g(\tau)e^{-\frac{v_1 x'}{2d_1}},$$

$$\bar{c}_1(n,\tau)\big|_{\tau=0}=0.$$

Її розв'язок знаходимо у формі [8]

$$\bar{c}_1(n,\tau)=e^{-(d_1 y_n^2+a_1)\tau}\int_0^{\tau}d_1 y_n^2\left[c_0-\frac{(-1)^n}{\lambda}g(\tau')\exp\left\{-\frac{v_1 x'}{2d_1}\right\}\right]\times$$
$$\times e^{(d_1 y_n^2+a_1)\tau'}d\tau'. \quad (14)$$

Тепер розглянемо крайову задачу конвективної дифузії в області Ω_2. Введемо скінченне інтегральне перетворення за змінною x таким чином:

$$\bar{c}_2(m,\tau)=\int_{x'}^{x_*}c_1^{(2)}(x,\tau)\exp\left\{-\frac{v_2(x-x')}{2d_2}\right\}\sin\left(y_m(x-x')\right)dx,$$
$$m=1,2,\ldots; \quad (15)$$

де $y_m=m\pi/\delta x$. Формула оберненого перетворення до (15) отримана у вигляді

$$c_1^{(2)}(x,\tau)=\frac{2}{\delta x}\exp\left\{\frac{v_2(x-x')}{2d_2}\right\}\sum_{m=1}^{\infty}\bar{c}_2(m,\tau)\,\sin\left(y_m(x-x')\right). \quad (16)$$

Застосовуючи перетворення (15), отримаємо крайову задачу в зображеннях в області Ω_2:

$$\frac{d\bar{c}_2}{d\tau}=-\left(d_2 y_m^2+a_2\right)\bar{c}_2+d_2 y_m\left[g(\tau)-(-1)^m c_* e^{-\frac{v_2\delta x}{2d_2}}\right],$$

$$\bar{c}_2(m,\tau)\big|_{\tau=0}=0,$$

розв'язок якої знаходимо у вигляді [8]

$$\bar{c}_2(m,\tau)=e^{-(d_2 y_m^2+a_2)\tau}\int_0^{\tau}d_2 y_m^2\left[g(\tau')-(-1)^m c_*\exp\left\{-\frac{v_2\delta x}{2d_2}\right\}\right]\times$$

$$\times e^{(d_2 y_m^2+a_2)\tau'}d\tau'. \quad (17)$$

У виразах (14) і (17) залишається невідомою функція $g(\tau)$. Шукаємо її з другої контактної умови рівності потоків на границі розділу областей (11). Після виконання обернених перетворень (13) і (16) виразів (14) і (17), підставляємо формули для концентрацій та їх похідних в точці $x=x'$ у співвідношення (10). Тоді одержимо рівняння

$$d_1\int_0^{\tau}\frac{2d_1}{x'}\sum_{n=1}^{\infty}y_n^2\left\{(-1)^n e^{\frac{v_1 x'}{2d_1}}c_0-\frac{g(\tau')}{\lambda}\right\}e^{-(d_1 y_n^2+a_1)(\tau-\tau')}d\tau'=$$

$$=d_2\int_0^{\tau}\frac{2d_2}{\delta x}\sum_{m=1}^{\infty}y_m^2\left\{(-1)^m g(\tau')-c_* e^{-\frac{v_2\delta x}{2d_2}}\right\}e^{-(d_2 y_m^2+a_2)(\tau-\tau')}d\tau', \quad (18)$$

яке розв'язуємо відносно функції $g(\tau')$.

Інтегральне рівняння (18) має неєдиний розв'язок, оскільки існують такі функції $F(\tau')$, що $\int_0^{\tau}F(\tau')d\tau'=0$. У той же час розв'язок вихідної задачі є єдиним незалежно від способу розв'язання рівняння (18), оскільки функція $g(\tau')$ у виразах для $\bar{c}_1(n,\tau)$ і $\bar{c}_2(m,\tau)$ фігурує тільки під знаком інтеграла за змінною τ'.

У результаті отримаємо вираз для функції $g(\tau')$:

$$g(\tau')=\frac{A_1 c_0\exp\left\{\frac{v_1 x'}{2d_1}\right\}\Sigma_n^- + A_2 c_*\exp\left\{-\frac{v_2\delta x}{2d_2}\right\}\Sigma_m^+}{A_1\lambda^{-1}\Sigma_n^+ + A_2\Sigma_m^-},$$

де $A_1=2d_1^2/x'$, $\Sigma_n^{\pm}=\sum_{n=1}^{\infty}(\pm1)^n y_n^2\,e^{-(d_1 y_n^2+a_1)(\tau-\tau')}$,

$A_2=2d_2^2/\delta x$, $\Sigma_m^{\pm}=\sum_{m=1}^{\infty}(\pm1)^m y_m^2\,e^{-(d_2 y_m^2+a_2)(\tau-\tau')}$.

Таким чином ми отримали точний розв'язок вихідної контактно-крайової задачі в інтегральному вигляді.

Отримання аналітичного розв'язку для концентрації забруднених частинок у водному поровому розчині в обох шарах фільтра дало можливість знайти за формулою (3) вирази для відповідних потоків маси. На цій основі розроблені програмні модулі з використанням методу чисельного інтегрування, а саме Ньютона-Котеса замкненого типу за 7 і 10 вузлами.

V. Метод чисельного інтегрування подвійного інтеграла зі змінними межами

Структура розв'язків для сорбованої домішки у шарі j у компактному вигляді має вигляд

$$c_2^{(j)}=F_j(x)-\frac{2d_2}{\delta x}e^{\frac{v_2(x-x')}{2d_2}}\times$$

$$\times \sum_{m=1}^{\infty} B_m(x) \int_0^{\tau} \int_0^{\tau'} g(\tau'') f(\tau' - \tau'') d\tau'' d\tau' \qquad (19)$$

Для знаходження подвійних інтегралів зі змінними верхніми межами у співвідношенні (19), запропоновано числовий метод для їх обрахунку.

Зазначимо, що область інтегрування підінтегральної функції (19) у просторі $O\tau\tau'\tau''$ має вигляд нахиленого трикутника (рис. 2).

Враховуючи, що τ - це час, тобто змінна неперервно змінюється від нуля до ∞, зі зміною τ область інтегрування в просторі $O\tau\tau'\tau''$ залишається трикутником. При цьому зміна τ призводить до зміни максимальних значень τ' і τ''. Для $\tau = \tau_1$ маємо $\max\tau' = \max\tau'' = \tau_1$, а для $\tau = \tau_2$ - $\max\tau' = \max\tau'' = \tau_2$. Ці області є подібними трикутниками, які знаходяться в паралельних площинах, тобто під однаковим кутом нахилу θ до $O\tau'\tau''$ (рис. 2). При цьому змінюється площа поверхні інтегрування. Для $\tau = 0$ область інтегрування є точкою $(0,0,0)$.

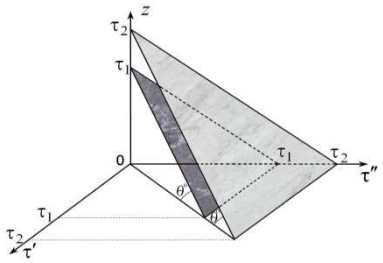


Рис. 2. Область інтегрування в просторі $O\tau\tau'\tau''$

Зробимо проекцію поверхні інтегрування на площину $O\tau'\tau''$ (рис. 3а). Тоді ми отримуємо множину областей інтегрування в просторі $O\tau'\tau''z$.

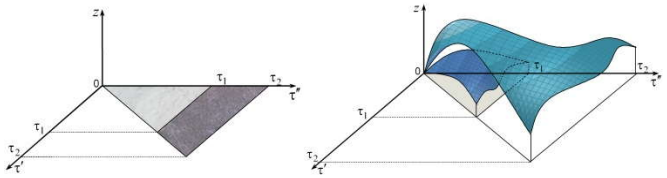


Рис. 3. Схематичне зображення множини областей інтегрування в просторі $O\tau'\tau''z$ (a) та відповідних підінтегральних функцій (b)

Зазначимо, що площа області P_a інтегрування є $P_a = \cos(\theta) P_{pr}$, де P_{pr} - площа проекції області інтегрування. Враховуючи, що для нашої задачі $\theta = 45^{\circ}$, маємо

$$P_{pr} = \frac{1}{\sqrt{2}} P_a .$$

В результаті ми отримаємо сімейство підінтегральних функцій, параметризованих змінною τ (рис. 3b).

На проекцію області інтегрування накладаємо квадратну сітку (рис. 4). Враховуємо адитивну властивість інтеграла. Тоді справедлива формула

$$\int_0^{\tau} \int_0^{\tau'} g(\tau'') f(\tau' - \tau'') d\tau'' d\tau' = \sum_{el} \iint_{(V_{el})} g(\tau'') f(\tau' - \tau'') d\tau'' d\tau' , \qquad (20)$$

де V_{el} - об'єм окремого елемента розбиття проекції області інтегрування. Сумування відбувається по всіх елементах $\bigcup_{el} V_{el} = [0,\tau'] \times [0,\tau]$.

Оскільки в результаті розбиття області інтегрування ми отримали квадратні і трикутні елементи, то формулу (20) можна подати у вигляді

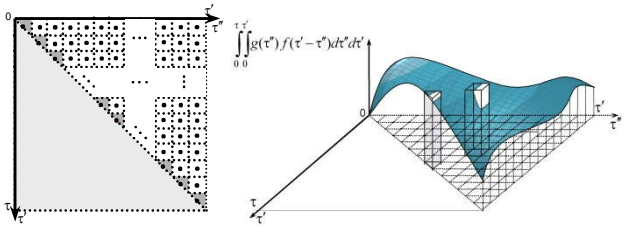


Рис. 4. Розбиття області інтегрування (a) та схема чисельного подвійного інтегрування зі змінними верхніми межами (b)

$$\sum_{el} \iint_{(V_{el})} g(\tau'') f(\tau' - \tau'') d\tau'' d\tau' = \sum_{sq_{el}} \iint_{(V_{sq_{el}})} g(\tau'') f(\tau' - \tau'') d\tau'' d\tau' +$$

$$+ \sum_{\mathrm{tr}_{el}} \iint_{(V_{\mathrm{tr}_{el}})} g(\tau'') f(\tau' - \tau'') d\tau'' d\tau' , \qquad (21)$$

де sq_{el} – квадратний елемент, tr_{el} – трикутний елемент; $V_{sq_{el}}$ і $V_{tr_{el}}$ – об'єми квадратного і трикутного елементів. У правій частині (21) перша сума знаходиться за всіма квадратними елементами sq_{el}, а друга – за всіма трикутними tr_{el}.

Відберемо з множини ті елементи V_{el}, які є квадратами (рис. 5а), що відмічені білим кольором на рис. 4а. І до цих елементів застосовуємо чисельне інтегрування в квадратурах [9, 10]. Тоді інтеграл (20) з квадратними елементами буде мати вигляд

$$\iint_{\bigcup_{\mathrm{sq_{el}}} V_{el}} g(\tau'') f(\tau' - \tau'') d\tau'' d\tau' = \sum_{sq_{el}} \iint_{(V_{sq_{el}})} g(\tau'') f(\tau' - \tau'') d\tau'' d\tau' =$$

$$= \sum_{i=1}^{n} \sum_{j=i-1}^{n} w^2 g(\tau''_j) f(\tau'_i - \tau''_j) .$$

Тут $w = \tau'/n$. Зазначимо, що висота виділеного елемента інтегрування з квадратною основою (рис. 4b) є $h_{ij} = g(\tau''_j) f(\tau'_i - \tau''_j)$, і це значення береться в центрі квадрата, що позначено крапками на рис. 4а.

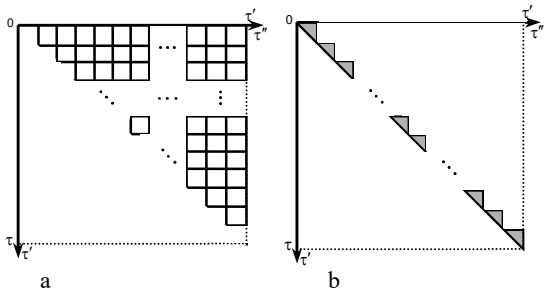


Рис. 5. Області, які складаються з квадратних (a) і трикутних (b) елементів

Враховуючи (21), потрібно знайти сумарний об'єм елементів, в основі яких є трикутники (рис. 5b), причому приймаючи, що сітка є рівномірною.

Суму об'ємів елементів із трикутною основою знаходимо у вигляді

$$\sum_{\mathrm{tr}_{el}} \iint_{\left(V_{\mathrm{tr}_{el}}\right)} g(\tau'')\, f(\tau'-\tau'')\, d\tau'' d\tau' = \frac{\sqrt{2}}{2} \sum_{i=1}^{n} w^2 g(\tau_i'')\, f(\tau_i'-\tau_i''), \quad (22)$$

де n - кількість трикутних елементів; $\tau'_{i+1} = \tau'_i + w$; τ'_i - значення τ' у i-му вузлі (трикутному елементі); τ'' - знаення τ'' у i-му вузлі (трикутному елементі), а висота виділеного елемента інтегрування з трикутною основою $h_{ij} = g(\tau''_j)\, f(\tau'_i - \tau''_j)$ обчислюється у точці, що знаходиться на діагоналі (рис. 4a).

У результаті отримаємо наступну формулу чисельного інтегрування для подвійного інтегралу зі змінними верхніми межами, з урахуванням в тому числі формул (21) та (22)

$$\int_0^{\tau}\int_0^{\tau'} f(\tau',\tau'')dxdy \approx \frac{1}{2}\sum_{i=1}^{n} w^2 g(\tau_i'')\, f(\tau_i'-\tau_i'') + \sum_{i=2}^{n}\sum_{j=i-1}^{n-1} w^2 g(\tau_j'')\, f(\tau_i'-\tau_j''), \quad (23)$$

де $\tau'_{i+1} = \tau'_i + w$; $\tau''_{i+1} = \tau''_i + w$.

Зазначимо, що кількість трикутних елементів на 1 менше ніж квадратних в найдовшому рядку.

З використанням формули (23) для чисельного інтегрування зі змінними верхніми межами аналітичний розв'язок для сорбованої домішкової речовини у шарі j можна записати у такому наближеному вигляді

$$c_2^{(j)}(t,x) \approx F_j(x) - \frac{2d_2}{\delta x} e^{\frac{v_2(x-x')}{2d_2}} \sum_{m=1}^{\infty} B_m(x) \times \times \left(\frac{1}{2}\sum_{i=1}^{n} w^2 g(\tau_i'')\, f(\tau_i'-\tau_i'') + \sum_{i=2}^{n}\sum_{j=i-1}^{n-1} w^2 g(\tau_j'')\, f(\tau_i'-\tau_j'') \right). \quad (24)$$

На рис. 5 показано обчислені за формулою (24) розподіли сумарної концентрації домішкової речовини, сорбованої на скелеті фільтра, у різні моменти безрозмірного часу $\tau = 0{,}1;\ 0{,}6;\ 1;\ 2$ (криві 1-4, рис. 6a) та для різних значень стрибка концентрації в розчині на границі контакту λ =0,5; 1; 2; 3 (криві 1-4, рис. 6b).

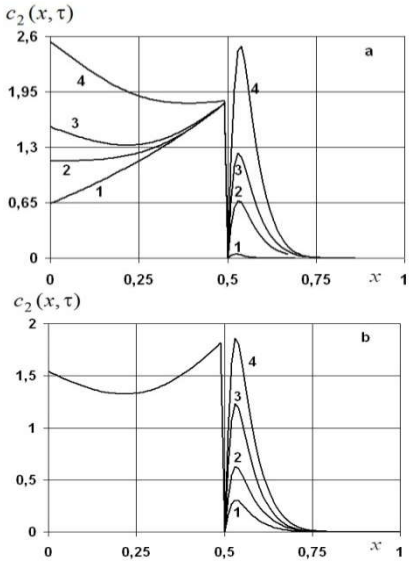


Рис. 6. Розподіли концентрації сорбованої домішкової речовини у двошаровому фільтрі в різні моменти часу (a) та для різних значень коефіцієнта λ (b)

Встановлено, що збільшення часу протікання процесу конвективної дифузії призводить до зростання концентрації сорбованої домішкової речовини у кожному з підшарів фільтра (рис. 6a). При цьому з часом змінюється поведінка функції концентрації в області Ω_1: так для малих часів $c_2^{(1)}(x,\tau)$ є монотонно зростаючою функцією, для часів близьких до стаціонарних $c_2^{(1)}(x,\tau)$ монотонно спадає до границі контакту підобластей. У випадку коефіцієнтів дифузії забруднення, більших у першому шарі ($d_1 > d_2$), збільшення стрибка функції концентрації сорбованих на скелеті частинок на границі контакту викликає зростання локального максимуму концентрації в цієї околі границі (рис. 6b). Водночас поведінка функції концентрації у першій області не змінюється.

VI. Довговічність роботи фільтрів

Якщо відоме максимальне значення концентрації частинок домішкової речовини, спроможної адсорбуватися скелетом тіла, то можна проаналізувати довговічність роботи фільтра. Зокрема, за умови $a_1 \ll a_2$ знайдемо час насичення τ_*, розв'язуючи нелінійне рівняння

$$\sup_{\xi\in\Omega_2} c_2^{(2)}(\xi,\tau_*) = N_2, \quad (25)$$

де N_2 - максимальна концентрація частинок домішки, здатних адсорбуватися на скелеті.

Для розв'язання рівняння (25) розвинуто метод на основі композиції методу простої ітерації та модифікації методу дихотомії. Алгоритм є наступним (рис. 7):

1. Задається початкове значення часу τ_0, кроку розбиття

h_0, точності обчислень ε та $k \in N$.

2. Перевіряється умова зупинки алгоритму $\left|N_2 - c_2^2(\xi, \tau_i)\right| < \varepsilon$. Якщо вона справджується, то поточне значення τ_i є наближеним розв'язком необхідної точності і алгоритм завершується. У протилежному випадку здійснюється перехід на наступний крок.

3. Визначається знак виразу $N_2 - c_2^2(\xi, \tau_i)$. Якщо $N_2 - c_2^2(\xi, \tau_i) > 0$, то $\tau_i = \tau_{i-1} + h_j$. Протилежний випадок означає, що алгоритм вийшов за область значень функції $c_2^{(j)}(\xi, \tau)$ (оскільки $c_2^{(j)}(\xi, \tau) > 0$), і повертаємось до попередньої точки τ_{i-1}. Після чого застосовуємо метод простої ітерації, зменшуючи крок розбиття $h_j = h_{j-1}/k$. Тоді $\tau_i = \tau_{i-2} + h_{j+1}$.

4. Обчислюється значення функції $c_2^{(j)}(\xi, \tau)$ у новій точці τ_i та повертаємось до кроку 2.

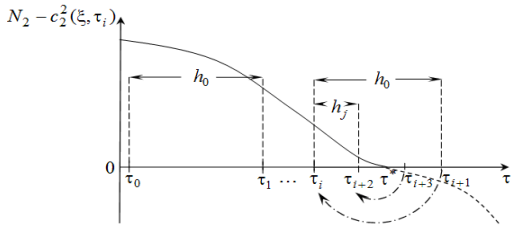


Рис. 7. Схема методу розв'язання нелінійного рівняння (25)

На рис.8 показані розподіли часу, за який відбувається насичення забрудненням другого шару фільтра, для різних

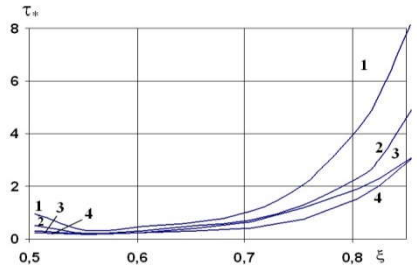


Рис. 8. Час насичення для різних значень відношення коефіцієнтів концентраційної залежності хімічних потенціалів в станах 1 і 2

значень відношення коефіцієнтів концентраційної залежності хімічних потенціалів в станах 1 і 2 λ =0,5; 1; 2; 3 (криві 1-4). Тут прийнято наступні значення параметрів задачі $d_1 = 1$; $d_2 = 0{,}02$; $a_1 = 1$; $a_2 = 1{,}2$; $v_1 = 1$; $v_2 = 0{,}01$

Зазначимо, що час насичення збільшується в околі границі контакту і значно зростає біля нижньої границі фільтруючої установки (рис. 8).

VII. Висновки

Таким чином:

- Сформульовано основні модельні співвідношення процесів конвективної дифузії у двошаровому фільтрі з урахуванням двох станів частинок – у конвективно рухомому розчині і на поверхні скелету.
- Розроблено метод знаходження точних розв'язків контактно-крайових задач масопереносу з урахуванням дифузійного і конвективного механізмів переносу та сорбції, який базується на застосуванні інтегральних перетворень окремо в контактуючих областях.
- Побудовано аналітичні розв'язки задачі конвективної дифузії для двошарового фільтра.
- Розроблено комплекс програм, модулі якого дозволяють розраховувати:
  - концентрацію забруднення, яка переноситься із водним розчином через двошаровий фільтр;
  - потік частинок через довільний переріз фільтра;
  - концентрацію домішкової речовини, яка сорбована на скелеті;
  - час насичення забруднюючою речовиною фільтра
- Встановлено основні закономірності процесу конвективної дифузії у двошаровій області та проаналізовано вплив фізичних і геометричних параметрів на досліджувані функції.

Література References

[1] Advanced Wastewater Treatment Technologies [Електрон. ресурс]. – Режим доступу: http://www.waterform.com.au/Products/10004/nextSand.aspx

[2] *М. Журба,* Основы процессов доочистки сточных вод фильтрированием // Тепло- и массообмен в капиллярнопористых телах. – 1965. – № 1. – С. 60-73.

[3] *V. Lytvy, D. Uhryn, A. Fityo,* Modeling of territorial community formation as a graph partitioning problem // Eastern European J. of Enterprise Technologies. – 2016. -**79**, No 1/4. – P.47-52.

[4] А. Бомб, А. Сафоник,Числово-асимптотичне наближення розв'язків просторових модельних задач процесу аеробного очищення стічної води. // *Вісник КрНу ім. М. Остроградського.* - 2011. - Вип. 6 (11), С. 31-36.

[5] *Є. Чапля*, О. *Чернуха,* Фізико-математичне моделювання гетеродифузного масопереносу. – Львів: СПОЛОМ, 2003. – 128 с.

[6] *Є. Чапля, О. Чернуха,* Математичне моделювання дифузійних процесів у випадкових і регулярних структурах. – Київ: Наукова думка, 2009. – 302 с.

[7] *Н. Мартыненко*, Л. *Пустыльников,* Конечные интегральные преобразования и их применение к исследованию систем с распределенными параметрами. – М.: Наука, 1985. – 304 с.

[8] *Э. Камке,* Справочник по обыкновенным дифференциальным уравнениям. – М.: Наука, 1985. – 304 с.

[9] Справочник по специальным функциям / Под ред. *М. Абрамовица* и *И. Стиган*. – М.: Мир, 1979. – 830 с.

[10] *А. Г. Блюмин, А. А. Федотов, П. В. Храпов*, Численные методы вычисления интегралов и решения задач для обыкновенных дифференциальных уравнений. — М.: МГТУ им. Н.Э.Баумана, 2008. - 74 с.

Фізико-Математична Модель Зв'язаних Теплових, Механічних і Дифузійних Процесів у Багатокомпонентних Тілах

В.Є. Гончарук
кафедра цивільної безпеки, Національний університет "Львівська політехніка";
відділ математичного моделювання нерівноважних процесів, Центр математичного моделювання Інституту прикладних проблем механіки і математики ім. Я.С.Підстригача НАН України, Львів, Україна
vegoncharuk@ukr.net

Physical-Mathematical Model of Coupled Thermal, Mechanical and Diffusive Processes in Multicomponent Bodies

V. Goncharuk
Department of Civil Security, Lviv Polytechnic National University;
Department of mathematical modeling nonequilibrium processes,
Centre of Mathematical Modeling of Pidstryhach Institute for Applied Problems of Mechanics and Mathematics,
Ukrainian Academy of Sciences
Lviv, Ukraine
vegoncharuk@ukr.net

***Анотація*—За континуально-термодинамічним підходом отримані вихідні співвідношення повної моделі механотермогетеродифузійних процесів з урахуванням розпаду мігруючої речовини. Проведена лінеаризація рівнянь стану та кінетичних співвідношень.**

***Abstract*—By the continuum-thermodynamic approach the basic relationships of the complete model of mechanothermoheterodiffusive processes with taking into consideration of decay of migrating particles are obtained. Linearization of both the state equations and kinetic relations is carried out.**

Ключові слова—континуально-термодинічний підхід; фізико-математична модель; механотермогетеродифузія; багатокомпонентне середовище; балансове рівнння

Keywords—continuum-thermodynamic approach; physical-mathematical model; mechanothermoheterodiffusion; multicomponent media; balance equation

I. Вступ

У багатьох випадках при математичному описі міграції домішкової речовини необхідно враховувати її природну деградацію та структуру матеріалу тіла [1-3]. Зокрема, у радіоактивних паливовмісних матеріалах, що виникли в результаті аварії на четвертому енергоблоці ЧАЕС, унаслідок розпаду утворюються мікродефекти, які істотно впливають на інтенсивність дифузії [1, 3]. У той же час математичне моделювання переносу забруднень безпосередньо пов'язано з певними схематичними уявленнями про пористе середовище, а відповідні прогнозні оцінки, наприклад, захищеності ґрунтових вод, повинні враховувати розпад дифундуючої речовини [4, 5].

Тому в роботі запропоновано побудову фізико-математичної моделі за допомогою континуально-термодинамічного підходу [6] для опису взаємозв'язаних механічних, теплових і гетеродифузійних процесів з урахуванням розпаду мігруючої речовини.

II. Об'єкт дослідження

Вважаємо, що мала довільно вибрана область складається з монокристалів мінералів різного типу, які утворюють деформівний пористий скелет. Пори заповнені водним розчином домішкових речовин. Деяка частина води під впливом приповерхневих електричних полів скелету знаходиться в поляризованому зв'язаному стані [7, 8]. У певній мірі умовно в ній виділяється міцнозв'язана складова безпосередньо біля поверхні монокристалів мінералів (для якої ряд авторів зауважують помітну зміну густини та інших властивостей) та слабко адсорбований шар води, що шляхом нагрівання до температури кипіння пе-

ретвориться в пароподібний стан. Ця частина води, як правило, знаходиться у капілярно-зв'язаному або рухомому стані, а при наявності мережі крупних тріщин і пор – у гравітаційно-рухомому вигляді [7].

Частинки домішкової речовини одного сорту у рамках фізично малого елемента тіла (рис. 1) можуть знаходитись в об'ємі і на поверхні скелету (у зв'язаних зі скелетом поляризованих долях води) чи бути в розчині (капілярно-зв'язаній і гравітаційно рухомій долях води). З фізичної точки зору локально вони перебувають у різних станах. При цьому в наслідок радіоактивного розпаду, хімічних реакцій тощо вони розпадаються і утворюють частинки інших домішкових речовин. Таким чином вважаємо, що кожен довільний фізично малий елемент тіла містить частинки води, твердого скелету грунту, розпадні домішкові частинки і частинки, які утворилися внаслідок розпаду, у вказаних вище станах. При цьому ці складові розглядаються як термодинамічні компоненти системи.

Приймемо, що тіло $\mathbf{K}^*$ (дискретна сукупність матеріальних частинок) є багатокомпонентним розчином і утворене частинками домішкової речовини у поровому розчині $\mathbf{K}_1^{*(1)}$, поверхні $\mathbf{K}_2^{*(1)}$ та об'ємі скелету $\mathbf{K}_3^{*(1)}$, які можуть розпадатися, та частинками порового розчину $\mathbf{K}_4^{*(0)}$ і скелету $\mathbf{K}_5^{*(0)}$. При розпаді речовини $\mathbf{K}_j^{*(1)}$ в стані $j=\overline{1,3}$ утворюються частинки інших домішкових речовин $\mathbf{K}_j^{*(2)}$ і $\mathbf{K}_j^{*(3)}$, причому частинки, що виникли внаслідок розпаду вже не розпадаються, тобто $\mathbf{K}_j^{*(1)} \to \mathbf{K}_j^{*(2)} + \mathbf{K}_j^{*(3)}$. Частинки домішкової речовини одного хімічного виду, перебуваючи у різних фазах, знаходяться в істотно різних фізичних станах, характеризуючись, зокрема, різною рухливістю (коефіцієнтами дифузії).

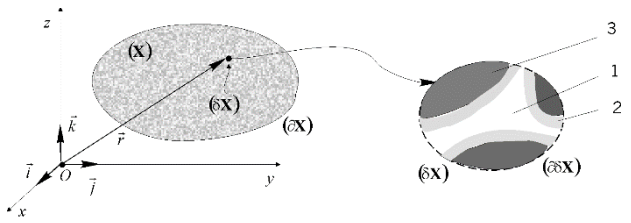


Рис. 1. Характерна структура фізично малого елемента тіла. Область 1 займає водний поровий розчин, область 2 – адсорбовані на скелеті грунту шари води, 3 – скелет грунту

При макроскопічному описі тіло $\mathbf{K}^*$ розглядається як термодинамічна система, яка вкладена в обмежену область евклідового простору $\mathbf{x}$, що віднесений до декартової системи кординат $\left\{x^\alpha\right\}$ [6, 9]. Кожній компоненті тіла (підсистемам частинок $\mathbf{K}_j^{*(0)}$, $j=4;5$, що утворюють скелет та розчин, а також частинкам розпадної домішкової речовини в різних станах $\mathbf{K}_j^{*(1)}$ і частинкам, які утворилися внаслідок розпаду, $\mathbf{K}_j^{*(i)}$, $j=\overline{1,3}$, $i=2,3$) співставляємо у відповідність континуми $\mathbf{K}_j^{(i)}$ ($i=\overline{0,3}$, $j=\overline{1,5}$).

Розглянемо також континуум центрів мас $\mathbf{K}_c$, точкам яких приписуємо швидкість $\vec{v}$ згідно із співвідношенням

$$\rho\vec{v} = \sum_{i=0}^{3}\sum_{j=1}^{5}\rho_j^{(i)}(\vec{r},t)\vec{v}_j^{(i)}(\vec{r},t), \qquad (1)$$

де

$$\rho_j^{(i)}(\vec{r},t) = \lim_{(\delta\mathbf{X})\to 0}\left(\frac{\delta m_j^{(i)}}{\delta V}\right)$$

- густини компонент $\mathbf{K}_j^{(i)}$, $\delta m_j^{(i)}$ - їхня маса в області $(\delta\mathbf{X})$ (рис. 1), об'єм якої δV:

$$\rho \equiv \rho(\vec{r},t) = \sum_{i=0}^{3}\sum_{j=1}^{5}\rho_j^{(i)}(\vec{r},t)\,\Phi \qquad (2)$$

- сумарна густина тіла; $\vec{v}_j^{(i)}$ - швидкості руху матеріальних точок континуумів $\mathbf{K}_j^{*(i)}$.

Вважаємо, що в умовах нестаціонарної взаємодії з зовнішнім середовищем, основними нерівноважними процесами, що протікають у тілі, є деформація, теплопровідність і дифузія домішкової речовини у кожному із вище згаданих станів, що супроводжуються взаємними переходами між ними та розпадом частинок.

III. Балансові рівняння

За вихідні співвідношення математичної моделі механотермогетеродифузії домішкової речовини двома шляхами у середовищі з пастками з урахуванням розпаду мігруючої речовини приймемо закони збереження і балансові рівняння для маси, імпульсу та енергії, які сформулюємо для кожної з компонент або континууму центрів мас.

Зміна маси компоненти $m_j^{(i)}$ відбувається за рахунок масових потоків і внутрішніх джерел, тоді мають місце рівняння балансу маси компоненти ij

$$\frac{\partial\rho_j^{(i)}}{\partial t} = -\vec{\nabla}\cdot\left(\rho_j^{(i)}\vec{v}_j^{(i)}\right) + w_j^{(i)} \quad (i=\overline{0,3},\ j=\overline{1,5}), \qquad (3)$$

де $\vec{\nabla}$ - набла-оператор Гамільтона; $w_j^{(i)}$ - густина внутрішнього джерела (або стоку) компоненти; крапкою позначений скалярний добуток.

Оскільки ми прийняли, що джерелом маси компоненти можуть бути процеси сорбції-десорбції частинок і розпад домішкової речовини, то потужність виробництва маси $w_j^{(i)}$ в загальному випадку можна подати у вигляді суми

$$w_j^{(i)} = \sum_{\substack{k=1\\k\neq j}}^{5}\omega_{jk}^{(i)} + \overline{w}_j^{(i)} \qquad (i=\overline{0,3},\ j=\overline{1,5}),$$

де $\omega_{jk}^{(i)}$ - потужність виробництва маси компоненти i у стані j у зв'язку з її переходом з континууму $\mathbf{K}_k^{(i)}$; $\overline{w}_j^{(i)}$ - потужність виробництва маси компоненти ij за

рахунок розпаду частинок компоненти $i=1$ ($i=2,3$, $j=\overline{1,3}$). Тоді

$$w_j^{(0)}=w_j^{(1)}=0\ \ (\forall j),\qquad \sum_{j=1}^{3}\sum_{i=2}^{3}w_j^{(i)}=0; \tag{4}$$

У тому числі

$$\omega_{jj}^{(i)}=0\ \ (\forall j),\ \omega_{jk}^{(i)}=-\omega_{kj}^{(i)}\ (\forall i,j,k),\sum_{j=1}^{5}\sum_{k=1}^{5}\omega_{jk}^{(i)}=0\ (\forall i).$$

Балансове рівняння (2) можна подати у вигляді

$$\frac{\partial\rho_j^{(i)}}{\partial t}=-\vec{\nabla}\cdot\vec{J}_j^{(i)0}+w_j^{(i)}\quad (i=\overline{0,3},\ j=\overline{1,5}), \tag{5}$$

де $\vec{J}_j^{(i)0}=\rho_j^{(i)}\vec{v}_j^{(i)}$ - масовий потік частинок компоненти i у стані j.

Просумуємо рівняння (3) або (5) за всіма індексами i, j та врахуємо, що сумарна густина маси вводиться співвідношенням (2), а швидкість $\vec{v}$ - рівнянням (1) або $\vec{v}=\sum_{j=1}^{5}\sum_{i=0}^{3}\rho_j^{(i)}\vec{v}_j^{(i)}\big/\rho$, тоді отримаємо

$$\frac{\mathrm{d}\rho}{\mathrm{d}t}=-\rho\vec{\nabla}\cdot\vec{v}, \tag{5а}$$

де $\frac{\mathrm{d}}{\mathrm{d}t}=\frac{\partial}{\partial t}+\vec{v}\cdot\vec{\nabla}$ - повна похідна за часом.

Оскільки $\rho=1/\mathrm{v}$, де v - питомий об'єм, то формула (5а) еквівалентна

$$\frac{\mathrm{dv}}{\mathrm{d}t}=\mathrm{v}\vec{\nabla}\cdot\vec{v}. \tag{5б}$$

Враховуючи, що $\rho_j^{(i)}=C_j^{(i)}\rho$, одержимо рівняння балансу для концентрацій компонент:

$$\rho\frac{\mathrm{d}C_j^{(i)}}{\mathrm{d}t}=-\vec{\nabla}\cdot\vec{J}_j^{(i)}+w_j^{(i)},\quad i=\overline{0,3},\ j=\overline{1,5}, \tag{6}$$

де $\vec{J}_j^{(i)}=\rho_j^{(i)}\left(\vec{v}_j^{(i)}-\vec{v}\right)$ - масовий потік компоненти ij; $C_j^{(i)}=\rho_j^{(i)}\big/\rho$ - масові концентрації компонент, які задовольняють умову нормування

$$\sum_{j=1}^{5}\sum_{i=0}^{3}C_j^{(i)}=1. \tag{7}$$

Якщо використати умову нормування (7), то балансове рівняння (6) можна записати у вигляді

$$\rho\frac{\mathrm{d}C_{\mathrm{j}}^{(\mathrm{i})}}{\mathrm{d}}=-\vec{\nabla}\cdot\vec{J}_j^{(i)}+w_j^{(i)},\ i=\overline{0,3},\ j=\overline{1,4},\ C_5^{(0)}=1-\sum_{j=1}^{4}\sum_{i=0}^{3}C_j^{(i)}.$$

Рівняння балансу імпульсу для тіла в цілому (континууму $\mathbf{K}_c$) має вигляд

$$\rho\frac{\mathrm{d}\vec{v}}{\mathrm{d}t}=\vec{\nabla}\cdot\hat{\Pi}+\sum_{j=1}^{5}\sum_{i=0}^{3}\rho_j^{(i)}\vec{F}_j^{(i)}, \tag{8}$$

де $\hat{\Pi}=\Pi^{\alpha\beta}\vec{i}^{\,\alpha}\otimes\vec{i}^{\,\beta}$ - симетричний тензор напружень Коші [10, 11], $\vec{F}_j^{(i)}$ - масова потенціальна ($\vec{F}_j^{(i)}=-\vec{\nabla}\psi_j^{(i)}$) і консервативна ($\partial\psi_j^{(i)}\big/\partial t=0$) сила, яка діє на i-ту компоненту у стані j, $\psi_j^{(i)}$ - потенціал сил (потенціальна енергія одиниці маси компоненти i у стані j).

Домножимо закон збереження маси компонент (3) на скалярний потенціал $\psi_i^{(j)}$, віднесений до одиниці маси компоненти, і врахуємо умову їх консервативності. Тоді знайдемо рівняння балансу потенціальної енергії компоненти. Підсумовуючи записане рівняння за індексами i та j, означуючи потенціальну енергію точок континууму $\mathbf{K}_c$ Ψ формулою

$$\rho\Psi=\sum_{j=1}^{5}\sum_{i=0}^{3}\rho_j^{(i)}\psi_j^{(i)}$$

і враховуючи $\vec{J}_j^{(i)0}=\vec{J}_j^{(i)}+\rho_j^{(i)}\vec{v}$, отримаємо рівняння балансу потенціальної енергії для тіла в цілому

$$\rho\frac{\mathrm{d}\Psi}{\mathrm{d}t}=-\vec{\nabla}\cdot\sum_{j=1}^{5}\sum_{i=0}^{3}\psi_j^{(i)}\vec{J}_j^{(i)}-\sum_{j=1}^{5}\sum_{i=0}^{3}\rho_j^{(i)}\vec{F}_j^{(i)}\cdot\vec{v}-$$
$$-\sum_{j=1}^{5}\sum_{i=0}^{3}\vec{J}_j^{(i)}\cdot\vec{F}_j^{(i)}+\sum_{j=1}^{5}\sum_{i=0}^{3}\psi_j^{(i)}w_j^{(i)}.$$

Домножимо скалярно рівняння балансу імпульсу (8) на $\vec{v}$ і використаємо тотожність $\vec{\nabla}\cdot\left(\hat{\Pi}\cdot\vec{v}\right)\equiv\vec{v}\cdot\left(\vec{\nabla}\cdot\hat{\Pi}\right)+$ $+\hat{\Pi}:\vec{\nabla}\otimes\vec{v}$, де знак «:» означає подвійну внутрішню згортку. Тоді маємо рівняння балансу кінетичної енергії

$$\rho\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{v^2}{2}\right)=\vec{\nabla}\cdot\left(\vec{v}\cdot\hat{\Pi}\right)-\hat{\Pi}:\vec{\nabla}\otimes\vec{v}+\sum_{j=1}^{5}\sum_{i=0}^{3}\rho_j^{(i)}\vec{F}_j^{(i)}\cdot\vec{v}.$$

Постулюємо, що повна енергія $e^{(j)}$ (з розрахунку на одиницю маси) задається виразом

$$\mathrm{e}\stackrel{*}{=}\rho\left(\mathrm{u}+\Psi+v^2/2\right) \tag{9}$$

і задовольняє закон збереження

$$\frac{\partial e}{\partial t}=-\vec{\nabla}\cdot\vec{J}_e, \tag{10}$$

де $u^{(j)}$ - питома внутрішня енергія j з розрахунку на одиницю маси; $\vec{J}_e^{(j)}$ - потік повної енергії:

$$\vec{J}_e = \rho\left(\mathrm{u} + \Psi + v^2/2\right)\vec{v} + \vec{J}_Q - \hat{\Pi}\cdot\vec{v} + \sum_{j=1}^{5}\sum_{i=0}^{3}\left(\psi_j^{(i)} + \mu_j'^{(i)}\right)\vec{J}_j^{(i)} \quad (11)$$

Тут $\vec{J}_Q^{(j)}$ - потік енергії у формі тепла, $\mu_j'^{(i)}$ - хімічний потенціал компоненти i у стані j [12].

Якщо повна енергія підпорядковується (9)-(11), тоді має місце рівняння балансу питомої внутрішньої енергії

$$\rho\frac{\mathrm{d}u}{\mathrm{d}t} = -\vec{\nabla}\cdot\vec{J}_Q - \vec{\nabla}\cdot\left(\sum_{j=1}^{5}\sum_{i=0}^{N}\mu_j'^{(i)}\vec{J}_j^{(i)}\right) + \sum_{j=1}^{5}\sum_{i=0}^{N}\vec{J}_j^{(i)}\cdot\vec{F}_j^{(i)} + \\ + \hat{\Pi} : \left(\vec{\nabla}\otimes\vec{v}\right)^T + \sum_{j=1}^{5}\sum_{i=0}^{N}\psi_j^{(i)}w_j^{(i)}, \quad (12)$$

де операція $(...)^T$ означає операцію транспонування.

IV. Рівняння Гіббса

Приймаємо гіпотезу локальної термодинамічної рівноваги. Нехай термодинамічний стан довільного макроскопічного елемента системи, який займає її малу підобласть, означений з допомогою таких спряжених параметрів як абсолютна температура T, питома ентропія s; тензор напружень Коші у рівновазі $\hat{\bar{\sigma}}$, тензор деформації $\hat{\varepsilon}$; хімічний потенціал $\mu_j'^{(i)}$ і масова концентрація $C_j^{(i)}$ компоненти i у стані j:

$$T - s, \quad \bar{\sigma}^{\alpha\beta} - \varepsilon_{\alpha\beta}, \quad \mu_j'^{(i)} - C_j^{(i)},$$

де $s = S/m$, S - ентропія системи в цілому; $\bar{\sigma}^{\alpha\beta}$ - контраваріантні компоненти тензора напружень Коші у рівновазі, $\varepsilon_{\alpha\beta}$ - коваріантні компоненти тензора деформації.

Питома внутрішня енергія системи є такою функцією стану $\mathrm{u} = \mathrm{u}\left(s, \mathrm{v}, \{C_j^{(i)}\}\right)$, яка задовольняє співвідношення

$$\mathrm{u} = Ts + \frac{1}{\rho}\Pi^{\alpha\beta}\varepsilon_{\alpha\beta} + \sum_{j=1}^{5}\sum_{i=0}^{3}\mu_j'^{(i)}C_j^{(i)} \text{ (рівняння Ейлера)}, \quad (13)$$

за умови, що в той же час інфінітизимально малі зміни масової густини ентропії ds, питомого об'єму $d\mathrm{v}$ та масових концентрацій $dC_j^{(i)}$ задовольняють рівняння

$$d\mathrm{u} = Tds + \frac{1}{\rho}\Pi^{\alpha\beta}d\varepsilon_{\alpha\beta} + \sum_{j=1}^{5}\sum_{i=0}^{3}\mu_j'^{(i)}dC_j^{(i)} \text{ (рівняння Гіббса)}. \quad (14)$$

Оскільки масові концентрації задовольняють умову (7), то рівняння Гіббса (14) можна переписати у вигляді

$$d\mathrm{u} = Tds + \frac{1}{\rho}\Pi^{\alpha\beta}d\varepsilon_{\alpha\beta} + \sum_{j=1}^{4}\sum_{i=0}^{3}\mu_j^{(i)}dc_j^{(i)}, \quad (15)$$

де $\mu_j^{(i)} = \mu_j'^{(i)} - \mu_5^{(0)}$.

Субстанціональна форма рівняння Гіббса (15) буде

$$\rho\frac{\mathrm{d}u}{\mathrm{d}t} = T\rho\frac{\mathrm{d}s}{\mathrm{d}t} + \Pi^{\alpha\beta}\frac{\mathrm{d}\varepsilon_{\alpha\beta}}{\mathrm{d}t} + \sum_{j=1}^{4}\sum_{i=0}^{3}\mu_j^{(i)}\frac{\mathrm{d}c_j^{(i)}}{\mathrm{d}t}. \quad (16)$$

V. Рівняння стану

Якщо питома внутрішня енергія u є функцією змінних s, $\varepsilon_{\alpha\beta}$, $\left\{c_j^{(i)}\right\}$ (де $\left\{c_j^{(i)}\right\}$ - набір $c_j^{(i)}$ для $\forall i, j$) і підпорядковується (13), (14), то рівняння стану мають вигляд

$$T = \left(\frac{\partial \mathrm{u}}{\partial s}\right)_{\varepsilon_{\alpha\beta}, c_j^{(i)}}, \quad \frac{1}{\rho}\bar{\sigma}^{\alpha\beta} = \left(\frac{\partial \mathrm{u}}{\partial \varepsilon_{\alpha\beta}}\right)_{s, c_j^{(i)}}, \quad \mu_j^{(i)} = \left(\frac{\partial \mathrm{u}}{\partial c_j^{(i)}}\right)_{T, \varepsilon_{\alpha\beta}, \left\{c_{l\neq j}^{(k\neq i)}\right\}}, \quad (17)$$

Конкретна структура залежностей (17) за макроскопічним описом матеріальної системи може бути встановлена з використанням експериментальних даних або методами фізики твердого тіла.

VI. Рівняння балансу ентропії

Домножимо рівняння Гіббса (15) на величину $\rho/\mathrm{d}t$. В отримане рівняння підставимо співвідношення (3), (5б) і (12). Також врахуємо, що

$$\frac{1}{T}\vec{\nabla}\cdot\vec{J}_Q = \vec{\nabla}\cdot\frac{\vec{J}_Q}{T} - \vec{J}_Q\cdot\vec{\nabla}\frac{1}{T}, \quad \vec{\nabla}\frac{1}{T} = -\frac{1}{T^2}\vec{\nabla}T,$$

Знехтуємо ефектами в'язкості і приймемо $\hat{\Pi} = -P\hat{I}$ ($\hat{I}$ - одиничний тензор). Маємо

$$\rho\frac{\mathrm{d}s}{\mathrm{d}t} = -\vec{\nabla}\cdot\frac{\vec{J}_Q}{T} + \vec{J}_Q\cdot\vec{\nabla}\frac{1}{T} + \frac{1}{T}\left[\sum_{j=1}^{5}\sum_{i=0}^{3}\vec{J}_j^{(i)}\cdot\left(\vec{F}_j^{(i)} - \vec{\nabla}\mu_j'^{(i)}\right) - \right. \\ \left. - \sum_{j=1}^{5}\sum_{i=0}^{3}\psi_j^{(i)}w_j^{(i)} - \sum_{j=1}^{5}\sum_{i=0}^{3}\mu_j'^{(i)}w_j^{(i)}\right]. \quad (18)$$

Позначимо $\vec{J}_s = \vec{J}_Q/T$ - потік ентропії; $\vec{X}_Q = -\vec{\nabla}T/T$ - векторна термодинамічна сила, що обумовлює теплопровідність і є спряженою до потоку тепла $\vec{J}_Q$; $\vec{X}_j^{(i)} = = -\vec{\nabla}\left(\mu_j'^{(i)} + \psi_j^{(i)}\right)$ - векторні термодинамічні сили дифузії, спряжені до векторних потоків дифузії $\vec{J}_j^{(i)}$.

Розглянемо дві останні суми рівності (18). Оскільки $\psi_j^{(0)} = \psi_j^{(1)} = ... = \psi_j^{(N)} = \psi$, то враховуючи (4), отримаємо

$$\sum_{j=1}^{5}\sum_{i=0}^{3}\psi_j^{(i)}w_j^{(i)} = \psi\sum_{j=1}^{5}\sum_{i=0}^{3}w_j^{(i)} = 0.$$

Приймемо до уваги тільки процеси переходу домішкових частинок між адсорбованими на скелеті шарами води і поровим розчином $\omega_{12}^{(i)}$ ($i = \overline{1,3}$) та поверхнею і об'ємом скелету $\omega_{23}^{(i)}$ ($i = \overline{1,3}$). Потужності виробництва маси $\overline{w}_j^{(i)}$ для компонент $i = \overline{1,3}$ підпорядковуються умовам

$$\overline{w}_j^{(i)} = \overline{w}_j^{(ii-1)} + \overline{w}_j^{(i2)} + \overline{w}_j^{(i3)},\ \overline{w}_j^{(i)} = 0\ (j = 4;5);$$
$$\overline{w}_j^{(i+1i)} = -\overline{w}_j^{(ii+1)}\ (i = \overline{1,3},\ j = \overline{1,3}),\ \overline{w}_j^{(ii)} = 0\ (\forall i);$$
$$\overline{w}_j^{(ii\pm l)} = 0\ (l = 2;3),\ \overline{w}_j^{(il)} = -\overline{w}_j^{(li)}\ (i = 2;3,\ j = \overline{1,3});$$
$$\overline{w}^{(i)} = \overline{w}_1^{(i)} = \overline{w}_2^{(i)} = \overline{w}_3^{(i)}\ (i = \overline{1,3}).$$

Тоді маємо

$$\sum_{j=1}^{5}\sum_{i=0}^{3}\mu_j'^{(i)}w_j^{(i)} = \sum_{j=1}^{5}\sum_{i=0}^{3}\mu_j'^{(i)}\left[\sum_{k=1}^{3}\omega_{jk}^{(i)} + \overline{w}_j^{(i)}\right]. \qquad (19)$$

Перший доданок правої частини (19) можна перетворити до вигляду

$$\sum_{j=1}^{5}\sum_{i=0}^{3}\sum_{k=1}^{5}\mu_j'^{(i)}\omega_{jk}^{(i)} = \sum_{i=1}^{3}\left[-\left(\mu_2'^{(i)} - \mu_1'^{(i)}\right)\omega_{12}^{(i)} - \left(\mu_3'^{(i)} - \mu_2'^{(i)}\right)\omega_{23}^{(i)}\right].$$

Позначимо

$$\omega_1^{(i)} = \omega_{12}^{(i)},\quad \omega_2^{(i)} = \omega_{23}^{(i)}\quad (i = \overline{1,3}) \qquad (20)$$

- скалярні масові потоки, які характеризують домішкових частинок відповідно між поровим розчином і поверхнею скелету та між адсорбованою водою і об'ємом скелету;

$$X_1^{(i)} = \mu_2'^{(i)} - \mu_1'^{(i)},\quad X_2^{(i)} = \mu_3'^{(i)} - \mu_2'^{(i)}\quad (i = \overline{1,3}) \qquad (21)$$

- скалярні термодинамічні сили, спряжені до відповідних масових потоків $\omega_j^{(i)}$.

Другий доданок правої частини рівності (19) можемо подати так

$$\sum_{i=0}^{3}\mu_j'^{(i)}\overline{w}_j^{(i)} = -\sum_{i=1}^{2}\left(\mu_j'^{(i+1)} - \mu_j'^{(i)}\right)\overline{w}_j^{(ii+1)}, \qquad (22)$$

де приймаємо, що

$$\overline{w}_j^{(ik)}\quad (i,k = \overline{1,3}) \qquad (23)$$

- скалярні масові потоки, які характеризують розпад частинок домішкової речовини, вважаємо відомими;

$$\overline{X}_j^{(i)} = \mu_j'^{(i+1)} - \mu_j'^{(i)}\quad (i = \overline{1,3},\ j = \overline{1,3}) \qquad (24)$$

- скалярні величини, спряжені до відповідних масових потоків $\overline{w}_j^{(ik)}$, які визначають супутнє тепловиділення в процесі розпаду.

Тоді з використанням позначень (16)-(18), (20)-(24) рівняння (14) перепишемо у вигляді

$$\rho\frac{\mathrm{d}s}{\mathrm{d}t} = -\vec{\nabla}\cdot\vec{J}_s + \frac{1}{T}\left[\vec{J}_Q\cdot\vec{X}_Q + \sum_{j=1}^{5}\sum_{i=0}^{3}\vec{J}_j^{(i)}\cdot\vec{X}_j^{(i)} + \sum_{k=1}^{2}\sum_{i=1}^{3}\omega_k^{(i)}X_k^{(i)} + \sum_{i=1}^{3}q^{(i)}\right]$$

або

$$\rho\frac{\mathrm{d}s}{\mathrm{d}t} = -\vec{\nabla}\cdot\vec{J}_s + \frac{\sigma_s}{T},$$

де

$$\sigma_s^{(j)} = \vec{J}_Q\cdot\vec{X}_Q + \sum_{j=1}^{5}\sum_{i=0}^{3}\vec{J}_j^{(i)}\cdot\vec{X}_j^{(i)} + \sum_{k=1}^{2}\sum_{i=1}^{3}\omega_k^{(i)}X_k^{(i)} + \sum_{j=1}^{3}q_j \geq 0$$

- потужність виробництва ентропії, $q_j = \sum_{i=1}^{3}\overline{X}_j^{(i)}\overline{w}_j^{(ii+1)}$ - локальні тепловиділення при розпаді компонент.

VII. Кінетичні співвідношення

Ми маємо сукупність спряжених величин – термодинамічні сили і відповідні термодинамічні потоки:

$$\vec{X}_Q \div \vec{J}_Q,\quad \vec{X}_j^{(i)} \div \vec{J}_j^{(i)},\quad X_k^{(i)} \div \omega_k^{(i)},$$

які у вихідному рівноважному стані приймають значення

$$\vec{X}_{Q0} \div \vec{J}_{Q0},\quad \vec{X}_{j0}^{(i)} \div \vec{J}_{j0}^{(i)},\quad X_{k0}^{(i)} \div \omega_{k0}^{(i)}.$$

Вважаємо, що термодинамічні потоки є функціями термодинамічних сил, структура функціональних залежностей яких задовольняє другий закон термодинаміки та умови взаємності Онзагера. Тобто мають місце:

I. Функціональні залежності

$$J_Q = J_Q\left(X_Q, X_{j=1}^{(0)},\ldots,X_{j=5}^{(3)}, X_{k=1}^{(1)}, X_{k=1}^{(2)}, X_{k=1}^{(3)}, X_{k=2}^{(1)}, X_{k=2}^{(2)}, X_{k=2}^{(3)}\right)$$
$$J_j^{(i)} = J_j^{(i)}\left(X_Q, X_{j=1}^{(0)},\ldots,X_{j=5}^{(3)}, X_{k=1}^{(1)}, X_{k=1}^{(2)}, X_{k=1}^{(3)}, X_{k=2}^{(1)}, X_{k=2}^{(2)}, X_{k=2}^{(3)}\right)$$
$$\omega_k^{(i)} = \omega_k^{(i)}\left(X_Q, X_{j=1}^{(0)},\ldots,X_{j=5}^{(3)}, X_{k=1}^{(1)}, X_{k=1}^{(2)}, X_{k=1}^{(3)}, X_{k=2}^{(1)}, X_{k=2}^{(2)}, X_{k=2}^{(3)}\right) \qquad (25)$$

Вирази (25) записані для скалярних компонент відповідних векторних потоків і сил.

II. Другий закон термодинаміки у вигляді

$$\sigma_s^{(j)} = \vec{J}_Q\cdot\vec{X}_Q + \sum_{j=1}^{5}\sum_{i=0}^{3}\vec{J}_j^{(i)}\cdot\vec{X}_j^{(i)} + \sum_{k=1}^{2}\sum_{i=1}^{3}\omega_k^{(i)}X_k^{(i)} + \sum_{j=1}^{3}q_j \geq 0 \qquad (26)$$

III. Умови взаємності Онзагера [11, 13]

$$\frac{\partial\vec{J}_Q}{\partial\vec{X}_j^{(i)}} = \frac{\partial\vec{J}_j^{(i)}}{\partial\vec{X}_Q},\ \frac{\partial\vec{J}_Q}{\partial X_k^{(i)}} = \frac{\partial\omega_k^{(i)}}{\partial\vec{X}_Q},\ \frac{\partial\omega_k^{(i)}}{\partial\vec{X}_j^{(i)}} = \frac{\partial\vec{J}_j^{(i)}}{\partial X_k^{(i)}}. \qquad (27)$$

Тут похідні означені формулами

$$\frac{\partial\vec{J}}{\partial\vec{X}} \equiv \overset{\alpha}{\underset{0}{\vec{i}}}\otimes\frac{\partial\vec{J}}{\partial X_0^{\alpha}},\qquad \vec{X} \equiv \overset{\alpha}{\underset{0}{\vec{i}}}\,X_0^{\alpha},$$

де $\overset{\alpha}{\underset{0}{\vec{i}}} = g_0^{\alpha\beta}\overset{0}{\underset{\beta}{\vec{i}}}$ - контраваріантні базисні вектори, $g_0^{\alpha\beta}$ - контраваріантні компоненти метричного тензора декартової системи координат.

Враховуючи вигляд функціональної залежності (26) виробництва ентропії від термодинамічних потоків і сил, можна ввести кінетичний потенціал [11, 14]

$$\Phi = \Phi\left(\vec{X}_Q, \left\{\vec{X}_j^{(i)}\right\}, \left\{X_k^{(i)}\right\}\right), \qquad (28)$$

диференціал якого набуває вигляду

$$\mathrm{d}\,\Phi = \vec{J}_Q \cdot \mathrm{d}\,\vec{X}_Q + \sum_{j=1}^{5}\sum_{i=0}^{3} \vec{J}_j^{(i)} \cdot \mathrm{d}\,\vec{X}_j^{(i)} + \sum_{k=1}^{2}\sum_{i=1}^{3} \omega_k^{(i)}\,\mathrm{d}\,X_k^{(i)} \,. \quad (29)$$

Співвідношення (28) і (29) дозволяють записати загальну форму кінетичних рівнянь у вигляді

$$\vec{J}_Q = \frac{\partial \Phi}{\partial \vec{X}_Q}, \quad \vec{J}_j^{(i)} = \frac{\partial \Phi}{\partial \vec{X}_j^{(i)}}, \quad \omega_k^{(i)} = \frac{\partial \Phi}{\partial X_k^{(i)}}.$$

Записані співвідношення складають повну систему рівнянь для знаходження введених у розгляд величин.

Зазначимо, що, якщо вільна енергія Гельмгольца $f = \mathrm{u} - Ts$, яка визначена в просторі температури T, коваріантних компонент тензора деформації $\left\{\varepsilon_{\alpha\beta}\right\}$ і концентрацій $\left\{c_j^{(i)}\right\}$, тобто $f = f\left(T, \left\{\varepsilon_{\alpha\beta}\right\}, \left\{c_j^{(i)}\right\}\right)$, є дійсною тричі диференційованою функцією своїх змінних, то мають місце лінійні рівняння стану

$$s - s_0 = T_0^{-1} c_V \mathrm{t} + \sum_{\alpha,\beta} \alpha^{\alpha\beta} \varepsilon_{\alpha\beta} - \sum_{j=1}^{4}\sum_{i=0}^{3} d_{tj}^{(i)} c_j^{(i)},$$

$$\frac{\overline{\sigma}^{\alpha\beta}}{\rho} = \alpha^{\alpha\beta}\mathrm{t} + \sum_{\gamma,\delta} \beta^{\alpha\beta\gamma\delta} \varepsilon_{\gamma\delta} + \sum_{j=1}^{4}\sum_{i=0}^{3} \gamma_j^{\alpha\beta(i)} c_j^{(i)},$$

$$\mu_j^{(i)} - \mu_{j0}^{(i)} = d_{tj}^{(i)}\mathrm{t} + \sum_{\alpha,\beta} \gamma_j^{\alpha\beta(i)} \varepsilon_{\alpha\beta j}^{(i)} + d_j^{(i)} c_j^{(i)}.$$

де $\mathrm{t} = T - T_0$ - відхилення абсолютної температури, $c_j^{(i)} = C_j^{(i)} - C_{j0}^{(i)}$ - відхилення концентрації.

Для ізотропного тіла кінетичний потенціал Ф (28) є дійсною тричі диференційованою функцією термодинамічних сил. Застосовуючи принцип Кюрі, одержимо лінійні кінетичні співвідношення:

$$\vec{J}_Q = L_{QQ}\vec{X}_Q + \sum_{j=1}^{4}\sum_{i=0}^{3} L_{Qj}^{(i)} \vec{X}_j^{(i)},$$

$$\vec{J}_j^{(i)} = L_{jQ}^{(i)}\vec{X}_Q + \sum_{m=1}^{4}\sum_{n=0}^{3} L_{jm}^{(in)} \vec{X}_m^{(n)}, \quad \omega_k^{(i)} = \sum_{m=1}^{2}\sum_{l=1}^{3} \lambda_{km}^{(il)} X_m^{(l)}.$$

Як наслідок умов взаємності Онзагера (27) коефіцієнти рівнянь (39) повинні задовольняти умови

$$L_{Qj}^{(i)} = L_{jQ}^{(i)}, \qquad L_{jm}^{(in)} = L_{mj}^{(ni)}, \qquad \lambda_{km}^{(il)} = \lambda_{mk}^{(li)}.$$

Наслідком другого закону термодинаміки (26) є певні обмеження на коефіцієнти кінетичних рівнянь, зокрема

$$L_{QQ},\, L_{jj}^{(ii)},\, \lambda_{kk}^{(ii)} \geq 0, \qquad L_{QQ} L_{jj}^{(ii)} - L_{Qj}^{(i)2} \geq 0,$$

$$L_{jj}^{(ii)} L_{mm}^{(nn)} \geq (L_{jm}^{(in)} + L_{mj}^{(ni)})^2 / 4, \quad \lambda_{kk}^{(ii)} \lambda_{mm}^{(ll)} \geq (\lambda_{km}^{(il)} + \lambda_{mk}^{(li)})^2 / 4.$$

Надалі необхідно враховувати ці умови та обмеження.

VIII. Висновки

Побудована математична модель взаємозв'язаних теплових, механічних і дифузійних процесів з урахуванням розпаду мігруючої речовини у середовищі з двома шляхами міграції та пастками на основі континуально-термодинамічного підходу [6, 15, 16]. Розглянуто багатокомпонентні системи, в яких кожній компоненті системи поставлено у відповідність континуум, з допомогою якого описано кінематичні та деформаційні властивості компонент. Сформульовано балансові співвідношення, які відображають закони збереження маси, імпульсу та енергії. Використовуючи концепцію локальної термодинамічної рівноваги побудовано рівняння стану та рівняння балансу ентропії, на основі якого записано кінетичні рівняння.

Література References

[1] *В. Г. Барьяхтар, В. В. Гончар, А. В. Жидков, А. А. Ключников,* Радиационные повреждения в лавообразных топливосодержащих материалах объекта «Укрытие» – Чернобыль, 1998, 18 с, (Препр. / НАН Украины. МНТЦ «Укрытие»; 98-12).

[2] *В. А. Каспаров, С. И. Зварич, В. П. Процак, М. А. Журба,* Кинетика растворения чернобыльских топливных частиц. I.Растворение топливных частиц в естественных условиях в почве // Радиохимия, 2000, **42**, №6, С. 533-541.

[3] *І. Р. Юхновський, О. Є. Кобрин, В. В. Токаревський, М. В. Токарчук* Проблеми взаємодії води з паливомісними масами в об'єкті "Укриття" ЧАЕС // Журнал фізичних досліджень, 1997, Т. 1, Число 2, С. 169-180.

[4] *Я. Бэр, Д. Заславски, С. Ирмей,* Физико-математические основы фильтрации воды, М.: Мир, 1971, 451 с.

[5] *Дж. Фрид,* Загрязнение подземных вод, М.: Недра, 1981, 304 с.

[6] *Я. Й. Бурак, Є. Я. Чапля, О. Ю. Чернуха,* Континуально-термодинамічні моделі механіки твердих розчинів, К.: Наукова думка, 2006, 272 с.

[7] Грунтоведение / Под ред. акад. *Е.С.Сергеева*, М.: Изд-во Московского ун-та, 1983, 389 с.

[8] *П. Я. Полубаринова-Кочинова,* Теория движения грунтовых вод.- М.: Недра, 1977, 664 с.

[9] *Л. И. Седов,* Механика сплошной среды. В 2-х т, М.: Наука, 1976, Т.1, 526 с, Т.2, 576 с.

[10] *К. Труздел,* Первоначальный курс рациональной механики сплошной среды, М.: Мир, 1975, 592 с.

[11] *И. Дьярмати* Неравновесная термодинамика, М: Мир, 1974, 304с.

[12] *А. Мюнстер,* Химическая термодинамика, М.: Мир, 1971, 295 с.

[13] *де С. П. Гроот , П. Мазур,* Неравновесная термодинамика, М.: Мир, 1964, 456 с.

[14] *И. Пригожин,* Введение в термодинамику необратимых процессов, М.: Изд-во иностр. литературы, 1960, 127 с.

[15] *Я. Бурак, Є. Чапля, О. Чернуха,* Про вертикальну міграцію радіонуклідів у грунті // Доп. НАН України, 1995, №11, С. 34-37.

[16] *Я. Й. Бурак, Є. Я. Чапля,* Континуальні моделі нелінійної термомеханіки бінарних систем // Фіз.-хім.мех.матеріалів, 1995, №4, С.7-15.

Математичне Моделювання та Аналіз Міцності Деревини

Софія Поберейко
кафедра інформаційних технологій
Національний лісотехнічний університет України
Львів, Україна
Sofi_23@i.ua

Ярослав Соколовський
кафедра інформаційних технологій
Національний лісотехнічний університет України
Львів, Україна
sokolowskyy@ukr.net

Mathematical Modeling and Analysis of the Strength of Wood

Sofia Pobereyko
dept. of Information Technology
National Forestry and Wood Technology University of Ukraine
Lviv, Ukraine
Sofi_23@i.ua

Yaroslav Sokolovskyy
dept. of Information Technology
National Forestry and Wood Technology University of Ukraine
Lviv, Ukraine
sokolowskyy@ukr.net

***Анотація*—синтезовано математичну модель для визначення міцності анізотропних матеріалів в умовах двовісного напруженого стану та проведено її апробацію на деревині листяних та хвойних порід. Зокрема, побудовано криві короткочасної міцності для деревини сосни та дуба. На основі їх аналізу виявлено, що запропонована модель задовільно відображає особливості граничного напруженого стану досліджуваних порід деревини у радіально-тангенціальній площині структурної симетрії.**

***Abstract*—the mathematical model was synthesized to determine the strength of anisotropic materials under biaxial stressed state and its approbation was held using broad – and needle – leaved trees. In particular, the short – term strength curved lines were made for wood of pine and oak. On the basis of their analysis it revealed that the offered model reflects the features of the limiting stressed state of the investigated timber species in radial and tangential plane of structural symmetry satisfactorily.**

Ключові слова—математична модель; міцність; напруження; деформації; відносна об'ємна деформація; густина потенціальної енергії

Keywords—mathematical model; strength; strain; warp; volumetric strain; potential energy density

I. ВСТУП

Підвищення показників якості продукції та зниження енергетичних витрат на її виготовлення є однією з головних умов успішного розвитку деревообробних підприємств. Визначальна роль у вирішенні цієї проблеми належить дослідженням короткочасної міцності деревини зі складним напруженим станом, бо міцність є одним із тих факторів, що істотно обмежують інтенсифікацію процесів зневоднення пиломатеріалів. Тривалість цих процесів повинна бути такою, щоб в усіх точках висушуваного матеріалу значення компонентів тензора напружень не перевищували граничних значень. Інакше матимуть місце залишкові напруження, які є основними чинниками зниження показників якості матеріалу. Для визначення граничних напружень у деревині з одновісним напруженим станом використовують експериментальні методи дослідження, а у деревині зі складним напруженим станом – метод математичного моделювання.

Застосування відомих критеріїв міцності [1] для моделювання кривих та поверхонь міцності деревини з двовісним, плоским та об'ємним напруженими станами є не достатньо обгрунтованим. Жодна із зазначених моделей не адаптована і не апробована на деревині хвойних та листяних порід. Складність вирішення цієї задачі полягає у тому, що вхідними даними цих моделей (критеріїв) є сталі міцності (наприклад, межі зсуву ортотропного матеріалу вздовж діагональних площин структурної симетрії), які на сьогодні не підлягають експериментальному визначенню.

На сьогоднішній день не існує технічних рішень щодо визначення напружень у виробах з деревини, а існуючі механічні теорії міцності з недостатньою повнотою описують їх міцність в умовах складного напруженого стану. Тому задача моделювання міцності деревини є актуальною. Відсутність емпіричних даних з визначення

меж міцності чистого зсуву вздовж головних та діагональних площин структурної симетрії для деревини різних порід не дає змоги підтвердити або заперечити достовірність відомих на сьогодні критеріїв міцності для досліджуваного матеріалу. Зокрема, не адаптовано та не апробовано жодну модель на деревині хвойних та листяних порід.

II. МАТЕМАТИЧНА МОДЕЛЬ ДЛЯ ВИЗНАЧЕННЯ МІЦНОСТІ АНІЗОТРОПНИХ МАТЕРІАЛІВ

Для вирішення цієї задачі побудовано нову математичну модель для визначення короткочасної міцності анізотропних матеріалів в умовах двовісного напруженого стану. Складовими цієї моделі є:

- умова міцності [1]:

$$\frac{1}{2\sqrt{E_{11}E_{22}}}\left(\left(1-\mu_{23}\mu_{32}\right)\sqrt{\frac{E_{22}}{E_{11}}}\sigma_{11}^{2}+\left(\mu_{13}\mu_{32}+\mu_{23}\cdot\right.\right. \quad (1)$$

$$\left.\cdot\,\mu_{31}-\mu_{12}-\mu_{21}\right)\sigma_{11}\sigma_{22}+\left(1-\mu_{13}\mu_{31}\right)\sqrt{\frac{E_{11}}{E_{22}}}\sigma_{11}^{2}\Bigg)-$$

$$-\frac{k}{E_{11}}\left(1-\mu_{23}\mu_{32}+\left(\mu_{23}\mu_{31}-\mu_{21}\right)\sqrt{\frac{E_{11}}{E_{22}}}+\left(\mu_{32}\mu_{21}-\right.\right.$$

$$\left.\left.-\mu_{31}\right)\cdot\sqrt{\frac{E_{11}}{E_{33}}}\right)\left(\sigma_{11}-a\right)-\frac{k}{E_{22}}\left(1-\mu_{13}\mu_{31}+\left(\mu_{13}\mu_{32}-\right.\right.$$

$$\left.\left.-\mu_{12}\right)\sqrt{\frac{E_{22}}{E_{11}}}+\left(\mu_{31}\mu_{12}-\mu_{32}\right)\sqrt{\frac{E_{22}}{E_{33}}}\right)\left(\sigma_{22}-b\right)-c=0$$

- залежності модулів пружності матеріалу від способу навантаження [2]:

$$\mathrm{E}_{\mathrm{ii}}=\begin{cases}E_{ii}^{P} \;\; для\;\sigma_{ii}>0\;i\;\sigma_{jj}-будь-якого,\; i\neq j;\\ E_{ii}^{C} \;\; для\;\sigma_{ii}<0\;i\;\sigma_{jj}-будь-якого,\; i\neq j;\end{cases} \quad (2)$$

- умова для визначення сталих *a, b* моделі (1): значення меж міцності розтягу σ_{ii}^{P-} та σ_{ii}^{P+} в умовах двовісних напружених станах $\left(\sigma_{ii}>0,\sigma_{jj}\to-0,\quad i\neq j\right)$ та $\left(\sigma_{ii}>0,\sigma_{jj}\to+0,\, i\neq j\right)$ є тотожно рівними

$$\sigma_{ii}^{P-}=\sigma_{ii}^{P+}=\sigma_{ii}^{P}. \quad (3)$$

Тут

$$c=\frac{1}{2\sqrt{E_{11}^{P}E_{22}^{P}}}\left(\left(1-\mu_{23}\mu_{32}\right)\sqrt{\frac{E_{22}^{P}}{E_{11}^{P}}}a^{2}+\left(\mu_{13}\mu_{32}+\right.\right. \quad (4)$$

$$\left.\left.+\,\mu_{23}\mu_{31}-\mu_{12}\mu_{21}\right)ab+\left(1-\mu_{13}\mu_{31}\right)\sqrt{\frac{E_{11}^{P}}{E_{22}^{P}}}b^{2}\right);$$

μ_{ij} – коефіцієнти Пуассона; E_{ii} – модулі пружності; σ_{ii} – компоненти тензора напружень; *a*, b і k – сталі коефіцієнти; $E_{\mathrm{ii}}^{\mathrm{P}}$ та $E_{\mathrm{ii}}^{\mathrm{C}}$ – значення модулів пружності розтягу та стиску в і-му напрямку анізотропії матеріалу.

Побудована модель на відміну від відомих дає змогу розрахувати граничні напружені стани як матеріалів із слабкою так і матеріалів із сильною асиметріями меж міцності. Тому оскільки деревина хвойних порід належить до матеріалів із сильною асиметрією меж міцності у тангенціальному та радіальному напрямках ортотропії, а деревина листяних порід – до матеріалів із слабкою асиметрією меж міцності у цих напрямках, то запропонована модель може бути використана для моделювання міцності як деревини листяних так і деревини хвойних порід у радіально-тангенціальній площині структурної симетрід

III. ВИЗНАЧЕННЯ КОЕФІЦІЄНТІВ A І B МАТЕМАТИЧНОЇ МОДЕЛІ

Для вирішення поставленої задачі проведемо дослідження кривої (1) у точках $\left(\sigma_{11}=\sigma_{11}^{P};\sigma_{22}=0;\right.$ $\left.\sigma_{33}=0\right)$ та $\left(\sigma_{11}=0;\sigma_{22}=\sigma_{22}^{P};\sigma_{33}=0\right)$. Для цього розглянемо такі випадки: $\left(\sigma_{11}>0;\,\sigma_{22}\to-0;\,\sigma_{33}\equiv0\right)$, $\left(\sigma_{11}>0;\sigma_{22}\to+0;\sigma_{33}\equiv0\right)$, $\left(\sigma_{11}\to-0;\sigma_{22}>0;\sigma_{33}\equiv0\right)$, $\left(\sigma_{11}\to+0;\,\sigma_{22}>0;\;\sigma_{33}\equiv0\right)$.

У випадку коли $\sigma_{11}>0$, $\sigma_{22}\to-0$, а $\sigma_{33}\equiv0$ значення модулів пружності E_{11}, E_{22} та E_{33}, згідно залежностей (2), дорівнюють відповідно E_{11}^{P}, E_{22}^{C} та E_{33}^{P}, а рівняння міцності (1) має вигляд:

$$\frac{1-\mu_{23}\mu_{32}}{2E_{11}^{P}}\left(\sigma_{11}^{P}\right)^{2}+\frac{k}{E_{11}^{P}}\left(1-\mu_{23}\mu_{32}+\left(\mu_{23}\mu_{31}-\mu_{21}\right)\sqrt{\frac{E_{11}^{P}}{E_{22}^{C}}}+\right. \quad (5)$$

$$\left.+\left(\mu_{32}\mu_{21}-\mu_{31}\right)\sqrt{\frac{E_{11}^{P}}{E_{33}^{P}}}\right)\left(\sigma_{11}^{P}-a\right)-\frac{k}{E_{22}^{C}}\left(\left(\mu_{13}\mu_{32}-\mu_{12}\right)\sqrt{\frac{E_{22}^{C}}{E_{11}^{P}}}+\right.$$

$$\left.+\left(1-\mu_{13}\mu_{31}\right)+\left(\mu_{31}\mu_{12}-\mu_{32}\right)\sqrt{\frac{E_{22}^{C}}{E_{33}^{P}}}\right)b-c=0.$$

Аналогічним чином визначаємо вигляд умови (1) для інших зазначених вище випадків. У результаті одержимо чотири рівняння, які з врахуванням тотожності (3) можна подати у вигляді системи двох рівнянь:

$$\begin{cases} \left(\mu_{23}\mu_{31}-\mu_{21}+\left(\mu_{32}\mu_{21}-\mu_{31}\right)\sqrt{\frac{E_{22}^P}{E_{33}^P}}+\right. \\ +\left(1-\mu_{23}\mu_{32}\right)\left(\sqrt{\frac{E_{22}^P}{E_{11}^P}}+\sqrt{\frac{E_{22}^P}{E_{11}^C}}\right)a+ \\ +\left(\mu_{13}\mu_{32}-\mu_{12}\right)b=\left(\mu_{13}\mu_{32}-\mu_{12}\right)\sigma_{22}^P; \\ \left(\mu_{13}\mu_{32}-\mu_{12}+\left(\mu_{31}\mu_{12}-\mu_{32}\right)\sqrt{\frac{E_{11}^P}{E_{33}^P}}+\right. \\ \left(1-\mu_{13}\mu_{31}\right)\left(\sqrt{\frac{E_{11}^P}{E_{22}^P}}+\sqrt{\frac{E_{11}^P}{E_{22}^C}}\right)b+ \\ +\left(\mu_{23}\mu_{31}-\mu_{21}\right)a=\left(\mu_{23}\mu_{31}-\mu_{21}\right)\sigma_{11}^P \end{cases} \quad (6)$$

З системи рівнянь (6) визначено сталі a і b математичної моделі короткочасної міцності анізотропних матеріалів (1) – (3):

$$a=B_2\frac{B_1\sigma_{22}^P-A_1\sigma_{11}^P}{A_2B_2-A_1B_2}; \quad (7)$$

$$b=A_1\frac{A_2\sigma_{11}^P-B_2\sigma_{22}^P}{A_2B_1-A_1B_2}, \quad (8)$$

де

$$A_1=\mu_{23}\mu_{31}-\mu_{21};\ B_2=\mu_{13}\mu_{32}-\mu_{12} \quad (9)$$

$$A_2=\mu_{23}\mu_{31}-\mu_{21}+\left(\mu_{32}\mu_{21}-\mu_{31}\right)\sqrt{\frac{E_{22}^P}{E_{33}^P}}+$$

$$+\left(1-\mu_{23}\mu_{32}\right)\left(\sqrt{\frac{E_{22}^P}{E_{11}^P}}+\sqrt{\frac{E_{22}^P}{E_{11}^C}}\right); \quad (10)$$

$$B_1=\mu_{13}\mu_{32}-\mu_{12}+\left(\mu_{31}\mu_{12}-\mu_{32}\right)\sqrt{\frac{E_{11}^P}{E_{33}^P}}+$$

$$+\left(1-\mu_{13}\mu_{31}\right)\left(\sqrt{\frac{E_{11}^P}{E_{22}^P}}+\sqrt{\frac{E_{11}^P}{E_{22}^C}}\right); \quad (11)$$

Значення сталої k математичної моделі визначається за формулою наведеною у роботі [1]

$$k=\frac{f\left(\sigma_{11}^P;0\right)-f\left(0;\sigma_{22}^P\right)}{\varepsilon_V\left(\sigma_{11}^P;0\right)-\varepsilon_V\left(0;\sigma_{22}^P\right)}, \quad (12)$$

де

$$f\left(\sigma_{11}^P;0\right)-f\left(0;\sigma_{22}^P\right)=\frac{1-\mu_{23}\mu_{32}}{2E_{11}^P}\left(\sigma_{11}^P\right)^2-$$

$$-\frac{1-\mu_{13}\mu_{31}}{2E_{22}^P}\left(\sigma_{22}^P\right)^2; \quad (13)$$

$$\varepsilon_V\left(\sigma_{11}^P;0\right)-\varepsilon_V\left(0;\sigma_{22}^P\right)=\frac{1}{E_{11}^P}\left(1-\mu_{23}\mu_{32}+\right.$$

$$+\left(\mu_{23}\mu_{31}-\mu_{21}\right)\sqrt{\frac{E_{11}^P}{E_{22}^P}}+\left(\mu_{32}\mu_{21}-\mu_{31}\right)\sqrt{\frac{E_{11}^P}{E_{33}^P}}\Bigg)\cdot$$

$$\cdot\sigma_{11}^P-\frac{1}{E_{22}^P}\left(1-\mu_{13}\mu_{31}+\left(\mu_{13}\mu_{32}-\mu_{12}\right)\sqrt{\frac{E_{22}^P}{E_{11}^P}}+\right.$$

$$\left.+\left(\mu_{31}\mu_{12}-\mu_{32}\right)\sqrt{\frac{E_{22}^P}{E_{33}^P}}\right)\sigma_{22}^P. \quad (14)$$

IV. РЕЗУЛЬТАТИ АПРОБАЦІЇ МАТЕМАТИЧНОЇ МОДЕЛІ НА ДЕРЕВИНІ ЛИСТЯНИХ ТА ХВОЙНИХ ПОРІД

Проведено апробацію моделі (1) – (11) на деревині хвойних та листяних порід. Для цього використано алгоритм [2], практична реалізація якого полягає у: 1) табулюванні функції $\varphi\left(\sigma_{11},\sigma_{22}\right)$ з кроками табулювання $\Delta\sigma_{11}$ та $\Delta\sigma_{22}$; 2) знаходженні для фіксованих значень аргументу σ_{11} таких значень аргументу σ_{22} для яких виконується хоча б одна із умов:

$$\varphi\left(\sigma_{11},\sigma_{22}\right)>0,\quad \text{а}\quad \varphi\left(\sigma_{11},\sigma_{22}+\Delta\sigma_{22}\right)<0 \quad (15)$$

або

$$\phi\left(\sigma_{11},\sigma_{22}\right)<0,\ \text{а}\ \phi\left(\sigma_{11},\sigma_{22}+\Delta\sigma_{22}\right)>0, \quad (16)$$

де $\varphi\left(\sigma_{11},\sigma_{22}\right)$ – функція, значення якої тотожно дорівнюють значенням виразу у лівій частині нелінійного рівняння (1); 3) виборі з пари чисел σ_{22} та $\sigma_{22}+\Delta\sigma_{22}$ такого числа для якого абсолютне значення функції $\varphi\left(\sigma_{11},\sigma_{22}\right)$ є мінімальним; 4) побудові кривих міцності за знайденими точками $\left(\sigma_{11},\sigma_{22}\right)$.

За допомогою цього алгоритму визначено граничні напружені стани деревини сосни та дуба, значення фізико-механічних характеристик яких наведено у табл. 1.

ТАБЛИЦЯ I. ФІЗИКО-МЕХАНІЧНІ ХАРАКТЕРИСТИКИ ДЕРЕВИНИ СОСНИ ТА ДУБА З ТЕМПЕРАТУРОЮ T = 20 °C ТА ВІДНОСНОЮ ВОЛОГІСТЮ W = 12% [3]: У ЧИСЕЛЬНИКУ – ЗНАЧЕННЯ РОЗТЯГУ, А В ЗНАМЕННИКУ – ЗНАЧЕННЯ СТИСКУ

Порода	Модулі пружності, ГПа			Межі міцності, МПа	
	E_a	E_r	E_t	σ_r	σ_t
Сосна	11,9 11,9	0,54 0,67	0,47 0,55	3,23 5,10	2,63 7,50
Дуб	14,2 14,2	1,18 1,40	0,91 1,01	8,0 -	6,5 -
Коефіцієнти Пуассона					
	μ_{ar}	μ_{at}	μ_{rt}	μ_{ra}	μ_{ta}	μ_{tr}
Сосна	0,03	0,037	0,38	0,49	0,41	0,79
Дуб	0,07	0,09	0,34	0,43	0,41	0,83

Побудовано криві міцності для цих матеріалів в умовах двовісного напруженого стану у радіально-тангенціальних площинах структурної симетрії (рис.1).

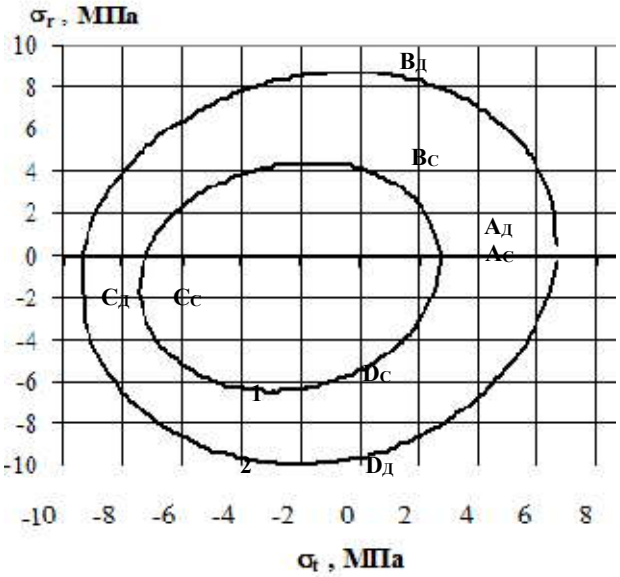


Рис. 1. Криві короткочасної міцності деревини з температурою T=20ºC, відносною вологістю W = 12% та двовісним напруженим станом у радіально-тангентальній площині структурної симетрії: 1 – сосна, 2 – дуб

Проведено аналіз координат таких точок побудованих кривих, які характеризують такий граничний напружений стан матеріалу, характеристики якого є емпірично визначеними. Цими точками є точки перетину цих кривих з осями координат A_C, C_C, B_C, D_C та $A_Д$, $C_Д$, $B_Д$, $D_Д$. Тому що, згідно рис.1, значення абсцис точок A_C і C_C та $A_Д$ і $C_Д$ – це значення меж міцності розтягу та стиску деревини сосни та дуба відповідно у радіальному напрямку ортотропії, а ординати точок B_C і D_C та $B_Д$ і $D_Д$ – це значення меж міцності розтягу та стиску зазначених порід деревини у тангенціальному напрямі, які на сьогоднішній день є відомими і наведені у довідниковій літературі [2]. Аналіз координат цих точок показав, що математична модель (1) – (4) задовільно відображає анізотропію та асиметрію міцності досліджуваних порід деревини вздовж головних напрямків симетрії. Дійсно, оскільки абсциса точки A_C є більшою за ординату точки B_C, а абсолютне значення абсциси точки C_C є меншим за абсолютне значення ординати точки D_C, то межа короткочасної міцності розтягу деревини сосни у радіальному напрямку є більшою за межу міцності розтягу у тангенціальному напрямку, а у випадку стиску навпаки: абсолютне значення межі міцності у радіальному напрямку є меншим за абсолютне значення межі міцності у тангентальному напрямі, що підтверджується результатами експериментальних досліджень [2]. Знайдене співвідношення меж міцності розтягу у радіальному і тангенціальному напрямках анізотропії для деревини сосни має місце і для деревини дуба.

Отже, з огляду на результати відомих на сьогодні експериментальних досліджень короткочасної міцності деревини з одновісними напруженими станами у напрямках анізотропії, математична модель (1) – (12) адекватно описує граничний напружений стан деревини сосни та дуба в умовах двовісного розтягу-розтягу, стиску-стиску та розтягу-стиску. Окрім цього, зазначимо, що характеристики побудованих кривих (рис.1) задовольняють основним евристичним вимогам побудови відомих феноменологічних механічних теорій міцності твердих тіл [4, 5]: 1) криві граничного напруженого стану (криві короткочасної міцності) повинні бути плавними та випуклими; 2) криві короткочасної міцності матеріалу повинні охоплювати початок декартової системи координат.

V. ВИСНОВКИ

Показано, що розрахункові криві короткочасної міцності деревини сосни та дуба з двовісним напружено-деформівним станом у радіально-тангентальній площині структурної симетрії задовільно описують результати експериментальних випробовувань матеріалу на одновісну міцність розтягу та стиску у радіальному та тангентальному напрямках анізотропії. Зокрема, підтверджено: а) межа короткочасної міцності розтягу деревини сосни у радіальному напрямку є більшою за межу міцності розтягу у тангентальному напрямку, а у випадку стиску навпаки: абсолютне значення межі міцності у радіальному напрямку є меншим за абсолютне значення межі міцності у тангентальному напрямі; б) абсолютні значення меж міцності розтягу у радіальному та тангентальному напрямках анізотропії деревини сосни є меншими за відповідні абсолютні значення меж міцності стиску.

ЛІТЕРАТУРА REFERENCES

[1] С. Поберейко, Я. Соколовський, “Математична модель міцності анізотропних матеріалів в умовах двовісного напруженого стану” in *Вісник НУ “Львівська політехніка”: Комп'ютерні науки та інформаційні технології,* Львів: НУ “Львівська політехніка”, 2016, № 604, С. 148–153.

[2] S. B. Pobereyko, O. M. Petriv, B. P.Pobereyko, “Developing software for modellinng the strength of biaxially stressed wood” in *Forestry, Forest, Paper and Woodworking Industry : міжвідомчий наук.-техн. збірник*, Львів : НЛТУ України, 2014, Вип. 40, С. 18 – 20.

[3] А. М. Боровиков, Б. Н. Уголев, Справочник по древесине, Москва: Лесная промышленность, 1989, 296 с.

[4] В. В. Божидарник, Г. Т. Сулим, Елементи теорії пластичності та міцності, Львів: Світ, 199, Т.2, 532 с.

[5] В. Ф. Яценко, Прочность композиционных материалов, К.: Выща школа Головное изд-во, 1988, 191 с.

Problems of Complexity in Cybernetics and Creation of Universal System of Knowledge

Petro Trokhimchuck
Department of theoretical and mathematical physics,
Lesya Ukrayinka East European National University,
Voly av. 13, Lutsk, Ukraine, 43025
trope@yandex.ua

Проблеми Складності в Кібернетиці та Створення Універсальної Системи Знань

Петро Трохимчук
Кафедра теоретичної та математичної фізики
Східноєвропейський національний університет імені Лесі Українки
Пр. Волі 13, Луцьк, Україна, 43025
trope@yandex.ua

***Abstract*—Problems of complexity in modern cybernetics and computer science are discussed. Basic ways of its resolution is analyzed. It shown that this problem has more global value. Polymetric analysis as variant resolution of problem simplicity – complexity is represented.**

***Анотація*—Обговорюється проблема складності в кібернетиці nf syajhvfnbws. Аналізуються основні шляхи її розв'язання. Показано, що ця проблема має більш глобальне значення. Наведений поліметричний аналіз як варіант розв'язання проблеми простоти – складності.**

Keywords—complexity; cybernetics; S. Beer; polymetric analysis; hybrid theory of systems.

Ключові слова—складність; кібернетика; С. Бір; поліметричний аналіз; гібридна теорія систем.

I. Introduction

Problem simplicity-complexity is central problem for each science. Roughly speaking it is the problem of optimal formalization of knowledge. This problem is beginning from famous Archimedes phrase: "Give me fulcrum and I'll reverse Universe" [1]. Therefore, this problem is basic for the creation new science. It was used by Aristotle, Descartes, Newton and other researches.

So problem of complexity is one of central problem in modern mathematics and cybernetics [1 – 8]. This problem is caused in synthetically sciences. Roughly speaking it has two aspects: system (problem of century in cybernetics according S. Beer [1, 3]) and computational (problem of computational complexity [1, 5]. Last problem is included in basic problems of modern mathematics (Smale problems) [1, 5].

As variant of resolution system aspect of problem complexity in cybernetics may be problem simplicity – complexity, which is included in Polymetric Analysis (PA) (universal system of analysis, synthesis and formalization of knowledge) as principle simplicity. Basic elements of this theory – functional numbers is generalizing of quadratic forms.

Hybrid theory of systems (HTS) as element of PA is created on the basis principles (criteria) of reciprocity and simplicity [1, 2, 9]. Only 10 minimal types of formalization system may be used. But number of real systems may be infinite. These systems are differed by step of its complexity. It is may be represented as answer on the one of basic question of modern theory of systems [1, 2, 9] about possible number of systems and its classification with point of simplicity – complexity [1, 2, 9].

Therefore HTS may be represented as variant of resolution the problem of century in cybernetics according S. Beer and may be used for the resolution problem of computational complexity (theory of informative calculations, TIC) [1, 2, 9].

Theory of informative calculations may be represented as variant of resolution of computational complexity [1, 2, 9].

II. Basic Results

We begin this chapter from phrase by S. Beer [3]: "Apparently, the complexity becomes the problem of the century, just as the ability to process natural materials has been a problem of life and death for our forefathers. Our tool must be computers, and their efficiency should be provided by science, able to handle large and complex systems of probabilistic nature. This science may be cybernetics - the science of management processes and communication. The basic thesis of cybernetics can be set forth as follows: there are natural laws behavior of the large manifold systems of any character submits that – biological, technical, special and economic." But nature of some problems may be improbable

too. Therefore this problem must be expanded on problem of formalization of all science and knowledge. But modern science is the realization of the R. Bacon – Descartian concept "Science is so science, how many mathematics is in her" [1]. Development of modern science practically isn't possible without computers.

Computational complexity theory [1, 5] is a branch of the theory of computation in theoretical computer science that focuses on classifying computational problems according to their inherent difficulty, and relating those classes to each other. A computational problem is understood to be a task that is in principle amenable to being solved by a computer, which is equivalent to stating that the problem may be solved by mechanical application of mathematical steps, such as an algorithm. Basic problem of computing complexity is connected with polynomial calculations.

Computational complexity is a concept in computer science and theory of algorithms, depending on the function of indicating the volume of work that is performed by some algorithm, the size of the input data [1, 5]. The science, which studies the computational complexity, was called the theory of computational complexity. The volume of work is usually measured by abstract concepts of time and space, called computational resources. The time is determined by the number of elementary steps needed to solve the problem, while the space defined by memory capacity, or space on the storage medium.

A problem is regarded as inherently difficult if its solution requires significant resources, whatever the algorithm used. The theory formalizes this intuition, by introducing mathematical models of computation to study these problems and quantifying the amount of resources needed to solve them, such as time and storage. Other complexity measures are also used, such as the amount of communication (used in communication complexity), the number of gates in a circuit (used in circuit complexity) and the number of processors (used in parallel computing). One of the roles of computational complexity theory is to determine the practical limits on what computers can and cannot do [1, 5].

Closely related fields in theoretical computer science are analysis of algorithms and computability theory. A key distinction between analysis of algorithms and computational complexity theory is that the former is devoted to analyzing the amount of resources needed by a particular algorithm to solve a problem, whereas the latter asks a more general question about all possible algorithms that could be used to solve the same problem. More precisely, it tries to classify problems that can or cannot be solved with appropriately restricted resources. In turn, imposing restrictions on the available resources is what distinguishes computational complexity from computability theory: the latter theory asks what kind of problems can, in principle, be solved algorithmically [1].

The question why the concept of computational complexity is hard for the verifiable mathematics was discussed by J. Hromkovič [8]. Therefore we must expand this problem on all science with help system with variable hierarchy or variable measure. And therefore we go to polymetric analysis.

III. Polymetric Analysis and Problem of Complexity

Polymetric analysis (PA) was created as alternative optimal concept to logical, formal and constructive conceptions of modern mathematics and theory of information [1, 10]. This concept is based on the idea of triple minimum: mathematical, methodological and concrete scientific [1, 9].

However, one of the main tasks of polymetric analysis is the problem of simplicity-complexity that arises when creating or solving a particular problem or science. It must be open system [1, 2, 9].

The polymetric analysis may be represented as universal theory of synthesis in Descartian sense. For resolution of this problem we must select basic notions and concepts, which are corresponded to PA. The universal simple value is unit symbol, but this symbol must be connected with calculation. Therefore it must be number. For the compositions of these symbols (numbers) in one system we must use system control and operations (mathematical operations or transformations). After this procedure we received the proper measure, which is corresponding system of knowledge and science.

Therefore the basic axiomatic of the polymetric analysis is was selected in the next form [1, 2, 9].

Definition 1. **Mathematical construction** is called set all possible elements, operations and transformations for resolution corresponding problem. The basic functional elements of this construction are called constructive elements.

Definition 2. The mathematical constructive elements $N_{x_{ij}}$ are called **the functional parameters**

$$N_{x_{ij}} = x_i \cdot \overline{x}_j, \tag{1}$$

where x_i , $\overline{x_j}$ – the straight and opposite parameters, respectively; $\cdot$ – corresponding respective mathematical operation.

Definition 3. The mathematical constructive elements $N_{\varphi_{ij}}$ are called the **functional numbers**

$$N_{\varphi_{ij}} = \varphi_i \circ \overline{\varphi}_j . \tag{2}$$

Where $\varphi_i\left(x_1,...x_n,\overline{x_1}...\overline{x_m},...N_{x_{ij}},\ldots\right)$,

$\overline{\varphi_j}\left(x_1,...x_n,\overline{x_1}...\overline{x_m},...N_{x_{ij}},...\right)$ are the straight and opposite functions, respectively; $\circ$ – respective mathematical operation.

Remark 1. Functions φ_I , $\overline{\varphi}_j$ may be have different nature: mathematical, linguistic and other.

Another example may be the orthogonal eigenfunctions of the Hermitian operator.

The theory of generalizing mathematical transformations is created for works on functional numbers [1].

Definition 4. **Qualitative transformations** on functional numbers $N_{\varphi_{ij}}$ (straight A_i and opposite $\overline{A}_j$) are called the next transformations. The straight qualitative transformations are reduced the dimension $N_{\varphi_{ij}}$ on *i* units for straight parameters, and the opposite qualitative transformations are reduced the dimension $N_{\varphi_{ij}}$ on *j* units for opposite parameters.

Definition 5. **Quantitative** (**calculative) transformations** on functional numbers $N_{\varphi_{ij}}$ (straight O_k and opposite $\overline{O}_p$) are called the next transformations. The straight calculative trasformations are reduced $N_{\varphi_{ij}}$ or corresponding mathematical constructive element on *k* units its measure. The opposite quantitative transformations are increased $N_{\varphi_{ij}}$ or corresponding mathematical constructive element on *l* units its measure, i.e.

$$O_k \overline{O}_l N_{\varphi_{ij}} = N_{\varphi_{ij}} - k \oplus l \,. \qquad (3)$$

Definition 8. **Left and right transformations** are called transformations which act on left or right part of functional number respectively.

Definition 9. The maximal possible number corresponding transformations is called **the rang of this transformation**

$$rang(A_i \overline{A}_j N_{\varphi_{ij}}) = \max(i, j); \qquad (4)$$

$$rang(O_k \overline{O}_p N_{\varphi_{ij}}) = \max(k, p) \,. \qquad (5)$$

Remark 2. The indexes *i,j, k,p* are called the steps of the corresponding transformations.

For this case we have finite number of generalizing transformations.

Basic elements of PA were called the generalizing mathematical elements or its various presentations – informative knots [1, 2, 9]. Generalizing mathematical element ${}^{stqo}_{nmab}M_{ijkp}$ is the composition of functional numbers (generalizing quadratic forms, including complex numbers and functions) and generalizing mathematical transformations, which are acted on these functional numbers in whole or its elements [1]. Roughly speaking these elements are elements of functional matrixes.

This element ${}^{stqo}_{nmab}M_{ijkp}$ may be represented in next form

$${}^{stqo}_{nmab}M_{ijkp} = A_i \overline{A}_j O_k \overline{O}_p A^r_s \overline{A}^r_t O^r_q \overline{O}^r_o A^l_n \overline{A}^l_m O^l_a \overline{O}^l_b N_{\varphi_{ij}} \,. \qquad (6)$$

Where $N_{\varphi_{ij}}$ – functional number; $O_k, O^r_q, O^l_a, \overline{O}_p, \overline{O}^r_o, \overline{O}^l_b$; $A_i, A^r_s, A^l_n, \overline{A}_j, \overline{A}^r_t, \overline{A}^l_m$ are quantitative and qualitative transformations, straight and inverse (with tilde), (*r*) – right and (*l*) – left..

We have only 15 minimal types of mathematical transformations, only six are mathematicval in classic sense.

Polyfunctional matrix, which is constructed on elements (6) is called informative lattice. For this case generalizing mathematical element was called knot of informative lattice [1, 2, 9]. Informative lattice is basic set of theory of informative calculations. This theory was constructed analogously to the analytical mechanics [1].

Basic elements of this theory are [1, 2, 9]:

1) Informative computability *C* is number of possible mathematical operations, which are required for the resolution of proper problem.

2) Technical informative computability $C_t = C \sum t_i$, where t_i – realization time of proper computation.

3) Generalizing technical informative computability $C_{t0} = k_{ac} C_t$, where k_{ac} – a coefficient of algorithmic complexity [1].

Basic principle of this theory is **the principle of optimal informative calculations** [1, 2, 9]: any algebraic, including constructive, informative problem has optimal resolution for minimum informative computability *C,* technical informative computability C_t or generalizing technical informative computability C_{to}.

For classification the computations on informative lattices hybrid theory of systems was created [1]. This theory allow to analyze proper system with point of view of its complexity,

The basic principles of hybrid theory of systems are next: 1) **the criterion of reciprocity;** 2) **the criterion of simplicity**.

The criterion of reciprocity is the principle of the creation the corresponding mathematical constructive system (informative lattice). The criterion of simplicity is the principle the optimization of this creation.

The basic axiomatic of hybrid theory of systems is represented below.

Definition 8. The set of functional numbers and generalizing transformations together with principles reciprocity and simplicity (informative lattice) is called **the hybrid theory of systems** (in more narrow sense the criterion of the reciprocity and principle of optimal informative calculations).

Criterion of the reciprocity for corresponding systems is signed the conservation in these systems the next categories:

1) the completeness;

2) the equilibrium;

3) the equality of the number epistemological equivalent known and unknown knotions.

Criterion of the simplicity for corresponding systems is signed the conservation in these systems the next categories:

1) the completeness;

2) the equilibrium;

3) the principle of the optimal calculative transformations.

For more full formalization the all famous regions of knowledge and science the **parameter of connectedness** σ_t was introduced. This parameter is meant the number of different bounds the one element of mathematical construction with other elements of this construction. For example, in classic mathematics $\sigma_t = 1$, in linguistics and semiotics $\sigma_t > 1$. The parameter of connectedness is the basic element for synthesis in one system of formalization the all famous regions of knowledge and science. It is one of the basic elements for creation the theory of functional logical automata too.

Thus we can receive next 10 types of hybrid systems [1, 2, 9]:

1. The system with conservation all positions the criteria of reciprocity and simplicity for all elements of mathematical construction ($N_{\varphi_{ij}}$ and transformations) is called the *simple system.*

2. The system with conservation the criterion of simplicity only for $N_{\varphi_{ij}}$ is called the *parametric simple system.*

Remark 4. Further in this classification reminder of criteria of reciprocity and simplicity is absented. It mean that these criteria for next types of hybrid systems are true.

3. The system with conservation the criterion of simplicity only for general mathematical transformations is called *functional simple system.*

4. The system with nonconservation the principle of optimal informative calculation and with $\sigma_t = 1$ is called the *semisimple system.*

5. The system with nonconservation the principle of optimal informative calculation only for $N_{\varphi_{ij}}$ and with $\sigma_t = 1$ is called the *parametric semisimple system.*

6. The system with nonconservation the principle of optimal informative calculation only for general mathematical transformations and with $\sigma_t = 1$ is called the *functional semisimple system.*

7. The system with nonconservation the principle of optimal informative calculation and with $\sigma_t \neq 1$ is called *complicated system.*

8. The system with nonconservation the principle of optimal informative calculation only for $N_{\varphi_{ij}}$ is called *parametric complicated system.*

9. The system with nonconservation the principle of optimal informative calculation only for general mathematical transformations and with $\sigma_t \neq 1$ is called *functional complicated system*.

10. The system with nonconservation the criteria of reciprocity and simplicity and with $\sigma_t \neq 1$ is called *absolute complicated system*.

With taking into account 15 basic types of generalized mathematical transformations we have 150 types of hybrid systems; practically 150 types of the formalization and modeling of knowledge and science. Only first four types of hybrid systems may be considered as mathematical, last four types are not mathematically. Therefore HTS may be describing all possible system of knowledge. Problem of verbal and nonverbal systems of knowledge is controlled with help of types the mathematical transformations and parameter connectedness [1]. HTS allows classifying all possible knowledge by step simplicity – complexity. It may be used for the representation evolution of development of concrete science from complexity to simplicity. PA may be represented as natural concept of foundation of mathematics. It may be represented as formalization Pythagorean phrase "Numbers rules of the World", Plato concept of three types of numbers: arithmetical (pure mathematics); sensitive (applied mathematics) and ideal (numerology), and variant resolution S. Beer problem in more widely sense [1].

IV. Conclusions

The problem of simplicity – complexity in modern science is discussed. Short historical analysis of this problem is represented. Problems of complexity in modern cybernetics, computer science and mathematics are analyzed. Necessity of creation universal theory for resolution of problem simplicity – complexity in modern science is shown. Basic structure of PA is represented. Basic applications of PA for the resolution of some problems of modern cybernetics and science are analyzed. The questions about the presentation PA as natural concept of foundations mathematics and foundation of science in Descartian sense are analyzed.

References

[1] P. P. Trokhimchuck, Mothematical foundations of knowledge. Polymetrical doctrine, 2nd ed. Lutsk: Vezha-Print, 2014, 624 p. (In Ukrainian)

[2] P. P. Trokhimchuck, "Problem of Simplicity-Complexity in Modern Science and Polymetrical Analysis", *International Journal of Engineering Research and Management*, vol.3, is.7, 2016. – P. 86-95.

[3] S. Beer, "We and complexity of modern world". In *Cybernetics today: problems and propositions.* Moscow: Znaniye, 1976, pp. 7 – 32 (In Russian)

[4] J. Castey, Large systems. Connectedness, complexity and catastrophes. Moscow: Mir, 1982. (In Russian)

[5] S. Smale, "Mathematical problems for the next century,". *Mathematics: frontiers and perspectives.* (American Mathematics Society), 2000, pp. 271–294.

[6] P. P. Trokhimchuck, "Problem of complexity in modern cybernetics and polymetrical analysis as possible way of its resolution", *Proc. Int. Conf. "Artificial Intelligence. Intelligent Transport System"*. Brest: Sytate Technical University Press, 2016, pp. 12-16.

[7] P. P. Trokhimchuck, "Problem of complexity in modern cybernetics and computing science and ways of their resolutions", *Bulletin of Khrson National Technical University*, No. 3(58), 2016, pp. 292-296

[8] J. Hromkovič, Why the Concept of Computational Complexity is Hard for Virifiable Mathematics. *Electroniuc Colloquiun on Computational Complexity.* Report No. 159. 2015. 14 p.

[9] P. P. Trokhimchuck, "Polymetrical analysis: retrospective and perspective," *Int. J. on Recent and Innovation Trends in Computing and Communications,* vol. 4, Jan. 2016, pp. 173 – 183.

[10] Yu. Bogachevskiy, "Formalism and intuitionism in mathematics", *Transactions of Shevchenko Dcientific Society*, vol. 31, 1937, pp. 81 – 98 (In Ukrainian)

Особливості Польоту Квадрокоптера на Малих Висотах

Б. Я. Благітко, Ю. С. Мочульський
Кафедра радіофізики та комп'ютерних технологій
ЛНУ імені Івана Франка
Львів, Україна
blagitko@gmail.com, mochulsky@ukr.net

І. М. Заячук
Центр математичного моделювання
ІППММ ім. Я.С. Підстригача НАН України
Львів, Україна
igorzaj@litech.lviv.ua

Features the Quadrocopter's Flight at Low Altitudes

B. Blagitk, Y. Mochulsky
Department of Radiophysics and Computer Technologies
Ivan Franko Lviv's National University
Lviv, Ukraine
blagitko@gmail.com, mochulsky@ukr.net

I. Zajachuk
Centre of Mathematical Modelling
Pidstryhach Institute for Applied Problems of Mechanics
and Mathematics Ukrainian Academy of Science
Lviv, Ukraine
igorzaj@litech.lviv.ua

***Анотація*—Обґрунтована необхідність аналізу і моделювання польотів квадрокоптера на малих висотах. Шляхом математичного моделювання визначені основні особливості польотів квадрокоптера на малих висотах. Запропоновані методи вимірювання висоти польоту квадрокоптера за допомогою ультразвукового далекоміра.**

***Abstract*—The necessity of analyzing and modeling process the quadrocopter's flight at low altitudes is described. Through mathematical modeling the basic features of the quadrocopter's flight at low altitudes are considered. The measuring methods of low altitudes of the quadrocopter's flight are proposed.**

Ключові слова—квадрокоптер; висотомір; 3-осьовий гіроскоп; 3-осьовий акселерометр; 3-осьовий магнітометр; термометр,керування

Keywords—quadrocopter; altimeter; 3-axis gyroscope; 3-axis accelerometer; 3-axis magnetometer; thermometer; control

I. Вступ

Політ квадрокоптера на малих висотах має суттєві відмінності від польотів на середніх та великих висотах [1,2].

Є дві основні особливості польоту квадрокоптера на малих висотах.

Перша особливість полягає в тому, що вимірювання висоти польоту на малих висотах не доцільно виконувати за допомогою барометра, як це робиться при польотах на великих висотах. Згідно із барометричною формулою біля поверхні Землі зміна висоти на 1 м викликає зміну атмосферного тиску на 0.09 мм. рт. ст. Тому точність вимірювання висоти барометричним висотоміром лежить в межах 10 м, що недостатньо при польоті на висоті кілька метрів. Використовування для вимірювання радіовисотоміра є недоцільним, оскільки тривалість часу проходження сигналу дуже мала (порядку десяти наносекунд). Тому похибка вимірювання висоти висотомірами такого типу становить десятки сантиметрів.

Друга особливість польоту квадрокоптера на малих висотах полягає в тому, що квадрокоптер при польоті повинен опустити "ніс" і летіти з від'ємним значенням тангажу. Якщо використати для вимірювання висоти інфрачервоний висотомір, то для правильного вимірювання висоти він повинен бути напрямлений перпендикулярно до поверхні Землі. Тут є дві альтернативи встановлення напряму прикріплення висотоміра. З одної сторони при вертикальному спуску (наприклад, в процесі приземлення) для правильного вимірювання висот висотомір необхідно встановити в напрямі перпендикулярному до площини квадрокоптера. Але при такому напрямку кріплення висотомір при горизонтальному польоті показуватиме відстань від нього до точки на Землі, яка розміщена не під ним, а дещо ззаду. Якщо ж встановити висотомір під гострим кутом в напрямку "носа", то покази висотоміра при посадці не будуть відповідати дійсній висоті, а даватимуть відстань від квадрокоптера до точки на Землі перед ним.

Для вирішення цих особливостей пропонується вимірювати висоту квадрокоптера при польоті на малих висотах за допомогою ультразвукового висотоміра, який має значно ширшу діаграму направленості. Кріпити ультразвуковий дальномір пропонується під певним кутом

до площини квадрокоптера. Математичному моделюванню можливих похибок при такому кріплені присвячена дана робота.

II. Ультразвуковий далекомір

Для вимірювання висоти польоту квадрокоптера над поверхнею Землі на квадрокоптері розмішений ультразвуковий далекомір HC-SR04. Далекомір працює за принципом сонара, який посилає ультразвуковий пучок хвиль і за тривалістю затримки відбитого від об'єкту сигналу визначає відстань до об'єкту

$$S = V \cdot t / 2 \quad (1)$$

де S – відстань до об'єкту, V – швидкість звуку в повітрі, t – час між відправленням сигналу і отриманням відбитого сигналу. Далекомір HC-SR04 дає на виході прямокутний імпульс (Echo), тривалість якого дорівнює часу t. На рис. 1 приведена часова діаграма роботи ультразвукового далекоміра HC-SR04 [3].

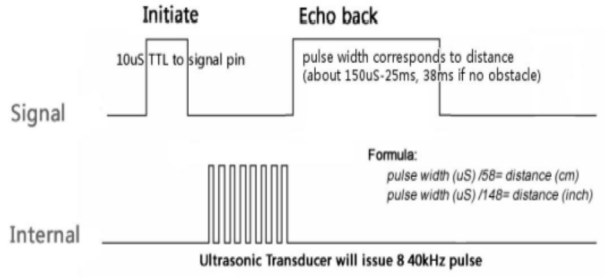


Рис. 1. Часова діаграма роботи ультразвукового далекоміра HC-SR04.

Залежність швидкості звуку V в м/с у повітрі від температури T описується формулою

$$V = 332 + 0{,}6 \cdot T \quad (2)$$

де T – температура в °С. З врахуванням цієї залежності при висоті польоту квадрокоптера 3.5 м тривалість імпульсу Echo при температурі повітря -30°С становитиме 22.4 м/с, а при температурі +30°С зменшиться до 20 мс. Якщо не враховувати залежність швидкості звуку від температури і прийняти її рівною 340 м/с (як описано в технічних даних), то при висоті квадрокоптера 3.5 м висотомір покаже при T=-30°С висоту 3.79 м, а при T=+30°С покази зменшаться до 3.4 м. З цього випливає, що для точнішого вимірювання висоти необхідно в квадрокоптері поставити термометр і в мікроконтролері враховувати залежність швидкості звуку від температури.

Оскільки в основі роботи пристрою використовується ультразвук, сонар погано надається для визначення відстані до звукопоглинальних об'єктів.

На рис. 2 приведена діаграма направленості ультразвукового далекоміра HC-SR04.

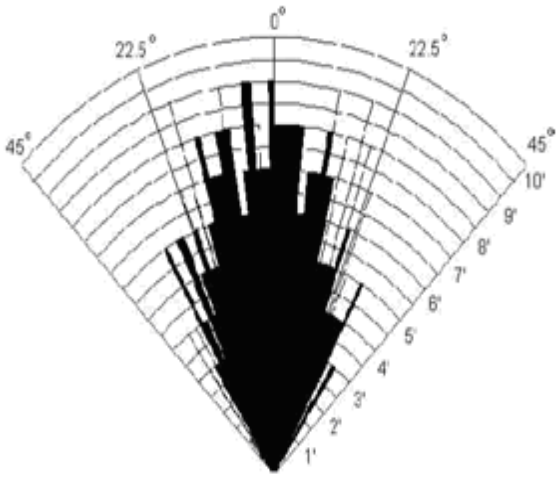


Рис. 2. Діаграма направленості ультразвукового далекоміра HC-SR04.

Ультразвуковий далекомір HC-SR04 має такі основні параметри:

вимірювання відстані – від 2 см до 4 м.

Ефективний кут вимірювання <15°.

Період вимірювання 50 мс.

Розглянемо перший варіант кріплення висотоміра (перпендикулярно до площини квадрокоптера). Тоді при польоті з кутом тангажу β головна вісь висотоміра буде відхилена від вертикалі на цей кут (рис. 3).

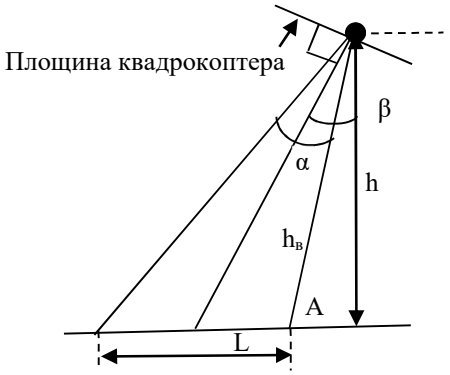


Рис. 3. Висотомір розміщений перпендикулярно до площини квадрокоптера.

Кут вимірювання висотоміра дорівнює α. Як видно з рис. 3, ультразвуковим імпульсом опромінюється ділянка Землі позаду квадрокоптера у формі еліпса з довжиною великої осі еліпса L. Відстань до об'єкта висотомір визначає за початком відбитого сигналу. Найшвидше прийде сигнал, відбитий від точки А, тобто висотомір покаже відстань $h_в$. Як видно з рис. 3,

$$h_в = h / cos(\beta - \alpha / 2) . \quad (3)$$

Відносна похибка вимірювання висоти дорівнюватиме

$$\delta = 1 - 1 / cos(\beta - \alpha / 2) . \quad (4)$$

Формули (3) і (4) справедливі лише при $\beta > \alpha / 2$. Як витікає з цього, висотомір даватиме правильні покази висоти квадрокоптера при куті тангажу $\beta \le \alpha / 2$. Для вибраного висотоміра цей граничний кут становитиме 7.5°. При моделюванні горизонтального польоту квадрокоптера із швидкістю 1.8 м/с кут тангажу становив 5.7°, тобто, існує деякий запас для точного вимірювання висоти.

Розглянемо другий випадок (вісь висотоміра направлена під кутом γ до перпендикуляру до площини квадрокоптера (рис. 4)).

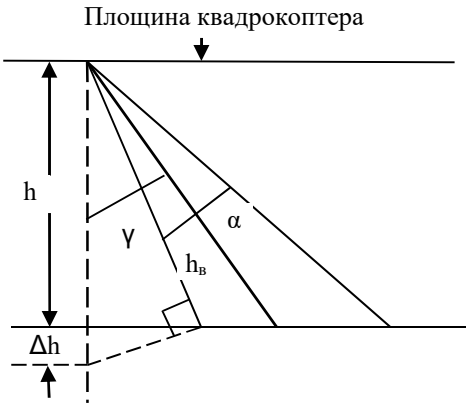


Рис. 4. Висотомір розміщений під кутом до перпендикуляру до площини квадрокоптера.

В цьому випадку при вертикальному приземленні висотомір покаже значення

$$h_в = h / cos(\gamma - \alpha / 2) , \quad (5)$$

яке на величину Δh більше від дійсної висоти. Відносна похибка вимірювання становитиме

$$\delta = 1 - 1 / cos(\gamma - \alpha / 2) . \quad (6)$$

З цих формул і рис.4 випливає, що для точного вимірювання висоти кут γ не повинен перевищувати α/2, тобто 7.5°.

Отже, при польоті з швидкістю 1.8 м/с, вертикальному наборі висоти і вертикальному приземленні похибка від напряму висотоміра дорівнюватиме нулю, якщо кут нахилу висотоміра не перевищуватиме 7.5°.

Для більшої швидкості польоту необхідно збільшити кут тангажу. Залежність кута тангажу від швидкості польоту V задається формулою

$$\theta = arctan\left(\frac{V \cdot A_x}{m \cdot g}\right), \quad (7)$$

де A_x – коефіцієнт опору повітря, m – маса квадрокоптера, g – прискорення вільного падіння.

Якщо для заданої швидкості кут тангажу перевищує 7.5°, то для будь-якого кута нахилу висотоміра при вертикальному приземлені виникатиме похибка. Як випливає з рівнянь (4) і (6), для того, щоб мінімізувати цю похибку необхідно висотомір направити під кутом

$$\gamma = \theta / 2 \quad (8)$$

III. МАТЕМАТИЧНА МОДЕЛЬ КВАДРОКОПТЕРА

Математична модель квадрокоптера в польоті з врахуванням особливостей сенсорів розглядається як дві суттєво різні математичні моделі: власне математичної моделі квадрокоптера, як безпілотного літального апарату, на вхід якого поступають значення напруг живлення електричних двигунів, а виходами є кутові швидкості обертання валів електричних двигунів; математичної моделі пропорційно-диференціального (ПД) регулятора, на вхід якого від сенсорів поступають виміряні значення, а мікроконтролер реалізує алгоритм ПД регулятора і видає на вихід значення напруг живлення двигунів.

Математична модель квадрокоптера описується системою диференціальних рівнянь 16-го порядку у формі Коші [4].

Для керування польотом квадрокоптера використовують регулятор, який задає необхідні напруги на двигуни пропелерів. В даній роботі вибрано пропорційно-диференційний регулятор, який регулює висоту польоту та кути нахилу квадрокоптера, змінюючи вертикальну складову сили тяги пропелерів та моменти сил, які обертають квадрокоптер.

Висоту Z визначають висотоміром. Вертикальну швидкість вичислюємо, чисельно продиференціювавши висоту квадрокоптера. Кутові швидкості ω_γ, ω_θ та ω_ψ, які повинен дати 3-осьовий гіроскоп, це відповідні змінні стану моделі квадрокоптера. Маючи покази гіроскопа за допомогою формул чисельного інтегрування вичислюємо кути тангажу, крену та рискання. Врахувавши при цьому, що гіроскоп має зміщення нуля Δ, отримаємо формулу для обчислення кута крену в момент часу t_{i+1} на основі значень цих кутів в момент часу t_i.

$$\gamma_{i+1} \approx \gamma_i + \omega_{\gamma,i} \cdot (t_{i+1} - t_i) + \Delta_\gamma \cdot (t_{i+1} - t_i) \quad (9)$$

Аналогічно обчислюються кути тангажу та рискання. За рахунок похибки нуля гіроскопа (третього доданку) похибка визначення кутів тангажу, крену та рискання

буде накопичуватися, що приводить до відхилення траєкторії польоту від заданої.

Для того, щоб провести корекцію кутів положення квадрокоптера використовують обчислення цих кутів за показами акселерометра. 3-осьовий акселерометр надає проекції (A_{xo}, A_{yo}, A_{zo}) суми вектора прискорення вільного падіння і вектора абсолютного прискорення об'єкту на осі X_o, Y_o, Z_o системи координат, зв'язаної з об'єктом. При нерухомому об'єкті або його рівномірному русі наведені вище проекції будуть проекціями вектора прискорення вільного падіння $\vec{G}$ на осі X_o, Y_o, Z_o квадрокоптера. Для корекції кута рискання використовують не акселерометр, а магнітометр. 3-осьовий магнітометр визначає проекції вектора магнітної індукції на осі X_o, Y_o, Z_o квадрокоптера (відповідно B_{xo}, B_{yo} та B_{zo}).

Для аналізу використовувався неявний метод Рунге-Кутта TR-BDF2 (на першому етапі рівняння інтегрують методом трапецій, а на другому етапі застосовують формулу диференціювання назад другого порядку) [5].

IV. МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ПОЛЬОТУ КВАДРОКОПТЕРА НА МАЛИХ ВИСОТАХ

Для прикладу проведене моделювання автономного польоту квадрокоптера на висоті 3.5м над ділянкою розміром 200х30 м..

На рис. 5 приведена траєкторія автономного польоту квадрокоптера на висоті 3.5м, яка отримана в результаті моделювання. Злет квадрокоптера з горизонтальної поверхні відбувається у вертикальному напрямі [6].

На горизонтальній ділянці знаходиться одна дзвоноподібна гірка висотою 3.0 м із координатами центра X=35 м,Y=40 м та діаметром 50 м. Також на ділянці знаходиться одна яма у формі відбитка дзвона глибиною 3.0 м із координатами центра X=90 м,Y=50 м та діаметром 50 м

На цьому рисунку відлік висоти Z ведеться від горизонтальної поверхні. Для кращої візуалізації траєкторії польоту квадрокоптера на рисунку проведені горизонталі через горбик і яму.

Траєкторія польоту задається координатами точок, через які повинен пролетіти квадрокоптер. Послідовність цих координат наступна: [0, 30 м], [120 м, 30 м], [130 м, 40 м], [0, 40 м], [10 м, 50 м], [120 м, 50 м], [110 м, 60 м], [0, 60 м].

Перший відрізок проходить по схилу горбика і через центр ями. На ньому видно невеликий підйом і великий спуск квадрокоптера. Третій відрізок проходить через центр горбика і схил ями. Відповідно квадрокоптер спочатку трохи спускається, а потім піднімається на 3 м. На наступних відрізках квадрокоптер проходить по схилах горбика і ями, а на останньому поза ямою. Окрім того, показаний вертикальний злет квадрокоптера з місця старту (координати X=0 м,Y=30 м).

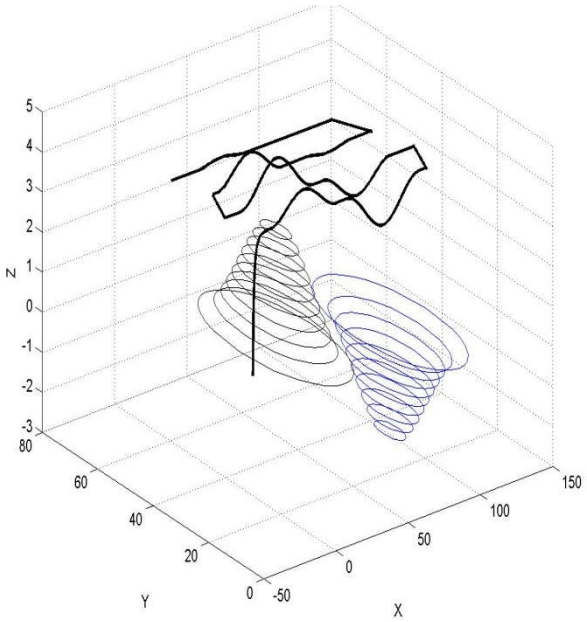


Рис. 5. Траєкторія автономного польоту квадрокоптера на висоті 3.5м.

V. ВИСНОВКИ

Результати проведених досліджень визначають ефективність та практичну значимість використання запропонованих методів вимірювання висоти польоту квадрокоптера. А саме, для вимірювання висоти польоту на малих висотах доцільно використовувати ультразвукові висотоміри. При використанні ультразвукового висотоміра HC-SR04 для зменшення похибки вимірювання висоти необхідно направити вісь висотоміра під кутом до перпендикуляра до площини квадрокоптера $\gamma = \theta_m / 2$, де θ_m – максимальний кут тангажу при польоті.

ЛІТЕРАТУРА REFERENCES

[1] H. Huang, G. M. Hoffmann, S. L. Waslander, C. J. Tomlin, "Aerodynamics and control of autonomous quadrotor helicopters in aggressive manoeuvring" in "IEEE International Conference on Robotics and Automation", May 2009, pp. 3277–3282.

[2] Becker Marcelo, In-Flight Collision Avoidance Controller Based Only on OS4 Embedded Sensors/ R.C.B.Sampaio, S.Bounabdallah, V.de Perrot, R.Siegwart// Journal of the Brazilian Society of Mechanical Sciences and Engineering. vol.34 no.3 Rio de Janeiro July/Sept. 2012 – March 2012. – Vol. XXXIV, No.3, July-September 2012, С. 294–107.

[3] http://www.robotshop.com/media/files/pdf/datasheet-sen026.pdf

[4] Б. Благітко, І. Заячук, Ю. Мочульський, Математичне моделювання впливу особливостей сенсорів на процес польоту безпілотного квадрокоптера, "Фізико-математичне моделювання та інформаційні технології", 2015, Вип. 21, С. 22–29.

[5] Ю. Мочульський, Matlab у фізичних дослідженнях: навч. - метод. посіб, Львів: ВЦ ЛНУ ім. Івана Франка, 2004, 121 с.

[6] Б. Благітко, І. Заячук, Л. Кіт. Ю. Мочульський, Моделювання процесу керування вертикальним злетом і приземленням безпілотного квадрокоптера, "Фізико-математичне моделювання та інформаційні технології", 2012, Вип.2, С.111–117

Програмний Комплекс для Дослідження і Порівняння Завадостійких Кодів

Юрій Гунченко
кафедра математичного забезпечення комп'ютерних систем
Одеський національний університет імені І.І.Мечникова
Одеса, Україна
7996445@mail.ru

Павел Ємельянов
кафедра математичного забезпечення комп'ютерних систем
Одеський національний університет імені І.І.Мечникова
Одеса, Україна
ipauler@gmail.com

Тетяна Щербакова
кафедра математичного забезпечення комп'ютерних систем
Одеський національний університет імені І.І.Мечникова
Одеса, Україна
ll_papilion@mail.ru

Вікторія Левчук
кафедра математичного забезпечення комп'ютерних систем
Одеський національний університет імені І.І.Мечникова
Одеса, Україна
levchuk.viktoria@stud.onu.edu.ua

Software for Research and Comparison of Noise-immune Codes

Yurii Gunchenko
department of mathematical support of computer systems
Odessa I.I. Mechnicov National University
Odessa, Ukraine
7996445@mail.ru

Pavel Emelyanov
department of mathematical support of computer systems
Odessa I.I. Mechnicov National University
Odessa, Ukraine
ipauler@gmail.com

Tetyana Shcherbakova
department of mathematical support of computer systems
Odessa I.I. Mechnicov National University
Odessa, Ukraine
ll_papilion@mail.ru

Viktoria Levchuk
department of mathematical support of computer systems
Odessa I.I. Mechnicov National University
Odessa, Ukraine
levchuk.viktoria@stud.onu.edu.ua

***Анотація*—В роботі запропоновано концепцію побудови програмного комплексу для дослідження завадостійких кодів з наявністю і без наявності помилок у повідомленні. Розглянуто необхідні модулі комплексу програмного забезпечення, які повинні здійснювати кодування, декодування повідомлень, внесення помилок, модулі зберігання, обробки та виводу даних, наведено умови і алгоритми їх функціонування. Показано доцільність використання OLAP-систем для зберігання і обробки отриманих даних. Наведено результати функціонування комплексу програмного забезпечення, створеного згідно концепції, що пропонується, на прикладі дослідження характеристик чотирьох кодів двох різних типів.**

***Abstract*—The paper introduces the concept of development a software system for research of noise-immune codes with and without errors in the message. This paper presents the necessary modules of software complex, that provide encoding, message decoding, introducing errors, data storage, processing and output modules and the operating conditions and algorithms as well. The expediency of using OLAP-systems for storing and processing of received data are shown in the work. The results of software complex operating, that was created under proposed concept, in terms of study the characteristics of four codes of two different types, is shown in the work.**

Ключові слова—завада, перешкода, завадостійкий код, програмне забезпечення, алгоритм побудови програмного модуля, база даних, OLAP-система

Keywords—noise; noise-immune code; software; software module creating algorithm; database; OLAP-system

I. INTRODUCTION

Obstacles and noises are undesirable factors of destructive character in different technical systems, because it interferes with the normal operation of such systems. It forces the creators of computer systems, to take certain actions to minimize the impact of noise on a particular system.

Because of the different nature of its appearance, noise can alter the data transmission between different computer systems or within the one system.

Noise immunity is a property of the system of performing its functions the most effectively if noise and distortion are available. Noise immunity is evaluated by the intensity of noises in which the dysfunction of the system (device) does not exceed the permissible limits. The stronger noise at which the system continues to operate properly and effectively is, the higher is its immunity.

II. THE REVIEW OF NOISE-IMMUNE DEVICES

A. Selecting a Template.

There is a wide range of outfits, devices and equipment, where resistance to interference and noise is essential for their proper operating. This may include personal computers, navigation systems, satellite communications, ballistic missiles.

Most often term “noise-immunity” is used to characterize the information transmission devices (for example, communication lines) or surveillance devices (for example, radars). For this reason in most cases, the term “signal” could be determined, and the noise-immunity evaluation can be made by considering the relation between interference and signal, at which the required quality of normal operation is provided. For example, in the radiolocation – the relation signal to the noise at which the specified detection reliability is provided (probability of correct detection at a certain probability of false alarm).

Building up of optimal devices that implement the required level of noise-immunity is usually very difficult, and their inevitable technical imperfections do not allow to reach this level in full. That is why the devices at the greatest simplicity which provide good approaching to the optimum level are applied for.

There are many ways to reduce the impact of noise on transmitted or stored data. The means of noise-immune coding based on certain noise-immunity codes are the most frequently used.

The work objective is to develop the concept and principles of design software to explore methods of encoding and decoding information in the presence and absence of errors. This software package will allow providing virtual research of systems that are being developed without their direct full-scale modeling or creation of prototypes.

III. REQUIREMENTS OF COMPLEX

Complex of software should be as a basis for analysis codes such as characteristics of code as follows: encoding, codeword size and length of the original character sequence. It should be noted that this system is designed to compare the different codes with each other. Its function capabilities should contain comparison of different codes in the same conditions. This approach allows getting adequate and correct information about the difference between the opportunities of codes, depending on their properties.

This complex is proposed to implement not as a monolithic system, but in the form of several discrete modules that are interconnected, where each module will perform its purpose. Interaction between the modules should be partially automated and partially supported by the operator. This solution allows you to make the software package more flexible, because such implementation facilitates rapid adjustments by changing the functionality in one or another part of the complex. Thus, the implementation of changes will cause relatively lower costs than the introduction of changes into a single monolithic system.

A. Algorithm Of Operation.

Let us describe the life cycle of a message that is processed by software complex, in the form of algorithm:

Message in a symbolic form is submitted in module that performs encoding (module "CODER").

Module "CODER" creates a binary representation of characters of the message based on their serial number in the ASCII table.

Module "CODER" performs apropriate to right code (for example Hemming code) algorithmic operations and writing data about the length of word that is encoded, and encode time in a log-file.

"CODER" sends the resulting encoded binary sequence to module "DECODER" through the communication channel.

Unit generating errors works where appropriate

Module "DECODER" processes received messages with necessary algorithms and writes data in log-file that are similar to data of encoding side.

Based on log-files of modules "CODER" and "DECODER", the recording into the database is implemented.

The database is processed using functional capabilities of OLAP-analyzer.

A graphical representation of the data is carried out.

Fig. 1 shows the diagram that demonstrates the interaction between the modules.

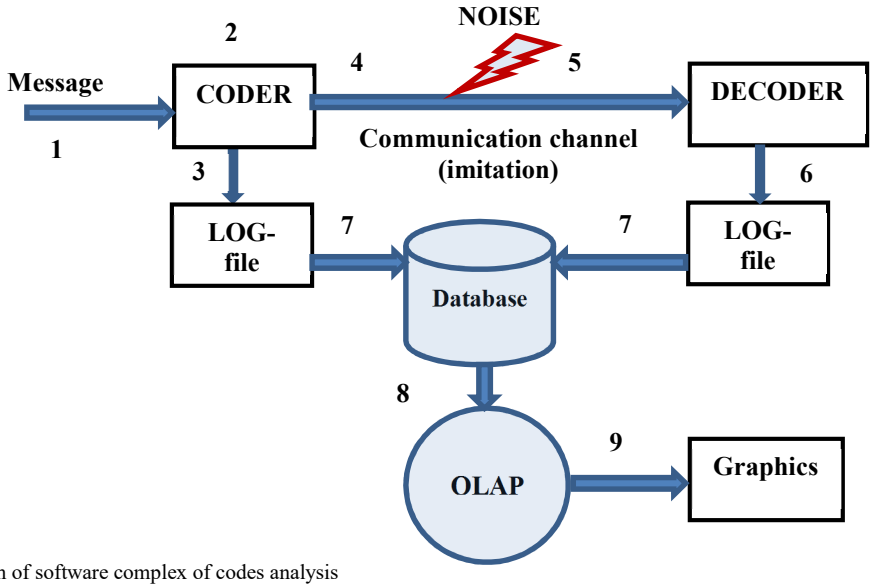


Fig. 1 Functional diagram of software complex of codes analysis

B. Design of "CODER" module

The module of coding messages for its intended use should perform the following functions:

translation of character sequences in a binary sequence consisting of symbols **1** and **0**;

coding of binary sequence using the appropriate algorithm and transferring encoded sequence through the communication channel that is imitated, precisely from the decoding side;

data output about the encoding sequence, and time spent on the process of coding this sequence.

Under the real conditions, encoding and decoding operations are often implemented using hardware. In the software complex, the analysis of different types of codes used for noise-immune coding will be investigated in software that imitates the hardware part.

It is assumed that the three above capabilities of this module will be implemented through the mechanism of using functions, such as writing discrete, logical coherent sections of a code, which can be called many times afterwards, transferring them the necessary data to process, or fully edit and modify.

This method will allow performing the whole spectrum of the module functionality in the form of software.

C. Design of generating errors module.

Under the conditions of software system, generating error module is an essential element for carrying out the code analysis. When transmitting information to the real conditions through the real communication channel, information can change due to interference, which in its turn can cause changes in the symbols in the transmitted sequence. As this complex does not involve information exchange between encoding and decoding parties by means of the communication channel, the coded sequence is devoid of risk to be changed.

It is good with hardware encoding, however, as the software system is developed for the analysis and comparison of codes and their capabilities, the wider range of information for analysis can be received having information about how codes correct errors.

Based on the above, there is a need to generate errors to test the time of algorithm work while correcting these errors. The module of generating errors will be considered as a generator of those errors. Processing of the transmitted coded message will be performed by generating errors block using the following algorithm:

an encoded sequence comes from the module "CODER" to block of generation of errors;

an error generating block changes the information bit of encoded message;

an error generating block transmits a coded binary sequence to the next module "DECODER"

This algorithm is quite simple, but we should note a few points. Using generation errors block will apply to encoded sequence strictly if it is necessary. So, it is assumed that each encoded sequence is transmitted without errors first, and then with error in the information bit that is generated as a result of the work of generating error module. Subsequently, the results of the code when decoding encoded binary sequence without errors and sequence with generated error will be recorded in the log-file at work of the module "CODER".

D. Design of "DECODER" module

The decoding module of encoded binary sequence is a module that receives data from the module "CODER" or generating error module.

Due to its appointment, module "DECODER" should perform the following functions:

receive the input encoded binary sequence from the generating error module or directly from the module "DECODER" passing generating error module;

implement decoding of the received encoded binary sequence consisting of symbols 1 and 0 according to the required decoding algorithm;

submit the output decoded sequence of characters and to input the required data in the log-file.

It should be noted that the decoding algorithm must comply with the encoding algorithm, not only by their type (e.g. Hemming code or CRC), but also by the key characteristics such as length codeword and the number of control symbols. Data recorded in the log file by module "DECODER" are completely the same as the data recorded in the log file of module "CODER" – precisely word length and decoding time. Subsequently, these data will be required to analyze and compare the effectiveness of various codes with different characteristics.

E. Design of database and its write action module

One of the key parts of the software complex is a database that will contain the information from log files of "CODER" and "DECODER" modules. One module should include both the database and the software that implements the data filling based on log-files.

The first step is to determine the database structure, such as its tables and fields in these tables. The database consists of five tables. These tables (for example) **Sequences**, **CodeTypes**, **Encodes**, **Decodes** and **DecodesEr**, intended for recording sequences data to encoding, algorithms and results of encoding, decoding and decoding with error, thereafter. The information required for the analysis should be entered into these tables.

The database structure is presented in Fig. 2.

Also, it is necessary to describe the functional capabilities of the program, intended to fill the database by available information. It can be made as an algorithm:

the program must read data from log files;

give the read data to the appropriate format to recording in a database;

initiate a connection to the database;

record in the appropriate fields of tables.

This algorithm will allow distributing data from log-files in appropriate table field most reliably. Also, the use of object-oriented language in this module will allow recording the necessary data in the database fast and on relatively high level.

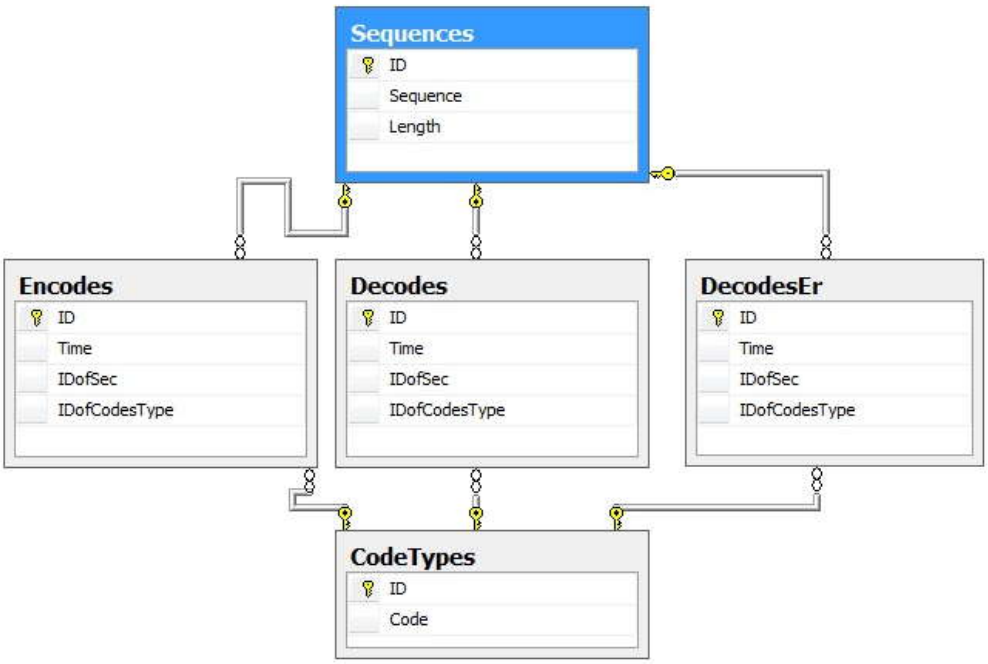


Fig. 2 The database structure

F. Design of OLAP tools

OLAP-systems allow working with data, precisely representing them in aggregated form for further convenience work with it. Also, OLAP-systems help to take the required decision that should be taken in the future, based on existing statistics. Often OLAP-systems are not separate solutions but integrated into other systems, such as decision support systems.

As data storage for OLAP-systems, data sources of different types can be used. The relational database MsSQL will be used for the proposed software module

G. *Design of visual data presentation tools*

The final step in the design of this software complex is the visual representation of data.

This step is the most simple in terms of implementation, because of the appropriate use of the finished software.

Part of the work that shows the different types of codes will be efficiently implemented by the form of graphs based on the received data.

Data that will be presented in graphs, will be received from the previous stage, where OLAP-component of this complex is implemented, which, in its turn, will handle the processing of information from the database.

This approach allowed implementing the software system, without negative affecting its quality, by the least costly way.

IV. Results of Operation and Test of Adequacy

To test the adequacy of the proposed approach of designing software complex, with its help, a research and analysis of the two types of codes that are widely used – Hamming code and Inverse code were conducted. Characteristics of the analyzed codes are shown in Table 6.

TABLE I. Characteristics of the Analyzed Codes

Code type	Hamming code	Hamming code	Inverse code	Inverse code
The number of information bits	8	16	8	16
General number of bits	12	21	16	32
Number of control bits	4	5	8	16
Specificity of work with errors	Corrects 1 error	Corrects 1 error	Checks for errors	Checks for errors
Denotation	HEM(8,12,4)	HEM(16,21,5)	INV(8,16,8)	INV(16,32,16)

Thus, four codes of two types are involved in comparison. Also, Hamming code has the ability to correct one error and inverse code operates on the principle of validation of the data.

General analysis of the code that is used further is a comparing of temporal characteristics of each code on such characteristics as time of encoding the information symbol sequence, decoding of coded sequence and time of decoding of symbol sequence with an error.

The result of the analysis becomes data shown graphically in Fig. 3 and Fig. 4.

Let us compare the four codes of two types, precisely Hamming code and Inverse code on parameters such as time of encoding, time of decoding and time of decoding with error. These operations were performed on eight information symbolic sequences of varying lengths – from 2 to 512 bits.

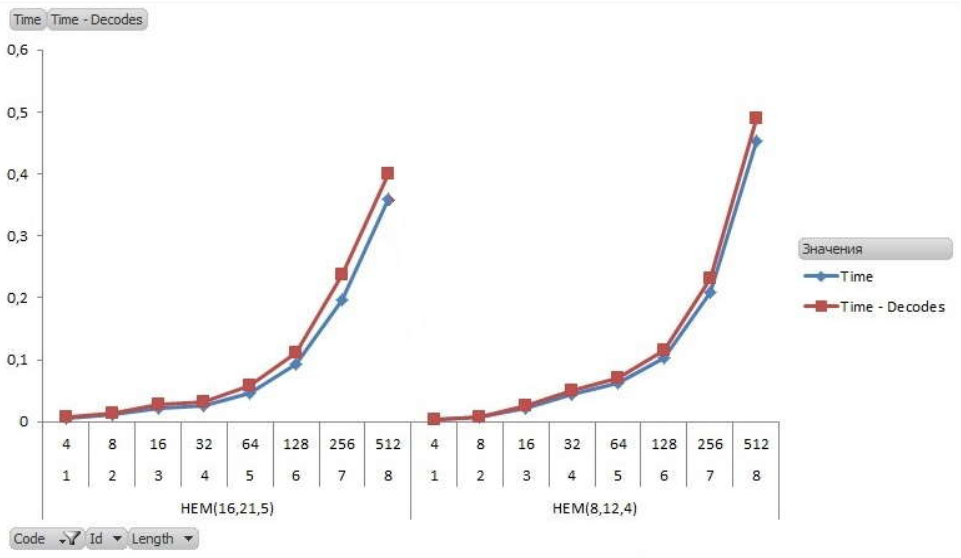


Fig. 3 Results of the study Hamming code

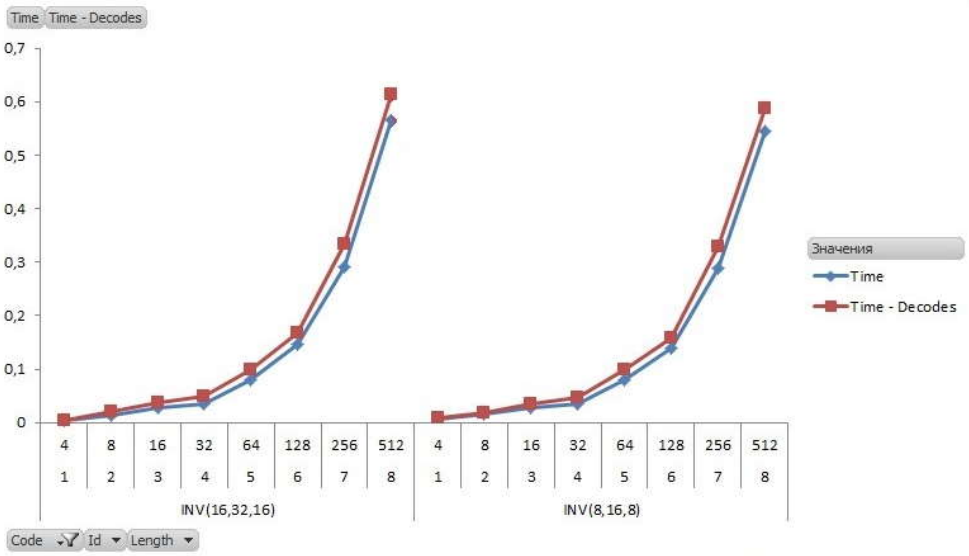


Fig. 4 Results of the study Inverse code

As a result of the software complex work, the advantages and disadvantages of each of the four codes were shown, such as code INV (16, 32, 16) by the number of bits of information and control bits twice those for the code INV (8, 16, 8). However, it performs the necessary operations in almost the same lap times as the code INV (8, 16, 8). This is caused by the unreasonableness of using this type of code in comparison with the Hamming code. The situation in which specific features of inversion code are necessary may be considered as the exception.

It was also determined that the Hamming code HEM (16, 21, 5) spends the least time on a work. This code in this situation can be called the most appropriate in terms of run-time encoding and decoding operations. It should also be noted that the code HEM (16, 21, 5) is a version of the code HEM (8, 12, 4), but with the increased numerical parameters. As the results of studies of these codes, code HEM (16, 21, 5) is much faster than code HEM (8, 12, 4). This allows suggesting that with increase characteristics in this type of code, the time of coding will be reduced. In other words, with the increase of number of information bits and control bits, productivity of the code will be increased too.

V. CONCLUSIONS

In this paper the concept of development was proposed and software complex to noise-immunity testing and analysis codes was implemented. The complex includes four components, and also two-party component in the form of software packages were involved.

This software complex allowed conducting a test comparison of the four codes of two types, precisely two implementations of the Hamming code and two implementations of Inversion code. During the testing and further analysis, the advantages and disadvantages of the investigated codes were identified. The advantage of this approach is the fact that the analysis was based on actual data obtained as a result of work of each code.

Also, it should be noted that this complex is only intended to emulate the actual processes of encoding and data transmitting, because encoding and decoding operations are implemented in software level. There is a possibility that the hardware implementation of encoding and decoding operations may vary from implementation presented in this paper.

However, with similar approaches in the implementation of both software and hardware, the similar nature of the relationships should be expected.

In addition, the results obtained in this work are based on the usage of software complex for code testing and analysis, are relative because they were obtained as a result of work of a specific computer with specific characteristics. Therefore it is fair to say that the time spent on encoding and decoding by similar codes, can change in one direction or another, depending on the technical characteristics of the platform.

One of the major advantages of the concept of software system is its versatility in terms of using other codes. It allows you to test, analyze and compare different implementations of different noise-immune codes under the same conditions. This result is achieved through the specific of call of modules functions that encode and decode: mechanism of function having almost nothing to do with the implementation of the function and allows you to give the same consistency to the input of many different functions.

ЛІТЕРАТУРА REFERENCES

[1] Lineynie blochnie kodi [Electronic resource] - Access mode: http://sernam.ru/book_tec.php?id=80.

[2] Kodi s obnarujeniem oshibok [Electronic resource] - Access mode: http://informkod.narod.ru/5_1item.htm.

[3] Kutuzaki A.S., Gunchenko Y.A. Analiz problem razbieniya na kadri dannih // Abstracts of the Tenth All-Ukrainian conference of students and young professionals "Informatics, information systems and technology". – Odessa: - 2013. – p. 193 – 195.

Розрахунок Сил Тертя у Викривлених Нафтовидобувних Свердловинах з Врахуванням Положення З'єднувальних Муфт

Ольга Євчук
кафедра інформаційно-телекомунікаційних технологій і систем
Івано-Франківський національний технічний університет нафти і газу
Івано-Франківськ, Україна
olga.yevchuk@gmail.com

Віктор Ровінський
кафедра інформатики
Прикарпатський національний університет
Івано-Франківськ, Україна
musicneutrino@gmail.com

Computation of Friction Forces in Deviated Oilwells Taking into Account Rod Coupling Positions

Olga Yevchuk
dept. of Information and Telecommunication Technologies and Systems
Ivano-Frankivsk National Technical University of Oil and Gas
Ivano-Frankivsk, Ukraine
olga.yevchuk@gmail.com

Victor Rovinsky
dept. of Computer Science
Precarpathian National University
Ivano-Frankovsk, Ukraine
musicneutrino@gmail.com

***Анотація*—Запропоновано спосіб врахування положення з'єднувальних муфт при визначенні контактних зон між штанговою колоною та колоною насосно-компресорних труб штангової глибиннонасосної установки. Вдосконалено алгоритм розрахунку сил тертя з врахуванням знайдених контактних зон.**

***Abstract*—Method of taking into account rod coupling positions has been proposed for contact zones detection between rod string and tubing of sucker-rod pumping unit. Algorithm of friction forces calculation has been improved using results of contact zone detection.**

Ключові слова—штангова глибиннонасосна установка, тертя викривлені свердловини, інтерполяція

Keywords— sucker-rod pumping unit, rod string, friction, deviated wells, interpolation

I. Вступ

Особливістю роботи нафтовидобувних свердловин, що мають викривлений профіль, є значний вплив сил тертя на роботу нафтовидобувної установки [1, 2]. Для свердловин, обладнаних штанговими глибиннонасосними установками (ШГНУ), наявність тертя призводить до збільшення навантаження на штангову колону при її русі вгору та відповідно зменшення при русі вниз. Це призводить до зміни форми динамограми (крива навантаження на полірований шток в залежності від переміщення полірованого штоку), що використовується як основна діагностична ознака при оцінці технічного стану підземної частини ШГНУ. Підвищення точності оцінки тертя покращує ефективність діагностування ШГНУ динамометричним методом, оскільки дозволяє більш адекватно сформувати еталонні динамограми для типових дефектів підземної частини ШГНУ.

Ряд робіт присвячено врахуванню в'язкого тертя між колоною штанг та рідиною, що видобувається [3] та колоною штанг і колоною насосно-компресорних труб (НКТ) [4,5,6], а також питанням визначення порогових навантажень, що призводять до спіралеподібної або синусоїдальної деформації колони під дією стискаючого навантаження [7,8,9]. Врахування положення з'єднувальних муфт, зокрема і під час руху штангової колони, не проводилося і залишається актуальною задачею.

II. Врахування положення муфт

Наявність з'єднувальних муфт ускладнює задачу оцінки сил тертя, оскільки характер зазору між штанговою колоною та колоною насосно-компресорних труб (НКТ) є нерівномірним. Для прикладу на рис. 1 показано фрагмент муфтового з'єднання штанг діаметром 25мм за допомогою муфти типу МШ25 (діаметром 56мм) всередині колони НКТ діаметром 73мм (відносні розміри всіх елементів збережено).

Рис. 1. Фрагмент муфтового з'єднання штанг Ø 25мм всередині колони НКТ Ø 73мм

Імовірність наявності контактної зони між штанговою колоною та колоною насосно-компресорних труб значно підвищується у місцях розташування муфтових з'єднань. У [6] пропонується вважати, що для викривленої свердловини всі штангові муфти мають контакт з колоною НКТ, а наявність та довжина контактної зони на ділянках між муфтами визначається із розв'язку рівняння рівноваги однієї штанги з врахуванням жорсткості штангової колони, прикладеного навантаження від ваги штанг у рідині та розтягуючої сили, що діє на колону штанг. Даний метод дає дещо завищені значення сил тертя, що можна пояснити невиконанням в загальному випадку припущення про контакт всіх муфт із колоною НКТ.

При використанні згладжуючої сплайн-інтерполяції для визначення положення точок штангової колони всередині колони НКТ за даними інклінометрії [10] допустиме відхилення задається як різниця між діаметрами штангової колони та колони НКТ. Пропонується для визначення наявності контакту у місцях муфтових з'єднань на першому етапі виконати інтерполяцію із значенням допустимого відхилення, що дорівнює різниці діаметрів з'єднувальної муфти та колони НКТ.

$$\Delta_{\max} = D_T - D_M,$$

де D_T – діаметр НКТ, D_M – діаметр муфт.

Слід зауважити, що для комбінованих колон, що складаються із ділянок труб та/або штанг з різними діаметрами, необхідно проводити окремий розрахунок для кожної комбінації діаметрів.

Умовою наявності контакту буде близькість різниці між інтерпольованими звичайним кубічним сплайном точками колони НКТ з координатами ($xs_{НКТ}$, $ys_{НКТ}$, $zs_{НКТ}$), та точками штангової колони з координатами (xs_M, ys_M, zs_M), інтерпольованими згладжуючим сплайном з допуском $\Delta_{\max}$, до величини зазору:

$$\sqrt{(xs_{НКТ} - xs_M)^2 + (ys_{НКТ} - ys_M)^2 + (zs_{НКТ} - zs_M)^2} - \Delta_{\max} < \varepsilon \qquad (1)$$

де ε – достатньо мала величина, яку можна прийняти рівною $0.01(D_T - D_M)$.

На наступному етапі для кожної знайденої імовірної контактної зони виконується перевірка наявності контакту виходячи із умови, що в даній зоні дійсно розташована з'єднувальна муфта. Для визначення положення з'єднувальних муфт необхідно знати, штанги якої довжини використовуються для компоновки штангової колони, а також мати інформацію про розташування укорочених штанг за умови їх наявності, довжину полірованого штоку $L_{ПШ}$ та відстань від гирла свердловини до точки підвісу штанг у крайньому верхньому положенні $H_П$:

$$L_M(0) = L_{ПШ} - H_П ,$$

$$L_M(i) = L_M(i\text{-}1) + l_Ш(i) + \Delta l_Ш(i),\ i=1...N,$$

де $L_M(i)$ – відстань до i-ї муфти у крайньому верхньому положенні полірованого штока, $l_Ш(i)$ – довжина i-ї штанги, N – загальна кількість штанг, $\Delta l_Ш(i)$ – величина розтягу i-ї штанги під дією навантаження від ваги W_i частини колони штанг, розташованої нижче даної штанги, та прикладеної до плунжера насоса ваги рідини всередині колони НКТ (обидві величини визначаються із врахуванням кривини по всій довжині стовбура свердловини):

$$W_i = \rho_Ш \cdot g \cdot \frac{\pi}{4} \cdot \sum_{k=0}^{ns} \Delta L_K \cdot D_{Шk}^2 \cdot \sum_{j=0}^{N_k} \cos(a_j)$$

де $\rho_Ш$ – густина матеріалу штанг, ns – кількість секцій довжиною ΔL з діаметром штанг $D_Ш$, N_k – кількість ділянок з середнім кутом нахилу α всередині k-ї секції, g – прискорення вільного падіння;

$$W_p = \rho_p \cdot g \cdot \frac{\pi}{4} \cdot \sum_{i=0}^{ns} \Delta L_i \cdot (D_{Ti}^2 - D_{Ші}^2) \cdot \sum_{j=0}^{Ni} \cos(a_j)$$

де ρ_P – густина рідини, що видобувається, ns – кількість секцій довжиною ΔL з відмінними комбінаціями діаметрів насосно-компресорних труб D_T та діаметрів штанг $D_Ш$, N_i – кількість ділянок з середнім кутом нахилу α всередині i-ї секції;

$$\Delta l_Ш(i) = \frac{4 \cdot L_i \cdot (W_i + W_p)}{\pi \cdot D_{Si}^2 \cdot E}$$

де E – модуль Юнга матеріалу штанг.

Остання формула є справедливою лише для верхньої мертвої точки, коли штангова колона навантажена вагою рідини всередині НКТ і перебуває в стані спокою. Оскільки на протязі періоду качання положення муфт змінюється, можна для кожного моменту часу визначити факт наявності відповідної контактної зони, враховуючи

положення муфти в цей момент часу. Найпростіше це зробити, вважаючи, що додатковий розтяг штангової колони під дією динамічного навантаження відсутній. Тоді положення муфт в часі описуються формулами:

$$L_M(0, t) = L_{ПШ} - H_П + S(t),$$

$$L_M(i, t) = L_M(i\text{-}1,t) + l_Ш(i) + S(t),$$

де $S(t)$ – закон руху точки підвісу штанг:

$$S(t) = \frac{S_{\max}}{2} \cdot \left[1 - \cos\left(2\pi \frac{nk}{60} t\right)\right]$$

де *Smax* – довжина ходу, *nk* – кількість качань за хвилину.

Для отримання більш точних результатів необхідно шукати розв'язок рівняння руху штангової колони [11]

$$\frac{\partial^2 y(x,t)}{\partial t^2} = v^2 \frac{\partial^2 y(x,t)}{\partial x^2} - c\frac{\partial y(x,t)}{\partial t}$$

де $y(x,t)$ – шукана функція, що описує рух штангової колони, v – коефіцієнт, що враховує модуль пружності та густину матеріалу штанг, c – коефіцієнт, що враховує коефіцієнт тертя між штанговою колоною та рідиною, що видобувається.

В результаті виконаних обчислень отримується інформація про наявність контактних зон між муфтами та колоною НКТ для довільного моменту часу на протязі циклу качання. Після цього слід визначити наявність контактних зон на ділянках між раніше знайденими зонами контакту. Цей розрахунок можна проводити за методикою, наведеною в [6], за умови що ділянка колони НКТ між даними точками не містить точок перегину.

Наявність і координати точок перегину можна визначити шляхом аналізу сплайн-коефіцієнтів, отриманих для відповідних ділянок колони НКТ [10]. Значення параметру t, для якого друга похідна дорівнює нулю для i-го інтерпольованого фрагменту (умова перегину):

$$t_{П,i} = -\frac{C_{2,i}}{6C_{3,i}}$$

де C – сплайн-коефіцієнти. За умови що 0 ≤ $_{П,i}$ ≤ , розраховуються координати точки перегину:

$$x_{П,i} = x\left(t_i + t_{П,i}\right),$$

$$y_{П,i} = y(t_i + t_{П,i}),$$

$$Z_{П,i} = z\left(t_i + t_{П,i}\right).$$

При наявності точки перегину пропонується виконати повторну інтерполяцію для даної ділянки за допомогою згладжуючого сплайну, причому допустиме відхилення, прийняти рівним різниці діаметрів штангової колони та колони НКТ.

$$\Delta_{\max} = D_T - D_Ш,$$

де D_T – діаметр НКТ, $D_Ш$ – діаметр штанг.

Після цього слід виконати повторний пошук зон контакту, використовуючи умову (1).

III. Алгоритм розрахунку сил тертя

Описаний в [10] алгоритм розрахунку сил, що діють на штангову колону з врахуванням тертя, передбачає наявність контакту по всій довжині колони, що призводить до систематично завищених значень сил тертя для умов роботи штангової глибиннонасосної установки навіть при значній викривленості стовбура свердловини. Для прикладу на рис. 3 наведено результат розрахунку сили тертя при русі штангової колони вгору по всій довжині стовбура свердловини Б-320 НГВУ «Надвірнанафтогаз», профіль якого наведено на рис. 2.

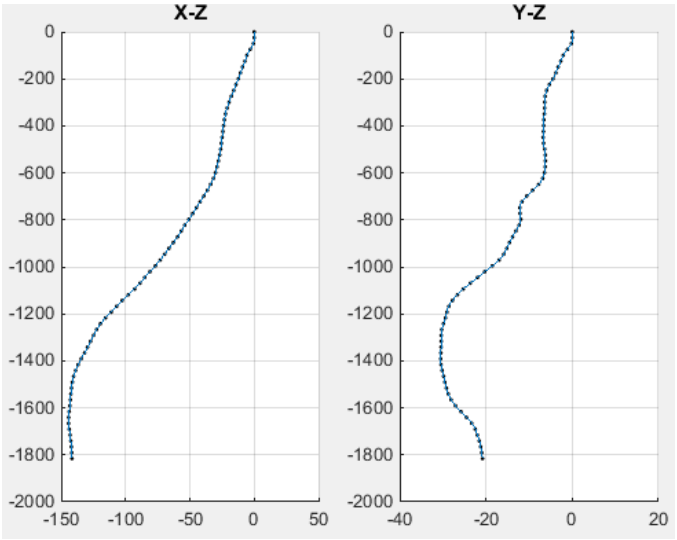


Рис. 2. Проекції профілю свердловини Б-320

Додаткове навантаження від сил тертя оцінюється величиною 30кН, що не відповідає значенням, які спостерігаються для даної свердловини.

З врахуванням того, що сили тертя присутні лише в зонах контакту, алгоритм набуває наступного вигляду.

Розрахунок починається від точки підвісу глибинного насосу. В даній точці штангу можна вважати центрованою всередині колони НКТ, а прикладене навантаження дорівнює:

при русі вниз

$$F_{0d} = W_{нас} + F_{кл} - F_{тП}$$

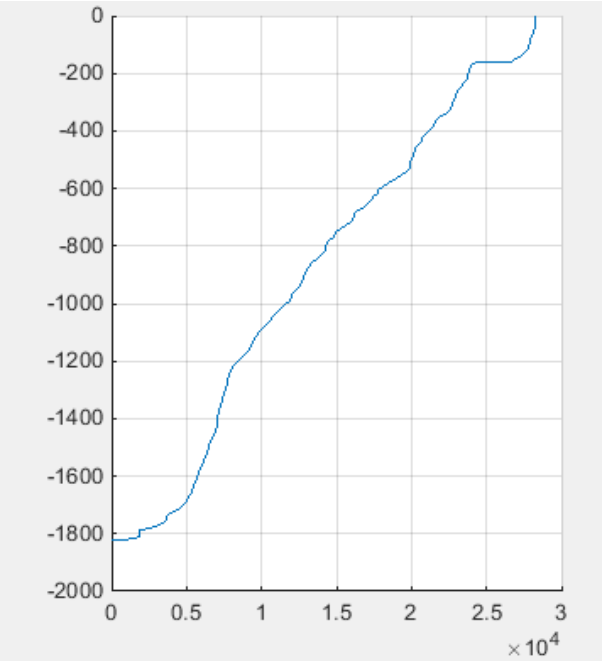


Рис. 3. Величина сили тертя при русі вгору для свердловини Б-320

де $F_{кл}$ – додаткове навантаження внаслідок гідравлічних втрат у клапанах, $F_{тП}$ – сила тертя між плунжером та циліндром;

при русі вгору

$$F_{0u} = W_{нас} + F_{кл} + F_{тП} + W_P,$$

де W_P дорівнює прикладеній до плунжера насоса вазі рідини всередині колони НКТ, яка визначається із врахуванням кривини по всій довжині стовбура свердловини.

При значній викривленості свердловини можна знехтувати складовими, що описують вагу плунжера глибинного насосу, гідравлічні втрати у клапанах та тертя у парі плунжер-циліндр. Тоді для руху вниз навантаження в точці підвісу насосу приймається нульовим, а для руху вгору – рівним W_P.

Для кожного інтервалу колони ΔL розраховується значення сили на верхньому кінці F_2 в залежності від сили, прикладеної на нижньому кінці F_1:

для прямолінійної ділянки

$$F_2 = F_1 + w \cdot \Delta L \cdot (\cos(\varphi) \pm \mu \cdot \sin(\varphi)),$$

де w – вага одиниці довжини колони в рідині, φ – кут нахилу;

для криволінійної ділянки

$$F_2 = F_1 \cdot e^{\pm\mu|\theta_1-\theta_2|} + w \cdot \Delta L \cdot \frac{\sin(\varphi_2)-\sin(\varphi_1)}{\varphi_2-\varphi_1}$$

де μ - коефіцієнт тертя, знак + береться для ходу вгору, мінус – для ходу вниз.

Для ділянок, де контактні зони відсутні, коефіцієнт тертя приймається рівним 0. Це приводить до формули

$$F_2 = F_1 + w \cdot \Delta L \cdot (\cos(\varphi))$$

для прямолінійної ділянки, та

$$F_2 = F_1 + w \cdot \Delta L \cdot \frac{\sin(\varphi_2)-\sin(\varphi_1)}{\varphi_2-\varphi_1}$$

для криволінійної ділянки.

Визначення складової, зумовленої тертям, виконується шляхом віднімання сили, зумовленої навантаженням від ваги колони штанг та ваги рідини.

IV. Висновки

Запропоновано алгоритм врахування положення з'єднувальних муфт при розрахунку сил тертя між штанговою колоною та колоною насосно-компресорних труб при русі штангової колони вгору та вниз, на сонові чого вдосконалено алгоритм розрахунку сил, що діють на штангову колону для викривлених нафтовидобувних свердловин, обладнаних ШГНУ.

Література References

[1] И.Т. Мищенко. Скважинная добыча нефти: Учебное пособие для вузов. — М.: Изд-во «Нефть и газ », 2003. - 816 с.

[2] Gabor Takacs, “Sucker-Rod Pumping Handbook: Production Engineering Fundamentals and Long-Stroke Rod Pumping”, Elsevier, 2015. – 598pp.

[3] Z. Ma, S. L. Peng, Z. Z. Qu, J. Li, "The Detailed Calculation Model of the Friction between Sucker Rod and the Liquid in the Sucker Rod Pump Lifting System of Heavy Oil", Applied Mechanics and Materials, Vol. 694, pp. 346-349, 2014

[4] К.Р. Уразаков, Б.М. Латыпов, Р.Р. Исмагилов. Экспериментальные исследования коэффициента трения элементов штанговой колонны винтовых насосных установок// Нефтегазовое дело: электрон. науч. журн. 2015. №3. С.256–270.

[5] А.Г. Газаров., С.А. Рекин, У.М. Абуталипов. Модифицированный параметр Зоммерфельда для исследования трения штанг о трубы: Сб. науч.тр. ДООО БашНИПИнефть. – Уфа, 2002.- Вып. 106.- С. 59-61.

[6] Насосная добыча высоковязкой нефти из наклонных и обводненных скважин / К. Р. Богомольный, Е.И. Уразаков. - М. : Недра-Бизнесцентр, 2003.

[7] F.J. Schuh, 1991. The Critical Buckling Force and Stresses for Pipe in Inclined Curved Boreholes, Paper SPE/IADC 21942 Presented at the SPE/IADC Drilling Conference, 11 -14 March, Amsterdam

[8] J. Wu and H.C. Juvkam-Wold. “The Effect of Wellbore Curvature on Tubular Buckling and Lockup”, J. Energy Resour. Technol 117(3), 214-218 (Sep 01, 1995)

[9] X. He, and A. Kyllingstad, “Helical Buckling and Lock-up Conditions for Coiled Tubing in Curved Wells”, SPE Paper 25370, 1995.

[10] О.В. Євчук, «Розрахунок сил тертя у викривлених свердловинах на основі параметричної сплайн-інтерполяції точок штангової колони», *Методи та прилади контролю якості*, Івано-Франківськ: ІФНТУНГ, 2016. - № 1. - С. 107-113

[11] S. Gibbs, “Rod pumping: Modern Methods of Design, Diagnosis, and Surveillance”, USA: BookMasters, Inc.; 2012. - 682pp.

Моделі та Методи Оптимізації Функціонування Розподільчих Систем Енергетичного Комплексу

Олена Кірік
Інститут прикладного системного аналізу
Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»
Київ, Україна
okirik@ukr.net

Models and Methods for the Functioning Optimization of Distribution Systems of the Energy Complex

Olena Kirik
Institute for Applied Systems Analysis
National Technical University of Ukraine " Igor Sikorsky Kyiv Polytechnic Institute"
Kyiv, Ukraine
okirik@ukr.net

***Анотація*—Побудовано серію моделей задач розподілу ресурсів, що враховують замовлення споживачів, змінну продуктивність джерел постачання та наявність тимчасових сховищ продукту. Запропоновано алгоритми нелінійного програмування для визначення оптимальних потоків та розподілу ресурсів в енергетичних системах.**

***Abstract*—Series models were built for the problems of resource distribution, taking into account consumer demand, productivity variable sources and availability of temporary storage of the product. Programming algorithms for finding optimal flows in power systems were proposed.**

Ключові слова—задачі розподілу потоків, мережеві моделі, методи нелінійного програмування, алгоритми оптимізації

Keywords—flow distribution problems, network models, nonlinear programming methods, optimization algorithms

I. Вступ

Математичне моделювання та розробка методів і алгоритмів оптимізації є важливою складовою частиною системних досліджень великих комплексів, до яких належать системи розподілу енергетичних ресурсів [1-3].

Загальні принципи побудови моделей розподілу ресурсів у розподільчих системах опишемо на прикладі газотранспортних мереж. Перспективним напрямком робіт з підвищення ефективності функціонування газотранспортної системи є розробка математичних моделей і чисельних методів, що дозволяють вирішувати сукупність взаємопов'язаних оптимізаційних завдань планування та оперативного управління. Весь комплекс видобутку, розподілу та транспортування газу треба розглядати як єдину систему і розв'язувати задачу управління газотранспортними мережами з точки зору системного підходу. Надійність транзитних поставок газу і безпека газопостачання внутрішніх споживачів в значній мірі забезпечується комплексом підземних сховищ газу.

Критеріями оптимізації функціонування всієї системи або окремих регіонів управління можуть бути досягнення максимальної продуктивності при видобутку обсягу газу не менше певного заздалегідь заданого рівня, забезпечення мінімальних сумарних експлуатаційних витрат в процесі задоволення споживачів з урахуванням заповнення або спорожнення сховищ, збереження при видобутку, відборі або закачуванні газу в сховища максимальної продуктивності.

Проблемам розвитку і функціонування об'єктів газопостачання від розгляду окремих питань до їх комплексного, взаємопов'язаного дослідження присвячена значна кількість публікацій.

Для постановки та вирішення галузевих завдань паливно-енергетичного комплексу пропонувалися відповідні моделі та методи [4,5]. Однак вони не завжди враховують належною мірою мережевого характеру систем газопостачання, вимог узгодженості у розвитку основних підсистем видобутку, трубопровідного транспорту та підземного зберігання газу.

Розглянемо нелінійну модель для розподільчих систем, що враховує видобуток, прокачування, зберігання та споживання енергетичних ресурсів. Розв‘язання задач

оптимального планування дозволяє погоджувати обсяги газу, що надходить з газових промислів з можливостями магістральних газопроводів і резервуарів тимчасового зберігання.

Метою даної роботи є побудова на основі дослідженої базової моделі розподілу потоків ряду моделей шляхом модифікації цільової функції та системи обмежень, а також створення ефективних алгоритмів нелінійного програмування для задач розрахунку мереж.

II. Побудова моделей

Розглядається система газопостачання, основними елементами якої є газові промисли, на яких відбувається видобуток газу; нитки газопроводів, що з'єднують газові промисли з вузлами, де розташовані споживачі; а також підземні сховища, що можуть використовуватися як резервуари тимчасового зберігання газу. Системи газопостачання повинні оптимальним чином задовольняти попит на газ при певних технологічних параметрах системи газопроводів. Ставиться така задача планування: при заданих обмеженнях на обсяги видобутку газу, обмеженнях на пропускні спроможності ниток газопроводу визначити оптимальні обсяги транспортування газу від газових промислів до споживачів або у тимчасові сховища при мінімальних витратах на транспортування.

Топологічно розподільча мережа представляє собою зв‘язний граф $G=(N,V)$, де N і V - пара скінчених множин. Елементи i множини N називаються вершинами графа G. Вони відображають розподіл споживачів газу, а також джерел та стоків ресурсів. Елементи v множини V називаються дугами (ділянками), причому кожному $v \in V$ співставлена впорядкована пара вершин (i,j), $i,j \in N$, що є, відповідно, початком і кінцем дуги $v=(i,j)$.

Нехай x_{ij} - це величина потоку речовини від вершини i в вершину j вдовж дуги $(i,j) \in V$. Величина d_i означає кількість речовини, що споживається ($d_i < 0$) або подається у мережу ($d_i \geq 0$) в вершині i Числа r_{ij}^+ і r_{ij}^- визначають верхнє та нижнє технологічне обмеження на потік вздовж дуги (i,j) і дозволяють фіксувати певний напрям потоків.

Якщо функція $f_{ij}(x_{ij})$ відображає умовну вартість проходження потоку вздовж дуги (i,j), тоді задача знаходження найкращого розподілу потоків вздовж мережі полягає у мінімізації функції

$$F = \sum_{(i,j)\in V} \int_0^{x_{ij}} f_{ij}(t)dt \qquad (1)$$

за обмежень

$$\sum_{j:(i,j)\in V} x_{ij} - \sum_{j:(j,i)\in V} x_{ji} = d_i, i \in N, \qquad (2)$$

$$r_{ij}^- \leq x_{ij} \leq r_{ij}^+, (i,j) \in V \qquad (3)$$

Важливою є умова збалансованості системи (2), тобто відповідність обсягів речовини, що подається в мережу і споживаного газу:

$$\sum_{i\in N} d_i = 0. \qquad (4)$$

Будемо вважати, що виділено деяку підмножину вершин $\overline{N} \subseteq N$, де розташовані джерела та стоки, в яких величина d_i не є фіксованою. Це означає, що вузлові параметри $d_i \in \overline{N}$ представляються як невідомі, тобто оптимізація відбувається не тільки за рахунок мінімізації вартості доставки газу споживачам, але і за рахунок ефективного перерозподілу навантаження джерел і стоків.

Припускаємо також, що для кожного ребра (i,j) функція $f_{ij}(x_{ij})$ є неперервною, строго монотонно зростаючою і існує таке $\overline{x}_{ij}$, що $f_{ij}(\overline{x}_{ij}) = 0$. Такі цілком природні припущення гарантують існування мінімуму сформульованої задачі.

Нехай $u_i, i \in N$ та u_0 - двоїсті (вартісні) оцінки, відповідно, обмежень (2) та (4). Умови оптимальності [6] формулюються таким чином. Якщо в задачі розподілу потоків параметри подачі ресурсу в мережу у вузлах $i \in \overline{N}$ не фіксовані і не обмежені, то для того, аби потоки x_{ij}, $(i,j) \in V$ були оптимальними, необхідно і достатньо, щоб знайшлися такі u_i, $i \in N \setminus \overline{N}$, для яких виконуються такі умови: якщо $x_{ij} = r_{ij}^-$, то $f_{ij}(x_{ij}) \geq u_j - u_i$; якщо $x_{ij} = r_{ij}^+$, то $f_{ij}(x_{ij}) \leq u_j - u_i$; якщо $r_{ij}^- < x_{ij} < r_{ij}^+$, то $f_{ij}(x_{ij}) = u_j - u_i$. Крім того, $u_i = u_0$ для $i \in \overline{N}$, причому величина u_0 може бути вибрана довільно. Це означає, що якщо на вузлові змінні не накладені обмеження, то джерела (стоки) рівноцінні і відповідні їм змінні двоїстої задачі збігаються.

Доведення отримується шляхом застосування теореми Куна-Таккера [6] до задачі (1)-(4), з урахуванням того, що умови, накладені на функції $f_{ij}(x_{ij})$, гарантують опуклість дугових складових функції Лагранжа

$$L = \sum_{(i,j)\in V} \left(\int_0^{x_{ij}} f_{ij}(t)dt - x_{ij}(u_j - u_i) \right) + \sum_{i\in N} (u_0 - u_i)d_i$$

.

для довільних значень двоїстих змінних.

Розв‘язувати задачу (1)-(4) зручно шляхом переходу до двоїстої задачі безумовної максимізації нелінійної безперервно диференційовної адитивно-сепарабельної функції Ф

$$max\ \Phi\ .$$

Можливість конкретизувати вид цієї неявно заданої цільової функції

$$\Phi = \sum_{(i,j)\in V} \min_{r_{ij}^- \le x \le r_{ij}^+} \left[\int_0^x f_{ij}(t)dt - x(u_j - u_i) \right] + \sum_{i\in N\setminus\overline{N}} (u_0 - u_i)d_i$$

та визначити її похідні за формулами [6]

$$\frac{d\Phi}{du_i} = \sum_{j:(i,j)\in V} x_{ij}(u_j - u_i) - \sum_{j:(j,i)\in V} x_{ji}(u_i - u_j) - d_i$$

дає простір для використання різних методів нелінійного програмування.

Запропонований підхід дозволяє не тільки перейти від задачі з великою кількістю обмежень до задачі без обмежень, але і підвищити ефективність процесу знаходження розв‘язку, врахувавши збіг двоїстих змінних у відповідних точках і виключивши при реалізації алгоритмів непотрібні обчислення. Знаючи оптимальні значення двоїстих змінних u_i, $i \in N\setminus\overline{N}$, і застосовуючи наведені вище умови оптимальності, можна визначити оптимальне значення потоків x_{ij}, $(i,j)\in V$, для всіх ділянок мережі.

Модель (1)-(4), шляхом модифікації цільової функції та системи обмежень, породжує набір моделей для розв‘язання ряду важливих задач функціонування мережевих систем.

A. Розподіл потоків з усталеними поставками продукту

Нехай $\overline{N} = \varnothing$. Це означає, що відомі замовлення споживачів та фіксована продуктивність джерел постачання продукту. У цьому випадку модель (1)-(4) дозволяє розрахувати усталений режим розподілу потоків з мінімальними затратами на доставку продукту.

B. Розподіл потоків з оптимальним перерозподілом постачальників продукту

Це модель (1)-(4), для $\overline{N} \neq \varnothing$. У вузлах графа задаються функції споживання (замовлення споживачів та план постачання продукту), причому подача продукту з різних джерел може бути або константами (детерміновані джерела) або змінними (недетерміновані джерела). Потрібно визначити такий план розподілу потоків вздовж мережі, що за рахунок можливого перерозподілу навантаження недетермінованих джерел задовольнить споживачів з найменшими загальними витратами на перевезення.

C. Розподіл потоків із врахуванням пріоритету джерел постачання продукту

Припустимо, що потрібно прагнути до певного рівня продуктивності газових промислів в процесі задоволення споживачів, який встановлено на рівні $\overline{d}_i$ для кожного з вузлів $i \in \overline{N}$. Врахування цих вимог можна здійснити введенням штрафу за великі ухилення від бажаних значень в цільову функцію у вигляді квадратичного доданку $\frac{1}{2}\sum_{i\in\overline{N}} (d_i - \overline{d}_i)^2$. Крім того, необхідно врахувати існування певних технологічних обмежень на продуктивність джерел $d_i^- \le d_i \le d_i^+$, $i \in \overline{N}$.

D. Оптимізація функціонування мережі в процесі наповнення і випорожнення сховищ газу

Підземні сховища газу виконують такі основні функції [7]: регулюють сезонні нерівномірності газопостачання; створюють резервні запаси газу; підвищують надійність роботи газотранспортних систем і газопостачання; гарантують рівномірність експортних поставок і транзит газу.

Нехай в вершинах $i \in \overline{N}$ мережі G розташовані сховища, причому y_i - питома кількість продукту, який заповнює i-е сховище, s_i^- и s_i^+ - мінімальний та максимальний об‘єми заповнення i-го сховища, $h_i(y_i)$ - експлуатаційні витрати цього сховища. Експлуатаційні витрати комплексу сховищ представляються в вигляді функції

$$H = \sum_{i\in N_0} \int_0^{y_i} h_i(\tau)d\tau$$

,

а загальні витрати на транспортування продукції з урахуванням його відбору (закачування) в сховища можна представити у вигляді суми функцій F і H.

Модель оптимізації функціонування мережі при мінімізації експлуатаційних витрат в процесі заповнення або спорожнення сховищ може бути виписана таким чином.

$$min(F + H)$$

$$\sum_{j:(i,j)\in V} x_{ij} - \sum_{j:(j,i)\in V} x_{ji} = d_i, i \in N\setminus\overline{N}$$

,

$$\sum_{j:(i,j)\in V} x_{ij} - \sum_{j:(j,i)\in V} x_{ji} = y_i, i \in \overline{N}$$

$$r_{ij}^- \le x_{ij} \le r_{ij}^+, (i,j)\in V\ ,\ s_i^- \le y_i \le s_i^+\ ,\ i \in \overline{N}\ .$$

Сумарний попит споживачів повинен знаходитися в балансі з подачею продукту в мережу, зокрема, з сумарною продуктивністю сховищ.

$$\sum_{i\in N\setminus\overline{N}} d_i + \sum_{i\in\overline{N}} y_i = 0$$

E. Збільшення інтенсивності роботи газодобувних підприємств

Актуальним в сфері розвитку газової промисловості залишається питання збільшення продуктивності газодобувних підприємств з метою зменшення витрат і збільшення віддачі від вкладених в цю галузь коштів [8].

Критеріями оптимізації функціонування газодобувного підприємства можуть бути: досягнення максимального прибутку підприємства при видобутку природного газу, забезпечення мінімальних експлуатаційних витрат в процесі відбору газу з родовища, збереження при відборі газу максимального потенціалу продуктивності підприємства.

Якщо припустити, що підприємство функціонує в умовах повної визначеності щодо промислових запасів, то в залежності від чотирьох перерахованих вище передумов можна побудувати чотири відповідні математичні моделі [9].

Розглянуті вище моделі дозволяють розв‘язувати комплекс взаємопов'язаних задач оптимізації експлуатації газотранспортної системи. Розв‘язавши задачі визначення оптимальних темпів видобутку для газових родовищ та заповнення сховищ і підставивши отримані значення в модель (1)-(4) можна розрахувати оптимальні транспортні потоки газу з урахуванням введення нових родовищ, використання підземного зберігання та у разі аварійних ситуацій.

III. Розв‘язання мережевих задач із застосуванням методів нелінійного програмування та апроксимаційних схем

Заміна складної вихідної задачі послідовністю більш простих задач або задач, які легко можуть бути розв‘язані із застосуванням спеціальних прийомів, є найбільш ефективним в обчислювальному відношенні підходом у теорії та методах оптимізації. Найвідомішими у нелінійному програмуванні є різноманітні методи, що застосовують лінеаризацію або квадратичну апроксимацію нелінійних функцій.

При застосуванні апроксимаційних методів відбувається заміна нелінійної оптимізаційної задачі послідовністю квадратичних задач, цільові функції яких є квадратичною апроксимацією цільової функції, або її лінійною апроксимацією з додаванням квадратичного штрафу за великі відхилення аргументу.

Перепишемо модель (1)-(3) у матричному вигляді. Нехай граф $G=(N,V)$ містить n вершин (множина N) та m дуг (множина V). Для кожної вершини i відоме споживання d_i, а кожній дузі графа v приписаний потік x_v, обмежений знизу та зверху величинами r_v^- і r_v^+, та функція вартості протікання потоку $F_v(x_v)$ вздовж дуги.

Задача знаходження розподілу потоків мінімальної вартості може бути сформульована таким чином

$$F(x)\equiv\sum_{v=1}^{m} F_v x_v \to min$$

$$Ax = d,$$

$$^{-}r \le x \le r^{+},$$

де $x=(x_1,x_2,...,x_m)^T$ — вектор невідомих дугових потоків, $d=(d_1,d_2,...,d_n)^T$ — вектор об‘ємів споживання у вузлах, $A(n\times m)$ — матриця інциденцій вузли-дуги. Кожен стовпець v цієї матриці відповідає дузі $v=(i,j)$, $i,j\in N$ та містить тільки два ненульових елементи: +1 у рядку i та –1 у рядку j.

Надалі вважатимемо, що функція $F(\cdot)$ є опуклою двічі неперервно диференційовною і $\left\langle v, F_v''(x_v)v\right\rangle > 0$ для всіх $v\in V, v\neq 0$. За умови існування розв’язку, остання умова гарантує його єдиність [6].

Загальна схема аппроксимуючих методів умовної оптимізації може бути представлена у вигляді ітераційного процесу

$$x^{k+1} = x^k + \alpha_k p^k,\ 0<\alpha_k\le 1 \qquad (5)$$

де $\alpha_k > 0$, а $p^k = p(x^k)$, $p^k\in R^m$, - розв‘язок допоміжної апроксимуючої задачі з лінійними обмеженнями:

$$F(x^k)+\left\langle F'(x^k), p\right\rangle+\frac{1}{2}\left\langle D_k p, p\right\rangle \to min$$

$$f_i(x^k)+\left\langle f_i'(x^k), p\right\rangle,\quad i=1,...,n,$$

де $f_i(\cdot)$ - функції обмежень вихідної задачі.

Для визначення крокового множника α_k у всіх апроксимаційних методах використовується релаксація вздовж вектора спуску p_k недиференційованої штрафної функції

$$\overline{F}(x,\Lambda)\equiv F(x)+\Lambda H(x),\quad \Lambda>0,$$

де $H(x)=max\{0,f_1(x),\ldots,f_n(x)\}$.

В задачі розподілу потоків всі функції обмежень є лінійними функціями. Якщо x^0 — початкове наближення до розв’язку нелінійної задачі (1)-(3), то для реалізації ітераційного процесу (5) у кожній точці x^k знаходимо

напрям зсуву $p^k = p(x^k) \in R^m$ як розв‘язок при фіксованому $x = x^k \in R^m$ допоміжної квадратичної задачі:

$$\frac{1}{2}\langle p, Dp\rangle + \langle c, p\rangle \to min \quad (6)$$

$$Ap = \tilde{d}\,, \quad (7)$$

$$\tilde{r}^- \le p \le \tilde{r}^+ . \quad (8)$$

Тут $c = F'(x)$ - градієнт функції $F(x)$, D - матриця її других похідних у методах послідовного квадратичного програмування або одинична матриця у випадку застосування методу лінеаризації Б.М.Пшеничного [10], $\tilde{d} = d - Ax$, $\tilde{r}^- = r^- - x$, $\tilde{r}^+ = r^+ - x$.

Припускаємо, що розв’язок допоміжних задач (6)–(8) існує для всіх точок процесу (5).

Критерієм завершення процесу (5) є виконання умови $p(x^k) = 0$.

Якщо обрати $D = F''(x)$ і таку формулу для обчислення крокового множника α_k

$$F(x^k + \alpha p^k) \le F(x^k) - \varepsilon\alpha\langle p^k, F''(x^k)p^k\rangle\,,\ 0 < \varepsilon < 1\,, \quad (9)$$

то виконується така теорема

Теорема [10]. Нехай цільова функція задачі (1)-(3) $F(x)$ є опуклою двічі неперервно-диференційовною і

$$\langle p, F''(x)p\rangle \ge \omega\|p\|^2,\ \omega > 0\,.$$

Тоді ітераційний процес алгоритму опуклого програмування (5), в якому допоміжною є задача (6)–(8), а для знаходження кроку використовується формула (9), збігається швидше довільної геометричної прогресії з довільного початкового наближення, що задовольняє обмеженням задачі.

Для розв’язання задачі (6)–(8) застосовувався метод двоїстого активного набору, який знаходить розв’язок задачі за скінчене число кроків [10]. Загальна ідея метода полягає у впорядкованому переборі граней допустимої множини та мінімізація квадратичної функції на цих гранях.

Як розрахунковий приклад апроксимуючого методу використовувався метод лінеаризації Б.М. Пшеничного, що є методом першого порядку. Ефективність цього методу була підвищена із застосуванням деяких спеціальних прийомів.

Для деяких класів задач доцільно використовувати апроксимуючі методи більш високого порядку, беручи у якості апроксимуючої матриці матрицю других похідних, або деякі наближення до цієї матриці, або функцію Лагранжа, як це робиться у методах послідовного квадратичного програмування.

Розрахунки газових мереж здійснювалися із застосуванням пакету Solver, Frontline Systems Inc. та спеціально створеної програми по методу лінеаризації[11]. Вони продемонстрували ефективність запропонованого підходу до розв‘язання мережевих задач.

IV. Висновки

Основними інструментами моделювання і практичних розрахунків на функціональному та техніко-економічному рівнях є методи математичного програмування і мережеві потокові моделі, реалізовані в сучасних системах управління.

Розглянуті моделі (модель розподілу потоків з усталеними поставками продукту, модель розподілу потоків з оптимальним перерозподілом постачальників продукту, модель розподілу потоків із врахуванням пріоритету джерел постачання продукту, модель оптимізації функціонування мережі в процесі наповнення і випорожнення сховищ газу) призначені для розв‘язання комплексу взаємопов'язаних задач оптимізації експлуатації газотранспортної системи. Вони дозволяють провести попередню оцінку ефективності системи при проектуванні мереж, розрахувати попередній план подачі газу в мережу при сезонній зміні потреб користувачів, гарантувати задоволення споживачів при аварійних відмовах деяких вузлів.

Література References

[1] М.З.Згуровский, Н.Д.Панкратова, Системный анализ. Проблемы, методология, приложения. 2-е изд. - К. : Наук. думка, 2011, 727 с.

[2] М.Н.Кулик, Методы системного анализа в энергетических исследованиях - К. : Наук. думка, 1987, 198 с.

[3] І.Ч.Лещенко, “Перспективи газотранспортної системи України в контексті змін у світовому газовому секторі”, *Проблеми загальної енергетики*, 2011, вип. 2 (25) . - С. 17-24

[4] М.Х.Прилуцкий, В.Е.Костюков, “Оптимизационные задачи планирования транспортировки газа”, *Информационные технологии и вычислительные системы* 2007, №2 .- С.67-73

[5] М.Притула, “Оптимізаційні задачі модернізації газотранспортної системи”. *Вісник Національного Університету “Львівська політехніка”. Серія: комп’ютерні науки та інформаційні технології*, 2016 , № 843.- С..308-316

[6] Б.Н.Пшеничный, Выпуклый анализ и экстремальные задачи. – М.: Наука, 1980. - 320с.

[7] Б.П.Савків, “Задачі оптимізації функціонування комплексів підземного зберігання газу в Україні”, *Нафтова і газова промисловість,* 1997, № 6. – С. 36 – 39.

[8] І.М. Карп, Д.О. Єгер, Ю.О. Зарубін, Л.М. Уніговський та ін., Стан і перспективи розвитку нафтогазового комплексу України. – Київ:. Наукова думка, 2006. – 310 с.

[9] Е.Е.Кирик, А.П. Яковлева, “Комплексные оптимизационные модели и задачи добычи, распределения и хранения газа”, *Кибернетика и системный анализ,* Киев, 2014, N 3. – С. 137-144.

[10] Б.Н. Пшеничный, Метод линеаризации. – М.: Наука, 1983. - 136 с.

[11] В.М.Александрова, О.Є Кірік. “Оптимізаційні моделі й алгоритми для мережевих задач розподілу ресурсів”, *Наукові вісті НТУУ “КПІ”.* - Київ, 2014, N 4. .С – 39-45.

Вплив Локального Зміщення Маси на Приповерхневі Явища в Кулі

Ольга Грицина
відділ математичних методів
обчислювального експерименту,
Центр математичного моделювання
Інституту прикладних проблем механіки і математики
ім. Я. С. Підстригача НАН України
Львів, Україна
gryt@cmm.lviv.ua, gryt045@gmail.com

Галина Мороз
відділ математичного моделювання
нерівноважних процесів,
Центр математичного моделювання
Інституту прикладних проблем механіки і математики
ім. Я. С. Підстригача НАН України
Львів, Україна
halynamoroz.ua@gmail.com

The Effect of Local Mass Displacement on Near-Surface Phenomena in Sphere

Olha Hrytsyna
department for mathematical methods
of computing experiment,
Center of Mathematical Modeling of Pidstryhach Institute
for Applied Problems of Mechanics and Mathematics,
National Academy of Sciences of Ukraine
Lviv, Ukraine
gryt@cmm.lviv.ua, gryt045@gmail.com

Halyna Moroz
department for mathematical modelling
of irreversible processes,
Center of Mathematical Modeling of Pidstryhach Institute
for Applied Problems of Mechanics and Mathematics,
National Academy of Sciences of Ukraine
Lviv, Ukraine
halynamoroz.ua@gmail.com

***Анотація*—Співвідношення градієнтного типу теорії деформації твердих тіл, що враховують взаємозв'язок процесу деформування з локальним зміщенням маси, застосовані для вивчення впливу останнього на приповерхневу неоднорідність напружено-деформованого стану в ізотропній суцільній кулі. Досліджено вплив кривини поверхні на приповерхневу неоднорідність напружень, поверхневий натяг та поверхневу енергію деформації.**

***Abstract*—The equations for the gradient-type theory of the deformation of solids, which take into account the interrelation between the deformation process and local mass displacement, were used to study the effect of local mass displacement on the near-surface inhomogeneity of a stress-strain state in an isotropic solid sphere. The effect of surface curvature on the near-surface inhomogeneous stresses, surface tension, and surface energy of deformation is studied.**

Ключові слова—локально градієнтна теорія; локальне зміщення маси; нелокальні визначальні співвідношення; приповерхневі ефекти, поверхнева енергія деформації.

Keywords—local gradient theory; local mass displacement; nonlocal constitutive equations; near-surface phenomena, surface enerdy of deformation.

I. Вступ

Дослідження напружено-деформованого стану тонких волокон, малих частинок, а також синтез нових нанопористих та нанокомпозитних матеріалів вимагає коректного врахування у модельному описі приповерхневих та розмірних ефектів. У межах одноконтинуумного підходу згадані ефекти коректно описують нелокальні теорії пружності [1-3]. Однією із таких є локально градієнтна теорія термопружних твердих тіл, яка ґрунтується на врахуванні взаємозв'язку процесів деформування й теплопровідності із локальним зміщенням маси, з яким пов'язують можливість зміни структури матеріалу у межах фізично малого елемента тіла [4, 5]. Такі зміни структури матеріалу характерні примежовим областям твердих тіл [6]. Замкнену систему співвідношень локально градієнтної термомеханіки одержано у працях [4, 5]. Метою пропонованого дослідження є вивчення впливу локального зміщення маси на приповерхневу неоднорідність зв'язаних полів у суцільній твердій ізотропній кулі, а також використання

одержаних результатів для аналізу впливу кривини поверхні на приповерхневу неоднорідність напружень, поверхневий натяг і поверхневу енергію деформації малорозмірних твердих тіл сферичної геометрії.

II. Система рівнянь моделі

Вивчимо вплив вільної поверхні на напружено-деформований стан вільної від зовнішнього навантаження суцільної кулі. З цією метою розглянемо рівноважний стан ізотропної кулі, яка у сферичній системі координат (r, φ, θ) займає область $r \le R$. Прийнявши ізотермічне наближення, обмежимось розглядом процесів деформування й локального зміщення маси.

Для ізотропних матеріалів лінеаризована система рівнянь нелокальної теорії деформації пружних тіл охоплює рівняння балансу механічного імпульсу й наведеної маси [5]

$$\nabla \cdot \bar{\sigma} + \rho_0 \mathbf{F} = \rho_0 \frac{\partial^2 \mathbf{u}}{\partial t^2}, \tag{1}$$

$$\frac{\partial \rho_{m\pi}}{\partial t} + \nabla \cdot \mathbf{J}_{ms} = 0, \tag{2}$$

лінійні рівняння стану

$$\bar{\sigma} = 2G\hat{\boldsymbol{e}} + \left[\left(K - \frac{2}{3}G\right)e - K\alpha_\rho \rho_m\right]\hat{\mathbf{I}}, \tag{3}$$

$$\mu'_\pi = \mu_{\pi 0} + d_\rho \rho_m - \frac{1}{\rho_0} K\alpha_\rho e, \tag{4}$$

$$\pi_m = -\chi_m \nabla \mu'_\pi, \tag{5}$$

а також геометричні співвідношення, які пов'язують тензор деформації $\hat{\boldsymbol{e}}$ і вектор переміщення $\mathbf{u}$

$$\hat{\boldsymbol{e}} = \frac{1}{2}\left[\nabla \mathbf{u} + (\nabla \mathbf{u})^{\mathrm{T}}\right]. \tag{6}$$

Тут $\bar{\sigma}$ — тензор напружень; $e = \hat{\boldsymbol{e}} : \hat{\mathbf{I}}$; $\hat{\mathbf{I}}$ — одиничний тензор; $\mathbf{J}_{ms} = \partial \boldsymbol{\Pi}_m / \partial t$; $\boldsymbol{\Pi}_m$ — вектор локального зміщення маси [4, 5]; $\pi_m = \boldsymbol{\Pi}_m / \rho_0$; $\rho_m = \rho_{m\pi}/\rho_0$, де $\rho_{m\pi} = -\nabla \cdot \boldsymbol{\Pi}_m$ — густина наведеної маси [5]; $\mu'_\pi = \mu_\pi - \mu$; μ_π — міра зміни внутрішньої енергії системи, зумовленої локальним зміщенням маси [5]; μ — хімічний потенціал; $\mu_{\pi 0}$ і ρ_0 — значення потенціалу μ'_π та густини маси у природному стані безмежного однорідного середовища; K — модуль об'ємного стиску за сталої густини наведеної маси; G — модуль зсуву; α_ρ — коефіцієнт об'ємного розширення, спричиненого локальним зміщенням маси; χ_m — коефіцієнт, який характеризує локальне зміщення маси, зумовлене градієнтом потенціалу μ'_π; d_ρ — ізохоричний коефіцієнт залежності потенціалу μ'_π від питомої густини наведеної маси; t — час; ∇ — оператор Гамільтона.

Розв'язувальну систему рівнянь отримаємо, підставивши у балансові рівняння (1), (2) визначальні та геометричні співвідношення (3)-(6). За ключові функції вибираємо вектор переміщень $\mathbf{u}$ і потенціал $\tilde{\mu}'_\pi$. Тоді для стаціонарних задач локально градієнтної механіки розв'язувальна система рівнянь є такою

$$\left(\bar{K} + \frac{1}{3}G\right)\nabla(\nabla \cdot \mathbf{u}) + G\Delta \mathbf{u} - K\frac{\alpha_\rho}{d_\rho}\nabla \tilde{\mu}'_\pi + \rho_0 \mathbf{F} = 0, \tag{7}$$

$$\Delta \tilde{\mu}'_\pi - \lambda_\mu^2 \tilde{\mu}'_\pi = \lambda_\mu^2 \frac{K\alpha_\rho}{\rho_0} \nabla \cdot \mathbf{u}. \tag{8}$$

У співвідношеннях (7), (8) введено позначення

$$\bar{K} = K - \frac{K^2 \alpha_\rho^2}{\rho_0 d_\rho}, \quad \lambda_\mu^2 = \frac{1}{d_\rho \chi_m}.$$

Зазначимо, що параметр $l_* = \lambda_\mu^{-1}$ має розмірність довжини і є характерною віддаллю для приповерхневих явищ.

Вважаємо, що задача характеризується центральною симетрією, а тому шукані функції залежать лише від радіуса: $\mathbf{u} = (u_r(r), 0, 0)$, $\tilde{\mu}'_\pi = \tilde{\mu}'_\pi(r)$. Тоді компоненти тензора деформації визначаються співвідношеннями:

$$e_{rr} = \frac{du_r}{dr}, \quad e_{\varphi\varphi} = \frac{u_r}{r}, \quad e_{\theta\theta} = \frac{u_r}{r} \tag{9}$$

Виходячи із рівнянь стану моделі (3)-(5) та співвідношень (9), для компонент тензора напружень, збурення $\tilde{\mu}'_\pi = \mu'_\pi - \mu_{\pi 0}$ модифікованого хімічного потенціалу та вектора локального зміщення маси маємо формули

$$\sigma_{rr} = \left(\bar{K} + \frac{4}{3}G\right)\frac{du_r}{dr} + 2\left(\bar{K} - \frac{2}{3}G\right)\frac{u_r}{r} - \frac{K\alpha_\rho}{d_\rho}\tilde{\mu}'_\pi, \tag{10}$$

$$\sigma_{\varphi\varphi} = \sigma_{\theta\theta} = \left(\bar{K} - \frac{2}{3}G\right)\frac{du_r}{dr} + 2\left(\bar{K} + \frac{1}{3}G\right)\frac{u_r}{r} - \frac{K\alpha_\rho}{d_\rho}\tilde{\mu}'_\pi, \tag{11}$$

$$\tilde{\mu}'_\pi = d_\rho \rho_m - \frac{K\alpha_\rho}{\rho_0}\left(\frac{du_r}{dr} + 2\frac{u_r}{r}\right),$$

$$\pi_m(r) = -\chi_m \frac{d\tilde{\mu}'_\pi}{dr}. \tag{12}$$

За відсутності дії масових сил, ґрунтуючись на співвідношеннях (7)-(12), для визначення компонент вектора переміщення u_r та потенціалу $\tilde{\mu}'_\pi$ маємо таку систему рівнянь

$$\frac{d}{dr}\left[\frac{1}{r^2}\frac{d(r^2 u_r)}{dr} - \frac{K\alpha_\rho}{d_\rho\left(\bar{K} + 4G/3\right)}\tilde{\mu}'_\pi\right] = 0, \tag{13}$$

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2 \frac{d\tilde{\mu}'_\pi}{dr}\right) - \lambda_\mu^2 \tilde{\mu}'_\pi = \lambda_\mu^2 \frac{K\alpha_\rho}{\rho_0 r^2}\frac{d(r^2 u_r)}{dr}. \tag{14}$$

Врахувавши, що $\frac{1}{r}\frac{d}{dr}\left(r^2\frac{d\tilde{\mu}'_\pi}{dr}\right)=\frac{d^2(r\tilde{\mu}'_\pi)}{dr^2}$, запишемо рівняння (14) так

$$\frac{d^2(r\tilde{\mu}'_\pi)}{dr^2}-\lambda_\mu^2 r\tilde{\mu}'_\pi=\lambda_\mu^2\frac{K\alpha_\rho}{\rho_0 r}\frac{d(r^2u_r)}{dr}. \qquad (15)$$

Загальний розв'язок рівнянь (13) і (15) має вигляд

$$u_r=\frac{K\alpha_\rho}{d_\rho\left(\bar{K}+4G/3\right)\tilde{\lambda}r}\left[B_1\left(ch\left(\tilde{\lambda}r\right)-\frac{sh\left(\tilde{\lambda}r\right)}{\tilde{\lambda}r}\right)+\right.$$
$$\left.+B_2\left(sh\left(\tilde{\lambda}r\right)-\frac{ch\left(\tilde{\lambda}r\right)}{\tilde{\lambda}r}\right)\right]+\frac{B_3 r}{3\left(1+\mathsf{M}\right)}+\frac{B_4}{r^2}, \qquad (16)$$

$$\tilde{\mu}'_\pi=\frac{1}{r}\left[B_1\,sh\left(\tilde{\lambda}r\right)+B_2\,ch\left(\tilde{\lambda}r\right)\right]-\frac{K\alpha_\rho}{\rho_0\left(1+\mathsf{M}\right)}B_3, \qquad (17)$$

де $B_i,\ i=\overline{1,4}$, — сталі інтегрування, які визначаємо з крайових умов, а

$$\tilde{\lambda}^2=\lambda_\mu^2\left(1+\mathsf{M}\right),\quad \mathsf{M}=\frac{K^2\alpha_\rho^2}{\rho_0 d_\rho}\Big/\left(\bar{K}+\frac{4}{3}G\right).$$

Розглядаємо рівноважний стан вільної від зовнішнього навантаження кулі. Вважаємо, що куля контактує з середовищем, яке описуємо у наближенні вакууму. Відтак, крайові умови запишемо так [5]:

$$\forall\ r=R:\ \sigma_{rr}\left(R\right)=0,\ \ u_r(0)=0,\ \tilde{\mu}'_\pi\left(R\right)=-\mu_{\pi 0}. \qquad (18)$$

До цих умов слід долучити умову обмеженості потенціалу $\tilde{\mu}'_\pi$ у разі $r\to 0$.

Розв'язок (16), (17), що задовольняє крайовим умовам (18), має вигляд

$$u_r(r)=-\frac{\mu_{\pi 0}RK\alpha_\rho}{\left(\bar{K}+4G/3\right)d_\rho}\left[\frac{1}{\tilde{\lambda}r\,sh(\tilde{\lambda}R)}\left(1+\frac{\mathsf{M}}{\zeta}\right)\left(ch(\tilde{\lambda}r)-\right.\right.$$
$$\left.\left.-\frac{sh(\tilde{\lambda}r)}{\tilde{\lambda}r}\right)+\frac{r}{3\zeta R}\right], \qquad (19)$$

$$\tilde{\mu}'_\pi(r)=-\mu_{\pi 0}\left[\left(1+\frac{\mathsf{M}}{\zeta}\right)\frac{sh(\tilde{\lambda}r)}{sh(\tilde{\lambda}R)}\frac{R}{r}-\frac{\mathsf{M}}{\zeta}\right], \qquad (20)$$

де $\zeta=\dfrac{K\tilde{\lambda}^2R^2}{4G\left(\tilde{\lambda}R\ cth(\tilde{\lambda}R)-1\right)}-\mathsf{M}$.

На основі формул (19), (20) та (9)-(12) для компонент тензора напружень та вектора локального зміщення маси отримаємо формули

$$\sigma_{rr}=\frac{4\mu_{\pi 0}KG\alpha_\rho}{d_\rho\left(\bar{K}+4G/3\right)}\left[\left(1+\frac{\mathsf{M}}{\varsigma}\right)\frac{R}{\tilde{\lambda}r^2\,sh\left(\tilde{\lambda}R\right)}\times\right.$$
$$\left.\times\left(ch\left(\tilde{\lambda}r\right)-\frac{sh\left(\tilde{\lambda}r\right)}{\tilde{\lambda}r}\right)-\frac{K}{4G\varsigma}\right], \qquad (21)$$

$$\sigma_{\varphi\varphi}=\sigma_{\theta\theta}=-\frac{\mu_{\pi 0}RK\alpha_\rho}{rd_\rho}\left[\left(1+\frac{\mathsf{M}}{\varsigma}\right)\frac{2G}{\bar{K}+4G/3}\frac{1}{\tilde{\lambda}r\,sh\left(\tilde{\lambda}R\right)}\times\right.$$
$$\times\left(ch\left(\tilde{\lambda}r\right)-\left(1+\tilde{\lambda}^2r^2\right)\frac{sh\left(\tilde{\lambda}r\right)}{\tilde{\lambda}r}\right)+$$
$$\left.+\left(\frac{\bar{K}}{\bar{K}+4G/3}+\mathsf{M}\right)\frac{r}{\varsigma R}\right] \qquad (22)$$

$$\pi_m(r)=\mu_{\pi 0}\chi_m\left(1+\frac{\mathsf{M}}{\zeta}\right)\frac{\tilde{\lambda}R}{r\,sh(\tilde{\lambda}R)}\left[ch(\tilde{\lambda}r)-\frac{sh(\tilde{\lambda}r)}{\tilde{\lambda}r}\right]. \qquad (23)$$

На рис. 1 показано розподіл напружень σ_{rr} та $\sigma_{\varphi\varphi}$ у радіальному напрямку. Напруження σ_{rr} є стискальними. Напруження $\sigma_{\varphi\varphi}$ у при поверхневій області кулі є розтягувальними. Ці напруження досягають максимальних значень на поверхні кулі. У міру віддалення від поверхні напруження $\sigma_{\varphi\varphi}$ стрімко спадають. У точці $r=R'$ вони набувають нульових значень ($\sigma_{\varphi\varphi}(R')=0$), а далі стають стискальними.

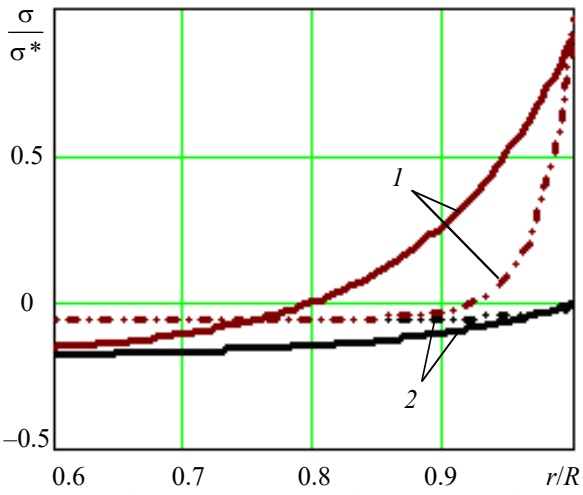


Рис. 1. Розподіл напружень $\sigma_{\varphi\varphi}$ (криві 1) та σ_{rr} (криві 2) у радіальному напрямку у кристалах хлористого натрію ($G/K=0.41$, $\mathsf{M}=10^{-3}$). Суцільна та штрих-пунктирна лінії відповідають $R=10l_*$ і $R==40l_*$, $\sigma^*=2GK\alpha_\rho\mu_{\pi 0}/d_\rho\left(\bar{K}+4G/3\right)$

Бачимо, що зі зменшенням радіуса кулі неоднорідність у розподілі напружень збільшується і охоплює ширші приповерхневі області.

Знаючи розподіл напружень у приповерхневій області тіла, згідно формули

$$F_n = \frac{1}{2\pi R}\int_{R'}^{R}\int_{0}^{2\pi} r\sigma_{\varphi\varphi}\, dr d\theta$$

можна обчислити поверхневий натяг F_n у кулі.

III. Вплив кривини поверхні на поверхневу енергію деформації

Використаємо розв'язок сформульованої задачі для обчислення поверхневої енергії деформації у кулі. Згідно [7] ця енергія визнається поверхневим значенням проекції вектора локального зміщення маси π_m на нормаль до поверхні (Σ), тобто

$$U_\Sigma = -\frac{1}{2}\rho_0 \mu'_{\pi 0}\mathbf{n}\cdot\pi_m\Big|_{\mathbf{r}\in(\Sigma)}. \qquad (24)$$

Ґрунтуючись на співвідношеннях (23) і (24), для поверхневої енергії деформації $U_\Sigma^{кулі}$ у суцільній кулі радіуса R маємо:

$$\frac{U_\Sigma^{кулі}}{U_\Sigma^{\infty}} = \frac{\tilde{\lambda}R\left[\tilde{\lambda}R\, ch(\tilde{\lambda}R) - sh(\tilde{\lambda}R)\right]}{(\tilde{\lambda}R)^2\, sh(\tilde{\lambda}R) - 4\dfrac{G}{K}\mathsf{M}\left(\tilde{\lambda}R\, ch(\tilde{\lambda}R) - sh(\tilde{\lambda}R)\right)}, \qquad (25)$$

де $U_\Sigma^{\infty} = -\dfrac{\rho_0}{2l_*}\mu'^2_{\pi 0}\,\chi_m$ — поверхнева енергія деформації у масивному тілі з цього ж матеріалу з плоскою поверхнею (півпросторі). Аналіз формули (25) показує, що при

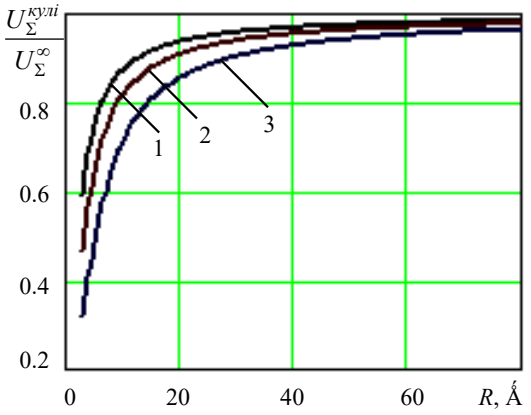


Рис.2. Залежність поверхневої енергії деформації від радіуса кривизни для різних матеріалів. Лінія 1: $G/K = 0.41$, $l_* = 1{,}3\cdot 10^{-10}$ м; лінія 2: $G/K = 0.45$, $l_* = 1{,}89\cdot 10^{-10}$ м, лінія 3: $G/K = 0.5$, $l_* = 3\cdot 10^{-10}$ м

$R \to +\infty$, $U_\Sigma^{кулі} \to -\dfrac{1}{2}\rho_0\mu^2_{\pi 0}\chi_m = U_\Sigma^{\infty}$.

Однак у наночастинках (кулях малих радіусів, сумірних із областю приповерхневої неоднорідності) поверхнева енергія деформації зменшується у міру збільшення кривини поверхні (див. рис. 2). Це узгоджується з результатами експериментальних спостережень [8-10]. Зазначимо, що криві 1 та 2 на рис. 1 відповідають хлористому натрію і хлористому калію. Характерні віддалі для цих матеріалів визначені з використанням теорії ґратки і наведені у працях [11, 12].

Отже формула (25) описує розмірний ефект поверхневої енергії деформації у твердих тілах сферичної геометрії.

IV. Висновки

Ґрунтуючись на співвідношеннях локально градієнтної пружності вивчено приповерхневу неоднорідність напружено-деформованого стану у суцільній ізотропній кулі. З використанням аналітичного розв'язку сформульованої крайової задачі математичної фізики визначено поверхневу енергію деформації та проаналізовано вплив на неї кривини поверхні. Показано, що зі збільшенням кривини поверхні абсолютне значення поверхневої енергії деформації зменшується, порівняно з тілом із плоскою межею.

Література References

[1] E. C. Aifantis, "Gradient effects at the micro-, macro- and nanoscales", *J. Mech. Behaviour of Materials*, vol. 5, no 3, pp. 355-375, 1994.

[2] A. Eringen, Nonlocal continuum field theories. Springer-Verlag, New York Inc., 2002.

[3] R. D. Mindlin, "Second gradient of strain and surface-tension in linear elasticity", *Int. J. Solids Struct.*, vol. 1, pp. 417–438, 1965.

[4] Я. Й. Бурак, Є. Я. Чапля, В. Ф. Кондрат, О. Р. Грицина, "Математичне моделювання термомеханічних процесів у пружних тілах із врахуванням локального зміщення маси", *Доп. НАН України*, № 6, с. 45-49, 2007.

[5] О. Грицина, "Крайові задачі нелокальної термопружності з урахуванням локального зміщення маси", *Фіз.-мат. моделювання та інформ. технології*, вип. 21, с. 79-88, 2015.

[6] И. Г. Марченко, И. М. Неклюдов, И. И. Марченко, "Коллективные процессы атомного упорядочения при низкотемпературном осаждении пленок", *Доп. НАН України*, № 10, с. 97-103, 2009.

[7] О. Грицина, "Визначення поверхневої енергії твердих тіл", *Фіз.-мат. моделювання та інформ. технології*, вип. 17, с. 43-54, 2013.

[8] Н. Т. Гладких, С. В. Дукаров, А. П. Крышталь и др, Поверхностные явления и фазовые превращения в конденсированных пленках. Под ред. Н. Т. Гладких, Харьков: ХНУ имени В. Н. Каразина, 276 с., 2004.

[9] В. М. Юров, Н. Х. Ибраев, С. А. Гученко, "Экспериментальное определение поверхностного натяжения наночастиц и нанопленок", *Известия высших учебных заведений. Физика,* № 1/3, с. 334-338, 2011.

[10] В. М.Юров, В. С. Портнов, Н. Х. Ибраев, С. А. Гученко, "Поверхностное натяжение твердых тел, малых частиц и тонких пленок", *Успехи современного естествознания. Физико-математические науки,* № 11, с. 55-58, 2011.

[11] A. Askar, P. C. Y. Lee, A. S. Cakmak, "The effect of surface curvature and discontinuity on the surface energy density and other induced fields in elastic dielectrics with polarization gradient", *Int. J. Solids and Struct.,* vol. 7, issue 5, pp. 523-537, 1971.

[12] R. D. Mindlin, Elasticity, piezoelectricity and crystal lattice dynamics. *J. Elast.,* vol. 2, pp. 217-282, 1972.

Автоматична Побудова Сематичної Мережі Тексту у Системах Запит-Відповідь

О. С. Волковський
кафедра автоматизованих систем обробки інформації
Дніпровський національний університет
імені Олеся Гончара
Дніпро, Україна
didivave@mail.ru

Є. Р. Ковилін
кафедра автоматизованих систем обробки інформації
Дніпровський національний університет
імені Олеся Гончара
Дніпро, Україна
kovilin.yegor@gmail.com

Automatic Construction of a Semantic Net of Text in the Request-Response Systems

O. Volkovskiy
Department of Automated information processing systems
Oles Honchar Dnipropetrovsk National University
Dnipro, Ukraine
didivave@mail.ru

Ye. Kovylin
Department of Automated information processing systems
Oles Honchar Dnipropetrovsk National University
Dnipro, Ukraine
kovilin.yegor@gmail.com

***Анотація*—Розглянуто структуру інтелектуальної системи запит-відповідь для отримання семантично зв'язних відповідей, генерованих на основі науково-технічних корпусів документів. Створено алгоритм автоматичної побудови семантичної мережі тексту на природній мові, розроблений в рамках завдання реалізації інтелектуальної системи запит-відповідь. Описано процес складання опорних міток для текстів, які характеризують семантичні зв'язки для тексту. На основі отриманого алгоритму була створена прикладна програмна система побудови графа смислових зв'язків науково-технічного тексту для мов слов'янської групи. Здійснено перевірку адекватності отриманої моделі на основі статистично нормалізованих текстів зі слабкими смисловими зв'язками.**

***Abstract*—The structure of the question-answer intellectual system for obtaining semantically connected answers which generated on the basis of scientific and technical corpus of documents is considered. An algorithm for the automatic construction of a semantic net of text in a natural language was developed for the task of implementing an intelligent request-response system. The process of compiling reference marks for texts characterizing semantic links for a text is described. On the basis of the obtained algorithm was created an applied software system for constructing a graph of the semantic links of the scientific and technical text for the languages of the Slavic group. The adequacy of the obtained model is checked on the basis of statistically normalized texts with weak semantic connections.**

Ключові слова—семантична мережа; система запит-відповідь; автоматична обробка тексту

Keywords—semantic network; question-answer system; automatic text processing

I. Вступ

Розробка прикладних програмних систем автоматичної обробки текстів (АОТ) має на увазі вибір того чи іншого механізму опису та реалізації моделі природної мови (ПМ), доступної ЕОМ. Оскільки, мова є досить не формалізованою системою з нестабільністю і неоднорідністю власних правил, то головною проблемою є складність опису семантичних характеристик тексту на рівні алгоритмічного уявлення. Оскільки ПМ це не просто набір слів, заснований на деяких граматичних складових (в задачах АОТ пріоритетним акцентом є отримання саме осмисленого тексту) це, в свою чергу, призводить багатьох розробників до необхідності врахування семантичних зв'язків не тільки між окремими словами, а й між реченнями і навіть між документами. Найчастіше, під семантикою і усвідомленням тексту машиною мається на увазі наступне: якщо ми ввели деякий текст в пам'ять ЕОМ і, скажімо, роздрукували його за допомогою принтера, про семантику годі й казати; але якщо над текстом виконується деяка обробка, в результаті якої користувач отримує зрозумілий і адекватний для нього новий текст (наприклад - переклад на іншу мову), то можна говорити про семантичне розуміння тексту комп'ютером. У цьому плані,

інтелектуальна генерація текстів є найбільш складним і цікавим завданням, оскільки при реалізації семантичної обробки шаблон синтезу тексту є більш нечітким, ніж, наприклад, при автоматичному перекладі. Тому в ході цієї роботи розглянута не тільки структура системи генерації тексту у вигляді відповіді на деякий запит, а й в першу чергу модель отримання семантичних зв'язків в тексті, доступна для прикладної програмної реалізації..

II. Система запит-відповідь

Загальна структура системи запит-відповідь що зараз розроблюється, зображена на рис.1. Найпершим вхідним параметром для системи є корпус документів, над яким проводиться операція розбиття на кластери, в результаті якої кожному отриманому кластеру відповідає набір семантичних міток, що характеризують семантичні властивості кожного документа, що міститься в кластері. Після цього виконується додавання розміченого таким чином кластера в загальний предметний корпус системи. При надходженні запиту користувача в систему, виконується його перетворення в набір опорних міток, на основі яких приймається рішення про формування визначаючого майбутню відповідь корпусу документів. По кожному з документів у визначальному корпусі виконується формування множини речень для включення їх в результуючу відповідь системи. Проаналізувавши отриману таким чином множину, система формує семантично пов'язану відповідь на запит, що надійшов.

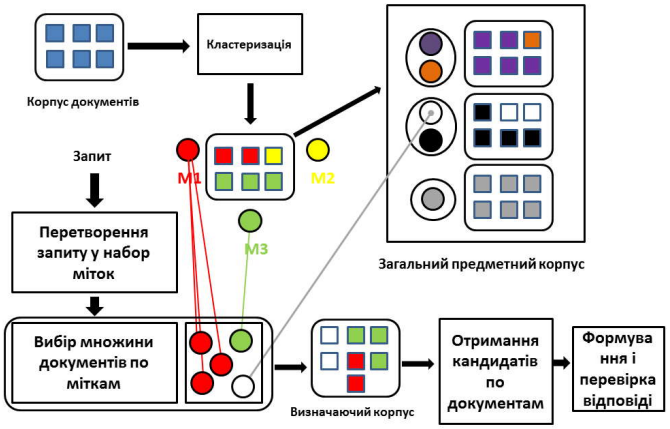


Рис. 1. Загальна структура роботи системи запит-відповідь

Одним з ключових моментів описаного алгоритму є отримання семантичних міток для документа, що має на увазі проведення семантичного аналізу всього документа і складання деякої опорної моделі для окремого тексту в корпусі. У нашому випадку, найкращим варіантом є побудова семантичної мережі документа, оскільки ця модель відображає смислові зв'язки між своїми елементами. Однак, програмна реалізація даної моделі викликає ряд складнощів, які є фундаментальними для більшості систем автоматичної обробки текстів. Йдеться і про відсутність чітких алгоритмів і математичних моделей, що описують саме поняття «семантична мережа документа» (і процес її отримання відповідно), і про самі властивості природної людського мови - встановлення зв'язків між елементами документа, написаного на мові флективно-багатої групи з вільним порядком слів (до якої відносяться всі слов'янські мови) значно складніше, ніж отримання аналогічних зв'язків, наприклад, в англомовному документі (де флексія набагато нижча і порядок слів жорстко фіксований граматичними правилами). Тому, складання семантичної мережі документа залишається відкритим та актуальним науковим питанням галузі автоматичної обробки текстів. В ході цієї роботи був складений і реалізований підхід до створення семантичних мереж документа.

III. Система побудови семантичної мережі документа

Алгоритм роботи системи зображений на рис.2. Першим етапом після завантаження документа і проведення синтаксичного аналізу (виділення речень і слів) стає визначення частини мови для кожного із знайдених слів. Для цього в систему була включена навчальна вибірка з 15 тис. слів і відповідних їм частин мови. На основі трьох типів закінчень (дві та три останні букви слова і закінчення отримане через стемінг за алгоритмом Портера) для кожного з елементів навчальної вибірки проводиться навчання наївного Баєсовського класифікатора, після чого на основі навченої моделі проводиться встановлення частини мови для кожного з виділених слів в тексті. Для підвищення точності в систему були включені кінцеві словники службових частин мови.

Рис. 2. Загальна структура роботи системи побудови семантичної мережі тексту

Після цього етапу проводиться визначення унікальних стем: для кожної пари слів проводиться відсікання приставки і закінчення за алгоритмом Портера - якщо довжина найбільшої загальної частини більше або дорівнює відстані Левенштейна для даної пари слів - то аналізоване слово замінюється на отриману стему. В результаті текст має такий вигляд (таб. 1) - для кожної із стем система визначає її частину мови та кількість входжень стеми в текст. На цьому етапі з тексту забираються всі слова помічені як службові частини мови.

ТАБЛИЦЯ I. РЕЗУЛЬТАТ СИНТАКСИЧНОЇ ОБРОБКИ

Текст до обробки	**Тест після обробки**
Сьогодні існують різні книги, відеокурси з програмування та інші способи швидко і відносно недорого навчитися писати програми і додатки.	|2|[годн{ADV}]|3|[існую{V}]|2|[різн{A}]|2|[книг{S}]|3|[відеокурс{S}]|22|программ{S}]|4|[інш{S}]7|[способ{S}]|2|[швидк{ADV}]|2|[орого{A}]|3|[навчити{V}]|8|[писа{V}]|22|[программ{S}] |2|[додатки{S}]

Для кожної із стем (загальною кількістю N) і кожного речення (загальною кількістю M) складається матриця $N*M$, значення якої визначаються кількістю входжень стем в речення. Над отриманої матрицею виконується операція сингулярного розкладання і проектування отриманих даних на площину. Оскільки сингулярне розкладання є стійким, ми можемо відкинути ті значення лівої і правої матриць, які відповідають низьким сингулярним значенням, залишивши тільки перші два, що представляють собою вектори координат двовимірної площині для стем і для речень. Отримана проекція зображена на рис.3.

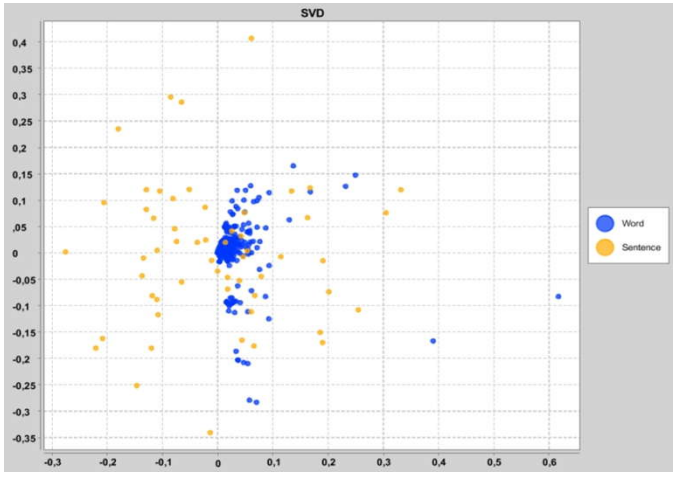


Рис. 3. Проекція сингулярного розкладання: Word – координати-стеми, Sentence – координати-речення

Наступним кроком стає кластеризація точок-координат для стем і для речень по алгоритму k-means. Кількість кластерів для стем і для речень cl визначається по формулі (1):

$$cl(W, W_U) = \frac{count(W)}{count(W_U)} \quad (1)$$

де W – слова, W_U – стеми. Центроїдами кластерів-стем є координати стем з найбільшою кількістю входжень в текст, що визначається за формулою (2):

$$Cst(W_U) = \max(W_0 ... W_{cl}) \quad (2)$$

де $W_0 \ldots W_{cl}$ – ваги стем. Центроїдами кластерів-речень є координати речень з найбільшим загальною вагою стем, яка визначається по формулі (3):

$$Cs(W_S) = \max\left(\sum_{i=0}^{SN} W_i\right) \quad (3)$$

де W_S – речення, W_i – вага стеми у речені, SN – кількість стем у реченні. Результат цих операцій можна побачити на рис. 4. (для стем) і на рис. 5 (для речень).

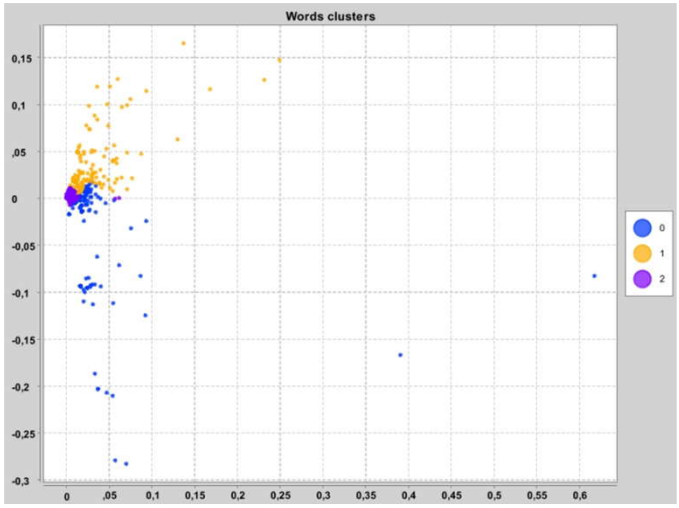


Рис. 4. Проекція кластерізації для стем. 0, 1, 2 – номера кластерів

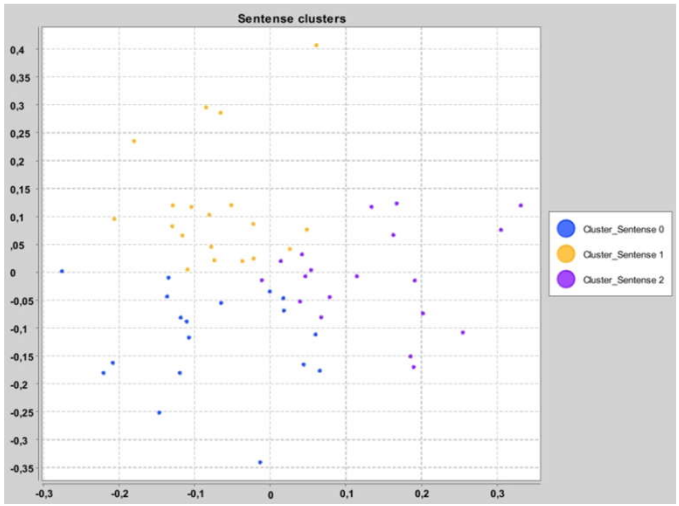


Рис. 5. Проекція кластерізації для реченнь. Cluster_Sentence 0, Cluster_Sentence 1, Cluster_Sentence 2 – номера кластерів

Заключним етапом стає формування семантичної мережі документа. На основі координат точок кожного кластера-стеми за алгоритмом Джарвіса будується контур опуклої фігури. Для кожного кластера-стеми визначається вага - кількість вхідних в нього стем, на основі чого будується семантичний граф зв'язків кластерів в порядку убування їх ваги. Для кожної фігури кластерів-стем, отриманої по алгоритму Джарвіса, перевіряється попадання точок що формують кожен кластер-речення.

Якщо такі точки знайдені - кластер речення приєднується в мережі до кластеру-стеми, де вага зв'язку - це кількість точок, що потрапили в контур кластера-стеми. Результат роботи системи зображений на рис. 6.

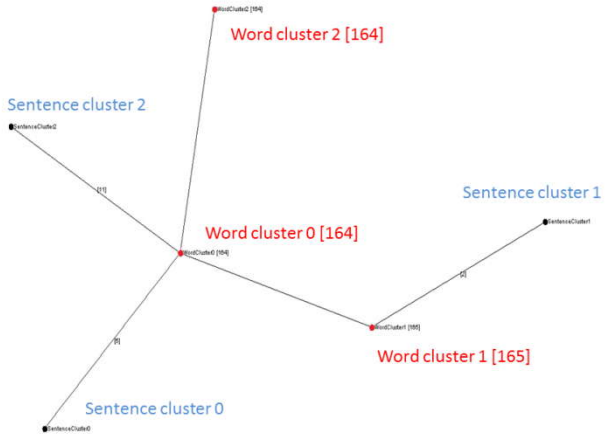


Рис. 6. Результат роботи системи – семантична мережа документа. WordCluster відповідають кластерам-стемам, SentenceCluste відповідаюсть кластерам-реченням.

IV. Аналіз отриманих результатів та висновки

Ґрунтуючись на отриманих результатах, стає можливим формування семантичної мітки документа в системах запит-відповідь. Так, мітка складається з тих кластерів-стем, які мають зв'язок з хоча б одним кластером-реченням, а для отримання кандидатів по мітці використовується перевірка входження мітки-запиту в мітку документа, і якщо таке входження знайдено - проводиться отримання кластера-речення з найбільшою вагою, пов'язаним зі знайденою міткою.

Для перевірки адекватності отриманої моделі система була перевірена на текстах, створених в результаті автоматичної генерації на основі патернів. Такі тексти є статично правильними, але мають слабкі смислові зв'язки між своїми частинами. Результат обробки такого тексту зображено на рис.7 (проекція кластерів стем та речень) та на рис.8 (результуюча семантична мережа).

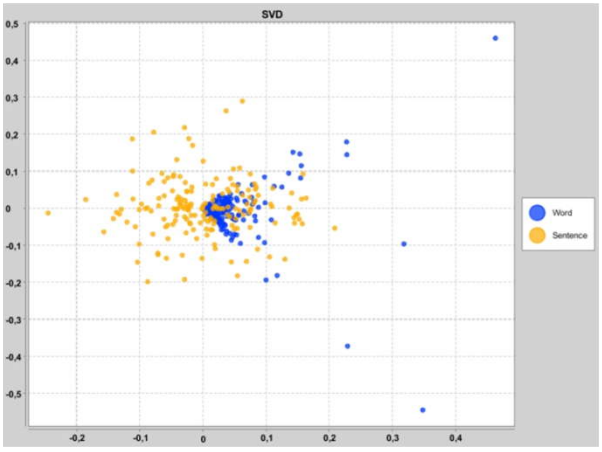


Рис. 7. Проекція сингулярного розкладання для тексту зі слабкими семантичними зв’язками: Word – координати-стеми, Sentence – координати-речення.

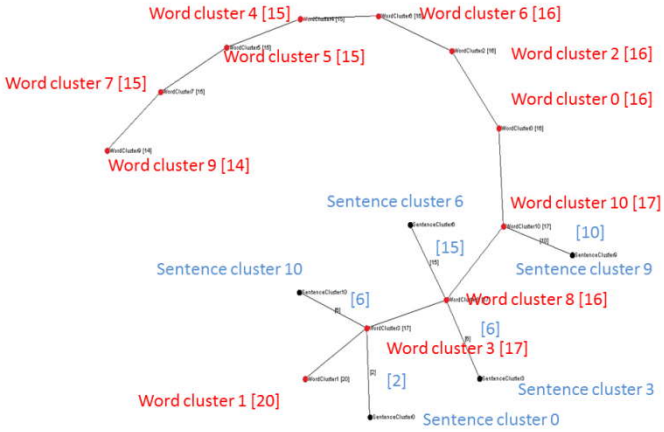


Рис. 8. Результат роботи системи для тексту зі слабкими семантичними зв’язками – семантична мережа документа. WordCluster відповідають кластерам-стемам, SentenceCluste відповідаюсть кластерам-реченням.

Незважаючи на те, що обсяг автоматично згенерованого тексту співпадав з прикладом, наведеним раніше, його семантична мережа має яскраво виражені відмінності. Це спостереження дозволяє зробити припущення не тільки про адекватність семантичної моделі, а так само і про можливості її застосування за межами завдання побудови систем запит-відповідь. Такі дані, як кількість кластерів-речень та кластерів-стем, кількість зв’язків та їх вага, ваги кластерів-слів, що містить в собі семантична мережа, можна використовувати для навчання моделей, які входять, наприклад, в системи автоматичного визначення плагіату або зв'язності тексту.

Література References

[1] Н.Н. Леонтьева. Автоматичне розуміння текстів. Системи, моделі, ресурси. [Текст]/Н.Н. Леонтьева//Москва – 2006

[2] М.В. Мозговой. Машиний семантичний аналіз російської мови и його застосування. [Текст]/Мозговой М. В.//СПбГУ, Санкт-Петербург –116с. – 2006г.

[3] N. Chomsky (2002) Syntactic Structures [Text]/Chomsky N//Berlin, New York: Mouton de Gruyter, 2002.

[4] Н.І. Гурін Семантична мережа электронного підручника для діалогу с віртуальним викладачем [Текст] / Н.І. Гурін, Я.А. Жук //Материали міжнародної научно-техничної інтернет конференції "Інформаційні технології в науці та виробництве"//Білоруський державний технологічний університет, Мінск, 2015 г..

[5] О.Н. Швецов. Система синтеза навчальних тестів заснована на формальних граматиках [Текст]/Швецов О.Н.//журнал «Программні продукти і системи», №2(102), 2013, с 181-185.

[6] I.A. Bolshakov. (2000) The Meaning ↔ Text Model: Thirty Years After. [Текст]/I.A. Bolshakov, A.F. Gelbukh//J. International Forum on Information and Documentation, N 1, 2000.

[7] D.S. Tarasov. (2015) - Natural language generation, paraphrasing and summarization of user reviews with recurrent neural networks [Text]/Tarasov D. S./ "Computer linguistics and Intellectual Technologies", No14(vol. 1), 2015, p.607-614//Materials of international conference "Dialog", 2015.

[8] Y.D. Apresyan. (1989) Linguistic support of the ETAP-2 system [Text]/Y. D. Apresyan, I. M. Bulavskiy, L.L. Iomdin//M.: Nauka, 1989.

[9] P. Bartlett, J. Shawe-Taylor. (1998) Generalization performance of support vector machines and other pattern classifiers//Advances in Kernel Methods. — MIT Press, Cambridge, USA, 1998.

Математичне Моделювання Кінетики Квазістатичного Фільтраційного Насичення Пористого Шару Природнім Газом

Василь Чекурін
Відд. математичних проблем механіки неоднорідних тіл
ІППММ ім. Я. С. Підстригача НАН України
Львів, Україна
v.chekurin@gmail.com

Зоя Притула
Відділ аналізу, геометрії та топології
ІППММ ім. Я. С. Підстригача НАН України
Львів, Україна
zoya777b@gmail.com

Mathematical Modelling of Kinetics of Quasistatic Filtrational Saturation of a Porous Layer by Natural Gas

Vasyl Chekurin
Dept. of Math. problems of mechanics of heterogen. solids
Pidstryhach IAPMM of NASU
Lviv, Ukraine
v.chekurin@gmail.com

Zoya Prytula
Dept. of analysis, geometry and topology
Pidstryhach IAPMM of NASU
Lviv, Ukraine
zoya777b@gmail.com

***Анотація*—Досліджено процеси перерозподілу тиску газу в пористому плоскому шарі за малих відхилень початкового тиску. За таких умов отримано лінеаризоване рівняння нестаціонарної фільтрації газу, коефіцієнти якого залежать від параметрів пористого середовища, термодинамічних властивостей газу та величини незбуреного тиску в шарі. Сформульовано задачу кінетики вирівнювання малих локальних збурень тиску в пористому шарі. З використанням функції Гріна отримано її аналітичний розв'язок. На основі отриманого розв'язку кількісно досліджено вплив коефіцієнта проникливості та величини незбуреного тиску на швидкість загасання тиску газу.**

***Abstract*—We investigate the processes of gas pressure redistribution in a porous plane layer under small perturbations of initial pressure. Under these conditions, we have obtained a linearized equation of nonstationary gas filtration with the coefficients depending on the parameters of the porous medium, the thermodynamic characteristics of gas and the value of unperturbed pressure. A problem of a kinetics of equalization of gas pressure small local perturbations in the porous layer has been formulated. Using the Green function, we have obtained an analytical solution of the problem. On the basis of obtained solution we have quantitatively investigated the influence of permeability coefficient and the value of unperturbed pressure on the rate of damping pressure.**

Ключові слова—пористий шар; фільтрація газу; мале збурення тиску газу; коефіцієнт проникливості; незбурений тиск; швидкість загасання.

Keywords—porous layer; gas filtration; small perturbation of gas pressure; permeability coefficient; rate of damping.

I. Вступ

Підземні сховища часто споруджують у пористих геологічних пластах [1, 2]. Експлуатаційні характеристики таких об'єктів, зокрема їхні ємнісні властивості, часові та енергетичні характеристики процесів відбору й нагнітання газу тощо істотно залежать від коефіцієнтів пористості, проникливості та інших фізичних параметрів середовища [3-8].

Під час експлуатації сховища процеси нагнітання та відбору газу чергуються із процесом зберігання. В ході зберігання тиск в пористому пласті вирівнюється, а при нагнітанні та відборі відбуваються локальні збурення тиску в околі робочих свердловин.

Дані моніторингу тиску в робочих та контрольних свердловинах можна використати для оцінювання експлуатаційних властивостей сховища, зокрема, їхніх пікових

характеристик, які визначають максимальну продуктивність процесів відбору та нагнітання газу.

Це можливо реалізувати шляхом розв'язування відповідних обернених задач нестаціонарної фільтрації в пористих шарах. Проте, на сьогодні такі задачі ще навіть не сформульовані математично. Це пояснюється, зокрема, тим, що фільтраційні процеси в пористих середовищах описуються суттєво нелінійними диференціальними рівняннями із частинними похідними, для яких не відомі аналітичні розв'язки. Друга причина полягає у тому, що фізичні параметри пористих пластів, в яких споруджені сховища, як правило вивчені недостатньо.

У доповіді представлено результати математичного моделювання кінетики насичення пористого пласта за повільних нагнітання або відбору газу через свердловину. Процес нагнітання (відбору) за таких умов моделюється стрибкоподібним збільшенням (зменшенням) тиску в привибійній зоні свердловини на величину, достатньо малу, порівняно із пластовим тиском. Оскільки за таких умов відхилення від рівноважного стану є малі, то ці процеси з достатньою точністю можна описати і за допомогою відповідних лінеаризованих рівнянь [9].

Тому у цій роботі розглянуто кінетику процесу вирівнювання тиску газу в пористому плоскому шарі за малих збурень. Для опису процесу фільтрації за таких умов проведено лінеаризацію рівняння нестаціонарної фільтрації. Коефіцієнти отриманого лінеаризованого рівняння залежать від характеристик пористого середовища, термодинамічних властивостей газу та незбуреного тиску.

В рамках такої математичної моделі кількісно досліджено процеси вирівнювання малих локальних збурень тиску газу в шарі залежно від коефіцієнта проникливості пористого середовища та величини незбуреного тиску.

II. ФОРМУЛЮВАННЯ ТА РОЗВ'ЯЗУВАННЯ ЗАДАЧІ КІНЕТИКИ ВИРІВНЮВАННЯ ТИСКУ ГАЗУ В ПОРИСТОМУ ШАРІ ЗА МАЛИХ ЗБУРЕНЬ

Розглянемо процес одновимірної ізотермічної фільтрації газу в пористому шарі, який описується таким нелінійним рівнянням [10, 11]

$$m\frac{\partial P}{\partial t}=\frac{k\sqrt{P}}{\mu}\frac{\partial^2 P}{\partial x^2}. \qquad (1)$$

Тут $P=p^2$, $p=p(x,t)$ – тиск, k – коефіцієнт проникливості, μ – динамічний коефіцієнт в'язкості газу, m – коефіцієнт пористості, x – просторова координата, t – час.

Поверхні шару $x=0$ та $x=L$ є непроникними. Тоді граничні умови відносно змінної P мають такий вигляд

$$\left.\frac{\partial P}{\partial x}\right|_{x=0}=0, \qquad \left.\frac{\partial P}{\partial x}\right|_{x=L}=0. \qquad (2)$$

Вважаємо, що задано розподіл тиску в початковий момент часу:

$$P(x,0)=P_0(x). \qquad (3)$$

Оскільки сформульована задача (1)-(3) є нелінійною, то проведемо дослідження кінетики перерозподілу тиску газу за малих відхилень початкового розподілу від однорідного.

Розглянемо випадок, коли кількість газу, яка є в шарі залишається незмінною, тобто

$$m\int_0^L \rho(x,t)\,dx=M=const.$$

Тут ρ – густина.

Нехай $p(x)$ – розподіл тиску в початковий момент часу, а $\overline{p}$ – однорідний тиск, який встановиться після вирівнювання. Нехай відхилення $p(x)-\overline{p}$ є достатньо малим: $\frac{p(x)-\overline{p}}{\overline{p}}<<1$. Оскільки процес перерозподілу газу є релаксаційним, то і для моменту часу $t>0$, має виконуватись умова: $\frac{p(x,t)-\overline{p}}{\overline{p}}<<1$. Ця умова дозволяє лінеаризувати рівняння (1).

Означимо збурення змінної P таким чином

$$\tilde{P}(x,t)=\frac{P(x,t)-\overline{P}}{\overline{P}},$$

де $\overline{P}=\overline{p}^2$.

З урахуванням малості збурення $\tilde{p}=\frac{p(x,t)-\overline{p}}{\overline{p}}$, отримуємо співвідношення:

$$\tilde{P}=2\tilde{p}.$$

Лінеаризоване рівняння відносно змінної P має такий вигляд

$$\frac{\partial\tilde{P}(x,t)}{\partial t}=\frac{k\overline{p}}{\mu m}\frac{\partial^2\tilde{P}(x,t)}{\partial x^2}, \qquad (4)$$

а граничні умови на поверхнях шару $x=0$, $x=L$ та початкова умова для збурення $\tilde{P}$ відповідні такі:

$$\left.\frac{\partial\tilde{P}}{\partial x}\right|_{x=0}=0, \qquad \left.\frac{\partial\tilde{P}}{\partial x}\right|_{x=L}=0, \qquad (5)$$

$$\tilde{P}(x,0)=\tilde{P}_0(x)\,. \qquad (6)$$

У результаті сформульовано задачу (4)-(6), яка описує кінетику вирівнювання тиску в шарі для випадку малих локальних збурень.

Задачу (4)-(6) розв'язуємо, використовуючи метод функцій Гріна.

Функція Гріна цієї задачі має вигляд [12]

$$G(x,\xi,t)=\frac{2}{L}\sum_{n=1}^{\infty}\cos\frac{\pi n x}{L}\cos\frac{\pi n\xi}{L}e^{-\frac{k\bar{p}\pi^2n^2}{\mu m L^2}t},$$

а навантажувальна функція відповідна така

$$\omega(x,t)=\tilde{P}_0(x)\delta(t)\,,$$

де $\delta(t)$ – функція Дірака.

З урахуванням цього, а також співвідношення між збуреннями $\tilde{P}$ та $\tilde{p}$, для збурення тиску $\tilde{p}$ в результаті отримуємо

$$\tilde{p}(x,t)=\frac{1}{L}\sum_{n=1}^{\infty}\cos\frac{\pi n x}{L}e^{-\frac{k\bar{p}\pi^2n^2}{\mu m L^2}t}\int_0^L\tilde{P}_0(\xi)\cos\frac{\pi n\xi}{L}d\xi\,. \qquad (7)$$

Таким чином, з використанням функції Гріна отримано аналітичний розв'язок задачі (4)-(6) про вирівнювання тиску в пористому шарі за малих збурень.

III. Дослідження впливу коефіцієнта проникливості та величини незбуреного тиску на швидкість загасання

На основі отриманого розв'язку (7) кількісно досліджено вплив коефіцієнта проникливості k пористого середовища, а також величини незбуреного тиску $\bar{p}$ на швидкість загасання тиску.

На рис. 1 показано кінетику загасання тиску в точці $\tilde{x}=0.5$ $\tilde{p}(0.5,t)$, нормованого на його значення в початковий момент часу $\tilde{p}_m=\tilde{p}(0.5,0)$: $\tilde{p}_n=\frac{\tilde{p}}{\tilde{p}_m}$, при різних значеннях k, а саме, $k=5\cdot10^{-12};\ 5\cdot10^{-13};\ 2\cdot10^{-13};\ 1{,}1\cdot10^{-13};\ 5\cdot10^{-14}\ м^2$ (криві 1-5 відповідно). При цьому $\bar{p}=50\cdot98066{,}5\ Па$; $\mu=1{,}283\cdot10^{-5}\ Па\cdot с$; $m=0{,}2$; $L=1000\ м$; $\tilde{x}=\frac{x}{L}$ – безрозмірна координата.

Зазначимо, що дані для фільтраційних характеристик відповідають суміші газів, в якій 96 % складає метан.

Функцію $\tilde{P}_0(\xi)$ вибрано у вигляді дельта-подібної функції:

$$\tilde{P}_0(\xi)=\frac{b}{1+\left(\frac{\xi}{a}\right)^2}+c\xi^2+d\,,$$

де коефіцієнти c та d мають такий вигляд:

$$c=\frac{a^2b}{\left(1+a^2\right)^2},$$

$$d=\frac{-\frac{1}{3}ab\left(3arctg\frac{1}{a}\left(1+a^2\right)^2+a\right)}{\left(1+a^2\right)^2}.$$

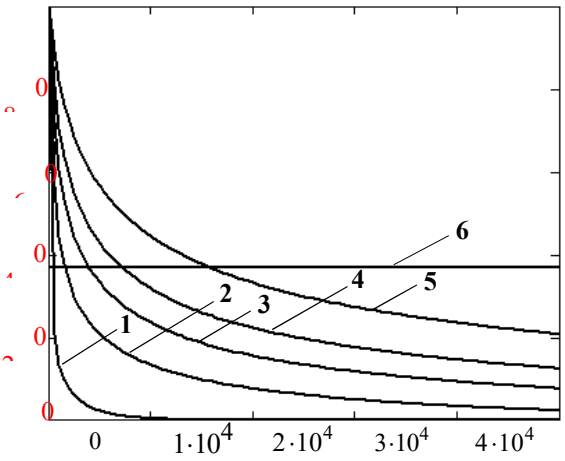


Рис. 1. Кінетика загасання тиску в точці $\tilde{x}=0.5$ за різних значень k

Проведено також оцінку часу релаксації на основі часу зменшення максимального значення тиску в точці $\tilde{x}=0.5$ в e раз.

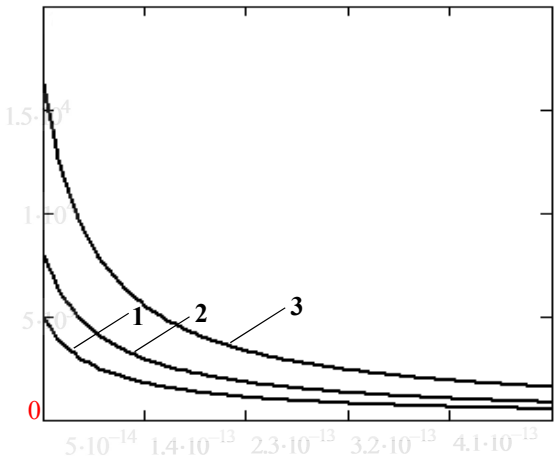


Рис. 2 Залежність часу загасання тиску від k за різних значень $\bar{p}$

Крива 6 на рис. 1 зображує характерне значення $1/e$. Для цього характерного значення показано залежності часу загасання τ від коефіцієнтів проникливості k, зокрема, це ілюструють криві 1-3 на рис. 2, які побудовано відповідно для $\bar{p} = 160 \cdot 98066{,}5;\ 100 \cdot 98066{,}5;\ 50 \cdot 98066{,}5$ Па.

Як видно із рис. 2, зі збільшенням коефіцієнта проникливості k, час загасання τ зменшується, причому чим більше значення незбуреного тиску $\bar{p}$ в шарі, тим менший час загасання.

На рис. 3 проілюстровано кінетику загасання тиску $\tilde{p}$ в точці $\tilde{x} = 0.5$ $\tilde{p}(0.5,t)$ за різних значень незбуреного тиску $\bar{p}$: $160 \cdot 98066{,}5;\ 100 \cdot 98066{,}5;\ 60 \cdot 98066{,}5;\ 30 \cdot 98066{,}5;\ 10 \cdot 98066{,}5$ Па (криві 1-5 відповідно). Крива 6 на цьому рисунку зображує характерне значення $1/e$.

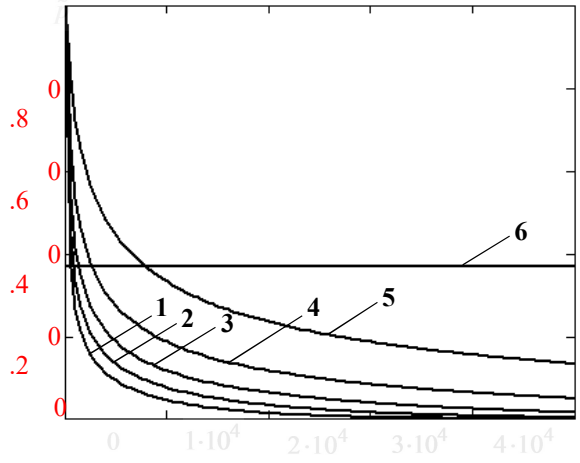


Рис. 3 Кінетика загасання тиску в точці $\tilde{x} = 0.5$ за різних значень k

На рис. 4 наведено залежності часу загасання τ від величини незбуреного тиску $\bar{p}$.

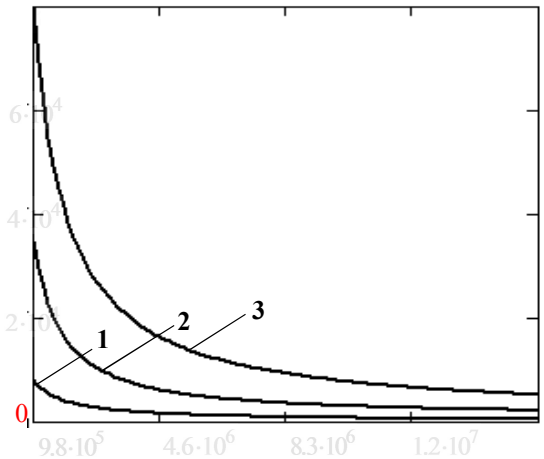


Рис. 3 Залежність часу загасання тиску від $\bar{p}$ за різних значень k

Криві 1-3 на рис. 4 відповідають таким значенням коефіцієнта k: $k = 5 \cdot 10^{-13};\ 1.1 \cdot 10^{-13};\ 5 \cdot 10^{-14}$ м^2.

Як видно з рис. 4, зі збільшенням незбуреного тиску $\bar{p}$, час загасання зменшується. При цьому, чим більший коефіцієнт проникливості, тим менший час загасання.

IV. Висновки

У роботі досліджено кінетику процесу вирівнювання тиску газу в пористому плоскому шарі за малих збурень. Для опису процесу фільтрації за таких умов проведено лінеаризацію рівняння нестаціонарної фільтрації. Отримано лінеаризоване рівняння, коефіцієнти якого залежать від характеристик пористого середовища, термодинамічних властивостей газу та незбуреного тиску. Одержане рівняння доповнено відповідними граничними умовами на поверхнях шару та початковою умовою.

Сформульовано задачу кінетики вирівнювання тиску в шарі для випадку малих локальних збурень. З використанням функції Гріна одержано аналітичний розв'язок сформульованої задачі.

Встановлено кількісні закономірності процесів вирівнювання малих локальних збурень тиску газу в шарі залежно від коефіцієнта проникливості та величини незбуреного тиску.

Література References

[1] К. С. Басниев, И. Н. Кочина, В. М. Максимов, Подземная гидромеханика. М.: Недра, 1993.

[2] Х. Азис, Э. Сеттари, Математическое моделирование пластовых систем. Москва-Ижевск: Институт компьютерных исследований, 2004.

[3] Н. М. Притула, Я. Д. Пянило, М. Г. Притула, Підземне зберігання газу (математичні моделі та методи). Львів: Видавництво «Растр-7», 2015.

[4] S. Y. Misyura, "The influence of porosity and structural parameters on different kinds of gas hydrate dissociation", *Scientific Reports,* vol. 6, p. 30324, July, 2016.

[5] T. P. Miroshnichenko, N. A. Lutsenko, V. A. Levin, "Gas filtration from an underground reservoir at a large initial pressure gradient," *J Appl. Mech. Tech Phy.*, vol. 56, no. 5, pp. 864–869, Sept., 2015.

[6] L. D. Eskin, "A self-similar solution to the equation of gas filtration in a spherically symmetric porous medium," *Russ. Math.*, vol. 52, no. 8, pp. 48–57, 2008.

[7] A. V. Akhmetzyanov, "Computational aspects in controlling filtration of fluids and gases in porous media," *Autom. Remote Control.*, vol. 69, no. 1, pp. 1–12, Jan., 2008.

[8] S. D. Algazin, "Numerical study of single-phase gas filtration in a porous medium," *J. Appl. Mech. Tech. Phys.*, vol. 52, no. 4, p. 615, July., 2011.

[9] В. Чекурін, З. Притула, "Математичне моделювання кінетики вирівнювання тиску газу в пористому шарі за малих збурень", *Вісник НУ «Львівська політехніка»: Комп. науки та інформ. техн.*, "подано до друку".

[10] Г. Б. Пыхачев, Р. Г. Исаев, Подземная гидравлика (учебное пособие). М.: Недра, 1973.

[11] Л. С. Лейбензон, Движение природных жидкостей и газов в пористой среде. М.: ОГИЗ, Гос. изд-во технико-теорет. лит., 1947.

[12] А. Г. Бутковский, Характеристики систем с распределенными параметрами (справочное пособие). М.: Наука, 1979.

Моделі Штучного Життя в Клітинному Просторі Ресурсів

В. Завертаний
Інститут прикладного системного аналізу
Національного технічного університету України
(Київський політехнічний університет ім.. І. Сікорського)
Київ, Україна
v.zavertanyy@gmail.com

О. Макаренко
Інститут прикладного системного аналізу
Національного технічного університету України
(Київський політехнічний університет ім.. І. Сікорського)
Київ, Україна
makalex51@gmail.com

Artificial Life Models in Cellular Resource Space

V. Zavertaniyy
Institute for Applied System Analysis at
National Technical University of Ukraine
(Igor Sykorski Kiev Politechnical Institute)
Kyiv, Ukraine
v.zavertanyy@gmail.com

A. Makarenko
Institute for Applied System Analysis at
National Technical University of Ukraine
(Igor Sykorski Kiev Politechnical Institute)
Kyiv, Ukraine
makalex51@gmail.com

***Анотація*—Розглядається підхід, заснований на моделюванні складних адаптивних систем, відомий як штучне життя. Такий підхід дозволяє роботу з внутрішньої адаптацією системи, з впливом організму на його навколишнього середовище та інші організми, зі зміною всієї біосфери і в кінцевому підсумку зі своєю власною можливістю існувати, тобто його власна придатність. В області досліджень штучного життя цифрових екосистем, такий підхід дає можливість простежити апостеріорну придатність, який можна розглядати як один з ознак виникнення властивостей системи, таких як чисельність населення, угруповання або стабільності виставленої поведінки. У цій статті розглянуто модель, схожу на класичні моделі штучного життя на просторової решітці і обговорено відносини між агресивною та мирною поведінкою через доступні ресурси в системі. Гетерогенний ландшафт ресурсів і його вплив на поведінку агента вводяться і розглядаються з точки зору стійкості видів. Також досліджувалася стійкість видів.**

***Abstract*—The individual-based approach in the modeling of complex adaptive systems, known as Artificial Life, is considered. Such approach allows dealing with the intrinsic adaptation of a system, with the influence of an organism on its environment and on other organisms, with altering the whole biosphere and eventually with its own possibility to exist, i.e. its own fitness. In the Artificial Life research field of digital ecosystems, such approach provides the ability to trace an a posteriori fitness, which can be treated as one of a system's emergent features like population size, grouping or stability of exhibited behavior. In this article, we explore the model similar to classic Artificial Life models on spatial lattice and discuss relation between combat and peaceful behavior due to available resource in the system. The heterogeneous resource landscape and its impact on agent's behavior is introduced and examined from the point of view of species sustainability. The species sustainability is also investigated.**

Ключові слова—моделі, штучне життя, колективна поведінка, клітинний простір, агресивні стратегії

Keywords—models, artificial life, collective behavior, cellular space, aggressive strategy

I. INTRODUCTION

Artificial Life (Alife) is an interdisciplinary research field, which tries to investigate and use the properties of living systems or systems which include a large number of living components (for example, individuals). Alife usually brings together biologists, philosophers, physicists, computer scientists, chemists, mathematicians, artists, engineers, and others. Alife applications are numerous and belong to different research fields such as artificial (digital) ecosystems, artificial society, evolutionary robotics, biology, origin of life (see for example [1], [2], [3], [5], [18], [21], [22], [24]), and many others. Alife systems have been implemented as software and as hardware (see recent reviews [11], [22]). One of the

important examples of the software Alife studies builds and explores digital ecosystems that provide novel methods to study evolution. These studies can be useful in answering questions about laws how evolution works and how to manage it. Traditional evolution in real biological systems is extremely slow for study. The computational Alife aims to put the evolution process into action on a computer, so the speed of evolution is only limited by processor performance. Embracing evolution instruments opens opportunities for researching a great variety of problems that are linked with it. Artificial evolving systems are used to build complex systems that expose intellectual behavior and to study the link between intellectuality and complexity [15]. Alife systems are plausible playground to explore the mechanisms of adaptation: general evolving system features such as speciation ([14], [15]), aging ([17]), cooperation ([19]), developmental processes in artificial systems [10], and learning.

Many models are developed for the purpose of studying social, ecological, swarming, artificial life and other topics. Despite the progress of other models, the interconnection between genotype and phenotype dynamic is still quite an unexplored issue; in current study we reveal an example of such unclearness that lurks in dynamic of the system. As one of the goals of the study, we want to concentrate on more detailed research of an agent phenotype sustainability. Further in this work, we discuss the dependency of combat interaction from input resource value and examine the sustainability of phenotypic assembly formation in homogeneous and heterogeneous spaces. These questions fit into the research field of Artificial Life determined by Bedau [5], and belong to a group of research areas that claim to: Determine predictability of evolutionary consequences of manipulating organisms and ecosystems, determine minimal conditions for evolutionary transitions from specific to generic response systems, determine what is inevitable in the open-ended evolution of life.

Alife consolidate different research fields, such as, for example, hardware and software Alife. It could be used to study the evolution of complexity, robotics, and digital organisms. One of the main approaches of constructing simulation models in Alife is multi-agent methodology that is broadly used in the study of complex adaptive systems. Individual-based approach surmounts difficulties of equation-based models by granting additional flexibility for both development and analysis of the model [11]. The popularity of multi-agent approach springs from early researches such as Sugarspace [12], Bugs [20], Echo [14] and Polyworld [23] models. One of the pioneering models of Artificial Life is the model of bugs on spatial lattice that was proposed by Norman H. Packard [20] denotes the importance of shift from extrinsic to intrinsic adaptation approaches in the modeling of evolutionary processes. The author of [20] defines the intrinsic adaptation of a system as a process of changes in interactions of all parts of the system aiming to fit it and permanently changing the environment. As a result of first simulations of his model, H. Packard introduced the notion of an a posteriori fitness function for the intrinsic adaptation evolutionary process and demonstrated with its help the emergence of specific behavior that is inherent for some individuals. This change in the concept of adaptation shifts the focus to the emerging characteristics of the system that can be treated as an aposteriori fitness function. The examples of such values could be population size over time, sustainability of emerging phenotypic assemblies under different factors such as environmental changes or arm races, and other system features. In particular work the size of agents' group with common phenotype (behavior strategy) is treated as the aposteriori fitness function.

Echo model is a Complex Adaptive System that was built with a purpose of extending genetic algorithms approach to ecological setting by adding geography (location), competition for resources and interaction among individuals (coevolution). The Echo model itself is intended to study different patterns of ecosystem behavior, which are in particular how resources flow through different kinds of ecologies takes place and how cooperation among agents can arise through evolution and arms races. The investigations allow identifying parameters or collections of parameters that are critical for emergence of specific behavior, i.e., to perform sensitivity analysis [14]. Simulation results and their analysis allow scientists to build deep intuitions about how different aspects of the digital ecosystem interact one with another, reveal important dependencies, and provide understanding of how evolution interacts with ongoing dynamics of the ecosystem [14].

In the study [14] Terry Jones reveals dynamics of system that is common for ecology systems. The conditions, under which distributions of this kind are seen, include early successional communities, environments perturbed by toxins or pollutants, and in appropriately sized samples [14]. In his study [13] by analyzing count of species in observed data, Preston showed that abundance of species in such areas has lognormal distributions. In his Echo model, he studied agents' species clustering based on genetic distance, stressed species abundance notion and showed that model exposes similar species abundance distribution characteristic to Peterson's lognormal distribution [13].

Continuing working with Echo model family P. Hraber and B. Milne [16] discovered the notion of the emergence of community assemblies. They showed the existence of agent groups that share common behavior that springs in order to response on interaction rules in model architecture. Certain genotype assemblies (complementary genotypes) were born and formed quasi-stable domination that was based on pairwise interaction between agents. In given work we consider digital ecosystem with such emergent feature and show that changing of system property such as space heterogeneity contributes to sustainability of complementary phenotypic assemblies over time.

The further development of digital ecosystem models are models where complex agent's behavior arises from the first principles: it never was predefined by fitness function and emerges through adaptation process. Remarkable examples of such models are Michael Burtsev's [7], [17] model and Robert Grass' [15] model. Further, Michael Burtsev proposed a model that resembles pioneer Artificial Life's Echo [14] and Bugs [20] models: the agents with simple behavior are acting in a simple space. In the study [7] the author develops latter models introducing kinship (by introducing culture affinity) and using the artificial neural network as a basis for agent's actions. In this model no agents have a predefined strategy, instead it

emerges as phenotype feature from agent's actions, defined by the neural network. Some of the strategies expose cooperative behavior, where agents adjust their behavior due to genotypic distance between each other. It was shown that in such model the strategies emerge that correspond to those in well-known game theory - dove-hawk-bourgeois, where dove acts like peaceful harvester, hawk demonstrates aggressive behavior attacking agents in neighborhood, and bourgeois that plays as dove when low on resource and displays hawk strategy in possession of it. Also, two new strategies of cooperative attack (when agents attack only non-relative ones) and defense (when agents gather in one location to defend themselves from aggression) emerged [9]. The similar results with different models are achieved by a research with novel artificial life model with predator-prey behavior in [15], where agents are driven by fuzzy cognitive map. Considering results of artificial life modeling it can be concluded that such approach is not being controversial to game theory but on the contrary is an extension that provides new research horizons, such as finding evolutionary stable strategy, designing an open-ended evolution, exploring new sophisticated agent's behavior, and analyzing system regularities, e. g. persistent emergence of the group behavior and arm races. Correlation between population density and frequency of fight action for the case of rich resources in the model is similar to the analogous correlation extracted from the ethnographic database [7]. By studying the model, Burtsev proposed a novel methodology to categorize agents' behavior into strategies and to trace population genotype dynamic [14].

Analysis of mentioned above researches of Alife models show that they open novel regularities and emergent behavior. Following study of the similar models discovers new aspects of agents' behavior dynamic. Evolution processes in the models of digital ecosystems are far from being clear and traceable, the interconnection between emergent features and system parameters are not yet properly established. In this work, we explore the correspondence of the phenotype sustainability to the heterogeneous or homogeneous resource distribution in space and discuss the dependency of aggressive and peaceful behavior on the amount of income resource. In the following part of this article we provide model description and its rules. We discuss the dynamics of aggressive and peaceful behavior in the next section noting some interesting regularities and displaying the dependency between the behavior and resource available in the system. The next section is dedicated to the description of experiments in the space where resource is heterogeneously distributed, that is, predator-prey, and strategy competition cases. Next, we demonstrate how variability of strategies changes on the type of resource space, concluding the article with discussions and future studies possibilities.

II. Model description

Here we describe as the important example the agent-based lattice foraging model with possible predator-prey behavior which is a development of classical artificial life models [12], [14], [23]. This model follows mainly the Burtsev's cellular automata approach [9], because of the similarities in the neural network type, environment rules, and culture affinity between agents. The major differences between his and ours model are agent's perception and agent's world arrangement. In Burtsev's model arbitrary number of agents can occupy single cell unlike ours model, where only one agent can live in the one cell. Perception in Burtsev's model is based on averaging of the parameters – the agent is aware only of mean attributes (culture affinity) of the whole agents' group in his cell. This follows to more generalized interactions. In our model, agent's perception is significantly different: each agent is aware of each neighbor and his culture affinity. Such architecture provides more individualized interactions and perception and is more in line with classical Alife digital ecosystem models [15], [16], and [23].

Each agent is characterized by culture affinity: 3-dimentional vector. Its coordinates can take possible integer values in [-2, 2] interval. Agents are treated as relatives if Euclidean distance between their culture affinity vectors are less than 0.2 threshold. This vector is also inherited by offspring from his parent with some mutations.

An agent occupies one grid cell in cellular space. He is driven by heading vector that defines a cell in front, where interaction may exist. The agent can perceive other agents in von Neuman neighborhood. Each neighbour cell refers to 2 neural network inputs, one corresponds to relative (whose culture vector similar enough in terms of Euclidian distance) and other to corresponds non-relative agent (except only one input for back cell that tracks non-relative agent). Therefore, the agent can differentiate whether a neighbor is relative or non-relative. Agent's accumulates energy. The agent gains resources by consuming resources of other agents.

Patches that grow at each iteration present the resource. The number of patches to grow is defined at startup and is fixed for the experiment. In the case of homogeneous resource distribution, resource appears in any cell with equal probability and the value of this resource is uniformly distributed in [0,500] interval. New resource is appended to the old one remaining in cell. We can track the average resource input count for each iteration (time step) and the average value of resource distributed per cell.

Each simulation run starts with 10 agents placed in the environment and equipped with 2000 points of resource. Agents have artificial neural network that has small predefined weights. They invoke the action eat when resource is in front cell and suppress move action when agent in in front cell (this action ends with except of losing resource). At first iteration, 1000 of resource patches is randomly distributed in environment.

The experiments were performed with various resource income rates (amount of resource patches that appears on each iteration), various parameter values, and heterogeneous resource landscape to display aggressive and peaceful behavior in system. The following set of experiments was performed on heterogeneous resource landscape to demonstrate predator-prey cycle and competition of strategies. Both homogeneous and heterogeneous resource landscape cases (for different resource income rates) were taken for the next set of experiments to display resource landscape influence on the phenotypic i. e. strategy sustainability.

III. Aggressive and peaceful behavior

Current model is a plausible background for studies of aggressive and peaceful behavior as depending on the number

of input resources. Also, the model displays different types of interaction between groups of related agents. The simulation results for different resource input rate are presented in this section.

In case of low amount of resources one of the most effective strategies is peaceful strategy. Agents either do not distinguish between relative and non-relative agents and prefer to stay at rest, or run away from the relatives in order to avoid competition for the resources, or escape the strangers feeling threatened, and other variations of these strategies are presented as well.

Number of cooperative strategies for this resource mode slightly surpass number of non-cooperative strategies (those that do not take culture affinity differences into account). Under the cooperative notion, we consider strategies that distinguish relative and non-relative agents using culture affinity and adjust their behavior to benefit from this. For example, when agents leave area filled with relatives, they reduce competition for local resources, or when agents attack only non-relative agents and make no harm to relatives.

Peaceful strategies dominate at this type of resource income. But this long-term pattern of peaceful strategies domination, can be seen rarely for this regime. Such simulation conditions commonly are not able to provide a peaceful strategy that survives after the first manifestations of aggressive behavior. So the peaceful phases change suddenly to much more volatile aggressive strategies.

For smaller amounts of resources, peaceful strategies play an important role. With an increase of volume of resources, almost all strategies show aggression. Peaceful strategies can no longer exist for significant periods of time as in the previous case. Almost all strategies exhibit aggressive behavior and are very volatile. Calculations display population dynamics with permanent aggressive behavior mentioned in the case of a small number of resources.

For a large number of resource inputs, agents have the ability to completely fill the grid space and competition between strategies becomes sluggish. Thus with a large volume of resources all cells are filled with aggressive and peaceful agents and peace strategies are stable and rarely changing each other.

IV. HETEROGENEOUS RESOURCE SPACE

The next step in the study of the behavior of the model was to investigate the strategy dynamics in the case of the heterogeneous resources landscape.

In this paper, among the whole set of computer experiments that were shown, several deserve the most attention. They illustrate the new modes of interaction between agents that adhere to cooperative behavior. Experiments with heterogeneous space show the modes with competition between strategies. In computer experiments most strategies that engaged in competition were mostly cooperative. We assume as antagonistic the peaceful (absence of the "attack" strategies in the vector) and aggressive (if there is at least one action "attack" strategies in the vector) strategies. System dynamic revealed that behavior strategies with similar behavior (peaceful or aggressive) compete with each other as well as strategies with antagonistic behavior (peaceful versus aggressive). It was found the competition between both antagonistic and between similar behavior strategies.

V. SUSTAINABILITY OF THE STRATEGIES

Considering strategies dynamic from the previous chapters, it can be noted that strategies dynamics from previous section are too volatile – lack sustainability. By encouraging agents' phenotypic assemblies, localization in space through heterogeneous resource landscape, we suppress strategies volatility. It can be useful in the search of evolutionarily stable strategy - the strategy that cannot be invaded by any other strategy in arbitrarily small amounts [4].

It should be noted that strategy does not exhibit evolutionary stability as it is familiar for Game Theory models [19]: a strategy that cannot be invaded by any other strategy present in arbitrarily small amounts.

The alternative classification of agent assemblies in research of population divergence [6] claims that agent population assembly belong to the space of stable instability if under the same conditions different experiments can give different population structure and it is not inevitable that this assembly will be presented in population. All (phenotypic) groups of agents in the given model experiments belong to the set of I groups because of great variability and instability of strategies. These populations are called unstable. In contrast, it is defined as "waist" population, the one that constricts to a single assembly [6]. In the model considered in this study, 'waist' behavior can be exposed by cooperative peaceful or cooperative aggressive phenotypic assemblies when the resources income is high. In this case the cooperative behavior that appeared first is fixed further for entire population. However, tracing strategies evolution is a complicated task due to many reasons such as complex model architecture, and computational performance restrictions.

VI. DISCUSSIONS

In this study, we considered phenotypic assemblies (grouping agents by strategies), their shared behavior and its sustainability under different resource environmental conditions: resource income rate and resource landscape distribution. One of the goals of investigations was to increase phenotypic assembly sustainability during time.

The given model architecture is connected mainly to Burtsev's cellular automata approach [9], but bears significant differences which provide more individualized interactions and make the model more familia to classical Alife digital ecosystem models [15], [16], and [23]. We have illustrated how aggressive and peaceful behavior is dependent from the volume of input resources. Such behavior correlates with resulting behavior in model [9]. This allows to speak about continuation of the model development without losing its emergent features. The goal of new experiments was to make strategies less volatile. And the simulation results for heterogeneous cellular space showed decreasing variability of strategies in such case and provided an opportunity to illustrate the emergence and development of strategies competition.

In experiments with heterogeneous space, we have identified modes of competitive agents' interaction with the complementary and similar behavior. So, paradoxically more diversity of conditions follows to the unification of behavior.

Other very interesting conclusion (or confirmation of intuition) is that the aggressive behavior is more inherent to population when the lack of resources is observed. The examples of illustrated strategy competitions state the question about the nature of phenotypic transition between strategies having both complementary (predator-prey) and antagonistic (predator-predator) interactions. It is to be clarified whether the transition occurs though combat or by peaceful genotype transition. The enhancements on model architecture and analysis could shed light on this question. In spite of powerful and demonstrative strategy analysis methodology, the given model demands precise tracking of agent's culture grouping and genotype dynamic. Culture group members can have various phenotypic (strategy) features. The emphasize on combat actions between agents should be done.

Agent's neural architecture is the very important factor for such kind of Artificial Life models. A simple artificial neural network with no hidden layer can achieve the scope of predator-prey behavior considered in this study as it is implemented in the model. More complex behavior such as group hunting and wandering could be simulated using more sophisticated methods of neuroevolution such as, for example, NEAT (neuroevolution of augmented topologies) [10]. Authors of [10] use the NEAT algorithm to investigate the evolution of effective predator group or group of collective foragers. The crucial advantage of NEAT for multiagent modeling is natural origin of agents' grouping by genotype affinity. This could provide useful insights on agent behavior emergence. Alife models could benefit from NEAT usage by inheriting its methodology of tracking genes evolution through historical markings. It is effective and sophisticated resolution of the speciation tracking problem.

As the possible development of given model and of other models of artificial life, we can pose the following problems: overcoming the great computational complexity of the experiments; improving accounting of interaction between the agent and the environment; replacing the discrete space type with continuous; introduction of new types of interaction between agents.

Software enhancement of models with high computation performance would give the possibility to observe long-term trends that provide valuable efforts for understanding of such type models. The application cross-over to high-performance computation environment is the pending task for many Artificial Life models [9], [15].

The important issue is the introduction and usage of novel analysis method for agent-based complex adaptive systems. For example, Burtsev proposed a promising methodology that considers evolving agents' population as a dynamic system in [8]. The open question is the study of the competition of groups and establishing the intensity of the impact of various factors, such as aggression and phenotype transition strategies, on the success in the competition. In this paper we formally speak in digital ecology terms. But it is possible to extend the models to other fields. One of them is investigation of real society by artificial society models. Other recent applications are the searching principles for arranging evolving teams of robots.

JIITEPATYPA REFERENCES

[1] Adamatskii A., Komosinski M. Artificial Life Models in Software. Berlin, Springer-Verlag. 2005.

[2] Adami C. Introduction to Artificial Life. Berlin, Springer. 2005.

[3] Aguilar W., Santamaria-Bonfil G., Froese T., Gershenson C. The Past, Present, and Future of Artificial Life, *Frontiers in Robotics and AI,* 10 October 2014, http: //dx.doi.org/10.3389/frobt.2014.00008.

[4] Aoki K, Feldman M. W. Evolution of learning strategies in temporally and spatially variable environments: A review of theory. *Theoretical population biology*, 2014. doi:10.1016/j.tpb.2013.10.004.

[5] Bedau M. A., McCaskill J. S., Packard N. H., Rasmussen, S., Adami C., Green .G. etal. Open problems in artificial life. *Artif. Life* 2000. 6*,* pp 363–376, doi:10.1162/106454600300103683, doi:10.3389/frobt.2014.00008.

[6] Bulgakova T. I., Kalugin O. S., Lyapunov A. A. On the stability of the simplest model biogeocoenose *Problems of Cybernetics*. Nauka, Moscow, 1968. pp 272-276.

[7] Burtsev M., Turchin P. Evolution of Cooperative Strategies From First Principles. *Nature*, 2006. vol. 440, pp 1041-1044.

[8] Burtsev M.S. Tracking the Trajectories of Evolution. *Artificial Life*. 2004. V. 10. 4, pp 397-411.

[9] Burtsev M.S. Artificial Life Meets Anthropology: A Case of Aggression in Primitive Societies. In: M. Capcarrere et al. (Eds.): *ECAL, LNAI,* 2005. pp 655 – 664.

[10] D'Ambrosio D. B., Stanley K. O. Generative encoding for multiagent learning. In *Proceedings of the Genetic and Evolutionary Computation Conference. New York: ACM* Press. 2008. doi: 10.1145/1389095.1389256.

[11] Dorin A. Biological Bits. A Brief Guide to the Ideas and Artefacts of Computational Artificial Life. Animaland, Melbourne. 2014.

[12] Epstein Joshua M., Robert Axtell Growing artificial societies: social science from the bottom up. Brookings Institution Press. 1996. Doi:10.2307/1060800.

[13] Preston F. W. The commonness, and rarity, of species. *Ecology*: *Ecological Society of America*, 1948. pp 185-215.

[14] Forrest S., Jones T. Modeling complex adaptive systems with Echo. *Complex Systems: Mechanisms of Adaptation*. IOS Press, Amsterdam, 1994. pp 3-21.

[15] Gras R, Golestani A, Hendry A. P., Cristescu M. E. Speciation without Pre-Defined Fitness Functions. *PLoS ONE,* 2015. 10(9): e0137838, doi:10.1371/journal.pone.0137838.

[16] Hraber P. T., Milne B. T. Community Assembly in a Model Ecosystem. *Ecological Modeling*, 1997. pp 267-285.

[17] Krivenko S., Burtsev M. Simulation of the evolution of aging: effects of aggression and kin-recognition. In: *Advances in Artificial Life, 9th European Conference, ECAL, Lecture Notes in Computer Science*, 2007. pp 84–92.

[18] Langton C.G. Artificial Life: An Overview. Cambridge, MA: MIT Press. 1997.

[19] Lindgren K., Nordahl N. G. Cooperation and Community structure in artificial ecosystems. *Artificial Life*, 1994. 1:15–37.

[20] Packard N., Langton C.G. Intrinsic adaptation in a simple model for evolution. *Artificial life*, Redwood City, Addison-Wesley, CA, 1989. pp 141-155.

[21] Penn A. Artificial Life and Society: Philosophies and Tools for Experiencing, Interacting and Managing Real World Complex Adaptive Systems. *Proc. Int. Conf. Acai – 2016,* Cancun, Mexico. 2016. Pp. 26-27.

[22] Taylor T., Auerbach J. E., Bongard J. teal. A Review of the First 21 Year of Artificial Life on the Web. *Artificial Life* (USA), MIT, Massacussets. 2016. Vol. 22. N.3, pp. 364-407.

[23] Yaeger L. Computational Genetics, Physiology, Learning, Vision, and Behavior or PolyWord: Life in a New Context. In Langton, C. G. (ed.): *Artificial Life III*. Addison-Wesley, 1994. pp 263-298.

[24] Wikipedia. Artificial Life. https://en.wikipedia.org/wiki/Artificial_Life.

Методика Вибору Оптимальної Геометричної Форми Порожнини в Зубі для його Лікування Ендокоронкою

В.Ф. Макєєв, Р.Р. Павличко
Кафедра ортопедичної стоматології
ЛНМУ імені Данила Галицького
Львів, Україна
ort_stom@meduniv.lviv.ua
romapavlychko@gmail.com

І.М. Заячук
Центр математичного моделювання
ІППММ ім. Я.С. Підстригача НАН України
Львів, Україна
igorzaj@mail.ru

Method of Choosing the Optimal Geometric Shape of Cavity in the Tooth for its Treatment by Endocrown

V.F. Makeyev, R.R. Pavlychko
Department of Prosthetic Dentistry
Danylo Halytskyy Lviv National Medical University
Lviv, Ukraine
ort_stom@meduniv.lviv.ua
romapavlychko@gmail.com

I. Zajachuk
Centre of Mathematical Modelling
Pidstryhach Institute for Applied Problems of Mechanics and Mathematics Ukrainian Academy of Science
Lviv, Ukraine
igorzaj@mail.ru

***Анотація*—Запропонована методика і обгрунтована ефективність її використання для формування порожнини зуба з подальшим його лікуванням ендокоронкою. Вибір геометричної форми порожнини зуба рекомендовано проводити шляхом математичного моделювання. В процесі побудови і реалізації марема-тичної моделі доцільно скористатися методом скінчен-них елементів і основними поняттями теорії пружності. Вказано на ефективність використання індивідуально розробленого програмного забезпечення для реалізації запропонованої методики.**

Abstract—Method is proposed and proved efficiency of its use for shaping tooth cavity with further treatment by endocrown. Choosing geometric shape of the tooth cavity is recommended to do by mathematical modeling. In the process of construction and implementation of mathe-matical models appropriate to use the finite element method and the basic concepts of theory of elasticity. Pointed on efficiency of using individual designed software for implementing the proposed method.

Ключові слова—математичне моделювання; ліку-вання ендокоронкою; метод скінченних елементів; програмне забезпечення; теорія пружності; числові методи.

Keywords—mathematical modelling; treatment by endocrown; method of finite elements; software.

I. Вступ

Довготерміновість експлуатаційної здатності ендодонтично лікованих зубів покритих ендокоронками залежить не тільки від виду матеріалу, з якого виготовлені такі реставрації та якості їх виготовлення ,а і силових напружень, що виникають у самому зубі під час його функціонально-оклюзійних навантажень.

Тому в останні роки досить широкого розповсюдження набула методика визначення таких напружень як у самому зубі, так і у відновних реставраціях методом скінченно-елементарного аналізу 3D моделі зуба.

Питання розподілу внутрішніх напружень у матеріалах коронки ти тканинах зуба обговорювалось у літературі вже давно, оскільки знання величин та напрямків напружень можуть допомогти зменшити пошкодження ендодонтично лікованих зубів та ендокоронок в умовах оклюзійної патології, а сам факт виникнення напружень в зубах під дією навантаження дискутується у науковій літературі протягом 40 років.

Дослідження біомеханічних систем доцільно проводити шляхом математичного моделювання [1]. Підтвердженням цього є висновки, зроблені авторами робіт [2,3,4,5]. Для побудови математичних моделей в ортопедичній стоматології доцільно використовувати метод скінченних елементів [6,7]. Правильність вибору такого методу можна аргументувати наступними твердженнями:

- перевага методу скінченних елементів в тому, що його можна використовувати не тільки для облас-тей з гладкою формою границі. Криволінійна об-ласть може бути апроксимована за допомогою прямолінійних відрізків;
- розмір елементів може мінятися, що дозволяє регулювати величину сітки розбиття області;
- метод скінченних елементів дає можливість роз-в'язувати задачі для пошуку поверхневого навантаження з точкою розриву, а також зі змішаними граничними умовами;
- властивості матеріалів суміжних елементів можуть бути різними. Отже, метод скінченних елементів можна застосовувати у випадку, якщо тіло сформоване з декількох складників.

Ефективність методики вибору оптимальної геометричної форми порожнини в зубі і застосування цієї методики для лікування хворих залежить від програмного забезпечення, яке використовується в процесі обчислювального експерименту. Існують універсальні програмні засоби для розрахунку за допомогою методу скінченних елементів [8, 9, 10] і індивідуально розроблені.

До універсальних можна віднести пакет прикладних программ COSMOS/M [18] за допомогою якого автоматично здійснюється розрахунок напружень і диформацій для заданої моделі. Вхідними даними для розрахунку за допомогою цього пакету є властивості матеріалів з яких складається модельований об'єкт (модуль Юнга, густина матеріалу і коефіцієнт Пуассона).

Індивідуально розроблене програмне забезпечення враховує конкретні умови і особливості лікування, які в деяких випадках не передбачені універсальними методами, а також легко модифікується за потребою лікаря.

Ризик руйнування та розшарування в зубі, реставрованому ендокоронкою та штифтовою конструктцією, досліджували в праці [19]. Під час розвитку руйнівних процесів у конструкції "зуб-ендокоронка" важливо першочергово проаналізувати ризик руйнування зуба. Адже якщо злам кореня відбувається до порушення зчеплення між елементами ендокоронки та зуба, то такий зуб потрібно вилучати і ставити імплантат. З іншого боку, якщо спочатку відбулось порушення зчеплення, то зуб можна реставрувати (відновити), видаляти його не потрібно. У цьому дослідженні не визначають порядок настання цих подій, а оцінюють ризик їх розвитку окремо.

Що стосується ризику зламу дентину, то максимальні напруження в дентині у штифтовій моделі є вищими, ніж в ендокоронці. Припускають, що це пов'язано з жорст-кістю матеріалу, яким заповнено кореневий канал: штифт зі скловолокна є більш жорстким, ніж м'який матеріал (гутаперча), яким заповнений кореневий канал в моделі з ендокоронкою. Цим пояснюється і той факт, що максимальні напруження в штифтовій моделі розташовуються нижче, ніж у моделі з ендокоронкою. Отже, модель з ендокоронкою є кращою щодо ризику зламу дентину.

Для визначення ризику порушення зчеплення розглядали контактні напруження у ділянках сполучення коронка-емаль, коронка-дентин, коронка-кореневий канал. Поверхня контакту між сполучним цементом і кореневим каналом більша в штифтовій моделі, ніж в моделі з ендокоронкою. Тому нормальні напруження у ділянці коронка-дентин вищі в моделі з ендокронкою для всіх рівнів висоти залишкової емалі. Проте ризик порушення зчеплення є і в інших зонах, де сполучний цемент кріпиться до композитної коронки (коронка-емаль і коронка-кореневий канал). Якщо залишкової емалі є багато, то ризик розшарування в моделі з ендокоронкою вищий, а для малої висоти емалі вищий ризик розшарування в штифтовій моделі.

Отже із наведеного вище бачимо, що у світовій стоматологічній практиці значну увагу приділено вивченню процесів руйнування і міцнісних характеристик матеріалів ендокоронок та відповідних типів конструкцій "зуб-ендокоронка".

II. МАТЕМАТИЧНА МОДЕЛЬ

В найпростіших тривимірних моделях, розроблених на початку 90-х років для біомеханічних систем, враховувалися основні закономірності поведінки зуба, як складової частини жувального аппарату. Структура будови самого зуба не приймалася до уваги.

Ускладнення математичної моделі з врахуванням сказанного здійснювалося шляхом використання математичного аппарату, зокрема фундаментальних понять з теорії пружності і числових методів. Для моделювання процессу дослідження напружень в порожнині зуба проектувалися детальні сітки скінченних елементів, які враховували як шаруватість будови зуба так і складні умови його закріплення [11].

Досліджувану систему в такому випадку доцільно представити у вигляді тримірної зв'язної області Ω з кусково гладкою границею $\Gamma \cup \Gamma$ [12].

Результати досліджень можна отримати шляхом розв'язку крайоваї задачі з теорії пружності. Тобто необхідно знайти вектор переміщень $= , , \in \Omega$, компоненти тензорів напружень і деформацій , такі, що задовольняють умовам рівноваги $/ = 0,$ закону Гука $=$,умовам навантаження $\cos(,) =$ і умовам закріплення $= 0$. Такі умови навантаження на зуб моделюють нормальний до поверхні тиск. Умови закріплення не враховують переміщень точок нижнього зрізу досліджуваної геометричної області і переміщень точок бокових зрізів у напрямі нормалі до їхньої площини.

Варіаційний аналог поставленої задачі включає метод скінченних елементів. Такий підхід повинен враховувати специфіку в лікуванні зуба, в даному випадку, ендокоронкою. Умови лікування необхідно враховувати в алгоритмі побудови сітки скінченних елементів для забезпечення оптимальної геометричної форми порожнини в зубі.

Аналіз результатів досліджень, проведених різними авторами за допомогою методик, основою яких є математичне моделювання, підтверджує ефективність такого підходу.

Зокрема, в статті [13] було доведено, що зруйновані корені фронтальних зубів нижче рівня ясен на ¼ довжини кореня можуть бути використані в клінічній стоматології. Рекомендовані штифтові конструкції шляхом математичного моделювання були вдосконалені, що підвищило їх надійність.

В роботі [14] подано результати експериментів з моделлю штифтової конструкції для фронтальних зубів з діаметром штифта 2 мм. Для встановлення такого штифта потрібно розширити канал зуба. Тоді така штифтова конструкція призведе до ослаблення міцності зуба і пришвидшить процесс його руйнування.

За результатами проведених досліджень було зроблено висновки і надано рекомендації що до ефективного лікування.

Авторами статті [15] проведено комп'ютерне дослідження 99 типів математичних моделей зубів, армованих культьовими штифтовими вставками. Моделі зубів відрізнялися формою, розмірами і властивостями конструкційного матеріалу. За результатами досліджень було показано, що форма і параметри культьової штифтової вставки є важливішими за властивості конструкційного матеріалу. Показано, що вибір штифтової конструкції залежить від клінічної ситуації: анатомо-топографічних особливостей зруйнованої коронки зуба і патології зубощелепно-лицевої системи.

В статті [16] описана методика математичного моделювання процесів взаємодії імплантів із кістковою тканиною. Запропонована модель і программа для розрахунку. Основою алгоритму програми є метод скінченних елементів. За допомогою цієї програми проведено аналіз напружено-деформованого стану кістки і оцінено гранич-но допустимі навантаження на імплант. За результатами досліджень встановлено, що макси-мально навантаженою зоною є шар компактної кістки, який безпосередньо контактує з імплантом в області його шийки. Такий висновок співпадає з результатами клінічних досліджень, згідно з якими в цій зоні достатньо високий ризик ускладнень. Також в [16] запропоновано шляхи і методи для оптимізації імплантів і протезних конструкцій.

Модель зуб-імплант пародонт є складною геометричною і фізичною системою. Розрахунок такої системи доцільно проводити числовими методами. Метод скінченних елементів найбільш прийнятний для таких розрахунків. Адже він дозволяє враховувати реальні особливості структури біомеханічного об'єкту і адаптується до зміни форми і зовнішніх навантажень.

Для моделювання напружено деформованого стану щелепної кістки з врахуванням довільних варіантів дентальних імплантних систем автори роботи [17] запропонували методику, яка базується на просторовій теорії пружності, методі скінченних елементів, цифровій обробці зображень, отриманих за допомогою комп'ютерного томографа. Результати досліджень за допомогою такої методики подані авторами у вигляді атласу з ілюстраціями.

В статті [18] автори поділилися результатами дослідженя в кістковій тканині в околі опори дентальних імплантів під дією статичного вертикального навантаження.

Об'єктом моделювання було вибрано мостовидні протези, закріплені на імплантах в боковому відділі нижньої щелепи. Для розрахунків були побудовані двомірні моделі різної конструкції і використаний метод скінченних елементів. Досліджено, що напруження в к4істковій тканині в околі імплантів спадають у випадку зменшкння віддалі між крайніми опорами і збільшенні кількості опорних елементів (за умови розміщення опори мостового протезу вздовж прямої лінії). У випадку моделювання консольних протезів спос-терігалось зростання напружень в порівнянні з тради-ційними мостовиднимими протезами. Причому значний ріст напружень у кістковій тканині в околі опорних елементів спостерігався у випадку зменшення відстані між крайніми опорами і видовженням консольного плеча.

Автори праці [18] шляхом математичного моделювання і подальшого числового розрахунку вказали на незначну ефективність процессу збільшення загальної кількості імплантних опор для досягнення мінімальних напружень в кістковій тканині.

Тривимірний аналіз методом скінченних елементів доцільно проводити для оцінки рівня напружень у кістці, імплантаті і його головці у випадку застосування золотих сплавів, керамічних або полімерних (пластмасових і композитних) протезів, що не знімаються [20]. В процесі розрахунку зростаюче навантаження було прикладене вздовж осі і у щоко-язиковому напрямку по центру мостовидного протезу. Для золотих сплавів і кераміки максимальні еквівалентні напруження спосте-рігалися на всіх частинах моделі. Практично для всіх випадків у моделях з пласт-масовими протезами напруження були одинакові або вищі за ті напруження, які спостерігалася у дослідженнях з двома попередніми матеріалами. Високий рівень напружень спостепігався у головках імплантантів у випадку осьового навантаження пластмаси. Захисна роль пластмаси для системи "імплантат-кістка" за умов даного дослідження не була підтверджена.

III. ВИСНОВКИ

Таким чином для формування ефективної методики лікування зуба ендокоронкою запропоновано застосовувати метематичне моделювання в процесі розробки потрібної методики.

Проведено огляд літературних джерел в яких вже представлено подібні задачі та їх вирішення з використанням засадничих понять з теорії пружності та числових методів.

Удосконалення та розробка даного напрямку дозволять підібрати оптимальну форму зуба і відповідно знизити кількість ускладнення внаслідок сколу та зламу твердих тканин зуба а також подовжити термін функціонування структури зуб-ендокоронка.

ЛІТЕРАТУРА REFERENCES:

[1] Чуйко А. Н., Бережная Е.О., Бахуринский Н.Ю. О современных воможностяхбиомеханического анализа в стоматологии. // Стоматолог. 2001. № 1-2. С.36_41. Н. Н. Мельников. – М., Химия, 1987. – 712 с.

[2] Воробьев В.А. Выбор конструкции зубных протезов при лечении больных с различными дефектами и имплантантных систем на основе программного математического моделирования зубных рядов. Дис. д-ра мед. Наук.Иркутск 1996.

[3] Маркин В. А. Прогнозирование осложнений при использовании металлокерамических протезов с помощью математического моделирования. Дис. канд. мед. наук. М., 1999.

[4] Олесова В.Н., Кленилин Е.С., Балгурина О.С. и др. Биомеханическое и клиническое обоснование штифтовых конструкций на основе стекловолокна. Паорама ортопедической стоматологии. 2001. № 2. С.4-5.

[5] Погожев И.Б. Применение математических моделей заболеваний в клинической практике. В кн. Вычислительные процессы и системы. Вып. 3. Под ред. Г.И. Марчука. М. 1985. С 108-165.

[6] Чуйко А.Н. О возможностях конечно- элементного моделирования в ортопедической стоматологии. Стоматолог. 2000. №3. С 37-38.

[7] Чуйко А.Н. Конечно элементный анализ в стоматологии –http: fem.bi.com..ua.index shtm.

[8] Канотина Т.Н., Семенюк В.М., Яовлев К.К., Гуц А.К., Панова Н.И. Матматическое обоснование к использованию культевой штифтовой вкладки с ”воротничком” при разрушении корней зубов ниже уровня десны. Вестник Омског ун-та. 1996. №2. С.17-19.

[9] Семенюк В.М., Гуц А.К., Канотина Т.Н., Вагнер В.Д., Яковлев К.К. Отдаленные результаты использования корней фронтальных зубов , разрушенных ниже уровня десны, в качестве опоры для штифтовых конструкций. Пути развития стоматологии:итоги и перспективы. Материалы конференции 25-27 апреля 1995. Екатеринбург, 1995. С. 204-206.

[10] Семенюк В.М.,Гуц А.К., Капотина Т.Н.,Вагнер В.Д. Выносливость опорно-удерживающего аппарата зуба, востановленного литой культевой штифтовой конструкцией, к вертикальной нагрузке. Перспективы развития современной стоматологии: проблемы уральского региона. Материалы конференции стоматологов 20-22 мая 1997. Екатеринбург: Уральская гос. Мед. Академия. 1997. С.211-213.

[11] Мандзюк Т., Вовк В. Огляд проблем комп’ютерного моделювання біомеханічних систем у стоматології. Вісник Львівського університету. Сер. прикл. матем. та інформ. 2008. Вип. 14. С.105-122.

[12] Мандзюк Т., Вовк В. До моделювання біомеханічних конструкцій з м'якими прошарками. Вісник Львівського університету. Сер. прикл. матем. та інформ. 2011. Вип. 17. С.85-93.

[13] Гуц А.К., Капотина Т.Н., Панова Н.И., Семенюк В.М., Файзулин Р.Т., Яковлев К.К. Математическое обоснование к использованию корней фронтальных зубов, разрушенных ниже уровня десны, под штифтовые конструкции. Деп. В ВИНИТИ 21.06.95. № 1790-И95. 22с.

[14] Семенюк В.М., Гуц А.К., Панова Н.И. Компютерные эксперименты с моделью штифтовой конструкции с большим диаметром штифта. Математические структуры и моделирование. 2002. Вып. 9

[15] Арутюнов С.Д., Чумаченко Е.Н., Лебеденко Н.Ю., Арутюнов А.С. Сравнительный анализ результатов математического моделирования напряженно-деформированного состояния различных конструкций штифтовых зубных протезов. Стоматология. 2001. № 2. С 11-16.

[16] Матвеева А.И., Иванов А.Г., Гветадзе Р.Ш., Гаврюшин С.С., Карасев А.В. Повышение эффективности ортопедического лечения больных на основе математического моделирования перспективных конструкций имплантов. Стоматология. 1997. № 5. С. 14-18.

[17] Атлас напряженных состояний дентальных имплантантных систем В.А. Воробьев, Т.М. Ратинер, А.С. Рубинов, В.В. Дудкин: под ред. М.З. Миргазизова. Ангарск.1997. 58с.

[18] Arataki T., Adaciri Y., Kishi M. Two-dimensional finiteclement analysis of the influence of bridge design on stress distribution in bone lissucs surroandiny fixlarcs of osscomtegrated implants in the lower molar region. Bull.Tokyo. Dent. Coll. 1998. V. 39. № 3. P. 199-209.

[19] Riera M. Study of the risk of fracture and debonding in a tooth restored by a post and a composite core and another by an endocrown methodology using FEA / Miquel Riera. – Universite Catholique de Louvain, 2014. – 93 p.

[20] Stegaroiu R, Kusakari H, Nishiyama S, Miyakawa O. Influence of prosthesis material on stress distribution in bone and implant: 3-dimensional finite element analysis // Int. J. Oral Maxillofac Implants. 1998. V.13, № 6. P. 781-790.

Модифікований Метод Граничних Елементів та Алгоритм Розв'язання Задач Індукованого Магнітного Поля Тонкостінних Конструкцій в Тривимірному Просторі

Володимир Михайлович Рябенький
кафедра теоретичної електротехніки та електронних систем
Національний Університет Кораблебудування ім. адмірала Макарова
Україна, м. Миколаїв
optron2@gmail.com

Ігор Іванович Чудайкін
кафедра теоретичної електротехніки та електронних систем
Національний Університет Кораблебудування ім. адмірала Макарова
Україна, м. Миколаїв
igorlu3141509@rambler.ru

Юлія Дмитрівна Таргунакова
кафедра теоретичної електротехніки та електронних систем
Національний Університет Кораблебудування ім. адмірала Макарова
Україна, м. Миколаїв
golubkajulya@gmail.com

Modified Boundary Elements Method and Algorithm for Solving Problems of the Thin Structures Induced Fields in Three Dimensional Space.

Volodymyr Myhalovych Ryabenkiy
Theoretical Electronic and Electronic Systems Department
Admiral Makarov National University of the Shipbuilding
Ukraine, Mykolaiv
optron2@gmail.com

Igor Ivanovych Chudaykin
Theoretical Electronic and Electronic Systems Department
Admiral Makarov National University of the Shipbuilding
Ukraine, Mykolaiv
igorlu3141509@rambler.ru

Julia Dmytrivna Targunakova
Theoretical Electronic and Electronic Systems Department
Admiral Makarov National University of the Shipbuilding
Ukraine, Mykolaiv
golubkajulya@gmail.com

***Анотація*—У статті описується розроблений модифікований метод граничних елементів та алгоритм для розрахунку індукованого магнітного поля тонкостінних конструкцій довільної форми в тривимірному просторі. Розроблений метод, на відміну від класичного методу граничних елементів, не вимагає наявності матриці фундаментальних рішень. Представлений модифікований метод, позбавлений основних недоліків методу граничних елементів і методу скінченних елементів. Наведені приклади розрахунку індукованого магнітного поля тонкостінних конструкцій.**

***Abstract*—The article describes developed Modified Boundary Elements Method (MBEM) for calculating the induced magnetic field of arbitrary shape thin-walled structures in a three-dimensional space. Presented MBEM is devoid of the Boundary Elements Method and the Finite Elements Method main drawbacks. Examples of the induced magnetic field calculating of the thin-walled structures are given.**

Ключові слова—магнітне поле; модифікований метод граничних елементів; алгоритми розрахунку індукованого магнітного поля тонкостінних конструкцій

Keywords——magnetic field; modified boundary elements method; algorithm for calculating the induced magnetic field

I. Вступ

У ряді важливих практичних застосувань потрібна інформація про розподіл магнітного поля в околиці феромагнітних тонкостінних конструкцій складної просторової конфігурації, які знаходяться в зовнішньому постійному магнітному полі (наприклад, в геомагнітному полі — ГМП). До таких додатків відносяться, зокрема, визначення магнітного поля морських суден, виявлення підводних човнів і мін, виявлення підводних копалин, процес неруйнівного контролю виробів магнітним методом та інше. Для отримання більш достовірної інформації про магнітні параметри подібних тіл є ефективним використання чисельних методів розрахунку.

Пашковський А.В. в своїй роботі [1] привів класифікацію методів розв'язку польових задач. Ця класифікація, котра була розширена і доповнена, зображена на рис. 1.

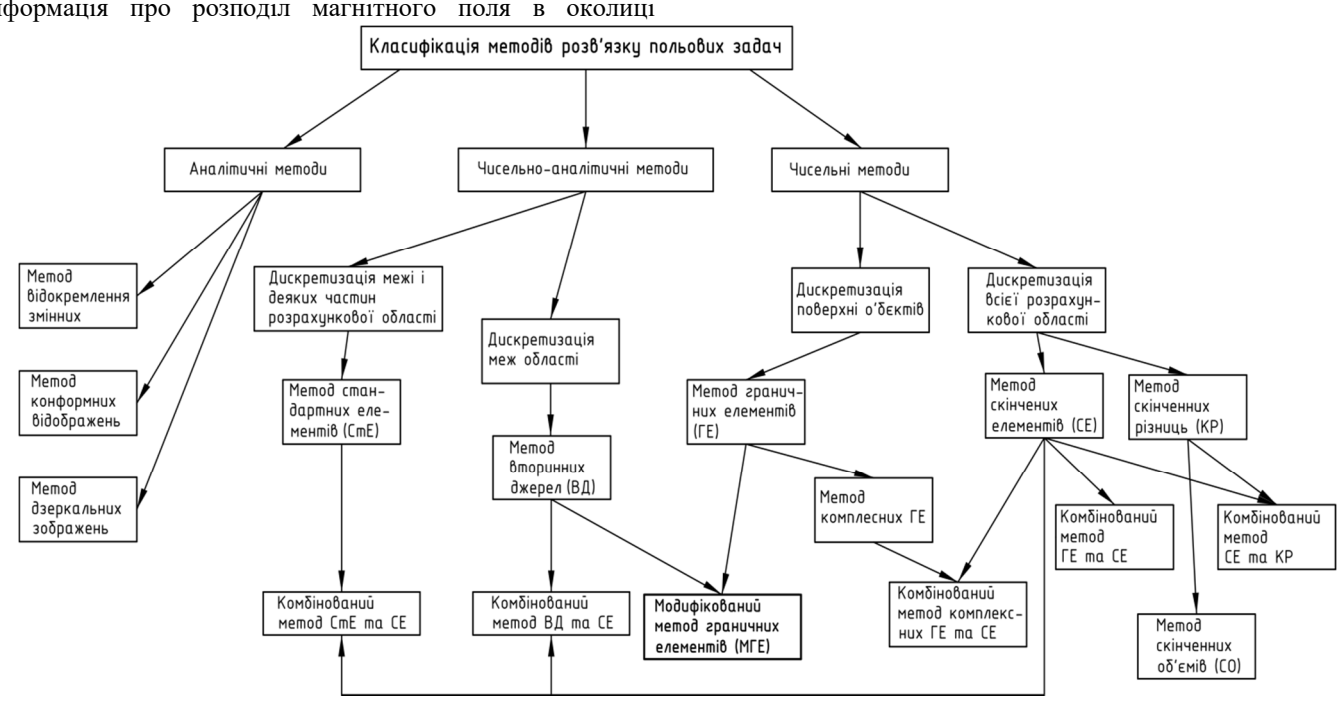


Рис. 1. Класифікація методів розв'язку польових задач.

Найбільш популярним обчислювальним методом є метод скінченних елементів (МСЕ). Понад 90% сучасних інженерних програм розрахунку ґрунтується на цьому методі. Однак у метода існує ще цілий ряд істотних недоліків — штучне обмеження області розрахунку, дискретизація навколишнього простору, виконання нової дискретизації при зміні положення елементів і ін. Це призвело до пошуків альтернативних методів, з яких найбільш перспективним є метод граничних елементів (МГЕ).

Кацикаделіс Дж.Т. в своїй книзі [2] описав переваги і недоліки МГЕ. Переваги МГЕ: 1. Дискретизації підлягає тільки межа тіла, що полегшує чисельне моделювання за допомогою МГЕ і на порядок зменшує число невідомих. Таким чином, спрощується і зміна моделі, пов'язана з урахуванням змін проекту. 2. МГЕ дозволяє обчислювати розв'язок та його похідні в будь-якій точці розглянутої області та в будь-який момент часу. Це можливо тому, що метод використовує інтегральне представлення розв'язку отриманого аналітичного виразу, яке можна диференціювати і використовувати як аналітичну формулу. Це неможливо при використанні МСЕ, так як розв'язок можна отримати тільки в вузлових точках. 3. Метод дуже ефективний для обчислень похідних польових функції. Можна з легкістю оперувати зосередженими силами і моментами як в області, так і на її межі. На поточній стадії розвитку МГЕ можна виділити наступні недоліки методу: 1. Чисельна реалізація МГЕ призводить до системи лінійних алгебраїчних рівнянь (СЛАР), у яких матриці коефіцієнтів повністю заповнені і несиметричні. При використанні МСЕ отримані матриці мають стрічкову структуру. Хоча, зазначений недолік МГЕ компенсується значно меншою розмірністю матриць. 2. Застосування МГЕ вимагає наявності фундаментальної

матриці розв’язків. Для побудови граничних елементів треба мати функцію Гріна для відповідної області. Такі функції знайдені не для всіх можливих областей. Цей метод не може використовуватися для задач, фундаментальний розв’язок яких не відомий чи не визначений. 3. Розроблені програми та підпрограми на базі МГЕ носять локальний характер і були створені лише для розв’язку конкретних задач в досліджуваній області, без особливих засобів візуалізації та повної автоматизації, що свідчить про необхідність розробки нової САПР, яка відповідає всім сучасним вимогам.

Філіппов Д.М. в своїй роботі [3] розвив метод вторинних джерел для моделювання плоскопаралельного магнітного поля електромагнітних систем, в якому при складанні СЛАР для знаходження густини вторинних джерел вперше запропоновано застосовувати інтегральні закони поля до елементів межі розділу середовищ.

II. Мета дослідження

Для тонких оболонок фундаментальна матриця розв'язків визначається складними громіздкими виразами, а для пологої сферичної оболонки виражається спеціальними функціями. Тому досліджень щодо розв’язку завдань з теорії пологих оболонок МГЕ небагато. У зв'язку з цим актуальною темою дослідження є розробка модифікованого методу граничних елементів (ММГЕ) і алгоритмів, заснованих на розробленому методі, для розв’язку лінійних і нелінійних задач теорії тонких оболонок, заснованих на застосуванні фундаментальних розв’язків, які визначаються простими аналітичними виразами. Для цього треба розробити аналітичні вирази для розв’язування задач тривимірного магнітного поля складної конфігурації на основі рівнянь для розв’язку задач плоскопаралельного магнітного поля [3]. Створити нову САПР, яка буде мати всі необхідні засоби візуалізації та буде повністю автоматизованою, що відповідає всім сучасним вимогам для розв’язку поставлених задач.

У запропонованому методі застосований інтегро-інтерполяційний спосіб дискретизації крайової задачі, заснований на дотриманні балансу магнітних потоків для контрольних об’ємів (КО) і представлення розв’язку у вигляді потенціалу в центрі КО.

Математичною основою збіжності методу є обмеженість операторів, зворотних проектуючим операторам та оператору граничних умов, а також густина в просторі неперервних функцій, побудована на основі дискретизації простого шару. При цьому консервативність схеми апроксимації граничних умов забезпечує «фізичність» розв’язку, отриманого при будь-якому, в тому числі і грубому, розбитті граничної поверхні. Остання обставина підтверджується, зокрема, наявним досвідом аналогового і аналого-цифрового моделювання поля намагнічення тонкостінних конструкцій на електричних сітках, при якому розв’язок може бути отриманий шляхом ітераційного підбору дискретних джерел струму (магнітного потоку), що задаються в вузлових точках однорідної R – сітки. Між моделюванням на електричних сітках і моделюванням чисельним методом, що пропонується, є майже повна аналогія. Відмінність полягає тільки в заміні розподілених джерел точковими.

III. Постановка задачі

У стаціонарному випадку (при впливі постійного в часі зовнішнього магнітного поля $\overline{H}_0$) потенціал φ_m шуканого поля намагнічування $H_m = -qrad\ \varphi_m$ тонкостінних конструкцій задовольняє [4,5] рівнянню Лапласа:

$$\varphi_m = 0 \tag{1}$$

Та наближеним («некласичним») граничним умовам на серединній поверхні конструкції S:

$$\left.\begin{aligned} &\varphi_m^{-} - \varphi_m^{+} = 0 \\ &\frac{\partial \varphi_m^{-}}{\partial n} - \frac{\partial \varphi_m^{+}}{\partial n} + div_s \mu_s^{*} grad \varphi_m = div_s \mu_s^{*} \overline{H}_{0s} \end{aligned}\right\} \tag{2}$$

де $\mu_s^{*} = \mu\delta$, $\mu^{*} = \frac{\mu}{\mu_0}$ та δ - відносна магнітна проникність і товщина конструкції відповідно.

Зовнішнє поле $\overline{H}_0$ вважається заданим. Потрібно знайти потенціал φ_m та значення напруженості $\overline{H}_m$ індукованого магнітного поля тонкостінних оболонок S.

Відомо [5], що розв'язок крайової задачі (1)—(2) може бути представлений у вигляді потенціалу простого шару (ППШ) з густиною g:

$$\varphi_m(t) = \int g(q)\Psi_m(t,q)dS_q \tag{3}$$

де t — точка спостереження; q — поточна точка поверхні S; $\Psi_m(t,q)$ — фундаментальний розв’язок Лапласа.

При цьому ППШ визначається з рішення інтегро-диференціального рівняння:

$$-g(t) + div_s \mu_s^{*} grad_s \int_S g(q)\Psi_m(t,q)dS_q = div_s \mu_s^{*} grad_s \overline{H}_{0s} \tag{4}$$

Перша основна ідея розробленого ММГЕ полягає в дискретизації рівняння (4) інтегро-інтерполяційним методом [6, 7] шляхом інтегрування по ділянці поверхні ΔS_j ($j = 1,...,N$). При цьому, переходячи від поверхневих інтегралів до контурних, одержуємо:

$$\begin{aligned} &-\int_S g(t)dS_t + \int_{L_j} \mu_s^{*}(l)\left[\frac{\partial}{\partial n_l}\int_S g(q)\Psi_m(t,q)dS_q\right]dl = \\ &= div_s \mu_s^{*} grad_s \overline{H}_{0s} \end{aligned} \tag{5}$$

де n_l - зовнішня нормаль до контуру L_j, що обмежує поверхню ΔS_j.

Рівняння (5) виражає умови балансу магнітного потоку для контрольного об'єму (елемента тонкостінної конструкції), обмеженого циліндричною поверхнею з твірною L_j.

В запропонованому методі апроксимується не сама функція φ_m, а простий шар, який її визначає, тобто формально використовується розкладання φ_m по неортогональним функціям [8]. Ця друга основна ідея методу має просту фізичну інтерпретацію: простий шар замінюється системою точкових джерел з інтенсивністю $Q_j = \int_{S_j} g(t)dS_t$, розміщених в вузлах q_j ($j = 1,...,N$). Так як потенціал $\varphi_m(t)$ є неперервною-диференційованою гладкою функцією на границі L_j всіх ділянок ΔS_j, при вказаному способі апроксимації (6) консервативність методу не порушується.

$$\varphi_m(t) = \sum_{i=1}^{N} Q_i \Psi_m(t, q_i) \qquad (6)$$

В результаті рівняння (4) зводиться до наступної СЛАР відносно значень Q_j:

$$-Q_j + \int_{L_j} \mu_s^*(l)[\frac{\partial}{\partial n_l} \sum_{i=1}^{N} Q_i \Psi_m(t, q_i)]dl = \int_{L_j} \mu_s^*(l) H_{0n}(l)dl, \quad j = 1,...,N \qquad (7)$$

Перепишемо рівняння (7) в компактному вигляді та поміняємо порядок додавання та інтегрування:

$$-Q_j + \sum_{i=1}^{N} a_{ij} Q_i = c_j, \quad j = 1,...,N, \qquad (8)$$

де

$$a_{ij} = \int_{L_j} \mu_s^*(l) \frac{\partial \Psi_m(t, q_i)}{\partial n_l} dl \; ; \; c_j = \int_{L_j} \mu_s^*(l) H_{0n}(l)dl \qquad (9)$$

Інтеграли в вище згаданому рівнянні зв'язують вузол q_j, де використовується фундаментальний розв'язок, з вузлом q_i ($i = 1,...,N$). a_{ij} та c_j часто називаються коефіцієнтами впливу. Точка q_j залишається незмінною (опорна точка) (рис. 2), в той час коли точка q_i змінюється по і-му елементу (точка інтегрування).

Контурний інтеграл (7) розраховується за формулою прямокутників на відрізках l_{jkl} ($k = 1,..., P_j$; $l = 1,2$).Тоді вираз для коефіцієнтів впливу в (8) матиме вигляд:

$$a_{ij} = \sum_{k=i}^{P_j} \sum_{i=1}^{2} \mu_s^*(t_{jkl}) l_{jkl} \left. \frac{\partial \Psi_m(t, q_i)}{\partial n_{jkl}} \right|_{t_{jkl}}, \qquad (10)$$

$$c_j = \sum_{k=i}^{P_j} \sum_{i=1}^{2} \mu_s^*(t_{jkl}) l_{jkl} H_{0n_{jkl}}, \qquad (11)$$

де l_{jk1} та l_{jk2} - відрізки, що з'єднують центр мас трикутника з серединою сторін його бічних граней; $\overline{n}_{jk1}$ та $\overline{n}_{jk2}$ - зовнішні (відносно ΔS_j) нормальні вектори до відрізків l_{jk1} та l_{jk2}, що лежать в площині Δ_{jk}.

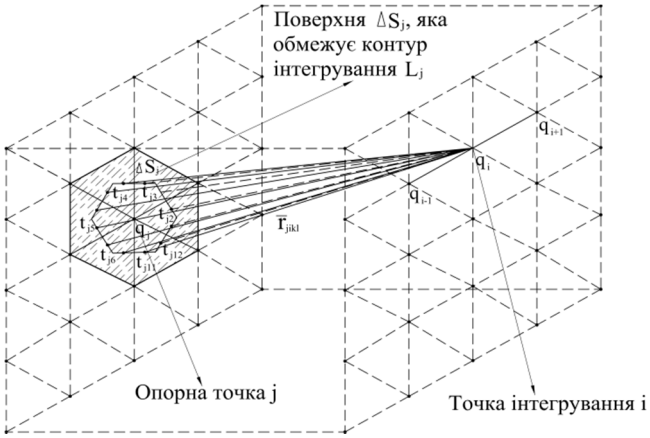


Рис. 2. Апроксимація тривимірної поверхні при розрахунку коефіцієнтів впливу

Після розв'язку СЛАР шуканий потенціал φ_m розраховується за формулою (6), а напруженість поля визначаються за формулами:

$$\overline{H}_m(t) = -\sum_{i=1}^{N} Q_i qrad_i \Psi_m(t, q_i) \qquad (12)$$

При $\mu_s = \infty$ вихідне рівняння (4) зводиться до інтегрального рівняння першого роду. Спектр оператору цього рівняння містить нульове власне число, яке є причиною нестійкості чисельного розв'язку СЛАР (7). Щоб запобігти нестійкості розв'язку СЛАР в ліву частину вводимо добавку γ [9], яка змінює спектр матричного оператора:

$$-Q_j + \sum_{i=1}^{N} a_{ij} Q_j + \gamma = c_j \quad j = 1,...,N \qquad (13)$$

В рівняння (13) додається умова рівності нулю сумарного магнітного заряду:

$$\int_S g(t)dt = \sum_{i=1}^{N} Q_j + \gamma = 0 \qquad (14)$$

СЛАР (14) в матричному вигляді:

$$[A]\{Q\} = \{C\} \tag{15}$$

Матриця A - це вироджена матриця. Для її розв'язку найефективніше використовувати метод найменших квадратів, а в якості ортогонального розкладання використовувати SVD (singular value decomposition) або сингулярне розкладання.

Після сингулярного розкладання матриця A матиме вигляд:

$$A = USV^T \tag{16}$$

де S — матриця такого ж розміру, як і матриця A, у котрій елементи, що лежать на головній діагоналі — це сингулярні числа (а всі елементи, які не лежать на головній діагоналі, дорівнюють нулю); U і V — це дві унітарні матриці, що складаються з лівих і правих сингулярних векторів відповідно.

СЛАР (15) з урахуванням (16) матиме вигляд:

$$[U][S]\{Q\} = \{C\}[V] \tag{17}$$

Тепер отриману СЛАР можна вирішити за допомогою методу найменших квадратів.

Розроблений ММГЕ складається з наступних кроків: на першому етапі відбувається тріангуляція поверхні заданого об'єкта; обчислення коефіцієнтів a_{ij} матриці A; обчислення коефіцієнтів c_j вектору C; сингулярне розкладання матриці (16); розрахунок значень інтенсивності Q_j в опорних точках q_j; розрахунок магнітного потенціалу та напруженості індукованого поля в будь-якій точці простору за формулами (18) і (19) відповідно:

$$\varphi_m = \Psi_m(t, q_i) = \\ = \frac{1}{4\pi}\sum_{i=1}^{P_j}\frac{Q_j}{\sqrt{(x_t - x_j)^2 + (y_t - y_j)^2 + (z_t - z_j)^2}} \tag{18}$$

$$\left.\begin{aligned} H_{mx} &= -\frac{1}{4\pi}\sum_{i=1}^{P_j}\frac{Q_j(x_t - x_j)}{\sqrt[3]{(x_t - x_j)^2 + (y_t - y_j)^2 + (z_t - z_j)^2}} \\ H_{my} &= -\frac{1}{4\pi}\sum_{i=1}^{P_j}\frac{Q_j(y_t - y_j)}{\sqrt[3]{(x_t - x_j)^2 + (y_t - y_j)^2 + (z_t - z_j)^2}} \\ H_{mz} &= -\frac{1}{4\pi}\sum_{i=1}^{P_j}\frac{Q_j(z_t - z_j)}{\sqrt[3]{(x_t - x_j)^2 + (y_t - y_j)^2 + (z_t - z_j)^2}} \end{aligned}\right\} \tag{19}$$

На базі розробленого ММГЕ була створена програма для розрахунку індукованого магнітного поля тонкостінних конструкцій в тривимірному просторі Mesh_lab, алгоритм якої наведено на рис. 3.

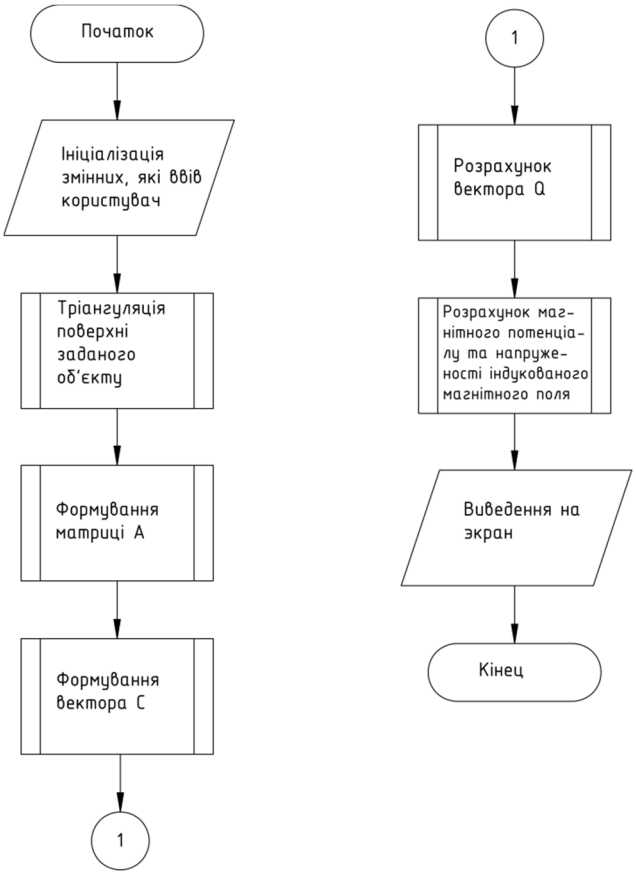


Рис. 3. Алгоритм роботи програми

Програма, реалізована мовою C++ з використанням Open Source бібліотеки GNU GSL. Алгоритм програми Mesh_lab складається з наступних основних блоків:

Тріангуляція поверхні заданого об'єкта (тонкостінної конструкції). Будується поверхнева конформна тріангуляція області за допомогою портативних бібліотек Ani3D [10]. Вхідна інформація: параметри об'єкта і крок тріангуляції. Вихідна інформація: файл mesh.out, що містить кількість точок розбиття (nV) і їх координати, кількість трикутників (faces) і номери вершин, що входять в кожен трикутник. Також, виводиться поверхнева трикутна сітка заданого об'єкта на екран.

Формування масивів для обробки результатів тріангуляції, які зберігаються в файлі mesh.out.

Формування матриці A. Функція get_A () обчислює елементи матриці A в кожній j-тій точці (опорна точка) розбиття поверхні і знаходить коефіцієнт впливу a_{ij} від цієї точки на інші і-ті точки (точки інтегрування) за формулою (10).

Формування матриці C. Функція get_C () обчислює елементи векторів C в кожній точці розбиття поверхні за формулою (11).

Визначення інтенсивності точкових джерел Q. Функція get_X () формує матриці і вектори для розрахунку інтенсивності точкових джерел Q_j і викликає функції sl_linalg_SV_decomp () і gsl_linalg_SV_solve () з бібліотеки GNU GSL; sl_linalg_SV_decomp (), розраховує сингулярне розкладання матриці (16), де S — діагональна матриця такого ж розміру, як матриця A, у якої елементи, що лежать на головній діагоналі - це сингулярні числа; U та V — це унітарні матриці, що складаються з лівих і правих сингулярних векторів відповідно. Функція gsl_linalg_SV_solve () вирішує СЛАР (16). Результати розрахунку зберігаються в вектор Q і записуються в файл result.txt.

Обчислення індукованого магнітного поля в заданій точці простору. Функція calc_field() обчислює значення потенціалу φ_m за формулою (18) і напруженості $\overline{H}_m$ за формулою (19) в заданій точці простору.

Детальніше зупинимося на алгоритмі формування матриці A, який зображений на рис. 4.

В першу чергу функція get_A() ініціалізує масив для зберігання елементів матриці A, розміром $nV \times nV$. Далі читаємо координати поточної вершини та розраховуємо коефіцієнт впливу між цією вершиною та усіма іншими вершинами за формулою (10). Для цього навколо кожної вершини розраховується елемент ΔS_j, який складається з трикутників, що примикають до цієї вершини. Межа контуру — це межа елемента ΔS_j.

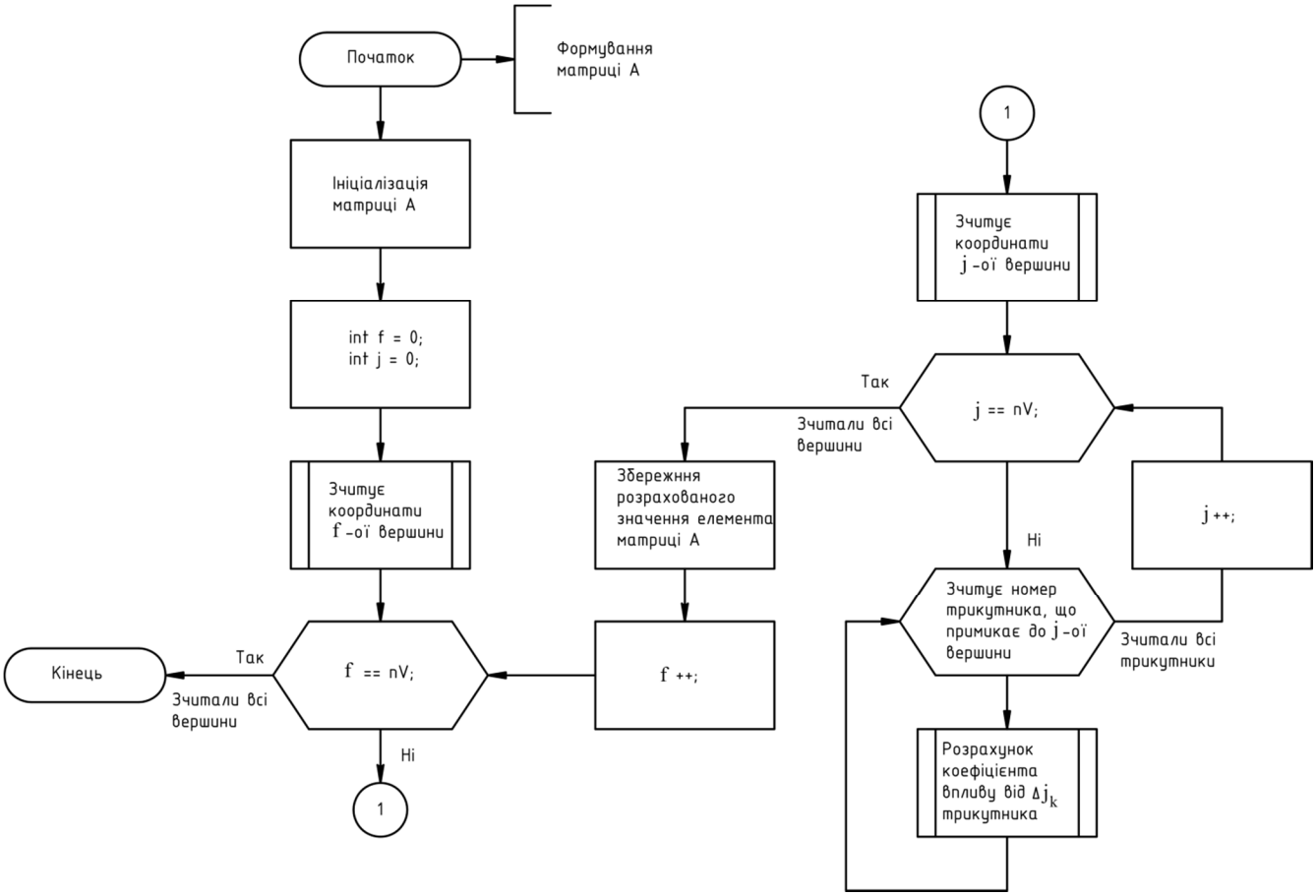
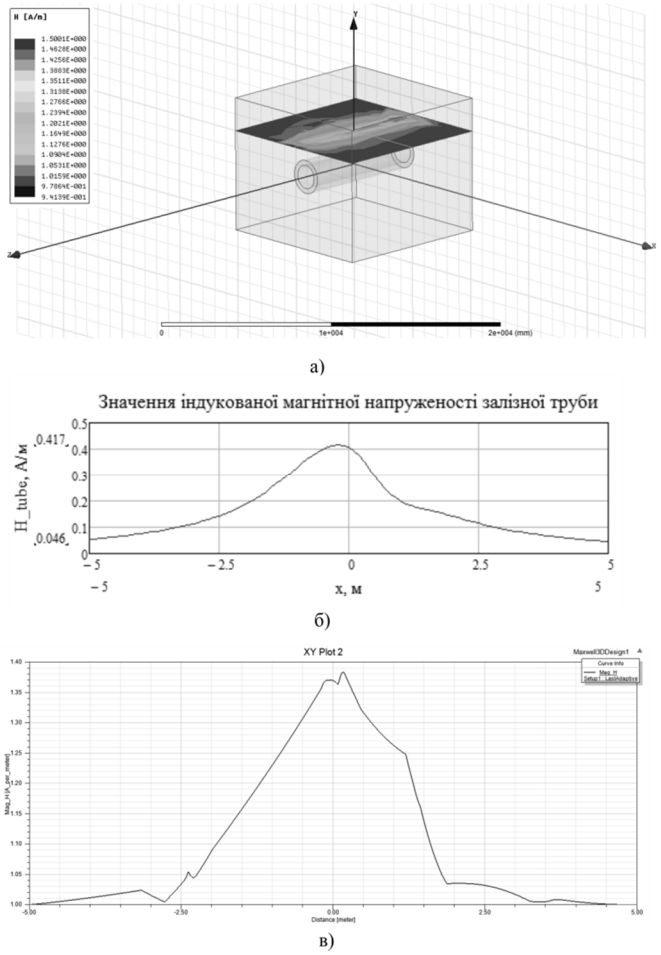


Рис. 4. Алгоритм формування матриці А.

Розглянемо роботу розробленого ММГЕ на прикладі. Розрахуємо магнітну індукцію залізної труби, зовнішній радіус $r_1 = 1м$, внутрішній радіус $r_2 = 0,7м$, довжина труби $l = 8м$, магнітна проникність заліза $\mu = 400$. Знаходиться в зовнішньому полі $H_{0x} = 0\ А/м$, $H_{0y} = 1\ А/м$, $H_{0z} = 0\ А/м$. Вісь труби співпадає з віссю z. Центр труби співпадає з початком координат.

Напруженість магнітного поля була розрахована на відстані 2 м від осі циліндру ($y = 2м, z = 0м$) на відрізку з координатами по осі абсцис [-5;5] метрів (рис. 5а). Значення індукованої магнітної напруженості $\overline{H}_m$ були розраховані за допомогою розробленого ММГЕ (рис. 5б). Отримані значення були перевірені за допомогою МСЕ в програмі ANSYS (рис. 5в). На рис. 5в зображено значення напруженості результуючого поля $\overline{H} = \overline{H}_0 + \overline{H}_m$. Як видно

на рис. 5, значення індукованої магнітної напруженості $\overline{H}_m$ тонкостінного циліндру, розрахованих на основі ММГЕ, були підтверджені результатами напруженості $\overline{H}$, отриманими в ANSYS.



а)

б)

в)

Рис. 5. Розрахунок магнітного поля тонкостінного циліндру.

IV. Висновки

Запропоновано новий актуальний модифікований метод граничних елементів та алгоритм розв'язання задач індукованого магнітного поля тонкостінних конструкцій. ММГЕ базуються на методі граничних елементів та методі вторинних джерел.

Застосування ММГЕ дозволяє підвищити точність та швидкість розрахунку індукованого магнітного поля будь-якої конфігурації. На базі розробленого ММГЕ було розроблено алгоритм, на базі якого була створена програма для розрахунку індукованого магнітного поля тонкостінних конструкцій в тривимірному просторі Mesh_lab. Необхідно підкреслити, що сфера застосування розробленого методу величезна. Це і визначення магнітного поля морських суден, виявлення підводних човнів і мін, виявлення підводних копалин, процес неруйнівного контролю виробів магнітним методом та інше.

Розроблений ММГЕ, на відміну від класичного МГЕ, не вимагає наявності матриці фундаментальних розв'язків. Для цього були розроблені універсальні коефіцієнти впливу, за допомогою яких можна розв'язувати завдачі з довільною геометрією. Представлений ММГЕ позбавлений вище зазначених недоліків МГЕ і МСЕ, тому, в порівнянні з ними, перспектива його використання набагато вище.

Розроблена програма для розрахунку індукованого поля тонкостінних конструкцій за допомогою ММГЕ може бути інтегрована в програмний комплекс Gmsh. Цей комплекс містить чотири модулі для опису геометрії, створення сітки і подальшої обробки розв'язку задачі (візуалізації результатів). Gmsh підтримує параметричне введення і володіє сучасними механізмами візуалізації. На базі цієї програми планується створити нову САПР, що володіє всіма засобами візуалізації і повної автоматизації, що відповідає всім сучасним вимогам.

Література References

[1] Пашковский А.И. Численно-аналитические методы стандартных элементов для моделирования стационарных физических полей в линейных кусочно-однородных и нелинейных средах: автореферат дис. доктор техн. наук. — Новочеркасск, 2014. — 36 с.

[2] Katsikadelis, John T The Boundary Element Method for Engineers and Scientists 2nd Edition. — Academic Press, 2016, 464 p.

[3] Филиппов Д.М. Развитие метода вторичных источников для моделирования плоскапараллельного магнитного поля электромагнитных систем: автореферат дис. канд. техн. наук. — Харьков, 2013. — 24 с.

[4] Цейтлин Л.А. Об определении магнитных и электрических полей тонких слоев и оболочек. — Журнал технической физики, 1958, т. 28, вып.6, с. 1326 – 1329.

[5] Краснов И.П. Орешении задач некоторых граничных задач теории гармонических функций. — Диффиренциальные уравнения, 1975, т. XI, № 11, с. 2052 – 2066.

[6] Марчук Г.И., Агошков В.И. Введение в проекционно-сеточные методы. — М.: Наука, 1981, 416 с.

[7] Патанкар С. Численные методы решения задач теплообмена и динамики жидкости. — М.: Энергоатомиздат, 1984, 152 с.

[8] Алексидзе М.А. Решение граничных задач методом разложения по неортогональным функциям. М.: Наука, 1978, 352 с.

[9] Белоцерковский С.М., Лифанов И.К. Численные методы в сингулярных интегральных уравнениях. — М.: Наука, 1978, 352 с.

[10] Данилов А. А. Способы построения трёхмерных поверхностных триангуляций и тетраэдральных сеток // Научно-технический вестник СПбГУ ИТМО. 2010. Т. 65, № 1. С. 87–92.

Інформаційно-Аналітична Система для Оптимального Планування Режимів Роботи Підземних Сховищ Газу

Назар Притула, Мирослав Притула
Відд. розробки систем опт. план. та прогн. режимів роботи газотранспортної системи
Філія «Науково-дослідний інститут транспорту газу» ПАТ «УКРТРАНСГАЗ»
Харків, Україна
nazar.prytula1@gmail.com, myroslav.prytula@gmail.com

Олег Гринів
Аспірантура
ІППММ ім. Я. С. Підстригача НАН України
Львів, Україна
oleg.gryniv@gmail.com

Information Analytical System for Optimal Scheduling Operating Modes of Underground Gas Storages

Nazar Prytula, Myroslav Prytula
Dept. for development of systems of optimal scheduling and forecasting operating modes of the gas transmission system
Research and Design Institute of Gas Transmission PJSC Ukrtransgas
Kharkiv, Ukraine
nazar.prytula1@gmail.com, myroslav.prytula@gmail.com

Oleg Gryniv
Postgraduate study
Pidstryhach IAPMM of NASU
Lviv, Ukraine
oleg.gryniv@gmail.com

***Анотація*—В роботі приведена змістовна постановка режимних задач функціонування технологічно пов'язаних підземних сховищ газу в складі газотранспортної системи та характеристика алгоритмічного забезпечення їх розв'язування. Крім цього обґрунтована важливість розв'язання задач оптимального планування процесів нагнітання/відбирання газу та побудови екстремальних характеристик роботи ПСГ в залежності від реальних режимних параметрів роботи газотранспортної системи. Приведені результати апробації розроблених програмних модулів.**

***Abstract*—The paper shows the problem statements of operation of technologically connected underground gas storages as the part of the gas transmission system and characteristics of algorithmic software for their solving. Besides this paper justify the importance of solving problems of optimal scheduling processes of gas injection and withdrawal and building extreme characteristics of the underground gas storages based on actual operational parameters of the transmission system. The results of testing the developed software modules are shown.**

Ключові слова—підземне сховище газу; математична модель; оптимальний режим роботи; газотранспортна система; програмний комплекс; піковість; компресорна станція.

Keywords—underground gas storage; mathematical model; optimal operating mode; gas transmission system; peak capacity; compressor station

I. Вступ

В складі української газотранспортної системи знаходиться 12 підземних газових сховищ (ПСГ) різної потужності, які нерівномірно розміщені по території країни. Більше 85% потужності знаходиться в західних областях України. Їх основна роль – забезпечення споживачів газом в умовах різкого похолодання, заміщення транзитних потоків в осінньо-зимовий період для надійного забезпечення газом східних регіонів України та забезпечення живучості системи у випадку появи нештатних ситуацій. Протягом опалювального сезону коливання рівня споживання газу багатократно міняється в значних межах. І тому найважливішою характеристикою ПСГ є її максимальна продуктивність, яку називають піковістю ПСГ. Фактори впливу на величину піковості залежать від

режиму функціонування ГТС та наповнюваності газом ПСГ. Газ в пластах – колекторах розділяють на буферний та активний. Буферний об'єм газу, як правило, є постійною величиною і служить "пружиною" для відбирання так званого активного газу. Газотранспортна система в останні роки є недовантаженою, що викликано суттєвим зменшенням загального споживання газу, і тому працює в непроектних режимах. Крім цього, із-за зміни джерел постачання газу в систему магістральних газопроводів, існують реверсні потоки.. Необхідність наповнювати всі сховища газом відпала. В зв'язку із сказаним виникла проблема розподілу необхідних об'ємів активного газу по сховищах на етапі його нагнітання та розподілу об'ємів відбирання в непікових режимах їх експлуатації. Важливою умовою експлуатації ПСГ є постійна підтримка газосховищ в стані максимальної піковості протягом всього сезону відбирання газу.

Рішення проблеми, яка виникла, пов'язане із ефективним розв'язанням таких задач:

- побудова функцій піковості в залежності від тиску на виході компресорної станції;
- розподіл об'ємів зберігання газу для досягнення піковості ПСГ на заданих інтервалах часу чи середньої піковості протягом всього сезону відбирання газу;
- розподіл об'ємів відбирання газу в такий спосіб, щоб залишки невідібраного активного газу могли забезпечити максимальну піковість ПСГ.

Розроблена система для розв'язування задач об'єднує декілька програмних модулів. Основними є модулі інформаційного забезпечення задач, моделювання та оптимізації. Інформаційне забезпечення формується із заміряних даних, а також даних отриманих в процесі розв'язання задач ідентифікації параметрів моделей та прогнозування балансу газу в системі. Математична модель системи об'єднує в єдину гідравлічну систему моделі всіх технологічних об'єктів, які беруть участь в процесі нагнітання та відбирання газу, це, зокрема, фільтраційна модель пласту – колектора та привибійної зони свердловин, моделі руху газу в свердловинах та мережі збору газу, моделі компримування газу газоперекачуючими агрегатами (ГПА) на компресорній станції (КС). Така інтегрована модель дозволила поставити та розв'язати повний набір режимних задач включно із оберненими задачами для ідентифікації параметрів моделей системи.

Всі розраховані режими роботи ПСГ на програмному комплексі є оптимальними за паливних газом [1-8]. Приводи ГПА на більшості сховищ є газотурбінного типу. Це й же комплекс дозволив поставити та розв'язати задачу на максимальну продуктивність сховища (піковість) в залежності від тиску на виході КС.

Всі газосховища мають різні характеристики за об'ємами зберігання газу, піковістю та затратами зберігання одиниці об'ємів газу. В умовах різкого зростання споживання газу основним критерієм є не оптимальність, а піковість ПСГ. І тому в умовах обмежених ресурсів газу важливо його розмістити в пластах ПСГ за критерієм досягнення максимальної піковості на прогнозованих інтервалах часу. Найчастіше, в умовах нечіткого прогнозування споживання газу, диспетчерські служби розміщають газ за критерієм максимальної середньої піковості на всьому інтервалі часу роботи всіх ПСГ в піковому режимі їх роботи. Як, легко бачити, для забезпечення критерію оптимальності необхідно розподілити об'єм активного газу в такий спосіб, щоб робота всіх газосховищ в піковому режимі починалася і завершувалася в один і той же час.

Стратегія експлуатації ПСГ є такою, щоб постійно підтримувати максимально можливу їх піковість. В цьому випадку необхідно розподіляти прогнозований об'єм відбирання газу між сховищами в такий спосіб, щоб наявний активний газ в сховищах забезпечував максимальну сумарну піковість.

Розроблене алгоритмічне та програмне забезпечення дозволяє розв'язувати і більш складні задачі, які забезпечують ефективну експлуатацію ПСГ. В процесі формування оптимального плану роботи ГТС є можливість вибирати тиск на виходах КС ПСГ для забезпечення оптимальної роботи сумісної роботи ГТС та ПСГ.

II. Про сумарну піковість та оптимальність

Оптимальна сумісна робота ПСГ і ГТС вимагає: узгодженого режиму роботи ПСГ з роботою МГ; у випадку водонапірного режиму – узгодження темпів відбирання та темпів руху газоводяного контакту (ГВК); при заданих об'ємах зберігання, в гідравлічно незв'язних пластах багато пластових пластів – колекторів, оптимально розподілити об'єми між ними; оптимального об'єму буферного газу, який пов'язаний з можливостями регулювання тиску в магістралі; використання в повній мірі періодів безкомпресорного нагнітання та відбирання газу є можливість їх встановлення при заданих планах нагнітання та відбирання газу; узгодженого тиску газу в магістральному газопроводі.

В останні роки постачальниками газу виступає декілька фірм. На встановлення об'ємів газу та їх постачальників впливають економічні чинники, що вимагає постійно проводити адаптацію режимів роботи ПСГ та ГТС до зміни об'ємів та джерел їх розміщення, для поступлення газу в систему газопроводів. Із різким зростання вартості газу на внутрішньому ринку, особливо для побутового використання, змінюються і об'єми його споживання. Існуюча система субсидій не стимулює проведення широкомасштабного впровадження систем тепло зберігання. Для розроблення надійного прогнозування споживання на даний час часові ряди ще є надто короткими. В найближчий час очікується впровадити систему надання субсидій в грошовій формі, що буде стимулювати до економії газу. В опалювальний сезон основним фактором впливу на об'єми відбирання газу із ГТС є температура зовнішнього повітря. Часто температурний фон по регіонах країни може суттєво відрізнятися, що вимагає оперативного диспетчерського втручання для вчасного балансування системи та її підсистем.

Будемо розрізняти піковість на заданому інтервалі часу і пікову характеристику ПСГ. Пікова характеристика ПСГ будується однозначно і при потребі корегується за умови зміни параметрів технологічного обладнання. Пікова характеристика встановлює зв'язок між об'ємами акумульованого активного газу в пластах ПСГ і максимальними відборами за умов його відбирання за мінімальний час, якщо в початковий момент відбирання газу пласт містив максимальний об'єм активного газу. Як бачимо, що піковість, на певних інтервалах часу, може дещо перевершувати пікову характеристику. Це пов'язано з тим, що тиск в робочій зоні може бути дещо вищим ніж отриманий за умов побудови пікової характеристики. Пікова характеристика залежить ще і від тиску в магістральному газопроводі. Якщо ДКС укомплектована ГПА з газотурбінними приводами, то пікова характеристика може мати розриви та скачки (див. рис. 1). Сумарна розрахована піковість представлена на рис. 2-3. Всі розрахунки в роботі проведені на програмному комплексі "ПСГ режим" [1-8].

III. Програмний комплекс. Веб-застосунок «ПСГ»

Означемо два поняття. Так поняття «**розподіл**» означає структуру даних, яка містить свою назву, тип, дату, користувача, який створив, а також масив відповідних значень для кожного ПСГ, або ж розподілів іншого типу. Поняття «форма» означає структуру даних, яка містить свою назву, тип задачі, результати якої представляє, дату, сезон, користувача, який створив, а також масив заповнених збережених розподілів, які використовуються даною формою.

Інформаційне забезпечення задач формується на основі даних, які містяться в журналах Центрального диспетчерського департаменту, а також необхідних даних для роботи застосунку, які збережені користувачами у вигляді «форм» як результати виконаних задач та «розподілів» і які сформовані як автоматично, так і створені безпосередньо за участю користувача.

Більшість «форм» представляються у табличному вигляді. Окремі типи «форм» для відповідних задач можуть результати розрахунків зберігати у зручному вигляді для їх графічного представлення. Кожна задача, а точніше її тип, прив'язані до певного режиму роботи ПСГ. Існують такі режими роботи: відбирання; нагнітання; відбирання/нагнітання. Передбачені в системі такі типи задач:

максимальна витрата на інтервалі часу;

нагнітання/відбирання, максимальна витрата на інтервалі часу;

мінімальний час повного відбирання/повного нагнітання ПСГ;

нагнітання/відбирання, мінімальний час повного відбирання/повного нагнітання ПСГ;

прогнозний та фактичний режим роботи;

прогнозний режим максимальної витрати газу;

оперативний прогноз;

максимальна добова поінтервальна продуктивність.

Розв'язування того чи іншого типу задач активізується для відповідних режимів роботи ПСГ. Розрахункові задачі містять поля для введення користувачем вхідних даних для розрахунку. анування режимів роботи ПСГ проходить на певному наборі Форм. Заповнення стовпці і рядків форм, при необхідності, проходить автоматично як даними з інформаційної бази так і даними, які отримані в результаті розв'язування вище перерахованих задач. Для кожної задачі можна створити нову форму, або завантажити для перегляду вже раніше збережену. При наявності у користувача прав, можна перейти у режим редагування форми. Після цього для нових форм є можливість їх збереження, а для завантажених – збереження (при наявності прав), збереження копії, видалення (при наявності прав), керування дозволами (для користувачів, які створили дану форму). Розроблено функціонал керування дозволами (перегляду, редагування, видалення) для форм та розподілів. Розроблена система запитів може надати дозвіл та їх схвалення або заборона власниками форм та розподілів.

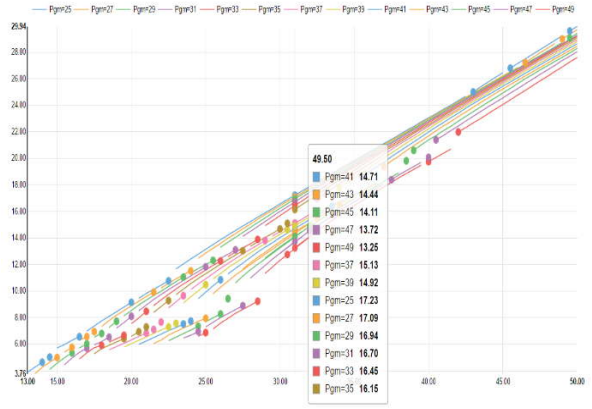


Рис. 1. Графік залежності максимальної продуктивності відбирання газу з ПСГ у млн.м^3/добу від пластового тиску у атм при різних тисках газу у магістральному газопроводі

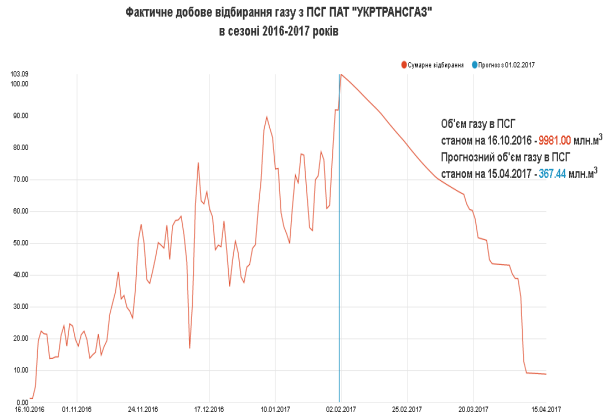


Рис. 2. Фактичне добове відбирання газу (до вертикальної лінії). Прогнозний графік максимального відбирання газу (після вертикальної синьої лінії)

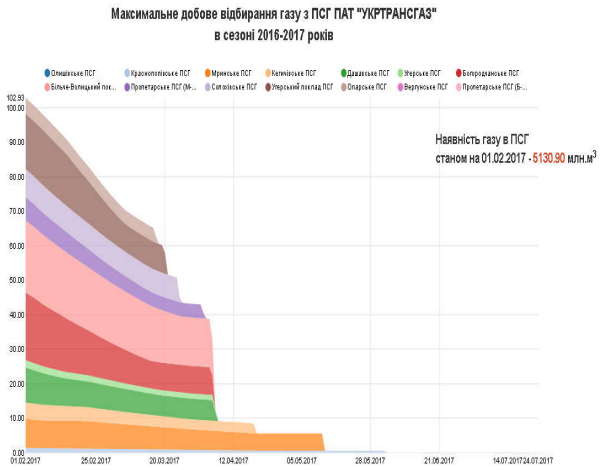


Рис. 3. Максимальне розрахункове щодобове відбирання газу протягом сезону по ПСГ (відмічене різними кольорами)

При створенні форм, автоматично відбувається створення і заповнення їх актуальними даними з БД розподілів. Частина даних є обов'язковими для даного типу форми, інша формується способом вичитування розподілів, за умовчанням, для певних типів розподілів. При бажанні, розподіли, які використовуються формою, можна замінити на інші, доступні користувачеві, з випадного списку для кожного типу розподілу. У режимі редагування форми з'являється можливість в таблиці редагувати значення певних доступних розподілів.

IV. Висновки

В процесі експлуатації розробленого програмного комплексу показано, що:

- на величину фільтраційних параметрів найбільш значний вплив має ближня зона вибою свердловин;
- для управління режимами роботи ПСГ не завжди обов'язковим є знання фільтраційних коефіцієнтів кожної свердловини для цього досить знати їх середнє значення, яке, на значних інтервалах часу, є досить стабільним;
- розподіл пластового тиску у всій області пласта-колектора формується протягом значного інтервалу часу і тому для його відтворення необхідно проводити моделювання роботи ПСГ протягом трьох-п'яти років;
- розрахунок розподілу пластового тиску вимагає проводити одночасно з ідентифікацією розподілених параметрів пласта, що забезпечує досягнення збігу розрахованих та заміряних середньо пластових тисків в області відборів та нагнітання газу;
- процесу проведення оптимізації роботи окремих та груп технологічно поєднаних ПСГ передує оцінка можливого його потенціалу;
- піковість газосховищ пов'язана із пропускною спроможністю об'єктів газосховищ;
- потенціал збільшення продуктивності пласта і ПСГ в цілому не завжди
- пов'язаний з потенціалом збільшення продуктивності свердловин;
- існує гранична межа економічної доцільності нарощення густини перфораційних каналів свердловин;
- розроблене алгоритмічне та програмне забезпечення розв'язування задач сезонного та оперативного планування режимів експлуатації ПСГ та груп технологічно поєднаних ПСГ вимагає автоматичної актуалізації системи побудови функції піковості газосховищ у процесі їх роботи.

Оптимальна робота ПСГ, груп технологічно поєднаних ПСГ та сумісної роботи ПСГ і ГТС забезпечується багаторівневою системою оптимального планування та оптимальною експлуатацією на всьому періоді нагнітання та відбирання газу. Оптимізація ПСГ проходить на етапах формування: оперативного балансу підсистем та системи в цілому ГТС; оптимального режиму ГТС і ПСГ; планування та процесу керування експлуатацією роботи груп технологічно поєднаних ПСГ; оптимізацією режимів роботи безпосередньо кожної ПСГ тощо. Реалізація потенціалу оптимізації роботи ПСГ проводиться вибором тисків на виходах ДКС, збільшенням часового інтервалу достовірного прогнозування, оптимізацією роботи ДКС, завчасною реакцією системи на зміни балансових показників роботи ГТС.

Література References

[1] Б. О. Клюк, Р. Л. Вечерик, Ю. Б. Хаєцький, Н. М. Притула, Я. Д. П'янило, М. Г. Притула, "Підземні сховища газу в системі забезпечення ефективної експлуатації газотранспортної системи: проблеми розвитку й експлуатації", *Нафт. і газ. промисл.,* № 6, С. 7–10, 2009.

[2] Н. Притула, М. Притула, Р. Боровий, О. Химко, "Математична модель Більче -Волицького сховища газу", *Вісник НУ «Львівська політехніка»: Комп. науки та інформ. техн.,* № 686, С. 192–198, 2010.

[3] І. Боярин, Р. Боровий, О. Гринів, Н. Притула, В. Ямнич, "Оптимізація роботи багато цехових компресорних станцій з різнотипними газоперекачуючими агрегатами", *Вісник НУ «Львівська політехніка»: Комп. науки та інформ. техн.,* № 672, С. 326–335, 2010.

[4] Н. Притула, "Математичні моделі заміщення буферного газу азотом у пластах газосховища", *Фіз.-мат. мод. та інформ. техн.,* вип. 14, С. 115–123, 2011.

[5] Н. М. Притула, Я. Д. Пянило, М. Г. Притула, Підземне зберігання газу (математичні моделі та методи). Львів: Видавництво «Растр-7», 2015.

[6] О. Гринів, Н. Притула, М. Притула, "Математичне моделювання та оптимізація сумісної роботи газосховищ", *Вісник НУ «Львівська політехніка»: Комп. науки та інформ. техн.,* № 744, С. 243–248, 2012.

[7] Н. М. Притула, М. Г. Притула, Р. Я. Шимко, С. В. Гладун, "Розрахунок режимів роботи Більче-Волицько-Угерського підземного сховища газу (програмний комплекс)", *Нафтогазова галузь України,* № 3, С. 36–41, 2013.

[8] Н. Притула, О. Гринів, Р. Вечерік, Р. Бойко, "Заміщення буферного газу азотом у пластах газосховища (моделі, методи, числові експерименти)", *Нафтогазова галузь України,* № 4, С. 32–39, 2013.

Оптимальне Планування Режимів Роботи Складних Газотранспортних Систем

Назар Притула, Мирослав Притула
Відд. розробки систем опт. план. та прогн. режимів роботи газотранспортної системи
Філія «Науково-дослідний інститут транспорту газу» ПАТ «УКРТРАНСГАЗ»
Харків, Україна
nazar.prytula1@gmail.com, myroslav.prytula@gmail.com

Вадим Фролов
Центр. диспетчерський департамент
ПАТ «УКРТРАНСГАЗ»
Київ, Україна
frolov-va@utg.ua

Optimal scheduling operating modes of gas transmission systems with complicated structure

Nazar Prytula, Myroslav Prytula
Dept. for development of systems of optimal scheduling and forecasting operating modes of the gas transmission system
Research and Design Institute of Gas Transmission PJSC Ukrtransgas
Kharkiv, Ukraine
nazar.prytula1@gmail.com, myroslav.prytula@gmail.com

Vadym Frolov
Central dispatching department
PJSC Ukrtransgas
Kyiv, Ukraine
frolov-va@utg.ua

***Анотація*—В роботі приведені змістовні постановки задач оптимального планування режимів роботи складних газотранспортних систем, моделями яких є системи нелінійних рівнянь з різнотипним математичним представленням. Запропоновані стійкі методи розв'язування таких системи за час, які забезпечили розв'язування оптимізаційних задач. Приведені результати апробації розробленого програмного забезпечення на реальних даних, які продемонстрували ефективність розробленого алгоритмічного забезпечення.**

***Abstract*—This work contains the problem statements of optimal optimal scheduling operating modes of gas transmission systems with complicated structure models of which have systems of nonlinear equations with different types of mathematical representation. Proposed stable methods for solving such systems, which provided solving optimization problems. There are the results of testing the developed software on real data that demonstrated the efficacy of the developed algorithmic software.**

Ключові слова—Газотранспортна система; оптимальний режим роботи; алгоритми оптимізації; програмний комплекс; компресорна станція.

Keywords—gas transmission system; optimal operating mode; optimization algorithm; software; compressor station.

I. Вступ

У роботі розвиваються методи теорії гідравлічних мереж з активними об'єктами, які використовуються для оптимізації розподілу потоків газу в системі трубопроводів – газотранспортних системах (ГТС). ГТС представляються у вигляді графів без петель, ребрами якого є технологічні об'єкти які володіють протяжністю, тобто мають вхід та вихід, а вершинами – точки їх з'єднання. Початковими чи крайовими умовами для розрахунку є значення тисків та витрати газу на входах та виходах системи. В теорії гідравлічних мереж розвиваються два основних методи гідравлічного розрахунку ГТС – метод вузлових потенціалів і метод контурних витрат [1]. Метод вузлових потенціалів передбачає побудову для графу ГТС системи рівнянь, невідомими в яких є значення тисків у всіх вузлах. Рівняння будують виходячи із необхідності виконання у всіх вузлах графу законів Кірхгофа. Він має обмежену сферу використання, так як в умовах існування об'єктів із суттєво різними пропускними здатностями, проявляються проблеми щодо його коректного використання.

Метод контурних витрат, невідомими в яких є витрата газу по контурах графу, грунтується на забезпеченні виконання одночасно обох законів Кірхгофа. Швидкість збіжності методу залежить від алгоритму виділення дерева

– остова в графі ГТС, який містить обов'язково вузол із заданим тиском. Таке виділення остова диктує однозначний початковий розподіл потоків газу. Він дозволяє суттєво скоротити число невідомих в системі, що приводить до зменшення обчислювальних затрат. Розрахунок лінійних та деревовидних структур проходить за одну ітерацію. В доступній літературі пропонуються методи розв'язування систем нелінійних контурних рівнянь Ньютонівського типу – градієнтні методи [2]. Цей тип методів є чутливим до точності початкового наближення. В [3] було запроновано метод неградієнтного типу, який забезпечив гарантовану збіжність процесу розв'язування задачі за умови коректної її постановки. Важливим є те, що розроблений метод не ставить особливих вимог стосовно початкового наближення та розмірності системи рівнянь.

Розрахунок та оптимізація усталених режимів роботи ГТС використовується для планування та замовлення затратних ресурсів, зокрема енергетичних, так і формування керуючих дій для переходу із поточного в прогнозований оптимальний режим. Керування перехідними процесами забезпечуються системами оптимального керування функціонуванням систем транспортування газу.

Пошук оптимального режиму є надзвичайно складною задачею. Це пов'язано з тим, що оптимальний режим для всієї системи не завжди забезпечується локально оптимальними режимами активних об'єктів та підсистем ГТС. Для ГТС, в реальних умовах неповної завантаженості, виникають проблеми вибору як активних об'єктів, які необхідно задіяти в режимі, так і встановлення їх потужності. Для ГТС України, яка включає 72 КС, 110 цехів, 702 ГПА, серед яких біля 30 різних типів – з газотурбінними, електричними та газомоторними приводами, побудувати оптимальний режим перебірними методами є неможливо. Більше того формування оптимального режиму включає ще вибір точок входу та розподіл між ними прогнозованих об'ємів транзитних та імпортних потоків газу.

Діючі методи, які б знайшли застосування в промислових системах планування розподілу потоків мережевого типу з компресорними станціями – активними об'єктами, з газопроводами із різним номінальним тиском газу в них в літературі є відсутніми [4]. Такі методи розроблялися для газотранспортних систем з простою технологічною схемою – одно- та багатониткових магістральних газопроводів. Оптимізація в однониткових системах магістральних газопроводах в основному проводиться способом вибору тиску близького до максимальних на виходах активних об'єктів. В загальному випадку, в умовах існування непроектних режимів, наявність в системі автоматичних систем підтримки заданих тисків та витрати, методи планування суттєво ускладняються.

II. Газотранспортна система як об'єкт планування

Всі об'єкти, які приймають участь в транспортуванні та зберіганні газу – технологічні об'єкти, об'єднані в єдиний технологічний процес трубопроводами з довжинами від кількох метрів до ста і більше кілометрів та діаметрами які мають від 100 до 1420 мм. Більшість ділянок газопроводів прокладені на певній глибині, а незначна частина проходять над поверхнею землі. Внутрішня поверхня трубопроводів є шершавою, а зовнішня поверхня труби теплопровідною (іде теплообмін із зовнішнім середовищем). Рельєф траси прокладання трубопроводів є змінним і на десятках кілометрів може мінятися до 800 і більше метрів. Робочий тиск в трубах доходить до 7.5 МПа. Компресорна станція (КС) може складатися із декількох цехів. Досить часто цехи відрізняються типом газоперекачуючих агрегатів (ГПА), як за потужністю (від 4.0 до 27.0 МВт) так і типом приводу (електропривід, газотурбінна установка) для відцентрового нагнітача. На багатьох підземних газосховищах працюють компресори поршневого типу. Технологічні схеми окремих КС можуть забезпечити її роботу в дві – три ступені. На ГТС є присутніми велика кількість різнотипної запірної та регулюючої арматури. Частина із них оснащена приводами для їх відкриття та закриття. Робота запірної арматури може моделюватися зміною площі поперечного перетину трубопроводу в зоні їх установки.

Вибір і підтримка найбільш раціонального вихідного режиму газотранспортної системи, при існуючій схемі взаємодії централізованої і розподіленої схеми керування та при наявності стійких відхилень фактичних значень параметрів від заданих розрахункових може здійснюватися способом корекції режиму роботи системи магістральних газопроводів. Результати корекції режиму верхнього рівня є вихідними даними для розрахунку режиму роботи окремих технологічних об'єктів.

Структура системи планування режимів роботи ГТС та його об'єктів включає підсистеми: актуалізації технологічних схем та даних; прогнозування та формування вхідних даних; ідентифікації параметрів моделей; розрахунку інтегральних параметрів режиму; пошуку оптимального розв'язку та представлення результатів моделювання.

Складність ідентифікації параметрів моделей та стану об'єктів пов'язаний із значним рівнем невизначеності режимної інформації, яка, в основному, проявляється у відсутності замірів витрат, тисків і температур в багатьох вузлах газотранспортної системи. Проведений аналіз видів невизначеностей, характерних для процесу управління складними системами, якими є газотранспортні системи, показав: низьку точність і неповноту оперативної інформації яка породжується значною похибкою вимірювання, ненадійною і малопродуктивною системою передачі даних, асинхронністю вимірювання і поступлення, недоступністю до частини інформації тощо. Крім цього неточність моделей пов'язана із допущенням про однорідність певних підсистем, спрощенням, яке пов'язане з неврахуванням в повній мірі рельєфу прокладання трубопроводів введення нечітких моделей типу термогідравлічних еквівалентів, неврахування в повній мірі суттєвої не лінійності та часткового використання паспортних характеристик, тощо).

Розроблені програмні інструменти забезпечують ідентифікацію термогідравлічних параметрів ділянок газопроводів, параметрів ГПА – приводу та ВН, параметрів пластів – колекторів, фільтраційних опорів при вибійних зон свердловин, параметрів запірної та регулюючої арматури [5].

III. Змістовна постановка задач

Задача оптимізації режиму функціонування ГТС ставиться як задача мінімізації цільової функції [6]. Найбільш природнім є критерій мінімізації енерговитрат. Якщо б всі ГПА були оснащені газотурбінними приводами, то критерієм була би мінімізація паливного газу. У випадку однакового к.к.д. у ГПА, то слід очікувати, що витрата паливного газу досягне свого мінімуму при мінімізації сумарної потужності задіяних в режимі ГПА.

Ускладнюють задачу оптимізації наявність в ГТС різних типів приводів – газотурбінні та електричні. Фінансові витрати на паливний газ та електроенергію для забезпечення одиниці потужності таких ГПА є різними та й суттєво вони різняться своїм к.к.д. Нерівномірним є і їх розміщення по системі магістральних газопроводів. Для задач оперативного оптимального планування часто вимагається наявність в певних підсистемах ГТС чи окремих ділянках магістральних газопроводів заданого об'єму акумульованого газу. Це пов'язано з тим, що поточний та прогнозований режим на декілька годин вперед за об'ємами акумульованого газу та його розподілом в системі, в основному, суттєво не різняться. Додатковою умовою для оптимізаційних задач часто служить вимога на формування стійких режимів та умов на перехідні режими – з поточного в прогнозований. До таких умов часто відносять умову на мінімальну кількість переключення силового обладнання на КС. Стійкість режимів ГТС пов'язана із стійкістю роботи окремих ГПА КС і вона оцінюється як за віддаленістю від помпажної зони робочих точок відцентрових нагнітачів так і за запасом по потужності приводів ГПА.

При постановці оптимізаційних задач режими роботи КС вважаються невідомими і для відомої витрати газу по магістральному газопроводі вони представляються як функції тиску на їх входах та виходах. В більш складних випадках режим роботи КС є також і функцією витрати газу.

На всіх входах та виходах системи транспортування газу задаються значення витрати газу і в одній із точок – тиск. Бажано, щоб з певною точністю виконувався і баланс газу – надходження в систему було рівним витокам із системи. На входах та виходах системи задаються, як обмеження, інтервали тисків, частина з яких визначаються контрактними умовами. Можливе задання обмежень на тиски на всіх, або окремих ребрах графу. На ребрах типу КС існують обмеження на тиски - мінімальні на входах та максимальні на виходах. Додатковими обмеженнями на КС можуть бути інтервали зміни внутрішньої потужності газоперекачуючих агрегатів та ступеня стиску газу.

Планування режимів роботи ГТС ґрунтується на певному прогнозі. Точність прогнозу пов'язана з його часовим періодом. Слід очікувати на задовільну точність при плануванні режимів в межах однієї, або декількох діб. При плануванні сезонних режимів користуються середніми прогнозними параметрами, часто і минулорічними. Такі режими формують, як правило, в оптимізаційній постановці.

IV. Планування оптимальних режимів

Практично всі традиційні методи оптимізації неможливо використати чи їх модифікувати для розв'язування сформульованої задачі. Це пов'язано з тим, що параметри функціонування КС мають як дискретні так і неперервні області складної форми які не є наперед фіксованими, а встановлюються на етапі розрахунку параметрів розподілу потоків по системі магістральних газопроводів. І тому забезпечити в процесі ітераційного процесу належність оптимальних параметрів роботи КС робочій області приводів та відцентрових нагнітачів є досить складно.

До системи планування режимів поставлені певні вимоги. Вона повинна забезпечити: планування прогнозованих режимів за заданими критеріями та із врахуванням наявності заданих об'ємів акумульованого газу та його розподілу в системі, високий рівень автоматизації процесу планування, врахування всіх наявних технологічних обмежень, в максимальній мірі врахування вимог диспетчера щодо топології ГТС, надання переваг певному типу технологічного обладнання для включення їх в процес експлуатації тощо.

Питання існування та єдиності розв'язку задач планування режимів для ГТС з компресорними станціями в повній мірі не є розв'язаними. Приведемо деякі необхідні умови існування розв'язку задач. Однією із необхідних умов існування розв'язку системи нелінійних рівнянь буде одна із вказаних нижче.

В кожній зв'язній компоненті граф-схеми ГТС, знайдеться тільки одна вершина, крім вершин які були вершинами входами і виходами КС, в якій є заданим абсолютне значення тиску.

З довільної вершини контуру існує шлях до вершини з абсолютним значенням тиску.

Для довільних двох вершин зв'язної компоненти знайдеться вершина яка належить одному із можливих шляхів між ними з абсолютним значенням тиску.

Підходи до оптимізації ГТС суттєво залежать від структури його підсистем і системи в цілому. Слід виділити такі основні типові підсистеми: КС з однотипними та різнотипними КС, лінійні і дерево видні ГТС, ГТС довільної структури.

Слід зазначити, що автоматизація процесу вибору оптимальної топології ГТС, як математична задача, в повній мірі для такої складної системи як ГТС України, є ще розв'язаною в неповній мірі. Вона добре автоматизується для багато ниткових газопроводів з мінімальним перебором варіантів, якщо попередньо проведене дослідження області оптимальності їх роботи.

Відомо, що для ГТС без активних об'єктів при коректній постановці задачі існує єдиність розв'язку, а початковий розподіл впливає тільки на швидкість збіжності методу. В умовах існування ГТС з активними об'єктами розрахунок режимів для різних початкових розподілів приводить до різних параметрів потоків і, відповідно, до різних паливно-енергетичних затрат. З точки зору оптимізації ГТС початковий розподіл грає важливу роль. На даний час не існує рекомендацій для ГТС із складною технологічною схемою щодо побудови оптимального початкового розподілу потоків і тому без організації переборів в таких випадках не обійтися. Задача полягає тільки в їх мінімізації.

Розв'язування оптимізаційних задач дискретними методами цільового перебору вимагають обмеження на критерій оптимальності – виконання властивості адитивності цільової функції. В цьому випадку можна суттєво зменшити складність перебірних методів. Для зменшення переборів та побудови швидких методів пошуку оптимальних режимів запропоновано декілька підходів. У всіх підходах вимагається представлення структури ГТС (рис.1) як об'єднання більш простих структур (рис.2). Для простих лінійних структур розроблені швидкі перебірні алгоритми побудови оптимального режиму (рис.3). Наступний крок – узгодження режимів простих структур, який також вимагає певного перебору варіантів. Чим простіші структури є в основі представлення графу ГТС тим більше переборів прийдеться провести для узгодження їх режимів. Тому розроблені швидкі методи оптимізації більш складних структур - деревовидних та променевих. Для більш швидкого розрахунку простих структур запропоновано так званий «ощадливий» алгоритм, який використовує певні знання фізики роботи систем з активними об'єктами, які суттєво зменшують складність алгоритмів. Різні підходи дають можливість оцінити якість результату оптимізації.

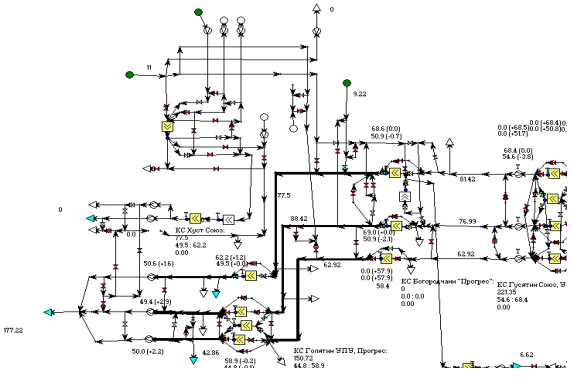


Рис. 1. Фрагмент технологічної схеми

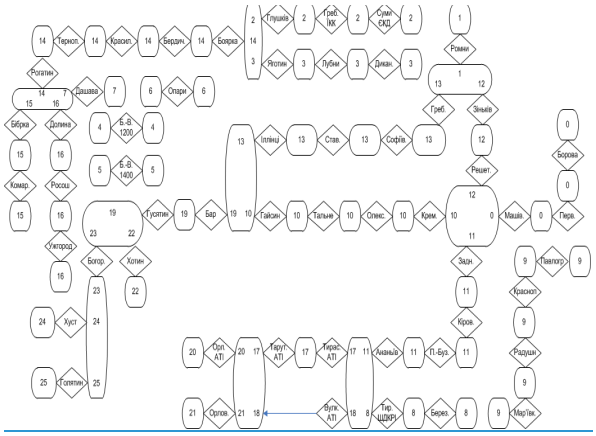


Рис. 2. Представлення технологічної схеми для розрахунку оптимального режиму

V. Висновки

Система актуалізація даних разом з системою знаходження оптимальних режимів роботи основних об'єктів ГТС та системи загалом забезпечила максимальну автоматизацію процесу розв'язування оптимізаційних задач з мінімальним втручанням користувача. Розроблений програмний комплекс дає можливість редагувати існуючі сценарії роботи ГТС, створювати нові та проводити аналіз сценаріїв, які були сформовані іншими користувачами. Дана розробка забезпечить ефективне розв'язання задач формування оптимальних параметрів керування газотранспортною системою.

Як показали проведені числові експерименти, що одночасне забезпечення в системі умови на завантаження на максимальну продуктивність перших по ходу руху газу КС та мінімізація їх кількості може привести до оптимальності режиму. Підтримка максимально можливого тиску в системі у весняно-літній період вимагає використання незначних об'ємів активного газу газосховищ, який є завжди в наявності, і цей потенціал оптимізації на даний час використовується не в повній мірі. У зимовий період зовсім не використовується потенціал оптимізації, який забезпечується в процесі керування тепловим режимом транспортування газу.

Література References

[1] А. П. Меренков, В. А. Хасилев, Теория гидравлических цепей. М.: Наука, 1965.

[2] А. Г. Евдокимов, А. Д. Тевяшев, Оперативное управление потокораспределением в инженерных сетях. Харьков: Вища школа, 1980.

[3] Н. Притула, "Розрахунок параметрів потокорозподілу в газотранспортній системі (стаціонарний випадок)", *Фіз.-мат. мод. та інформ. техн.,* вип. 5, С. 146–157, 2007.

[4] В. С. Кулик, "Алгоритм оптимизации транспорта газа через разветвлённую ГТС", *Труб. транспорт: теория и практика,* № 2, С. 22–25, 2014.

[5] С. Гладун, Н. Притула, Б. Землянський, О. Химко, "Розрахунок гідродинамічних параметрів стану об'єктів транспорту газу", *Вісник НУ «Львівська політехніка»: Комп. науки та інформ. техн.,* № 629, С. 174–201, 2008.

[6] І. Боярин, Р. Боровий, О. Гринів, Н. Притула, В. Ямнич, "Оптимізація роботи багатоцехових компресорних станцій з різнотипними газоперекачуючими агрегатами", *Вісник НУ «Львівська політехніка»: Комп. науки та інформ. техн.,* № 672, С. 326–335, 2010.

Методи Розв'язування Крайових Задач з Використанням Дробових Похідних за Часом

Ярослав П'янило
Відділ математичних методів обчислювального експерименту Центру математичного моделювання Інституту прикладних проблем механіки і математики ім. Я. С. Підстригача НАН України
Львів, Україна
danylo794@gmail.com

Олег Браташ
Відділ математичних методів обчислювального експерименту Центру математичного моделювання Інституту прикладних проблем механіки і математики ім. Я. С. Підстригача НАН України
Львів, Україна
olebra31@gmail.com

Галина П'янило
Відділ математичних методів обчислювального експерименту Центру математичного моделювання Інституту прикладних проблем механіки і математики ім. Я. С. Підстригача НАН України
Львів, Україна
danylo794@gmail.com

Валентина Собко
Інститут прикладних проблем механіки і математики ім. Я. С. Підстригача НАН України
Львів, Україна
vg_sobko@ukr.net

Methods for Boundary Value Problems Solving Using Fractional Time Derivatives

Yaroslav Pyanylo
Department for Mathematical Methods of Computing Experiment in Center of Mathematical Modeling of Pidstryhach Institute for Applied Problems of Mechanics and Mathematics, National Academy of Sciences of Ukraine
Lviv, Ukraine
danylo794@gmail.com

Oleh Bratash
Department for Mathematical Methods of Computing Experiment in Center of Mathematical Modeling of Pidstryhach Institute for Applied Problems of Mechanics and Mathematics, National Academy of Sciences of Ukraine
Lviv, Ukraine
olebra31@gmail.com

Halyna Pyanylo
Department for Mathematical Methods of Computing Experiment in Center of Mathematical Modeling of Pidstryhach Institute for Applied Problems of Mechanics and Mathematics, National Academy of Sciences of Ukraine
Lviv, Ukraine
danylo794@gmail.com

Valentyna Sobko
Pidstryhach Institute for Applied Problems of Mechanics and Mathematics, National Academy of Sciences of Ukraine
Lviv, Ukraine
vg_sobko@ukr.net

***Анотація*—в даній роботі побудовано спектральний метод дослідження математичних моделей фізичних процесів з використанням похідних за часом дробових порядків.**

***Abstract*—The spectral method for investigation of mathematical models of physical processes using fractional time derivatives are constructed in this paper.**

Ключові слова—спектральні методи, дробові похідні, фільтрація газу, пористі середовища

Keywords—spectral methods, fractional derivatives, gas filtration, porous media

I. Вступ

Процеси руху газу в технологічних об'єктах в системах його транспорту описують, як правило, диференціальними рівняннями в частинних похідних. За теперішніх умов задання вхідної інформації (стосовно її точності та кількості) таких підхід до математичного моделювання забезпечує розв'язання прикладних задач. Однак в деяких випадках є необхідність в штучному заданні деяких параметрів для того, щоб забезпечити фізичну відповідність математичної моделі фізичному процесові. Це, зокрема, має місце у випадку побудови математичної моделі процесу фільтрації газу в пластах підземних сховищ, в яких значення тиску в кожній точці в заданий момент часу залежить від значення його в попередні значення часів. Одним із підходів до розв'язання цього питання є застосування похідних дробових порядків.

Метою роботи є побудова спектрального методу дослідження математичних моделей фізичних процесів, в яких для розв'язування задач математичної фізики використовуються похідні дробових порядків.

II. Визначення дробових похідних

Оператор дробової похідної у термінах Капуто визначається так [6-12]:

$$ {}^{c}D_{\tau}^{\alpha} = \frac{{}^{c}\partial^{\alpha}}{\partial\tau^{\alpha}}\varphi(\tau) := \frac{1}{\Gamma(m+1-\alpha)}\int_{0}^{\tau}\frac{\frac{\partial^{m+1}}{\partial\xi^{m+1}}\varphi(\xi)}{(\tau-\xi)^{\alpha-m}}d\xi , \qquad (1) $$

де $m=[\alpha]$; $[\cdot]$ – ціла частина дійсного числа, а в термінах Ріммана-Ліувіля –

$$ D_{\tau}^{\alpha} = \frac{{}^{c}\partial^{\alpha}}{\partial\tau^{\alpha}}\varphi(\tau) := \frac{1}{\Gamma(m+1-\alpha)}\frac{\partial^{m+1}}{\partial\xi^{m+1}}\int_{0}^{\tau}\frac{\varphi(\xi)}{(\tau-\xi)^{\alpha-m}}d\xi . \qquad (2) $$

Між похідними Капуто і Ріммана-Ліувіля має місце наступний зв'язок [12]

$$ {}^{C}D_{\tau}^{\alpha}\varphi = D_{\tau}^{\alpha}\varphi - \sum_{k=0}^{m}\frac{\tau^{k-\alpha}}{\Gamma(k-\alpha+1)}\frac{\partial^{k}}{\partial\tau^{k}}\varphi . \qquad (3) $$

III. Методи розв'язування

1. Аналітичні.

Інтегральні перетворення

Якщо $\Im$ оператор перетворення Лапласа-Карсона, то має місце рівність

$$ \Im\left(D_{a+}^{\alpha}f(t)\right) == s^{\alpha}\left[F(s)-f(0)\right], $$

де $F(s)$ - зображення Лапласа-Карсона оригіналу $f(t)$.

Для знаходження оригіналів зображень з параметром перетворення s^{α} використовується теорема Ефроса.

Теорема Єфроса. Якщо оператор $\bar{a}(p)$ можна подати у вигляді $\bar{a}(p) = \dfrac{\bar{\Phi}(q(p))}{pq(p)}$, причому:

◊) $\bar{\Phi}(p) = p\int_{0}^{\infty}\Phi(t)e^{-pt}dt$ $\int◊_{0}^{\infty}\left|\Phi(t)\right|e^{-\gamma_0 t}dt < \infty$, ™$\sum$ ©$_0$ >0;

б) $q(p)$ аналітична в півплощині $Re\ p>\gamma_0\geq 0$ функція яка задовільняє в півплощині умову $Re\ q(p)\geq\gamma_0$, тоді оператор $\bar{a}(p)$ зводиться до функції $\bar{a}(p) = \int_{0}^{\infty}L(t,\xi)\Phi(\xi)d\xi$, де

$$ L(t,\xi) = \frac{1}{p}e^{-\xi q(p)} . $$

Числові. Схема Грюнвальда-Летнікова.

Розклад дробової похідної $\dfrac{\partial^{\alpha}p}{\partial t^{\alpha}}$ за схемою Грюнвальда-Летнікова:

$$ {}^{GL}D_{\tau}^{\alpha}p := \lim_{\Delta t\to 0}(\Delta t)^{-\alpha}\sum_{j=0}^{[\tau/\Delta t]}(-1)^{j}\binom{\alpha}{j}p(\tau-j\Delta t) . $$

Оператор Грюнвальда-Летнікова (11) апроксимується на проміжку $[0, \tau]$ з підінтервальним кроком Δt як

$$ {}^{GL}D_{\tau}^{\alpha}p(\tau) \approx \sum_{j=0}^{[\tau/□t]}c_{j}^{(\alpha)}p(\tau-j□t) , $$

де $c_j^{(\alpha)}$ – коефіцієнти Грюнвальда-Летнікова, які визначаються як

$$ c_{j}^{(\alpha)} = (\Delta t)^{-\alpha}(-1)^{j}\binom{\alpha}{j} . $$

Використавши рекурентне співвідношення [6]

$$ c_{j}^{(\alpha)} = (\Delta t)^{-\alpha}, \quad c_{j}^{(\alpha)} = \left(1-\frac{1+\alpha}{j}\right)c_{j-1}^{(\alpha)} , $$

ми можемо обчислити коефіцієнти $c_j^{(\alpha)}$. Для $j = 1$ маємо $c_1^{(\alpha)} = -\alpha(\Delta t)^{-\alpha}$.

2. Застосування квадратурних формул. Метод скінченних елементів. Особливості застосування

До числа відкритих і невирішених завдань, пов'язаних з дробовим обчисленням слід віднести:

створення ефективних методів, алгоритмів і програм вирішення нелінійних інтегро-диференціальних рівнянь нецілого порядку;

розвиток числових методів, адаптованих до вирішення практичних завдань та їх реалізація;

розробка методів вирішення та комп'ютерного моделювання інтегро-диференціальних рівнянь змінних нецілих порядків із заданими початковими і крайовими умовами;

розвиток методів структурної і параметричної ідентифікації динамічних систем, математичні моделі яких містять інтегро-диференціальні оператори нецілих порядків, тощо.

Значний вплив має значення параметру порядку похідної (задача Даніловської).

IV. Модель фільтрації газу в пористому середовищі

Процес масопереносу в пористих середовищах розглядається на прикладі фільтрації газу та рідини, яка описується рівнянням із дробовою похідною за часовою змінною [2-4]

$$\frac{\partial}{\partial x}\left(\frac{kh}{\mu\chi}\frac{\partial p^l}{\partial x}\right)+\frac{\partial}{\partial y}\left(\frac{kh}{\mu\chi}\frac{\partial p^l}{\partial y}\right)+\frac{\partial}{\partial z}\left(\frac{kh}{\mu\chi}\frac{\partial p^l}{\partial z}\right)=2mh\left(\frac{\partial^\alpha}{\partial t^\alpha}\left(\frac{p}{\chi}\right)+2qp_{at}\right). \quad (4)$$

В останньому рівнянні $l=2$ для газу та l=1 для нестисливої рідини; $\alpha\epsilon(0,2]$ – степінь дробової похідної; $k=k(x,y,z,t)$, $m=m(x,y,z)$ та $h=h(x,y,z)$ коефіцієнти проникності, пористості та товщина середовища відповідно; μ – динамічна в'язкість речовини, p_{at} – атмосферний тиск, q – густина відбору, χ – коефіцієнт стисливості газу.

Рух газу в трубопроводах з використанням дробових похідних порядку α за часом, описується системою

$$\frac{\partial^\alpha}{\partial t^\alpha}\omega(x,t)+\frac{\partial p}{\partial x}+a\omega-bp=\Theta(x,t),$$

$$\frac{\partial\omega}{\partial x}+\frac{1}{c^2}\frac{\partial^\alpha}{\partial t^\alpha}p(x,t)=\Psi(x,t).$$

Якщо $k=0$, то отримуємо похідну Рімана-Ліувіля, а при $k=1$ – Капуто.

V. Обчислювальний експеримент. Модель фільтрації газу в пористому середовищі

Задача полягає в знаходженні розв'язку $p=p(x,y,z,t)$ рівняння (4) за відомими значеннями тиску $p(x_i, y_i, z_i, t_i)$ у заданих точках середовища та умовою непроникності на контурі середовища. При цьому необхідно, щоб виконувалась умова балансування маси газу в сховищі

$$M=\int_V \rho\, dv.$$

Інтегрування проводиться по об'єму сховища V, M – маса газу в сховищі, ρ – густина газу, яка пов'язана з тиском рівнянням стану $p=\rho\chi RT$. Тут R – газова стала, T – абсолютна температура газу.

Подамо функції $p(x,t)$ та $\omega(x,t)$, які входять у розв'язок задачі, у ряд Фур'є за многочленами Лагерра $L_m(t)$.

$$p(x,t)=\sum_{m=0}^{\infty}p_m(x)L_m(t),\quad \omega(x,t)=\sum_{m=0}^{\infty}\omega_m(x)L_m(t). \quad (5)$$

Використовуючи формулу (6) та розклавши функцію $k(t)$ яка є ядром інтегралу, у ряд Фур'є за многочленами $L_m(t)$, отримаємо

$$\frac{d}{dt}\int_0^t k(t-\tau)p(\tau)d\tau=\sum_{n=0}^{\infty}k_n\sum_{m=0}^{\infty}p_mL_{n+m}(t)=\sum_{n=0}^{\infty}c_nL_n(t), \quad (6)$$

$$\begin{cases}\dfrac{1}{\Gamma(1-\alpha)}\displaystyle\sum_{n=0}^{\infty}c_n(x)L_n(t)+\sum_{n=0}^{\infty}p'_n(x)L_n(t)+\\ +a\displaystyle\sum_{n=0}^{\infty}\omega_n(x)L_n(t)-b\sum_{n=0}^{\infty}p_n(x)L_n(t)=0,\\ \displaystyle\sum_{n=0}^{\infty}\omega'_n(x)L_n(t)+\frac{1}{c^2}\frac{1}{\Gamma(1-\alpha)}\sum_{n=0}^{\infty}d_n(x)L_n(t)=0.\end{cases} \quad (7)$$

В останній формулі

$$c_n(x)=\sum_{m=0}^{n}k_m\omega_{n-m}(x)=\sum_{m=0}^{n}\omega_m(x)k_{n-m},$$

$$d_n(x)=\sum_{m=0}^{n}k_mp_{n-m}(x)=\sum_{m=0}^{n}p_m(x)k_{n-m}.$$

Розв'язок останньої системи будемо шукати у вигляді експоненціальних рядів Фур'є виду

$$p_j(x)=\sum_{n=-\infty}^{\infty}p_{jn}(\tau)e^{n\pi ix/l},\qquad x\in(0,l), \quad (8)$$

де

$$p_{jn}=\frac{1}{l}\int_0^l p_j(x)e^{-n\pi ix/l}\,dx.$$

Запропонована схема розв'язування диференціального рівняння апробована в ході обчислювального експерименту на основі даних пористого середовища, площею $S=16$ *млн* $м^2$, за наступних початкових параметрів:$\mu=0{,}000011$ Па·с, $h=18{,}2$ м, $R=506{,}7$ $Дж/кг\,°K$, $T=293\ °K$, $z=0{,}87$, $m=0{,}31$, $k=1{,}8e-12$ $м^2$.

На рис. 1-4 приведені обчислені значення тисків газу в околі свердловин (рис. 1), в примежовій зоні пласта (рис. 2), в середині області пласта (рис. 3) та середньопластових тисків (рис. 4) для різних значень порядку дробової похідної α.

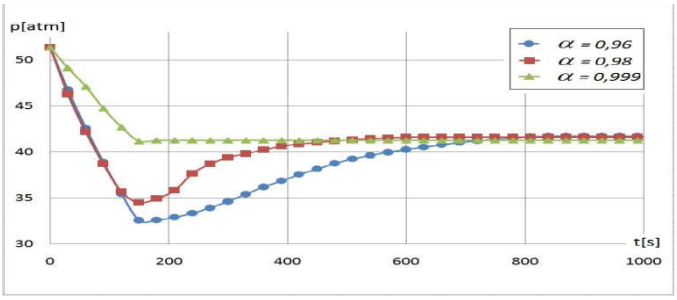


Рис. 1. Значення тисків газу в околі свердловин для різних значень порядку дробової похідної α.

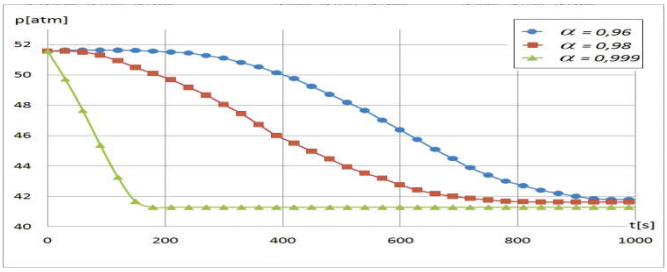


Рис. 2. Значення тисків газу в примежовій зоні пласта для різних значень порядку дробової похідної α.

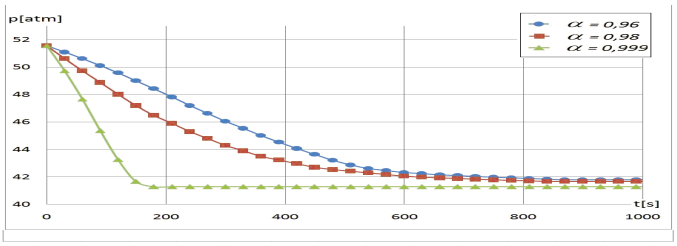


Рис. 3. Значення тисків газу в середині області пласта для різних значень порядку дробової похідної α.

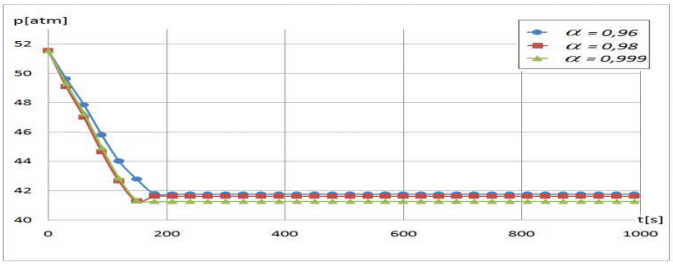


Рис. 4. Значення середньопластових тисків для різних значень порядку дробової похідної α

VI. Висновок

Застосування спектральних многочленів дає можливість побудувати ефективний автоматизований алгоритм для розв'язування крайових задач. Зокрема, застосування многочленів Лагерра дозволяє згортку двох функцій звести до сумування згортки рядів. Тому уникається використання процедури дискретизації, яка вносить значну похибку в процес обчислень. Аналіз результатів, поданих в рисунках, показує, що порядок дробової похідної має значний вплив на розв'язок вихідної задачі. Звідси випливає необхідність апріорної інформації для визначення порядку дробової похідної при моделювання фізичних процесів. Разом з тим, як випливає з обчислювального експерименту модельної задачі, застосування спектрального методу для задач такого типу дозволяє знаходити параметричні подання інтегральних перетворень, ядрами яких є многочлени Лагерра.

Література References

[1] Самко С. Г., Килбас А. А., Маричев О. И. Интегралы и производные дробного порядка и некоторые их приложения. — Минск: Наука и техника, 1987. — 688 с.

[2] Васильев В. В., Симак Л. А. Дробное исчисление и аппроксимационные методы в моделировании динамических систем. — Киев: Научное издание НАН Украины, 2008 — 256 с.

[3] Псху А. В. Уравнения в частных производных дробного порядка. — Научно-исследовательский ин-т приклад, математики и автоматизации Кабардино-Балкарского научного центра РАН. — Москва: Наука, 2005. — 199 с.

[4] Нахушев А. М. Дробное исчисление и его применение. — Москва: Физматлит, 2003. — 272 с.

[5] Нахушева В. А. Некоторые классы дифференциальных уравнений математических моделей нелокальных физических процессов. — Нальчик: КБНЦ РАН, 2002. — 100 с.

[6] Диткин В. А., Прудников А. П. Операционное исчисление. — Москва: Высшая школа, 1975.— 407 с.

[7] Диткин В. А., Прудников А. П. Справочник по операционному исчислению. — Москва: Высшая школа, 1965. — 466 с.

[8] П'янило Я., Васюник М., Васюник І. Використання многочленів Лагерра до спектрального методу розв'язування рівнянь у дробових похідних за часом // Фізико-математичне моделювання та інформаційні технології. – 2013. – Вип.17. – С. 163-167

[9] П'янило Я.Д., Собко В.Г. Побудова та дослідження біортогональних поліномів на базі многочленів чебишева . – Прикл. проблеми мех.. і мат. – 2013.-Вип. 11.- С. 181-189.

Модель Неусталеного Руху Крові в Судинах з Врахуванням Пружності

Ярослав П'янило
Відділ математичних методів обчислювального експерименту Центру математичного моделювання Інституту прикладних проблем механіки і математики ім. Я. С. Підстригача НАН України
Львів, Україна
danylo794@gmail.com

Володимир Череватий
кафедра диференціальних рівнянь та прикладної математики
Прикарпатський національний університет ім. В. Стефаника
Івано-Франківськ, Україна
volodia4392@gmail.com

The Unstabled Blood Flow Model in the Vessels Considering the Walls Elasticity

Yaroslav Pyanylo
Department for Mathematical Methods of Computing Experiment in Center of Mathematical Modeling of Pidstryhach Institute for Applied Problems of Mechanics and Mathematics, National Academy of Sciences of Ukraine
Lviv, Ukraine
danylo794@gmail.com

Volodymyr Cherevatyy
dept. of Computer Science
Precarpathian National University
Ivano-Frankovsk, Ukraine
volodia4392@gmail.com

***Анотація*—В праці розглядається неусталений рух крові в судинах з врахуванням пружності стінок. Вивчається модель руху крові у нестаціонарному випадку. Показано, що розподіл тиску в крові залежить як від параметрів судини, так і від моментів часу.**

***Abstract*—In the labor considered movement blood in the vessels taking into account the elastic walls. We study the model of blood flow in the unsteady case. It is shown that the distribution of blood pressure depends on the parameters of the vessel and from points in time.**

Ключові слова—математична модель, нестаціонарний рух крові, медоди розв'язування, еластичність трубок

Keywords—mathematical model, unsteady movement of blood medody solving, flexibility tubes

I. Вступ

На практиці досить часто є необхідність в розрахунку параметрів руху рідини в гнучких трубках, зокрема судинах крові. Внутрісудинний тиск крові є одним з основних параметрів, за яким судять про функціонування серцево-судинної истеми [3-5]. Між артеріальним тиском, об'ємною швидкістю крові та опором судини існує певна функціональна залежність. Очевидно, що поряд з цими параметрами на процес руху крові мають вплив і інші параметри, зокрема, сила гравітації, еластичність судин, траєкторія руху, тощо. Відмінною особливістю характеристики серцево-судинної системи є вимога обчислювати всі складові параметри в кількісному виді. Для цього будуються адекватні математичні моделі процесу руху крові в судинах. Однією із задач, які можуть бути розв'язані на базі математичних моделей, є введення лікарських препаратів

Формулювання задачі. В цей час в літературі відомо багато математичних моделей руху крові в живих організмах. Всі вони мають місце в деяких просторово-часових інтервалах і з різною адекватністю описують цей процес. Одним із таких підходів може бути застосування математичних моделей руху газів або рідин в трубопроводах. Однак особливістю процесу руху крові в судинах живих організмів є те, що вони мають невелику довжину та їх стінки є еластичними, що необхідно враховувати при розрахунку відповідних параметрів. Однією з таких моделей може бути наступна.

II. Модель усталеного руху рідини за наявності джерел та відборів.

З врахуванням сили тертя і впливу сили тяжіння розподіл тиску рідини в трубках за усталеного руху та усереднених параметрами можна описати диференціальним рівнянням

$$\frac{dp}{dx}+\frac{\lambda\rho q^2}{2DA^2}+\rho g\frac{dh}{dx}=0, \qquad (1)$$

де $p=p(x)$ — розподіл тиску вздовж трубки; ρ — густина рідини, D — внутрішній діаметр трубки; x — біжуча координата $x\in[0,l]$, де l — довжина трубки; g — прискорення вільного падіння; $h=h(x)$ — крива, що описує рельєф трубки (траєкторію руху рідини); $q=\upsilon A$ об'ємна витрата рідини (υ — швидкість); $A=\pi D^2/4$; λ - гідравлічний опір.

1.1. Випадок неперервного рівномірного відбору. Нехай q_n - потік рідини, що входить, а q_k - потік, виходить з трубки. Тоді, $\Delta q=q_n-q_k$ - кількість рідини, яка відбирається вздовж трубки. Будемо вважати, що $q(x)=q_n-\Delta q\,x/l$. Тоді диференціальне рівняння (1) буде мати вигляд

$$\frac{dp}{dx}+\frac{\lambda\rho(q_n-\frac{x}{l}\Delta q)^2}{2DA^2}+\rho g\frac{dh}{dx}=0. \qquad (2)$$

З рівняння (2) отримується розподіл тиску вздовж трубки у випадку рівномірного відбору рідини, який обчислюється за формулою

$$p(x)-p_0=\frac{l\lambda\rho}{6DA^2\Delta q}\left[(q_n-\frac{x}{l}\Delta q)^3-q_n^3\right]-\rho g(h(x)-h_0). \qquad (3)$$

1.*2. Випадок зосереджених відборів.*

Нехай вздовж трубки довжини l є I - зосереджених відборів рідини маси m_i в точках x_i. Розподіл тиску $p(x)$ на кожному з проміжків обчислюється за формулою [1]

$$p(x)-p_0=-\frac{\lambda m^2}{2DA^2}\frac{x}{\rho_c}-\rho_c g h(x). \qquad (4)$$

Використання формули (4) для визначення кінцевих тисків на кожному i - му проміжку веде до наступної формули для обчислення розподілу тиску

$$p_n-p_0=-\frac{1}{2\rho_c}\sum_{i=1}^{n}\frac{\lambda_i m_i^2 l_i}{D_i A_i^2}-g\sum_{i=1}^{n}\rho_i h_i.$$

Остання формула дає можливість змінювати геометричні розміри трубки та параметри рідини на кожному з проміжків. Якщо вважати, що гідравлічний опір λ, густина ρ та діаметр D трубки на кожному проміжку приймають ті самі значення, то

$$p_n-p_0=\frac{\lambda}{2\rho_c DA^2}\sum_{i=1}^{n}m_i^2 l_i+\rho g\sum_{i=1}^{n}h_i.$$

Якщо задано величини масових відборів в кожному з вузлів, тобто вузлу x_i відповідає відбір $-\delta_i$, то

$$p_n-p_0=\frac{1}{2\rho_c}\sum_{i=1}^{n}\frac{\lambda_i l_i}{D_i A_i^2}(m_0-\sum_{j=1}^{i}\delta_j)^2+g\sum_{i=1}^{n}\rho_i h_i.$$

1.3. Неперервний відбір з врахуванням залежності гідравлічного опору від чисел Рейнольдса $\mathrm{Re}=\upsilon_c D/\nu$, ν - кінематична в'язкість, υ_c - середнє значення лінійної швидкості руху рідини на $[0,x]$.

У випадку ламінарного руху $\lambda=\lambda_0\mu A/\rho q d$, де λ_0 - деяка стала, яка визначається експериментально, та

$$\frac{\lambda\rho q^2}{2DA^2}=\frac{\rho q^2}{2DA^2}\cdot\frac{\lambda_0\mu A}{\rho q d}=\frac{\lambda_0\mu}{2DA}q.$$

У випадку рівномірного відбору

$$\frac{\lambda\rho q^2}{2DA^2}=\frac{\lambda_0\mu}{2DA}\left[q_n-\frac{x}{l}\Delta q\right].$$

Таким чином, якщо врахувати залежність числа Рейнольдса від величини потоку, то отримаємо рівняння

$$\frac{dp}{dx}+\frac{\lambda_0\mu}{2DA}\left[q_n-\frac{x}{l}\Delta q\right]+\rho g\frac{dh}{dx}=0,$$

розв'язок якого має вигляд

$$p(x)-p_0=\frac{\lambda_0\mu}{2DA}\cdot\frac{l}{2\Delta q}\left[(q_n-\frac{x}{l}\Delta q)^2-q_n^2\right]-\rho g(h(x)-h_0).$$

III. Вплив параметрів судини на процес усталеного руху крові в усталеному режимі.

Постановка задачі. В рівнянні (1) перейдемо від швидкості руху крові до об'ємної витрати q за формулою $q=\upsilon A$, де $A=\pi D^2/4$. Тоді отримаємо

$$dp=-\frac{\lambda m^2}{2D\rho A^2}dx-\rho g dh. \qquad (5)$$

В рівнянні (5) $m=\rho q$ - масова витрата крові. Оскільки кров вважається нестисливою рідиною, то з рівняння (5) отримаємо

$$p(x)-p_0=-\frac{\lambda m^2}{2DA^2}\frac{x}{\rho}-\rho gh(x) \qquad .(6)$$

В літературі модель процесу руху крові в судинах в усталеному випадку за сталого поперечного перерізу описується рівнянням

$$\frac{dp}{dx}+\frac{2\pi\upsilon r}{\delta}\frac{m}{A^2}+\frac{m^2}{\rho A}\frac{d}{dx}\frac{1}{A}. \qquad (7)$$

Інтегрування рівняння (7) дає співвідношення

$$p_0-p(x)=\frac{2\pi\upsilon r}{\delta}\frac{m}{A^2}x. \qquad (8)$$

В рівняннях (7) та (8) позначено: $D=2r$, δ - товщина судини, ν - кінематична в'язкість крові.

Очевидно, що на рух крові впливає як барометричний перепад тиску, так і деформація судин. Тому доцільно побудувати математичну модель руху крові в судинах з врахуванням згаданих факторів.

Розв'язування задачі. Скористаємося рівнянням нерозривності

$\rho\upsilon A\equiv const$. (9)

Очевидно, що $A=\pi r^2$ та $A_0=\pi r_0^2$. Якщо враховувати пружні властивості судин, то зміна їх радіуса в залежності від тиску крові виражається формулою :

$$r=r_0\left(1+\frac{r_0}{E\delta_0}p\right). \qquad (10)$$

Тут δ_0 та r_0 - товщина і радіус судини в початковому стані. З рівності (9) отримуємо $\upsilon_0 A_0=\upsilon A$, звідки

$$\upsilon=\upsilon_0\frac{A_0}{A}=\upsilon_0\frac{\pi r_0^2}{\pi r^2}=\frac{\upsilon_0}{(1+\frac{r_0}{E\delta_0}p)^2}. \qquad (11)$$

Якщо позначити $\beta=\frac{r_0}{E\delta_0}$, тоді рівність (11) запишеться:

$$\upsilon=\upsilon_0(1+\beta p)^{-2}. \qquad (12)$$

Продиференціюємо рівняння (9) і отримаємо наступне співвідношення:

$$d\upsilon/dx=-2\beta\upsilon_0(1+\beta p)^{-3}\,dp/dx. \qquad (13)$$

З рівняння (1), враховуючи рівність (12), будемо мати:

$$\left(1-\frac{2\beta\rho_0\upsilon_0}{(1+\beta p)^3}\right)\frac{dp}{dx}+\frac{\lambda\rho}{2D}\frac{\upsilon_0^2}{(1+\beta p)^4}+\rho g\frac{dh}{dx}=0. \qquad (14)$$

Введемо позначення:

$$\psi_1=\left(\frac{\lambda\rho}{2D}\frac{\upsilon_0^2}{(1+\beta p)^4}\right)\left(1-\frac{2\beta\rho_0\upsilon_0}{(1+\beta p)^3}\right)^{-1};\ \psi_2=\rho g\left(1-\frac{2\beta\rho_0\upsilon_0}{(1+\beta p)^3}\right)^{-1}.$$

Тоді рівняння (14) буде запишеться так:

$$dp/dx+\psi_2\,dh/dx=-\psi_1. \qquad (15)$$

Рівняння (12) є нелінійним за тиском. Для його розв'язування можна використати числові методи, або побудувати ітераційний алгоритм, який полягає в наступному. У величинах ψ_1 та ψ_2 значення тиску будемо вважати постійним, рівним значенню, знайденому на попередньому кроці. Тоді отримаємо рівняння з відокремленими змінними

$$d(p+\psi_2 h)=-\psi_1 dx,$$

розв'язок якого є

$$p(x)=p_0-\psi_2 h(x)-\psi_1 x. \qquad (16)$$

Алгоритм розв'язування такого рівняння наступний.

1. Для початкового значення тиску p_0 обчислюються величини ψ_1 та ψ_2.
2. Для початкового значення x_z визначається відповідне значення тиску $p(x_z)$.
3. Уточнюються величини ψ_1 та ψ_2 та визначається уточнене значення тиску.
4. Процес продовжується до того часу, поки різниця між двома сусідніми ітераціями буде менше заданої величини.

IV. Вплив параметрів судини на 424роцесс усталеного руху крові в неусталеному режимі.

В ізотермічному випадку поширеною математичною моделлю руху крові в судині є система взаємопов'язаних диференціальних рівнянь у частинних похідних [5,6]

$$\begin{cases}\dfrac{\partial p}{\partial x}+\upsilon\alpha\dfrac{\partial}{\partial x}\left(\dfrac{\upsilon^2}{2}\right)+\rho g\dfrac{\partial h}{\partial x}+\dfrac{\lambda\rho\upsilon^2}{2D}+\dfrac{\partial(\rho\upsilon)}{\partial t}=0,\\ \dfrac{\partial(\rho\upsilon)}{\partial x}+\dfrac{1}{c^2}\dfrac{\partial p}{\partial t}=\Psi(x,t).\end{cases} \qquad (17)$$

Тут ρ, v, p – відповідно, густина, швидкість руху і тиск крові; λ – коефіцієнт гідравлічного опору; D – діаметр

судини; t – час; x – біжуча координата, $x \in [0, L]$; L – довжина судини; c – швидкість звуку в рідині. Функція

$$\Psi(x,t) = \frac{4}{\pi D^2} \sum_{j=1}^{J} \varpi_j(t)\, \delta\ (x - x_j)\left[\eta\ (t - t_{1i}) - \eta(t - t_{2i})\right] \quad (18)$$

моделює наявність на вздовж судини в точках $x_i\ \left(i = \overline{1,J}\right)$ наявність відводів–закачування деяких речовин, t_{1i} та t_{2i} – відповідно часи включення та виключення масових відводів, $\eta(t)$ – одинична функція Хевісайда, $\delta(x)$ – дельта–функція Дірака. Будемо вважати кров нестискуючою рідиною. Очевидно, що на рух крові впливає як барометричний перепад тиску, так і деформація судин. Тому доцільно побудувати математичну модель руху крові в судинах з врахуванням згаданих факторів.

Розв'язування задачі. Для цього використаємо рівняння нерозривності

$$\rho \upsilon s = c \equiv const\ .$$

Тут s - площа поперечного перерізу судини. Очевидно, що вона залежить від тиску крові і для її обчислення можна використати емпіричну формулу

$$s = s(p) = s_0(1 + \beta(p - p_0))\ .$$

Тут h та D_0 - товщина судини і її діаметр, E - модуль Юнга.

$$\beta = D_0 / Eh$$

Індексом 0 будемо позначати стан крові в діастолічній фазі роботи серця. Оскільки

$$\frac{\partial}{\partial x}\left(\frac{1}{s}\right) = \frac{1}{s_0}\frac{\partial}{\partial x}\left(\frac{1}{1 + \beta(p - P_d)}\right) = \frac{\beta}{s_0\left(1 + \beta(p - P_d)\right)^2}\frac{\partial p}{\partial x}\ ,$$

то рівняння нерозривності

$$\partial(\rho\upsilon)/\partial x + c_z^{-2}\,\partial p/\partial t = \Psi(x,t)$$

Буде мати вигляд

$$q_1 \frac{\partial p}{\partial x} + \frac{1}{c_z^2}\frac{\partial p}{\partial t} = \Psi(x,t)\ .\quad q_1 = \frac{c\beta}{s_0\left(1 + \beta(p - P_d)\right)^2}\ . \quad (19)$$

Рівняння (18) є нелінійним. Отримати його розв'язок в аналітичному вигляді не завжди можливо. Як випливає з аналізу літературних джерел, ефективним способом розв'язування задач такого типу є побудова ітераційних процедур.

Відомо, що параметр β є малою величиною. У зв'язку з тим на першому етапі параметр q_1 будемо вважати сталим. Розв'язок рівняння (19) будемо шукати операційним методо в базисі перетворення Лапласа-Карсона. В цьому випадку в зображеннях Лапласа-Карсона рівняння (19) буде мати вигляд

$$q_1 c_z^2 \overline{p}' + s\overline{p} = sp(x,0) + c_z^2 \overline{\Psi}(x,s) \quad (20)$$

Розв'язок рівняння (20) має вигляд

$$\overline{p}(x,s) = e^{-\frac{sx}{q_1 c_z^2}}\left(\overline{p}(0,s) + \frac{1}{q_1 c_z^2}\int_0^x \left[sp(y,0) + c_z^2\overline{\Psi}(y,s)\right] e^{\frac{sy}{q_1 c_z^2}} dy\right)\ .$$

За початковий розподіл тиску можна взяти вираз

$$p(x,0) = p_c + \Delta p_c \cos(\alpha_c x + \gamma_c)\ ,$$

параметри якого задаються за заміряними даними частоти сердечних скорочень (ЧСС). Аналогічний вигляд має і граничний розподіл тиску

$$p(0,t) = p_t + \Delta p_t \cos(\alpha_t x + \gamma_t)\ .$$

Такий вигляд крайових умов пояснюється тим, що робота серця має періодичний характер з певною частотою та амплітудою

Переходячи від зображень до оригіналів, отримуємо рішення сформульованої задачі:

V. Висновок

Отримані результати дозволяють розраховувати розподіл тиску за довжиною трубки у випадку, коли є стоки та джерела, а також аналізувати вплив еластичності як гнучких трубок, так і великих кровоносних суден на процес розподілу тиску в них. Отримані результати підтверджують необхідність врахування розміщення трубки відносно горизонту, оскільки розподіл тиску у вертикальних і горизонтальних судинах істотно відрізняється. Як показує числовий експеримент, на рух рідини має істотний вплив і пружність трубки, яка характеризується модулем Юнга.

Література References

[1] Hotra O., P'yanylo Y. On Approach to the Modeling Process of the Flow of Blood in Vessels. // Intelligent Information and Electronic Technology, 2008.-IIET 2/- P.103-119.

[2] П'янило Я.Д., Лопатьєв А.О., Готра О.З., Трач В.М., П'янило Г.М. Прямі та обернені задачі моделювання рухуу речовини в об'єктах складної структури // Теорія та методика фізичного виховання.- 2009.-№7.- с.11-15.

[3] Педли Т. Гидродинамика крупных кровеносных сосудов. М.: Мир, 1983. 400 с.

[4] Левтов В.А., Регирер С.А., Шадрина Н.Х. Реология крови. М.: Медицина, 1982. 272 с.

[5] Каро К., Педли Т., Шротер Р., Сид У. Механика кровообращения. М.: Мир, 1981. 624 с..

[6] Лойцянский Л.Г. Механика жидкости и газа. М.: Наука, 1978. 736

Система Оперативного Контролю Органіки у Природній і Питній Водах

Олександр Малько
Кафедра інформатики
Прикарпатський національний університет
Івано-Франківськ, Україна
malko.pochta@gmail.com

Анастасія Малько
Кафедра математичних методів в інженерії
Івано-Франківський національний технічний університет нафти і газу
Івано-Франківськ, Україна
malko@pochta.ru

System of Operative Control Organic Matter in Natural and Drinking Water

Oleksandr Malko
dept. of Informatics
Precarpathian National University
Ivano-Frankivsk, Ukraine
malko.pochta@gmail.com

Anastasija Malko
dept. of of Mathematical Methods in Engineering
Iv.-Frankivsk National Technical University of Oil and Gas
Ivano-Frankivsk, Ukraine
malko@pochta.ru

***Анотація*—Наведені прикладні аспекти математичного моделювання процесу пульсації рідинного меніска на предмет контролю органіки у природній і питній водах**

***Abstract*—The applied aspects of mathematical modelling of the pulsation process of a liquid meniscus for control of organic matter in natural and drinking water**

Ключові слова—рідинний меніск, капілярна поверхня, гістерезис, пульсація, поверхневий натяг, система.

Keywords—liquid meniscus, capillary surface, hysteresis, ripple, surface tension system

I. Вступ

Виняткова важливість отримання достовірної та оперативної інформації про загальний зміст органічних домішок в питних, природних і стічних водах в поєднанні з недоліками стандартного методу визначає актуальність даного дослідження, спрямованого на вивчення експресного методу контролю забруднення вод органічними речовинами.

Поверхнево-активні властивості мають більшість органічних сполук (поверхнево-активні речовини - ПАР): спирти, кислоти, аміни, гидроперекиси, кетони, ефіри, солі карбонових кислот та ін [1]. Таким чином, інформація про вміст у воді ПАР може також служити інтегральної оцінкою ступеня її чистоти на предмет органічних включень.

Для контролю органіки шляхом вимірювання динамічного або рівноважного поверхневого натягу (ПН) на границі рідина-газ, в даний час, використовується міжфазна тензометрія яка є вельми чутливим методом аналізу малих концентрацій ПАР [2].

Найбільш поширеними на теперішній час є методи, засновані на аналізі форми краплі або бульбашки, практично не мають обмежень за часом контролю процесу адсорбції, що дає можливість визначення мікроконцентрацій органічних включень [3]. Однак даний метод має ряд істотних недоліків: вплив випаровування рідини на зміну обсягу меніска бульбашки при тривалому процесі дослідження; неможливість оперативного контролю концентрації органіки; неможливість використання в автоматизованих системах контролю.

За результатами математичного моделювання процесу пульсації рідинного меніска в околиці максимального тиску в системі замкнутої щодо кількості газової фази [4] запропоновано новий метод визначення динамічного поверхневого натягу за виміряним тиску в пульсуючому меніску.

II. Метод пульсуючого меніску

Сутність методу полягає в тому, що шляхом прямого і зворотного подачі газу в пляшечку, реалізується процес пульсації меніска в околиці максимального тиску в ньому, значення якого однозначно визначає поверхневий натяг на кожному циклі. Ефект полягає в тому, що бульбашка при

кожному циклі не зачиняє після проходження максимального тиску, а переходить в новий стійкий стан рівноваги, тобто процес адсорбції поверхнево-активних речовин здійснюється на одну і ту ж поверхню розділу фаз. Основні теоретичні положення, на яких ґрунтується запропонований метод дослідження динамічного ПН рідин, а саме: суть методу пульсуючого меніска, математичний опис процесу зростання меніска на торці вертикально зануреного в рідину капіляра, особливості зміни параметрів меніска при переході через екстремальні стану наведені в працях [4]. Можливість реалізації запропонованого методу теоретично обґрунтована шляхом розробки математичної моделі процесу пульсації бульбашки в околиці максимального тиску в системі замкнутої щодо кількості газової фази.

На рис.1 наведено фрагмент залежності $V(\Delta V)$ з вираженим вигином і відповідна їй візуалізація положень об'єму меніска для , що пояснює процес гістерезису. При витискуванні газу об'єм меніска зростає у напрямі 1-2. У точці 2 відбувається перехід 2-3, тобто стрибкоподібний ріст об'єму меніска. Наступна подача газу приводить до росту меніска у напрямі 3-4. Зворотній хід поршня від точки 4 відбувається по шляху 4-3-5. У точці 5 відбувається різке зменшення об'єму меніска по напрямку 5-6. При подальшому зменшенню об'єму системи (втягування газу) процес йде по шляху 6-1. При циклічному процесі витискування – втягування відбувається виражений ефект гістерезису по всіх характеристикам меніска (об'єму, тиску, площі поверхні тощо). Причому амплітуда петлі гістерезису залежить від радіуса вимірювального капіляра, початкового об'єму системи подачі і ПН досліджуваної рідини.

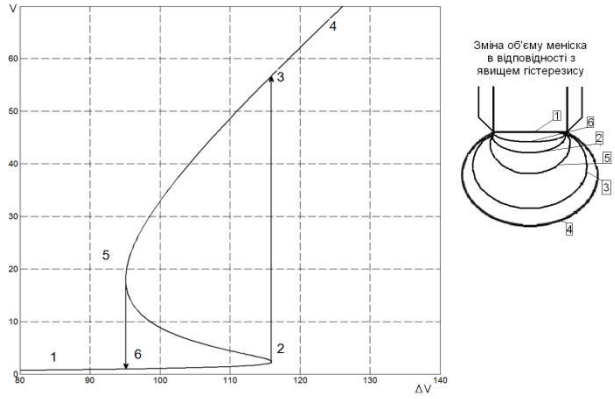


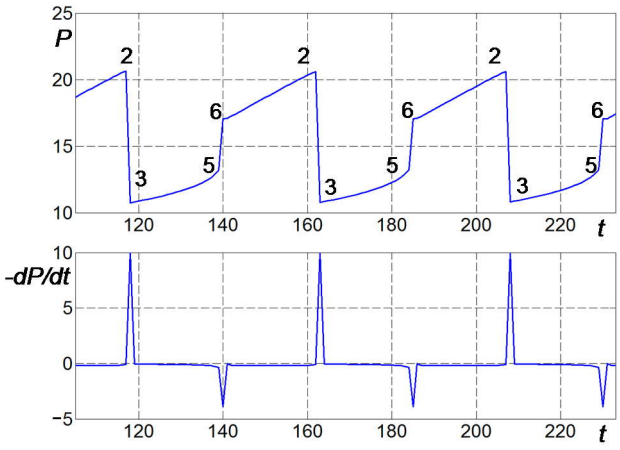
Рис. 1. Гістерезис зміни об'єму бульбашки від об'єму витискування і втягування поршнем (приріст 1-2-3-4, спад 4-3-5-6-1)

Для імітації процесу пульсації меніска з розгорткою у часі (швидкість подачі об'єму ΔV стала) шляхом сплайн апроксимації з урахуванням гістерезису було здійснено приведення всіх характеристик меніска (4)÷(6) до рівномірної дискретизації по ΔV.

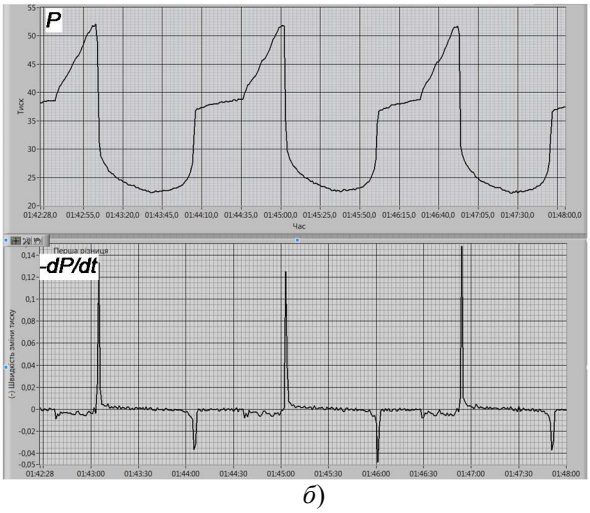
Імітація часової залежності тиску і першої похідної тиску (з оберненим знаком) у пульсуючому меніску наведені на рис.2, *а*. В даному випадку режим пульсації визначається реверсом подачі повітря в моменти стрибкоподібної зміни тиску у меніску. Відповідність між характеристикою $V(\Delta V)$ петлі гістерезису (рис.1), і зміни тиску у меніску на періоді пульсації відслідковується за номерами характерних точок кривих. Так на інтервалах 6 – 2 здійснюється подача повітря, а на інтервалах 3 – 5 всмоктування, відповідний реверс в точках 3 і 6.

Результати фізичної реалізації описаного режиму пульсації тиску у бульбашці та його похідної наведені на рис.7, *б*. Очевидна якісна подібність часових характеристик отриманих експериментально і в результаті математичного моделювання, що говорить про адекватність розробленої моделі. Закругленості і зломи експериментальній характеристиці обумовлені негерметичністю системи подачі, наявністю мертвої зони при реверсі подачі газу, інерційністю системи.

Описаний режим пульсації є основним режимом роботи вимірювальної системи, що реалізує метод пульсуючого меніска при визначенні динамічного ПН. При його реалізації, у системі управління процесом вимірювання інформаційним сигналом є швидкість зміни тиску, тобто перша похідна по тиску в часі. Як видно з характеристик (рис.2) амплітуда сигналів імпульсів похідної добре виражена, що дає можливість їх використання в контурі управління процесом дослідження динамічного ПН розчинів ПАР методом пульсуючого меніска.



а)



б)

Рис.2. Часові характеристики пульсації меніска з реверсом подачі в моменти стрибкоподібної зміни тиску (а – результати моделювання; б – результати експерименту)

III. Структура оперативного контролю органіки у природній і питній водах

Структура системи для дослідження ПН зображена на рис.4 і містить вузли, конструктивні характеристики яких базуються на результатах математичного моделювання[4]:

1 – вимірювальний капіляр з ножовим торцем, який занурюється у рідину 3 і сполучений із системою подачі повітря і одним із пліч диференціального давача тиску 4;

2 – гідростатичний компенсатор тиску з широким торцем, який занурюється в рідину 3 на однакову глибину з капіляром 1 і сполучений із протилежним плечем диференціального давача тиску 4;

3 – досліджуваний розчин;

4 – диференційний давач тиску, який вимірює різниці тисків у капілярі 1 і гідростатичному компенсаторі тиску 2;

5 – термометр опору, який вимірює температуру рідини 3 і служить для корекції результатів вимірювання ПН по температурі;

6 – вузол подачі повітря, який здійснює подачу повітря в капіляр 1 шляхом поступального переміщення плунжера у шприцевій системі;

7 – механізм переміщення плунжера вузла подачі повітря;

8 – блок обміну даними NI USB-6009, який з'єднаний через USB порт з персональним комп'ютером 9 і призначений для обробки сигналів диференційного давача тиску 4, термометра опору 5 та керуванням процесом вимірювання (запуском та реверсом механізм переміщення плунжера 7);

9 – персональний комп'ютер.

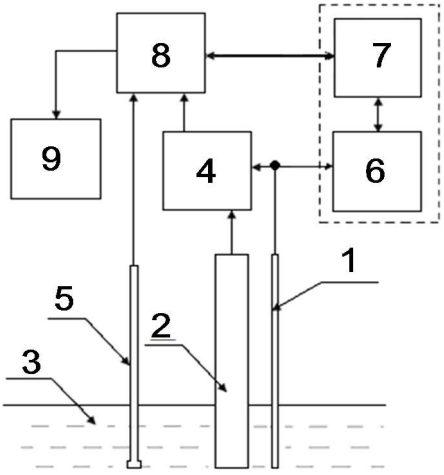


Рис. 3. Структурна системи дослідження поверхневого натягу на предмет визначення концентрації поверхнево активних речовин

IV. АЛГОРИТМ УПРАВЛІННЯ ПРОЦЕСОМ ВИМІРЮВАННЯ ТА ОБРОБКИ ДАНИХ

Узагальнена блок-схема алгоритму програми керування процесом вимірювання та обробки даних системи, з врахуванням специфіки функціонування блоку збору даних NI USB–6009, показана на рис.5. Задана тактова частота виконання простих операцій програми в даному випадку рівна 40 кГц.

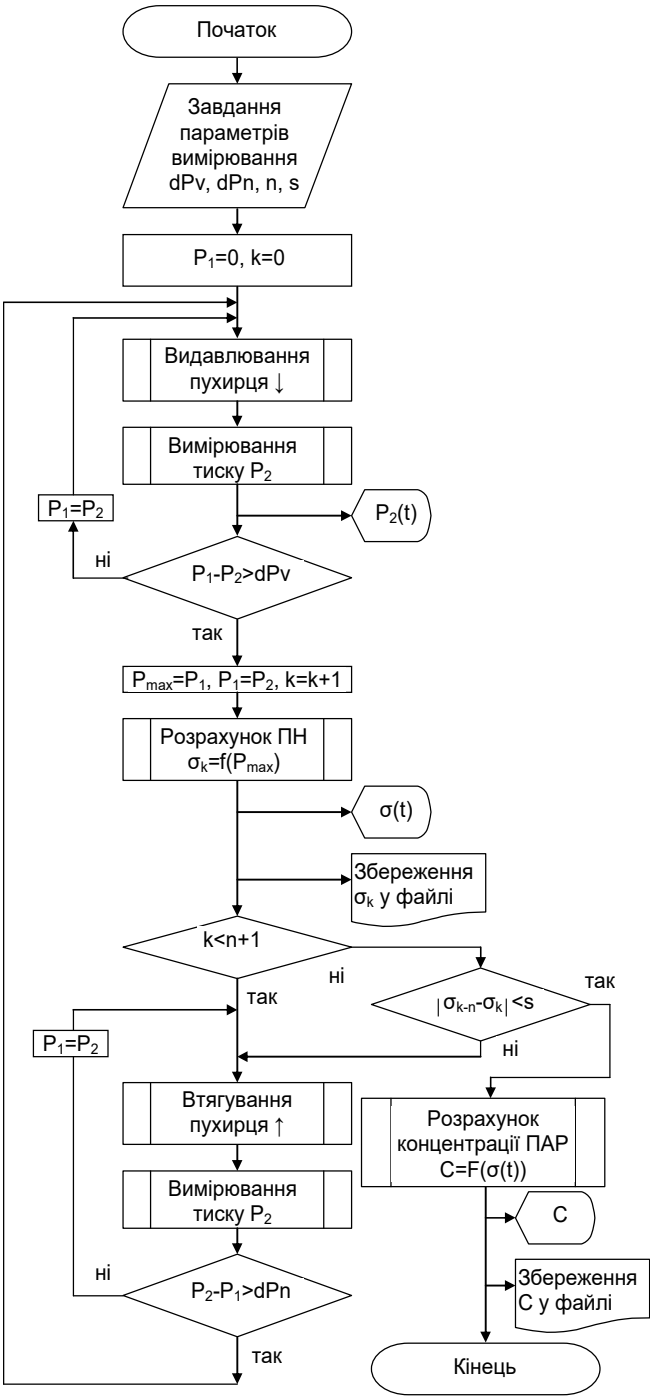


Рис. 4. Блок схема алгоритму програми керування процесом вимірювання та обробки даних

Послідовність операцій алгоритму програми є наступною.

1) На початку функціонування програми необхідно задати основні параметри, які визначають критерії виконання керуючих функцій алгоритму. Їх значення обумовлені апріорною інформацією про область зміни динамічного ПН досліджуваного розчину ПАР. До даних параметрів слід віднести: приблизний час досягнення рівноважного ПН, радіус вимірювального капіляра, що обумовлює чутливість вимірювань, наявність шумів

сигналу вимірювань і т. п. Завдання параметрів може здійснюватись: введенням з клавіатури або зчитуватися з файлу, який попередньо введено у програмний модуль. В даному випадку задаються наступні параметри:

– dP_v – значення критерію досягнення максимального тиску у бульбашці в процесі зростання її об'єму, який визначає мінімальну величину стрибкоподібного спаду тиску після проходження максимуму (перехід 2–3, рис.1);

– dP_n – значення критерію досягнення моменту стрибкоподібного спадання об'єму бульбашки в процесі зменшення її об'єму, який визначає мінімальну величину стрибкоподібного зростання тиску (перехід 5–6, рис.1);

– n – значення зсуву номерів у послідовності вимірювань ПН для визначення умови досягнення рівноважного поверхневого натягу;

– s – значення критерію досягнення рівноважного поверхневого натягу, який визначає максимальну різницю значень у послідовності вимірювань ПН з різницею номерів n.

2) Здійснюється обнулення лічильника кількості циклів вимірювання – k і значення попереднього виміру P_1 у послідовності P_1, P_2 – вимірів надлишкового тиску у рідинному меніску, з метою коректності їх подальшого порівняння.

3) Процес вимірювання починається з команди на витискування бульбашки, що реалізується поданням імпульсу логічної одиниці на відповідний цифровий вихід блоку збору даних NI USB–6009.

4) На кожному такті здійснюється вимірювання тиску у рідинному меніску P_2, що реалізується усередненням вибірки з п'яти послідовних вимірів.

5) Результати вимірювання тиску P_2, за бажанням користувача, відображаються на моніторі у вигляді осцилограми розгорнутої в часі.

6) Після кожного наступного вимірювання тиску P_2 здійснюється порівняння з попереднім значенням P_1, у разі стрибкоподібного спадання тиску на величину більше ніж dP_v (досягнення максимального тиску), процес видавлювання бульбашки припиняється, у протилежному випадку, після переприсвоєння послідовності вимірів P_1, P_2, процес витискування продовжується (перехід до п.3).4

7) При досягненні максимального тиску фіксується його значення P_{max}, проводиться переприсвоєння значень послідовності вимірів P_1, P_2, та здійснюється операція декрементування лічильника k.

8) За отриманим значенням P_{max} здійснюється розрахунок поверхневого натягу σ побудовані для конкретного радіуса вимірювального капіляра.

9) Результати обчислень значення ПН σ виводяться на монітор у вигляді осцилограми розгорнутої в часі і, за бажанням користувача, записуються у файл для можливості наступної обробки.

10) Здійснюється оцінка на можливість перевірки досягнення рівноважного стану – кількість проведених послідовних вимірів ПН k повинна бути більша від n.

11) При виконанні попередньої умови п.10 здійснюється порівняння значень двох вимірів ПН σ у послідовності вимірювань ПН з різницею номерів n.

12) Якщо неможливо зробити перевірку досягнення рівноважного стану ($k<n+1$), або стан рівноважного поверхневого натягу є не досягнутий, то подається команда на втягування бульбашки, що реалізується поданням імпульсу логічної одиниці на відповідний цифровий вихід блоку збору даних NI USB–6009.

13) Аналогічно як в п.4 проводиться вимірювання тиску у бульбашці P_2, що реалізується усередненням вибірки з п'яти послідовних вимірів.

14) Після кожного вимірювання тиску P_2 здійснюється порівняння з попереднім значенням P_1, у разі стрибкоподібного зростання тиску на величину більше ніж dP_n (стрибкоподібне зменшення об'єму бульбашки) процес втягування бульбашки припиняється, і проводиться перехід до режиму витискування (перехід до п.3).

15) У протилежному випадку, після переприсвоєння послідовності вимірювань P_1, P_2 процес втягування продовжується (перехід до п.12).

16) Якщо модуль різниці цих вимірювань є меншим значення критерію досягнення рівноважного поверхневого натягу s, то вважається, що досягнуто стан рівноважного поверхневого натягу.

17) За отриманою динамічною характеристикою σ(t) розраховується концентрація ПАР.

18) Результати розрахунку виводяться на екран комп'ютера та можуть бути записані у заданий файл.

19) Процес вимірювання закінчено.

V. Висновки

На основі моделювання квазистатики капілярних поверхонь типу лежача крапля була запропонована нова методика та синтезована інформаційно-вимірювальна система оперативного контролю мікровключень органічних речовин у водному середовищі за динамічним поверхневим натягом. Результати випробувань показали, що дана система дає можливість ідентифікувати за зміною динамічного поверхневого натягу дистилят, кранову, та природну води.

Література References

[1] В. Б Файнерман, В.Я. Уманський, Б.С. Горелік, Д.О. Ластков, О.Г.Козаков, О контроле содержания органических соединений в питьевой и природной воде методом межфазной тензиометрии. - Вестн. гиг. епид. – 2006. – Т. 10, № 1. – С. 181 - 185.

[2] J. Krägel, A.M. Stortini, N. Degli-Innocenti, G. Loglio, R. Miller Dynamic Interfacial Properties of Marine Microlayers // Colloids Surfaces A. – 1995. – Vol. 101, № 1. – P. 129 - 135.

[3] V.B. Fainerman, S.A. Zholob, R. Miller, G. Loglio, Cini R. Measurement of Dynamic Surface Tension of Low-Concentration Surfactant Solutions Using a Stopped-Flow Maximum Bubble Pressure Method // Tenside Surfactants Detergents. – 1996. – Vol. 33, № 3. – P. 452 – 456.

[4] О. Г. Малько, А. О. Малько, Математичне моделювання процесу пульсації рідинного меніска в околі максимального тиску: Міжнародна науково-практична конференція ІТКІ - 15, Прикарпатський національний університет, Івано-Франківськ, 2015. – С. 214 – 216.

[5] V. B Fainerman, B. J. Umanskiy, B. S. Gorelik, D. O., Lastkov, O. G. Kozakov, On the control of the content of organic compounds in drinking and natural water by the method of interfacial tensiometry. Bull. of hyg. Epid. – 2006 – V. 10, №1. – P. 181-185

[6] O. G. Malko, A. J. Malko, Mathematical modeling pulsation process of a liquid meniscus in the vicinity of the maximum pressure: International Scientific and Practical Conference ITCM-2015: Precarpath. Nation. Univ., Ivano-Frankivsk, 2017, P. 214 -216.

.

СЕКЦІЯ 8.

Прикладні методи дослідження дискретно-неперервних математичних моделей

SECTION 8.

Applied methods of studying discrete-continuous mathematical models

Загальні Крайові Задачі для Гіперболічного Рівняння із Сумовними Коефіцієнтами та Правими Частинами

Р.М. Тацій
кафедра прикладної математики і механіки
ЛДУ безпеки життєдіяльності
Львів, Україна

О.О. Карабин
кафедра прикладної математики і механіки
ЛДУ безпеки життєдіяльності
Львів, Україна
tosjakarabyn@gmail.com

О.Ю. Чмир
кафедра прикладної математики і механіки
ЛДУ безпеки життєдіяльності
Львів, Україна
o_chmyr@yahoo.com

The Total Boundary Value Problems for Hiperbolic Equation with Summable Coefficients and Right Parts

R.M. Tatsij
Department of Applied Mathematics and Mechanics
Lviv State University of vital activity safety
Lviv, Ukraine

O.O. Karabyn
Department of Applied Mathematics and Mechanics
Lviv State University of vital activity safety
Lviv, Ukraine
tosjakarabyn@gmail.com

O.Yu. Chmyr
Department of Applied Mathematics and Mechanics
Lviv State University of vital activity safety
Lviv, Ukraine
o_chmyr@yahoo.com

***Анотація*—Вперше запропоновано та обґрунтовано нову формальну схему розв'язування загальних крайових задач для гіперболічного рівняння із сумовними коефіцієнтами та правими частинами. В основу схеми розв'язування покладено концепцію квазіпохідних, сучасну теорію систем лінійних диференціальних рівнянь, а також класичний метод Фур'є та метод редукції.**

***Abstract*—The main methods for solving nonstationary boundary value problems are the separation of variables method, Green's function method, method of integral transforms, approximate and numerical methods.**

The scheme proposed in this article belongs to the direct methods for solving boundary value problems. In the basis of the solving scheme is the concept of quasi-derivatives that allows to bypass the problem of multiplication of generalized functions.

In this direction, a mixed problem for the heat equation with piecewise continuous coefficients by the general boundary conditions of the first kind was solved first.

This article examines the general boundary value problems for a hyperbolic type equation with summable coefficients and right parts. With the use of the reduction method, solving of such problems is reduced to finding a solution of the quasi-stationary inhomogeneous boundary value problem with the initial

boundary conditions and the mixed problem with zero boundary conditions for an inhomogeneous equation.

The practical implementation of this scheme depends on the structure of the coefficients of the original differential equation by the spatial variable. Thus, for example, in the case of piecewise constant coefficients there is a possibility to examine the longitudinal (torsional) oscillations of the stepwise pivots with piecewise-variable distribution of parameters.

Ключові слова—квазідиференціальне рівняння; крайова задача; матриця Коші; задача на власні значення; метод Фур'є та метод власних функцій.

Keywords—kvazidifferential equation; the boundary value problem; the Cauchy matrix; the eigenvalues problem; the method of Fourier and the method of eigenfunctions.

I. Вступ

Методи розв'язування нестаціонарних крайових задач можна поділити на прямі, основа яких становить метод відокремлення змінних, метод джерел (метод функції Ґріна), метод інтегральних перетворень, наближені та числові методи.

Запропонована в даній роботі схема належить до прямих методів розв'язування крайових задач. В основу реалізації цієї схеми покладено концепцію квазіпохідних [4], що дозволяє "обійти" проблему множення узагальнених функцій.

В цьому напрямку першою була розв'язана загальна мішана задача для рівняння теплопровідності з кусково-неперервними коефіцієнтами за крайових умов першого роду [6].

В цій роботі досліджують загальні крайові задачі для гіперболічного рівняння із сумовними коефіцієнтами та правими частинами. За допомогою методу редукції розв'язування таких задач зведено до знаходження розв'язку двох задач: квазістаціонарної неоднорідної крайової задачі з вихідними крайовими умовами та мішаної задачі з нульовими крайовими умовами для певного неоднорідного рівняння.

II. Основні позначення, формулювання задачі

Нехай $[0;l]$ – відрізок дійсної осі; $m(x) \in L[0;l]$, $m(x) > 0$; $\dfrac{1}{a(x)}$ – обмежена і вимірна на $[0;l]$, $a(x) > 0$; $f(x) \in L[0;l]$.

Введемо позначення: $u^{[1]} = au_x'$ – квазіпохідна.

Розглянемо загальну крайову задачу для гіперболічного рівняння

$$m(x)\frac{\partial^2 u}{\partial t^2} = \frac{\partial}{\partial x}\left(a(x)\frac{\partial u}{\partial x}\right) + f(x), \qquad (1)$$

$$x \in (0;l),\ t \in (0;+\infty),$$

з крайовими умовами

$$\begin{cases} p_{11}u(0,t) + p_{12}u^{[1]}(0,t) + q_{11}u(l,t) + q_{12}u^{[1]}(l,t) = \psi_0(t), \\ p_{21}u(0,t) + p_{22}u^{[1]}(0,t) + q_{21}u(l,t) + q_{22}u^{[1]}(l,t) = \psi_1(t), \end{cases}$$
$$(10)\ t \in [0;+\infty)$$

та початковими умовами

$$\begin{cases} u(x,0) = \varphi_0(x), \\ \dfrac{\partial u}{\partial t}(x,0) = \varphi_1(x), \end{cases} \quad x \in [0;l], \qquad (2)$$

де $\psi_0(t)$, $\psi_1(t) \in C^2(0;+\infty)$, $\varphi_0(x)$, $\varphi_1(x)$ – абсолютно - неперервні на $[0;l]$.

Метод редукції відшукання розв'язку задачі детально описаний, наприклад, в [1, 2]. Згідно з цим методом розв'язок задачі (1) – (2) шукаємо у вигляді суми двох функцій

$$u(x,t) = w(x,t) + v(x,t). \qquad (3)$$

Одну з функцій, наприклад $w(x,t)$ виберемо спеціальним способом, тоді функцію $v(x,t)$ вже визначимо однозначно.

III. Побудова функції $w(x,t)$

Визначимо функцію $w(x,t)$ як розв'язок крайової задачі

$$(a(x)w_x')_x' = -f(x), \qquad (4)$$

$$\begin{cases} p_{11}w(0) + p_{12}w^{[1]}(0) + q_{11}w(l) + q_{12}w^{[1]}(l) = \psi_0(t), \\ p_{21}w(0) + p_{22}w^{[1]}(0) + q_{21}w(l) + q_{22}w^{[1]}(l) = \psi_1(t), \end{cases} \qquad (5)$$

$$t \in [0;+\infty).$$

Зауважимо, що змінна t тут вважається параметром.

В основі методу розв'язування задачі (4), (5) лежить концепція квазіпохідних [3].

Введемо вектори $\overline{W} = \begin{pmatrix} w \\ w^{[1]} \end{pmatrix}$, де $w^{[1]} = aw_x'$, $\overline{F} = \begin{pmatrix} 0 \\ -f \end{pmatrix}$

.

За таких позначень квазідиференціальне рівняння (4) зводиться до еквівалентної системи диференціальних рівнянь першого порядку

$$\overline{W}_x' = A \cdot \overline{W} + \overline{F}, \qquad (6)$$

де $A = \begin{pmatrix} 0 & \dfrac{1}{a(x)} \\ 0 & 0 \end{pmatrix}$.

Під розв'язком системи (6) розуміємо вектор-функцію $\overline{W}(x,t)$, що за змінною x є абсолютно-неперервна та справджує систему (6) майже всюди (див. [3]).

Крайові умови (5) теж запишемо у векторній формі

$$P\cdot\overline{W}(0,t)+Q\cdot\overline{W}(l,t)=\overline{\Gamma}(t), \qquad (7)$$

де $P=\begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}$, $Q=\begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix}$, причому $rang(P\,|\,Q)=2$, $\overline{\Gamma}(t)=\begin{pmatrix} \psi_0(t) \\ \psi_1(t) \end{pmatrix}$.

Матриця Коші $B(x,s)$ такої системи має вигляд $B(x,s)=\begin{pmatrix} 1 & K(x,s) \\ 0 & 1 \end{pmatrix}$, де $K(x,s)=\int\limits_s^x \frac{1}{a(z)}dz$ (див. [7]).

Розв'язок системи (6) має вигляд

$$\overline{W}(x,t)=B(x,0)\cdot\overline{W}_0(t)+\int\limits_0^x B(x,s)\cdot\overline{F}(s)\,ds, \qquad (8)$$

де $\overline{W}_0(t)$ – початковий (невідомий) вектор [6].

Для знаходження $\overline{W}_0(t)$ використовуємо крайові умови (7), в яких покладемо $\overline{W}(0,t)\overset{def}{=}\overline{W}_0(t)$. Тоді $[P+Q\cdot B(l,0)]\overline{W}_0(t)+Q\int\limits_0^l B(l,s)\cdot\overline{F}(s)\,ds=\overline{\Gamma}(t)$, звідки одержуємо

$$\overline{W}_0(t)=[P+Q\cdot B(l,0)]^{-1}\cdot\left(\overline{\Gamma}(t)-Q\int\limits_0^l B(l,s)\cdot\overline{F}(s)\,ds\right). \qquad (9)$$

Підставляючи (9) в (8), отримаємо зображення вектор - функції $\overline{W}(x,t)$

$$\overline{W}(x,t)=B(x,0)\cdot[P+Q\cdot B(l,0)]^{-1}\times$$
$$\times\left(\overline{\Gamma}(t)-Q\int\limits_0^l B(l,s)\cdot\overline{F}(s)\,ds\right)+\int\limits_0^x B(x,s)\cdot\overline{F}(s)\,ds. \qquad (10)$$

Перша координата вектора $\overline{W}(x,t)$ в (10) і є шуканою функцією $w(x,t)$.

IV. Побудова функції $v(x,t)$

Запишемо мішану задачу для функції $v(x,t)$. Підставляючи (3) в (1) та враховуючи, що функція $w(x,t)$ задовольняє (4), одержуємо неоднорідне рівняння

$$m(x)\frac{\partial^2 v}{\partial t^2}-\frac{\partial}{\partial x}\left(a(x)\frac{\partial v}{\partial x}\right)=-m(x)\frac{\partial^2 w}{\partial t^2},$$
$$x\in(0;l),\ t\in(0;+\infty). \qquad (11)$$

Підставимо (3) в початкові умови (4). Одержимо для функції $v(x,t)$ початкові умови

$$\begin{cases} v(x,0)=\Phi_0(x), \\ \dfrac{\partial v}{\partial t}(x,0)=\Phi_1(x), \end{cases} x\in[0;l], \qquad (12)$$

де $\Phi_0(x)\overset{def}{=}\varphi_0(x)-w(x,0)$, $\Phi_1(x)\overset{def}{=}\varphi_1(x)-\frac{\partial w}{\partial t}(x,0)$.

Оскільки функція $w(x,t)$ справджує крайові умови (5), то із (3) випливають крайові умови для функції $v(x,t)$

$$\begin{cases} p_{11}v(0)+p_{12}v^{[1]}(0)+q_{11}v(l)+q_{12}v^{[1]}(l)=0, \\ p_{21}v(0)+p_{22}v^{[1]}(0)+q_{21}v(l)+q_{22}v^{[1]}(l)=0, \end{cases} \qquad (13)$$
$$t\in[0;+\infty).$$

Отже, за умови, що розв'язок $w(x,t)$ задачі (4), (5) є відомим, функція $v(x,t)$ є розв'язком мішаної задачі (11) – (13).

V. Метод Фур'є та задача на власні значення

Для рівняння (11) розглянемо відповідне однорідне рівняння

$$m(x)\frac{\partial^2 v}{\partial t^2}=\frac{\partial}{\partial x}\left(a(x)\frac{\partial v}{\partial x}\right). \qquad (14)$$

Знайдемо його нетривіальні розв'язки у вигляді

$$v(x,t)=\sin(\omega t+\varepsilon)\cdot X(x), \qquad (15)$$

де ω – параметр, ε – константа, $X(x)$ – поки що невідома функція [1], що справджує крайові умови (13).

Підставимо (15) в рівняння (14). Одержимо квазідиференціальне рівняння

$$\left(a(x)X'(x)\right)'+\omega^2 m(x)X(x)=0. \qquad (16)$$

Підставивши (15) в умови (13), одержимо крайові умови

$$\begin{cases} p_{11}X(0)+p_{12}X^{[1]}(0)+q_{11}X(l)+q_{12}X^{[1]}(l)=0, \\ p_{21}X(0)+p_{22}X^{[1]}(0)+q_{21}X(l)+q_{22}X^{[1]}(l)=0. \end{cases} \qquad (17)$$

Як і вище, під розв'язком рівняння (16) розуміємо абсолютно-неперервну на $[0;l]$ функцію $X(x)$, що справджує його майже всюди.

Ввівши квазіпохідну $X^{[1]} \overset{def}{=} aX'$, вектор $\overline{X} = \begin{pmatrix} X \\ X^{[1]} \end{pmatrix}$ та матрицю $A = \begin{pmatrix} 0 & \frac{1}{a} \\ -m\,\omega^2 & 0 \end{pmatrix}$, запишемо задачу (16), (17) в матричному вигляді

$$\overline{X}' = A \cdot \overline{X}, \tag{18}$$

$$P\overline{X}(0) + Q\overline{X}(0) = \overline{0}. \tag{19}$$

Щоб отримати характеристичне рівняння задачі (18), (19), шукатимемо нетривіальний розв'язок $\overline{X}(x,\omega)$ системи (18) у вигляді $\overline{X}(x,\omega) = \overline{B}(x,0,\omega) \cdot \overline{C}$, де $\overline{C} = \begin{pmatrix} C_1 \\ C_2 \end{pmatrix}$ – деякий ненульовий вектор, $\overline{B}(x,0,\omega)$ – матриця Коші системи (18).

Вектор - функція $\overline{X}(x,\omega)$ має задовольняти крайові умови (19), тобто

$$\left(P \cdot \overline{B}(0,0,\omega) + Q \cdot \overline{B}(l,0,\omega)\right) \cdot \overline{C} = \overline{0},$$

врахувавши, що прийдемо до рівності

$$\left(P + Q \cdot \overline{B}(l,0,\omega)\right) \cdot \overline{C} = \overline{0}. \tag{20}$$

Для існування ненульового вектора $\overline{C}$ в (20) необхідно і досить виконання умови

$$\det\left(P + Q \cdot \overline{B}(l,0,\omega)\right) = 0. \tag{21}$$

Власні значення задачі (18), (19) не завжди дійсні. Вкажемо достатні умови дійсності коренів характеристичного рівняння (21).

Перепишемо систему (18) у вигляді

$$\overline{X}' = (B_0 + \omega^2 \cdot A_0) \cdot \overline{X},$$

де $B_0 = \begin{pmatrix} 0 & \frac{1}{a} \\ 0 & 0 \end{pmatrix}$, $A_0 = \begin{pmatrix} 0 & 0 \\ -m & 0 \end{pmatrix}$.

Введемо косоермітову матрицю $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, яка володіє такими властивостями: $J^T = -J$, $J \cdot J^T = J^T \cdot J = E$, $J^2 = -E$.

Відомо [9], що, коли виконуються умови

$$\begin{aligned} &1.\ B_0{}^T = JB_0J;\ A_0{}^T = JA_0J; \\ &2.\ rang(P \mid Q) = 2; \\ &3.\ PJP^T = QJQ^T, \end{aligned} \tag{22}$$

то всі власні значення задачі (18), (19) дійсні. В даній роботі умови (22) будемо вважати виконаними.

Позначимо $\overline{X}_k(x,\omega_k)$ – нетривіальний власний вектор, що відповідає власному значенню ω_k.

VI. Побудова розв'язку $v(x,t)$ мішаної задачі (11) – (13)

Для розв'язання задачі (11) – (13) застосуємо метод власних функцій [2], який полягає в тому, що розв'язок задачі (11) – (13) шукаємо у вигляді

$$v(x,t) = \sum_{k=1}^{\infty} T_k(t) \cdot X_k(x,\omega_k), \tag{23}$$

де $T_k(t)$ – невідомі функції, які визначимо далі.

Оскільки $\frac{\partial^2 w}{\partial t^2}$ входить в праву частину рівняння (11), то розвинемо її в ряд Фур'є за власними функціями $X_k(x,\omega_k)$ крайової задачі (16), (17)

$$\frac{\partial^2 w}{\partial t^2} = \sum_{k=1}^{\infty} w_k(t) \cdot X_k(x,\omega_k). \tag{24}$$

Підставляючи вираз (23) у (11) та враховуючи (24), отримаємо рівність

$$m(x) \cdot \sum_{k=1}^{\infty} T_k''(t) \cdot X_k(x,\omega_k) = \sum_{k=1}^{\infty} T_k(t) \cdot \left(a(x) X_k'(x,\omega_k)\right)' -$$

$$-m(x) \cdot \sum_{k=1}^{\infty} w_k(t) \cdot X_k(x,\omega_k).$$

Враховуючи, що власні функції $X_k(x,\omega_k)$ задовольняють рівняння (16), приходимо до рівності

$$m(x) \cdot \sum_{k=1}^{\infty} T_k''(t) \cdot X_k(x,\omega_k) = -m(x) \sum_{k=1}^{\infty} \omega_k^2 \cdot X_k(x,\omega_k) T_k(t) -$$

$$-m(x) \cdot \sum_{k=1}^{\infty} w_k(t) \cdot X_k(x,\omega_k),$$

$$\sum_{k=1}^{\infty} \left(T_k''(t) + \omega_k^2 \cdot T_k(t) + w_k(t)\right) \cdot m(x) \cdot X_k(x,\omega_k) = 0. \tag{25}$$

Помножимо ліву і праву частини (25) на $X_j(x,\omega_j)$ та проінтегруємо за змінною x на проміжку $[0;l)$, врахувавши ортогональність власних функцій, приходимо до сукупності диференціальних рівнянь

$$T_k''(t)+\omega_k^2\cdot T_k(t)=-w_k(t),\ k=1,2,3,\ldots. \quad (26)$$

Загальний розв'язок кожного з диференціальних рівнянь (26) має вигляд

$$T_k(t)=a_k\cos\omega_k t+d_k\sin\omega_k t-\frac{1}{\omega_k}\int_0^t\sin\omega_k(t-s)\cdot w_k(s)\,ds \quad ,(27)$$

де a_k , d_k – невідомі сталі [8].

Позначимо $I(t)=\frac{1}{\omega_k}\int_0^t\sin\omega_k(t-s)\cdot w_k(s)\,ds$.

Зауважимо, що $I(0)=0$, $I_t'(0)=0$ [5].

Для визначення сталих a_k , d_k розвинемо в ряди Фур'є за власними функціями $X_k(x,\omega_k)$ праві частини початкових умов (12)

$$\Phi_0(x)=\sum_{k=1}^{\infty}\Phi_{0\,k}\cdot X_k(x,\omega_k), \quad (28)$$

$$\Phi_1(x)=\sum_{k=1}^{\infty}\Phi_{1\,k}\cdot X_k(x,\omega_k), \quad (29)$$

де $\Phi_{0\,k}$, $\Phi_{1\,k}$ – відповідні коефіцієнти Фур'є.

З (27) випливає, що

$$T_k(0)=a_k, \quad (30)$$

$$T_k'(t)=-a_k\omega_k\sin\omega_k t+d_k\omega_k\cos\omega_k t-I_t'(t),$$

звідки

$$T_k'(0)=d_k\omega_k. \quad (31)$$

З (23), першої умови в (12), та врахувавши (28), одержуємо $\sum_{k=1}^{\infty}T_k(0)\cdot X_k(x,\omega_k)=\sum_{k=1}^{\infty}\Phi_{0\,k}\cdot X_k(x,\omega_k)$.
Звідки, використовуючи (30),маємо

$$T_k(0)=a_k=\Phi_{0\,k}.$$

Аналогічно з (23), другої умови в (12), врахувавши (29) маємо $\sum_{k=1}^{\infty}T_k'(0)\cdot X_k(x,\omega_k)=\sum_{k=1}^{\infty}\Phi_{1\,k}\cdot X_k(x,\omega_k)$. Звідки, використовуючи (31), знаходимо

$$T_k'(0)=d_k\omega_k=\Phi_{1\,k}\text{, або } d_k=\frac{\Phi_{1\,k}}{\omega_k}.$$

Отже, остаточно отримуємо розв'язок мішаної задачі (11) – (13) у вигляді ряду

$$v(x,t)=\sum_{k=1}^{\infty}\left(\Phi_{0\,k}\cos\omega_k t+\frac{\Phi_{1\,k}}{\omega_k}\sin\omega_k t-\right.$$

$$\left.-\frac{1}{\omega_k}\int_0^t\sin\omega_k(t-s)\cdot w_k(s)\,ds\right)\cdot X_k(x,\omega_k).$$

VII. Висновки

Теорема про розвинення за власними функціями адаптована для випадку диференціальних рівнянь із сумовними (за Лебегом) коефіцієнтами.

Отримано формули для обчислення розв'язку та його квазіпохідної для замкненого інтервала дійсної осі.

Практична реалізація цієї схеми залежить від структури коефіцієнтів вихідного диференціального рівняння за просторовою змінною. Так, наприклад, у випадку кусково-сталих коефіцієнтів виникає можливість досліджувати поздовжні (крутильні) коливання ступінчатих стрижнів з кусково-змінним розподілом параметрів.

Література References

[1] В.Я. Арсенин. Методы математической физики. - М.: Наука, 1974. - 432 с.

[2] А.Н. Тихонов, А.А. Самарский. Уравнения математической физики. - М.: Наука, 1977. - 735 с.

[3] Р.М. Тацій. Узагальнені квазідиференціальні рівняння / Р.М. Тацій, М.Ф. Стасюк, В. Мазуренко, О.О. Власій - Дрогобич. Коло, 2011. - 297 с.

[4] Р.М. Тацій. Дискретно-неперервні крайові задачі для найпростіших квазідиференціальних рівнянь другого порядку / Р.М. Тацій, М.Ф.Стасюк, О.О.Власій // Вісник НУ «Львівська політехніка»: Серія «Фіз. - мат. науки». 2011. - №718. - С. 61 - 69.

[5] В.С. Мартыненко. Операционное исчисление: Учеб. пособие. - 4 - е изд., перераб. и доп. - К.: Выща школа, 1990. - 359 с.

[6] Р.М. Тацій. Загальна перша крайова задача для рівняння теплопровідності з кусково-змінними коефіцієнтам / Р.М. Тацій, О.О. Власій, М.Ф. Стасюк // Вісник Національного університету "Львівська політехніка". Фізико-математичні науки . - 2014. - № 804. - С. 64-69.

[7] Ю.К. Рудавський. Збірник задач з диференціальних рівнянь: Навч. посібник. / Ю.К. Рудавський, П.І. Каленюк, Р.М. Тацій та ін. - Л. Вид. Національного університету "Львівська політехніка", 2001. - 244 с.

[8] П.І. Каленюк. Диференціальні рівняння: Навч. посібник. / П.І. Каленюк, Ю.К. Рудавський, Р.М. Тацій, І.Ф. Клюйник, В.М. Колісник, П.П. Костробій, І.Я. Олексів - Л. Видавництво Львівської політехніки, 2014. - 380 с.

[9] В.В. Мазуренко. Про звідність дискретно-неперервної крайової задачі до узагальненої схеми Аткінсона // Доповіді НАН України. - 2001. - № 8 - С. 19-22.

Прямий Метод Дослідження Температурного Поля в Багатошаровому Трубопроводі за Умов Пожежі

Роман Тацій
кафедра прикладної математики і механіки
Львівський державний університет безпеки життєдіяльності
Львів, Україна
roman.tatsiy@gmail.com

Олеся Власій
кафедра інформатики
Прикарпатський національний університет імені Василя Стефаника
Івано-Франківськ, Україна
olesia.vlasii@pu.if.ua

Марта Стасюк
кафедра прикладної математики і механіки
Львівський державний університет безпеки життєдіяльності
Львів, Україна
marta_stasiuk@yahoo.com

Олег Пазен
кафедра наглядово-профілактичної діяльності
Львівський державний університет безпеки життєдіяльності
Львів, Україна
opazen@gmail.com

Direct Method of Research Temperature Field in a Multilayer Conduit for Fire Conditions

Roman Tatsiy
Department of Applied Mathematics and Mechanics
Lviv State University of Life Safety
Lviv, Ukraine

Olesia Vlasii
Department of Computer Science
Vasyl Stefanyk Precarpathian National University
Ivano-Frankivsk, Ukraine
olesia.vlasii@pu.if.ua

Marta Stasiuk
Department of Applied Mathematics and Mechanics
Lviv State University of Life Safety
Lviv, Ukraine
marta_stasiuk@yahoo.com

Oleg Pazen
Department of Monitoring and Fire Prevention
Lviv State University of Life Safety
Lviv, Ukraine
opazen@gmail.com

***Анотація*—Запропоновано та обґрунтовано конструктивну схему дослідження розподілу нестаціонарного температурного поля по товщині багатошарового трубопроводу за умов пожежі. В основі даної схеми покладено прямий метод дослідження крайових задач теплопровідності з кусково-неперервними коефіцієнтами та стаціонарною неоднорідністю.**

***Abstract*—A constructive and reasonable scheme of study unsteady temperature field distribution in thickness multilayer pipe for fire conditions. At the heart of the scheme put direct method of research boundary value problems of heat conduction with piecewise continuous coefficients and stationary heterogeneity.**

Ключові слова—теплопровідність; прямий метод; багатошаровий трубопровід

Keywords—thermal conductivity; direct method; multilayer conduit

I. Вступ

У запропонованій роботі розглянуто конструктивну схему побудови розв'язку мішаної задачі для рівняння теплопровідності з кусково-сталими коефіцієнтами, що

виникає при дослідженні нестаціонарного температурного поля у багатошаровому трубопроводі. В основу схеми покладено метод редукції, концепцію квазіпохідних, модифікований метод Фур'є та метод власних функцій. В якості числового прикладу розглянуто модельну задачу про поширення температурного поля за товщиною чотиришарового трубопроводу в умовах пожежі.

II. Постановка задачі та її математична модель

Розглядається мішана задача для рівняння теплопровідності [1]

$$c\rho\frac{\partial t(r,\tau)}{\partial\tau}=\frac{1}{r}\frac{\partial}{\partial r}\left(r\lambda\frac{\partial t(r,\tau)}{\partial r}\right), \tag{1}$$

з системою крайових умов третього роду:

$$\begin{cases}\alpha_0 r_0 t(r_0,\tau)-t^{[1]}(r_0,\tau)=\alpha_0 r_0\psi_0(\tau),\\ \alpha_n r_n t(r_n,\tau)+t^{[1]}(r_n,\tau)=\alpha_n r_n\psi_n(\tau),\end{cases} \tag{2}$$

і початковою умовою:

$$t(r,0)=\varphi(r) \tag{3}$$

Тут $r_0<r_1<\ldots<r_n$ довільне розбиття проміжку $[r_0,r_n]$ дійсної осі OR на n частин, θ_i - характеристична функція відрізка $[r_i,r_{i+1}]$, $t^{[1]}\overset{df}{=}r\lambda t'$ – квазіпохідна [2]. Покладемо, що $\lambda(r)=\sum_{i=0}^{n-1}\lambda_i\theta_i$, $c(r)\rho(r)=\sum_{i=0}^{n-1}c_i\rho_i\theta_i$.

Розв'язування задачі (1), (2), (3) проводитимемо за наступною схемою [3]:

1. Розв'язок $t(r,\tau)$ шукається методом редукції у вигляді

$$t(r,\tau)=u(r,\tau)+v(r,\tau) \tag{4}$$

2. Для однієї з функцій (наприклад для $u(r,\tau)$) розв'язується квазістаціонарна крайова задача

$$\frac{1}{r}\frac{d}{dr}\left(r\lambda\frac{du(r,\tau)}{dr}\right)=0, \tag{5}$$

з крайовими умовами (2) для функції $u(r,\tau)$, тобто

$$\begin{cases}\alpha_0 r_0 u(r_0,\tau)-u^{[1]}(r_0,\tau)=\alpha_0 r_0\psi_0(\tau),\\ \alpha_n r_n u(r_n,\tau)+u^{[1]}(r_n,\tau)=\alpha_n r_n\psi_n(\tau),\end{cases} \tag{6}$$

2.1. Розв'язування крайової задачі (5), (6).

Введемо вектор $\mathbf{u}=\left(u,\ u^{[1]}\right)^T$ та матрицю $A(r)=\begin{pmatrix}0 & \frac{1}{r\lambda(r)}\\ 0 & 0\end{pmatrix}$.

Диференціальне рівняння (5) зводиться до еквівалентної системи диференціальних рівнянь 1-го порядку

$$\mathbf{u}'=A\mathbf{u} \tag{7}$$

Крайові умови (6) запишемо у векторній формі [2]

$$P\cdot\mathbf{u}(r_0)+Q\cdot\mathbf{u}(r_n)=\mathbf{\Gamma}(\tau), \tag{8}$$

де P, Q и $\mathbf{\Gamma}(\tau)$ мають вигляд

$$P=\begin{pmatrix}\alpha_0 r_0 & -1\\ 0 & 0\end{pmatrix},\quad Q=\begin{pmatrix}0 & 0\\ \alpha_n r_n & 1\end{pmatrix},\ \mathbf{\Gamma}(\tau)=\begin{pmatrix}\alpha_0 r_0\psi_0(\tau)\\ \alpha_n r_n\psi_n(\tau)\end{pmatrix}. \tag{9}$$

На кожному з проміжків система (7) має вигляд

$$\mathbf{u_i}'=A_i\mathbf{u_i},\quad A_i=\begin{pmatrix}0 & \frac{1}{r\lambda_i}\\ 0 & 0\end{pmatrix}. \tag{10}$$

Матриця Коші $B_i(r,s)$ системи (10) зображується наступним чином

$$B_i(r,s)=\begin{pmatrix}1 & \frac{\ln r-\ln s}{\lambda_i}\\ 0 & 1\end{pmatrix}. \tag{11}$$

Для довільного $k\geq i$ позначимо

$$B(r_k,r_i)\overset{df}{=}B_{k-1}(r_k,r_{k-1})\cdot B_{k-2}(r_{k-1},r_{k-2})\cdots B_i(r_{i+1},r_i). \tag{12}$$

Структура (11) матриці $B_i(r,s)$ дає можливість встановити структура матриці (12), а саме

$$B(r_k,r_i)=\begin{pmatrix}1 & \sum_{m=i}^{k-1}\frac{\ln r_{m+1}-\ln r_m}{\lambda_m}\\ 0 & 1\end{pmatrix}. \tag{13}$$

На кожному з проміжків $[r_i,r_{i+1})$ розв'язок задачі (5), (6) має вигляд

$$\mathbf{u_i}(r,\tau)=B_i(r,r_i)\cdot B(r_i,r_0)\cdot\mathbf{P_0}(\tau), \tag{14}$$

де

$$\mathbf{P_0}(\tau)=\left(P+Q\cdot B(r_n,r_0)\right)^{-1}\cdot\mathbf{\Gamma}(\tau)=$$

$$=\left[\begin{pmatrix}\alpha_0 r_0 & -1\\ 0 & 0\end{pmatrix}+\begin{pmatrix}0 & 0\\ \alpha_n r_n & 1\end{pmatrix}\begin{pmatrix}1 & \sum_{i=0}^{n-1}\frac{\ln r_{i+1}-\ln r_i}{\lambda_i}\\ 0 & 1\end{pmatrix}\right]^{-1}\begin{pmatrix}\alpha_0 r_0\psi_0(\tau)\\ \alpha_n r_n\psi_n(\tau)\end{pmatrix}=$$

$$=\begin{pmatrix}\frac{(\alpha_n r_n\sigma_n+1)\alpha_0 r_0\psi_0(\tau)+\alpha_n r_n\psi_n(\tau)}{\alpha_n r_n+1+\alpha_0 r_0(\alpha_n r_n\sigma_n+1)}\\ \frac{\alpha_0\alpha_n r_0 r_n(\psi_n(\tau)-\psi_0(\tau))}{\alpha_n r_n+1+\alpha_0 r_0(\alpha_n r_n\sigma_n+1)}\end{pmatrix}.$$

Вираз (14) дозволяє записати розв'язок $\mathbf{u}(r,\tau)$ на проміжку $[r_0,r_n]$ за допомогою характеристичних функцій

θ_i у вигляді

$$\mathbf{u}(r,\tau)=\sum_{i=0}^{n-1}\mathbf{u_i}(r,\tau)\theta_i. \qquad (15)$$

3. Для функції $v(r,\tau)$ отримуємо мішану неоднорідну задачу

$$c\rho\frac{\partial v(r,\tau)}{\partial\tau}=\frac{1}{r}\frac{\partial}{\partial r}\left(r\lambda\frac{\partial v(r,\tau)}{\partial r}\right)-c\rho\frac{\partial u(r,\tau)}{\partial\tau}. \qquad (16)$$

з нульовими крайовими умовами для функції $v(r,\tau)$,

$$\begin{cases}\alpha_0 r_0 v(r_0,\tau)-v^{[1]}(r_0,\tau)=0,\\ \alpha_n r_n v(r_n,\tau)+v^{[1]}(r_n,\tau)=0,\end{cases} \qquad (17)$$

і початковою умовою

$$v(r,0)=f(r)=\varphi(r)-u(r,0). \qquad (18)$$

3.1.Відокремлення змінних та задача на власні значення.

Розв'язок однорідного диференціального рівняння

$$c\rho\frac{\partial v(r,\tau)}{\partial\tau}=\frac{1}{r}\frac{\partial}{\partial r}\left(r\lambda\frac{\partial v(r,\tau)}{\partial r}\right) \qquad (19)$$

шукатимемо у вигляді

$$v(r,\tau)=e^{-\omega\tau}\cdot R(r), \qquad (20)$$

де ω – параметр, а $R(r)$ – невідома функція.

Підставляючи (20) у (19), отримаємо (квазі)диференціальне рівняння

$$(r\lambda R')'+\omega c\rho rR=0, \qquad (21)$$

з крайовими умовами

$$\begin{cases}\alpha_0 r_0 R(r_0)-R^{[1]}(r_0)=0,\\ \alpha_0 r_0 R(r_n)+R^{[1]}(r_n)=0.\end{cases} \qquad (22)$$

Задача (21), (22) – класична задача на власні значення.

Ввівши квазіпохідну $R^{[1]}\overset{df}{=}r\lambda R'$, вектор $\mathbf{R}=\begin{pmatrix}R, & R^{[1]}\end{pmatrix}^T$ і матрицю $A(r)=\begin{pmatrix}0 & \dfrac{1}{r\lambda(r)}\\ -\omega c\rho r & 0\end{pmatrix}$, зведемо (квазі)диференціальне рівняння (21) до еквівалентної системи диференціальних рівнянь 1-го порядку

$$\mathbf{R}'=A\mathbf{R}. \qquad (23)$$

На кожному з проміжків $[r_i,r_{i+1})$ система (23) матиме вигляд

$$\mathbf{R}'_i=A_i\cdot\mathbf{R}_i,\quad A_i(r)=\begin{pmatrix}0 & \dfrac{1}{r\lambda_i}\\ -\omega c_i\rho_i r & 0\end{pmatrix}. \qquad (24)$$

Матриця Коші $B_i(r,s,\omega)$ системи (24) має наступний вигляд

$$B_i(r,s,\omega)=\begin{pmatrix}\beta_{11} & \beta_{12}\\ \beta_{21} & \beta_{22}\end{pmatrix}. \qquad (25)$$

де, $\beta_{11}=\dfrac{\pi\lambda_i\alpha_i s\left(J_1(\alpha_i s)\cdot N_0(\alpha_i r)-J_0(\alpha_i r)\cdot N_1(\alpha_i s)\right)}{2\lambda_i}$

$$\beta_{12}=\frac{\pi\left(J_0(\alpha_i s)\cdot N_0(\alpha_i r)-J_0(\alpha_i r)\cdot N_0(\alpha_i s)\right)}{2\lambda_i}$$

$$\beta_{21}=\frac{\pi\lambda_i^2\alpha_i^2 rs\left(J_1(\alpha_i r)\cdot N_1(\alpha_i s)-J_1(\alpha_i s)\cdot N_1(\alpha_i r)\right)}{2\lambda_i}$$

$$\beta_{22}=\frac{\pi\lambda_i\alpha_i s\left(J_1(\alpha_i r)\cdot N_0(\alpha_i s)-J_0(\alpha_i s)\cdot N_1(\alpha_i r)\right)}{2\lambda_i}$$

Тут J_0, J_1 и N_0, N_1 – функції Бесселя і Неймана нульового та першого роду відповідно.

Позначимо:

$$B(r_i,r_0,\omega)\overset{df}{=}\prod_{j=0}^{i}B_{i-j}(r_{i-j+1},r_{i-j},\omega), \qquad (26)$$

$$B(r,r_0,\omega)\overset{df}{=}\sum_{i=0}^{n-1}B_i(r,r_i,\omega)\cdot B(r_i,r_0,\omega)\cdot\theta_i, \qquad (27)$$

$$\left[P+Q\cdot B(r_n,r_0,\omega)\right]\overset{df}{=}\begin{pmatrix}b_{11}(\omega) & b_{12}(\omega)\\ b_{21}(\omega) & b_{22}(\omega)\end{pmatrix}. \qquad (28)$$

Розв'язок системи (23) шукатимемо у вигляді

$$\mathbf{R}(r,\omega)=B(r,r_0,\omega)\cdot\mathbf{C}, \qquad (29)$$

де, $\mathbf{C}=\begin{pmatrix}C_1, & C_2\end{pmatrix}^T$ – деякий ненульовий вектор.

Застосувавши до рівності (29) крайові умови (22) та після перетворень отримаємо

$$\left[P+Q\cdot B(r_n,r_0,\omega)\right]\cdot\mathbf{C}=0. \qquad (30)$$

Для існування вектора $\mathbf{C}$ в (30) необхідно и достатньо виконання умови

$$\det\left[P+Q\cdot B(r_n,r_0,\omega)\right]=0. \qquad (31)$$

Характеристичне рівняння задачі на власні значення (21), (22) має вигляд (31).

Виходячи з умови (30), ненульовий вектор $\mathbf{C}$ має вигляд

$$\mathbf{C} = \left(\frac{1}{\alpha_0 r_0}, \quad 1 \right)^T. \tag{32}$$

Власні вектори системи диференціальних рівнянь (23) з крайовими умовами (22) мають наступну структуру

$$\mathbf{R}_k(r,\omega_k) = \tilde{B}(r,r_0,\omega) \cdot \left(\frac{1}{\alpha_0 r_0}, \quad 1 \right)^T, \quad k = 1,2,3,\ldots. \tag{33}$$

Власні функції $R_k(r,\omega_k)$, як перші координати власних векторів $\mathbf{R}_k(r,\omega_k)$, можна зобразити наступним чином

$$R_k(r,\omega_k) = (1, \quad 0) \cdot \tilde{B}(r,r_0,\omega_k) \cdot \left(\frac{1}{\alpha_0 r_0}, \quad 1 \right)^T, \quad k = 1,2,3,\ldots. \tag{34}$$

Оскільки $R_k(r,\omega_k) = \sum_{i=0}^{n-1} R_{ki}(r,\omega_k)\theta_i$, то із (34) випливає, що

$$R_{ki}(r,\omega_k) = \\ = (1, \quad 0) \cdot \tilde{B}_i(r,r_i,\omega_k) \cdot \tilde{B}(r_i,r_0,\omega_k) \cdot \left(\frac{1}{\alpha_0 r_0}, \quad 1 \right)^T. \tag{35}$$

3.2. *Метод власних функцій.*

Розв'язок мішаної задачі (16)-(18) будемо шукати у вигляді

$$v(r,\tau) = \sum_{k=1}^{\infty} T_k(\tau) \cdot R_k(r,\omega_k), \tag{36}$$

де $T_k(\tau)$ – невідомі функції.

Оскільки $\frac{\partial u}{\partial \tau}$ входить у праву частину рівняння (16), розвинемо її в ряд Фур'є за власними функціями $\mathbf{R}_k(r,\omega_k)$.

$$\frac{\partial u}{\partial \tau} = \sum_{k=1}^{\infty} u_k(\tau) \cdot R_k(r,\omega_k) \tag{37}$$

Слід зауважити, що розвинення функції $g(r)$ в ряд Фур'є за власними функціями $R_k(r,\omega_k)$ задачі (21), (22) має вигляд

$$g(r) = \sum_{k=1}^{\infty} g_k \cdot R_k(r,\omega_k), \tag{38}$$

де коефіцієнти Фур'є g_k знаходяться за формулою

$$g_k = \frac{1}{\|R_k\|^2} \int_{r_0}^{r_n} c\rho g(r) R_k(r,\omega_k) r dr = \\ = \frac{1}{\|R_k\|^2} \sum_{i=0}^{n-1} c_i \rho_i \int_{r_i}^{r_{i+1}} g_i(r) R_{ki}(r,\omega_k) r dr. \tag{39}$$

Зауважимо, що $\|R_k\|^2$ – квадрат норми власних функцій R_k

$$\|R_k\|^2 = \int_{r_0}^{r_n} c\rho r R_k^2(r,\omega_k) dr = \sum_{i=0}^{n-1} c_i \rho_i \int_{r_i}^{r_{i+1}} R_{ki}^2(r,\omega_k) r dr. \tag{40}$$

Підставляючи (36) у (16) одержуємо рівність

$$c\rho \cdot \sum_{k=1}^{\infty} T_k'(\tau) \cdot R_k(r,\omega_k) = \\ = \frac{1}{r} \sum_{k=1}^{\infty} T_k(\tau) \cdot \left(r\lambda R_k'(r,\omega_k) \right)' - c\rho \sum_{k=1}^{\infty} u_k(\tau) \cdot R_k(r,\omega_k). \tag{41}$$

Враховуючи рівність (20), після перетворень одержимо співвідношення

$$\sum_{k=1}^{\infty} \left[T_k'(\tau) + \omega_k T_k(\tau) + u_k(\tau) \right] \cdot R_k(r,\omega_k) = 0. \tag{42}$$

Прирівнюючи коефіцієнти Фур'є ряду (42) до нуля, приходимо до нескінченної сукупності диференціальних рівнянь

$$T_k'(\tau) + \omega_k T_k(\tau) + u_k(\tau) = 0, \quad k = 1,2,3,\ldots. \tag{43}$$

Загальний розв'язок кожного з цих рівнянь має вигляд

$$T_k(\tau) = C_k \cdot e^{-\omega_k \tau} - \int_0^{\tau} e^{-\omega_k(\tau - s)} \cdot u_k(s) ds, \tag{44}$$

де C_k – невідома стала величина. Для її знаходження використаємо розвинення функції $f(r)$ з початкової умови (18) в ряд Фур'є

$$v(r,0) = f(r) = \sum_{k=1}^{\infty} f_k R_k(r,\omega_k). \tag{45}$$

Після порівнянь відповідних коефіцієнтів Фур'є, приходимо до висновку, що $C_k = f_k$.

Отже розв'язок мішаної задачі (16)-(18) одержуємо у вигляді ряду

$$v(r,\tau) = \sum_{k=1}^{\infty} \left[f_k \cdot e^{-\omega_k \tau} - \int_0^{\tau} e^{-\omega_k(\tau - s)} u_k(s) ds \right] \cdot R_k(r,\omega_k) = \\ = \sum_{i=0}^{n-1} v_i(r,\tau) \cdot \theta_i, \tag{46}$$

Враховуючи зображення (4) та формули (15) і (46) розв'язок мішаної задачі (1)-(3) отримаємо у вигляді

$$t(r,\tau)=\sum_{i=0}^{n-1}\left[u_i(r,\tau)+v_i(r,\tau)\right]\cdot\theta_i. \qquad (47)$$

III. Модельна задача

В якості модельної задачі розглянемо чотиришаровий трубопровід, який виконаний з ізотропних матеріалів радіусами: $r_0=0{,}15\,м$, $r_1=0{,}154\,м$, $r_2=0{,}164\,м$, $r_3=0{,}214\,м$, $r_4=0{,}216\,м$. Необхідно визначити розподіл нестаціонарного температурного поля та густину теплового потоку по товщині трубопроводу, якщо з зовнішньої сторони відбувається пожежа, температура якої змінюється за законом $t_{зов.}(\tau)=660\left(1-0{,}687e^{-0{,}32\tau}-0{,}313e^{-3{,}8\tau}\right)+18$, а всередині протікає рідина з температурою 18^0C. В початковий момент часу температура конструкції становить 18^0C. Теплотехнічні характеристики матеріалів наведено у таблиці 1.

ТАБЛИЦЯ I. Теплофізичні характеристики

Параметр	*Шар 1*	*Шар 2*	*Шар 3*	*Шар 4*
Коефіцієнт теплопровідності $\lambda,\frac{Вт}{м\cdot К}$	58	0,27	0,056	209
Питома теплоємність $c,\frac{Дж}{кг\cdot К}$	470	1680	940	894
Густина $\rho,\frac{кг}{м^3}$	7800	1000	200	2680
Коефіцієнт теплообміну $\alpha,\frac{Вт}{м^2\cdot К}$	Всередині $-\ \alpha_0=4$, Ззовні $-\ \alpha_n=25$			

Використовуючи запропонований авторами метод, та програмне забезпечення Maple 13 одержуємо розв'язок поставленої задачі про розподіл нестаціонарного температурного поля та густини теплового потоку у вигляді графіків (Рис. 1 і Рис. 2) та таблиць 2 і 3.

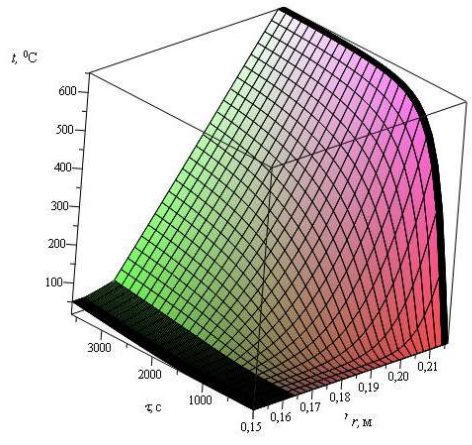


Рис. 1. Розподіл температурного поля

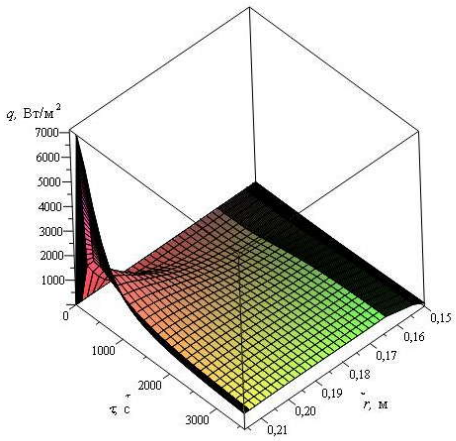


Рис. 2. Розподіл густини теплового потоку

ТАБЛИЦЯ II. Розподіл температурного поля

час, с	Координати, м						
	0,15	*0,154*	*0,164*	*0,17*	*0,2*	*0,214*	*0,216*
120	18	18	18	19	21,5	151	151
300	18	18	18,3	20,6	65	345	346
600	18	18,1	18,4	21,3	180	521	521
1200	19,1	19,2	24	47,7	340	617	618
1700	22,7	22,8	32,5	75,1	406	635	636
3600	51,5	51,5	70,2	138	488	651	651

ТАБЛИЦЯ III. Розподіл густини теплового потоку

час, с	Координати, м						
	0,15	*0,154*	*0,164*	*0,17*	*0,2*	*0,214*	*0,216*
120	0	0	0	12	65	1467	7208
300	0	0	6	8.6	446	2020	6021
600	0	0	24	28	881	1836	3449
1200	6	62	212	241	954	1218	1484
1700	20	159	390	413	855	957	1049
3600	134	369	640	639	662	645	652

Література References

[1] А.В. Лыков. Теория теплопроводности / А.В. Лыков. – М: Высшая школа, 1967. – 559 с.

[2] Р.М. Таций. Общие краевые задачи для уравнения теплопроводности с кусочно-непрерывными коэффициентами / Р.М. Таций, О.Ю. Пазен // Инженерно-физический журнал. — 2016. — Том 89, №2. — С. 350-362.

[3] Р.М. Тацій. Загальна третя крайова задача для рівняння теплопровідності з кусково-неперервними коефіцієнтами та стаціонарною неоднорідністю / Р.М. Тацій, М.Ф. Стасюк, О.Ю. Пазен // Інформаційні технології та комп'ютерна інженерія : матер. V Міжнар. наук. конф. (27-29 травня). – Івано-Франківськ – Вінниця, 2015. – С. 201–202.

Розв’язання Третьої Крайової Задачі для Рівняння Лапласа в Прямокутнику

Р.М. Тацій
кафедра прикладної математики і механіки
Львівський державний університет безпеки життєдіяльності
Львів, Україна
roman.tatsiy@gmail.com

М.Ф. Стасюк
кафедра прикладної математики і механіки
Львівський державний університет безпеки життєдіяльності
Львів, Україна
marta_stasiuk@yahoo.com

О.М. Трусевич
кафедра прикладної математики
Львівський державний університет безпеки життєдіяльності
Львів, Україна

The Solution of Third Boundary Problem for Laplace’s Equation Incide a Rectangle

R. Tatsiy
Department of Applied Mathematics and Mechanics
Lviv State University of Life Safety
Lviv, Ukraine
roman.tatsiy@gmail.com

M. Stasiuk
Department of Applied Mathematics and Mechanics
Lviv State University of Life Safety
Lviv, Ukraine
marta_stasiuk@yahoo.com

O. Trusevych
Department of Applied Mathematics and Mechanics
Lviv State University of Life Safety
Lviv, Ukraine

***Анотація*—Розв’язана задача про поширення двовимірного стаціонарного температурного поля в прямокутній пластинці за умов конвективного теплообміну з навколишнім середовищем. Дослідження відповідних задач на власні значення здійснюється шляхом зведення їх до еквівалентних задач для систем диференціальних рівнянь першого порядку. Розв’язок отримано у вигляді суми рядів.**

***Abstract*—There was solved the problem of expansion of two-dimensional stationary temperature field inside a rectangular plate under conditions of convective heat transfer with the surrounding environment. The investigation of the corresponding eigenvalue problems has been conducted by means of their reduction to the equivalent problems for systems of first order differential equations. The solution has been obtained in the form of sum of series.**

Ключові слова—крайова задача; метод Фур’є; матриця Коші

Keywords—boundary problem; Fourier method; Cauchy matrix

I. Вступ

В монографії [1] детально описаний спосіб побудови розв’язку задачі Діріхле в прямокутнику. Цей результат використано в статті [2] при описанні прямого методу розв’язання першої загальної крайової задачі для рівняння теплопровідності в прямокутнику при реалізації методу редукції.

В даній роботі розв’язана крайова задача для рівняння Лапласа в прямокутнику з умовами третього роду на його сторонах. На відміну від роботи [1], тут не вдається знайти власні значення та власні функції відповідних задач в явній формі. Натомість, для дослідження задач на власні значення, після відокремлення змінних, використовується метод зведення їх до еквівалентних задач для систем диференціальних рівнянь першого порядку [3].

II. Постановка задачі та її розв'язання.

Необхідно розв'язати рівняння Лапласа в прямокутнику $0 < x < p, 0 < y < q$ (див. Рис.1):

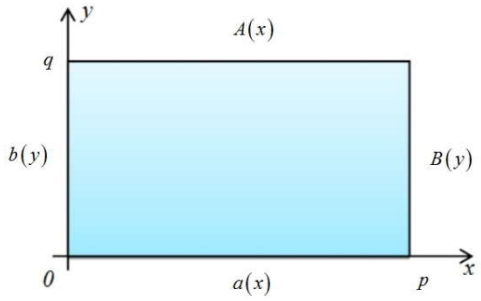


Рис. 1. Прямокутна область знаходження розв'язку крайової задачі (1),(2),(3)

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \tag{1}$$

з системою крайових умов третього роду:

$$\begin{cases} \alpha_0 \, u\big|_{x=0} - \lambda \dfrac{\partial u}{\partial x}\Big|_{x=0} = b(y), \\ \alpha_l \, u\big|_{x=p} + \lambda \dfrac{\partial u}{\partial x}\Big|_{x=p} = B(y), \end{cases} \tag{2}$$

$$\begin{cases} \beta_0 \, u\big|_{y=0} - \lambda \dfrac{\partial u}{\partial x}\Big|_{y=0} = a(x), \\ \beta_l \, u\big|_{y=q} + \lambda \dfrac{\partial u}{\partial x}\Big|_{y=q} = A(x), \end{cases} \tag{3}$$

де $a(x), A(x)$ – задані неперервні на $[0, p]$ функції від x; $b(y), B(y)$ – задані неперервні на $[0, q]$ функції від y; $\alpha_0 > 0, \alpha_l > 0, \beta_0 > 0, \beta_l > 0, \lambda > 0$ – деякі числа.

Якщо числа $\alpha_0, \alpha_l, \beta_0, \beta_l$ інтерпретувати, як відповідні коефіцієнти теплообміну з навколишнім середовищем, а параметр λ – як коефіцієнт теплопровідності матеріалу прямокутної пластинки, то задача (1), (2), (3) має наступний фізичний зміст: вона описує процес поширення двовимірного стаціонарного температурного поля в пластинці за умов конвективного теплообміну з навколишнім середовищем.

Знайдемо спочатку розв'язки [1] рівняння (1), що мають вигляд

$$u(x, y) = X(x) \cdot Y(y) \tag{4}$$

і справджують нульові крайові умови (2) при $b(y) = 0, B(y) = 0$. Підставляючи (4) в (1), дістанемо умову відокремлення змінних

$$\frac{X''}{X} = -\frac{Y''}{Y} = -\omega, \tag{5}$$

де λ – параметр. Для визначення функції $X(x)$ отримуємо задачу ні власні значення

$$X'' + \omega X = 0, \tag{6}$$

$$\begin{cases} \alpha_0 X(0) - \lambda X'(0) = 0, \\ \alpha_p X(p) + \lambda X'(p) = 0. \end{cases} \tag{7}$$

Для дослідження цієї задачі зведемо рівняння (6) до еквівалентної системи диференціальних рівнянь першого порядку [3]

$$\bar{X}' = A \cdot \bar{X}, \tag{8}$$

де $\bar{X} = (x, x')^T$, $A = \begin{pmatrix} 0 & 1 \\ -\omega & 0 \end{pmatrix}$.

Крайові умови (7) також запишемо у матричному вигляді

$$\begin{pmatrix} \alpha_0 & -\lambda \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x(0) \\ x'(p) \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ \alpha_p & \lambda \end{pmatrix} \begin{pmatrix} x(p) \\ x'(p) \end{pmatrix} = \bar{0}. \tag{9}$$

Безпосередньою перевіркою переконуємося, що матриця Коші $B(x, s)$ системи (8) має вигляд

$$B(x, s) = \begin{pmatrix} \cos\sqrt{\omega}\,(x-s) & \dfrac{\sin\sqrt{\omega}\,(x-s)}{\sqrt{\omega}} \\ -\sqrt{\omega}\sin\sqrt{\omega}\,(x-s) & \cos\sqrt{\omega}\,(x-s) \end{pmatrix}. \tag{10}$$

Розв'язок задачі (8),(9) будемо шукати у вигляді [3]

$$\bar{X}(x, \omega) = B(x, 0, \omega) \cdot \bar{C}, \tag{11}$$

де $\bar{C}$ – деякий нетривіальний вектор. Для його визначення підставимо зображення (11) в крайові умови (9) і, після елементарних спрощень, отримаємо рівність

$$\left[\begin{pmatrix} \alpha_0 & -\lambda \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ \alpha_p & \lambda \end{pmatrix} \cdot B(p, 0, \omega) \right] \cdot \bar{C} = \bar{0}. \tag{12}$$

Ця рівність – рівняння для визначення нетривіального вектора $\bar{C} = (c_1, c_2)^T$. Для існування вектора $\bar{C}$ необхідно і досить виконання умови

$$Det\left[\begin{pmatrix}\alpha_0 & -\lambda\\ 0 & 0\end{pmatrix}+\begin{pmatrix}0 & 0\\ \alpha_p & \lambda\end{pmatrix}\cdot B(p,0,\omega)\right]=$$

$$\begin{vmatrix}\alpha_0 & -\lambda\\ \left(\alpha_p\cos\sqrt{\omega}\,p-\lambda\sqrt{\omega}\sin\sqrt{\omega}\,p\right) & \left(\alpha_p\dfrac{\sin\sqrt{\omega}\,p}{\sqrt{\omega}}+\lambda\cos\sqrt{\omega}\,p\right)\end{vmatrix}=$$

$$\left(\alpha_0\lambda+\alpha_p\lambda\right)\cos\sqrt{\omega}\,p+\left(\frac{\alpha_0\alpha_p}{\sqrt{\omega}}-\lambda\sqrt{\omega}\right)\sin\sqrt{\omega}\,p=0. \quad (13)$$

Відомо [1], що корені ω_k, $k=1,2,...$ характерристичного рівняння (13), як власні значення задачі (11), (12) або еквівалентної до неї задачі на власні значення (6), (7) – додатні та різні. Підставивши ці корені в (12), прийдемо до однорідної системи лінійних рівнянь для визначення невідомих координат c_1, c_2 вектора $\bar{C}$:

$$\begin{cases}\alpha_0\cdot c_1-\lambda c_2=0,\\ \left(\alpha_p\cos\sqrt{\omega_k}\,p-\lambda\sqrt{\omega_k}\sin\sqrt{\omega_k}\,p\right)\cdot c_1+,\end{cases}$$

$$+\left(\alpha_p\frac{\sin\sqrt{\omega_k}\,p}{\sqrt{\omega_k}}+\lambda\cos\sqrt{\omega_k}\,p\right)\cdot c_2=0, \quad (14)$$

звідки отримуємо $c_2=1,\quad c_1=\dfrac{\lambda}{\alpha_0}$.

Отже,

$$\bar{C}=\left(\frac{\lambda}{\alpha_0},1\right)^T. \quad (15)$$

На основі зображення (11), з точністю до сталого множника, отримуємо вираз для власного вектора $\bar{X}_k(x,\omega_k)$:

$$\bar{X}_k(x,\omega_k)=B(x,0,\omega_k)\cdot\begin{pmatrix}\dfrac{\lambda}{\alpha_0}\\ 1\end{pmatrix}=$$

$$\begin{pmatrix}\cos\sqrt{\omega_k}\,x & \dfrac{\sin\sqrt{\omega_k}\,x}{\sqrt{\omega_k}}\\ -\sqrt{\omega_k}\sin\sqrt{\omega_k}\,x & \cos\sqrt{\omega_k}\,x\end{pmatrix}\cdot\begin{pmatrix}\dfrac{\lambda}{\alpha_0}\\ 1\end{pmatrix}=$$

$$\begin{pmatrix}\dfrac{\sin\sqrt{\omega_k}\,x}{\sqrt{\omega_k}}+\dfrac{\lambda}{\alpha_0}\cos\sqrt{\omega_k}\,x\\ -\dfrac{\lambda\sqrt{\omega_k}}{\alpha_0}+\cos\sqrt{\omega_k}\,x\end{pmatrix}. \quad (16)$$

Власні функції $\bar{X}_k(x,\omega_k)$ задачі (6), (7) – це перші координати власних векторів $\bar{X}_k(x,\omega_k)$:

$$\bar{X}_k(x,\omega_k)=\frac{\sin\sqrt{\omega_k}\,x}{\sqrt{\omega_k}}+\frac{\lambda}{\alpha_0}\cos\sqrt{\omega_k}\,x. \quad (17)$$

Підставивши ω_k в рівність (5) замість ω, для визначення $Y_k(y)$, $k=1,2,...$, отримаємо диференціальні рівняння:

$$Y_k''-\omega_k Y_k=0,\quad k=1,2,... \quad (18)$$

Загальний розв'язок рівняння (18) має вигляд

$$Y_k(y)=\alpha_k ch\sqrt{\omega_k}\,y+\beta_k sh\sqrt{\omega_k}\,y,\quad k=1,2,..., \quad (19)$$

де α_k, β_k – деякі сталі.

Підставляючи функції X_k, Y_k $k=1,2,...$ в (4) та сумуючи всі розв'язки такого вигляду, дістанемо розв'язок рівняння (1), що справджує нульові крайові умови при $x=0$ і $x=p$:

$$U_1(x,y)=\sum_{k=1}^{\infty}\left(\alpha_k ch\sqrt{\omega_k}\,y+\beta_k sh\sqrt{\omega_k}\,y\right)\cdot X_k(x,\omega_k). \quad (20)$$

Доберемо тепер коефіцієнти α_k, β_k в (20) так, щоб цей розв'язок справджував умови (2). Для цього припустимо, що функції $a(x)$ і $A(x)$ розвиваються в рівномірно збіжні ряди Фур'є за власними функціями $X_k(x,\omega_k)$:

$$a(x)=\sum_{k=1}^{\infty}a_k X_k(x,\omega_k),\quad A(x)=\sum_{k=1}^{\infty}A_k X_k(x,\omega_k). \quad (21)$$

Покладаючи в рівності (20) $y=0$ та використовуючи умову (3) і розвинення в ряд Фур'є $a(x)$ з (21), послідовно отримуємо:

$$U_1(x,0)=\sum_{k=1}^{\infty}\alpha_k X_k(x,\omega_k);$$

$$\frac{\partial U_1}{\partial y}=\sum_{k=1}^{\infty}\left(\alpha_k ch\sqrt{\omega_k}\,y+\beta_k sh\sqrt{\omega_k}\,y\right)X_k(x,\omega_k);$$

$$\left.\frac{\partial U_1}{\partial y}\right|_{y=0}=\sum_{k=1}^{\infty}\beta_k\sqrt{\omega_k}\cdot X_k(x,\omega_k);$$

$$\left(\beta_0 U_1 - \lambda \frac{\partial U_1}{\partial y}\right)\Big|_{y=0} = \beta_0 \sum_{k=1}^{\infty} \alpha_k X_k(x,\omega_k) -$$

$$-\lambda \sum_{k=1}^{\infty} \beta_k \sqrt{\omega_k} X_k(x,\omega_k) =$$

$$\sum_{k=1}^{\infty} (\beta_0 \alpha_k - \lambda \beta_k \omega_k) X_k(x,\omega_k) = a(x). \qquad (22)$$

З (22) отримуємо

$$\beta_0 \alpha_k - \lambda \beta_k \omega_k = a_k. \qquad (22')$$

Аналогічно, покладаючи в (20) $y = q$ та використовуючи другу з умов (3) і рівняння для $A(x)$, з (21) маємо:

$$U_1(x,q) = \sum_{k=1}^{\infty} \left(\alpha_k ch\sqrt{\omega_k} q + \beta_k sh\sqrt{\omega_k} q\right) X_k(x,\omega_k);$$

$$\frac{\partial U_1}{\partial y}\Big|_{y=q} = \sum_{k}^{\infty} \left(\alpha_k \sqrt{\omega_k} sh\sqrt{\omega_k} q + \beta_k \sqrt{\omega_k} ch\sqrt{\omega_k} q\right) X_k(x,\omega_k);$$

$$\left(\beta_q U_1 + \lambda \frac{\partial U_1}{\partial y}\right)\Big|_{y=q} = A(x),$$

або

$$\sum_{k=1}^{\infty} \left(\beta_q \alpha_k + \lambda \beta_k \sqrt{\omega_k}\right) \cdot ch\sqrt{\omega_k} q \cdot X_k(x,\omega_k) +$$

$$+\sum_{k=1}^{\infty} \left(\beta_q \beta_k + \lambda \alpha_k \sqrt{\omega_k}\right) \cdot sh\sqrt{\omega_k} q \cdot X_k(x,\omega_k) =$$

$$= \sum_{k=1}^{\infty} A_k X_k(x,\omega_k). \qquad (23)$$

Прирівнюючи в (23) коефіцієнти Фур'є, отримуємо:

$$\left(\beta_q \alpha_k + \lambda \beta_k \sqrt{\omega_k}\right) \cdot ch\sqrt{\omega_k} q +$$

$$+\left(\beta_q \beta_k + \lambda \alpha_k \sqrt{\omega_k}\right) \cdot sh\sqrt{\omega_k} q = A_k. \qquad (23')$$

Рівності $(22')$ і $(23')$ утворюють систему рівнянь для визначення α_k і β_k. Розв'язуючи цю систему, остаточно отримуємо, що

$$\begin{cases} \alpha_k = \dfrac{a_k + \lambda \sqrt{\omega_k} \cdot \beta_k}{\beta_0}, \\ \beta_k = \dfrac{\beta_0 A_k - a_k \left(\lambda \sqrt{\omega_k} sh\sqrt{\omega_k} q + \beta_q ch\sqrt{\omega_k} q\right)}{L}, \end{cases} \qquad (24)$$

де

$$L = \lambda \sqrt{\omega_k} \left(\lambda \sqrt{\omega_k} sh\sqrt{\omega_k} q + \beta_q ch\sqrt{\omega_k} q\right) +$$

$$+\beta_0 \left(sh\sqrt{\omega_k} q + \lambda \sqrt{\omega_k} ch\sqrt{\omega_k} q\right), \quad k = 1,2,...$$

Отже, $U_1(x,y)$ при $b(y) = B(y) = 0$ подається у вигляді (20), де α_k і β_k визначаються формулами (24).

Якби ми шукали розв'язок задачі (1), (2), (3) при $a(x) = A(x) = 0$, то x і y помінялися б місцями, і в цьому випадку ми дістали б розв'язок у вигляді

$$U_2(x,y) = \sum_{k=1}^{\infty} \left(\gamma_k ch\sqrt{\nu_k} x + \delta_k sh\sqrt{\nu_k} x\right) \cdot Y_k(y,\nu_k), \qquad (25)$$

де ν_k і $Y_k(y,\nu_k)$ – власні значення та власні функції відповідно задачі:

$$Y'' + \nu Y = 0, \quad \begin{cases} \left(\beta_0 Y - \lambda Y'\right)\Big|_{y=0} = 0, \\ \left(\beta_q Y + \lambda Y'\right)\Big|_{y=q} = 0, \end{cases}$$

а коефіцієнти γ_k і δ_k визначаються за формулами, що аналогічні до (24).

Той факт,що розв'язки $U_1(x,y), U_2(x,y)$ справджують нульові крайові умови в кутових точках прямокутника, не обмежує загальності проведених вище міркувань. Дійсно, побудуємо гармонічну функцію

$$W(x,y) = A + Bx + Cy + Dxy \qquad (26)$$

і доберемо сталі A,B,C,D так, щоб виконувались певні умови узгодженості у вершинах прямокутника, як це зроблено, наприклад, у роботі [2].

III. Висновки

Розв'язана третя крайова задача для рівняння Лапласа в прямокутнику у вигляді розвинення в ряди Фур'є за власними функціями відповідних задач на власні значення. До дослідження таких задач вперше застосовано метод зведення відповідної задачі на власні значення для диференціального рівняння другого порядку до еквівалентної системи диференціальних рівнянь першого порядку.

Література References

[1] Г. М. Положій. Рівняння математичної фізики. Київ: Радянська школа, 1959, 478 с.

[2] Р. М. Таций, О. М. Трусевич. "Прямий метод розв'язування першої загальної крайової задачі для рівняння теплопровідності в прямокутнику" в *Вісник ЛДУ БЖД*", 2016. №13, С. 149-154.

[3] Р. М. Тацій, М. Ф. Стасюк, О. О. Власій. "Загальна перша крайова задача для рівняння теплопровідності з кусково-змінними коефіцієнтами" в *Вісник НУ «Львівська політехніка»: Сер. «Фізико-математичні науки,* 2014, №804, С. 64–69.

Оптимізація на сферично розташованих комбінаторних множинах

О.С. Пічугіна
кафедра прикладної математики
Харківський національний
університет радіоелектроніки
Харків, Україна
pichugina_os@mail.ru

Optimization over spherically located combinatorial sets

O.S. Pichugina
Department of Applied Mathematics
Kharkiv National University
of Radio Electronics
Kharkiv, Ukraine
pichugina_os@mail.ru

***Анотація*—В роботі представлено підхід до розв'язання оптимізаційних задач на вершинно розташованих множинах, що вписані у гіперсферу та дозволяють ефективне розв'язання лінійних задач. Продемонстровано застосовність даного методу до нелінійних задач на таких комбінаторних класах як загальна множина переставлень, множина розміщень з повтореннями, що індукована двома числами, множини парних переставлень та парних булевих векторів. Показано шляхи розповсюдження результатів на інші класи вершинно розташованих множин.**

***Abstract*—In the paper, an approach to optimization over vertex located sets, lying on a hypersphere and enabling efficient solving linear problems, is presented. Applicability of this method to nonlinear optimization over combinatorial classes such as the general permutation set, the set of partial permutations with repetitions induced by two numbers, the even permutation set and a vertex set of a co-cube is demonstrated. Ways to extend the results to other classes of vertex located sets are shown.**

Ключові слова—комбінаторна оптиміація, лінійна оптимізація, метод умовного градієнта, проекція на множину, вершинно розташована множина, сферично розташована множина, добре описана множина, загальна множина переставлень, множина розміщень з повтореннями, парні переставлення, парні бульові векторів

Keywords—combinatorial oprimization, linear oprimization, the conditional gradient method, projection onto a set, a vertex located set, a spherically located set, a well-described set, the general permutation set, the set of partial permutations with repetitions, the even permutation set, a co-cube vertex set.

I. Вступ

Дискретні задачі оптимізації полягають у виборі найкращого об'єкта чи варіанта дії із скінченої чи зліченої кількості варіантів. Вони виникають у таких областях як математика, економіка, фінанси, геометричне проектування, штучний інтелект, програмної інженерія, тощо. В класі оптимізаційних комбінаторні задач традиційно вважаються найскладнішими і значна кількість серед них є NP-складними. Тому актуальним є виокремлення класів дискретних, зокрема комбінаторних, задач, для яких можливо створення ефективних методів оптимізації на базі дослідження властивостей допустимих множин.

На сьогоднішній момент найбільш досліджена тема – різноманітні евристичні підходи, які досить часто на практиці працюють достатньо добре. Проблемою же є те, що оцінити точність цих результатів зазвичай неможливо. Тому цікавим напрямком досліджень є також розробка наближених методів комбінаторної оптимізації із оцінкою точності. Особливу важливість мають спеціальні методи оптимізації, що використовують специфіку конкретної комбінаторної множини та властивості конкретних цільових функцій на цих множинах.

II. Постановка задачі

Розглянемо задачу оптимізації

$$f(x) \to min, \quad (1)$$

$$x \in E \subset R^n, \quad (2)$$

де E - скінченна множина й функція $f(x)$, визначена на ній.

У якості E виберемо дискретну множину одного з наступних класів [1]-[5]:

$E_{nk}(G)$ - загальна множина перестановок із n-мультимножини G, яка містить k різних елементів:

$$G = \{g_i\}_{i \in J_n},\ g_i \le g_{i+1},\ i \in J_{n-1} = \{1,...,n-1\};$$

$\overline{E}_2^n(G)$ - множина розміщень із повтореннями, що індукована двома елементами;

$E_n^e(G)$ - множина парних перестановок із множини G;

B_n^h - множина парних булевих n-векторів;

$E_{nk}^{\pm}(G)$ - множина, сформована відображенням $E_{nk}(G)$ відносно всіх координатних площин;

Ми пропонуємо загальний підхід до розв'язання задачі (1), (2) на п'яти зазначених класах комбінаторних множин:

$$E \in \left\{ E_{nk}(G), \overline{E}_2^n(G), E_n^e, B_n^h, E_{nk}^{\pm}(G) \right\}. \quad (3)$$

Отже, розглядається задача (1)-(3) і пропонується метод її наближеного розв'язання з оцінкою точності, що ґрунтується на застосуванні наступних загальних характеристик даних класів комбінаторних множин: сферичне розташування (Властивість 1), вершинне розташування (Властивість 2); добре описуваність (Властивість 3).

Уведемо необхідні поняття.

Множину E будемо називати <u>сферично розташованою</u>, якщо вона вписана в деяку сферу [6]:

$$\mathrm{E} \subseteq \mathrm{S}_r(a), \quad (4)$$

де

$$\mathrm{S}_r(a) = \left\{ x \in R^n : \|x-a\|_2^2 = r^2 \right\}. \quad (5)$$

Множина точок E називається <u>вершинно розташованою</u>, якщо вона збігається з множиною вершин своєї опуклої оболонки [7]:

$$E = vert\ P, \quad (6)$$

де

$$P = conv\ \mathrm{E}\ . \quad (7)$$

Множина точок E - добре описувана (a well described set) [8], якщо лінійна задача на цій множині

$$\mathbf{LP(E,c)}: z^{lin,E} = min\ \mathrm{c}^T x,\ x^{lin,E} = arg\,min\ \mathrm{c}^T x.$$

ефективно розв'язна (тобто може бути розв'язана за поліноміальний час).

Якщо для E виконані умови (4) і (6), неважко бачити, що вона утворюється у перетині многогранника (7) зі сферою (5):

$$E = \mathrm{S}_r(a) \cap P. \quad (8)$$

Тому множини вигляду (8) називають поліедрально-сферичними [6].

У п. IV пропонується загальний підхід (далі **Метод А**) до розв'язання задач (1), (2) на добре описуваних поліедрально-сферичних множинах (далі **Задача 1**). Даний метод суттєво використовує можливість уопуклення [7],[9],[10] цільової функції, тому п. III головним чином присвячений теорії опуклих продовжень з таких множин. Нарешті, у п. V показано, що множини (3) мають Властивості 1-3, тобто є одночасно поліедрально-сферичними та добре описуваними, отже Метод А застосовний до оптимізації на них. У п. VI наведено інші класи комбінаторних множин, для яких цей метод також застосовний, а також показано напрямки формування нових класів поліедрально-сферичних та добре описуваних множин із множин вигляду (3).

III. Властивості поліедрально-сферичних множин

Як було зазначено в постановці задачі, будемо вважати, що нам задано множину E та деяку функцію $f(x)$, визначену на ній. Якщо до того ж E - вершинно розташована, то при розв'язанні задачі (1), (2) з'являється можливість її зведення до оптимізації опуклої функції в результаті уопуклення вихідної цільової функції. Це обґрунтовується наступними теоретичними фактами:

Визначення. Функція $F(x)$ називається продовженням функції $f(x)$ з E на $X \supset E$, якщо

$$F(x) \text{ визначена на } X \quad (9)$$

і збігається з $f(x)$ на E:

$$F(x):\ F(x) \underset{E}{=} f(x). \quad (10)$$

Теорема 1 [7]. Якщо для E виконана умова (6), то існує опукле диференційоване продовження $F(x)$ функції $f(x)$ з E на довільний опуклий компакт $X \supset E$

Це означає існування $F(x)$, яка задовольняє (9), (10):

$$F(x) \text{ - опукла, диференційована на } X. \quad (11)$$

Наслідок 1. Якщо для E виконано (6), то існує неперервно диференційоване продовження $F(x)$ функції $f(x)$ з E на довільний опуклий компакт $X \supset E$.

Якщо E не тільки вершинно розташована, а ще й поліедрально-сферична, з неї існують двічі неперервно диференційовані продовження. Дійсно, мають місце наступні теореми:

Теорема 2 [11]. Якщо $f(x) \in C^2(R^n)$ й E - сферично розташована, то опукле (сильно опукле з параметром $\rho > 0$) продовження функції $f(x)$ з E на опуклий компакт $X \supset E$ існує у формі:

$$F(x) = f(x) + \lambda\left((x-a)^2 - r^2\right), \text{ де } \lambda \in R.$$

Теорема 3. Якщо E - поліедрально-сферична множина, то існує двічі неперервно диференційоване опукле продовження $F(x)$ функції $f(x)$ з E на довільний компакт $X \supset E$.

Доведення. Згідно з наслідком 1, існує

$$\varphi(x):\ \varphi(x) \underset{E}{=} f(x),\ \varphi(x) \in C^1(R^n). \qquad (12)$$

Розглянемо функцію $\xi(x):\ \xi(x) = \varphi^2(x)$. З одного боку, $\xi(x) \underset{E}{=} f(x)$ і $\xi(x)$ також визначена на X, тобто $\xi(x)$ є продовженням $f(x)$ з E на X. З іншого боку, $\nabla\xi(x) = 2\varphi(x)\cdot\nabla\varphi(x)$, звідки, згідно (12), слідує, що $\xi(x) \in C^2(R^n)$. Отже, $\xi(x)$ - двічі неперервно диференційоване продовження $f(x)$ з E на X.

Для $\xi(x)$ виконані припущення теореми 2, отже, існує двічі неперервно диференційоване опукле продовження функції $f(x)$ у формі:

$$F(x) = \xi(x) + \lambda\left((x-a)^2 - r^2\right), \text{ де } \lambda \in R. \qquad (13)$$

Таким чином, ми довели можливість побудови двічі неперервно диференційованих продовжень із поліедрально-сферичних множин за допомогою рівняння описаної сфери (5).

Тут ми тільки вкажемо на можливість одержання нижньої оцінки z^u для z^*:

$$z^* = \min_E f(x),\ x^* = \arg\min_E f(x). \qquad (14)$$

за допомогою опуклого продовження $F(x)$ цільової функції.

Формування нижніх оцінок z^* забезпечується опуклістю $F(x)$ та еквівалентністю задач (1), (2) та (2),

$$F(x) \to min. \qquad (15)$$

Так, наприклад, якщо $X = P$, тобто будується опукле продовження $f(x)$ на многогранник (7), то $z^u \geq z^P$, де

$$z^P = \min_P F(x),\ x^P = \arg\min_P F(x). \qquad (16)$$

Задача (16) являє собою поліедральну релаксаційну задачу (далі **Задача 2**) до дискретної задачі (2), (15) (далі **Задача 3**) й може бути розв'язана опуклими методами гладкої оптимізації. Якщо до того E - добре описувана, для її розв'язання може бути застосована модифікація методу умовного градієнта, наведена в п. IV.

Те ж саме стосується і вибору кулі в якості X, наприклад, якщо $X = C_r(a) = conv\, S_r(a)$, то $z^u \geq z^C = \min_{C_r(a)} F(x)$.

Ще одна – сферична релаксація – формується ослабленням в (8) умови належності многограннику. Якщо функція $f(x)$ дозволяє знаходження розв'язку $z^S = \min_S F(x)$, де $S = S_r(a)$, значення z^S можна обрати в якості нижньої оцінки, $z^u \geq z^S$, яка в багатьох випадках точніше за z^C, z^P.

Якщо E має Властивості 1-3, тобто, є поліедрально-сферичною та добре описуваною, то квадратична задача пошуку проекції довільної точки на цю множину розв'язна за поліноміальний час.

Теорема 4. Якщо E - добре описувана поліедрально-сферична множина, то задача пошуку проекції на неї довільної точки $x \in R^n$ може бути розв'язана за поліноміальний час.

Доведення. Нехай $y = Pr_E x$. Згідно визначення проекції точки на множину, y - це найближча точка E до x, тобто вона є розв'язком квадратичної задачі:

$$h(x,y) = (y-x)^2 \to \min_{y \in E}. \qquad (17)$$

Спростимо h, враховуючи сферичне розташування E:

$$h(x,y) = (y-x)^2 = \left((x-a)+(a-y)\right)^2 =$$
$$= (x-a)^2 + 2(a-x)\cdot(y-a) + (a-y)^2 \underset{E}{=}$$
$$= \left((x-a)^2 + r^2 - 2a(a-x)\right) + 2(a-x)\cdot y.$$

Перший доданок у правій частині – константа, тому що a, x - фіксовані вектори. Таким чином, задача (17) еквівалентна задачі $(a-x)\cdot y \to \min_{y \in E}$. Остання – це $LP(E, a-y)$, яка, за припущенням, розв'язна за поліноміальний час.

IV. Метод А оптимізації на поліедрально-сферичних множинах

Викладемо Метод А розв'язання задачі (1), (2) на множині вигляду (8), для якої відомий поліноміальний алгоритм *MLP(E,c)* розв'язання *LP(E,c)* (далі **Задача 4**):

Етап 1. Побудувати опукле диференційоване продовження $F(x)$ (наприклад, у формі (13)) і перейти від Задачі 1 до еквівалентної Задачі 3 з опуклою цільовою функцією.

Етап 2. Перейти до поліедральної релаксації Задачі 3 та розв'язати отриману Задачу 2.

Зауваження 1. Оскільки Задача 2 - опукла, її глобальний розв'язок може бути знайдений методами опуклої оптимізації, такими як метод проекції градієнта [12]. Проблема ж полягає в тому, що ці методи оперують із лінійної системою обмежень многогранника, тобто з H -представленням многогранника P . Таким чином, для їх застосування потрібно попередньо розв'язати задачу пошуку такого представлення, яка сама по собі є нетривіальною проблемою. Саме тому остання є центральною проблемою у поліедральній комбінаториці [13],[14]. Але навіть якщо ця проблема вирішена, обмежень комбінаторних многогранників звичайно настільки багато, що використання стандартних методів оптимізації стає практично неможливим. Так, наприклад, серед множин (3) тільки гіперкуб $\overline{\Pi}_2^n(G) = conv\overline{E}_2^n(G)$ містить поліноміальну кількість обмежень, інші ж класи включають множини, H -представлення яких задаються неполіноміальним числом обмежень [1]-[4],[15].

Обчислювальна схема розв'язання Задачі 2 заснована на диференційованості $F(x)$, застосуванні Властивості 3 та V -представлення (7) многогранника P .

Скористаємося модифікацією [16] методу умовного градієнта [12], суть якої полягає у тому, що розв'язання допоміжних лінійних Задач 4 здійснюється за допомогою *MLP(E,.)* .

У ході ітераційного процесу розв'язання Задачі 2 формується дві послідовності:

1) наближених розв'язків Задачі 2:

$$\left\{x^k\right\}_{k\in Z_+} : x^k \underset{k\to\infty}{\to} x^P, \mathrm{z}^{Pk} = F\left(x^k\right) \underset{k\to\infty}{\to} z^P;$$

2) наближених розв'язків Задач 1, 3:

$$\left\{y^k\right\}_{k\in Z_+} : y^k = Pr_E\, x^k, \mathrm{z}^k = \min_{i\in J_k^0} F\left(y^k\right) = \min_{i\in J_k^0} f\left(y^k\right)$$

де $J_K^0 = J_K \cup \{0\}$.

Відзначимо, що при цьому розв'язок Задачі 2 покращується на кожній ітерації, у той час як розв'язок Задач 1,3 – по мірі того як будуть знайдені нові рекорди цільової функції.

Для застосування методу умовного градієнта, потрібна допустима точка x^0 області в якості початкового наближення. Ми пропонуємо вибирати за x^0 центр описаної сфери (5) мінімального радіуса: якщо $a \in P$, то

$$\mathrm{x}^0 = a\ . \tag{18}$$

Інакше можна обрати довільну точку E .

Процес закінчується по досягненні заданої точності ε :

$$K: \left|z^{Pk} - z^{P,k-1}\right| > \varepsilon\ \forall \mathrm{k} \in \mathrm{J}_{K-1},\ \left|z^{PK} - z^{P,K-1}\right| \le \varepsilon\ .$$

$\left(x^K, z^{PK}\right)$ - наближений розв'язок Задачі 2, найкращий з $\left(y^K, z^K\right)$ - Задач 3 і, відповідно, Задачі 1:

$$\left(x^P, z^P\right) \approx \left(y^K, z^{PK}\right),$$
$$\left(x^*, z^*\right) \approx \left(x^{k^*}, z^{k^*}\right) : z^{k^*} = \min_{k\in J_K^0} z^k, x^{k^*} = arg\ min\ z^{k^*}.$$

Відносна похибка розв'язку Задачі 1 - $\delta = \left|\frac{z^{PK} - z^K}{z^K}\right|$.

Якщо $x^P \in int\, P$, оцінка точності може бути поліпшена за рахунок z^S (див. п. III).

V. Обґрунтування застосовності методу А для оптимізації на множинах (3)

Обґрунтуємо можливість використання Методу А до розв'язання Задачі 1 на множинах (3). Для цього достатньо показати, що всі вони мають Властивості 1-3.

A. Властивість 1.

При демонстрації сферичного розташування кожної з множин E вигляду (3), наведемо параметри сфери мінімального радіуса $S^{min}(E)$, що описана навколо множини E . Для цих класів комбінаторних множин центри даних гіперсфер є також допустимими точками многогранника, тому вони можуть бути використані в якості x^0 (див. (18)).

1) $E_{nk}(G)$ - загальна множина перестановок з мультимножини $\mathrm{G}=\left\{\mathrm{e}_1^{n_1},\ldots,\mathrm{e}_\mathrm{k}^{n_\mathrm{k}}\right\} : \sum_{\mathrm{i}=1}^{\mathrm{k}} n_\mathrm{i} = n$, $e_\mathrm{i} < \mathrm{e}_{\mathrm{i}+1}$, $\mathrm{i} \in \mathrm{J}_{k-1} = \{1,\ldots,k-1\}$.

$S^{min}\left(E_{nk}(G)\right)$ має параметри [2]:

$$a = \left(\left(\frac{S_1}{n}\right)^n\right)^T,\ r = \sqrt{S_2 - \frac{S_1^2}{n}}, \tag{19}$$

де $S_1 = \sum_{i=1}^{n} g_i$, $S_2 = \sum_{i=1}^{n} x_i^2$.

2) $\overline{E}_2^n(G)$ - множина розміщень із повтореннями з $\mathrm{G}=\{\mathrm{e}_1^n,\mathrm{e}_2^n\}$, $e_1 < \mathrm{e}_2$: параметри $S^{min}\left(\overline{E}_2^n(G)\right)$ -

$$a=\left(\frac{e_1+e_2}{2}\right)^n,\ r=\frac{e_2-e_1}{2}\sqrt{n}. \qquad (20)$$

3) $E_n^e(G)$ - множина парних перестановок з n-елементної множини G, тобто перестановок з множини G з парним числом інверсій. $E_n^e(G)$ - підмножина множини перестановок $E_n(G)$ з тієї ж самою сферою мінімального радіуса (19), що й $E_n(G)$.

4) B_n^h - множина парних булевых векторів, тобто таких елементів булевої множини $B_n=\{0,1\}^n$, сума координат яких – парна. У цьому випадку $S^{min}\left(B_n^h\right)=S^{min}\left(B_n\right)$, а параметри $S^{min}\left(B_n\right)$ визначаються з (20), враховуючи що $B_n=\overline{E}_2^n\left(\{0^n,1^n\}\right)$:

$$a=(0.5)^n,\ r=\frac{\sqrt{n}}{2}.$$

5) $E_{nk}^{\pm}(G)$ відноситься до класу композиційних образів розміщень і перестановок і є множиною перестановок розміщень [17]. Для її формування на першому етапі з кожної з мультимножин $G^i=\{\pm g_i\},\ i\in J_n$ обирається по одному елементу, а потім зі сформованої n-мультимножини формується загальна множина перестановок.

Параметри $S^{min}\left(E_{nk}^{\pm}(G)\right)$ наступні:

$$a=\mathbf{0},\ r=\sqrt{\sum_{i=1}^{n} g_i^2}. \qquad (21)$$

Зауваження 2. Співпадіння описаних мінімальних сфер для пар $E_n^e(G)$ і $E_n(G)$, а також для B_n^h й B_n випливає з того, що вимірність відповідних многогранників не зменшується при переході від перестановок (булевих векторів) до парних перестановок (парних булевих векторів).

B. Властивість 2.

Вершинне розташування всіх перерахованих вище комбінаторних множин обґрунтуємо в наступній теоремі:

Теорема 5. Довільна скінченна сферична множина – вершинно розташована.

Доведення. Припустимо, що $E\subset R^n$ задовольняє (4) й:

$$|E|<\infty. \qquad (22)$$

Гіперсфера $\mathrm{S}_r(a)$ являє собою границю опуклого тіла – кулі $\mathrm{C}_r(a)$. Легко бачити, що $\mathrm{S}_r(a)$ - множина крайніх точок $\mathrm{C}_r(a)$, тобто жодна із точок $\mathrm{S}_r(a)$ не представляється опуклою лінійною комбінацією інших точок $\mathrm{S}_r(a)$. Це ж можна сказати й про точки E, оскільки виконане включення (4). У силу (22), опукла оболонка множини E є многогранником. Множина E буде множиною його крайніх точок, а отже і його вершин. Теорему доведено.

Усі множини (3) – скінчені, адже вони є містять вибірки зі скінчених мультимножин. З іншого боку, усі вони – сферично розташовані. Таким чином, згідно Теореми 5, вони є вершинно розташованими.

C. Властивість 3.

Для множин $E_{nk}(G)$, $\overline{E}_2^n(G)$ розв'язки лінійних задач відомі [3],[15].

Теорема 6. Якщо

$$\left(i_j\right)_{j\in J_n}:\ \left\{i_j\right\}_{j\in J_n}=\mathrm{J}_n,\ \mathrm{c}_{i_j}\ge \mathrm{c}_{i_{j+1}},\ \mathrm{j}\in \mathrm{J}_{n-1},$$

то

$$x_{i_j}^{lin,E_{nk}(G)}=g_j,\ \mathrm{j}\in \mathrm{J}_n,\ z^{lin,E_{nk}(G)}=\sum_{j=1}^{n} c_{i_j} g_j. \qquad (23)$$

Теорема 7.

$$\forall i\in J_n\ x_i^{lin,\overline{E}_2^n(G)}=\begin{cases} e_1, \text{ якщо } \mathrm{c}_i\ge 0 \\ e_2, \text{ якщо } \mathrm{c}_i<0 \end{cases}, \qquad (24)$$

$$z^{lin,\overline{E}_2^n(G)}=e_1\sum_{i:c_i\ge 0} c_i+e_2\sum_{i:c_i<0} c_i\ .$$

Таким чином, $E_{nk}(G)$, $\overline{E}_2^n(G)$ - добре описувані множини. Метод $MLP(E_{nk}(G),c)$ зводиться до упорядкування координат вектору c із наступним присвоюванням (23). $MLP(\overline{E}_2^n(G),c)$ полягає у порівнянні координат вектору c із нулем і присвоюванні (24).

Для $E_n^e(G)$, B_n^h запропонуємо $MLP(E_n^e(G),c)$, $MLP(B_n^e,c)$, виходячи зі способу їх формування з множин E_n, B_n. Неважко бачити, що у многограннику $\Pi_n(G)$ усі $n-1$ перестановки, суміжні до парної, – непарні й навпаки. Те ж стосується одиничного гіперкуба $PB_n=conv\ B_n$ й напівкуба $PB_n^h=conv\ B_n^h$: у многограннику PB_n кожний парний булевий вектор має n суміжних непарних булевих векторів і навпаки.

Окрім цього, критерії суміжності $\Pi_n(G)$, PB_n відомі [1],[15]:

- у многограннику перестановок $\Pi_n(G)$ суміжними з $x \in E_n$, є точки, отримані перестановкою g_i й g_{i+1} тільки вони;

- в одиничному гіперкубі PB_n суміжними з $x \in B_n$, є точки, отримані із x заміною однієї нульової координати одиницею або однієї одиничної координати на 0.

Перераховані особливості дозволяють запропонувати єдину схему розв'язання $LP(E,c)$ для

$$E = E_n^e, B_n^h. \quad (25)$$

1) Схема розв'язання $LP(E,c)$:

Крок 1. Розв'язати $LP(E',c)$ на надмножині E' множини E, $E' \supset E$. Якщо $x^{lin,E'} \in E$, задачу $LP(E,c)$ розв'язано: $x^{lin,E} = x^{lin,E'}$, інакше перейти до Кроку 2:

Крок 2. Знайти множину $N_{P'}\left(x^{lin,E'}\right)$ суміжних вершин до $x^{lin,E'}$ у многограннику $P' = convE'$, вибрати мінімальне зі значень $c^T x$ на $N_{P'}\left(x^{lin,E'}\right)$:

$$z^{lin,E} = \min_{x \in N_{P'}\left(x^{lin,E'}\right)} c^T x,\ x^{lin,E} = \arg\min_{x \in N_{P'}\left(x^{lin,E'}\right)} c^T x.$$

Ясно, що даний алгоритм є поліноміальним, якщо E' - добре описувана й кількість суміжних вершин до кожної вершини – поліноміально залежить від n. При застосуванні до множин (25), він перетворюється на $MLP(E_n^e(G),c)$, $MLP(B_n^e,c)$ відповідно.

Отже, ми повністю обґрунтували можливість застосування Методу А до п'яти множин (3).

VI. ОБГОВОРЕННЯ РЕЗУЛЬТАТІВ

Метод А може бути безпосередньо використаний для інших підмножин та окремих випадків $E_{nk}(G)$, $\overline{E}_2^n(G)$, на яких легко розв'язати лінійну задачу. Серед підмножин $E_{nk}(G)$ це, наприклад, множина поліперестановень, серед підмножин $\overline{E}_2^n(G)$ - загальна множина розміщень [15], індукована двома числами, серед підмножин B_n - множина переставних матриць Π_n [18], серед спеціальних класів $E_{nk}(G)$ та B_n - множина *0 – 1* -переставлень $B_n(m)$ [16], тощо.

Слід відзначити, що Декартів добуток вищеозначених множин зберігає Властивості 1-3, тож у різних комбінаціях множин (3) ми отримуємо нові класи комбінаторних множин, до оптимізації на яких можна застосувати також Метод А.

Метод А може бути узагальнено у декількох напрямках. Наприклад, якщо множина E – вершинно розташована, але не є поліедрально-сферичною, тоді, побудувавши для неї опукле диференційоване продовження (яке існує згідно Теореми 1) і замінивши задачу проектування на комбінаторну множину E задачею комбінаторного заокруглення до елемента E, ми отримаємо одне таке узагальнення (далі **Модифікація 1**).

Якщо легко знайти наближений розв'язок $LP(E,c)$, можна замінити ним точку $x^{lin,E}$ в Методі А, а також використати цей прийом наближеного розв'язку лінійної задачі при проектуванні на множину. В результаті отримаємо ще одне узагальнення, яке в дає наближений розв'язок задач 1-3. Нажаль оцінки точності в такому випадку дати неможливо.

Нарешті, якщо про множину E мало що відомо і навіть наближений розв'язання лінійної задачі викликає труднощі, тоді, замість поліедральній релаксації, можна перейти до релаксації на кулі $C_r(a)$, адже її межа – сфера $S_r(a)$ - є добре описуваною множиною. В результаті такої заміни буде одержано

$$z^C = \min_C F(x),\ x^C = \arg\min_C F(x)$$

так наближений розв'язок задачі 1, причому z^C служитиме нижньої оцінкою для z^*.

Представлені модифікації дозволяють розповсюдити результаті статті ще на один клас множин – це прямі суми поліедрально-сферичних множин. В результаті цієї операції формується множна, що є еліпсоїдальною (тобто такою, що вписана в еліпсоїд), опукле продовження з якої можна побудувати за допомогою рівняння еліпсоїда.

Це означає, що до довільної послідовності прямих сум і добутків вищеозначених множин застосовна Модифікація 1. Так, наприклад, цілий клас множин вершин многогранників Ханнера [19] утворюється в результаті послідовностей прямих сум і Декартових добутків бінарної множини $\{-1,1\}$, що є окремим випадком $\overline{E}_2^1(G)$ для $G = \{-1,1\}$.

Перерахуємо ще деякі класи вершинно розташованих множин, дослідження екстремальних властивостей лінійних функцій на яких дозволить використати Метод А та його модифікації до оптимізації на них. Ці множини являють собою підмножини та спеціальні класи множин $E_{nk}(G)$, B_n, тож за конструкцією вони є поліедрально-сферичними.

Так для $E_{nk}(G)$ серед таких множин:

1. $E_n(G)$ - підмножина – множина $E_n^c(G)$ циклічних

переставлень, тобто таких, що містять єдиний цикл [20],[21];

2. $E_{nk}(G)$-підмножина – множина почережних переставлень $E_n^p(G)$, координати яких формують послідовність, у якій на парних місцях елементи, що не перевищують двох сусідніх, на непарних – що не менше обох сусідніх [21];
3. $E_{nk}(G)$-підмножина – множина повних переставлень $E_{nk}(G,x)$ до заданого переставлення $x \in E_{nk}(G)$. Вона включає ті переставлення, жодна координата яких не збігається з відповідною координатою x [21];
4. Підмножиною $E_n(G)$ та узагальненням множини $E_n^e(G)$ є множина $E_n(G,m)$ переставлень, що розташовані у $E_n(G)$ на відстані $m \geq 1$. Зокрема $E_n^e(G) = E_n(G,1)$.

Також відомі численні B_n-підмножини, опуклі оболонки яких називаються 0-1-політопами [22]-[24], відповідно, ці множини будуть множинами вершин 0-1-політопів. Серед них відзначимо наступні:

1. Множина 0-1-розміщень $B_n(m_1,m_2)$ ($0 \leq m_1 < m_2 \leq n$), кількість одиничних координат яких обмежена числами m_1, m_2 [16];
2. Множина вершин простого 0-1-політопа, тобто такого, вершини якого утворені у перетині гіперграней, кількість яких збігається з вимірністю многогранника [23]. Цей клас множин має таку властивість, що він утворюється як декартів добуток множин $B_n(1)$;
3. Множина вершин дворівневого 0-1-многогранника, тобто такого, вершини якого розкладаються по двох паралельних гіпергранях в напрямку кожної своєї гіперграні [24]. Це дуже широкий клас, адже відомо, що кожен 2-рівневий многогранник афінно-еквівалентний деякому 2-рівневому 0-1-многограннику. Серед таких многогранників, окрім многогранників Ханнера, Біркгофа (що є опуклою оболонкою Π_n), є також многогранники порядку, Хансена і т.п. [24].

VII. Висновки

В даній роботі представлено метод розв'язання нелінійних задач на широкому класі поліедрально-сферичних комбінаторних множин. Обґрунтовано застосовність даного методу та його модифікацій до оптимізації на довільних поліедрально-сферичних множинах, зокрема до загальної множини переставлень, множини розміщень із повтореннями, індукованої двома числами, а також на їхніх Декартових добутках, прямих сумах та деяких підмножинах.

Література References

[1] V. A. Yemelichev, M. M. Kovalev, M. K. Kravtsov, Polytopes, graphs and optimization, Cambridge: Cambridge University Press, 1984.

[2] Y. G. Stoyan, S. V. Yakovlev, Mathematical models and optimization methods in Geometric Design, Kiev: Naukova Dumka, 1986.

[3] Y. G. Stoyan, O. A. Yemets, Theory and methods of Euclidean combinatorial optimization, Kiev: ISSE, 1993.

[4] A. Postnikov, "Permutohedra, associahedra, and beyond," IMRN: International Mathematics Research Notices. 6, pp. 1026-1106, 2009.

[5] R. M. Green, "Homology representations arising from the half cube, II," Journal of Combinatorial Theory: Series A, 117(8), pp. 1037-1048, 2010.

[6] O.S. Pichugina, S.V. Yakovlev, "On the continuous representations and functional extensions in combinatorial optimization," Cybernetics and Systems Analysis, 52(6), pp. 102-113, 2016.

[7] S. V. Yakovlev, "The theory of convex continuations of functions on vertices of convex polyhedra," Computational Mathematics and Mathematical Physics, 34(7), pp. 1112-1119, 1994.

[8] Y. Berstein, J. Lee, S. Onn, R. Weismantel, "Parametric nonlinear discrete optimization over well-described sets and matroid intersections,"Mathematical Programming, 124(1/2), pp. 233-253, 2010.

[9] S. V. Yakovlev, "Bounds on the minimum of convex functions on Euclidean combinatorial sets," Cybernetics and Systems Analysis, 25(3), pp. 385-391, 1989.

[10] H. D. Sherali, W. P. Adams, A reformulation-linearization technique for solving discrete and continuous nonconvex problems, P. M. Pardalos (eds.), Dordrecht: Kluwer Academic Publishers, 1999.

[11] Y. G. Stoyan, S. V. Yakovlev, O. A. Emets, O. A. Valuiskaya, "Construction of convex continuations for functions defined on a hypersphere," Cybernetics and Systems Analysis, 34(2), pp. 176-184, 1998.

[12] D. P. Bertsekas, Nonlinear Programming, 2nd edn. Belmont: Athena Scienific, 1999.

[13] W. R. Pulleyblank, "Edmonds, matching and the birth of polyhedral combinatorics," Documenta Mathematica, Extra Volume ISMP, pp. 181-197, 2012.

[14] P. M. Pardalos, D. Du, R. L. Graham, Handbook of combinatorialoptimization, New York: Springer, 2013.

[15] S. V. Yakovlev, I. V. Grebennik, "Some classes of optimization problems on a set of arrangements and their properties," Izv.Vyssh. Uchebn. Zaved. Mat., 35(11), pp. 74-86, 1991.

[16] O. Pichugina, S. Yakovlev, "Convex extensions and continuous functional representations in optimization, with their applications," J. Coupled Syst. Multiscale Dyn, 4, pp. 129-152, 2016.

[17] Y. G. Stoyan, I. V. Grebennik, "Combinatorial species for the enumeration of combinatorial configurations with special properties," Dopov. Nats.Akad. Nauk Ukr. Mat. Prirodozn. Tekh. Nauki, 7, pp. 28-32, 2010.

[18] R. A. Brualdi, Combinatorial matrix classes, T. Britz (eds.), Cambridge: Cambridge University Press, 2006.

[19] R. Sanyal, A. Werner, G. M. Ziegler, "On Kalai's Conjectures Concerning Centrally Symmetric Polytopes," Discrete & Computational Geometry, 41(2), pp. 183–198, 2008.

[20] J. F. Korsh, P. S. LaFollette, "Loopless array generation of multiset permutations, " The Comp. Journ., 47 (5), pp. 612–621, 2004.

[21] E. W. Weisstein, CRC Concise Encyclopedia of Mathematics, Second Edition, Boca Raton: CRC Press, 2002.

[22] G. M. Ziegler, "Lectures on 0/1-Polytopes," in Polytopes — Combinatorics and Computation, G. Kalai and G. M. Ziegler, Eds. Birkhäuser Basel, 2000, pp. 1–41.

[23] V. Kaibel and M. Wolff, "Simple 0/1-Polytopes," European Journal of Combinatorics, vol. 21, no. 1, pp. 139–144, Jan. 2000.

[24] A. Bohn, Y. Faenza, S. Fiorini, V. Fisikopoulos, M. Macchia, K. Pashkovich, "Enumeration of 2-Level Polytopes," in Algorithms - ESA 2015, N. Bansal and I. Finocchi, Eds. Berlin, Heidelberg: Springer. — 2015, pp. 191–202.

Шаблон Статті МНПК ІТКМ (стиль: ***paper title***)

Підзаголовок, за необхідності ***(paper subtitle)***

Любомир Петришин (*Author*)
кафедра компютерних наук (*Affiliation*)
AGH науково-технологічний університет (*Affiliation*)
Краків, Польща (*Affiliation*)
p.name@mail.com (*Affiliation*)

Артем Ізмайлов (*Author*)
кафедра інформатики (*Affiliation*)
Прикарпатський національний університет
Івано-Франківськ, Україна (*Affiliation*)
a.name@mail.com (*Affiliation*)

ISPC ITCM Article Template (use style: ***paper title***)

Subtitle, as needed ***(paper subtitle)***

Lubomyr Petryshyn (*Author*)
dept. of Computer Science (*Affiliation*)
AGH University of Science and Technology (*Affiliation*)
Krakow, Poland (*Affiliation*)
p.name@mail.com (*Affiliation*)

Artem Izmailov (*Author*)
dept. of Computer Science (*Affiliation*)
Precarpathian National University (*Affiliation*)
Ivano-Frankovsk, Ukraine (*Affiliation*)
a.name@mail.com (*Affiliation*)

***Анотація*—Даний електронний документ є шаблоном, в якому визначено уклад та стилі компонентів статті згідно вимог ©IEEE [назва, текст, заголовки, і т.д.]. *Не використовувати у назві статті чи анотації символи, спеціальні знаки, або математичні вирази*. (*Abstract*)**

***Abstract*—This electronic document is a "live" template and already defines the components of your paper [title, text, heads, etc.] in its style sheet. *CRITICAL: Do Not Use Symbols, Special Characters, or Math in Paper Title or Abstract*. (*Abstract*)**

Ключові слова—компонент; форматування; стиль; оформлення у визначеному стилі (key words)

Keywords—paper; component; formatting; style; styling; insert (key words)

I. Вступ (*Heading 1*)

Шаблон уможливлює авторам застосування стандартних стилів форматування при підготовці електронних версій статей згідно вимог ©IEEE в форматі "Microsoft Word" [1]. Матеріал статті рекомендовано форматувати згідно чинного шаблону з метою:

- спрощення форматування статей авторським колективом, (*bullet list*)
- автоматичного дотримання вимог, що полегшує загальне формування матеріалів у електронній формі, (*bullet list*)
- відповідності стилів усіх статей в збірнику матеріалів конференції. (*bullet list*)

У шаблоні задані ширина колонок, міжрядковий інтервал та формати шрифтів. Шаблон вміщує стилі титулу статті, афіліації, анотації, ключових слів, тексту, заголовків, таблиць, рисунків, формул, літератури. Стилі, які застосовано у даному документі, позначені у відповідних абзацах в дужках (*курсивом*). Авторам необхідно здійснювати форматування рівнянь, графіки і таблиць з урахуванням наступних вказівок.

II. Використання Шаблону (*Heading 1*)

A. Формат сторінки шаблону (Heading 2)

Розміри сторінки матеріалів публікації повинні відповідати формату А4.

B. Підтримка цілісності вимог (Heading 2)

Шаблон має на меті дотримання єдиного форматування статті та стилів тексту публікації. Розміри полів, ширина колонок, міжрядковий інтервал та шрифти задані шаблоном. Верхнє поле у цьому шаблоні має розміри, які відрізняються від загальноприйнятих, що встановлені на основі вимог, згідно яких стаття є окремою складовою цілісних матеріалів конференції, а не самостійним документом.

Вимога не змінювати жодних з налаштувань шаблону.

III. Підготовка статті перед форматуванням

При оформленні матеріалів рекомендовано підготувати текст статті та графіку робочим файлом, попередньо

Вкажіть тут наявного спонсора. Якщо спонсор відсутній, то видаліть цей текстовий блок (*sponsors*).

застосувавши відповідні стилі форматування цього шаблону, або ж сам шаблон. При підготовці статті увімкніть відображення недрукованих символів (знак ¶) та меж робочого поля, що дозволить автору бачити та правильно застосовувати форматування матеріалу. Відредагуйте текст та структуру статті до початку остаточного форматування шаблону. На основі робочого файлу заповніть шаблон та остаточно відформатуйте матеріал, застосовуючи відповідні стилі шаблону згідно авторського укладу. **Важливо**: до редакції подається авторський файл, який є наслідком збереження заповненого матеріалом статті шаблону, а не робочий файл, сформатований стилями шаблону. Шаблоном не передбачено застосування табуляції, тому прохання її не використовувати. Явні переноси рядка за необхідності обмежити одним переносом у кінці абзацу згідно замислу автора. У жодному місці статті не використовувати розбиття тексту на сторінки. Не нумерувати заголовки у тексті - у шаблоні це робиться автоматично шляхом застосування відповідного стилю.

При перевірці орфографії та граматики дотримуватись наступних вимог:

A. Абревіатури та скорочення

Застережено використання чи визначення абревіатур та скорочень у назві статті, анотації чи заголовках. Визначення абревіатур та скорочень здійснюється при першій їх зустрічі в тексті. Загальноприйняті абревіатури, як IEEE, SI, MKS, CGS, sc, dc та rms визначати не потрібно.

B. Одиниці вимірювання

- Для визначення одиниць вимірювання застосовуйте систему СІ (МКС), або, у виключних випадках, СГС. За необхідності можна використовувати англійську систему мір шляхом додаткового визначення одиниць вимірювання (у дужках). Дозволено використання англійської системи мір як комерційних ідентифікаторів, наприклад "3.5-дюймовий привід". (*bullet list*)
- Уникати спільного використання одиниць СІ та СГС, наприклад, струм у амперах, а магнітне поле – у ерстедах, що спричиняє невідповідність розмірностей. (*bullet list*)
- Не змішувати повне написання та скорочення одиниць вимірювання: "Wb/m^2" або "вебер на квадратний метр", але не "вебер/м^2". Повністю визначайте назву одиниці вимірювання у тексті: "...кілька генрі," але не "...кілька Н." (*bullet list*)
- Використовуйте нуль перед десятковою крапкою: "0.25," а не ".25." Використовуйте "cm^3," а не "cc." (*bullet list*)

C. Формули

При написанні формул необхідно застосовувати виключно *похилений* шрифт *Times New Roman* та регулярний Symbol згідно стилю (*equation*). *Курсивом* відображається аналітика, латинські символи величин та змінних, але не символи грецького алфавіту. У форматі формул, за наявності, текст пояснень повинен бути сформатований регулярним шрифтом Times New Roman. Якщо у Вас інші налаштування формульного процесора MS-Equation_Editor_3.1, ніж наведені на рис. 1 та рис. 2, змініть їх до рекомендованих на рисунках значень.

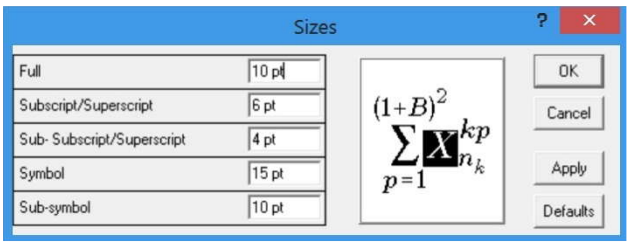


Рис. 1. Рекомендовані налаштування розмірів шрифту в формульному процесорі MS-Equation_Editor_3.1. (*figure caption*)

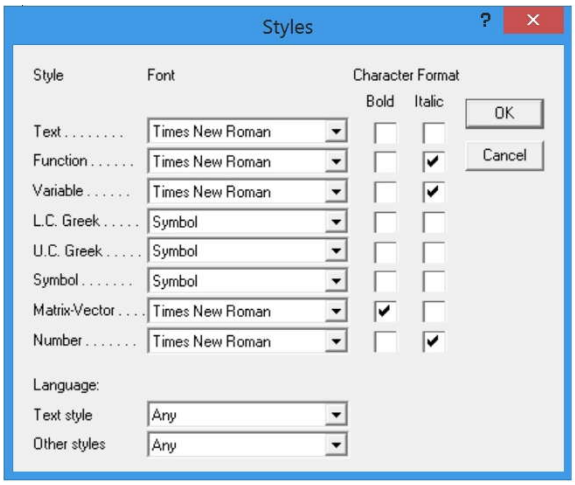


Рис. 2. Рекомендовані налаштування стилів шрифтів в формульному процесорі MS-Equation_Editor_3.1. (*figure caption*)

Застосовуйте центрування формули з використанням табуляції по центру в стилі (*equation*).

$$\sum_{p=1}^{(1+B)^2} aX_{n_k}^{kp} \qquad (1)$$

.

Здійснюйте наскрізну в статті (а не за розділами) нумерацію тільки тих формул, на які є посилання в тексті статті. Номери формул, які зазначаються у дужках, слід вирівнювати вправо, як у прикладі (1), використовуючи відступ вправо з табуляцією. За наявності простих формул, що не вимагають застосування формульного процесора, уникайте його використання, наприклад

$$N = a_{n-1}\, q^{n-1} + \ldots + a_i\, q^i + \ldots + a_0\, q^0, \qquad (2)$$

де q - основа; $a_i = \{q-1, 0\}$ - коефіцієнти.

Для знака мінус використовуйте довге тире –, а не дефіс -. Коли формули є складовими частинами речення, їх в кінці слід відділити комами або крапками, наприклад (1)

та (2). Символи у рівнянні необхідно визначати до (1) або одразу після формули (2). Застосовуйте посилання “(2),” а не “Eq. (2)” чи “equation (2),” за винятком, коли посилання відкриває речення: “Equation (2) is ...”

Для написання програмних кодів рекомендовано застосовувати шрифт із однаковою шириною символів `Courier New`.

D. *Деякі розповсюджені помилки*

- У якості над- чи підрядкових індексів у прикладі (2), та інших розповсюджених констант, наприклад μ_0, слід застосовувати число “0” (нуль) з відповідним форматуванням, а не малу літеру “o”.
- У американському варіанті англійської мови такі розділові знаки, як коми, крапки з комою, крапки, знаки запитання та оклику записуються всередині лапок лише у випадку цитування закінченої думки або імені, наприклад, у випадку назви або повної цитати. Коли лапки використовуються для виділення слова або фрази, пунктуація повинна знаходитись поза лапками. Розділові знаки для фрази або висловлювання у дужках у кінці речення виносяться за межі дужок (як тут). (Розділові знаки у реченні, яке взяте у дужки, залишаються у дужках.)
- У латинському скороченні “et al” після “et” крапка не ставиться. тощо
- Скорочення “i.e.” означає “тобто,” а “e.g.” означає “наприклад.”

Рекомендовано посібник [9] по стилю написання наукових статей.

IV. Використання шаблону

Після завершення редагування статті в робочому файлі необхідно перенести матеріал до шаблону. Або ж здійснити редагування безпосередньо в шаблоні. Для цього копії файлу шаблону за допомогою команди «Зберегти як», присвоїти латиною ім'я файлу статті відповідно із прізвищем, іменем першого автора та скороченою темою статті, наприклад Prizvysche_Imya_Title.doc . Здійснити остаточне форматування матеріалу статті, використовуючи випадаюче вікно зі стилями панелі інструментів шаблону MS Word.

A. *Автори та афіліації*

Шаблон розроблено таким чином, щоб афіліація авторів не повторювалась у випадку, коли співавтори є співпрацівниками одного підрозділу та організації. Прохання вказувати афіліацію якнайкоротше. Даний шаблон розроблено для двох типів афіліацій:

1) Для автора (ів) з однією афіліацією (Heading 3): Щоб змінити налаштування за замовчуванням, змініть шаблон наступним чином:

a) Селекція (Heading 4): Виділити рядки з автором та афіліацією.

b) Змініть число колонок: Вибрати пункт меню MS Word «Формат\Колонки», а наступним вибрати пункт «одна».

c) Видалення: Видалити рядки з автором та афіліацією для другої афіліації.

2) Для автора (ів) з різними афіліаціями: Щоб змінити налаштування за замовчуванням, змініть шаблон наступним чином:

a) Селекція: Виділити рядки з автором та афіліацією.

b) Змінити число колонок: Вибрати пункт меню MS Word «Формат\Колонки», наступним вибрати пункт «одна».

c) Виділити усі рядки з автором та афіліацією 1 і скопіювати це виділення.

d) Форматування: Вставити жорстке перенесення одразу після останньої літери останнього рядка афіліації. Далі вставити копію афіліації 1. У випадку необхідності повторити чинності для кожної додаткової афіліації.

e) Повернення числа колонок: Помістити курсор справа від останньої літери останнього рядка афіліації з парним номером (наприклад, якщо є три афіліації, помістити курсор у кінці другої афіліації). Перемістити курсор вверх, щоб виділити усі рядки з автором та афіліацією. Відкрити пункт меню «Формат\Колонки» та вибрати «дві». Якщо у статті непарна кількість афіліацій, то остання опиниться у центрі сторінки, а попередні будуть розташовані у двох колонках.

f) Рекомендація не перевищувати кількості чотирьох співавторів.

B. *Визначення заголовків*

Заголовки – це структурні елементи, які допомагають читачеві орієнтуватись у структурі статті. Визначено два типи заголовків: структурні та текстові.

Структурні заголовки визначають різні частини статті та тематично не залежать один від одного. Прикладами таких структурних заголовків є ПОДЯКА та ЛІТЕРАТУРА, тому правильним стилем для них є (*Heading 5*). Початкові заголовки “*Анотація*” і “*Ключові слова*” вимагають застосування додаткового форматування (у даному випадку, *курсиву*) закладеного стилю з випадаючого списку для того, щоб відрізнити заголовок від тексту.

Текстові заголовки формують ієрархічну інфраструктуру тематики статті. Назва статті є головним текстовим заголовком, згідно якої здійснено викладку матеріалу статті. Якщо зміст матеріалу передбачає наявність двох чи більше підтем, то необхідно використовувати заголовки нижчого рівня (які позначаються римськими цифрами), і, навпаки, за відсутності принаймні двох підтем, то підзаголовки вводити не слід. Стилі “Heading 1,” “Heading 2,” “Heading 3,” та “Heading 4” прописані у шаблоні за замовчуванням.

C. Рисунки і таблиці

Розташування рисунків і таблиць. Рисунки і таблиці вкладати після їх згадування у тексті. Великі рисунки і таблиці можуть займати формат шириною цілої сторінки, тобто обох колонок. Заголовки таблиць розташовувати над таблицями. Для назв таблиць використовуйте стиль (*Table Head*).

TABLE II. ЗАГОЛОВОК ТАБЛИЦІ (*TABLE HEAD*)

Заголовок таблиці (*table col head*)	**Заголовок стовпця таблиці (*table col head*)**		
	Підзаголовок стовпця таблиці (table col subhead)	***Підзаголовок (table col subhead)***	***Підзаголовок (table col subhead)***
Текст (*table copy*)	текст в таблиці[a] (*table copy*)	текст в таблиці[b] (*table copy*)	текст в таблиці (*table copy*)

[a] Приклад посилання у Таблиці. *(Table footnote)*

[b] Приклад посилання у Таблиці. *(Table footnote)*

Використовуйте скорочення “Рис. 1”, в тому числі і на початку речення. Підписи до рисунків розташовувати під рисунками. Для підписів рисунків використовуйте стиль (*figure caption*).

Для вставки рисунків (растровий файл формату BMP, GIF, PNG, TIFF, JPEG, JPG або EPS з роздільною здатністю 300 dpi зі всіма включеними шрифтами) рекомендовано використовувати графічний текстовий бокс, як на цьому прикладі. Щоб зробити рамку невидимою, у меню MS Word “Формат\Напис” виберіть опції “Немає заливки” та “Немає ліній”.

Рис. 3. Приклад підпису до рисунку. (*figure caption*)

З метою уникнення плутанини використовуйте слова замість символів або скорочень при підписуванні координатних осей на рисунку. Наприклад, вказуйте величину “Magnetization” або “Magnetization, M.” але не “M.” Рекомендовано у підписах включити одиниці вимірювання, подаючи їх (у дужках). Не підписувати осі координат тільки одиницями вимірювання. Наприклад, використовуйте “Magnetization (A/m)”, але не “A/m.” Не підписуйте координатні осі, використовуючи відношення величин і одиниць вимірювання. Наприклад, використовуйте “Temperature (°K),” але не “Temperature / °K.”

ПОДЯКА (HEADING 5)

Подяка спонсору наводиться у ненумерованому посиланні на першій сторінці.

ЛІТЕРАТУРА REFERENCES (*HEADING 5*)

Шаблон передбачає послідовне нумерування цитованих джерел у квадратних дужках [1]. Розділові знаки у реченні слідують за квадратними дужками [2]. Посилатись на джерело необхідно як [3], а не як запис виду “Ref. [3]” або “reference [3]”, за винятком початку речення: “У джерелі [4] перше ...”

Нумеруйте посилання окремо, з використанням надрядкових індексів. Помістіть конкретне посилання внизу тієї колонки, у якій воно згадується. Не розміщуйте посилання у списку літератури. Використовуйте букви для посилання у таблиці, як у вище наведеному прикладі табл. 1.

Статті, які не були опубліковані, навіть якщо вони були відправлені у журнал, повинні згадуватись як “unpublished” («неопубліковані») [5]. Статті, які були прийняті до публікування, повинні згадуватись як “in press” («подано до друку») [2, 6]. Якщо у публікації менше шести співавторів, не використовуйте скорочення “et al” («та ін.»), а наводьте повний список [7]. З великої літери пишіть лише перше слово у назві статті, за винятком власних назв, імен та символів елементів.

Для статей, які опубліковано у перекладених виданнях, прохання спочатку навести джерело цитування англійською, а потім мовою оригіналу [2, 8]. Форматування списку літератури здійснити згідно вимог IEEE [9]. Застережено автоцитування.

Стрічку «Література» обов’язково здублювати англійською «References».

ЛІТЕРАТУРА REFERENCES (HEADING 5)

[1] IEEE. (2014). “Manuscript tmplates for conference proceedings,” [Online]. Available: www.ieee.org/publications_standards/publications/conferences/2014_04_msw_a4_format.doc (*references*)

[2] L. Petryshyn, A. Izmailov, “ISPC ITCM article template,” – “Шаблон статті МНПК ІТКМ,” *Iformation Technologies and Computer Modelling, Інформаційні технології та компютерне моделювання:* International Scientific and Practical Conference ITCM-2017: Precarpathian National University, Ivano-Frankivsk, 2017, in press.

[3] J. K. Author, “Title of chapter in the book,” in *Title of His Published Book, x*th ed. City of Publisher, Country if not USA: Abbrev. of Publisher, year, ch. *x*, sec. *x*, pp. *xxx–xxx.*.

[4] J. K. Author, “Name of paper,” *Abbrev. Title of Periodical*, vol. *x,* no. *x,* pp. *xxx-xxx,* Abbrev. Month, year.

[5] J. K. Author, “Title of paper,” unpublished.

[6] R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.

[7] A. Author, B. Author, C. Author, D. Author and E. Author, “Title of conference technical articles,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740-741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[8] L.B. Petryshyn, M.L. Petryshyn, “Synthesis of discrete recursive functions and properties of recursive data coding systems,” — “Синтез системы дискретных рекурсивных функций и свойства систем рекурсивного кодирования данных.” *Information Processing Systems, Системи обробки інформації*, vol. 3 no. 2, pp. 46–51. 2013.

[9] D. Graffox. (2009). IEEE citation reference [Online]. Available: http://www.ieee.org/documents/ieeecitationref.pdf

З М І С Т

СЕКЦІЯ 1. ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ В ТЕХНІЧНИХ ТА СИСТЕМАХ СПЕЦІАЛЬНОГО ПРИЗНАЧЕННЯ..... 3

SECTION 1. INFORMATION TECHNOLOGIES IN TECHNICAL SYSTEMS AND SYSTEMS OF SPECIAL PURPOSE..... 3

Спосіб Формування Множини Оціночних Сигналів Для Дослідження Функціонального Простору Об'єкта 4
Юрій Стрілецький, Віктор Ровінський

Технологічні Можливості Інформаційного Серверу Hyperwave 8
Бойко Наталія Іванівна, Копач Ольга

Радіовимірювальний Перетворювач Концентрації Газу на Транзисторній Структурі з Від'ємним Опором 12
О.В. Осадчук, В.С. Осадчук, Я.О. Осадчук

Застосування Частотного Критерію Інформативності Ознак в Задачах Інтелектуального Аналізу Тексту Багаторівневого Інформаційного Моніторингу 16
М.С. Голуб

Структурно-Параметрична Ідентифікація Правил Бази Модельних Знань 20
Ірина Жирякова, Сергій Голуб

Комп'ютерна Система Стереозору для Візуального Визначення Відстаней 24
Олег Прокопчук, Сергій Вовк, Михайло Церр

Моделювання Структури Програмного Комплексу для Управління Газотранспортними Системами.....28
Ольга Химко, Василь Чекурін

Принципи Побудови Систем Інформаційної Підтримки Прийняття Рішень на Авіаційних Підприємствах.....32
Володимир Темніков, Андрій Темніков

Cancellation Patterns and Term Rewriting 36
Bartosz Zieliński, Paweł Maślanka

Fast Gender Classification Using Geometrical Features and SVM 40
Piotr Milczarski, Zofia Stawska

Стохастичний Аналог Задачі Лінійного Програмування з Ефектом Галуження.....44
Фахріддін Мірзоахмедов

СЕКЦІЯ 2.ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ В СОЦІУМІ, ОСВІТІ, МЕДИЦИНІ, ЕКОНОМІЦІ, УПРАВЛІННІ, ЕКОЛОГІЇ ТА ЮРИСПРУДЕНЦІЇ 48

SECTION 2.INFORMATION TECHNOLOGIES IN SOCIETY, EDUCATION, MEDICINE, ECONOMICS, MANAGEMENT, ECOLOGY AND JURISPRUDENCE 48

Пошук Облич з Допомогою Поділу Зображення та Параметричних Статистичних Ознак 49
Р. Мельник, Ю. Каличак, С. Дахній

Моделювання Плану Ремонту Доріг в Межах Територіальної Громади на Основі Модифікації Алгоритму Пріма .53
В. Литвин, З. Рибчак, Д. Угрин

Оптична Система для Фототерапії з Контролем Поглинутої Енергії 57
Олексій Яненко, Костянтин Шевченко, Роман Ткачук, Василь Кузь

Моделювання Елементів Мікросистем-на-Кристалі зі Структурою КНІ для Неінвазивних Глюкометрів.....61
Ігор Когут, Віктор Голота, Віктор Довгий, Михайло Котик

Побудова Системи Моніторингу та Керування Станом Навколишнього Середовища на Основі Технології IoT 65
Н. Бондаренко

Формалізація Етапів Діагностично-Лікувальних Заходів при Проектуванні Систем Підтримки Прийняття Рішень в Медицині 69
Анатолій Поворознюк, Ганна Філатова

The Approach to Drawing up a List of Professional Competences in the Field of Software Engineering for Harmonization of Educational Standards 73
Igor Turkin, Vitaliy Narozhnyy, Yuliya Vykhodets
Методи Розпізнавання Облич: Короткий Огляд 78
І. Голуб'як
Формування Масиву Вхідних Даних для Мамографічних Досліджень 82
С. Палаш, О. Харченко
Використання Фрактальних Характеристик Щільності Хроматину Ядра Клітини для Визначення Стану Інтерфази. 86
О. Шумейко, С. Герасименко, Г. Шевченко
Моделювальний Комплекс для Симуляції Дефіциту Енергії в Клітинах Людини 91
Т. Аксьонова
Застосування Дискретної Оптимізації для Автоматизації Вибору Курсів Студентами 95
Г. Добровольський, О. Тодоріко, Є. Кушнір
Геоінформаційна Підсистема Моніторингу Концентрації Радіоактивних і Небезпечних Хімічних Речовин Внаслідок Надзвичайної Ситуації 99
Петро Гучек, Олена Литвиненко, Ігор Астіоненко
Застосування Зростаючої Пірамідальної Мережі для Вирішення Задачі Ідентифікації Людини по її ЕКГ 104
Віталій Вишневський, Віталій Величко, Лілія Кізуб
Using Gamification in Mobile Activation Learning System 108
V. Emets, J. Rogowski, J. Krasiukianis
The Comparison of the Web Application Development Frameworks 112
J. Rogowski
Застосування Багатомодельного Класифікатора в Інформаційних Системах Кризового Моніторингу 116
А. Авраменко, С. Голуб
Технологія Проектування Електронних Дидактичних Ресурсів 120
Олеся Власій, Ольга Дудка, Орест Гейко
Інформаційна Технологія Моніторингу Якості Загальноосвітнього Навчального Закладу 124
Володимир Лужецький, Олег Білик, Богдан Білусяк
Старовікові Насадження Природного Заповідника "Ґорґани", їх Просторова Структура та Ідентифікація Супутниковими Знімками 128
Анна Цуняк, Олег Часковський, Микола Король, Ольга Токар
Про Квантове Моделювання Інформації в Фінансовій Математиці 132
Анатолій Прикарпатський
Гейміфікація Навчання 136
Віталій Горєлов, Даріуш Саля
Інформаційні Технології в Діяльності Керівника Навчального Закладу 140
Юрій Іляш, Олеся Власій, Богдан Білусяк

СЕКЦІЯ 3.ТЕОРІЯ ІНФОРМАЦІЇ, КОДУВАННЯ ТА ПЕРЕТВОРЕННЯ ФОРМИ ІНФОРМАЦІЇ 144

SECTION 3.INFORMATION THEORY, CODING AND TRANSFORMATION OF INFORMATION FORM 144

Апаратно-Мінімальні Паралельні Коди CRC 145
Василь Семеренко
On some Class of Quantum Finite Automata 149
Volodymyr G. Skobelev
Багатоканальні Системи Збору Даних 153
Марія Семаньків, Даріуш Саля
Оптимізація Пошукового Алгоритму Боєра-Мура-Хорспула для Бінарних Даних 157
Ігор Завадський, Василь Остапчук

New Probabilistic Model of Stochastic Context-Free Grammars in Greibach Normal Form 161
Yevhen Hrubiian

СЕКЦІЯ 4.ЦИФРОВА ОБРОБКА ТА УЩІЛЬНЕННЯ ІНФОРМАЦІЇ 165

SECTION 4.DIGITAL INFORMATION PROCESSING AND COMPRESSION 165

Arbitrary Meshes to Simulate Natural Video Objects 166
Olena Osharovska

Multivariate Outlier Detection Technique Based on Normalizing Transformations for Non-Gaussian Data 170
Sergiy Prykhodko, Natalia Prykhodko, Lidiia Makarova, Kateryna Pugachenko

Апроксимація Параметричними Сплайнами для Ущільнення Інформації Електрокардіограми у Трьох Ортогональних Відведеннях 175
В.В. Вишневський, Т.М. Романенко

Алгоритми Побудови Ортогональних Базисів на Основі Рекурсивних Функцій Ґалуа 179
Наталія Превисокова, Любомир Петришин

The Dynamic Information in the Images 183
Vitaliy Boyun

Застосування Трійкових Симетричних Функцій для Цифрової Обробки Інформації на Основі Ортогональних Перетворень 187
Артем Ізмайлов

СЕКЦІЯ 5.ЗАХИСТ ІНФОРМАЦІЇ В ІНФОРМАЦІЙНО-ТЕЛЕКОМУНІКАЦІЙНИХ СИСТЕМАХ 191

SECTION 5.INFORMATION SECURITY IN INFORMATION AND TELECOMMUNICATION SYSTEMS 191

Математичні Моделі Модифікованої Несиметричної Крипто-Кодової Системи Мак-Еліса на Модифікованих Еліптичних Кодах 192
Євсеєв Сергій

Development of Data Exchange Protocol Prototype Between Intrusion Detection and Prevention Systems 197
Oleksandra Yeremenko, Oleksandr Lemeshko, Anatoliy Persikov

Огляд Видів Аналізу Соціальних Мереж для Забезпечення Інформаційної Безпеки 201
Тамара Радівілова, Людмила Кіріченко, Даніїл Рудченко

Вдосконалення Методу Моніторингу Цифрових Слідів Інцидентів Інформаційної Безпеки у Сегментах Національної Інформаційної Інфраструктури 207
Надія Казакова, Олексій Фразе-Фразенко

The Authentication Method in Wireless Sensor Network Based on Neighbor Trust Model 212
Olesia Voitovych, Olga Shulyatitska

Методи Байт-Орієнтованого Хешування Даних Низькоресурсної Криптографії 216
Володимир Лужецький, Світлана Слободян, Дмитро Кисюк, Олександр Ювковецький

Концепція Сучасної Архітектури Менеджменту Процесів Інформаційної Безпеки Soapa 220
Копиченко Іван

Оцінка Ризиків Інформаційної Безпеки в Мережах Wi-Fi на Основі Апарату Нечіткої Логіки 224
Леонід Куперштейн, Татарчук Артем, Олеся Войтович, Ратников Назар

Access Control Method Based on Workstation Binding 228
Yurii Baryshev, Kristina Neuimina

Перевірка Модельних Припущень у Криптоаналізі Arx-Шифрів 232
Олександр Деркач

Асимтотичні Розподіли Імовірностей Змішаних Диференціалів Випадкових S-Блоків 236
Всеволод Бахтігозін, Сергій Яковлєв

Реалізація Методу Виділення Зон Управління в Захищеній Комп'ютерній Мережі 240
Любомир Пархуць, Марина Костяк

Оцінки Імовірностей Диференціалів Безключової R-Схеми Блокового Шифрування 244
Євсюкова Яна, Яковлєв Сергій
Ланцюги Підпросторів Калина-Подібних Шифрів 248
Коляда Марія, Яковлєв Сергій
Формування Сеансовиих Ключів на Базі Концепції Прийняття Рішень 252
Глущенко В.Є., Петришин М.Л.
Семантичний Аналіз Великих Даних в Задачах Кібербезпеки 256
Вероніка Островська, Олеся Войтович

СЕКЦІЯ 6.АРХІТЕКТОНІКА ТА КОМПОНЕНТИ КОМП'ЮТЕРНИХ СИСТЕМ ТА МЕРЕЖ 260

SECTION 6.COMPUTER NETWORKS, SYSTEMS AND COMPONENTS ARCHITECTONICS 260

Комп'ютерна Система Визначення Реологічних Характеристик Суднових Технічних Рідин 261
Віталій Нікольський, Кирило Бережний, Марк Нікольський
Система контролю завантаження контейнеровозів "ShipLoad" 265
Юрій Накул, Владислав Стовманенко, Марк Нікольський
Скорочення необхідного ресурсу ПЛІС для обчислювальних систем на кристалі 269
Валерій Жабін, Валентина Жабіна
Глибокі Нейронні Мережі для Вирішення Завдань Розпізнавання і Класифікації Зображення 274
Віктор Синєглазов, Олена Чумаченко
Deep Learning Classifier Based on NEFCLASS and NEFPROX Neural Networks 278
Olena Chumachenko
Ітераційний метод самокалібрування передатних характеристик вимірювальних каналів біомедичної аналогово-цифрової системи 282
Леонід Крупельницький, Дмитро Куций
Апаратна реалізація алгоритму сортування на основі просторово-часових графів 287
В.М. Грига, І.Т. Когут, В.І. Голота, Л.В. Николайчук

СЕКЦІЯ 7.МАТЕМАТИЧНЕ ТА КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ СКЛАДНИХ СИСТЕМ 291

SECTION 7.MATHEMATICAL AND COMPUTER MODELLING OF COMPLEX SYSTEMS 291

Застосування Алгоритму Імітації Відпалу для Розподілення Даних при Паралельному Моделюванні Неізотермічного Напружено-Деформованого Стану з Урахуванням Фільтрації 292
В.О. Богаєнко
Про Природу Вхідних Даних в Деяких Задачах Штучного Інтелекту 297
Н.Тимофієва
Застосування Віртуалізованої ІТ-Інфраструктури для Збільшення Рівня Екологічної Безпеки Технологічних Систем 301
Вавулін Петро
Комп'ютерне Моделювання Дифузійних Потоків Водню у Шаруватому Композиті Мідь-Залізо за Арксинус-Розподілу Включень 307
А.Є. Чучвара
Parametric Analysis of Input Circles of Measuring Transducers Impedance of Potentiometric Type 311
G. Barylo, R. Holyaka, V. Virt, F. Vezyr
Математичне Моделювання Процесу Графітування Вуглецевих Виробів 315
Олексій Жученко, Марія Волощук
Метод Змінних Радіусів в Задачі Розміщення Нерівних Кіл 319
С.В. Яковлев, О.В. Карташов, К.П. Коробчинський
Двовимірна Математична Модель Реакції Окислення Чадного Газу на Поверхні Платинового Каталізатора 323
Петро Костробій, Ірина Рижа

The Movable Cellular Automata as a Tool to Simulation Biological Process and Systems 327
Kateryna Gazdiuk, Volodymyr Zhikharevich, Sergiy Ostapov
Поверхнева Енергія Деформації і Поляризації у Межах Градієнтної Теорії
Пружних Діелектриків з Квадрупольними Електричними Моментами 331
Ольга Грицина, Василь Кондрат
Математичне Моделювання Дифузії у Середовищі з Пастками за Каскадного Розпаду Домішок 336
Ю.І Білущак
Математичне та Комп'ютерне Моделювання Процесів Конвективної Дифузії у Двошарових
Засипних Фільтрах Води 342
О. Ю. Чернуха, Ю. І Білущак
Фізико-Математична Модель Зв'язаних Теплових, Механічних і Дифузійних Процесів
у Багатокомпонентних Тілах 349
В.Є. Гончарук
Математичне Моделювання та Аналіз Міцності Деревини 355
Софія Поберейко, Ярослав Соколовський
Problems of Complexity in Cybernetics and Creation of Universal System of Knowledge 359
Petro Trokhimchuck
Особливості Польоту Квадрокоптера на Малих Висотах 363
Б. Я. Благітко, Ю. С. Мочульський, І. М. Заячук
Програмний Комплекс для Дослідження і Порівняння Завадостійких Кодів 367
Юрій Гунченко, Тетяна Щербакова, Павел Ємельянов, Вікторія Левчук
Розрахунок Сил Тертя у Викривлених Нафтовидобувних Свердловинах з Врахуванням Положення
З'єднувальних Муфт 373
Ольга Євчук, Віктор Ровінський
Моделі та Методи Оптимізації Функціонування Розподільчих Систем Енергетичного Комплексу 377
Олена Кірік
Вплив Локального Зміщення Маси на Приповерхневі Явища в Кулі 382
Ольга Грицина, Галина Мороз
Автоматична Побудова Сематичної Мережі Тексту у Системах Запит-Відповідь 386
О. С. Волковський, Є. Р. Ковилін
Математичне Моделювання Кінетики Квазістатичного Фільтраційного Насичення
Пористого Шару Природнім Газом 390
Василь Чекурін, Зоя Притула
Моделі Штучного Життя в Клітинному Просторі Ресурсів 394
В. Завертаний, О. Макаренко
Методика Вибору Оптимальної Геометричної Форми Порожнини в Зубі для Його Лікування Ендокоронкою 399
В.Ф. Макєєв, Р.Р. Павличко, І.М. Заячук
Модифікований Метод Граничних Елементів та Алгоритм Розв'язання Задач Індукованого
Магнітного Поля Тонкостінних Конструкцій в Тривимірному Просторі 403
Володимир Рябенький, Юлія Таргунакова, Ігор Чудайкін
Інформаційно-Аналітична Система для Оптимального Планування Режимів Роботи Підземних Сховищ Газу 410
Назар Притула, Мирослав Притула, Олег Гринів
Оптимальне Планування Режимів Роботи Складних Газотранспортних Систем 414
Назар Притула, Мирослав Притула, Вадим Фролов
Методи Розв'язування Крайових Задач з Використанням Дробових Похідних за Часом 418
Ярослав П'янило, Галина П'янило, Олег Браташ, Валентина Собко
Модель Неусталеного Руху Крові в Судинах з Врахуванням Пружності 422
Ярослав П'янило, Володимир Череватий

Система Оперативного Контролю Органіки у Природній і Питній Водах 426
Олександр Малько, Анастасія Малько

СЕКЦІЯ 8.ПРИКЛАДНІ МЕТОДИ ДОСЛІДЖЕННЯ ДИСКРЕТНО-НЕПЕРЕРВНИХ МАТЕМАТИЧНИХ МОДЕЛЕЙ..... 430

SECTION 8.APPLIED METHODS OF STUDYING DISCRETE-CONTINUOUS MATHEMATICAL MODELS..... 430

Загальні Крайові Задачі для Гіперболічного Рівняння із Сумовними Коефіцієнтами та Правими Частинами 431
Р.М. Тацій, О.О. Карабин, О.Ю. Чмир

Прямий Метод Дослідження Температурного Поля в Багатошаровому Трубопроводі за Умов Пожежі 436
Роман Тацій, Марта Стасюк, Олеся Власій, Олег Пазен

Розв'язання Третьої Крайової Задачі для Рівняння Лапласа в Прямокутнику 441
Р.М. Тацій, О.М. Трусевич, М.Ф. Стасюк

Оптимізація на сферично розташованих комбінаторних множинах 445
О.С. Пічугіна

Шаблон Статті МНПК ІТКМ (стиль: *PAPER TITLE*).....452
Любомир Петришин, Артем Ізмайлов

З М І С Т.....456

Наукове видання

"ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ТА КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ"

матеріали
Міжнародної науково-практичної конференції
15-20 травня 2017 року
Івано-Франківськ

"INFORMATION TECHNOLOGIES AND COMPUTER MODELLING"

proceedings
of the International Scientific Conference
2017, May, 15th to 20th
Ivano-Frankivsk

Науковий редактор: Л.Б. Петришин
Технічний редактор: В.З. Максимець
Коректор: Ю.Ю. Іляш
Комп'ютерна верстка: О.М. Голіней

Прикарпатський національний університет
імені Василя Стефаника
76018, Івано-Франківськ, вул. Шевченка, 57,
ПНУ, адміністративний корпус

Формат. Папір офсетний. Ум. Друк арк.
Тираж 100 прим.

Друк: підприємець Голіней О.М.
тел. (0342) 58 04 32

Свідоцтво внесення до державного реєстру
ДК №

Для нотаток

Для нотаток