

# Інтелектуальний пошук відомостей технічними системами інформаційного моніторингу

М.С. Голуб

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

## Intelligent search for details in technical systems information monitoring

M. Holub

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

**Анотація** — Описана інформаційна технологія інтелектуального пошуку відомостей технічними системами. Для формування вирішуючих правил запропоновано використати методи багаторівневого моделювання. Експериментально підтверджено можливість та доцільність застосування описаної технології для автоматизації процесу пошуку текстів заданого змісту.

**Abstract** — Described 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.

**Ключові слова** — Пошук тексту за змістом, багаторівневе моделювання, технічна система, інформаційний моніторинг.

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

### I. ВСТУП

Інтелектуальні властивості технічного пристрою визначається його здатністю здобувати нову інформацію на основі вже відомої. Системи інформаційного моніторингу, на відміну від інших, Моніторингові інформаційні системи призначені для забезпечення відомостями про властивості об'єктів процеси прийняття рішень шляхом обробки та перетворення результатів

неперервних спостережень, поданих у вигляді текстових повідомлень, аудіо – та відео-файлів .

Необхідність контекстного пошуку інформації технічними системами викликана потребою підвищити ефективність пошуку відомостей, які потрібні саме для синтезу моделі, що призначена для профілювання текстів із заданими характеристиками. Тобто потрібно знаходити тексти заданого змісту, а не тільки за ключовими словами.

Поставлене завдання вирішується шляхом використання методів багаторівневого моніторингу [1]. Забезпечується ієрархічне поєднання багатопараметричних моделей, що навчені еталонними текстами заданого змісту [2].

Автором розвивається напрямок багаторівневого моніторингу в предметній області інтелектуального аналізу текстових повідомлень.

### II. ПОСТАНОВКА ЗАДАЧІ ДОСЛІДЖЕНЬ

В цій роботі розв'язується задача класифікації тестових повідомлень за заданим змістом. Наперед відомою інформацією є перелік класів та зміст повідомлень, що належать до цих класів. Завданням є пошук тих ділянок текстів серед множини текстових повідомлень, зміст яких відповідає наперед заданому змістові одного із класів.

### III. ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ КОНТЕКСТНОГО ПОШУКУ ВІДОМОСТЕЙ

Поставлена задача розв'язувалась шляхом побудови множини вирішуючи правил, що дозволяли відносити виділені ділянки тексту до наперед заданих змістовних класів. Правила синтезувались у вигляді багаторівневих функціональних залежностей, що поєднують в своїй структурі індуктивні моделі, нейромережі, генетичні алгоритми, гібридні моделі, отримані за завершеними алгоритмами. Для розв'язання задачі узгодження взаємодії цих моделей використовується метод висхідного синтезу елементів [1]. Ці методи, засоби та алгоритми синтезу моделей (АСМ) поєднані в інформаційну технологію інтелектуального пошуку відомостей технічними системами інформаційного моніторингу (ТСІМ). Вона містить такі етапи.

1. Підготовчий. Формуються моделі змісту повідомлення. Проводиться аналіз тексту, перетворення текстового повідомлення в масив його чисельних характеристик, синтез та випробування моделей.

Текст аналізується шляхом виділення окремих ділянок (вікон) із послідовно розміщених в тексті речень. Перетворення текстового повідомлення проводиться шляхом розрахунку частотних характеристик наперед заданих ознак тексту. В результаті розрахунків отримуємо вектор ознак — точку спостереження в багатовимірному просторі ознак первинного опису об'єкта моделювання. Після оцінки інформативності деякі ознаки, що мають незадовільні частотні характеристики, текстів видаляються із первинного опису. Точки спостереження, що сформовані за вектором інформативних ознак утворюють масив вхідних даних (МВД).

Належність ділянки тексту до заданого змістовного класу відображалась шляхом позначення точки спостереження як «Свій». Таким чином утворюється послідовність точок спостереження, які відображають в своїй структурі зміст повідомлення, що належить до заданого класу. Іншим точкам спостереження надавався статус «Чужий». Вони відображають зміст повідомлення, які належать до інших класів.

Синтез моделей забезпечує розв'язок задачі групування точок спостереження. Кожна модель використовується як вирішуючи правило, що дозволяє віднести кожному точці до одного із заданих класів. Для формування цього правила забезпечується ієрархічне поєднання багатопараметричних моделей [1] відповідно до індуктивного методу [3]

За результатами випробування моделі визначалося порогове значення результатів моделювання, вище якого модельоване спостереження позначалося як «Свій». Після цього модель заноситься до бази модельних знань.

2. Оперативне використання. Аналізований текст подається на вхід ТСІМ. За результатами перетворення МВД паралельно кількома вирішуючими правилами бази модельних знань приймається рішення про відповідність дослідженого фрагменту тексту заданому змісту.

При випробуванні описаної технології визначалась кількість вірно розпізнаних фрагментів тексту відповідно до наперед заданого змісту та залежність результатів контекстного пошуку від виду опорної моделі. В таблиці 1. подані види опорних моделей, що приймали участь в процесі синтезу вирішуючого правила.

ТАБЛИЦЯ 1. ВИДИ ОПОРНИХ МОДЕЛЕЙ

| Модель                                       | Вид моделі |
|--|------------|
| $x_1 + x_2$                                  | 1          |
| $x_1^2 + x_2^2 + x_1x_2$                     | 2          |
| $x_1^2 x_2 + x_2^2 x_1 + x_1x_2 + x_1 + x_2$ | 3          |
| $x_1^2 + x_2^2 x_1 + x_1x_2 + x_1 + x_2$     | 4          |

На рис. 1 подані результати випробувань описаної технології.

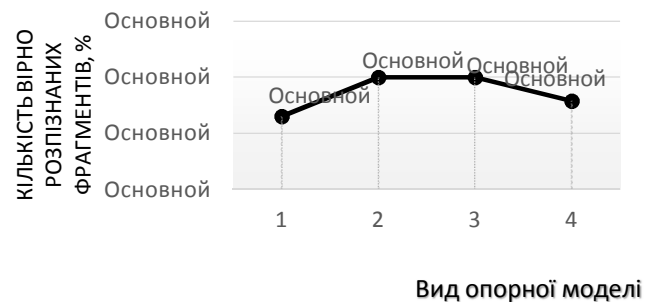


Рис. 1. Результати контекстного пошуку відомостей

### ВИСНОВКИ

Розв'язано задачу інтелектуального пошуку текстових повідомлень заданого змісту за технологією багаторівневого моніторингу. Ефективність інтелектуального пошуку залежить від інформативності заданих ознак тексту та вигляду опорної моделі. Після параметричної оптимізації процесу інтелектуального пошуку відомостей досягнуто прийнятної надійності розв'язку поставленої задачі.

### ЛІТЕРАТУРА REFERENCES

- [1] S. V. Holub Bagatorivneve modeluvanya v tehnologiah monitoringu otchujuchogo seredovischa / S.V. Holub. – Cherkasy: Vidavnichiy viddil Cherkaskogo natsionalnogo universitetu imeni Bohdana Hmel'nitskogo, 2007. – 220 s.
- [2] S. V. Holub Vidobragenia vlastivostey avtora tekstu v strukturі bagatoparametrichnoi modeli / S.V. Holub, O.V. Konstantinovska, M.S. Holub // Sistemi obrobky informatsii: Zbirnik naukovih prats. – Harkiv: Harkivskii universitet povitrianih sil imeni Ivana Kozeđuba, 2014. – Vipusk. 9 (125). – S. 82-87
- [3] A. G. Ivahnenko Induktivniy metod samoorganizatsii modeley sloĝnih sistem / A.G. Ivahnenko. – Kiiv. : Naukova dumka, 1981. – 296 s.