Розробка комунікаційної стратегії із застосуванням аналізу даних соціальних медіа

О.М. Степанова
кафедра менеджменту і адміністрування
Східноукраїнський національний університет
імені Володимира Даля
Северодонецьк, Україна
stepelen@gmail.com

А.А. Волков
Інститут комунікацій
Вестфальський університет
імені Вільгельма
Мюнстер, Німеччина
a_volk07@uni-muenster.de

Communication strategy formulation based on the analysis of social media data

E. Stepanova
Department of Management and Administration
Volodymyr Dahl East Ukrainian
National University (SNU)
Severodoneck, Ukraine
stepelen@gmail.com

A. Volkov
The Department of Communication (IfK)
Westphalian Wilhelm University of Münster
Muenster, Germany
a_volk07@uni-muenster.de

***Анотація*— У статті відзначено, що Інтернет став важливою частиною сучасної комунікації і має бути інтегрований у процес розробки стратегії комунікації. Автори торкаються питань розширення інформаційної бази з метою отримання якісного аналізу середовища, цільових аудиторій та конкурентів, який надалі може бути використаний у процесі планування комунікації. Соціальні медіа, та зокрема, соціальна мережа Фейсбук, яка надає доступ до даних своїх користувачів за допомогою додатків програмного інтерфейсу (API), є одними з найцікавіших джерел інформації. Автори надають огляд низки методів збору даних у соціальних медіа та певних інструментів збору та обробки Інтернет-даних. У статті зауважено, що R (мові програмування) притаманні певні риси, завдяки яким вона стає особливо привабливим засобом у цій галузі. Автори наводять приклад використання R для типового аналізу Facebook -сторінок конкурентів.**

***Abstract*— This article emphasizes the fact that the Internet has become an important part of today`s communication and, as such, should be integrated in the process of communication strategy development. It addresses issues of expanding the information base in order to get a more thorough analysis of the environment, target audiences and competitors which could be later used in the process of communication planning. Social media and particularly the social network Facebook which provides access to the user data through the application programming interface (API) are among the most promising sources of information. The Authors overview different methods of social media mining and special tools which could be used for retrieving and processing Internet-data. They argue that the R (programming language) has a number of features which make it particularly attractive for those purposes. The Authors give then an example of R being used for a typical first-step analysis of competitors` Facebook-pages.**

Ключові слова—комунікаційна стратегія; Інтернет; соціальні мережі; аналіз даних; API methods; R.

Keywords— communication strategy; Internet; social networks; data analysis; методи API; R.

I. INTRODUCTION

The rapid development of information and communication technologies, which resulted in the formation of the worldwide web, plays an increasingly important role in the modern society, enriching the existing business techniques and methods, forming new channels of communication between a customer and a company and opening new ways for public relations.

However, despite the popularity of the Internet and its role as the latest communications` environment, most of Ukrainian companies have lesser understanding and thus effectiveness in it if compared to those of the developed western countries.

In 2013 Ukrainian office of Google together with "the GfK Ukraine", presented the results of the study "Impact of the Internet on the economy of Ukraine" [1]. This study confirms that Internet tools are being used uncoordinated and unsystematic. The choice of instruments is rather being affected by modern trends, than by a careful analysis of needs and capabilities of the company. Furthermore, only a handful of companies realize the necessity and possible benefits of the complex communication strategy and its constant evaluation.

In many ways, poorly developed methodological framework leads to the fact that the majority of companies using online communication tools, has no clear strategy for their online presence and do not rationally adopt their budgets for online communications, marketing and advertising. This prevents exploitation of the Internet`s full potential from establishing an effective interaction with an audience, reducing it to some rudimentary attribute [2].

Further improvement of methodical and practical aspects of the strategy`s formation in the field Internet-communication is, therefore, a highly relevant field of studies.

Given the dynamics of the environment, continuous renewal, differentiation and increase in volume of the required data information aspect gains a particular importance and the improvement of methodological and applied aspects of the strategic directions development in the field of communications has a high relevance.

The purpose of this paper is to offer tools for the development of the information base needed for the formation of an effective communication strategy with the use of the Internet.

II. Strategic Approach To Communication

The communication strategy is a result of the selection and formation of the organizations communication direction; a concept, a program, a general course of the managing subject focused on identifying and achieving main goals of the organization in the field of strategic management; communication policy that ensures achievement of the organizations long-term goals.

The communication strategy should be developed according to the actual needs of the market, taking into account the target audiences preferences and other factors. The main objective pursued in developing of a communication strategy is a management of the customer-brand relationship, formation of an effective and predictable consumer-brand communication, improvement of the brands image according to the customers' needs, current trends and market conditions.

In the process of developing a communication strategy necessary tasks are analyzed, communicative message and purpose are generated and basic elements, which could be most effective in mediating a message to an audience, are defined.

A strategic approach to communication is widely represented in the foreign studies [3-5]. These works present significantly different numbers of planning stages, their sequence and essence, but each of them, one way or another, highlights the necessity of researching target audiences, competitors and a current situation. Addressing these issues involves the use of a wide range of information sources and an intensive analytical work.

Traditionally, the required data could be obtained from a variety sources: published reports, speeches of management representatives, business press, employees busy in a distribution system, customers and suppliers common to the company and its competitors, company`s technical personal, information gathered from competitors` former managers and other employees, and so forth [5].

The rapid growth of the social networks users numbers on the Internet creates additional opportunities to expand the required information base, which will enable a whole new level of possibilities to ensure the formation of a company`s communication strategies. Given these possibilities, let us consider the peculiarities of researching target audiences, competitors and current situation.

The first of the mentioned researches should be directed at the identification and the segmentation of a company`s target audiences on the Internet, identification and classification of communication platforms with each target group, criteria development for the evaluation of the each type of platforms, determination of optimal communication formats, identification of opinion leaders, analysis of the interaction process with customers and partners via social media.

All tasks dealing with the integrated analysis of competitors` presence in social media should be reduced to the analysis of their marketing activities in social networks, methods and platforms they use, their announced plans and the overall analysis of their communicative activities. This is necessary for the development of original ideas and ways of promotion that haven`t been used by competitors yet. Quantity and nature of critical points, meaning those mostly significant elements of the environment, vary from company to company. Analysis of competitor's communicative activity is limited by the following time frames. In a short-term period, it is usually possible to focus only on those elements that have a critical impact on current operations of a company. On the other hand, in the long-term period, opportunity arises to investigate the general nature of competitors` communications.

One of the important tasks in the communication strategy development is the evaluation of a company's existing resources and therefore its capacity to maintain some degree of Internet presence. This capacity should not be overrated as the higher variety of instruments a company chooses to use the higher amount of financial and human resources as well as time it should need. Therefore, such an assessment should be based on a sufficient analysis of a previous communication experience, state of the current information field and analysis of an audiences` image of a company and its products.

The sources for the necessary research in the Internet environment are: online media, industry and business platforms, thematic portals, search engines, comment sections on websites, blogosphere, social networks and forums.

III. Social Media Mining

The data analysis of social networks is rapidly gaining popularity around the world. Various institutions and companies around the globe such as Carnegie Mellon University, Stanford, Oxford, INRIA, as well as Facebook, Google, Yahoo!, LinkedIn and many others are actively collecting and studying this data [6-9].

Social Networking Services (or social networks) are online services designed to organize and maintain network communications. In addition to personal information contained in users' profiles (age, location, interests, education, hobbies, and so on.) it is possible to retrieve the data about users` visits to the pages of other users, their participation in the formation of networks and groups with the same interests, and so on. Furthermore, social networks, forums, news and entertainment portals and blogs contain many valuable materials which can be used to obtain information about preferences and characteristics of different people and companies.

Experts from research centers and companies around the world use social media data for the modeling of social, economic, political and other processes from the individual and up to the governmental level in order to develop mechanisms which could help to influence these processes.

However, processing of social networks data also requires the development of an appropriate infrastructure and algorithmic solutions that would take into account their dimension. Collecting and presenting data from social networks in a convenient form is still a non-trivial task.

It is of course possible to collect and process that data by simply using search engines and manually adding it to data tables. Unfortunately, that would be the least efficient method of all and by the time a sufficient amount of data could be gathered it might already become out-of-date, thus losing its relevance.

Another more effective way to collect and process data is to use a variety of existing systems for monitoring social media, such as the Youscan, IQbuzz, Kribrum, Wobot and Brandspotter, which differ in their focuse on specific tasks. Studying functionality of such systems leads to the conclusion that none of the tools is able to give a full picture for completing specific tasks, mentioned above.

Apart from that, data presented by those programs is often not directly compatible with the formats required for analytical programs that are used in a further analysis. Therefore, a programming environments is needed to convert the raw data into some usable data frame. At present, the most popular languages suited for such tasks are Python, Perl, and Java. In addition, there are some programs specifically designed for this purpose which are already available to researchers.

Companies that own online social networking services (e.g. Facebook, Twitter) actively invest in the development of improved infrastructural (Cassandra, Presto, FlockDB, Thrift) and algorithmic (new algorithms for searching and recommending users, products and services) solutions for handling large volumes of user data. Therefore, the third way to obtain information from a social network is to use its software interface or API (Application Programming Interface), provided by almost all of the popular social networks.

Based on the API methods, developers can create applications that are compatible in the context of different social networks for which the APIs differ in sets of available data, restrictions on the number of requests and the cost of access to the interfaces [10,11].

The advantages of this method are the possibility of obtaining information about one client in a structured format (JSON or XML) and the ease of integrating API protocols into your own application.

IV. Analyzing Facebook With R

The social network “Facebook” is among the most popular platforms for extensive studies. The Facebook API can be easily used to create applications which then become available for other users of the social network. Facebook API allows applications to utilize connections and individual information of a user to make it more involving, and to publish activities to the news feed and profile pages of Facebook, subject to individual users’ privacy settings. With the API, users can add social context to their applications by utilizing profile, friend, Page, group, photo, and event data [10,11]. Facebook API uses RESTful protocol and responses in JSON format.

In addition to that, API allows an adjustment of the data flow not only to the external applications but also in the opposite direction.

As a tool for framing and analytical processing of data obtained through API could be conveniently used R. It is a programming language for statistical data processing and its graphical representation and which has a rapidly growing community that now includes more than two million users.

Compared to other software, R allows experts to analyze data quicker and using more advanced techniques, which explains the rapid growth in popularity of R in the various sectors of the economy [12].

Along with simple methods of statistical analysis, such as frequency analysis, calculation of statistical characteristics, contingency tables, correlation and graphing, this module includes a t-tests and a large number of other non-parametric tests as well as more sophisticated techniques like regression analysis, discriminant analysis, factor analysis, cluster analysis, analysis of variance, analysis of suitability (reliability analysis) and multidimensional scaling analysis of multidimensional tables, spatial analysis, creation of cartograms and more.

Among the technical advantages of the system are:

- Distribution of the program under the GNU Public License;
- Source code and binary modules available through the extensive network of repositories CRAN (The Comprehensive R Archive Network);
- Installation packages for Microsoft Windows (works on both 32-bit and 64-bit versions), Mac OS X, UNIX-,

Linux-, FreeBSD-based operation systems and OS Solaris;

- Statistical procedures programming language R has virtually became a standard. For example, it is now fully supported in the new IBM SPSS Statistics Developer system;
- Ability to exchange data with spreadsheets;
- Ability to save the entire history of computing for documentation purposes;
- Excellent imaging capabilities.

A typical problem is that most companies choose to perform only a partial analysis which allows to get just an intuitive understanding of the current competitors' communication strategies along with their advantages and weaknesses. Therefore, to illustrate the instruments proposed above, a typical task faced by companies in the process of developing a communication strategy should be presented: comparison of competitors` Facebook-page popularity growth.

V. Example

In the early 90s Skateboarding and BMX industries got into marketing competition with comparatively younger and weaker aggressive inline manufacturers. That, together with the identity crisis inside aggressive inline community itself marginalized the sport, effectively denying them previous public attention, together with the access to the best known worldwide extreme sports events[16].

Consequently, Skateboarding and BMX became a mainstream, thus losing their unique identities in the past two decades. This potentially opens an opportunity for the aggressive inline industry to communicate its new image to the focus groups, winning back its reputation. The development of the internet and adaptation of social networks to the needs of corporate communication could help in exploring this opportunity.

In the following example, the media activity on the official Facebook profiles of the companies involved in the aggressive inline industry will be to be analyzed and compared to those of the BMX and Skateboard manufacturers to see how their popularity progresses.

Popularity of the Facebook-pages will be compared by the number of Page-followers, number of posts and likes those posts get. Though, only the number of likes could be exactly traced over time.

"Rfacebook" package of "R" (programming language) provides an interface to the Facebook API and its function "getPage" can automatically retrieve and frame information from public Facebook pages[14].

The retrieved data will be plotted using the "ggplot2" package for R, which can graphically present the smoothed conditional mean for the plotted data with a function "geom_smooth". This function could be used to compare how the number of likes for posts on the Facebook-pages progresses over time [13, 15].

The smoothed conditional mean graphic of likes per post over time for all the companies combined, shown on the Fig. 1, indicates that all of the chosen Skateboard manufacturers have dominantly a higher popularity than that of BMX and Rollerblading manufacturers.

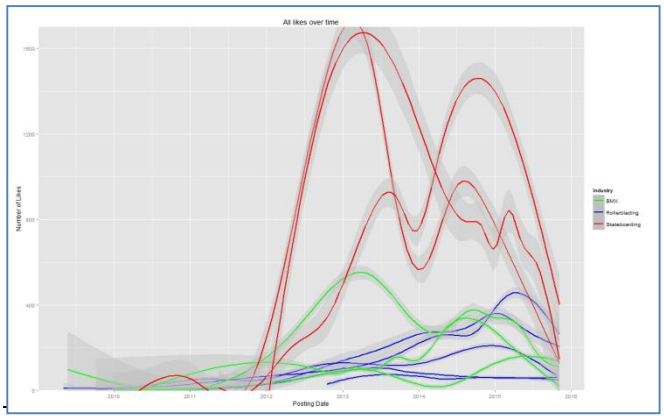

Fig. 1. Likes per post over time

But only one BMX manufacturer ("Wethepeople") has a mostly higher popularity than that of Rollerblading. Popularity of the two other BMX companies lies in the similar limits as that of rollerblading [16].

The progression of the skateboard manufacturers' popularity shows a rather strong fluctuation with a roughly similar pattern and has a resemblance with that of "Wethepeople" (BMX manufacturer). All these companies show extensive growth in popularity in the period between 2012 and 2013 with a strong fall between 2013 and 2014. Others show either a slow progression between 2012 and 2014 or a similar, however, much smoother pattern, which becomes more noticeable between 2014 and 2016. Thus, almost all companies' Facebook-pages could represent the same pattern of the popularity rate in 2014 – 2015.

The comparison of the number of post, shown on the Fig. 2, indicates that 4 out of 5 rollerblades manufacturers place last with the range of almost 500 to almost 1400 post. On the other hand, the 5th rollerblades' manufacturer "Razors" is on the 3rd place with almost 4000 posts.

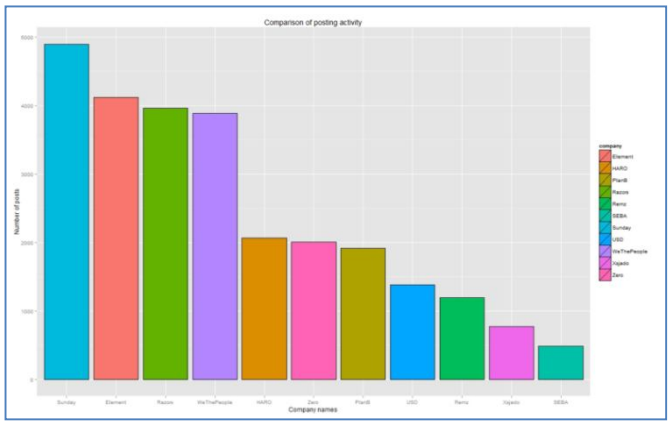

Fig. 2. Comparison of posting activity

This basic study could be used by any company active in one of these three extreme sports` brunches. As an example three of the following rollerblades` manufacturers could be compared: SEBA, Razors and USD.

SEBA is a relative newcomer both to Facebook as a communication platform and to the industry itself is at the bottom of all the represented graphics. Being rather unpopular, SEBA has the second smallest number of followers (as shown on the Fig. 3) and the smallest number of posts.

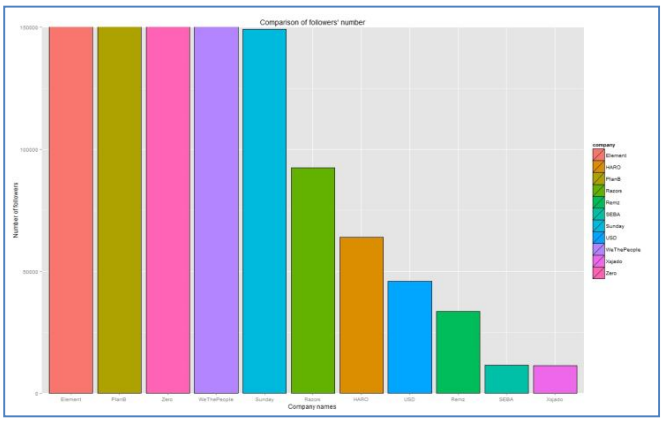

Fig. 3. Comparison of followers` number

Razors, which on the other hand is almost 8 times more active and have roughly 7 times more followers has a practically identical popularity. Whereas USD, being roughly in the middle by its activity and the number of followers, is the leader among rollerblade manufacturers and shows a successful, constant growth of popularity.

Without a corresponding analysis the nature of correlation between popularity/followers/activity is uncertain. However, it could be considered enough for practical use as in this case. It is thus possible to speculate that Razors has troubles generating content of their numerous posts and does not reach its audience. The presented graphics could be then used to determine which of the company`s competitors` Facebook-pages are suitable for a deeper content-analysis, which in turn could provide practical advice for the posting tactics.

CONCLUSIONS

The Internet should be used for the expansion of the information base in the communication strategy formulation. In this paper, we propose a method of using the Internet to expand the range of information sources for analytical work.

Deep web analytics should be the basis of any communication strategy in the Internet. This means using not only such tools as basic metrics of Google Analytics or reports of the Yandex.Metrika Webvisor, but foremost, such advanced tools for automatic data mining and processing as custom-designed R-applications adopted to specifics needs of a company.

This paper overviews capabilities given by modern online social networks` APIs. It provides an example of an R-developed application purposed for a typical task of a first-step data mining and analysis using Facebook API as an access point to the Facebook-pages of the several competitive companies.

PR-managers and marketing strategists can use the ideas presented in this paper to formulate more effective Internet-based communication strategies.

REFERENCES

[1] Евгений Куликов. (2013, November 21). Интернет и экономика, Компьютерное Обозрение [Online]. Available: http://ko.com.ua/internet_i_jekonomika_102575.

[2] Gregory G., M. Karavdic and Sh. Zou, “The Effects of E-Commerce Drivers on Export Marketing Strategy,” Journal of International Marketing, Vol. 15, No. 2, pp. 30-57, Jan. 2007.

[3] S. Summerfield and S. Benninghoven. (2009, September). “Managing More Effectively With a Strategic Communications Plan” [Online]. Available: http://www.westerncity.com/Western-City/September-2009/Managing-More-Effectively-With-a-Strategic-Communications-Plan/

[4] M. Porter. “Competitive Strategy. Techniques or Analyzing Industries and Competitors,” New York: Free Press, 2004.

[5] M. Porter, "Strategy and the Internet," Harvard Business Review 79, no. 3, pp. 62–78. March 2001.

[6] M. Najork and J. Wiener, “Breadth-first crawling yields high-quality pages,” in Proceedings of the 10th international conference on World Wide Web, New York, USA, ACM, 2001, pp. 114-118.

[7] J. Leskovec and C. Faloutsos, ”Sampling from large graphs,” in Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining, New York, USA, ACM, 2006, pp. 631-636.

[8] H. LeHong and J. Fenn. (2012, September 18). Key Trends to Watch in Gartner 2012 Emerging Technologies Hype Cycle. [Online]. Available: http://www.forbes.com/sites/gartnergroup/2012/09/18/key-trends-to-watch-in-gartner- 2012-emerging-technologies-hype-cycle-2/

[9] M. Häsel, “Opensocial: an enabler for social applications on the web,” Communications of the ACM, vol. 54 issue 1,. pp. 139–144, Jan. 2011.

[10] N. Ellison, C. Steinfield and C. Lampe, “The benefits of Facebook “friends:” Social capital and college students` use of online social network sites”, Journal of Computer Mediated Communication, Vol. 12, Issue 4, C. 1143–1168., Jul. 2007.

[11] R. Ackland, “Using Facebook as a data source and platform for e-researching social networks,” in refereed paper presented at the Fourth International Conference on e-Social Science, University of Manchester, pp.18-20, June 2008.

[12] D. Smith, “How Companies Use R to Compete in a Data,”(2014, June 24), Data Informed Big Data and Analitics in the Enterprise, [Online]. Available: http://data-informed.com/companies-use-r-compete-data-driven-world/

[13] W. Chang, “R Graphics Cookbook,” Sebastopol: O'Reilly Media Inc, 2012.

[14] P. Barbera and M. Piccirilli. (2015, August 7). Package 'Rfacebook'. Retrieved from Rfacebook: Access to Facebook API via R [Online]. Available: https://cran.r-project.org/web/packages/Rfacebook/Rfacebook.pdf

[15] M. Pedro. (2015, March 4). Part 3a: Plotting with ggplot2. Retrieved from R-Bloggers [Online]. Available: http://www.r-bloggers.com/part-3a-plotting-with-ggplot2/

[16] D.Urquhart, (Director). (2006). Barely Dead [Motion Picture]. Available: http://everything.explained.today/Barely_Dead/