Model Selection of the Target Audience in Social Networks in Order to Promote the Product

Olena Piatykop [0000-0002-7731-3051], Olga Pronina [0000-0001-7085-8027]

State Higher Educational Institution "Priazov State Technical University", University str., 7, Mariupol, 87555, Ukraine

pee_pstu@ukr.net, pronina.lelka@gmail.com

Abstract. Social networks today is a new type of social relations in the form of a platform for advertising and promotion of goods and services. The paper analyzes the models of target audience formation. Based on the analysis, a mathematical model for the formation of the target audience has been developed. This model takes into account segmentation criteria, customer preferences, their actions regarding products or services. Using the model, it is possible to form a rating of users of a social network for further advertising. The mathematical model formed the basis of the recommendation system for generating recommendations regarding the target audience. The system allows to search for a target audience by criteria and to rank the result by user rating, for further analysis. The system will allow not only quality to be the target audience at the request of the marketer, but also save money on advertising, thereby increasing profits.

Keywords: social networks, target audience, user selection criteria, integrated rating indicator, recommendation system

1 Introduction

In recent years, the world of modern information and communication technologies has been actively developing mobile communications and data transfer technologies. The combination of these technologies provides mobile access to Internet resources. Thanks to this, the public is increasingly using mobile Internet access. According to "We Are Social" reports, more than 4.5 billion people worldwide use the Internet, with most users gaining access to selected web resources and platforms via mobile devices [1].

Humanity is actively using the Internet at home and at work - almost all the time being on-line. The availability of mobile Internet has radically changed the forms, content, mechanisms, functions of social communications. One of such manifestations has become social networks, which have received today the status of an integral attribute of life. More than 3 billion people around the world use social networks [1]. Social networks attract people who pursue different goals: maintaining contact with

old friends and searching for new ones, finding a job, exchanging information and media content with other users, promoting their business.

Significantly increased the role of social networks in sales and marketing. This is due to the fact that social networks give direct access to the buyer. Social networks store and process data on millions of users. Marketers use the information to develop and timely change the brand promotion strategy on social platforms. This helps marketers set up targeted ads, and businesses can drive sales. [2-4]. Advertising campaigns on social networks are developing rapidly. Actively studying the questions of choosing the optimal social networking site for advertising [4].

Many companies, institutions and organizations of various fields of activity are engaged in the promotion of their goods and services on social networks [6-8]. In the study [6], the use of heterogeneous social networks is considered as auxiliary information to increase the effectiveness of hotel recommendations. The article [7] presents a social approach for recommendation systems in the field of tourism, which creates a group profile by analyzing not only user preferences, but also social relations between group members. The main contribution of the document [8] is the provision of a personalized recommendation system, which is suitable for use in systems of the social Internet of things (SIoT).

Many number of factors affect the success of an advertising campaign or an increase in sales [2-5]. The papers by many authors is devoted to the study of these questions [2-8].

In the study [2], the factors that determine the choice in relation to age groups - generations Y- people who were born from 1981 to 2000 were studied. The authors believe that different generations are characterized by unique behavioral patterns and factors that influence the decision to purchase. The article proposes a model of the influence of social networks on the adoption of purchasing decisions of generation Y. Work [5] is devoted to a method for analyzing social networks and identifying significant participants. The model is proposed to detecting actors who can effectively disseminate information to others. Analysis of other publications [3,7] also showed that the number of potential customers and audience growth are important indicators that affect the effectiveness of advertising campaigns on social networks.

Thus, social networks are an effective tool for optimizing trade subject to high-quality advertising in online communities and targeted mailing. The last parameter is especially important. Using thematic data from a social network, it is possible to collect and group information for marketers in narrow areas. Thus, it is possible to track user preferences and offer them related products or services. The collected user data is used to display personalized ads. In this case, everyone will see that advertisement, which, based on the analytics of the social network, allows one to correctly guess his interests. In relation to business, such an approach to advertising is more useful and many times more effective than non-directional advertising, where everyone is shown the same products. To organize such an approach, it is necessary to form a target audience for specific areas.

The target audience is understood as the allocation of a circle of specific people who have common interests and signs that are distinguished from the total mass of potential consumers.

The target audience is a combination of potential and real customers (consumers of a product or service) with an interest in a product or service and united by a certain number of common characteristics, criteria. For example, the criteria include: gender, age, profession, marital status, field of activity, place of residence (big city, rural area, etc.), profit level, family composition, education.

Once a potential target audience is identified, it needs to be classified into a primary target audience and a second target audience. The primary target audience will not be the largest group, but it will bring the most profit and the largest sales.

n social networks, potential participants in the main target audience are active users who are already subscribed to the brand profile, actively like, comment on and buy products.

It is very important to highlight the desired target group. To do this, segmentation (clustering) is performed - dividing customers into groups with similar properties, identifying the needs of the group and forming the target segment.

To segment the target audience, the «5W» method of Mark Sherrington is used [9]. This is the most popular method for determining the target audience and psychological characteristics inherent to customers. The "5W" methodology involves identifying the target audience by searching for answers to the following five questions: what? (segmentation by type of product) who? (consumer segmentation) why? (segmentation by type of motivation to make purchases and consumption) when? (segmentation by time of need), where? (segmentation by place of purchase)

After answering these questions, a portrait of the consumer (a typical user profile) is compiled based on the criteria: demographic (gender, age, nationality, marital status), socio-economic (education and profession, employment status, income and savings), geographic (place, region, country, city, work location), psychographic (character and temperament, lifestyle, system of values, the consumption frequency, marketing activities susceptibility).

Orientation target audience allows to create more personalized advertising for consumers from different segments. The main advantage of this method of determining the target audience is the presentation of a product or service, based on user requests based on their behavior, desires and lifestyle.

Therefore, the stage of gathering the target audience is very important in the stage of promoting a product or service. This is the problem that the current study is devoted to. The purpose of the work is to increase the efficiency of selecting the target audience to promote products (services) in the social network on the basis of an appropriate model that will automate the process. This article is devoted to the formation of a model (method) of target audience selection.

2 Literature Review

Researchers are actively working on algorithms and methods for analyzing data from social networks in order to formulate recommendations [10-12]. Today, recommender systems make a significant contribution to the processing of big data and the provision of relevant information, services, and items to users. Such systems make it possible to automate the provision of recommendations to users on the basis of already completed actions (purchases, ratings, visits, etc.) and feedback results (orders in

stores, following links, etc.) [13]. These systems form recommendations through the use of personal information in social networks, including user preferences and the attractiveness of elements (objects). The main goal is to recommend objects, predicting absent or unobservable interactions between participants in a social network, while identifying various types of objects and links [12].

Systems of social recommendations use various factors to formulate recommendations: the historical behavior of users, social connections between them. For example, if two users are friends, they are likely to have similar preferences [12, 14]. Many recommendation systems use user reviews, ratings, likes, reposts to offer new elements to the user [14-15]. For example, in publication [15], a thematic model of the hidden distribution of pre-placement preferences (RPLDA) model was proposed in order to understand the preferences of mass user repost in relation to different content. To measure the degree of preference of individual users, a thematic preference metric is proposed. And the forecasting reposting function is formulated to identify the target audience.

To determine the similarity of users and their preferences, methods of collaborative filtering can be used [16-17]. The main idea of collaborative filtering algorithms is to propose new elements for a specific user based on his previous preferences or the opinions of other like-minded users. In some cases, social connections can replace the similarity between users in user-oriented color filtering methods [17]. But this approach is not enough to select the target audience of the subject advertising company. It is also necessary to consider the features of the social network. The authors' studies [18] are devoted to the analysis and classification of Facebook profiles based on their demographic, psychometric, lifestyle and value, as well as geographical information that can be obtained from their profiles. This is necessary so that marketers can properly form a target group for their advertising campaigns.

Thus, the selection of the target group from users of the social network is quite an urgent task, which still requires research.

3 Model of Target Audience Formation

To promote a product or service on the Internet, the user needs to know the target audience for which this product is offered. To solve the problem of selecting the target audience, a recommendation system is proposed.

The recommendation system that forms the target audience for promoting a product or service on a social network is presented in the form of a model (1):

$$RS = \langle \{U\}, \{C\} \rangle \quad (1)$$

where $\{U\}$ – set of social network users, $\{C\}$ – set of criteria.

In turn, a set of users $\{U\}$ is represented by elements of the following form:

$$U_i = (Ra, UN, B, Web, Fs, Fg, Pr, R), \forall i \in [1, n] \quad (2)$$

where n is a number of users in the target audience; Ra is a user number on the list; UN is a last name of the patronymic; B is a biography of the user: hashtags, geotags, headings and more; web is the URLs of the user's site; Fs is a followers, subscribers

to the account of the investigated user; *Fg* is a following, the users to which the researched user is subscribed; *Pr* is a private; *R* is a rating of the user in the list.

A set of users *{C}* elements presented to the following:

$$C_j = (CN, W), \forall j \in [1, m] \tag{3}$$

where *m* is a number of criteria by which users are selected; *CN* is a name of the criterion; *W* is a weight of the criterion.

At the beginning of work with the system, the user (marketer) sets a set of criteria, for each of them hears his subjective weight

$$W_j \in [0,1], \forall j \in [1, m] \tag{4}$$

where *m* is the number of criteria by which users are selected, and *m* is the weight of the criterion.

For each criterion, a relative rating is evaluated, based on the following indicators: the number of likes, comments, posts that he has left. Ratings are estimated by formulas (5-7):

$$P_i^{C_j^{Like}} = \begin{cases} \frac{count\left(K_i^{Like}\right)}{MAX_{f=1,n}(K_f^{Like})}, \forall i \in [1,n], j \in [1,m] \\ 0, if\ K_i^{Like} = 0 \end{cases} \tag{5}$$

where *j* is a criterion number, n is a number of users in the target audience; K_i^{Like} is a number of likes.

$$P_i^{C_j^{comment}} = \begin{cases} \frac{count\left(K_i^{comment}\right)}{MAX_{f=1,n}(K_f^{comment})}, \forall i \in [1,n], j \in [1,m] \\ 0, if\ K_i^{comment} = 0 \end{cases} \tag{6}$$

where *n* is a number of users in the target audience; $K_i^{comment}$ is a number of comments.

$$P_i^{C_j^{post}} = \begin{cases} \frac{count\left(K_i^{post}\right)}{MAX_{f=1,n}(K_f^{post})}, \forall i \in [1,n], j \in [1,m] \\ 0, if\ K_i^{post} = 0 \end{cases} \tag{7}$$

where *n* is a number of users in the target audience; K_i^{post} is a number of publications on the network.

Based on the obtained relative ratings, we determine the integral index of the user rating in the list R_i by the formula (8):

$$R_i = \sum_{j=1}^{m} \left(W^j \cdot P_i^{C_j^{Like}} + W^j \cdot P_i^{C_j^{comment}} + W^j \cdot P_i^{C_j^{post}} \right), \forall i \in [1,n] \tag{8}$$

where n is the number of users in the target audience; m is the number of criteria; $P_i^{C_j^{Like}}$ is the relative rating by the number of likes of the i-th user of the j-th criterion; $P_i^{C_j^{comment}}$ is the relative rating by the number of comments of the i-th user of the j-th criterion; $P_i^{C_j^{post}}$ is the relative rating by number of posts of the i-th user of the j-th criterion; W_j is the weight of the j-th criterion.

Based on the calculations, the list of users has been supplemented by an integral index of rating R_i, through which we can organize our list by descending of this indicator according to the formula (9):

$$f(U_i`) \geq f(U_j),\ f: \begin{cases} R_i \geq R_j \\ i,j = 1,2 \dots n \\ i < j \end{cases} \quad (9)$$

Using the resulting list of users, you can reach the target audience using either a boundary value R_i:

$$R_i \leq R_{lim}, \forall i \in [1,n]. \quad (10)$$

You can enter a limit on the number of users who are in the target audience:

$$\forall i \in [1,k],\ \text{де}\ k \leq n. \quad (11)$$

4 Experiments and Further Work

Based on the proposed mathematical model for calculating user ratings to highlight the target audience, a prototype system has been developed. Work with the system begins with a form for entering criteria and their weights. After that, a list is formed and displayed on the screen, as shown in the figure 1.

There are also commands in this window: “Search Queries” for calling a form for entering / editing data according to criteria; “Collect Profiles”; updating data in the user list; “Calculate Rating”; calculating user ratings based on the proposed mathematical model for target audience formation. The resulting list can be further filtered by indicating the boundaries or the number of users, or the rating value.

A study was conducted on the allocation of the target audience for the social network Instagram. After that, a list of users of the potential target audience will be formed taking into account the selected parameters.

To conduct experimental research, it is necessary to identify the target audience. For this, a verbal description of the target audience was compiled. Service - premium engagement rings. The first general description of the target audience: wealthy men from 18 to 45 years old.

A more detailed description of the target audience: men, age from 28 to 45 years, income level - from 80 thousand per month, resident of the metropolis, are in relation-

ships about 1.5 years and above, they are used to looking for information on the network, they prefer Facebook or Instagram from social networks.

The target audience only on the profile of a man does not end there. Since in the modern world a woman performs many social roles. Continued description of the target audience: women, age from 25 to 35, income level - average, from 50-80 thousand per month, residents of the metropolis working in the office have been in a relationship for about 3 years, they prefer Instagram from social networks.

This example assumes that the proposal that is being analyzed is designed not only for men as a prospective buyer, but also for women who can choose this product or buy it.

The following selection criteria for potential customers are selected from the verbal description: gender - male; income is stable; place of residence - Mariupol, Kyev, Odessa, Dnipro, Zaporizhzhia, Lviv; female gender; place of residence - Mariupol, Kyev, Odessa, Dnipro, Zaporizhzhia, Lviv; participates in holiday organization groups, subscribed to photographers and wedding designers.

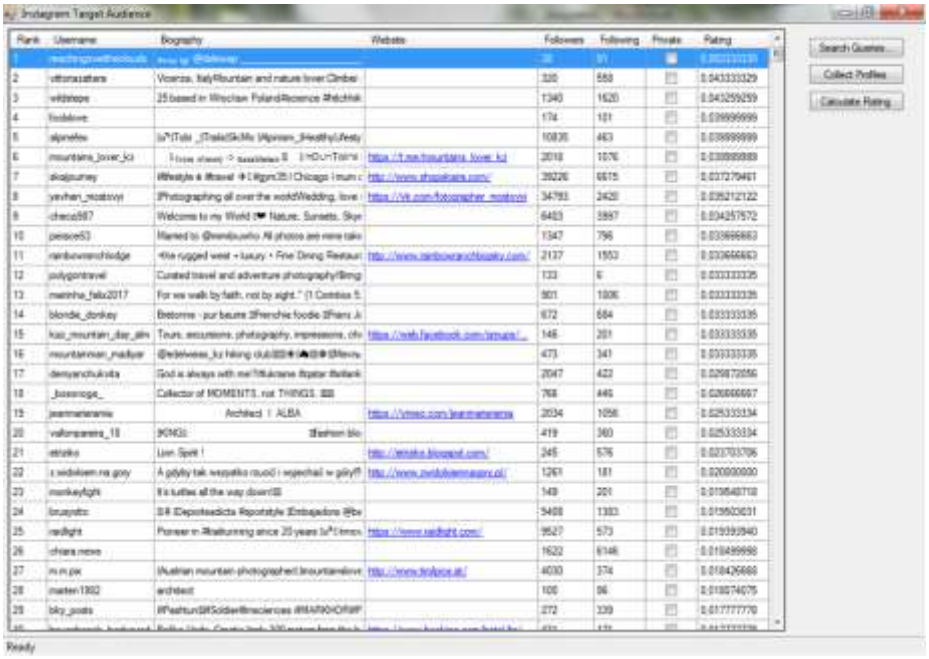

Fig. 1. Target audience for the Instagram network

A description of the target audience is necessary in order to highlight the main groups in which the user can belong, his hobbies, hashtags, as well as possible geolocations. Based on the portrait, parameters for parsing are selected.

A comparative analysis of the existing target audience was carried out using the developed recommendation system and external systems SocialKit [19] and Zengram [20]. The result is shown in figure 2.

In the first experiment, the target audience was formed. The growth dynamics of users who fit the description of the target audience, taking into account the criteria, are presented in the figure 2.

In the experiment, one criterion was first established, the narrowest, after which criteria were added that characterize the target audience. All the added criteria did not narrow the target audience, but expanded it. As can be seen from figure 2, the more criteria are added, the more users as a result. The user growth rate is exponential. The developed recommendation system shows results close to professional systems.

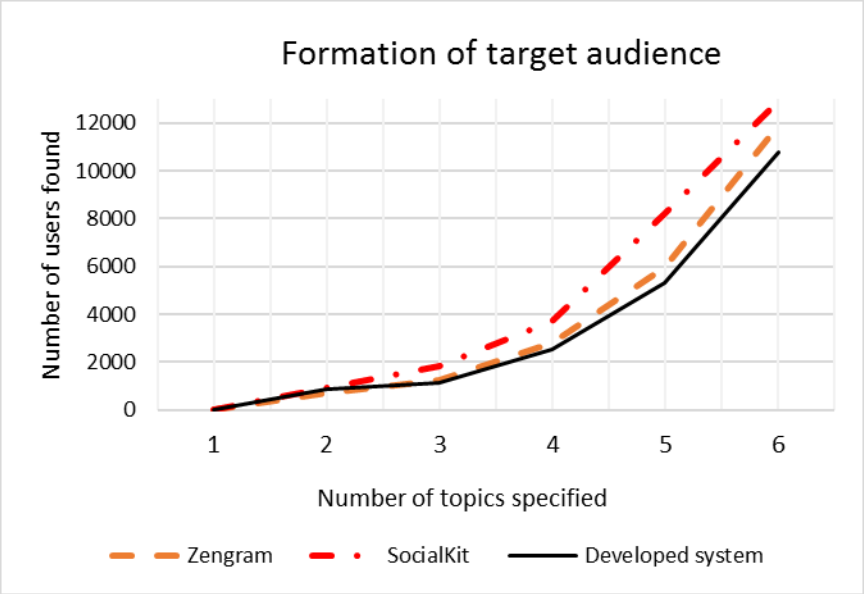

Fig. 2. Comparison of systems for the formation of the target audience

The second experiment was the installation of additional restrictions. Users were selected with the exception of verified profiles (stars, politicians, athletes, brands), which are officially confirmed on Instagram. And also unique profiles were excluded (profiles without repetition of content), profiles without an avatar. Due to these actions, the number of potential users has been reduced.

As a result of selection according to the criteria that were formed during the oral description in the developed system of recommendations, about 9000 records of potential customers were received. When setting additional parameters, the similarity of the behavior of software products becomes apparent.

For each client, a rating was calculated according to formula (8). Since after calculating the rating, it is necessary to analyze the results, it was decided to establish a threshold rating value. The threshold value of the rating eliminates those users who do not pass, thereby reducing the final sample. For the experiment, a threshold rating value was established, which was at least 0.8. This made it possible to reduce the number of potential customers to 1008 units.

As a result of the developed system, the target audience of users of the social network Instagram was obtained according to the selected search criteria.

In the future, it is planned to improve the selection of the target audience according to the specified criteria based on the distribution assessment. Such an assessment will determine the quality of the obtained sample after setting the criteria. Since the number of users cannot be an accurate indicator of the correct operation of systems of this type.

Due to the fact that it is necessary to evaluate the binary distribution, regarding the correctness of the selection of users in the target audience. For these purposes, the use of errors of the type I and type II. Errors of the type I show that the selected user is part of the target audience, and this is a “false positive”. In fact, the analyzed object is not part of the target audience. Errors of the type II implies an “event skipping”, which means that the necessary user was rejected, but in fact was part of the target audience.

If a user is selected according to a query according to established criteria, despite the irrelevance of the search query, the structure of the final selection will change. Moreover, this situation is more preferable than in the case of errors of the second kind, when there is a shortage of users. Since the error of the second kind is the shortage of the user, which means the loss of estimated income.

On the other hand, when advertising is launched with respect to a selected target audience, click-through of impressions with the help of competitive bots plays an important role. In this case, when analyzing this set of target audience, errors of the first kind, namely the exclusion of presumably non-users, will lead to the fact that all bot users can be removed. This means that advertising resources will be saved.

5 Conclusions

The analysis of models of formation of the target audience is carried out, which allows to conclude that the effective methods are the establishment of selection criteria taking into account the experience of the marketer, classification of users of social network, segmentation of the base target audience, as well as drawing a portrait of the target audience.

The development of a recommendation system that allows the user to select the target audience in accordance with the parameters set by the marketer is an urgent task. Since the search for the target audience, taking into account the developed model and the possibility of ranking the results, adjusted for the threshold value of the integrated rating, can increase the effectiveness of the marketer. It also helps reduce advertising costs.

The developed mathematical model of target audience formation takes into account segmentation criteria, customer preferences, their actions on products or services, forms a rating selection of social users using the Instagram network example.

References

1. Digital 2020 Global Overview Report https://wearesocial.com/digital-2020
2. Davidaviciene, V., Meidute-Kavaliauskiene, I., Paliulis, R.: Research on the Influence of Social Media on Generation Y Consumer Purchase Decisions. Marketing and Management of Innovations, vol. 4, pp. 39-49. (2019). doi: 10.21272/mmi.2019.4-04
3. Raudeliuniene, J., Davidaviciene, V., Tvaronaviciene, M., and Jonuška, L.: Evaluation of advertising campaigns on social media networks. Sustainability, 10(4), pp. 1-14. (2018). doi: 10.3390/su10040973
4. Calli, L.: Selection of Social Media Sites for Advertising: Literature Review and a Model Proposal. In: Proceedings of MAC-MME 2016, pp.253-261. (2016)
5. Litterio, A.M., Nantes, E.A., Larrosa, J.M., Gomez, L.J.: Marketing and social networks: a criterion for detecting opinion leaders, European Journal of Management and Business Economics, vol. 26 Issue: 3, pp.347-366, (2017). doi: 10.1108/ EJMBE-10-2017-020
6. Chang, J., Tsai, C., Chiang, J.: Using Heterogeneous Social Media as Auxiliary Information to Improve Hotel Recommendation Performance, In: IEEE Access, vol. 6, pp. 42647-42660, (2018). doi: 10.1109/ACCESS.2018.2855690
7. Christensen, I. A., Schiaffino, S. N., Armentano, M. J.: Social group recommendation in the tourism domain. Journal of Intelligent Information Systems, 47(2): pp. 209–231, (2016). doi: 10.1007/s10844-016-0400-0
8. Cheng, W. K., Ileladewa, A. A., Tan, T. B.: A Personalized Recommendation Framework for Social Internet of Things (SIoT), In: International Conference on Green and Human Information Technology (ICGHIT-2019), Kuala Lumpur, Malaysia, pp. 24-29. (2019). doi: 10.1109/ICGHIT.2019.00013
9. Sherrington, M., Added Value: The Alchemy of Brand-Led Growth, Palgrave Macmillan UK, 2003. Doi: 10.1057/9780230513488
10. Shokeen, J., Rana, C.: Social recommender systems: techniques, domains, metrics, datasets and future scope. Journal of Intelligent Information Systems. pp. 1-35. (2019). doi: 10.1007/s10844-019-00578-5.
11. Schall, D.: Social Network-Based Recommender Systems. (2015). DOI: 10.1007/978-3-319-22735-1.
12. Slokom M., Ayachi R.: A New Social Recommender System Based on Link Prediction Across Heterogeneous Networks. In: Czarnowski I., Howlett R., Jain L. (eds) Intelligent Decision Technologies 2017. IDT 2017. Smart Innovation, Systems and Technologies, vol. 73. (2017). doi: 10.1007/978-3-319-59424-8_31
13. He, J., Chu, W.W.: Design Considerations for a Social Network-Based Recommendation System (SNRS). In: Pardede E. (eds) Community-Built Databases. Springer, Berlin, Heidelberg (2011). doi: 10.1007/978-3-642-19047-6_4.
14. Nguyen, T.D.A., Vu, T.N., Le, T.D.: The Implicit Effect of Items Rating on Recommendation System. In: Dang T., Küng J., Takizawa M., Bui S. (eds) Future Data and Security Engineering. FDSE 2019. Lecture Notes in Computer Science, vol. 11814. Springer, Cham (2019). doi: 10.1007/978-3-030-35653-8_47.
15. Luo, J., Pan, X., Wang, S., Huang, Y.: Identifying target audience on enterprise social network, Industrial Management & Data Systems, vol. 119 No. 1, pp. 111-128, (2019). doi: 10.1108/IMDS-01-2018-0007
16. Pyatikop, E.E.: The research of collaborative filtering method based on neighborhood elements, Scientific papers of Donetsk National Technical University. Series: Informatics, Cybernetics and Computer Science, vol 2 (18), pp. 109-114 (2013)

17. Yang, X., Guo, Y., Liu, Y., Steck, H.: A survey of collaborative filtering based social recommender systems. Computer communications, 41, pp. 1-10. (2014). doi: 10.1016/j.comcom.2013.06.009
18. Risius, M., Aydingül, O.: Facebook User Segmentation to Enable Targeted Social Advertisement. Conference: 39th International Conference on Information Systems, At San Francisco, USA pp. 1-8. (2018).
19. Socialkit is a fully functional Instagram account promotion tool. https://socialkit.ru/
20. Zengram - Instagram account promotion. https://zengram.ru/