

DOI: 10.15593/2499-9873/2021.2.06

УДК 32.019.51

А.Н. Рабчевский¹, Л.Н. Ясницкий^{2,3}, В.С. Заякин²

¹ ООО «СЕУСЛАБ», Пермь, Россия

² Национальный исследовательский университет
«Высшая школа экономики», Пермский филиал, Пермь, Россия

³ Пермский государственный национальный исследовательский
университет, Пермь, Россия

СРАВНЕНИЕ МЕТОДОВ ВЫЯВЛЕНИЯ РОЛЕЙ ПОЛЬЗОВАТЕЛЕЙ В ОНЛАЙНОВЫХ СОЦИАЛЬНЫХ СЕТЯХ

Развитие социальных сетей привело к тому, что они стали использоваться как инструмент пропаганды и мобилизации пользователей для участия в протестном движении и политических акциях, направленных на подрыв устоев общества и свержение действующей власти. Воздействие на социальные сети со стороны организаторов протестных движений становится все более целенаправленным и организованным. В контексте обеспечения общественной безопасности и противодействия деструктивным воздействиям на социальные сети все более актуальной становится задача выявления структуры целенаправленного воздействия на социальную сеть. Важными элементами такой структуры являются роли, которые выполняют пользователи социальной сети, участвующие в протестном движении.

Представлены данные исследования информации о пользователях социальной сети «ВКонтакте» Пермского края, опубликовавших материалы протестной тематики в течение 2020 г. Представлены описания ролей пользователей социальной сети на основе данных об их публикационной активности. Описаны существующие методы выявления ролей пользователей онлайн-социальных сетей на основе кластеризации и нейросетевой классификации. Указаны проблемы, связанные с подготовкой датасетов для качественного обучения нейронных сетей.

Авторами выполнены исследования по выявлению ролей пользователей с использованием различных методов кластеризации, а также предложены оригинальные методы численной оценки ролей пользователей и экспертной нейросетевой классификации ролей пользователей на основе искусственно синтезированных датасетов. Представлены результаты сравнения различных методов кластеризации, метода численной оценки и метода экспертной нейросетевой классификации, указаны их достоинства и недостатки. Показана высокая корреляция между методом численной оценки и методом экспертной нейросетевой классификации. Отмечена более высокая эффективность метода экспертной нейросетевой классификации ролей пользователей социальных сетей по сравнению с различными методами кластеризации. В заключении указаны оптимальные области применения предложенных методов классификации ролей пользователей социальных сетей и намечены направления дальнейших исследований.

Ключевые слова: социальная сеть, роли пользователей, классификация, кластеризация, численные оценки, наиболее влиятельные пользователи, нейронная сеть, экспертная система, искусственный интеллект, протестные движения.

A.N. Rabchevsky¹, L.N. Yasnitsky^{2,3}, V.S. Zayakin²

¹LLC "SEUSLAB", Perm, Russian Federation

²National Research University "Higher School of Economics",
Perm Branch, Perm, Russian Federation

³Perm State University, Perm, Russian Federation

COMPARISON OF METHODS FOR IDENTIFYING USER ROLES IN ONLINE SOCIAL NETWORKS

The development of social media has led to its use as a tool for propaganda and mobilising users to participate in protest movements and political actions aimed at undermining the foundations of society and overthrowing the current government. The impact on social media by the organisers of protest movements has become increasingly targeted and organised. In the context of ensuring public safety and countering destructive influences on social media, it is becoming increasingly important to identify the structure of purposeful impact on social media. Important elements of this structure are the roles played by social network users who participate in the protest movement.

The paper presents the data of a survey of the social network VKontakte users in Perm region, who have published protest-related materials during the year 2020. Descriptions of the roles of social network users based on data on their publication activity are presented. Existing methods of identifying the roles of online social network users based on clustering and neural network classification are described. Problems associated with the preparation of datasets for qualitative training of neural networks are indicated.

The authors have researched user roles using different clustering methods, and proposed original methods of numerical evaluation of user roles and expert neural network classification of user roles based on artificially synthesized datasets. The results of comparison of different clustering methods, numerical estimation method and expert neural network classification method are presented, their advantages and disadvantages are indicated. High correlation between numerical evaluation method and expert neural network classification method is shown. It is noted that the effectiveness of expert neural network classification of roles of users in social networks is higher than that of various clustering methods. In conclusion, the optimum areas of application of the proposed methods for classifying the roles of social network users are indicated and the directions for further research are outlined.

Keywords: social network, user roles, classification, clustering, numerical method, most influence users, neural network, expert system, artificial intelligence, protest movements.

Introduction

The current level of development of social media and the extent of its penetration in everyday life has resulted in social media becoming a powerful tool in organizing political actions and other protest phenomena. Over the past 10 years, the most prominent examples of political actions using social media are the Arab Spring events in 2010–2011, the #Occupy movement in the US in 2011, protests in Turkey, Brazil and Hong Kong (2013–2014), as well as the recent presidential elections in Belarus (2020) and political actions around Navalny's arrest and "Putin's palace" in 2021.

The study of the mechanisms and degree of influence of social networks on people's behavior has generated a great deal of scientific interest. In [1–3], the authors argue that all protest movements are inextricably

linked to the creation of autonomous communication networks supported by the Internet, while in [4, 5] the authors explicitly point to the significant impact of social networks on the level of people's mobilization for active action. According to the authors [6], the use of distributed unorganized social networks facilitates the dissemination of information among protest groups through interpersonal connections and increases the ability to mobilize participants for specific actions.

However, recent protests show that the impact on social networks is becoming more and more organised and targeted, manifesting as information waves triggered by various newsworthy causes. No matter where the news is coming from, whether it is a real event or a fake one created to achieve a certain goal, it is being embedded in social networks, spread among as many online users as possible, amplified through multiple discussions, and supported by a large number of users' approval. In a targeted attack on a network, specific users with specific tasks or roles are behind the actions. Thus, the main challenge in countering targeted influences on social networks is to develop methods and software to identify the roles of users and their level of influence on social networks. This paper explores methods for identifying user roles in online social networks.

1. Overview of methods for identifying user roles

The social roles of users are expressed in the forms of their online activity. In order to analyse and categorise users into roles, data on the number of publications, types of publications, patterns of their behavior etc. are usually used. The sum of such data can be interpreted as creating a conditional user profile. Clustering or classification methods are used to divide such profiles into groups with similar parameter values.

When clustering methods are used, sets of users are divided into groups with the closest parameters. A description and comparison of the best known clustering methods are presented in [7, 8] et al. The authors [9–11] use the K-Means clustering method to partition users into groups, while [12] uses the N-gram clustering method. A combination of several clustering algorithms of K-Means, SOM and DBSCAN is presented in [13]. A combination of centrality and clustering metrics is proposed in [14, 15].

In addition to clustering methods, it is also possible to use neural networks to divide sets into groups. For example, in [16], the authors use a neural network as one of the tools for image classification and in [17] they pro-

pose a hybrid neural network for text classification to detect users' intentions. The use of deep neural networks to classify Twitter users' sentiments is presented in [18]. Judging by the dates of these publications, classification of sets using neural networks is becoming increasingly popular.

Before exploring techniques for identifying users' social roles in social media, it is necessary to define the term "social role". In [19] it is proposed to standardise the use of the term 'social role' in online communities as a set of social, psychological, structural and behavioral attributes, and suggested strategies for identifying users' social roles in some online communities. However, the authors do not propose a strict classification of users' social roles, as the set and definition of social roles depends both on the type of social community and on the context in which user roles are considered.

2. Identifying user roles

In the context of countering the influence of social media on citizens' protest activity, it is of interest to identify the following types of roles:

1. A poster is an idea generator, a content creator, often an opinion leader and, with a lot of connections, can rally a large number of users around him or her.

2. Reporter – a distributor of ideas, rarely creates content, mostly reposts existing publications, aims to maximise dissemination of others' publications.

3. Commentator – does not create content, does not repost, but leaves lots of comments, engages in condemnations and arguments. Often creates redundant comments to increase the popularity of discussion topics.

4. A passive member is a user who is not very active online in terms of creating content, reposts or comments, but regularly visits various pages in the social network. Is the recipient of all information created by Posters, Reposters and Commentators.

3. Identification of user roles by numerical evaluation

When analysing the roles of the users in this paper, data on 1,793 Perm Krai users involved in protest activity in the social network VKontakte, who published protest-related materials in 2020, was used. The list of data collected for each user is presented in Table 1.

Table 1

General statistics of the Perm Krai protest sample for 2020

№ п/п	Parameter name	Min	Max
1	Age of account (days)	86	5 170
2	Number of friends	1	10 000
3	Number of subscribers	1	12 828
4	Number of subscriptions	1	5 481
5	Number of posts created	0	16 221
6	Number of reposts created	0	36 209
7	Number of comments created	0	2 012
8	Number of other people's posts on the wall	0	506
9	Number of reposts received	0	4 299
10	Number of comments received	0	41 108

To determine the roles of users by numerical evaluation, the parameters of the level and form of their activity were used (items 5–7 of Table 1), because it is these parameters that determine the nature of user behaviour on the network.

The first task was to divide the array of users into active and passive users. The division was based on the assumption that active users published significantly more material than passive users. The ratio of materials published by active and passive users was assumed to be 70/30. To identify active users an activity rating was calculated by ranking in descending order of the sum of the number of publications (posts, reposts and comments) for each user. The total amount of content published by users, starting with the leader of the ranking, was then calculated sequentially down the ranking and compared with the total amount of content published. When the sum of materials reached a value of 70 % of the total sum of publications, the counting was stopped. That is, active users published a total of 70 % of the total number of publications, while passive users published the remaining 30 %. Next, a list of users whose sum of publications was 70 % of the total amount of published materials was determined. Mathematically, it can be expressed as follows:

If p_i – the number of posts published by the user i ,

r_i – the number of reposts published by the user i ,

k_i – number of comments published by the user i ,
 the level of activity a_i user i can be expressed as

$$a_i = p_i + r_i + k_i, \tag{1}$$

then the total activity A of all users will be

$$A = \sum_{i=1}^I a_i, \tag{2}$$

where I – total number of users.

Total number of materials A_n , published by all active users can be expressed as

$$A_n = \sum_{i=1}^n a_i = 0,7 \cdot A, \tag{3}$$

where n – user number in the activity rating for which the equality is fulfilled (3). Thus, activity rating users with numbers $\{1,2, \dots n\}$ are active and all others are passive. In our case the number of active users is $n = 243$. The remaining 1,550 users are defined as passive users. A graph of the user activity rating with the boundary between active and passive users is shown in Figure.

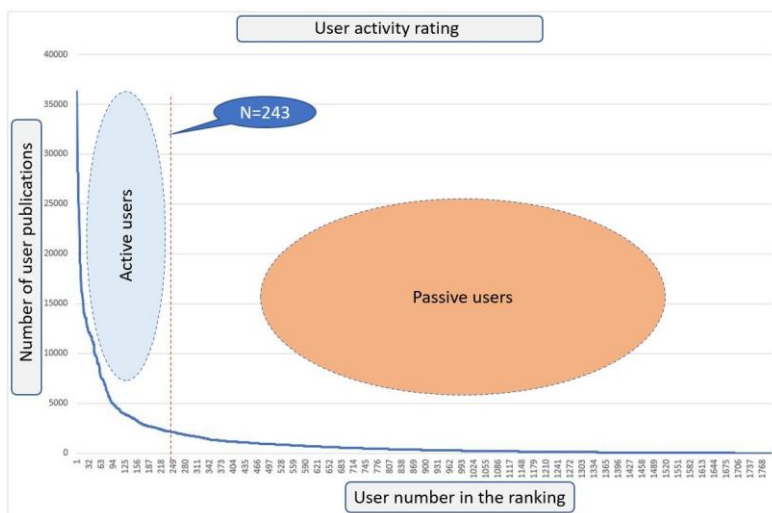


Fig. Activity ranking with boundary of division into active and passive users

As shown in Figure 1, a small number of users with high levels of activity fell into the active category, while there were many more passive users. It is these users that were identified at this stage as users who are Passive participants.

The next task was to classify active users into Poster, Reposter and Commentator roles. For this classification it was assumed that a user performs a pronounced role if one type of activity accounts for at least 60 % of their total activity. In other words, a Poster would have at least 60 % of their published posts out of their total published content, a Reposter would have at least 60 % of their reposts and so on. In other words,

if $p_i \geq 0,6 \cdot a_i$ – the user role is defined as Poster,

if $r_i \geq 0,6 \cdot a_i$ – the user role is defined as Reposter,

if $k_i \geq 0,6 \cdot a_i$ – the user role is defined as Commentator.

As a result of performing calculations of the proportion p_i , r_i and k_i in activity level a_i for each active user, a separation of users by role was performed. Most of the users we identified as active users satisfy the cut-off rule of 60 % for an activity. At the same time, among the active users we identified those who did not satisfy the cut-off rule of 60 % for any of the activities. Such users we have classified as Universal. Exploring the function of Universals is a topic for a separate study. The identified user roles are presented in Table 2.

Table 2

General statistics on the identified roles of users
on the Perm region's protest theme for 2020

№ п/п	User role	Number of users
1	Poster	29
2	Reposter	195
3	Commentator	2
4	Universal	17
5	Passive	1550
6	Total	1793

Thus, using the numerical evaluation method, we identified the roles of the users Poster, Reposter, Commentator, Universal and Passive Participant.

As Table 2 shows, the main number of users in the submitted sample are Passive participants. Given that in the sample presented the number of Commentators is very low and the role of Universalists is yet to be explored, further calculations will be performed for the roles of Poster and Reposter. The parameter ranges for items 1–7 in Table 1 for each role are shown in Table 3.

Table 3

Parameter value ranges identified for the user roles Poster, Reposter, Passive Participant

Parameter		Posters		Reposters		Passive member	
		Min	Max	Min	Max	Min	Max
1	Age of account (days)	311	4382	0	4881	86	5170
2	Number of friends	1	9993	6	8117	1	9999
3	Number of subscribers	1	5568	1	1948	1	12 828
4	Number of subscriptions	7	1148	7	3634	1	5481
5	Number of posts	1579	16 221	0	3319	0	1950
6	Number of reposts	0	5195	1396	36 209	0	2119
7	Number of comments	0	341	0	1123	0	1005
8	Number of other people's posts on the wall	0	87	0	506	0	149
9	Number of reposts received	0	2839	0	4299	0	2027
10	Number of comments received	0	41 108	0	16 459	0	4296

4. Expert neural network classification of user roles

In order to perform neural network classification of user roles, high-quality datasets including training, test and validation sets are needed. The preparation of such datasets is a separate problem. Some authors, such as [20, 21], use ready-made datasets, while others prepare their own datasets for training their neural networks. The use of ready-made datasets is convenient, but can be associated with both difficulties in obtaining them and

incomplete correspondence of the ready dataset to the subject area for which it is planned to be used. In addition, neural networks trained on some networks may be unsuitable for other networks. Pointing this out, the authors of [22] suggest using transfer learning to analyse users' social roles.

Usually, real data from a specific subject area are used to prepare datasets. The data sets obtained in this way are partitioned into specific classes by experts. Although this is the most common method of preparing datasets, it is associated with some problems. For qualitative training of a neural network, it is necessary to ensure a uniform number of examples for each represented class, and the number of examples should be large enough. At the same time, there are always outliers, irregularities and skewness towards one or another class in the real data. Therefore, in order to prepare a high-quality dataset, it is necessary to mark up a very large number of examples, which is associated with high labour costs.

Thus, the authors of [23] prepared a special dataset of 1,000 user profiles to train a neural network that determines the social roles of Twitter users. The authors [24] prepared a dataset based on 740 thousand messages, while the authors of [25] used a dataset consisting of more than 1.2 million text messages extracted from the online community of higher education in Australia. To train their neural network, the authors [26] prepared a dataset based on content analysis of 350 million messages on Twitter.

In our case, there were no ready-made datasets, so we had to create them ourselves, which, in turn, was associated with high labour costs for collecting and marking up the necessary data. To solve this problem, a method of expert neural network classification was applied, based on the use of datasets artificially synthesised based on the knowledge of experts in the given subject area.

The essence of the method is that an expert with deep knowledge of the subject area can easily describe or reconstruct the most characteristic variants of behavior of the object under study or modeled and give examples of the so-called extreme cases. However, the main problem with this method is that the expert is often unable to provide the required number of examples of behavior of the simulated object for the qualitative training of a neural network. At the same time the expert may not indicate specific values of input parameters, but rather allowable ranges within which a change of one or another parameter, in the expert's opinion, will not lead to a significant change in output parameters. Using the ranges given by the expert, it is possible to randomly generate additional examples and bring the training set to the size necessary for high-quality training of the neural network.

The expert neural network was used to identify the roles of Poster, Reposter and Passive Participant. We did not identify the roles Commentator and Universal, as according to Table 2 their number is too small to assess the quality of the neural network. These three roles we have subsequently used as the output parameters of the neural network model. It was assumed that the first output parameter Y1 is equal to one if the user is a Poster and equal to zero if the user is not a Poster. The second output parameter Y2 is equal to one if the user is a Poster and equal to zero if the user is not a Poster, etc. As input parameters X1X7 of the neural network model it was accepted to use characteristics of item 17 of Table 3.

For the generation of sets as expert ranges, the minimum and maximum values of the parameters of item 17 of Table 3 were taken for the roles Poster, Reposter and Passive Participant, identified earlier by numerical evaluation method. Further, 400 examples were synthesized randomly according to the method [27] for each role with equal steps from the maximum to the minimum value X1 and indicating the attribute of belonging to the role Y1, Y2, Y3. In the process of synthesizing the sets, the rule of identifying passive users, according to formula (3), and the predominant form of activity for the roles Poster and Reposter were taken into account:

if $p_i \geq 0,6 \cdot a_i$ – the user role is defined as Poster,

if $r_i \geq 0,6 \cdot a_i$ – the user role is defined as Reposter.

All the sets were then combined into a common set and shuffled randomly. A total of 1200 examples were obtained, which were split into three sets:

- training – 1,000 examples,
- validating – 100 examples,
- testing – 100 examples.

A neural simulator was used to design the neural network Nsim5-10¹ (access www.LbAi.ru). The formula used to estimate the error is:

$$E = \frac{\sqrt{\sum_{n=1}^N (d_n - y_n)^2}}{|\max(d_n) - \min(d_n)|} 100 \% , \quad (4)$$

where N – number of sample elements, d_n – is the declared role of the n th user, and y_n is its role estimated by the neural network.

¹ Neurosimulator 5.0: state registration certificate for computer programme No. 2014618208 dated 12.07.2014 / F.M. Cherepanov, L.N. Yasnitsky (RF).

The best result was shown by the perceptron neural network with seven input neurons, one hidden layer with six neurons and three output neurons. The hyperbolic tangent was used as the activation functions of all neurons. The error of the neural network calculated by the formula (4) for each role is presented in Table 4.

Table 4

Neural network model error values for different roles

Set name	Error Poster (Y1)	Error Reposter (Y2)	Error Passive (Y3)
Training	0,0 %	4,5 %	3,1 %
Testing	0,0 %	0,7 %	5,0 %
Validating	0,0 %	0,0 %	0,8 %

Table 4 shows that the neural network has learned the patterns of user role definition. This means that the neural network behaves in the same way as the simulated users of the social network would behave. Therefore, the neural network can be used to identify the roles of users in the Perm region of the social network VKontakte. Recall that the neural network was trained on an artificially synthesized set of.

A real set classification was then performed using a trained neural network, the results of which were compared with the results of identifying user roles using the numerical evaluation method (Table 5).

Table 5

Correlation between the results of the numerical estimation method and the neural network classification method

№ п/п	Pearson correlation coefficient between the numerical estimate and the neural network classification method	Value
1	Poster	0,97
2	Reposter	0,96
3	Passive	0,98

As can be seen from Table 5, the results obtained by the numerical evaluation method and the neural network expert classification method are

very similar and both methods have proved to be applicable to the task of classifying user roles.

5. Clustering of users

Another method that could be used to select groups of similar users is clustering. K-Means, spectral clustering, and hierarchical clustering algorithms were used to perform clustering, for which the number of clusters to be searched could be specified. The aim of the clustering was to identify the roles Poster, Reposter and Passive Participant, i.e. the number of clusters to be searched is 3. The inputs for the clustering were the same parameters (item 1–7 of Table 3) as for the neural network classification.

Table 6 shows the correlation between the numerical results and the different clustering methods.

Table 6

Correlation between numerical results and clustering methods
for Poster and Reposter roles

Input data	Clustering method	Correlation between numerical estimation and clustering methods Poster	Correlation between numerical estimation and Reposter clustering methods
X1–X7	KMeans	0,20	0,20
X1–X7	Spectral clustering	0,04	0,04
X1–X7	Hierarchical clustering	0,22	0,23

As can be seen from Table 6, classical clustering methods did not produce a positive result when trying to solve our problem.

Conclusions

As we can see from Tables 5 and 6, the results obtained by numerical and neural-network methods almost completely matched, while clustering proved to be inapplicable for our task. Analytical studies of the profiles of the most influential and popular Posters and Reposters showed that most of them are ardent opponents of the current government, actively promoting a protest agenda, having high influence in the social network VKontakte in

the Perm Krai segment. This confirms that we got two equally effective methods to classify user roles.

The use of the numerical estimation method is recommended for ad hoc analyses of user arrays in unsteady or temporary social phenomena. The neural network classification method is more suitable if the social phenomenon is stable and experts can predict adequate values of ranges and ratios of subject matter parameters. An important feature of the neural network classification method is that it can be used in online applications of streaming estimation of user parameters. Such applications can be used to prepare information for deciding to block, suppress activity, or put key users on social network for monitoring. In addition, the use of artificially synthesised sets can be proposed when real-world examples for training neural networks are insufficient and the ranges of variation of the input data are known.

It is envisaged that the results of this work will be used to identify the structure of social network influence, examine patterns of structure behaviour over the life cycles of different information occasions, and develop a methodology for identifying signs of targeted social network influence.

Список литературы

1. Castells M. Networks of outrage and hope. Social movements in the Internet age. – Cambridge. Polity, 2012. – 328 p.
2. Faris D.M. Dissent and revolution in a digital age. – I.B. Tauris Media, 2013. – 267 p. DOI: 10.5040/9780755607839
3. Gerbaudo P. Tweets and the streets. Social media and contemporary activism. – London: Pluto Books, 2012. – 208 p.
4. Tindall D.B. From metaphors to mechanisms: Critical issues in networks and social movements research // *Social Networks*. – 2007. – Vol. 29, iss. 1. – P. 160–168. DOI: 10.1016/j.socnet.2006.07.001
5. Bennett W.L., Segerberg A. The logic of connective action // *Information, Communication & Society*. – 2012. – Vol. 15, iss. 5. – P. 739–768. DOI: 10.1080/1369118X.2012.670661
6. Juris J.S. Reflections on #occupy everywhere: social media, public space, and emerging logics of aggregation // *American Ethnologist*. – 2012. – Vol. 39, iss. 2. – P. 259–279. DOI: 10.1111/j.1548-1425.2012.01362.x
7. Kotyrba M., Volna E., Kominkova Oplatkova Z. Comparison of modern clustering algorithms for two-dimensional data // *Proceedings of 28th European Conference on Modelling and Simulation (ECMS 2014) European Council for Modeling and Simulation, Brescia, Italy, 27–30 May 2014 / Ed. by F. Squazzoni,*

F. Baronio, C. Archetti, M. Castellani. – Brescia, Italy, 2014. – 6 p. DOI: 10.7148/2014-0346

8. Clustering algorithms: A comparative approach / M.Z. Rodriguez, C.H. Comin, D. Casanova, O.M. Bruno, D.R. Amancio, L.dF. Costa, F.A. Rodrigues // PLOS ONE. – 2019. – Vol. 14, iss. 1. – Art. e0210236. DOI: 10.1371/journal.pone.0210236

9. User roles and Contributions in innovation-contest communities / J. Füller, K. Hutter, J. Hautz, K. Matzler // Journal of Management Information Systems. – 2014. – Vol. 31, iss. 1. – P. 273–308. DOI: 10.2753/MIS0742-1222310111

10. Brandtzaeg P.B., Heim J. A typology of social networking sites users // International Journal of Web Based Communities. – 2011. – Vol. 7, iss. 1. – P. 28–51. DOI: 10.1504/IJWBC.2011.038124

11. Çiçek M., Erdoğan E. Social media marketing: exploring the user typology in Turkey // International Journal of Technology Marketing. – 2013. – Vol. 8, iss. 3. – P. 254–271. DOI: 10.1504/IJTMKT.2013.055343

12. Arularasan A.N., Suresh A., Seerangan K. Identification and classification of best spreader in the domain of interest over the social networks // Cluster Computing. – 2019. – Vol. 22, iss. 22. – P. 4035–4045. DOI: 10.1007/s10586-018-2616-y

13. Identification and characterisation of Facebook user profiles considering interaction aspects / P.H.B. Ruas, A.D. Machado, M.C. Silva, M.R.G. Meireles, A.M.P. Cardoso, L.E. Zárte, C.N. Nobre // Behaviour & Information Technology. – 2019. – Vol. 38, iss. 8. – P. 858–872. DOI: 10.1080/0144929X.2019.1566498

14. Identifying influential nodes in complex networks with community structure / X. Zhang, J. Zhu, Q. Wang, H. Zhao // Knowledge-Based Systems. – 2013. – Vol. 42. – P. 74–84. DOI: 10.1016/j.knosys.2013.01.017

15. Lu P., Dong C. Ranking the spreading influence of nodes in complex networks based on mixing degree centrality and local structure // International Journal of Modern Physics B. – 2019. – Vol. 33, no 32. – Art. 1950395. DOI: 10.1142/S0217979219503958

16. Srinivasa Rao T.Y., Chenna Reddy P. Content and context based image retrieval classification based on firefly-neural network // Multimedia Tools and Applications. – 2018. – Vol. 77, iss. 24. – P. 32041–32062. DOI: 10.1007/s11042-018-6224-x

17. A hybrid neural network RBERT-C based on pre-trained RoBERTa and CNN for user intent classification / Y. Liu, H. Liu, L.-P. Wong, L.-K. Lee, H. Zhang, T. Hao // Communications in Computer and Information Science. – 2020. – Vol. 1265. – P. 306–319. DOI: 10.1007/978-981-15-7670-6_26

18. Abdelhade N., Soliman T.H.A., Ibrahim H.M. Detecting Twitter users' opinions of arabic comments during various time episodes via deep neural net-

work // *Advances in Intelligent Systems and Computing*. – 2018. – Vol. 639. – P. 232–246. DOI: 10.1007/978-3-319-64861-3_22

19. Gleave E., Welser H.T., Lento T.M. A Conceptual and operational definition of “Social Role” in online community // 2009 42nd Hawaii International Conference on System Sciences, Waikoloa, HI, USA, 5–8 January 2009. – IEEE, 2009. – 11 p. DOI: 10.1109/HICSS.2009.6

20. Jabłońska M.R., Zajdel R. Artificial neural networks for predicting social comparison effects among female Instagram users // *PLOS ONE*. – 2020. – Vol. 15, iss. 2. – Art. e0229354. DOI: 10.1371/journal.pone.0229354

21. What your Facebook profile picture reveals about your personality / C. Segalin [et al.] // *Proceedings of the 25th ACM international conference on Multimedia*, Mountain View, CA, USA, 23–27 October 2017. – New York: Association for Computing Machinery, 2017. – P. 460–468. DOI: 10.1145/3123266.3123331

22. Jun S., Jerome K., Steffen S. Predicting user roles in social networks using // 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW), Barcelona, Spain, 12–15 December 2016. – IEEE, 2016. – P. 128–135. DOI: 10.1109/ICDMW.2016.0026

23. Sunghwan M.K., Stephen W., Cecile P. Detecting social roles in Twitter // *Proceedings of The Fourth International Workshop on Natural Language Processing for Social Media*, Austin, TX, USA, 1 November 2016. – USA: Association for Computational Linguistics, 2016. – P. 34–40. DOI: 10.18653/v1/W16-6206

24. Matsumoto K., Yoshida M., Kita K. Classification of emoji categories from tweet based on deep neural networks // *Proceedings of the 2nd International Conference on Natural Language Processing and Information Retrieval (NLP IR 2018)*, Bangkok, Thailand, 7–9 September 2018. – New York: Association for Computing Machinery, 2018. – P. 17–25. DOI: 10.1145/3278293.3278306

25. Automated detection of social roles in online communities using deep learning / P. Wijenayake, D. de Silva, D. Alahakoon, S. Kirigeegamage // *Proceedings of the 3rd International Conference on Software Engineering and Information Management*, Sydney, Australia, 12–15 January 2020. – New York: Association for Computing Machinery, 2020. – P. 63–68. DOI: 10.1145/3378936.3378973

26. User-level psychological stress detection from social media using deep neural network / H. Lin, J. Jia, Q. Guo, Y. Xue, Q. Li, J. Huang, L. Cai, L. Feng // *Proceedings of the 22nd ACM international conference on Multimedia*, Orlando, FL, USA, 3–7 November 2014. – New York: Association for Computing Machinery, 2014. – P. 507–516. DOI: 10.1145/2647868.2654945

27. Экспертный способ формирования обучающих выборок на примере создания нейросетевой системы классификации пользователей социальных сетей / Е.А. Рабчевский, А.Н. Рабчевский, В.С. Заякин, Л.Н. Ясницкий // *Нейрокомпьютеры: разработка, применение*. – 2020. – Т. 22, № 5. – С. 54–63. DOI: 10.18127/j19998554-202005-05

References

1. Castells M. Networks of Outrage and Hope. Social Movements in the Internet Age. Cambridge, Polity, 2012, 328 p.
2. Faris D.M. Dissent and Revolution in a Digital Age. Social Media, Blogging and Activism in Egypt. I.B.Tauris Media, 2013, 267 p. DOI: 10.5040/9780755607839
3. Gerbaudo P. Tweets and the Streets. Social Media and Contemporary Activism. London: Pluto Books, 2012, 208 p.
4. Tindall D.B. From metaphors to mechanisms: Critical issues in networks and social movements research. *Social Networks*, 2007, vol. 29, iss. 1, pp. 160–168. DOI: 10.1016/j.socnet.2006.07.001
5. Bennett W.L., Segerberg A. The logic of connective action. *Information, Communication & Society*, 2012, vol. 15, iss. 5, pp. 739–768. DOI: 10.1080/1369118X.2012.670661
6. Juris J.S. Reflections on #Occupy Everywhere: Social media, public space, and emerging logics of aggregation. *American Ethnologist*, 2012, vol. 39, iss. 2, pp. 259–279. DOI: 10.1111/j.1548-1425.2012.01362.x
7. Kotyrba M., Volna E., Oplatkova Z.K. Comparison Of Modern Clustering Algorithms For Two-Dimensional Data. Proceedings of 28th European Conference on Modelling and Simulation (ECMS 2014). [Proceedings edited by: Flaminio Squazzoni, Fabio Baronio, Claudia Archetti, Marco Castellani European Council for Modeling and Simulation, Brescia, Italy, May 27th – 30th, 2014], 2014. – 6 p. DOI: 10.7148/2014-0346
8. Rodriguez M.Z., Comin C.H., Casanova D., Bruno O.M., Amancio D.R., Costa L.dF., Rodrigues F.A. Clustering algorithms: A comparative approach. *PLOS ONE*, 2019, vol. 14, iss. 1, art. e0210236. DOI: 10.1371/journal.pone.0210236.
9. Füller J., Hutter K., Hautz J., Matzler K. User Roles and Contributions in Innovation-Contest Communities. *Journal of Management Information Systems*, 2014, vol. 31, iss. 1, pp. 273–308. DOI: 10.2753/MIS0742-1222310111
10. Brandtzaeg P.B., Heim J. A typology of social networking sites users. *International Journal of Web Based Communities*, 2011, vol. 7, iss. 1, pp 28–51. DOI: 10.1504/IJWBC.2011.038124
11. Çiçek M., Erdoğan rem E. Social media marketing: exploring the user typology in Turkey. *International Journal of Technology Marketing*, 2013, vol. 8, no. 3, pp. 254–271. DOI: 10.1504/IJTMKT.2013.055343
12. Arularasan A.N., Suresh A., Seerangan K. Identification and classification of best spreader in the domain of interest over the social networks. *Cluster Computing*, 2019, vol. 22, iss. 22, pp. 4035–4045. DOI: 10.1007/s10586-018-2616-y

13. Ruas P.H.B., Machado A.D., Silva M.C., Meireles M.R.G., Cardoso A.M.P., Zárate L.E., Nobre C.N. Identification and characterisation of Facebook user profiles considering interaction aspects. *Behaviour & Information Technology*, 2019, vol. 38, iss. 8, pp. 858–872. DOI: 10.1080/0144929X.2019.1566498.
14. Zhang X., Zhu J., Wang Q., Zhao H. Identifying influential nodes in complex networks with community structure. *Knowledge-Based Systems*, 2013, vol. 42, pp. 74–84. DOI: 10.1016/j.knosys.2013.01.017
15. Lu P., Dong C. Ranking the spreading influence of nodes in complex networks based on mixing degree centrality and local structure. *International Journal of Modern Physics B*, 2019, vol. 33, no. 32, art. 1950395. DOI: 10.1142/S0217979219503958
16. Srinivasa Rao T.Y., Chenna Reddy P. Content and context based image retrieval classification based on firefly-neural network. *Multimedia Tools and Applications*, 2018, vol. 77, pp. 32041–32062. DOI: 10.1007/s11042-018-6224-x
17. Liu Y., Liu H., Wong L.-P., Lee L.-K., Zhang H., Hao T. A Hybrid Neural Network RBERT-C Based on Pre-trained RoBERTa and CNN for User Intent Classification. *Communications in Computer and Information Science* [International conference Neural Computing for Advanced Applications (NCAA 2020)]. Springer, Singapore, 2020, vol. 1265, pp. 306–319. DOI: 10.1007/978-981-15-7670-6_26
18. Abdelhade N., Soliman T.H.A., Ibrahim H.M. Detecting Twitter Users' Opinions of Arabic Comments During Various Time Episodes via Deep Neural Network. *Advances in Intelligent Systems and Computing* [Proceedings of the International Conference on Advanced Intelligent Systems and Informatics 2017 (AISI 2017), In: Hassanien A., Shaalan K., Gaber T., Tolba M. (eds)]. Springer, Cham, 2018, vol. 639, pp 232–246. DOI: 10.1007/978-3-319-64861-3_22
19. Gleave E., Welser H.T., Lento T.M. Smith M.A. A Conceptual and Operational Definition of “Social Role” in Online Community. *2009 42nd Hawaii International Conference on System Sciences* [Waikoloa, HI, USA, 5-8 Jan. 2009]. IEEE, 2009, 11 p. DOI: 10.1109/HICSS.2009.6
20. Jabłońska M.R., Zajdel R. Artificial neural networks for predicting social comparison effects among female Instagram users. *PLOS ONE*, 2020, vol. 15, iss. 2, art. e0229354. DOI: 10.1371/journal.pone.0229354
21. Segalin C., Celli F., Polonio L., Kosinski M., Stillwell D., Sebe N., Cristani M., Lepri B. What your Facebook Profile Picture Reveals about your Personality. *Proceedings of the 25th ACM international conference on Multimedia*. [Mountain View, CA, USA, October, 23–27, 2017]. New York, Association for Computing Machinery, 2017, pp. 460–468. DOI: 10.1145/3123266.3123331
22. Jun S., Jerome K., Steffen S. Predicting User Roles in Social Networks using. *2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW)* [Barcelona, Spain 12-15 Dec. 2016]. IEEE, 2016, pp. 128–135. DOI 10.1109/ICDMW.2016.0026

23. Sunghwan M.K., Stephen W., Cecile P. Detecting Social Roles in Twitter. *Proceedings of The Fourth International Workshop on Natural Language Processing for Social Media* [Austin, TX, 1 November 2016]. Association for Computational Linguistics, 2016, pp. 34–40. DOI 10.18653/v1/W16-6206

24. Matsumoto K., Yoshida M., Kita K. Classification of Emoji Categories from Tweet Based on Deep Neural Networks. *Proceedings of the 2nd International Conference on Natural Language Processing and Information Retrieval (NLPPIR 2018)* Bangkok, Thailand, 7–9 September 2018. New York, Association for Computing Machinery, 2018, pp 17–25. DOI: 10.1145/3278293.3278306

25. Wijenayake P., Silva D. de, Alahakoon D., Kirigeegamage S. Automated Detection of Social Roles in Online Communities using Deep Learning. *Proceedings of the 3rd International Conference on Software Engineering and Information Management (ICSIM-2020)* [Sydney, Australia, 12–15 January 2020]. New York, Association for Computing Machinery, 2020, pp. 63–68. DOI: 10.1145/3378936.3378973

26. Lin H., Jia J., Guo Q., Xue Y., Li Q., Huang J., Cai L., Feng L. User-level psychological stress detection from social media using deep neural network. *Proceedings of the 22nd ACM international conference on Multimedia* [Orlando, FL, USA, 3–7 November 2014]. New York, Association for Computing Machinery, 2014, pp. 507–516. DOI: 10.1145/2647868.2654945

27. Rabchevskii E.A., Rabchevskii A.N., Zaiakin V.S., Yasnitsky L.N. Экспертный способ формирования обучающих выборок на примере создания нейросетевой системы классификации пользователей социальных сетей [Expert method of training samples forming on the example of creating a neural network classification system for social network users]. *Нейрокомп'ютеры: разработка, применение*, 2020, vol. 22, no. 5, pp. 54–63. DOI 10.18127/j19998554-202005-05

Статья получена: 29.03.2021

Статья принята: 09.06.2021

Сведения об авторах

Рабчевский Андрей Николаевич (Пермь, Россия) – директор по науке ООО «СЕУСЛАБ» (614066, Пермь, шоссе Космонавтов, 111, e-mail: andrey@ranat.ru).

Ясницкий Леонид Нахимович (Пермь, Россия) – доктор технических наук, профессор, профессор кафедры «Информационные технологии в бизнесе», Национальный исследовательский университет «Высшая школа экономики», Пермский филиал; профессор кафедры «Прикладная математика и информатика», Пермский государственный национальный исследовательский университет (614068, Пермь, ул. Букирева, 15, e-mail: yasn@psu.ru).

Заякин Виктор Сергеевич (Пермь, Россия) – студент магистратуры, Национальный исследовательский университет «Высшая школа экономики», Пермский филиал (614070, Россия, Пермь, ул. Студенческая, 38, e-mail: vszayakin@yandex.ru).

About the authors

Andrey N. Rabchevsky (Perm, Russian Federation) – Director of Science, LLC “SEUSLAB” (111, Shosse Kosmonavtov, Perm, 614066, e-mail: andrey@ranat.ru).

Leonid N. Yasnitsky (Perm, Russian Federation) – Dr. Habil. in Engineering, Professor, Department of Information Technologies in Business, National Research University “Higher School of Economics”, Perm Branch; Professor, Department of Applied Mathematics and Informatics, Perm State University (17, Bukireva st., Perm, 614068, e-mail: yasn@psu.ru).

Victor S. Zayakin (Perm, Russian Federation) – Master Student, National Research University “Higher School of Economics”, Perm Branch (38, Studencheskaya st., Perm, 614070, e-mail: v.zayakin@seuslab.ru).

Библиографическое описание статьи согласно ГОСТ Р 7.0.100–2018:

Rabchevsky, A.N. Comparison of methods for identifying user roles in online social networks = Сравнение методов выявления ролей в социальных сетях / A. N. Rabchevsky, L. N. Yasnitsky, V. S. Zayakin. – текст : непосредственный. – DOI: 10.15593/2499-9873/2021.2.06 // Прикладная математика и вопросы управления = Applied Mathematics and Control Sciences. – 2021. – № 2. – С. 91–109. – Ст. на англ. языке.

Цитирование статьи в изданиях РИНЦ:

Rabchevsky A.N., Yasnitsky L.N., Zayakin V.S. Comparison of methods for identifying user roles in online social networks // Прикладная математика и вопросы управления. – 2021. – № 2. – С. 91–109. DOI: 10.15593/2499-9873/2021.2.06

Цитирование статьи в references и международных изданиях

Cite this article as:

Rabchevsky A.N., Yasnitsky L.N., Zayakin V.S. Comparison of methods for identifying user roles in online social networks. *Applied Mathematics and Control Sciences*, 2021, no. 2, pp. 91–109. DOI: 10.15593/2499-9873/2021.2.06