SOCIETY FOR SCIENCE AND EDUCATION
UNITED KINGDOM

TNC TRANSACTIONS ON NETWORKS AND COMMUNICATIONS

VOLUME 7, NO. 5
ISSN: 2054 -7420

Regularities in the Social Network's User Distribution by the Number of Mutual Contacts

[1,2]Ibragim E Suleimenov, [2]Akhat S Bakirov

[1] Almaty University of Power Engineering and Telecommunications, Almaty, Kazakhstan;

[2]National Academy of Engineering of the Republic of Kazakhstan, Almaty, Kazakhstan;

esenych@yandex.ru; axatmr@mail.ru

ABSTRACT

This paper describes a study in which the pattern of distribution of the number of friends among users of the Vkontakte social network among residents of eight large cities of the CIS (Moscow, St. Petersburg, Almaty, Novosibirsk, Tashkent, Kiev, Yekaterinburg, Pavlodar) was studied. Experimental data show that this distribution is of a similar nature for all selected cities. A semi-empirical model was built, on the basis of which an explicit form of theoretical dependence was obtained, describing the nature of the distribution of the number of mutual contacts ("friends") of users of social online networks. It is shown that this theoretical dependence agrees with satisfactory accuracy with experimental data for a sufficiently large sample of cities. It is established that the Dunbar number, which is included in the dependencies considered as a control parameter, is a characteristic of the communication environment of each specific city and correlates with the population of the city.

Keywords: Online social networks, Dunbar number, City communication environment.

1 Introduction

The study of the communication space formed by the rapid development of telecommunication technologies is of considerable interest in several aspects.

Thus, a significant part of campaigns to promote products on the market is currently conducted in online social networks [1-5], and such networks are already used not as an auxiliary means of campaigning, but as the main one.

Social online networks connecting a large number of people are an ideal platform for conducting various informational influences of a mass nature, including in the framework of information wars [6-11].

At the same time, modern telecommunication technologies make it possible to receive education online, which creates a number of advantages (access to educational courses of universities around the world, saving travel time, convenience for people with advanced requirements) [12-16]. Massive open online courses (MOOCs) are beginning to gain great popularity [17-21].

The consistent use of social online networks for various purposes requires their adequate theoretical description on a quantitative level.

DOI: 10.14738/tnc.75.7552
Publication Date: 15th December 2019
URL: http://dx.doi.org/10.14738/tnc.75.7552

In this paper, we analyze the distribution of users of social online networks by the number of friends (mutual contacts) and propose a model that describes this distribution.

2 Research method

The data used in the work was collected on the basis of direct collection of information regarding the number of friends of users of social online networks VKontakte.

A random sample of users living in the following cities was created: Almaty, Moscow, Novosibirsk, Kiev, Pavlodar, Yekaterinburg, St. Petersburg, Tashkent. The number of friends each user has was counted. Based on these data, the dependencies presented in Fig. 1 - 8 (dots). The abscissa axis shows the numbers n falling at the center of the intervals on which the axis is divided. The partition scale is selected uneven. The ordinates show the ratio of the total number of users, whose number of friends falls within this interval, to the length of this interval. Such a construction approximately corresponds to the density of distribution of users by the number of friends.

It can be seen that the curves obtained for various cities in which the Runet is actively exploited have a similar character.

In general, the distribution of the number of users by the number of friends can be found on the basis of the following equation:

$$\frac{dn_j}{dt} = n_{j-1} \sum_{i=0} w_{j-1,i} n_i - n_j \sum_{i=0} w_{j,i} n_i - \frac{n_j}{\tau} \tag{1}$$

nj – the number of users whose friends are j, $w_{j,i}$ — the frequency of the formation of a new relationship between people with j and i friends, τ - network user lifetime (actual time of using the resource).
Assuming that the functions under consideration with a change in the number j change slowly, we can proceed to the continuous form of writing equation (1). Replacing the discrete numbers j and i with continuous variables x and y, we can write

$$\frac{d}{dt} n(x) = n(x - \Delta x) \int w(x - \Delta x, y) n(y) dy - n(x) \int w(x, y) n(y) dy - \frac{n(x)}{\tau} \tag{2}$$

Applying the Taylor series expansion up to quadratic terms, we have

$$\frac{d}{dt} \frac{n(x)}{\Delta x} = -n(x) \int \frac{\partial}{\partial x} w(x, y) n(y) dy - \frac{dn}{dx} \int w(x, y) n(y) dy - \frac{1}{\tau} \frac{n(x)}{\Delta x} \tag{3}$$

Whence we get that for the stationary case

$$\frac{df}{dx} = f(x) K(x) \tag{4}$$

where

$$K(x) = \frac{\left[\int \frac{\partial}{\partial x} W(x,y) f(y) dy - \frac{1}{\tau}\right]}{\int W(x,y) f(y) dy} \tag{5}$$

It is essential that the functions included in the factor for the function f (x) in expression (4) are determined by the entire profile of the distribution under consideration.

Expression (4) allows us to interpret the establishment of an equilibrium distribution of users by the number of friends through a stream directed towards their increase. This, in particular, allows us to use heuristic considerations in finding the function K (x).

These considerations are as follows. The literature currently uses the Dunbar number [22-25], which represents the maximum number of constant social contacts (ties) that one person can support. Initially, this number was obtained by studying the influence of the size and development of the neocortex on the size of a flock of monkeys [22-23], but now it is increasingly used to describe the communication space [26-29].

It can be assumed that the Dunbar number describes the potential causing the flow in a positive direction along the x axis in the model under consideration. If the number of friends of this user is less than the Dunbar number, then we can expect that the derivative in formula (4) will have a positive sign (and vice versa).

The simplest record satisfying this requirement has the form

$$\frac{df}{dx} = (a - x)f^{\alpha} \tag{6}$$

where the factor α takes into account the fractal nature of the communication space.

This nature of the communication space follows from visual considerations. The formation of a relationship between two persons is facilitated if they have a common friend, i.e. there is a certain chain linking them. The inclusion of such chains leads to a power-law dependence in formula (6) and the fractal nature of the communication space as a whole.

Equation (6) is of the first order and is easily integrated; we have

$$-(1 - \alpha)f^{1-\alpha} = (a - x)^2 + C \tag{7}$$

where C is the integration constant.

From here

$$f = \frac{A}{((a-x)^2 + C)^{\beta}} \tag{8}$$

where the normalization factor A is entered in the record,

$$\beta = \frac{1}{\alpha - 1}; \ \ \alpha = 1 + \frac{1}{\beta} \tag{9}$$

A comparison of the experimental results with theoretical calculations based on the proposed model is presented in Fig. 1-8.

The theoretical curves (solid lines) in these figures were obtained based on the model described above. It is seen that there is good agreement between the experimental and theoretical data.

The control parameters defining the theoretical curve were determined on the basis of experimental data using the least squares method. The obtained numerical values, using which theoretical curves are plotted, shown in Fig. 1 to 8 are summarized in Table 1.

It can be seen that analogues of the Dunbar number obtained on the basis of the analysis of the presented dependencies fit into the range (100-230) indicated in the literature [22-25]. It is also seen that the analogue of the Dunbar number for Moscow is noticeably higher than for Almaty and Novosibirsk, which could be expected from general considerations.

This means that the analogue of the Dunbar number included in the dependencies used as a parameter can be considered as a characteristic of the communication space for each specific city.

The most significant result is that the fractal dimension of the communication space is the same for all eight cities, and the value is 2/3 with high accuracy. It can be assumed that this value reflects some fundamental features of the formation of the communication space.

Table 1. Empirically determined constant formulas (8) for various settlements and the population of cities.

City	A	C	a	β	Population, thousand people
Almaty	820	5500	95	0,654	1806,6
Moscow	1290	43000	300	0,657	12506,5
Novosibirsk	1120	14000	95	0,651	1612,8
Kiev	1100	20000	130	0,656	2934,5
Pavlodar	800	3000	110	0.648	334,9
Yekaterinburg	800	9800	120	0,.649	1468,8
St. Petersburg	716	9350	205	0.65	5351,9
Tashkent	200	1500	140	0.654	2424,1

It should be noted that the obtained values of the Dunbar number (considered as parameters of the theoretical dependence) correlate with the population of the city (Fig. 9).

This figure shows that the Dunbar number is not an absolute indicator, but significantly depends on the characteristics of the communication environment, mainly the population of the city.

This circumstance cannot be surprising - the larger the city, the greater the number of communications a particular person is involved. He, on average, is better informed, more sociable, etc. Of course, these factors are not the only ones, but we can speak with all certainty about the influence of the nature of the communication environment on the value of the Dunbar number.

3 Conclusion

In conclusion, we note that based on heuristic considerations and literature data on the nature of the Dunbar number, we can propose a simple semi-empirical model that adequately describes the experimental distributions of users of social online networks by the number of friends.

This model is based on the assumption of the fractal nature of the communication space. This assumption seems to be justified also because the fractal dimension obtained for various cities in which the Runet is actively used remains constant with high accuracy

REFERENCES

[1]. Chen W., Wang C., Wang Y. (2010). Scalable influence maximization for prevalent viral marketing in large-scale social networks. In Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM. P. 1029-1038.

[2]. Doyle S. (2007). The role of social networks in marketing. Journal of Database Marketing & Customer Strategy Management. T. 15. №. 1. P. 60-64.

[3]. Stelzner M. (2014). 2014 Social Media Marketing Industry Report. Social media examiner. P. 1-52.

[4]. Chaffey D. (2016). Global social media research summary 2016. Smart Insights: Social Media Marketing.

[5]. Ilyashenko S.B. (2016). Social networks internet as a tool to improve e-commerce. In Proceedings of the International Scientific and Practical Conference. P. 105-107.

[6]. Hunter E., Pernik P. (2015). The challenges of hybrid warfare. International Centre for Defence and Security, 2015.

[7]. Klausen J. (2015). Tweeting the Jihad: Social media networks of Western foreign fighters in Syria and Iraq. Studies in Conflict & Terrorism. T. 38. №. 1. P. 1-22.

[8]. Kamalipour Y.R., Friedrichsen M. (2017). Introduction: Digital Transformation in a Global World. Digital Transformation in Journalism and News Media. Springer, Cham. C. 1-4.

[9]. Sergeev I.V. (2015). Social networks on the Internet as a means of implementing operations of information and psychological warfare. International Scientific Research Journal. №. 9(40) Part 4. P. 101-104.

[10]. Chernova, T.A., Slepovronskaya, K.Yu. (2015). Historical memory in the information war. How are social networks used in the ideological confrontation between Ukraine and Russia. Philosophical Sciences. №. 5. P. 16-23.

[11]. Zheltukhina M.R., Pavlov P.V. (2016). Social network "Facebook" in the XXI century: from the communication tool to the information war tool. Philology. Theory and practice. №. 7-3. P. 89-93.

[12]. De-Marcos L. et al (2014). An empirical study comparing gamification and social networking on e-learning. Computers & Education. T. 75. P. 82-91.

[13]. De-Marcos L. et al (2016). Social network analysis of a gamified e-learning course: Small-world phenomenon and network metrics as predictors of academic performance. Computers in Human Behavior. T. 60. P. 312-321.

[14]. De Meo P. et al (2017). Combining trust and skills evaluation to form e-Learning classes in online social networks. Information Sciences. T. 405. P. 107-122.

[15]. Bukayeva A.A., Magzumova A.T. (2015). The use of social networks in the educational process. Innovations in Science. №42 p. 120-126.

[16]. Kalmykova S.V., Krasnoshchekov A.V. (2017). Approaches to e-learning organization through the social networks. Modern education: content, technology, quality. T. 1. P. 195-198.

[17]. Makhmutova M.V., Makhmutov G.R. (2015). Models and platforms for the implementation of mass open online courses. Modern information technologies and IT education. T. 1. №. 11. P. 486-496.

[18]. Kurzaeva L.V., Grigoriev A.D. (2015). Mass open online courses: essence, specific characteristics. New information technologies in education. P. 250-253.

[19]. Bremer C., Weiß D. (2013). How to analyze participation in a (C) MOOC?. Univ.-Bibliothek Frankfurt am Main. P. 992-1002.

[20]. Khalil M., Brunner H., Ebner M. (2015). Evaluation grid for xMOOCs. International Journal of Emerging Technologies in Learning (iJET). T. 10. №. 4. P. 280-314.

[21]. Colas J.F., Sloep P.B., Garreta-Domingo M. (2016). The effect of multilingual facilitation on active participation in MOOCs. The International Review of Research in Open and Distributed Learning. T. 17. №. 4. P. 4-11.

[22]. Dunbar R.I.M. (1993). Coevolution of neocortical size, group size and language in humans. Behavioral and brain sciences. T. 16. №. 4. P. 681-694.

[23]. Dunbar R. (1998). Grooming, gossip, and the evolution of language. Harvard University Press.

[24]. Dunbar R.I.M. (2009). The social brain hypothesis and its implications for social evolution. Annals of human biology. T. 36. №. 5. P. 562-572.

[25]. Shultz S., Dunbar R.I.M. (2012). The social brain hypothesis: an evolutionary perspective on the neurobiology of social behavior. I Know What You Are Thinking: Brain Imaging and Mental Privacy. P. 13-28.

[26]. Gilyazova A.A. (2014). On the estimation of the number of Dunbar and its influence on the interaction of users in the social network Facebook. In XII All-Russia Conference on Control Problems ARCCP-2014. P. 6292-6300.

[27]. Chesnokov V.O. (2017). The use of the algorithm for isolating communities in the information confrontation in social networks. Cybersecurity issues. №. 1 (19). P. 37-44.

[28]. Katasev A.S., Kirpichnikov A.P., Ramazanova R.I. (2015). Method of analyzing the security of social network accounts from malicious content. Bulletin of Kazan Technological University. T. 18. №. 18. P.195-198.

[29]. Rzayeva M.E. (2015). Evaluation of the effectiveness of special projects in Russian online media through the determination of the price of contact. The media sphere: problems and points of development.

