

METHOD FOR DETERMINATION OF THE SOCIAL GRAPH ORIENTATION BY THE ANALYSIS OF THE VERTICES VALENCE IN THE CONNECTIVITY COMPONENT

E.V. Garin, R.V. Meshcheryakov

Tomsk University of Control Systems and Radioelectronics, Tomsk, Russian Federation

E-mail: mrv@tusur.ru

The work is the continuation of the authors' work on the simulation of the structure of the society by the method of random graphs generation. In their previous works, the authors experimentally proved that the social graph has a strict orientation of the information direction propagation from the shares of the graph with high bonds density to the graph parts having a lower bond density. The authors proposed a method for determining the orientation of the social graph by investigating the ratio of the number of outgoing links to the incoming ones. This method found its application in the analysis of postal mailing, but it turned out to be practically inapplicable in the analysis of social networks, since the counting of incoming and outgoing communications required the compilation of the announced social graphs which was a time-consuming and computational resources task that does not have a solution for polynomial time at present.

Since the issue of the social networks analysis is not only of a strictly scientific, but also of a practical interest, the authors developed and tested a technique for determining the orientation of social graphs by the method of analyzing the numerical characteristics of a graph. The new method does not require a detailed analysis of the correspondence of users of social networks, but operates with open user information that is a list of friends (Friend List). The ultimate goal of this work is to develop simple and effective methods for analyzing social networks to identify "opinion leaders", ways of disseminating information, including propagandizing deviant and dependent forms of behavior, identifying anti-systems and separate closed network communities, and general monitoring of the state of social systems.

Keywords: social graph; random graphs; numerical characteristics of graphs; social graph orientation.

Introduction

The present work is devoted to the social graphs methods and researches. Two interrelated areas: the theory of generating of random graphs [1–7] and practical methodologies for studying of social and technical networks intersect in this field [8], remote diagnostics of a person's state using methods of mathematical modeling [9]. Both of these areas complement each other, and they are used for network monitoring: identifying channels of distribution of narcotic and psychotropic substances, the spread of radical literature and ideology, determining critical situations, to find out "opinion leaders" influencing on the nearest social community and forming the public opinion. Methods of identifying the orientation of the social graph that is identifying the direction of information propagation in social networks from the "opinion leaders" to the audience of readers presented in this paper.

1. An epidemic as a model of information dissemination in a network

In 2010 Grant Schoenbeck showed that if the epidemic distributes across the network of a random Watts–Strogatz graph [10] we will get a network that satisfies a power law, fig. 1.

The original Watts–Strogatz model did not satisfy the power law, because initially it was developed in accordance with the conditions of having a small diameter in the social graph. As well the Watts–Strogatz model imitated the presence of "linking rings" in the Internet web rather well. Linking rings represent definite motives of the social graph, namely the looping banner links of the advertising sites

with each other in order to increase the rate of citation artificially by indexing cloud search systems. Despite the artificial formation of these motives of the social graph, actually they reflect the presence observed in the closed ways of disseminating information.

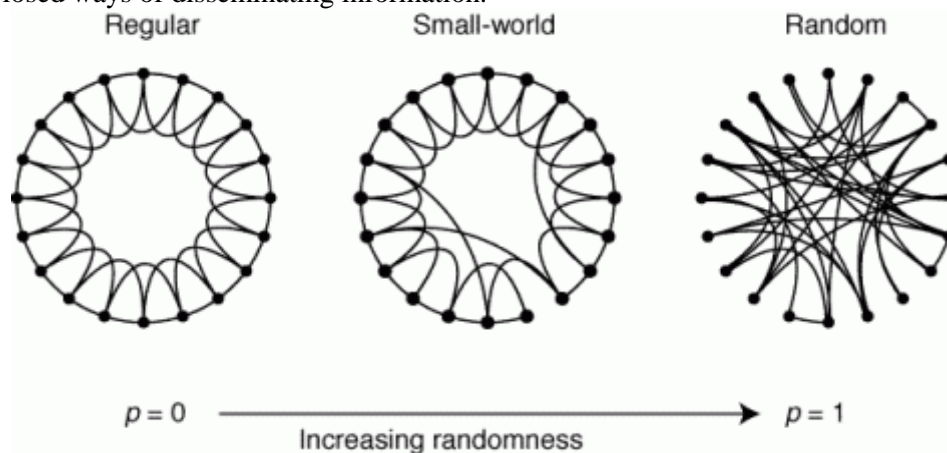


Fig. 1. Watts–Strogatz model of a random graph

In a series of studies, the authors showed that in addition to the closed nature of information dissemination in social graph, there exists a strict orientation of information dissemination [11, 12]. In this case, the authors developed a method of detecting the social graph orientation by the percolation coefficient analysis, where the percolation coefficient is the ratio of outbound links to inbound links. The percolation coefficient ratio refers increasingly to the entire society as a whole, or to certain groups of the society, such as, for example, the audience of a particular TV, radio, or the Internet media. It is hard to determine the percolation coefficient for a single person, moreover, all studies of the information dissemination in the society were aimed constantly at the studying of the group coefficient of percolation. This method is applicable to the analysis of the e-mail distribution when we know who sends the letter and from whom the letter is sent. Such method was not practically applicable in social network analysis, because it requires the individual review pages for each user of the social graph investigated to identify correlations of received and sent messages. It requires a more simple method of identifying the graph orientation by the available open data in the friend lists of users for the analysis of social networks.

2. A method of determining the orientation of the graph by the valence of the vertices (“a method to identify opinion leaders”)

To determine the distribution properties of the information in the society the authors conducted several experiments in social networks LinkedIn and Facebook. The oriented graph of information dissemination was constructed based on the results of the experiments. Initially, knowing the information about orientation of a graph, it was necessary to determine what easy calculation numerical characteristic of graphs correspond to their orientation.

The most obvious authors’ initial assumption was the conformity of the orientation distribution of the graph vertices according to the number of friends in friend lists (fig. 2).

The compliance verification of previously discovered graph orientation and the distribution according to the number of friends in friend lists shows serious discrepancies (right up to the fact, that the vertices on which closed oriented graph gets to the top of the distribution, because they have 5 000 friends). Besides, it is clear from figure 2 that almost 30 people out of 107 have a maximum value for the number of friends ($\approx 5\,000$), so it becomes impossible to identify the parameter values friend list of “opinion leaders” because there are too many potential candidates (28 %).

This distribution (fig. 2) has a strong divergence with a power distribution of vertices in the social graph (Pareto Rule). This difference has a trivial explanation as the social network Facebook is built with an additional constraint that does not exist in the real society, that is the restriction of the vertices valences of the graph to 5,000 links. The discrepancy in the orientation of the graph to the total number of friends has one explanation: a large number of friends is not an indicator of the source of information in the society, as all 5,000 human connections can be inbound.

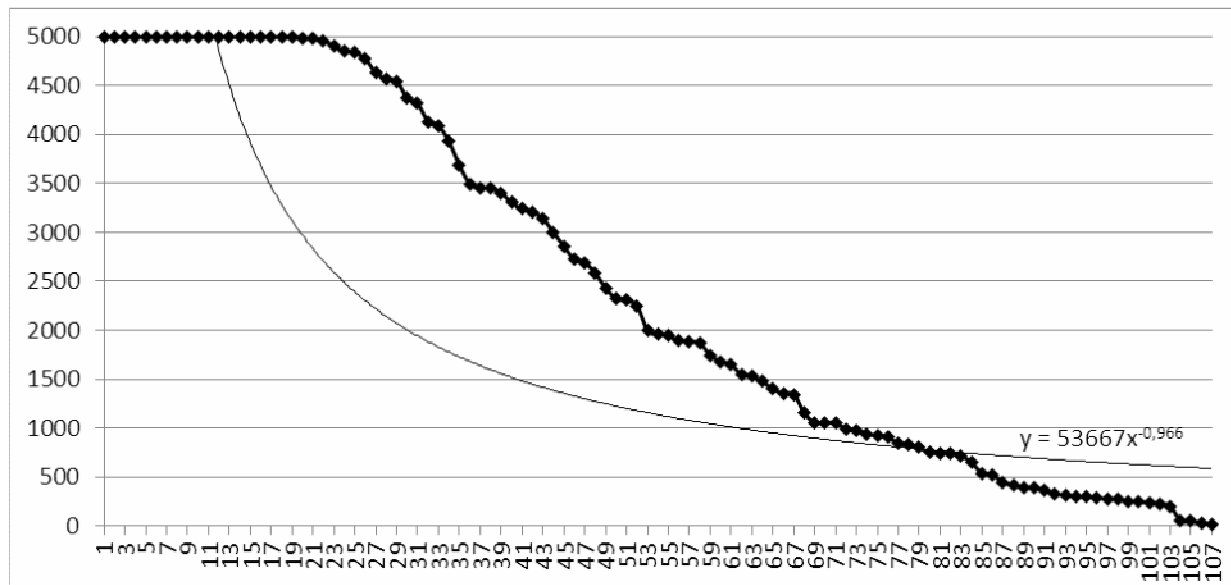


Fig. 2. The social graph distribution according to the number of friends in friend lists

The authors of this research had to complicate the task, and to investigate the valence of each vertex for a bounded social graph including only the participants of the experiment. This was done by comparative analysis of all the friend-lists of participants in the experiment. After that, all the vertices were located in the descending order valence, and communication between vertices were visualized, – fig. 3:

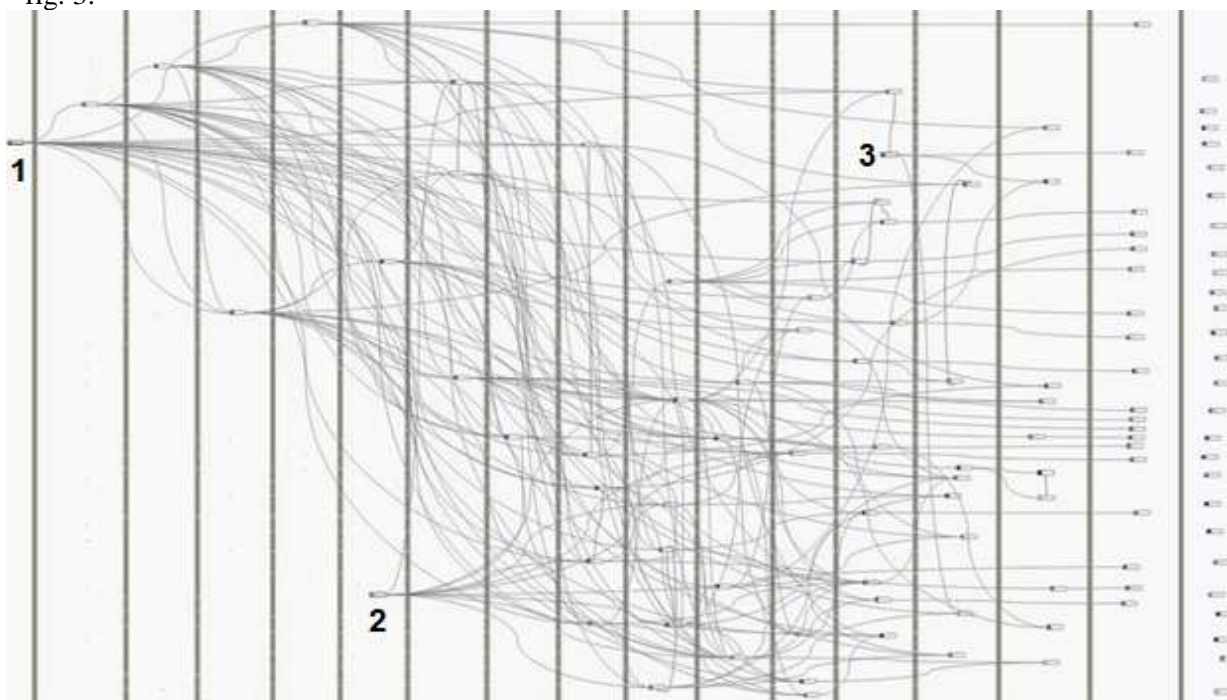


Fig. 3. The social graph with descending vertices valences orientation

The comparison of the resulting graph with the previously constructed oriented graph showed almost a complete matching, that is up to 88 % and it allowed us to draw some conclusions:

1. Social graph orientation actually coincides with the valences distribution of graph vertices.
2. Graph arrangement according to the vertices valences (fig. 3.) is more informative than the “simple” arrangement by the number of friends in friend-list, or even by the number of common friends. This arrangement allows to identify at once the clear “opinion leaders”, who are the primary sources of information dissemination. “Opinion leaders” are labeled as vertices 1, 2, 3 in fig. 3. These vertices differ

from the others in the graph by a complete lack of relations with vertices having more than their valence. At the same time all the three vertices have different indices of valence. This fact agrees well with the field observations: there exist the “leaders of opinions” not only in large, but even in small groups.

3. the vertices distribution in the graph by valence (degree of vertex) correlates with the clustering coefficient of this vertex, that is the ratio of closed triplets in the investigated graph to which this vertex bounds, to the open triplets – fig. 4:

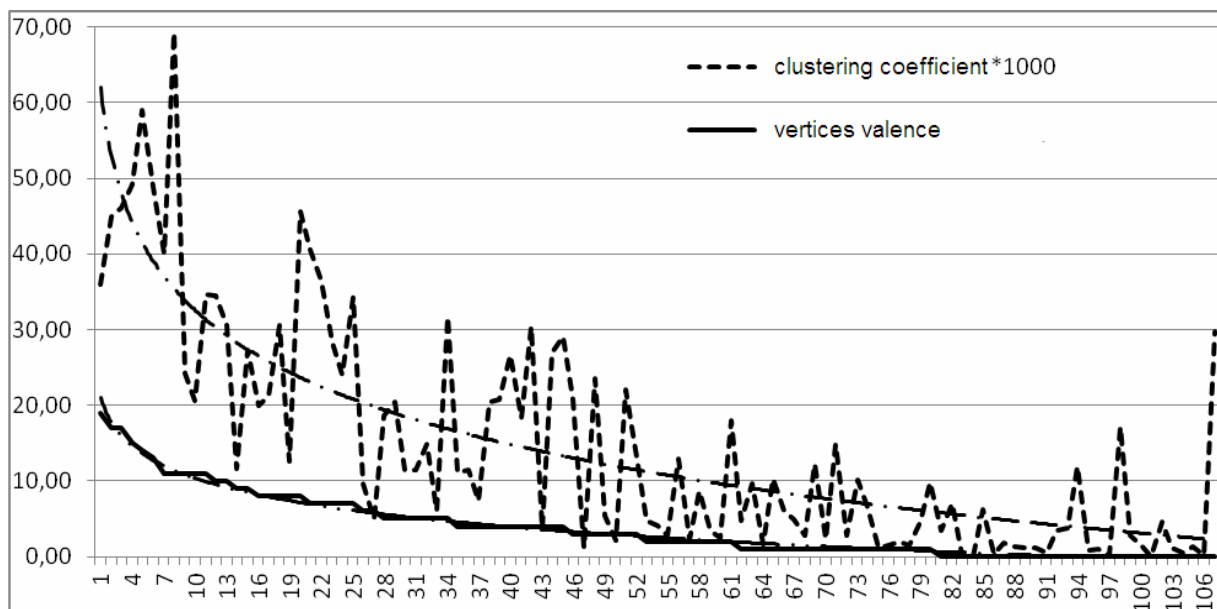


Fig. 4 The correlation of vertices graph distribution with clustering coefficient

In addition, it can be seen from figure 3 that 80 (74,77 %) people in the social graph are linked in a connectivity component, and 27 (by 25,23 %) components are separated from each other. If we remove all the vertices with the highest valences (>10) from the graph in figure 3, then the connectivity of the component will lose only 2,7 % of the remaining vertices. In other words, if “removing” the vertices of the social graph breaks down into individual clicks – this is one of the signs that “leader of opinions” is in the vertex of the social graph.

Correlated distribution coefficient of clustering and the distribution of valences of the vertices allow us to analyze social graphs to identify centers of clustering. While a large scatter (fig. 4) does not allow to carry out individual analysis of the graph vertices (people) in terms of clustering, but allows us to define how close a group of people is to the center of clustering.

Thus, the most important is the analysis of an index of the valence of the vertices in the task of identifying the orientation of the social graph and associated with the task application for individual analysis of the users of social networks and the identification of “opinion leaders”. In the definition of valences it is necessary to focus not only on the measure of friend-list and on the indicator of the vertices valence in a subgraph of the network with connectivity component.

3. Not oriented links in the society (method of detecting a critical situation)

The construction of a random graph in descending order of valence in fig. 3 as well as in the Watts–Strogatz model reproduces the cyclic closure of the social graph in the central part in the form of equitable relations, that is the relations between vertices with the same valence, and for this reason having no a pronounced orientation.

Fig. 5 shows the distribution of equitable relations to the parts of the graph with different valence.

It is clear from fig. 5 that the number of equal relations in the society gradually increases with increasing degree of vertices reaches a certain maximum and begins to decrease with increasing degrees of vertices. It is shown in fig. 6, that a gradual increase and a gradual decline show the ratio of the number of equitable relations to the number of vertices in the parts of graph with a certain valence of the vertices more evidently.

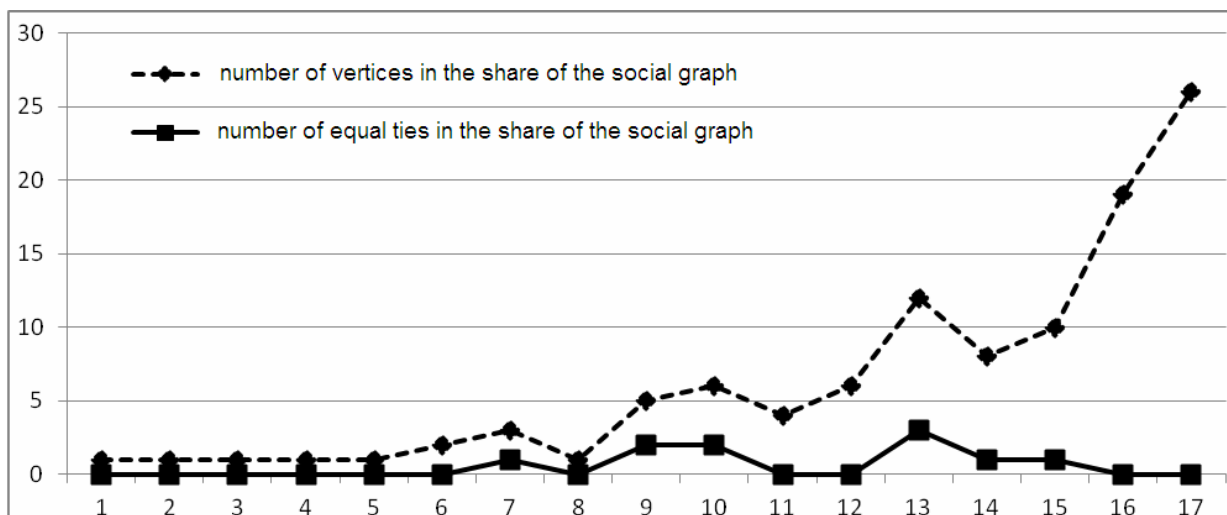


Fig. 5. The correlation between the quantity of equitable relations and the vertex valence in the parts of the graph

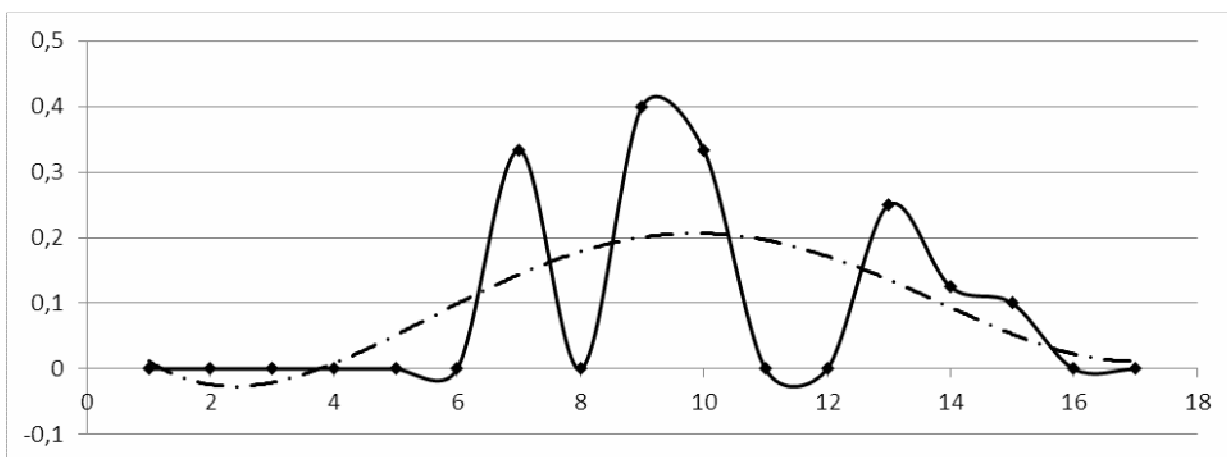


Fig. 6. The distribution of the ratio index of the number of equitable relations to the number of vertices in the parts of graph with a certain valence of the vertices

This distribution shows the evolution stages of relations formation of vertices in the social graph:

- 1) initially any vertex in the society has only some incoming links;
- 2) gradually the increasing of the number of incoming links process is growing;
- 3) the first equitable relations are forming;
- 4) the process of increasing equitable relations is growing;
- 5) the growth of the number of equitable relations changes to the growth of the number of outgoing relations;
- 6) equitable relations are disappearing, incoming relations are reducing, the growth of outgoing relations is taking place;
- 7) missing outgoing relations are missing completely, only the outgoing relations are in the vertex;
- 8) the number of outgoing relations is decreasing;
- 9) vertex is becoming completely isolated from the society.

The same process can capture the separate clusters of the society: new countries, new languages, isolated subcultures are supposed to be born through “the spin-off from the giant components”. Precisely this process of “spin-off” the subgraph from the connectivity components can be identified due to mass “breaking” relations, that is, by the reduction of the clustering coefficient and by the growth of the sparseness of the social graph. Such “spin-off” processes describe the critical situation and, in fact, they are their main manifestation. Checking the relation chains tearing is the fact of emergence of a critical situation, at the stage when nothing can be changed, because the society breaks down into separate

components. Knowing the evolutionary path of building the relations of vertices and parts of the society it is possible to track deviations from its normal development.

In this case, the most important role of the indicator in the pre-critical situation play just equitable relations between vertices with the same valence orientation, which are weakly expressed. If the experimentally determined norm of such equitable relations in the society does not exceed $\approx 5\%$, a significant excess of this norm may indicate the loss of the society equilibrium and the emergence of pre-critical situation. This occurs from the fact that equitable relations do not have a fixed orientation and for that reason the growth of their number deprives the social graph of the strict orientation and leads to the loss of controllability over the parts of the society with the highest index of valence.

References

1. Barabási L.-A., Albert R. Emergence of scaling in random networks. *Science*, 1999, Vol. 286, Issue 5439, pp. 509–512. DOI: 10.1126/science.286.5439.509
2. Barabási L.-A., Albert R., Jeong H. Scale-free characteristics of random networks: the topology of the world-wide web. *Physica A: Statistical Mechanics and its Applications*, 2000, Vol. 281, Issues 1–4, pp. 69–77. DOI: 10.1016/S0378-4371(00)00018-2
3. Bollobás B., Riordan O. Mathematical results on scale-free random graphs. *Handbook of Graphs and Networks: From the Genome to the Internet*, Weinheim: Wiley-VCH, 2003. pp. 1–34. DOI: 10.1002/3527602755.ch1
4. Stoimenow A. Enumeration of chord diagrams and an upper bound for Vassiliev invariants. *Journal of Knot Theory and Its Ramifications*, 1998, Vol. 7, Issue 1, pp. 93–114. DOI: 10.1142/S0218216598000073
5. Bollobás B., Riordan O. The diameter of a scale-free random graph. *Combinatorica*, 2004, Vol. 24, Issue 1, pp. 5–34. DOI: 10.1007/s00493-004-0002-2
6. Bollobás B., Riordan O., Spencer J., Tusnády G. The degree sequence of a scale-free random graph process. *Random Structures Algorithms*, 2001, Vol. 18, no. 3, pp. 279–290. DOI: 10.1002/rsa.1009
7. Kumar R., Raghavan P., Rajagopalan S., Sivakumar D., Tomkins A., Upfal E. Stochastic models for the web graph. *Proc. 41st Annual Symposium on Foundations of Computer Science*, 2000. DOI: 10.1109/SFCS.2000.892065
8. Kukushkin A.A., Bogomolov A.V., Ushakov I.B. Mathematical providing of estimation of material systems state. *Informacionnye tehnologii* (Information technologies), 2004, no. 7 (appendix), 32 p.
9. Bogomolov A.V., Gridin L.A., Kukushkin Yu.A., Ushakov I.B. *Diagnostika sostoyaniya cheloveka: matematicheskie podkhody* (Diagnosis of the human condition: mathematical approaches). Moscow, Meditsina Publ., 2003, 464 p.
10. Newman M.E.J., Watts D.J. Scaling and percolation in the small-world network model. *Phys. Rev. E*, 1999, Vol. 60, pp. 7332–7342. (arXiv: 9904419v2) DOI: 10.1103/physreve.60.7332
11. Garin E.V., Meshcheryakov R.V. Social graph of the society. *News of higher educational institutions. Sociology. Economy. Policy*, 2015, no. 4(47), pp. 11–14. (in Russ.).
12. Garin Y., Meshcheryakov R. Model of Socium Governance. *Problems of Management Theory and Practice – an international journal of the Research Institute of Management Problems*, 2017, no. 1, pp. 105–121. (in Russ.).

Received September 16, 2017

УДК 519.173.5, 519.175.4

DOI: 10.14529/mmph170401

МЕТОД ОПРЕДЕЛЕНИЯ ОРИЕНТАЦИИ СОЦИАЛЬНОГО ГРАФА ПО АНАЛИЗУ ВАЛЕНТНОСТИ ВЕРШИН В КОМПОНЕНТЕ СВЯЗНОСТИ

Е.В. Гарин, Р.В. Мещеряков

Томский университет систем управления и радиоэлектроники, г. Томск,
Российская Федерация
E-mail: mrv@tusur.ru

Данная работа является продолжением работ авторов по проблематике моделирования строения социума методом генерации случайных графов. В своих предыдущих работах авторы экспериментально доказали, что социальный граф имеет строгую ориентацию направления распространения информации от долей графа с высокой плотностью связей к долям графа, имеющим более низкую плотность связей. Авторами предлагался метод определения ориентации социального графа путем исследования соотношения количества исходящих связей к входящим. Этот метод нашел применение в анализе почтовых рассылок, но оказался практически не применим в анализе социальных сетей, так как подсчет входящих и исходящих связей требовал составления возмещённых социальных графов – затратной по времени и вычислительным ресурсам задачи, не имеющей на сегодняшний день решения за полиномиальное время.

Так как вопрос анализа социальных сетей представляет собой не только строго научный, но и практический интерес, авторами была разработана и апробирована методика определения ориентации социальных графов методом анализа числовых характеристик графа. Новый метод не требует детального анализа переписки пользователей социальных сетей, а оперирует открытой информацией пользователей – списком друзей (френд-лист).

Конечной целью данной работы авторов является разработка простых и действенных методов анализа социальных сетей на предмет выявления «лидеров мнения», путей распространения информации, в том числе пропагандирующих девиационные и зависимые формы поведения, выявления антисистем и отдельных закрытых сетевых сообществ, общего мониторинга состояния социальных систем.

Ключевые слова: социальный граф; случайные графы; числовые характеристики графов; ориентация социального графа.

Литература

1. Baraba'si, L.-A. Emergence of scaling in random networks / L.-A. Baraba'si, R. Albert // Science. – 1999. – Vol. 286, Issue 5439. – P. 509–512.
2. Baraba'si, L.-A. Scale-free characteristics of random networks: the topology of the world-wide web / L.-A. Baraba'si, R. Albert, H. Jeong // Physica A: Statistical Mechanics and its Applications. – 2000. – Vol. 281, Issues 1–4. – P. 69–77.
3. Bolloba's, B. Mathematical results on scale-free random graphs / B. Bolloba's, O. Riordan // Handbook of Graphs and Networks: From the Genome to the Internet. – Weinheim: Wiley-VCH, 2003. – P. 1–34.
4. Stoimenow, A. Enumeration of chord diagrams and an upper bound for Vassiliev invariants / A. Stoimenow // Journal of Knot Theory and Its Ramifications. – 1998. – Vol. 7, Issue 1. – P. 93–114.
5. Bolloba's, B. The diameter of a scale-free random graph / B. Bolloba's, O. Riordan // Combinatorica. – 2004. – Vol. 24, Issue 1. – P. 5–34.
6. The degree sequence of a scale-free random graph process / B. Bolloba's, O. Riordan, J. Spencer, G. Tusna'dy // Random Structures Algorithms. – 2001. – Vol. 18, № 3. – P. 279–290.
7. Stochastic models for the web graph / R. Kumar, P. Raghavan, S. Rajagopalan *et al.* // Proc. 41st Symposium on Foundations of Computer Science. – 2000.

8. Кукушкин, Ю.А. Математическое обеспечение оценивания состояния материальных систем / Ю.А. Кукушкин, А.В. Богомолов, И.Б. Ушаков // Информационные технологии. – 2004. – № 7 (приложение). – 32 с.

9. Диагностика состояния человека: математические подходы / А.В. Богомолов, Л.А. Гридин, Ю.А. Кукушкин, И.Б. Ушаков. – М.: Медицина, 2003. – 464 с.

10. Newman, M.E.J. Scaling and percolation in the small-world network model / M.E.J. Newman, D.J. Watts // Phys. Rev. E. – 1999. – Vol. 60. – P. 7332–7342. (arXiv: 9904419v2).

11. Гарин, Е.В. Социальный граф общества / Е.В. Гарин, Р.В. Мещеряков // Известия высших учебных заведений. Социология. Экономика. Политика. – 2015. – № 4(47). – С. 11–14.

12. Гарин, Е.В. Модель управления социумом / Е.В. Гарин, Р.В. Мещеряков // «Проблемы теории и практики управления» – международный журнал научно-исследовательского института проблем управления: раздел «Управление социальным развитием». – 2017. – № 1. – С. 105–121.

Поступила в редакцию 16 сентября 2017 г.