Applied problems of nonlinear oscillation and wave theory

Official English translation

Mathematical modelling of the network of professional interactions

E. N. Pitsik, M. V. Goremyko, V. V. Makarov, A. E. Hramov

Yuri Gagarin State Technical University of Saratov
77, Politechnicheskaya street, Saratov 410054, Russia
E-mail: pitsikelena@gmail.com, gormv67@sstu.ru,
vladmak404@gmail.com, hramovae@gmail.com
Received 2.09.2017, revised 28.10.2017

Description of real-world systems of interacting units by the means of network model is an effective method of research both in macro- and microscale. In addition, using the simple one-layer networks with one type of connections between the nodes when describing real-world networks is inefficiently because of their complex structural and dynamical nature. Besides, presence of similar features in real networks that are fundamentally different by their nature provided a wide spread of proposed model in many fields of science for the acquisition of new fundamental knowledge about functioning of the real network structures. For this reason the object of this article is modelling of multiplex network build on the basis of real data about professional interactions in world-wide musical community. The changes in characteristics in in proposed model reflects structural and dynamical features of real network, such as scale-free connection structure and clusters formation. Results obtained for multiplex network shows that after uniting the isolated systems their topologies undergo noticeable changes. In particular, significant changes in centrality values and in cluster formation inside the network were obtained. Besides, the correlations between the characteristics and dynamics of these correlations in process of uniting the isolated systems in general network. Obtained results confirm the effectiveness of multiplex network model for studying structural and dynamical processes of many real systems.

Key words: complex network, multiplex network, mathematical modelling, social system.

https://doi.org/10.18500/0869-6632-2018-26-1-21-32

References: Pitsik E.N., Goremyko M.V., Makarov V.V., Hramov A.E. Mathematical modelling of the network of professional interactions. *Izvestiya VUZ, Applied Nonlinear Dynamics,* 2018, vol. 26, iss. 1, pp. 21–32. https://doi.org/10.18500/0869-6632-2018-26-1-21-32

© *Pitsik E.N., Goremyko M.V., Makarov V.V., Hramov A.E.*

Interaction

Many phenomena characterized by interaction of a lot of elements coupled in any way, may be described with the help of network models. There is an unlimited number of such systems in macro- and microscales: brain neural networks [1–3], complicated biological systems in separate cells, as well as in the whole organism [4, 5], transport nets [6, 7], networks of wireless mobile communications [8], social networks [9], systems of science works citation [10–12], computer nets [13], etc. The fact deserving special attention is that in spite of global difference many real networks have some common fundamental properties: freely scaled structure of elements coupling and the tendency of nodes to unite into clusters [14, 15]. The first property means that a net contains some nodes (i.e. hubs or concentrators), whose degree (a number of coupling) far exceeds the degree of all the other nodes. This can be explained by the additivity effect in real systems, i.e. the ability of high-degree nodes to attract more coupling than low-degree nodes.

The second property is connected with dynamical processes, which influence on the net structure, and with homeostasis principle [16], which is known of social networks as Dunbar's number, i.e. a suggested cognitive limit to the number of people with whom one can maintain stable social relationships [17–19]. All these processes lead to the association of social network members by different common features. This results in appearance of cluster structures (communities) in the network. The nodes inside the clusters are connected closer than with the nodes of some other clusters (communities).

The description of real systems with the help of traditional single-layer network (in which the nodes are coupled by the same type of connections) sometimes can't consider all the aspects of elements interaction [20]. The examples of such connections are different kinds of mutual relations between the network members: friendship, professional collaboration, etc. If we try to model such processes as spreading rumors [21–24] or epidemics [25], considering the social system as a traditional single-layer network, then the obtained results will be very different from real life. Thus the researches, who describe real systems, usually apply the model of multiplex network which implies that the nodes have a possibility to interact with each other on different levels. Every level can be considered as separate layer of multiplex network, which has an identical set of nodes, but its own unique structure. Herewith every node has not only the relations inside its layer, but also the connection with its «prefiguration» on other layers. The scheme of such network is shown in Fig. 1.

In this work we are modelling the complicated social network based in statistical data concerning professional activity of musicians. The analysis of characteristics of this network allows to highlight the degree of influence of different types of relations upon the peculiarities of social network, which are principally important for studying the mechanisms of distribution of different processes in society and understanding of regularities of these processes: spread of diseases; public opinion formation; communities formation, etc. Thus the understanding of structural and dynamical peculiarities of real systems is achieved.

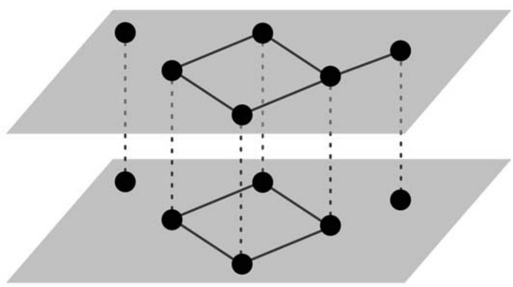

Fig. 1. Schematic illustration of two-layer multiplex network structure

The researched data

As a private example of social interaction we chose the all-world musicians community. The statistical data was taken from Allmusic.com [26], the database containing the largest musical archive with complete information about musical genres, composes and performers. The obtained data had two sets of parameters. The first of them reflected the performers' genre attachment, where two elements were coupled if the corresponding musicians had performed the compositions belonging to the same genres. The second data set was the list of professional connections between individuals, i.e. if the musicians had worked on one or more albums together. The initial data set concerning the genres of the musicians contained 32377 nodes and 117621 connections, the data set concerning collaboration contained 34724 nodes and 123082 connections. The further data processing involved the selection of the network main component, i.e. the greatest set of related nodes [27]. It was taken into account that every node could belong to the main component (1) of the both layers, (2) to anyone of the layers, or (3) to none of them (Fig. 1). In this work we considered only the nodes belonging to (1) and (2) groups. Deleting all the other nodes from the sets we selected the main coupled component which consisted of $N = 8279$ nodes in every layer, and worked with it. This approach is typical for studying real network structures, including social ones. It allows to avoid distortions of system characteristics caused by the large number of nodes without connections, without loosing information about basic observed tendencies.

Mathematical model

For the purposes of study we have constructed a model of multi-layer network. The first layer is based on data of musicians genre (further G – genre), the second layer represents the collaboration between them (further C – collaboration). The distributions of nodes degree (dotted lines in Fig. 2) indicate that the network structure has the property of free scaling: the dependencies may be approximated by the law $P(D) = D^{-\gamma}$, where D is the node degree, $\gamma \approx 2.4$. (In Fig. 2 the approximation is shown by lines). Thus the properties of the constructed network agree with really existing systems, which is an important criterium of authenticity of the results, which are obtained with the help of this system.

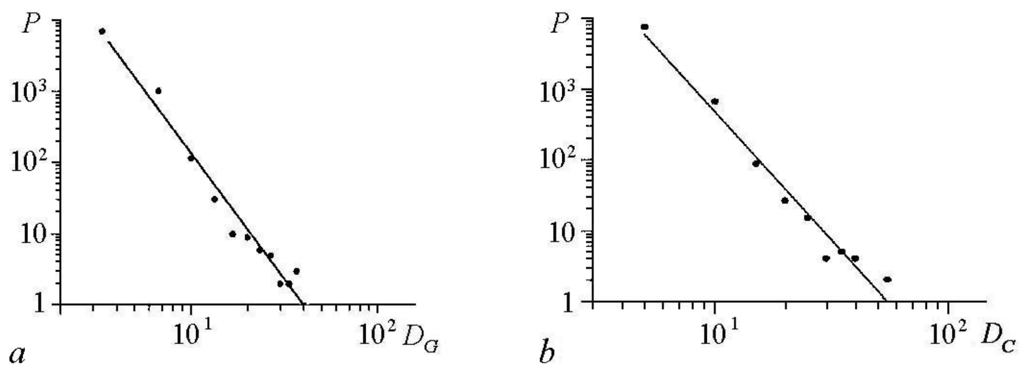

Fig. 2. Degree distribution for the networks G (*a*) and C (*b*) in logarithmic scale

To study the network structure peculiarities and their mutual influence, we have calculated several characteristics for unconnected layers and in presence of inter-layer connections. These characteristics are: firstly, the nodes degrees for both cases, i.e., the quantity of the connections of every node. Secondly, an important characteristic is the betweenness centrality B_k, which represents the node loading and is equal to the quantity of the shortest ways from every node to all the other nodes passing through the considered node [28].

$$B_k = \sum_{i \neq k \neq j} \frac{\sigma_{ij}(k)}{\sigma_{ij}}, \tag{1}$$

where σ_{ij} is the main quantity of the ways from node i to node j, and $\sigma_{ij}(k)$ is the quantity of the ways passing through the considered node k.

One of the interesting network properties is the ability of its nodes to form groups – clusters or modules. Detection of such structures in the network has an important practical meaning in different areas of science, including the possibility to track the formation of groups and communities in real social networks [29–31]. In order to characterize the degree of attachment of a node to tightly bounded group, we use local clustering coefficient.

$$C_i = \frac{t_i}{q_i(q_i - 1)/2}, \tag{2}$$

where i is the considered node, q_i is the number of nearest neighbours of this node, t_i is the number of connections between the neighbours. This parameter characterizes the possibility that two nearest neighbours of the considered nodes are also the nearest neighbours for each other. The number t_i is interpreted as total number of triangles attached to the node i, and $q_i(q_i - 1)/2$ is the maximum possible number of triangles. The clustering coefficient is equal to zero if the neighbouring nodes can't have such connection – the example of such structure is hierarchical tree. If all the neighbours of the considered element have connections, then the clustering coefficient is equal to one.

However, to determine the number of mesoscopic structures in every layer and there characteristics with the appearance of inter-layer connections, we have calculated a more complex parameter, namely, network modularity [32]

$$M = \frac{1}{2m} \sum_{ij} \left(A_{ij} - \gamma \frac{k_i k_j}{2m} \right) \delta_{g_i g_j}, \tag{3}$$

where m is the total network connections; A_{ij} is the adjacency matrix; $\delta_{g_i g_j}$ denotes the Kronecker delta and is the indicator of equality of the groups g_i and g_j, to which the i and j elements belong; k_i is the i node degree and γ is the parameter denoting the value of the communities and their number.

As one more characteristic we considered the E_i eigenvector centrality – another type of centrality, defining the influence of the node upon the network [33]. Each node is given an estimate based on the assumption that its connections with high-centrality nodes contribute into its estimate more than the connections with low-centrality nodes. Thus, the eigenvector centrality depends not only from the number of node connections but also from the centrality of the connected nodes. This allows to pick out a small group of nodes which has a significant impact on the network as a whole. At the same time, we exclude from this group the nodes, which have a big number of connections, but whose influence on the network does not extend beyond their own neighbors, i.e., isolated clusters with no impact in the scale of whole network. Thus, the eigenvector centrality represents the eigenvector corresponding to the largest eigenvalue of the adjacency matrix

$$E_i = \frac{1}{\lambda} \sum_{j \in M} w_{ij} E_j. \quad (4)$$

Looking ahead it's worth noting that because of their specificity, the values of this parameter will be very different for the cases of connected and isolated layers.

Results of the analysis of the model network of professional relations

During the study we have analysed correlation relationships between all the parameters of the net. In this section we discuss only those dependencies, which most clearly reflect the structural features of the system.

The characteristics shown in the previous section have been firstly calculated separately for each network and then were united into a multiplex network for each layer of the derived structure. This approach allows to determine how the layers within one complex network affect on structural and dynamical characteristics of each other.

First of all we have analysed the dependencies between the nodes characteristics inside the isolated layers, which are shown in Fig. 3. In the result of calculations it has been found that the nodes betweenness centrality is in inverse correlation with the clustering coefficient as in the network G, so as in the network C (see Fig. 3, *a, b*). This indicates that if there are many shortest paths passing through the node, then the node has a low clustering coefficient, i.e., is beyond the structure cluster. We can observe similar

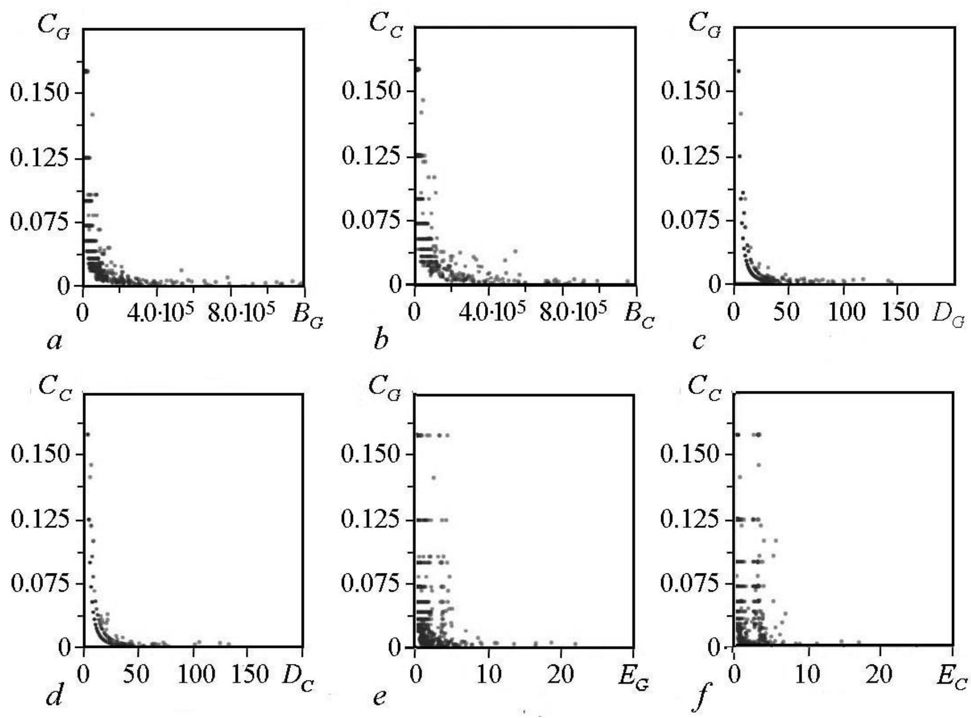

Fig. 3. Correlations of nodes characteristics for layer G (*a, c, e*) and C (*b, d, f*). Figures show the correlations between clustering coefficient and betweenness centrality (*a, b*), node degree (*c, d*) and eigenvector centrality (*e, f*)

dependance between the node degree and its clustering coefficient (see Fig. 3, *c, d*). This indicates that the cluster elements has a limited number of connections with their neighbours.

At the same time the nodes having larger betweeness centrality and situated beyond the clusters, can be considered as structure hubs. From the point of realistic view we have two variants. At first, the more genres a musician performs, the less of his neighbours play in one genre. This conclusion is logical, because there are too mane different genres among the musicians, connected with the considered musician. At second, the more active is the communication of the considered performer with the other ones, the less is the communication of his neighbours with each other.

The inverse correlation between eigenvector centrality and clustering coefficient (see Fig. 3, *e, f*) indicates the presence of a large number of clusters, having a limited number of inside links and showing a strong influence on the network because of large eigenvector centrality of the included nodes. This result points to a heterogeneous structure of a network, in which nodes with a high clustering coefficient are outside the central cluster.

Now let's discuss the characteristics of a network after combining it into a multi-layer network P. Here the network of genres G is considered as the first layer and the collaboration network C – as the second layer of network P.

In Fig. 4 one can see the changing of mediation centrality and eigenvector centrality within the transition from isolated layer (y-axis) to multilayer network (x-axis). We can

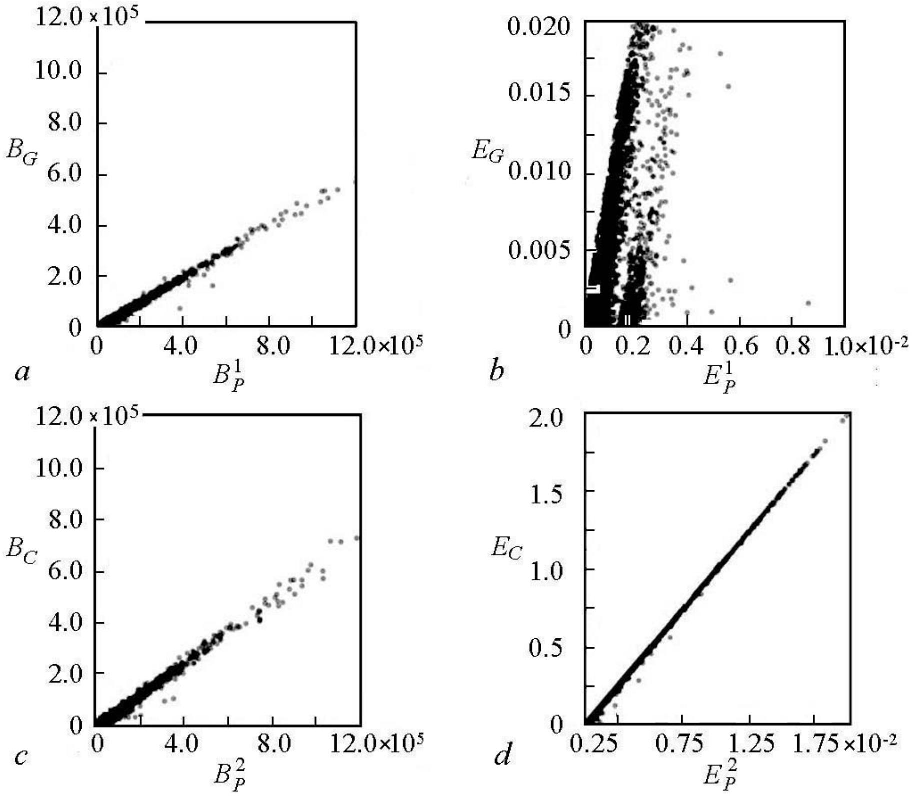

Fig. 4. Correlation between centralities of the networks G (*a, b*) and C (*c, d*) before and after uniting of the network; B_p^1, B_p^2, E_p^1 и E_p^2 – centralities of the nodes of first and second layer of multiplex network P, respectively

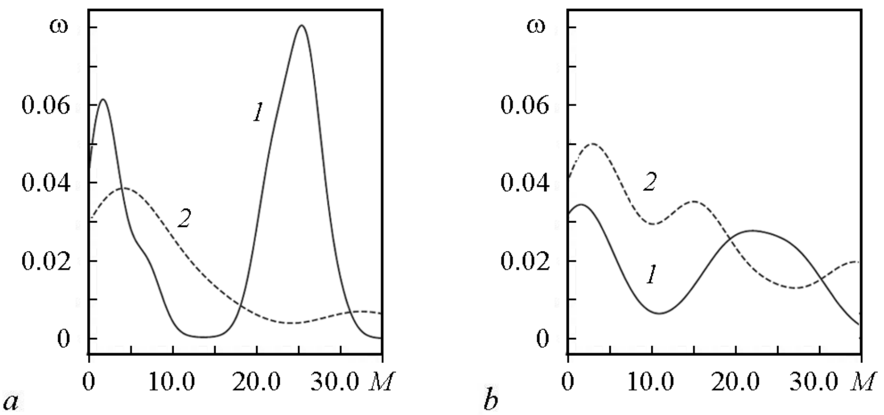

Fig. 5. Probability density for modularity parameter before uniting in general network (*a*) and after (*b*). *1* – modularity of the network G, *2* – modularity of the network C. X-axis and Y-axis are marked with calculated communities and the values of probability density function, respectively

mark that for both layers the mediation centrality doesn't demonstrate any changes within the transition to multiplex network. (see Fig. 4, *a, b*). It indicates that, in spite of very different criteria of layers construction, the same set of nodes has the largest mediation centrality. Nevertheless, if we look at eigenvector centrality (see Fig. 4, *c, d*), we find out that after the transition to multilayer network the value of this characteristic greatly reduced for nodes that are in the genre layer. It is mostly connected with the specificity of calculation of this characteristic. In the case of multiplex network such dynamics indicates optimal structure of collaboration layer.

In Fig. 5 one can see the curves for probability density $\omega(M)$ for modularity parameter M. The probability density function means that the probability that the node q belongs to the community m is equal to the square of the figure bordered by the interval $[m-1, m]$ and the curve $\omega(M)$. Thus, the more is y-coordinate in the point, the more numerous is the corresponding community.

The graphs show noticeable changes in the formation of communities within networks: whereas in the isolated case, strong heterogeneity of structure elements is clearly visible, the appearance of interlayer connections leads to the formation of connections between clusters intersecting at different layers of a multiplex network. As a consequence, the network structure becomes more homogeneous.

Conclusion

In our study, we have built a multiplex network model on the basis of real statistical social data concerning professional interaction of musicians. The main structural characteristics have been calculated both for isolated networks and for each layer of a multiplex network. For two isolated networks certain similarities in the obtained dependencies between the characteristics have been found, which indicates that there are general regularities in the processes of network topologies formation.

The results obtained for multiplex network, indicate that when isolated networks are combined, the characteristics of their nodes undergo noticeable changes. In particular, we have carried out the changes in the nodes centrality values and in the formation of communities within the network. These results confirm the efficiency of using the multiplex network model for studying structural-dynamic processes in many real-world systems.

The work is supported by the grant of the President of Russian Federation for leading science schools (project НШ-2737.2018.2).

References

1. Sporns O., Chialvo D.R., Hilgetag C.C. Organization, development and function of complex brain networks. *Trends Cogn. Sci.*, 2004, vol. 8, pp. 418–425.
2. Maksimenko V.A., Lüttjohann A., Makarov V.V., Goremyko M.V., Koronovskii A.A., Nedaivozov V., Runnova A.E., van Luijtelaar G., Hramov A.E., Boccaletti S. Macroscopic and microscopic spectral properties of brain networks during local and global synchronization. *Phys. Rev. E.*, 2017, vol. 96, no. 1.
3. Kirsanov D.V., Nedaivozov V.O., Makarov V.V., Goremyko M.V., Hramov A.E. Study of pattern formation in multilayer adaptive network of phase oscillators in application to brain dynamics analysis. *Proc. SPIE,* 2017, vol. 10337, pp. 103370Z-1–7.
4. Kai S., Gonçalves J.P., Larminie C., Pržuli N. Predicting disease associations via biological network analysis. *BMC Bioinformatics,* 2014, vol. 15.
5. Sharan R. and Ideker T. Modeling cellular machinery through biological network comparison. *Nat. Biotechnol.*, 2006, vol. 24, pp. 427–433.
6. Ma X., Yu H., Wang Y., Wang Y. Large-Scale transportation network congestion evolution prediction using deep learning theory. *PLoS ONE,* 2015, vol. 10, no. 3.
7. He X., Liu X.H. Modeling the day-to-day traffic evolution process after an unexpected network disruption. *Transp. Res. Part B Methodol.*, 2012, vol. 46, no. 1, pp. 50–71.
8. Wang L., Kuo G.S. Modeling for network selection in heterogeneous wireless networks. A tutorial. *IEEE Commun. Surveys Tuts.*, 2013, vol. 15, no. 1, pp. 271–292.
9. Yang S., Yang X., Zhang C., Spyrou E. Using social network theory for modeling human mobility. *IEEE Netw.*, 2010, vol. 24, no. 5, pp. 6–13.
10. Dawson S., Gaševič D., Siemens G., Joksimovic S. Current State and Future Trends: A Citation Network Analysis of the Learning Analytics Field, ACM, 2014, pp. 231–240.
11. Érdi P., Makovi K., Somogyvári Z., Strandburg K., Volf P., Zalányi L. Prediction of emerging technologies based on analysis of the US patent citation network. *Scientometrics,* 2013, vol. 95, pp. 225.
12. Ta-Shun C., Hsin-Yu S. Using social network theory for modeling human mobility. *IEEE Netw.*, 2010, vol. 24, no. 5, pp. 6–13.
13. Shakkottai S., Srikant R. Network optimization and control. *Found. Trends. Network.*, 2008, vol. 2, no. 3, pp. 271–379.
14. Battiston F., Nicosia V., Latora V. Structural measures for multiplex networks. *Phys. Rev. E,* 2014, vol. 89, no. 3.
15. Makarov V.V., Koronovskii A.A., Maksimenko V.A., Hramov A.E., Moskalenko O.I., Buldú J.M., Boccaletti S. Emergence of a multilayer structure in adaptive networks of phase oscillators. *Chaos, Solitons, Fractals,* 2016, vol. 84, pp. 23–30.

16. Turrigiano G.G., Nelson S.B. Homeostatic plasticity in the developing nervous system. *Nat. Rev. Neurosci.*, 2004, Vol. 5, pp. 97–107.
17. Arnaboldi V., Conti M., Passarella A., Pezzoni F. Ego networks in Twitter: An experimental analysis. *IEEE CCW,* 2013, pp. 229–234.
18. Gonçalves B., Perra N., Vespignani A. Modeling users' activity on twitter networks: Validation of Dunbar's number. *PLoS ONE,* 2011, vol. 6, no. 8, pp. 1–5.
19. de Ruiter J., Weston G., Lyon S.M. Dunbar's Number: Group Size and Brain Physiology in Humans Reexamined. *Am Anthropol.*, 2011, vol. 113, no. 4, pp. 557–568.
20. Menichetti G., Remondini D., Panzarasa P., Mondragón R.J., Bianconi G. Weighted multiplex networks. *PLoS ONE,* 2014, vol. 9, no. 6, pp. 1–8.
21. Ru-Ya T., Xue-Fu Z., Yi-Jun L. SSIC model: A multi-layer model for intervention of online rumors spreading. *Physica A,* 2015, vol. 427, pp. 181–191.
22. Zhao L., Wang J., Chen Y., Wang Q., Cheng J., Cui H. SIHR rumor spreading model in social networks. *Physica A: Statistical Menics and its Applications,* 2012, vol. 391, no. 7, pp. 2444–2453.
23. Hu Q., Gao Y., Ma P., Yin Y., Zhang Y., Xing C. A New Approach to Identify Influential Spreaders in Complex Networks / Web-Age Information Management: 14th International Conference. Springer Berlin Heidelberg, 2013, pp. 99–104.
24. Roshani F., Naimi Y. Effects of degree-biased transmission rate and nonlinear infectivity on rumor spreading in complex social networks. *Phys. Rev. E,* 2012, vol. 85, no. 3.
25. Buono C., Alvarez-Zuzek L.G., Macri P.A., Braunstein L.A. Epidemics in partially overlapped multiplex networks. *PLoS ONE,* 2014, vol. 9, no. 3, pp. 1–5.
26. Allmusic.com
http://www.allmusic.com/
27. Callaway D.S., Newman M.E.J., Strogatz S.H., Watts D.J. Network robustness and fragility: Percolation on random graphs. *Phys. Rev. Lett.*, 2000, vol. 85, no. 25, pp. 5468–5471.
28. Brandes U. A faster algorithm for betweenness centrality. *J. Math. Sociol.*, 2001, vol. 25, no. 3, pp. 163–177.
29. Papadopoulos S., Kompatsiaris Y., Vakali A., Spyridonos P. Community detection in Social Media. *Data Min Knowl Discov.*, 2012, vol. 24, no. 3, pp. 515–554.
30. Aggarwal C.C. An Introduction to Social Network Data Analytics. Social Network Data Analytics. Springer US, 2011, pp. 1–15.
31. Leskovec J., Lang Kevin J., Mahoney M. Empirical comparison of algorithms for network community detection. *Proceedings of the 19th International Conference on World Wide Web. ACM*, 2010, pp. 631–640.
32. Newman M.E.J. Equivalence between modularity optimization and maximum likelihood methods for community detection. *Phys. Rev. E,* 2016, vol. 94, no. 5.
33. Newman M.E.J. Mathematics of Networks. The New Palgrave Dictionary of Economics. Palgrave Macmillan, 2008.