

Сибирское отделение
Российской академии наук

ISSN 2073-0667

ПРОБЛЕМЫ ИНФОРМАТИКИ

ПРОБЛЕМЫ ИНФОРМАТИКИ № 3 2021



ТЕОРЕТИЧЕСКАЯ И СИСТЕМНАЯ ИНФОРМАТИКА
ПРИКЛАДНЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

3
—
2021

ПРОБЛЕМЫ ИНФОРМАТИКИ № 3 (52) 2021 г.

Журнал выходит ежеквартально, издается с 2008 г.

Учредитель журнала — Институт вычислительной математики и математической геофизики СО РАН при поддержке Института информационных и вычислительных технологий МОН РК.

Редакционный совет

Председатель — акад. НАН РК М. Н. Калимолдаев,

Акад. РАН А. Л. Асеев, акад. АН РУ **Т. Ф. Бекмуратов** (Республика Узбекистан), проф.

В. А. Васенин, акад. РАН С. Н. Васильев, проф. В. М. Вишневский, акад. РАН С. С. Гончаров,

акад. РАН Ю. Л. Ершов, акад. РАО А. А. Кузнецов, акад. РАН Н. А. Кузнецов, акад. РАН

А. П. Кулешов, проф. А. Г. Марчук, проф. Б. Я. Рябко, проф. Н. А. Семенов, акад. РАН

И. А. Соколов, проф. А. Н. Сотников, чл.-кор. РАН Ю. А. Флеров, проф. П. С. Чубик.

Редколлегия

Главный редактор — проф. В. Э. Малышкин,

Д. Ж. Ахмед-Заки, А. Г. Вострецов, **В. П. Гергель**, Б. С. Гольдштейн, В. И. Гужов,

Ю. А. Загоруйко, С. Д. Каракозов, М. М. Каримов, В. Н. Касьянов, О. В. Кибис, В. В. Корнеев,

И. В. Котенко, Т. П. Любимова, А. И. Ляхов, М. А. Марченко, В. В. Окольников,

Б. В. Поллер, А. С. Родионов (зам. гл. редактора), М. А. Сонькин, В. В. Шахов (зам. гл.

редактора), М. С. Хайретдинов, Ph. D. Moonseong Kim (Korea), Prof. Dr.-Eng. V. D. Nguyen (Vietnam), Michele Pagano (Italy).

Редакция: отв. секретарь М. С. Делидович, системный администратор В. А. Перепелкин, верстка Д. В. Лазуткин, логист Л. В. Трофимова.

Адрес редакции, издателя: 630090, г. Новосибирск, просп. Академика Лаврентьева, д. 6, ИВМ и МГ СО РАН

тел. (383) 330-96-43; e-mail: problem-info@sscc.ru, <http://www.problem-info.sssc.ru>.

Журнал зарегистрирован в Федеральной службе по надзору в сфере массовых коммуникаций, связи и охраны культурного наследия. Свидетельство ПИ № ФС77-32088 от 27 мая 2008 г.

Журнал распространяется по подписке. Оформление подписки на сайте «Объединенного каталога „Пресса России“» https://www.ppressa-rf.ru/cat/1/edition/y_e69980/, подписной индекс 69980, и через интернет-магазин «Пресса по подписке»

https://www.akc.ru/itm/problemny_i-informatiki/. Цена свободная. Журнал распространяется на территории России.

Журнал включен в Перечень ведущих рецензируемых научных журналов, рекомендованных для публикаций Высшей аттестационной комиссией.

Все права авторов сохранены. Использование материалов журнала возможно только с разрешения редакции и авторов.

Отпечатано в типографии „АЛЕКСПРЕСС“ ИП Малыгин Алексей Михайлович.

Адрес: 630090, Новосибирск, пр-т Академика Лаврентьева, 6/1, оф. 104, тел. +7 (383) 217-43-46.

Формат 60 × 84 1/8. Усл. печ. л. 9,3. Уч.-изд. л. 6,81. Печать офсетная.

Тираж 60 экз. Заказ № 676. Подписано в печать 15.09.2021 г. Выход в свет 26.09.2021 г.

© Институт вычислительной математики и математической геофизики СО РАН, 2021

© Институт информационных и вычислительных технологий МОН РК, 2021

JOURNAL “PROBLEMS OF INFORMATICS”. No. 3 (52) 2021

Publisher: Institute of Computational Mathematics and Mathematical Geophysics of Siberian Branch of Russian Academy of Sciences with the support of the Institute of Information and Computing Technologies of the Ministry of Education of the Republic of Kazakhstan.

Editorial Council

Chairman Academician of the National Academy of Sciences of the Republic of Kazakhstan

M. N. Kalimoldayev

Full Member of the RAS A. L. Aseev, Full Member of the Academy of Sciences of Uzbekistan

T. F. Bekmuratov, Professor V. A. Vasenin, Full Member of RAS C. N. Vassilyev, Professor V. M. Vishnevsky, Full Member of RAS S. S. Goncharov, Full Member of RAS Yu. L. Ershov, Corr. Member of RAE A. A. Kuznetsov, Full Member of RAS N. A. Kuznetsov, Full Member of RAS A. P. Kuleshov, Professor A. G. Marchuk, Professor B. Y. Ryabko, Professor N. A. Semenov, Full Member of RAS I. A. Sokolov, Professor A. N. Sotnikov, Corr. Member RAS Y. A. Flerov, Full Member of NAS KR J. Sh. Sharshenaliev.

Editorial board

The Editor-in-Chief Professor V. E. Malyshkin

Associate Editors-in-Chief: A. S. Rodionov, V. V. Shakhov

D. Zh. Akhmed-Zaki, A. G. Vostretsov, V. P. Gergel, B. S. Goldstein, V. I. Guzhov, Y. A. Zagorulko, S. D. Karakozov, M. M. Karimov, V. N. Kasyanov, O. V. Kibis, V. V. Korneev, I. V. Kotenko, T. P. Lyubimova, A. I. Lyakhov, M. A. Marchenko, V. V. Okolnishnikov, B. V. Poller, Y. G. Soloveichik, M. A. Sonkin, M. S. Khairtdinov, Moonseong Kim (Korea), Van Duc Nguyen (Vietnam), Michele Pagano (Italy).

Editorial staff: Managing Editor M. S. Delidovich, System Administrator V. A. Perepelkin, Maker-up D. V. Lazutkin, Logistician L. V. Trofimova.

Address of the editorial office: 630090, pr. Lavrentieva, 6, Novosibirsk, Russia, Institute of Computational Mathematics and Mathematical Geophysics of SB RAS.

Phone: +7 (383) 330-96-43; e-mail: problem-info@sscc.ru, <http://www.problem-info.sccc.ru>.

The journal has been registered in accordance with Legislation of the Russian Federation. Certificate of Mass Media Registration: ПИ № ФС77-32088, of 27 May, 2008, ISSN 2073-0667. The journal is distributed in Russia.

The journal “Problems of Informatics” is in the List of Peer-Reviewed Scientific Journals for publication of scientific results of Ph.D. and Dr. of Sci. theses in three scientific specialties and corresponding scientific disciplines in which academic degrees are awarded:

05.13.11 — Mathematical Support and Software for Computers, Computing Complexes and Computer Networks,

05.13.17 — Theoretical foundations of Informatics.

05.13.18 — Mathematical Modeling, Numerical Methods and Software Complexes.

All rights reserved. The journal contents may only be used by the permission of editors and authors.

© Institute of Computational Mathematics and Mathematical Geophysics of Siberian Branch of Russian Academy of Sciences, 2021

© Institute of Information and Computing Technologies of the Ministry of Education of the Republic of Kazakhstan, 2021

СОДЕРЖАНИЕ

Теоретическая и системная информатика

<i>Кальней А. М.</i> Модели многоуровневых сетей (краткий обзор)	5
<i>Мигов Д. А., Коротков А. Н.</i> Использование структурных разрезов в моделировании пространства каскадных отказов в электроэнергетических сетях	21

Прикладные информационные технологии

<i>Акишин В. А.</i> Когнитивная модель оценки клиентского опыта в структуре инфокоммуникационного ландшафта оператора связи	34
<i>Бредихин С. В., Ляпунов В. М., Щербакова Н. Г.</i> Структура и параметры невзвешенной сети соавторства на основе данных БД RePES	56
<i>Миронов В. В.</i> Новый наукометрический индекс издательской активности автора	68
Правила представления и подготовки рукописей для публикации в журнале „ПРОБЛЕМЫ ИНФОРМАТИКИ“	79

На сайте «Объединенного каталога "Пресса России"» www.pressa-rf.ru
можно оформить подписку на 2022 год на печатную версию журнала
«Проблемы информатики» по подписному индексу 69980,
а также подписаться через интернет-магазин
«Пресса по подписке» www.akc.ru

CONTENTS

Theoretical informatics

<i>Kalney A. M.</i> Layered Network Models (Overview)	5
<i>Migov D. A., Korotkov A. N.</i> Cuts using for modeling the propagation of cascading failures in electrical power grids	21

Applied information technologies

<i>Akishin V. A.</i> Cognitive model for evaluating customer experience in the structure of infocommunication landscape of a telecom operator	34
<i>Bredikhin S. V., Lyapunov V. M., Scherbakova N. G.</i> The structure and parameters of the unweighted co-authorship network based on DB RePEc data	56
<i>Mironov V. V.</i> New scientific index of the author's publishing activity	68
Rules of presentation and preparation of manuscripts offered for publication	79

LAYERED NETWORK MODELS (OVERVIEW)

A. M. Kalney

Technograd Plus Ltd,
630087, Novosibirsk Russia,
Institute of Computational Mathematics and Mathematical Geophysics SB RAS,
630090, Novosibirsk, Russia

DOI: 10.24412/2073-0667-2021-3-5-20

Network science is an essential tool for describing and analyzing complex systems in the social, biological, physical, information and engineering sciences. Initially, almost all studies on networks used an abstraction, in which systems are represented by an ordinary graph: the „vertices“ (or „nodes“) of the graph represent some entity or agent, and the connection between a pair of nodes is represented by an „edge“ (or „link“). Loops and multi-edges are usually ignored. Although this approach is rather naive, it has been extremely successful.

With the development of research on complex systems, it became necessary to move towards more complex and realistic models than a simple graph. For example, different heterogeneous properties of edges: they can be directional, have different strengths (i. e., „weights“), and exist only between nodes that belong to different sets (for example, bipartite networks), or be active only at a certain time. Much later, more and more efforts were made to investigate networks with multiple types of connections and the so-called „networks of networks“. Such systems were explored decades ago in disciplines such as sociology and design, but relatively recently serious research has been carried out on multi-level complex networks and generalizations of terminology and tools in this area. One such generalization is the multilayer network model.

In telecommunication networks, problems naturally arise that are solved at several levels of the network. For example, the task of routing in a circuit-switched data network with several logical layers (different technologies) and different interfaces, which can lead to invalid paths. This paper shows a negative example of a graph with edge properties. The tasks of designing a multi-level WLAN structure, a two-level SDH / WDM network were solved, a scheme was developed to protect (restore) a two-level optical network. To assess the distribution of traffic, a two-level model (LCN) was introduced, consisting of physical and logical layers. All of these models are not universal (i. e., they either depend on a specific technology, or are applicable to specific types of networks, or only take into account connections between neighboring layers). However, for modeling multi-level embedded networks of various natures, for more than 30 years several universities in Russia, Kyrgyzstan and Kazakhstan have been using the model of a hyper net and its development. Hyper nets make it possible to adequately describe multi-level networks with an arbitrary number of levels.

The hyper net (or S-hyper net) model consists of a physical layer and a logical layer(s) and is thus an abstraction of computer networks. Probably the largest number of applications of the theory of hyper nets and S-hyper nets are in telecommunications and transport. Nevertheless, the theory of S-hyper nets is applicable to the analysis and synthesis of other systems of network structure. The multilayer network model can be used to represent most types of complex systems (for example, in sociology, epidemiology, biomedicine, etc.) that consist of several networks or include disparate and / or multiple interactions between objects.

Modeling real networks with a more complex model than a graph has long been a necessity. However, questions about the unification of models, and most importantly, terminology began to arise only in the last decade. This article presents two of the most common layered networking models to date. It is shown that the choice of this or that model (even when solving the same problem) is based on the features of the network. For example, the application of the hyper-network model will most likely be appropriate when modeling a multi-layer telecommunication or transport network. A list of the studied literature was also given, as well as the author's works in the table, which showed the model's belonging to one or another class of multi-level networks.

Each of the above models describes only a subset of the set of multi-layer networks. It is possible to develop both existing models and introduce new ones for networks that are not yet integrated into the general theory. Therefore, research in this area will remain relevant and have many applications.

Key words: multilevel networks, multilayer networks, modeling, hyper nets.

References

1. Newman, M. E. J. *Networks: An Introduction*. Oxford: Oxford University Press, 2010.
2. Wasserman, S., Faust, K. *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press, 1994.
3. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.-U. Complex networks: structure and dynamics // *Phys. Reps.* 2006. N 424. P. 175–308.
4. Bollobas, B. *Modern Graph Theory*. Berlin: Springer, 1998.
5. Barabasi, A.-L., Albert, R. Emergence of scaling in random networks // *Science*. 1999. N 286. P. 509–512.
6. Clauset, A., Shalizi, C. R., Newman, M. E. J. Power-law distributions in empirical data // *SIAM Rev.* 2009. N 51. P. 661–703.
7. Watts, D. J., Strogatz, S. H. Collective dynamics of ‘small-world’ networks. *Nature*. 1998. N 393. P. 440–442.
8. Porter, M. A. Small-world network // *Scholarpedia*. 2012. N 7. P. 1739.
9. Porter, M. A., Onnela, J.-P., Mucha, P. J. Communities in networks // *Not. Am. Math. Soc.* 2009. N 56. P. 1082–1097, 1164–1166.
10. Fortunato, S. Community detection in graphs // *Phys. Reps.* 2010. N 486. P. 75–174.
11. Lancichinetti, A., Kivela, M., Saramaki, J., Fortunato, S. Characterizing the community structure of complex networks. *PLOS ONE*. 2010. N 5. P. e11976.
12. Bang-Jensen, J., Gutin, G. *Digraphs: Theory, Algorithms and Applications*, 2nd edn. Berlin: Springer, 2008.
13. Barrat, A., Barthélemy, M., Pastor-Satorras, R., Vespignani, A. The architecture of complex weighted networks // *Proc. Natl Acad. Sci. USA*, 2004. N 101. P. 3747–3752.
14. Newman, M. E. J. Analysis of weighted networks // *Phys. Rev. E*, 2004. N 70. P. 056131.
15. Breiger, R. L. The duality of persons and groups // *Social Forces*, 1974. N 53. P. 181–190.
16. Holme, P., Saramaki, J. Temporal networks // *Phys. Reps.*, 2012. N 519. P. 97–125.
17. Holme, P., Saramaki, J. (eds). *Temporal Networks*. Berlin: Springer, 2013.
18. Mikko Kivela, Alex Arenas, Marc Barthélemy, James P. Gleeson, Yamir Moreno, and Mason A. Porter. (2014), Multilayer networks // *Journal of Complex Networks* 2, 2014. N 3. P. 203–271.
19. D’Agostino, G., Scala, A. *Networks of Networks: The Last Frontier of Complexity*. Berlin: Springer, 2014.
20. Boccaletti, S., Bianconi, G., Criado, R., del Genio, C. I., Gomez-Gardenes, J., Romance, M., Sendina-Nadal, I., Wang, Z., Zanin, M. The structure and dynamics of multilayer networks // *Phys. Reps.*, to appear. 2014.
21. Krackhardt, D. (1987) Cognitive social structures // *Soc. Netw.* 1987. N 9. P. 109–134.

22. Padgett, J. F., Ansell, C. K. Robust action and the rise of the medici, 1400–1434 // *Am. J. Soc.* 1993. N 98. P. 1259–1319.
23. Scott, J. *Social Network Analysis*. New York, NY: SAGE Publications, 2012.
24. Roethlisberger, F., Dickson, W. *Management and the Worker*. Cambridge: Cambridge University Press, 1939.
25. Gluckman, M. *The Judicial Process Among the Barotse of Northern Rhodesia*. Manchester: Manchester University Press, 1955.
26. Verbrugge, L. M. Multiplexity in adult friendships // *Soc. Forces*, 1979. N 57. P. 1286–1309.
27. Craven, P., Wellman, B. (1973) The network city // *Sociol. Inquiry*. 1973. N 43. P. 57–88.
28. Pattison, P., Wasserman, S. Logit models and logistic regressions for social networks: II. Multivariate relations // *Br. J. Math. Stat. Psychol.* 1999. N 52. P. 169–193.
29. Lusher, D., Koskinen, J., Robins, G. *Exponential Random Graph Models for Social Networks*. Cambridge: Cambridge University Press, 2013.
30. Krackhardt, D., Carley, K. M. A PCANS model of structure in organization // *International Symposium on Command and Control Research and Technology*, Monterey, CA, 1998. P. 113–119.
31. Carley, K. M., Hill, V. *Structural change and learning within organizations. Dynamics of Organizational Societies: Models, Theories and Methods* (A. Lomi, ed.). Cambridge, MA: MIT Press/AAAI Press/Live Oak, 2001.
32. Lorrain, F., White, H. C. Structural equivalence of individuals in social networks // *J. Math. Sociol.* 1971. N 1. P. 49–80.
33. Boorman, S. A., White, H. C. Social structure from multiple networks. II. Role structures. *Am. J. Sociol.* 1976. N 81. P. 1384–1446.
34. Breiger, R. L., Pattison, P. E. Cumulated social roles: the duality of persons and their algebras // *Soc. Netw.* 1986. N 8. P. 215–256.
35. Doreian, P., Batagelj, V., Ferligoj, A. *Generalized Blockmodeling*. Cambridge, UK: Cambridge University Press, 2004.
36. White, H. C., Boorman, S. A., Breiger, R. L. Social structure from multiple networks. I. Blockmodels of roles and positions // *Am. J. Sociol.* 1976. P. 730–780.
37. Winship, C., Mandel, M. Roles and positions: a critique and extension of the blockmodeling approach // *Sociol. Method.* 1983–1984. N 14. P. 314–344.
38. Pattison, P. Social networks, algebraic models for. *Encyclopedia of Complexity and Systems Science* (R. A. Meyers ed.). New York: Springer, 2009. P. 8291–8306.
39. Dunlavy, D. M., Kolda, T. G., Kegelmeyer, W. P. Multilinear algebra for analyzing data with multiple linkages. *Graph Algorithms in the Language of Linear Algebra* (J. Kepner & J. Gilbert eds). *Fundamentals of Algorithms*. Philadelphia: SIAM, 2011. P. 85–114.
40. Kolda, T. G., Bader, B. W. Tensor decompositions and applications. *SIAM Rev.* 2009. N 51. P. 455–500.
41. Acar, E., Yener, B. Unsupervised multiway data analysis: a literature survey // *IEEE Trans. Knowl. Data Eng.*, 2009. N 91. P. 6–20.
42. Martin, C. D., Porter, M. A. The extraordinary SVD // *Am. Math. Monthly*. 2012. N 119. P. 838–851.
43. Kolda, T., Bader, B.W. The TOPHITS model for higher-order web link analysis // *Proceedings of the SIAM Data Mining Conference Workshop on Link Analysis, Counterterrorism and Security*, Bethesda, MD, 2006.
44. Kolda, T. G., Bader, B. W., Kenny, J. P. Higher-order web link analysis using multilinear algebra // *Proceedings of the 5th IEEE International Conference on Data Mining (ICDM 2005)*, Houston, TX, 2005. P. 242–249.

45. Zhou, D., Orshanskiy, S. A., Zha, H., Giles, C. L. Co-ranking authors and documents in a heterogeneous network // Seventh IEEE International Conference on Data Mining (ICDM 2007), Omaha, NE, 2007. P. 739–744.
46. Cai, D., Shao, Z., He, X., Yan, X., Han, J. Community mining from multi-relational networks // Proceedings of the 9th European Conference on Principles and Practice of Knowledge Discovery in Databases, Porto, Portugal, 2005.
47. Sun, Y., Han, J. Mining heterogeneous information networks: a structural analysis approach // ACM SIGKDD Explor. Newslett. 2013. N 14. P. 20–28.
48. Carley, K. M. Dynamic network analysis. Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers. Committee on Human Factors, National Research Council, 2003. P. 133–145.
49. Chang, S. E., Seligson, H. A., Eguchi, R. T. Estimation of the economic impact of multiple lifeline disruption: memphis light, gas and water division case study. Technical Report NCEER-96-0011, Multidisciplinary Center for Earthquake Engineering Research (MCEER), Buffalo, NY, 1996.
50. Little, R. G. Controlling cascading failure: understanding the vulnerabilities of interconnected infrastructures // J. Urban. Tech. 2002. N 9. P. 109–123.
51. Rosato, V., Issacharoff, L., Tirittico, F., Meloni, S., Porcellinis, S., Setola, R. Modelling interdependent infrastructures using interacting dynamical models // Int. J. Crit. Infra. 2008. N 4. P. 63–79.
52. Khapugin, S., Rodionov, A. S. Modeling grouped failures in network reliability analysis // B ACM IMCOM 2015 - Proceedings [a44] (ACM IMCOM 2015 - Proceedings). 2015. Association for Computing Machinery, Inc. <https://doi.org/10.1145/2701126.2701218>.
53. Leicht, E. A., D'Souza, R. M. Percolation on interacting networks. arXiv:0907.0894 [condmat. dis-nn]. 2009.
54. Buldyrev, S. V., Parshani, R., Paul, G., Stanley, H. E., Havlin, S. (2010) Catastrophic cascade of failures in interdependent networks // Nature. 2010. N 464. P. 1025–1028.
55. Vespignani, A. The fragility of interdependency // Nature. 2010. N 464. P. 984–985.
56. Dijkstra, F., Andree, B., Koymans, K., van der Hama, J., Grosso, P., de Laat, C. A multi-layer network model based on ITU-T G.805 // Comput. Netw. 2008. N 52. P. 1927–1937.
57. Rotaru, Sandra. Cross-layer design in wireless local area networks (WLANs) issues and possible solutions. 2018. 10.13140/RG.2.2.13913.77924.
58. Koster, A. M. C. A., Orlowski, S., Raack, C., Baier, G., Engel, T., Belotti, P. Branch-and-cut techniques for solving realistic two-layer network design problems // Graphs and Algorithms in Communication Networks. Springer, Heidelberg, 2009. P. 95–118.
59. Chigan, C., Atkinson, G., Nagarajan, R. On the modeling issue of joint cross-layer network protection/restoration // Proceedings of Advanced Simulation Technologies Conference (ASTC '04), 2004. P. 57–62.
60. Kurant, M., Thiran, P. Layered complex networks // Phys. Rev. Lett. 96, 2006. 138701-1–138701-4.
61. Popkov V. K. Matematicheskie modeli svyaznosti. Novosibirsk: IVMiMG SO RAN, 2006.
62. Popkov V. K., Sokolova O. D. Application of Hyper net Theory for the Networks Optimization Problems // 17th IMACS World Congress, July 2005, Paper T4-I-42-011.
63. Rodionov A. S., Sokolova O., Yurgenson A., Choo H. On Optimal Placement of the Monitoring Devices on Channels of Communication Network // ICCSA 2009, Part II, LNCS, Vol. 5593. 474487.
64. Rodionov A. S., Choo H., Nechunaeva K. A. Framework for Biologically Inspired Graph Optimization // Proceedings of ICUIMC 2011, Seoul, Republic of Korea, 2011. P. 2.5, 4 pages.
65. Popkov V. K. Primenenie teorii S-gipersetej dlya modelirovaniya sistem setevoy struktury // Problemy informatiki. 2010. N 4.

66. Popkov V. K. *Matematicheskie modeli zhivuchesti setej svyazi*. Novosibirsk: VC SO AN SSSR, 1990.
67. Popkov V. K. *Metody optimizacii struktur zonovyh setej svyazi* / V.K. Popkov, S. B. Kaul', M. I. Nechepurenko. Novosibirsk: VC SO AN SSSR, 1983.
68. Popkov V. K. *Matematicheskie modeli i metody optimizacii gorodskih transportnyh sistem* // *Materialy 2-j Vseros. konf. „Problemy optimizacii i ekonomicheskie prilozheniya”*, Omsk, 29 iyunya 4 iyulya 2009 g. Omsk: B. i., 2009. S. 80–81.
69. Kalney A., Migov D., Rodionov A., Nasibullina T. Designing of optimal power supply networks for the equipment of multifunctional safety systems // Novosibirsk, Russia, The 2019 15th International Asian School-Seminar Optimization Problems of complex systems.
70. Kal'nej A. M., Rodionov A. S. Analiz nadezhnosti mnogourovnevnyh setej s nenadezhnymi vershinami // *Problemy informatiki*. 2020. N 2. S. 5–15. DOI: 10.24411/2073-0667-2020-10005.
71. Rodionov A., Rodionova O. Random Hypernets in Reliability Analysis of Multilayer Networks. In: Mastorakis N., Bulucea A., Tsekouras G. (eds) *Computational Problems in Science and Engineering* // *Lecture Notes in Electrical Engineering*. 2015. V. 343. Springer, Cham. <https://doi.org/10.1007/978-3-319-15765-8-17>.
72. Garbuzov, A. S. Rodionov. Some Problems of Fuzzy Modeling of Telecommunications Networks // *Proc. of the 9 Int. Conf. on Ubiquitous Information Management and Communication, (IMCOM)*, Indonesia, Bali, 8–10 Jan. New York : ACM, 2015. P. 12.1–12.5. DOI: 10.1145/2701126.2701216. ISBN: 978-1-4503-3377-1. S. 418–422.
73. Hammoud, Zaynab, Kramer, Frank. Multilayer networks: aspects, implementations, and application in biomedicine. *Big Data Analytics*. 2020. 5. 10.1186/s41044-020-00046-0.
74. Kinsley Amy C., Rossi Gianluigi, Silk Matthew J., VanderWaal Kimberly. Multilayer and Multiplex Networks: An Introduction to Their Use in Veterinary Epidemiology. *Frontiers in Veterinary Science*, 2020, P. 596.
75. Sole-Ribalta, Albert & Gomez, Sergio & Arenas, Alex. Congestion Induced by the Structure of Multiplex Networks // *Physical Review Letters*. 2016. 116. 10.1103/PhysRevLett.116.108701.
76. Zhang, Shuai & Liang, Man-Gui & Li, Hui-Jia. Method to enhance traffic capacity for two-layer complex networks // *Canadian Journal of Physics*. 2014. N 92. P. 1599–1605. 10.1139/cjp-2013-0711.
77. Wu, Jiexin, Pu, Cunlai & Li, Lunbo & Cao, Guo. Traffic dynamics on multilayer networks // *Digital Communications and Networks*. 2018. N 6. 10.1016/j.dcan.2018.10.011.

МОДЕЛИ МНОГОУРОВНЕВЫХ СЕТЕЙ (КРАТКИЙ ОБЗОР)

А. М. Кальней

ООО „Техноград плюс“,
630087, Новосибирск, Россия,
Институт вычислительной математики и математической геофизики СО РАН,
630090, Новосибирск, Россия

УДК 519.718

DOI: 10.24412/2073-0667-2021-3-5-20

Ограниченность модели графа при моделировании реальных сетей давно уже стала очевидна. Как следствие, стало появляться огромное количество трудов, учитывающих многоуровневый характер реальных сетей, однако вопросы унификации моделей и терминов стали подниматься лишь в прошлом десятилетии. В данной статье представлен исторический обзор развития моделей многоуровневых сетей. Описаны несколько наиболее развитых моделей на сегодняшний момент, а также логика их построения. Приведена таблица, показывающая примеры применения рассмотренных моделей.

Ключевые слова: многоуровневые сети, многослойные сети, моделирование, гиперсети.

Введение. Наука о сетях (Network science) является важным инструментом для описания и анализа сложных систем в социальных, биологических, физических, информационных и инженерных науках [1–3]. Изначально почти все исследования о сетях использовали абстракцию, в которой системы представлены обычным графом [4]: „узлы“ (или „вершины“) графа представляют некоторую сущность или агента, и связь между парой узлов представляется „ребром“ (или „связью“). Петли и мульти-ребра обычно игнорируются. Хотя этот подход довольно наивен, он был крайне успешен [1, 2, 5–11].

По мере развития исследований сложных систем стало необходимо двигаться к более сложным и реалистичным моделям, чем простой граф. Например, различные неоднородные свойства ребер: они могут быть направленными [1, 2, 12], иметь различную силу (т. е. „веса“) [2, 13, 14], существовать только между узлами, которые принадлежат различным множествам (например двудольные сети) [1, 2, 15] или быть активными только в определенное время [16, 17]. Намного позже стали прилагаться все большие усилия по исследованию сетей с множественными типами связей и так называемые „сети сетей“ [18]. Такие системы были исследованы десятилетия назад в таких дисциплинах как социология и проектирование, но относительно недавно были проведены серьезные исследования многоуровневых сложных сетей и обобщения терминологии и инструментов в этой области [19, 20].

В социальных сетях можно классифицировать ребра в зависимости от характера отношений или действий которые они представляют [2, 21, 22]. Упрощение социальной системы до сети, в которой действующие лица связаны попарно только одним типом отношений является очень грубым приближением реальности. В результате социологи десятилетия назад осознали важность изучения социальных систем путем конструирования множества

социальных сетей, использующих различные связи среди одинаковых множеств индивидов [2, 23]. Например, рассмотрим социолограммы, которые были нарисованы в 1930-е, чтобы представить социальные сети в bank-wiring room [24]. Эти социолограммы отображают отношение между 14 индивидами через 6 различных типов социальных взаимодействий. В социологической литературе сети, в которых каждое ребро классифицируется его типом, называются „мультиплексные сети“ [25, 26] или „многосвязные сети“ [2]. Социальные сети также часто включают несколько типов узлов (например, мужчины и женщины) или иерархические структуры (например, частные лица являются частью организаций), которые изучались, используя „многоуровневые сети“. Определение „сети сетей“ также датируется, по крайней мере, еще в 1973 году [27]. Инструменты, разработанные для исследования многоуровневых социальных сетей, включают модели экспоненциального случайного графа (ERGMs) [28, 29], мета-сети и мета-матрицы [30, 31] и методы выявления социальных ролей с использованием блочного моделирования и реляционных алгебр [32–38].

В сообществах компьютерных наук и вычислительной линейной алгебры методы тензорной декомпозиции [39, 40] и многосторонний анализ данных [41] использовались для изучения различных типов многоуровневых сетей. Эти типы методов основаны на представлении многоуровневых сетей как тензоров смежности „ранга“, большего, чем 2 (например, порядка больше, чем 2), а затем на применении механизмов, разработанных для тензорных разложений. Возможно, наиболее распространенные методы, которые используют этот подход, являются обобщениями сингулярного разложения (SVD) [42], и эти и другие инструменты были крайне успешны во многих приложениях [40]. Например, методы тензорной декомпозиции и многостороннего анализа данных могут использоваться для извлечения сообществ (например наборы узлов, которые плотно соединены друг с другом) [39] или для классифицирования узлов [43, 44] в многоуровневых сетях.

Сетевые системы, которые не могут быть представлены как традиционные графы, также были изучены с точки зрения сбора данных. Например, неоднородные (информационные) сети были разработаны как общая структура чтобы, учесть множество типов узлов и ребер [45–47]. Аналогично, можно использовать мета-матрицы, чтобы провести анализ динамических сетей [48], который включает временную и пространственную информацию, узловые атрибуты и типы и другие типы данных о социальных сетях в той же общей структуре. Мета-матрицы применялись в контексте „организационной теории“ в качестве организации, людей, ресурсов и всех типов взаимосвязанных сущностей [30, 31].

Взаимосвязанные системы рассматривались в технической литературе как источник каскадных отказов [49–52]. По аналогии с понятием „системный риск“ в финансовых системах, увеличение связности, в том числе взаимосвязанности различных систем в инфраструктуре, может привести к увеличению масштабов событий. Позже эти идеи были формализованы взаимодействующими сетями (и взаимозависимыми сетями) [53, 54]. Например, было показано (особенно с использованием процессов просачивания) что взаимосвязанные системы могут реагировать на случайные отказы способом, который отличается от „моноплексных“ (т.е. одноуровневых) сетей. Для некоторых типов процессов каскадного отказа взаимозависимая система может демонстрировать фазовый переход „первого рода“ (т.е. прерывистый) вместо фазовых переходов „второго рода“ (т.е. непрерывных), которые типичны для моноплексных систем [54, 55].

В телекоммуникационных сетях естественным образом возникают задачи, решаемые на нескольких уровнях сети. Например, задача маршрутизации в сети передачи данных

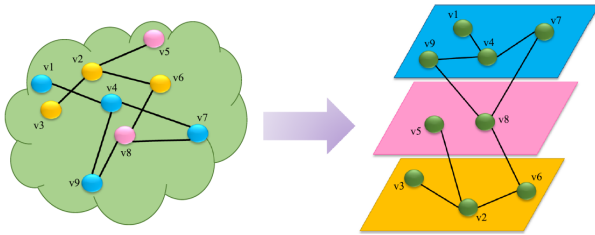


Рис. 1. Раскраска вершин графа. Узлы, имеющие схожие цвета, сгруппированы в одном слое. Ребра те же, что и в исходной сети

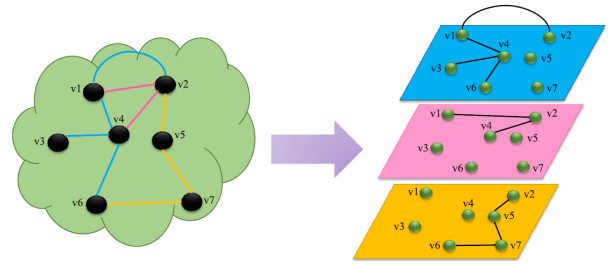


Рис. 2. Раскраска ребер графа. Узлы исходной сети повторяются на различных слоях. Слои представляют различные аспекты отношений в сети. Ребра разделены по слоям на основе их цветов

с коммутацией каналов с наличием нескольких логических слоев (разных технологий) и различных интерфейсов [56], что может приводить к невалидным путям. В данной работе показан негативный пример графа с реберными свойствами. Решались задачи проектирования многоуровневой структуры WLAN [57], двухуровневой SDH/WDM сети [58], разрабатывалась схема для защиты (восстановления) двухуровневой оптической сети [59]. Для оценки распределения трафика вводилась двухуровневая модель (LCN), состоящая из физического и логического уровней [60]. Все эти модели не являются универсальными (т.е. либо зависят от конкретной технологии, либо применимы к конкретным типам сетей, либо учитывают связи только соседних уровней). Однако, для моделирования многоуровневых встроенных сетей различной природы, уже более 30 лет в нескольких университетах России, Киргизии и Казахстана используется модель гиперсети и ее развития [61–72]. Гиперсети позволяют адекватно описывать многоуровневые сети с произвольным количеством уровней.

В последующих двух главах будут кратко описаны несколько наиболее общих моделей, которые описывают два больших класса многоуровневых сетей. После будет приведена таблица с исследованными статьями и их принадлежностью к определенному классу.

1. Многослойная сеть. В данной главе вместо понятия уровень будет применяться близкое по значению понятие слой, как в оригинальной статье [19]. В основе перехода от однослойной сети к сети с большим уровнем слоев используется несколько способов: 1) раскраска вершин графа (см. рис. 1, заимствованный из [73]); 2) раскраска ребер графа (см. рис. 2, заимствованный из [73]).

Как следствие таких рассуждений была создана общая модель, представленная на рис. 3, заимствованном из [19]. Многослойная сеть обладает набором узлов V как обычная сеть (т.е. граф), а также может обладать любым количеством d аспектов, и можно определить последовательность $\mathbf{L} = \{L_a\}_{a=1}^d$ наборов элементарных слоев, таких что есть один набор элементарных слоев L_a для каждого аспекта a .

Имея последовательность наборов элементарных слоев, мы можем сконструировать набор слоев в многослойной сети путем сборки множества всех комбинаций элементарных слоев, используя прямое произведение $L_1 \times \dots \times L_d$. Поскольку узлы могут отсутствовать в некоторых слоях, во-первых вводятся множество $V \times L_1 \times \dots \times L_d$ всех комбинаций и подмножество $V_M \subseteq V \times L_1 \times \dots \times L_d$, которое содержит только те комбинации узлов-слоев, в которых узел присутствует в соответствующем слое. Множество связей между этими узлами обозначим $E_M \subseteq V_M \times V_M$. Таким образом,

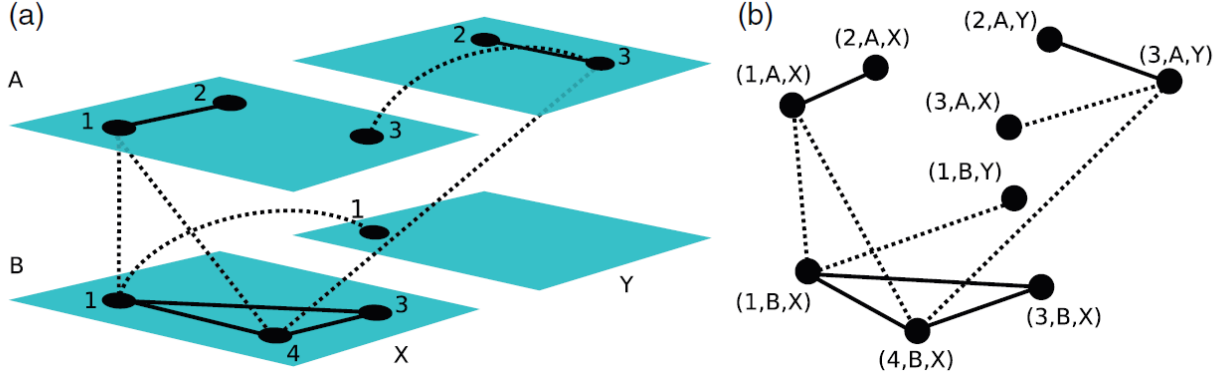


Рис. 3. (а) Пример наиболее общего типа многослойной сети, $M = (V_M, E_M, V, L)$, рассматриваемой в статье [19]. (б) Граф $G_M = (V_M, E_M)$ той же многослойной сети

многослойная сеть определяется $M = (V_M, E_M, V, L)$. Понятия смежности и инцидентности вводятся аналогично сетям с одним слоем. В примере сеть M имеет 4 узла $V = \{1, 2, 3, 4\}$ и два аспекта, которые имеют соответствующие элементарные наборы слоев $L_1 = \{A, B\}$ и $L_2 = \{X, Y\}$. Таким образом, имеем в общем четыре различных слоя: (A, X) , (A, Y) , (B, X) и (B, Y) . Каждый слой содержит некоторый набор узлов множества V ; для этого примера, множествами пар узлов-слоев является $V_M = \{(1, A, X), (2, A, X), (3, A, X), (2, A, Y), (3, A, Y), (1, B, X), (3, B, X), (4, B, X), (1, B, Y)\} \subseteq V \times L_1 \times L_2$. Узлы могут быть соединены друг с другом попарно как в одном слое, так и в разных слоях. Узлы, которые остаются внутри слоя (т.е. внутрислойные ребра), показаны как сплошные линии и ребра, которые пересекают слоя (т.е. межслойные ребра) как пунктирные линии. В графе G_M внутрислойные ребра показываются как сплошные линии и межслойные ребра пунктирными линиями. Матрица смежности графа (или „сверхграфа“) является матрицей сверхсмежности многослойной сети.

2. Гиперсети. Перейдем к описанию гиперсетевых моделей. Двухуровневая или простая гиперсеть $H = (X, V, R; P, F, W)$ состоит из следующих объектов (см. рис. 4, заимствованный из [71]): $X = (x_1, \dots, x_n)$ — набор вершин; $V = (v_1, \dots, v_m)$ — набор ветвей (ребер графа первичной сети); $R = (r_1, \dots, r_g)$ — набор ребер (ребер графа вторичной сети); $P : V \rightarrow X \times X$ — отображение, которое определяет граф $PN = (X, V)$, который называется первичной сетью; $W : R \rightarrow X \times X$ — отображение, которое определяет граф $SN = (X, R)$, который называется вторичной сетью; $F : R \rightarrow 2^V$ — отображение ребер в маршруты PN .

Вводя множественное вложение вторичных сетей в первичную сеть, вложение вторичных сетей друг в друга, а также отображение нескольких вершин вторичной сети в узел первичной, приходим к формальному определению S-гиперсети [65].

Пусть задано множество графов (гиперграфов) $G_0 = (X^0, V), G_1 = (X^1, U^1), \dots, G_k = (X^k, U^k)$ и корневое дерево $T_0 = (Z, R)$, где $Z = z_0, z_1, \dots, z_k, R = r_1, \dots, r_k$ — определяющее вложение графов G_j в G_i ($i < j$) аналогично вложениям, определяемым в гиперсетях, за исключением того, что вершины x_k^i и x_l^j графов G_i и G_j не тождественны, а инциденты. Очевидно, что одной и той же вершине x_k^i могут быть инциденты несколько вершин $X_k^j = \{x_{k_1}^{j_1}, x_{k_2}^{j_2}, \dots, x_{k_l}^{j_l}\}$ из l графов $\{G_{j_s}\}$, $s = 1, \dots, l$. На множестве вершин X_l^j определяется

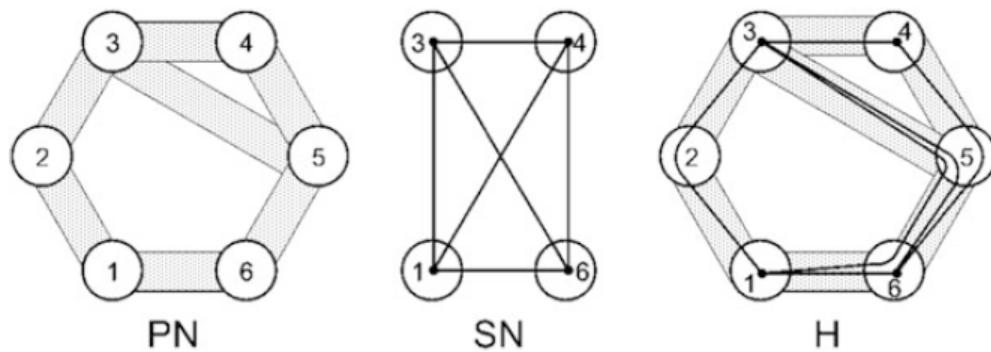


Рис. 4. Иллюстрация к определению гиперсети: PN — это первичная сеть; темные вершины $\{1, 3, 4, 6\}$ из набора вершин относятся к SN ; SN — это вторичная сеть (в нашем случае является полным графом); H — это гиперсеть (SN отображается в PN)

$L^j = (X_k^j, E)$. Вершины $x_{k_j}^{j_i}$ и $x_{k_s}^{j_s}$ смежны в L^j , если соответствующие графы G_{j_i} и G_{j_s} в вершине x_k^i имеют некоторую системообразующую связь $l(x^{j_i}, x^{j_s})$. В противном случае эти вершины не связаны. Так же, как и в гиперсетях, ребру $u_l^j \in G_j$ в графе G_i сопоставляется цепь или некоторая связная часть между соответствующими вершинами из G_i . На рис. 5 (заимствованном из [65]) приведен пример такой гиперсети. Системообразующие связи типа $\{l(x, y)\}$ могут иметь различную природу и, как правило, существенно зависят от времени.

Определенную таким образом S-гиперсеть с помощью приведенных в [65] матриц можно задать с точностью до изоморфизма и даже с точностью до нумерации вершин и ребер. В [71] рассмотрены различные варианты гиперсетей с ненадежными, а в [72] — нечеткими элементами.

3. Сравнение. Вышеприведенные модели описывают несколько крупных классов многоуровневых сетей. Модель гиперсети (или S-гиперсети) состоит из физического слоя и логического(-их) слоя(-ев) и тем самым является абстракцией компьютерных сетей. Вероятно, наибольшее число приложений теории гиперсетей и S-гиперсетей приходится на электросвязь и транспорт. Тем не менее, теория S-гиперсетей применима для анализа и синтеза других систем сетевой структуры. Другая же модель может использоваться для представления большинства типов сложных систем (например, в социологии, эпидемиологии, биомедицине и т. д.), которые состоят из нескольких сетей или включают несопоставимые и/или множественные взаимодействия между объектами. Приведем таблицу с изученными статьями, где будет сравниваться принадлежность рассматриваемой в статье модели к тому или иному классу многоуровневых сетей.

В обзоре многослойных сетей [19] резюмируют различные наименования этих сетей в таблице из 26 различных наименований (а также еще большее количество ссылок на статьи, которые можно было бы вписать в таблицу в качестве моделей многослойных сетей) и вводят общее определение многослойных сетей. Также эти наименования классифицируются на основе типов ограничений, которые налагаются на это определение многослойной сети.

В части статей, которые приписываются к классу гиперсетей [56–59] не используется модель гиперсети, однако ее применение вполне возможно. В [60, 75, 76] решаются задачи,

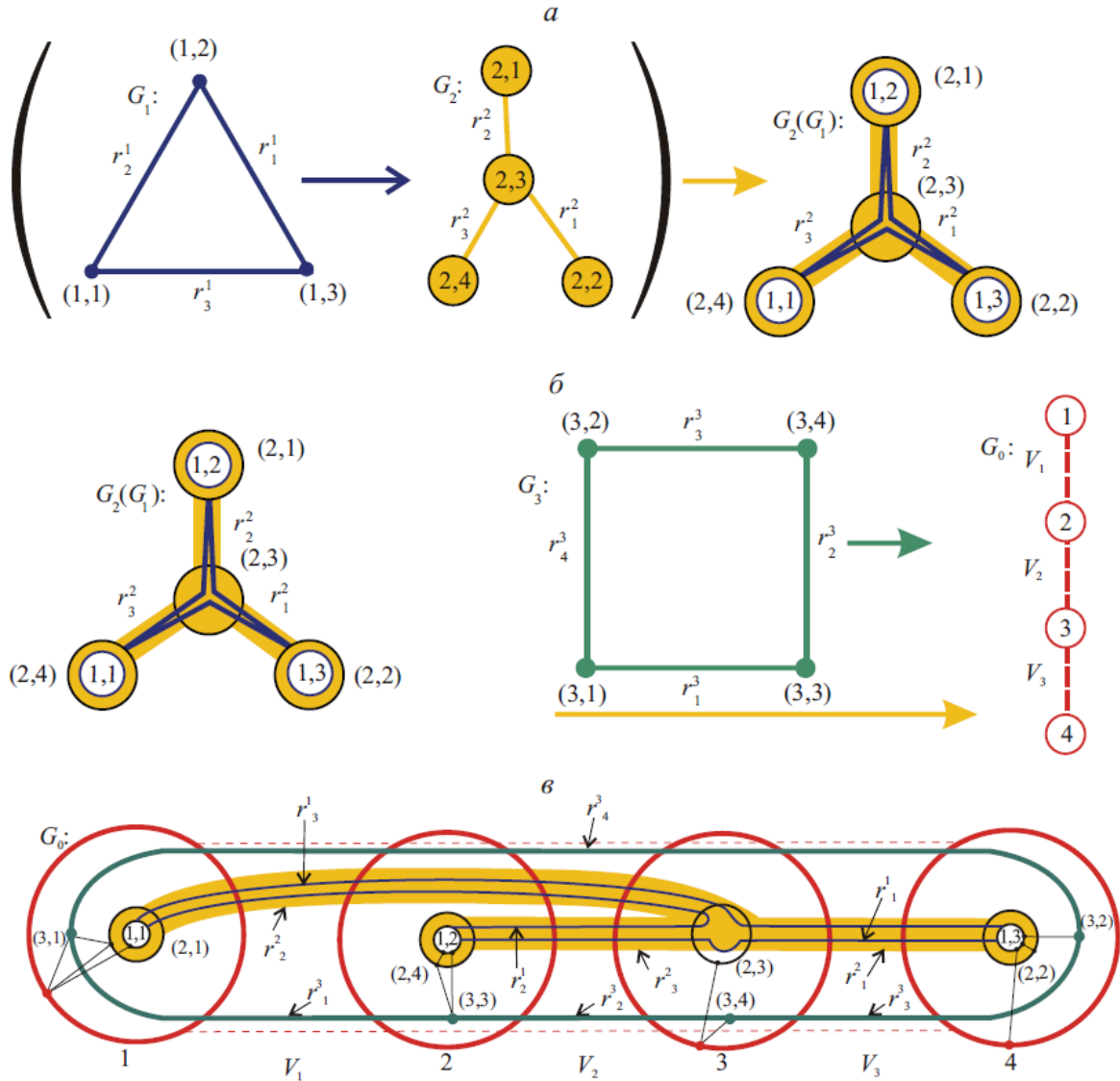


Рис. 5. Гиперсеть и ее составляющие: a — отображение графа вторичной сети G_1 в граф первичной сети G_2 с образованием гиперсети $G_2(G_1)$; b — отображение гиперсети $G_2(G_1)$ и графа G_3 в граф G_0 с образованием S-гиперсети $H = (G_0(G_3, G_2(G_1)))$; v — гиперсеть $H = (G_0(G_3, G_2(G_1)))$, в которой определены внутриузловые связи (графы L_i)

связанные с оценкой, динамикой трафика в сети. Однако из-за особенностей сетей может понадобиться применение той или иной модели [77].

Заключение. Моделирование реальных сетей более сложной моделью, чем граф, давно уже стало необходимостью. Однако вопросы насчет унификации моделей, а самое главное, терминологии начали возникать только в прошлом десятилетии. В данной статье представлены две наиболее общие модели многоуровневых сетей на текущий момент. Показано, что выбор той или иной модели (даже при решении одной и той же задачи) происходит на основе особенностей сети. Например, применение модели гиперсети будет, скорее всего, уместным при моделировании многоуровневой сети электросвязи или транспорта.

Таблица

Принадлежность моделей к рассматриваемым классам многоуровневых сетей

Статья	(S-) Гиперсеть	Многослойная сеть
A multi-layer network model based on ITU-T G.805 [56]	✓	
Cross-layer design in wireless local area networks (WLANs) issues and possible solutions [57]	✓	
Branch-and-cut techniques for solving realistic two-layer network design problems [58]	✓	
On the modeling issue of joint cross-layer network protection/restoration [59]	✓	
Layered complex networks [60]	✓	
Математические модели связности [61]	✓	
Application of Hyperneet Theory for the Networks Optimazation Problems [62]	✓	
On Optimal Placement of the Monitoring Devices on Channels of Communication Network [63]	✓	
Framework for Biologically Inspired Graph Optimization [64]	✓	
Математические модели живучести сетей связи [66]	✓	
Методы оптимизации структур зоновых сетей связи [67]	✓	
Математические модели и методы оптимизации городских транспортных систем [68]	✓	
Multilayer and Multiplex Networks: An Introduction to Their Use in Veterinary Epidemiology [74]		✓
Multilayer networks: aspects, implementations, and application in biomedicine [73]		✓
Congestion Induced by the Structure of Multiplex Networks [75]		✓
Method to enhance traffic capacity for two-layer complex networks [76]	✓	
Designing of optmal power supply networks for the equipment of multifunctional safety systems [69]	✓	
Анализ надежности многоуровневых сетей с ненадежными вершинами [70]	✓	
Random Hypernets in Reliability Analysis of Multilayer Networks [71]	✓	
Some Problems of Fuzzy Modeling of Telecommunications Networks [72]	✓	

Приведен список изученной литературы, а также работ автора в таблице, где показывалась принадлежность модели к тому или иному классу многоуровневых сетей.

Каждая приведенная выше модель описывает лишь некоторое подмножество множества многоуровневых сетей. Можно развивать как существующие модели, так и вводить новые для сетей, которые еще не встраиваются в общую теорию. Поэтому исследования в данной области будут оставаться актуальными и иметь множество приложений.

Список литературы

1. Newman, M. E. J. Networks: An Introduction. Oxford: Oxford University Press, 2010.
2. Wasserman, S., Faust, K. Social Network Analysis: Methods and Applications. Cambridge: Cambridge University Press, 1994.
3. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.-U. Complex networks: structure and dynamics // Phys. Reps. 2006. N 424. P. 175–308.
4. Bollobas, B. Modern Graph Theory. Berlin: Springer, 1998.

5. Barabasi, A.-L., Albert, R. Emergence of scaling in random networks // *Science*. 1999. N 286. P. 509–512.
6. Clauset, A., Shalizi, C. R., Newman, M. E. J. Power-law distributions in empirical data // *SIAM Rev.* 2009. N 51. P. 661–703.
7. Watts, D. J., Strogatz, S. H. Collective dynamics of ‘small-world’ networks. *Nature*. 1998. N 393. P. 440–442.
8. Porter, M. A. Small-world network // *Scholarpedia*. 2012. N 7. P. 1739.
9. Porter, M. A., Onnela, J.-P., Mucha, P. J. Communities in networks // *Not. Am. Math. Soc.* 2009. N 56. P. 1082–1097, 1164–1166.
10. Fortunato, S. Community detection in graphs // *Phys. Reps.* 2010. N 486. P. 75–174.
11. Lancichinetti, A., Kivela, M., Saramaki, J., Fortunato, S. Characterizing the community structure of complex networks. *PLOS ONE*. 2010. N 5. P. e11976.
12. Bang-Jensen, J., Gutin, G. *Digraphs: Theory, Algorithms and Applications*, 2nd edn. Berlin: Springer, 2008.
13. Barrat, A., Barthélemy, M., Pastor-Satorras, R., Vespignani, A. The architecture of complex weighted networks // *Proc. Natl Acad. Sci. USA*, 2004. N 101. P. 3747–3752.
14. Newman, M. E. J. Analysis of weighted networks // *Phys. Rev. E*, 2004. N 70. P. 056131.
15. Breiger, R. L. The duality of persons and groups // *Social Forces*, 1974. N 53. P. 181–190.
16. Holme, P., Saramaki, J. Temporal networks // *Phys. Reps.*, 2012. N 519. P. 97–125.
17. Holme, P., Saramaki, J. (eds). *Temporal Networks*. Berlin: Springer, 2013.
18. Mikko Kivela, Alex Arenas, Marc Barthélemy, James P. Gleeson, Yamir Moreno, and Mason A. Porter. (2014), Multilayer networks // *Journal of Complex Networks* 2, 2014. N 3. P. 203–271.
19. D’Agostino, G., Scala, A. *Networks of Networks: The Last Frontier of Complexity*. Berlin: Springer, 2014.
20. Boccaletti, S., Bianconi, G., Criado, R., del Genio, C. I., Gomez-Gardenes, J., Romance, M., Sendina-Nadal, I., Wang, Z., Zanin, M. The structure and dynamics of multilayer networks // *Phys. Reps.*, to appear. 2014.
21. Krackhardt, D. (1987) Cognitive social structures // *Soc. Netw.* 1987. N 9. P. 109–134.
22. Padgett, J. F., Ansell, C. K. Robust action and the rise of the medici, 1400–1434 // *Am. J. Soc.* 1993. N 98. P. 1259–1319.
23. Scott, J. *Social Network Analysis*. New York, NY: SAGE Publications, 2012.
24. Roethlisberger, F., Dickson, W. *Management and the Worker*. Cambridge: Cambridge University Press, 1939.
25. Gluckman, M. *The Judicial Process Among the Barotse of Northern Rhodesia*. Manchester: Manchester University Press, 1955.
26. Verbrugge, L. M. Multiplexity in adult friendships // *Soc. Forces*, 1979. N 57. P. 1286–1309.
27. Craven, P., Wellman, B. (1973) The network city // *Sociol. Inquiry*. 1973. N 43. P. 57–88.
28. Pattison, P., Wasserman, S. Logit models and logistic regressions for social networks: II. Multivariate relations // *Br. J. Math. Stat. Psychol.* 1999. N 52. P. 169–193.
29. Lusher, D., Koskinen, J., Robins, G. *Exponential Random Graph Models for Social Networks*. Cambridge: Cambridge University Press, 2013.
30. Krackhardt, D., Carley, K. M. A PCANS model of structure in organization // *International Symposium on Command and Control Research and Technology*, Monterey, CA, 1998. P. 113–119.
31. Carley, K. M., Hill, V. Structural change and learning within organizations. *Dynamics of Organizational Societies: Models, Theories and Methods* (A. Lomi, ed.). Cambridge, MA: MIT Press/AAAI Press/Live Oak, 2001.
32. Lorrain, F., White, H. C. Structural equivalence of individuals in social networks // *J. Math. Sociol.* 1971. N 1. P. 49–80.

33. Boorman, S. A., White, H. C. Social structure from multiple networks. II. Role structures. *Am. J. Sociol.* 1976. N 81. P. 1384–1446.
34. Breiger, R. L., Pattison, P. E. Cumulated social roles: the duality of persons and their algebras // *Soc. Netw.* 1986. N 8. P. 215–256.
35. Doreian, P., Batagelj, V., Ferligoj, A. *Generalized Blockmodeling*. Cambridge, UK: Cambridge University Press, 2004.
36. White, H. C., Boorman, S. A., Breiger, R. L. Social structure from multiple networks. I. Blockmodels of roles and positions // *Am. J. Sociol.* 1976. P. 730–780.
37. Winship, C., Mandel, M. Roles and positions: a critique and extension of the blockmodeling approach // *Sociol. Method.* 1983–1984. N 14. P. 314–344.
38. Pattison, P. Social networks, algebraic models for. *Encyclopedia of Complexity and Systems Science* (R. A. Meyers ed.). New York: Springer, 2009. P. 8291–8306.
39. Dunlavy, D. M., Kolda, T. G., Kegelmeyer, W. P. Multilinear algebra for analyzing data with multiple linkages. *Graph Algorithms in the Language of Linear Algebra* (J. Kepner & J. Gilbert eds). *Fundamentals of Algorithms*. Philadelphia: SIAM, 2011. P. 85–114.
40. Kolda, T. G., Bader, B. W. Tensor decompositions and applications. *SIAM Rev.* 2009. N 51. P. 455–500.
41. Acar, E., Yener, B. Unsupervised multiway data analysis: a literature survey // *IEEE Trans. Knowl. Data Eng.*, 2009. N 91. P. 6–20.
42. Martin, C. D., Porter, M. A. The extraordinary SVD // *Am. Math. Monthly.* 2012. N 119. P. 838–851.
43. Kolda, T., Bader, B.W. The TOPHITS model for higher-order web link analysis // *Proceedings of the SIAM Data Mining Conference Workshop on Link Analysis, Counterterrorism and Security*, Bethesda, MD, 2006.
44. Kolda, T. G., Bader, B. W., Kenny, J. P. Higher-order web link analysis using multilinear algebra // *Proceedings of the 5th IEEE International Conference on Data Mining (ICDM 2005)*, Houston, TX, 2005. P. 242–249.
45. Zhou, D., Orshanskiy, S. A., Zha, H., Giles, C. L. Co-ranking authors and documents in a heterogeneous network // *Seventh IEEE International Conference on Data Mining (ICDM 2007)*, Omaha, NE, 2007. P. 739–744.
46. Cai, D., Shao, Z., He, X., Yan, X., Han, J. Community mining from multi-relational networks // *Proceedings of the 9th European Conference on Principles and Practice of Knowledge Discovery in Databases*, Porto, Portugal, 2005.
47. Sun, Y., Han, J. Mining heterogeneous information networks: a structural analysis approach // *ACM SIGKDD Explor. Newslett.* 2013. N 14. P. 20–28.
48. Carley, K. M. Dynamic network analysis. *Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers*. Committee on Human Factors, National Research Council, 2003. P. 133–145.
49. Chang, S. E., Seligson, H. A., Eguchi, R. T. Estimation of the economic impact of multiple lifeline disruption: memphis light, gas and water division case study. Technical Report NCEER-96-0011, Multidisciplinary Center for Earthquake Engineering Research (MCEER), Buffalo, NY, 1996.
50. Little, R. G. Controlling cascading failure: understanding the vulnerabilities of interconnected infrastructures // *J. Urban. Tech.* 2002. N 9. P. 109–123.
51. Rosato, V., Issacharoff, L., Tiriticco, F., Meloni, S., Porcellinis, S., Setola, R. Modelling interdependent infrastructures using interacting dynamical models // *Int. J. Crit. Infra.* 2008. N 4. P. 63–79.
52. Khapugin, S., Rodionov, A. S. Modeling grouped failures in network reliability analysis // *B ACM IMCOM 2015 - Proceedings [a44] (ACM IMCOM 2015 - Proceedings)*. 2015. Association for Computing Machinery, Inc. <https://doi.org/10.1145/2701126.2701218>.

53. Leicht, E. A., D'Souza, R. M. Percolation on interacting networks. arXiv:0907.0894 [condmat. dis-nn]. 2009.
54. Buldyrev, S. V., Parshani, R., Paul, G., Stanley, H. E., Havlin, S. (2010) Catastrophic cascade of failures in interdependent networks // *Nature*. 2010. N 464. P. 1025–1028.
55. Vespignani, A. The fragility of interdependency // *Nature*. 2010. N 464. P. 984–985.
56. Dijkstra, F., Andree, B., Koymans, K., van der Hama, J., Grosso, P., de Laat, C. A multi-layer network model based on ITU-T G.805 // *Comput. Netw.* 2008. N 52. P. 1927–1937.
57. Rotaru, Sandra. Cross-layer design in wireless local area networks (WLANs) issues and possible solutions. 2018. 10.13140/RG.2.2.13913.77924.
58. Koster, A. M. C. A., Orlowski, S., Raack, C., Baier, G., Engel, T., Belotti, P. Branch-and-cut techniques for solving realistic two-layer network design problems // *Graphs and Algorithms in Communication Networks*. Springer, Heidelberg, 2009. P. 95–118.
59. Chigan, C., Atkinson, G., Nagarajan, R. On the modeling issue of joint cross-layer network protection/restoration // *Proceedings of Advanced Simulation Technologies Conference (ASTC '04)*, 2004. P. 57–62.
60. Kurant, M., Thiran, P. Layered complex networks // *Phys. Rev. Lett.* 96, 2006. 138701-1–138701-4.
61. Попков В. К. Математические модели связности. Новосибирск: ИВМ и МГ СО РАН, 2006. 490 с.
62. Popkov V. K., Sokolova O. D. Application of Hyper net Theory for the Networks Optimization Problems // 17th IMACS World Congress, July 2005, Paper T4-I-42-011.
63. Rodionov A. S., Sokolova O., Yurgenson A., Choo H. On Optimal Placement of the Monitoring Devices on Channels of Communication Network // *ICCSA 2009, Part II, LNCS*, Vol. 5593. 474487.
64. Rodionov A. S., Choo H., Nechunaeva K. A. Framework for Biologically Inspired Graph Optimization // *Proceedings of ICUIMC 2011*, Seoul, Republic of Korea, 2011. P. 2.5, 4 pages.
65. Попков Владимир Константинович Применение теории S-гиперсетей для моделирования систем сетевой структуры // *Проблемы информатики*. 2010. № 4.
66. Попков В. К. Математические модели живучести сетей связи. Новосибирск: ВЦ СО АН СССР, 1990. 233 с.
67. Попков В. К. Методы оптимизации структур зоновых сетей связи / В.К. Попков, С. Б. Кауль, М. И. Нечепуренко. Новосибирск: ВЦ СО АН СССР, 1983. 182 с.
68. Попков В. К. Математические модели и методы оптимизации городских транспортных систем // *Материалы 2-й Всерос. конф. „Проблемы оптимизации и экономические приложения“*, Омск, 29 июня — 4 июля 2009 г. Омск: Б. и., 2009. С. 80–81.
69. Kalney A., Migov D., Rodionov A., Nasibullina T. Designing of optimal power supply networks for the equipment of multifunctional safety systems // *Novosibirsk, Russia, The 2019 15th International Asian School-Seminar Optimization Problems of complex systems*.
70. Кальней А. М., Родионов А. С. Анализ надежности многоуровневых сетей с ненадежными вершинами // журнал „Проблемы информатики“, 2020, № 2, с.5-15. DOI: 10.24411/2073-0667-2020-10005
71. Rodionov A., Rodionova O. Random Hypernets in Reliability Analysis of Multilayer Networks. In: Mastorakis N., Bulucea A., Tsekouras G. (eds) *Computational Problems in Science and Engineering // Lecture Notes in Electrical Engineering*. 2015. V. 343. Springer, Cham. <https://doi.org/10.1007/978-3-319-15765-8-17>.
72. Garbuzov, A. S. Rodionov. Some Problems of Fuzzy Modeling of Telecommunications Networks // *Proc. of the 9 Int. Conf. on Ubiquitous Information Management and Communication, (IMCOM)*, Indonesia, Bali, 8–10 Jan. New York : ACM, 2015. P. 12.1–12.5. DOI: 10.1145/2701126.2701216. ISBN: 978-1-4503-3377-1. S. 418–422.

73. Hammoud, Zaynab, Kramer, Frank. Multilayer networks: aspects, implementations, and application in biomedicine. Big Data Analytics. 2020. 5. 10.1186/s41044-020-00046-0.
74. Kinsley Amy C., Rossi Gianluigi, Silk Matthew J., VanderWaal Kimberly. Multilayer and Multiplex Networks: An Introduction to Their Use in Veterinary Epidemiology. Frontiers in Veterinary Science, 2020, P. 596.
75. Sole-Ribalta, Albert \& Gomez, Sergio \& Arenas, Alex. Congestion Induced by the Structure of Multiplex Networks // Physical Review Letters. 2016. 116. 10.1103/PhysRevLett.116.108701.
76. Zhang, Shuai \& Liang, Man-Gui \& Li, Hui-Jia. Method to enhance traffic capacity for two-layer complex networks // Canadian Journal of Physics. 2014. N 92. P. 1599–1605. 10.1139/cjp-2013-0711.
77. Wu, Jiexin, Pu, Cunlai \& Li, Lunbo \& Cao, Guo. Traffic dynamics on multilayer networks // Digital Communications and Networks. 2018. N 6. 10.1016/j.dcan.2018.10.011.



Кальней Артем Макси-мович получил бакалаврскую и магистерскую степени по направлению „Математика и компьютерные науки“ в Новосибирском государственном университете в 2017 и в 2019 годах. Сейчас специализируется в Институте вычислительной математики и математической геофизики СО РАН.

Его основными научными интересами являются дискретная математика и надежность сетей.

Kalney Artyom Maksimovich received a bachelor's and master's degree in mathematics and computer science from Novosibirsk State University in 2017 and 2019. Now specializes in the Institute of Computational Mathematics and Mathematical Geophysics SB RAS. His main research interests are discrete mathematics and network reliability.

Дата поступления — 05.04.2021

CUTS USING FOR MODELING THE PROPAGATION OF CASCADING FAILURES IN ELECTRICAL POWER GRIDS

D. A. Migov, A. N. Korotkov*

Institute of Computational Mathematics and Mathematical Geophysics SB RAS,
630090, Novosibirsk, Russia,

*Center for Project Development Petersburg Real Estate LLC,
196066, St. Petersburg, Russia

DOI: 10.24412/2073-0667-2021-3-21-33

Analysis of network reliability is extremely important for their design and operation. For various types of networks, various models have been proposed that take into account network particular features, within which different indicators of network reliability are considered. As a rule, random graphs in various modifications are taken as a basis. Usually, the probability of connectivity of the corresponding random graph in the case of unreliable edges that fail independently and absolutely reliable nodes is used as an indicator of network reliability.

The problems of exact calculating of various reliability indicators are NP-hard. When network elements are subject to dependent failure, reliability analysis becomes a much more time consuming task. A typical example of dependent network failures is cascading failures in power networks. The initiating event of the failure propagation process is caused by external circumstances: it can be a fallen tree, a strong gust of wind, a line break due to overload, etc. If its failure caused overloading of other lines or equipment, then this, in turn, can generate new outages, etc. Thus, a sequence of dependent failures occurs.

An important property of cascade outages in power grids is both their locality and their non-locality, as practice shows. The examples of real cascading outages show that there is a failure of lines along the sections that cut off certain subnets. Such scenarios for the propagation of cascading failures are explained by the fact that the failure of a power transmission line leads to an almost instantaneous redistribution of electricity to other power transmission lines, primarily to those lines that are included in the cut with the failed one. This paper proposes a model for the propagation of dependent failures in a network along its structural cuts.

As a structural model of a power grid, we consider an undirected graph $G = (V, E)$, where V is the set of vertices, and E is the set of edges of the graph G . Let the presence probability be given for each edge. We will interpret this value as the probability of failure-free operation of the corresponding transmission line within a given time interval. If a failure occurs, then a cascade failure begins along the network cuts, the development of which is described by influence graphs. In this case, it is assumed that the vertices are absolutely reliable, i.e. are present with probability equal to 1.

In such conditions, several characteristics are considered as indicators of power network reliability: the probability of network connectivity, the probability of each consumer can connect to any power center, the probability of the availability of any power source for a given proportion of consumers. The last indicator can be more informative than the previous ones when considering a power grid of large dimension, for example, on a national scale, or several countries, if the corresponding networks are interconnected.

The article proposes an algorithm for the accurate calculation of reliability indices, based on the use of the total probability formula, and an estimation algorithm, based on the Monte Carlo approach.

In addition, a method of cumulative updating of the bounds of reliability indicators is proposed, which makes it possible to make a conclusion about the sufficient reliability (or unreliability) of a network in relation to a given threshold.

The pseudocodes of the proposed algorithms and the results of numerical experiments are given.

Key words: power grid, reliability, cascading failure, dependent failure, influence graph.

References

1. Manov N. A., Xoxlov M. V., Chukreev Yu. Ya., Shumilova G. P., Uspenskij M. I., Chukreev M. Yu., Polubotko D. V., Gotman N. E., Starceva T. B. *Metody i modeli issledovaniya nadezhnosti e'lektroe'nergeticheskix sistem*. Komi nauchny'j centr UrO RAN. Sy'kty'vkar, 2010.
2. Shilin A. N., Soshinov A. G., Elfimova O. I., Shilin A. A. *Ocenka nadyozhnosti vozduzhny'x linij e'lektroperedachi po topologicheskim sxemam s uchyotom vliyaniya pogodny'x uslovij dlya realizacii informacionno-izmeritel'noj sistemy*. VolgGTU, KTI (filial) VolgGTU. Volgograd, 2017.
3. Krupenev D., Boyarkin D., Iakubovskii D. Improvement in the computational efficiency of a technique for assessing the reliability of electric power systems based on the Monte Carlo method // *Reliability Engineering and System Safety*. 2020. Vol. 204. ID: 107171. DOI: 10.1016/j.res.2020.107171.
4. Metodicheskie voprosy issledovaniya nadezhnosti bol'shix sistem e'nergetiki. Vy'p. 61. *Problemy issledovaniya i obespecheniya nadezhnosti liberalizovanny'x sistem e'nergetiki* / Otv. red. N. I. Voropaj, A. D. Tevyashev. Irkutsk: ISE'M SO RAN, 2011.
5. Kaplunovich P. A., Turitsyn K. S. Fast selection of $N - 2$ contingencies for online security assessment // *Proc. of 2013 IEEE Power and Energy Society General Meeting (PES)*, 21–25 July 2013, Vancouver. IEEE Press. P. 1–5. DOI: 10.1109/PESMG.2013.6672792.
6. Colbourn Ch. J. *The combinatorics of network reliability*. N. Y., Oxford Univ. Press, 1987.
7. Cancela H., El Khadiri M., and Petingi L. Polynomial-time topological reductions that preserve the diameter constrained reliability of a communication network // *IEEE Trans. on Reliability*. 2011. Vol. 60, N 4. P. 845–851.
8. Rodionov A. S., Rodionova O. K. Kumulyativny'e ocenki srednej veroyatnosti svyaznosti pary' verшин sluchajnogo grafa // *Problemy informatiki*. 2013. Vol. 2. P. 3–12.
9. Rodionov A. S., Khapugin S. A. Modelirovanie gruppovy'x otkazov v analize nadezhnosti setej // *Problemy informatiki*. 2015. Vol. 2. P. 31–43.
10. Chen W., Huang N., Kang R. A reliability model with the dependent failures for telecommunication network // *Proc. of the 8th International Conference on Reliability, Maintainability and Safety, Chengdu*, 2009. IEEE Press. P. 1129–1132. DOI: 10.1109/ICRMS.2009.5270059.
11. Barrera J., Cancela H., Eduardo M. Topological optimization of reliable networks under dependent failures // *Oper. Res. Let.* 2015. 43(2). P. 132–136.
12. Chang L., Wub Z. Performance and reliability of electrical power grids under cascading failures // *International Journal of Electrical Power & Energy Systems*. Oct. 2011. Vol. 33. N 8. P. 1410–1419.
13. Hines P., Balasubramaniam K., Sanchez E. Cascading failures in power grids // *IEEE Potentials*. Sept. 2009. Vol. 28. N 5. P. 24–30.
14. Hines P. D. H., Dobson I., Rezaei P. Cascading power outages propagate locally in an influence graph that is not the actual grid topology // *IEEE Transactions on Power Systems*. March 2017. Vol. 32. N 2. P. 958 — 967.
15. Page L. B., Perry J. E. A practical implementation of the factoring theorem for network reliability // *IEEE Trans. on Reliability*. 1998. Vol. 37. N 3. P. 259–267.
16. Won J.-M., Karray F. Cumulative update of all-terminal reliability for faster feasibility decision // *IEEE Trans. Reliability*. 2010. V. 59. N 3. P. 551–562.
17. Rodionov A., Migov D., and Rodionov O. Improvements in the efficiency of cumulative updating of all terminal network reliability // *IEEE Trans. Reliability*. 2012. V. 61. N 2. P. 460–465.

18. Migov D. A., Vins D. V. Parallel'naya realizaciya i imitacionnoe modelirovanie ocenki nadyozhnosti seti metodom Monte-Karlo // Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie vychislitel'naja tehnika i informatika [Tomsk State University Journal of Control and Computer Science]. 2019. Vol. 47. P. 66–74. DOI: 10.17223/19988605/47/8.

19. Crucitti P., Latora V., Marchiori M. Locating critical lines in high-voltage electrical power grids // Fluctuation and Noise Letters. 2005. Vol. 5. N 2. P. 201–208. DOI 10.1142/S0219477505002562.

ИСПОЛЬЗОВАНИЕ СТРУКТУРНЫХ РАЗРЕЗОВ В МОДЕЛИРОВАНИИ РАСПРОСТРАНЕНИЯ КАСКАДНЫХ ОТКАЗОВ В ЭЛЕКТРОЭНЕРГЕТИЧЕСКИХ СЕТЯХ

Д. А. Мигов, А. Н. Коротков*

Институт вычислительной математики и математической геофизики СО РАН,
630090, Новосибирск, Россия,

*ООО „Центр развития проектов «Петербургская Недвижимость»“,
196066, Санкт-Петербург, Россия

УДК 621.311.1+519.17

DOI: 10.24412/2073-0667-2021-3-21-33

Предлагается модель распространения зависимых отказов в сети по ее структурным разрезам. Подобное может происходить в электроэнергетических сетях: отказ линии электропередачи приводит к практически мгновенному перераспределению электроэнергии на другие линии электропередачи, в первую очередь — на те линии, которые входят в разрез с отказавшей. Для показателей надежности сетей в таких условиях — вероятности связности сети, вероятности связности каждого потребителя с центром питания — разработаны методы их точного и приближенного расчета надежности, а также метод кумулятивного уточнения границ надежности.

Ключевые слова: электроэнергетические сети, надежность, каскадный отказ, зависимый отказ, граф влияний.

Введение. Анализ надежности сетей крайне важен при их проектировании и эксплуатации, исследования на эту тему проводятся уже давно, активно продолжаются и в настоящее время, в том числе и на тему анализа надежности и живучести электроэнергетических сетей [1–5]. При этом для различных типов сетей предложены различные, учитывающие именно их особенности модели, в рамках которых рассматриваются разные показатели надежности сети. Как правило, за основу берутся случайные графы в различных модификациях. Одной из наиболее изученных моделей сети с ненадежными элементами является случайный граф Эрдеша-Реньи [6]. В большинстве случаев используется в качестве показателя надежности сети вероятность связности соответствующего случайного графа в случае ненадежных ребер, отказывающих независимо, и абсолютно надежных узлов. Однако рассматриваются и другие показатели, например — вероятность связности с ограничением на диаметр [7], математическое ожидание числа несвязных пар вершин случайного графа [8] и другие.

Задачи точного расчета указанных показателей надежности являются NP -трудными. Самый очевидный способ расчета — это перебор всех 2^M реализаций графа, где M — количество ненадежных элементов. Однако практически этот метод не используется, в пользу приближенных методов или других точных методов, которые позволяют сократить, иногда значительно, объем вычислений с помощью различных приемов.

Работа поддержана РФФИ в рамках проекта № 18-07-00460.

В случае, когда элементы сети подвержены зависимым отказам [9–11], анализ надежности становится гораздо более трудоемкой задачей. В этих случаях используются, как правило, модели, основанные на марковских процессах. Обычно предполагается, что изначально элементы сети имеют определенные значения надежности, независимые друг от друга. Однако при отказе какого-либо элемента дальнейшие отказы элементов уже носят зависимый характер, т. е. отказ одного элемента провоцирует отказ других. Также предполагается, что отказы происходят за достаточно малые промежутки времени, так чтобы за один такой промежуток происходило не более одного.

Типичный пример зависимых отказов в сети — это каскадные отказы в электроэнергетических сетях [12–14], которые и исследуются в данной статье. Иницилирующее событие процесса распространения отказа определяется по внешним обстоятельствам: это может быть упавшее дерево, сильный порыв ветра, обрыв линии из-за перегрузки и т. д. Если же ее отказ вызвал перегрузки других линий или оборудования, то это, в свою очередь, может порождать новые отключения и т. д. — то есть, возникает последовательность зависимых отказов. Одним из значимых результатов в данной области стало построение различных видов так называемых графов зависимостей или графов влияния (*influence graph*) [14]. Вершины этого графа — ненадежные элементы исходной сети; дуга связывает две вершины, если отказ первого элемента приводит к отказу второго; вес дуги равен вероятности соответствующего отказа. Эти данные (вероятностей зависимых отказов) получаются в основном при помощи имитационного моделирования распространения отказов.

Ниже мы предлагаем новую модель распространения каскадных отказов, учитывая их локальный и нелокальный характер. В ее рамках предлагаются к рассмотрению несколько показателей надежности — вероятность связности сети, вероятность связности каждого потребителя с центром питания и другие. На основе аппарата теории надежности сетей разработаны методы точного и приближенного расчета надежности.

1. Модель распространения зависимых отказов. Важным свойством каскадных отключений в энергосетях является, как показывает практика, как их локальность, так и их нелокальность. На схеме (рис. 1) представлена последовательность отключений в Западной ЭЭС США, 1996 г. Первые три отказа были локальными — по смежным элементам сети, но уже четвертый — через множество устоявших линий, на расстоянии в несколько сотен километров. На этом примере также видно, что характерен отказ линий по разрезам, отсекающим определенные подсети. Например, отказы с номерами 8, 9, 10 или 12, 13, 14, 15, 16, 17, 20. Аналогичная картина была и при другом отказе на востоке США тем же летом.

Такие же закономерности обнаруживаются и при структурном анализе других каскадных отключений в ЭЭС. Объясняются подобные сценарии распространения каскадных отказов тем, что отказ линии электропередачи производит к практически мгновенному перераспределению электроэнергии на другие линии электропередачи, в первую очередь — на те линии, которые входят в разрез с отказавшей.

Трудоемкость задач анализа надежности сетей с зависимыми отказами, в том числе ЭЭС, растет, в общем случае, как факториал от числа ненадежных элементов. Для уменьшения вычислительных затрат предлагается сузить количество возможных вариантов развития каскадов — от всех, включая практически нереализуемые, до наиболее вероятных: по разрезам графа и по смежным элементам. При необходимости можно дополнить пространство возможных траекторий отказов также и другими вариантами, полученными, например, из статистических данных, экспертных оценок или имитационного

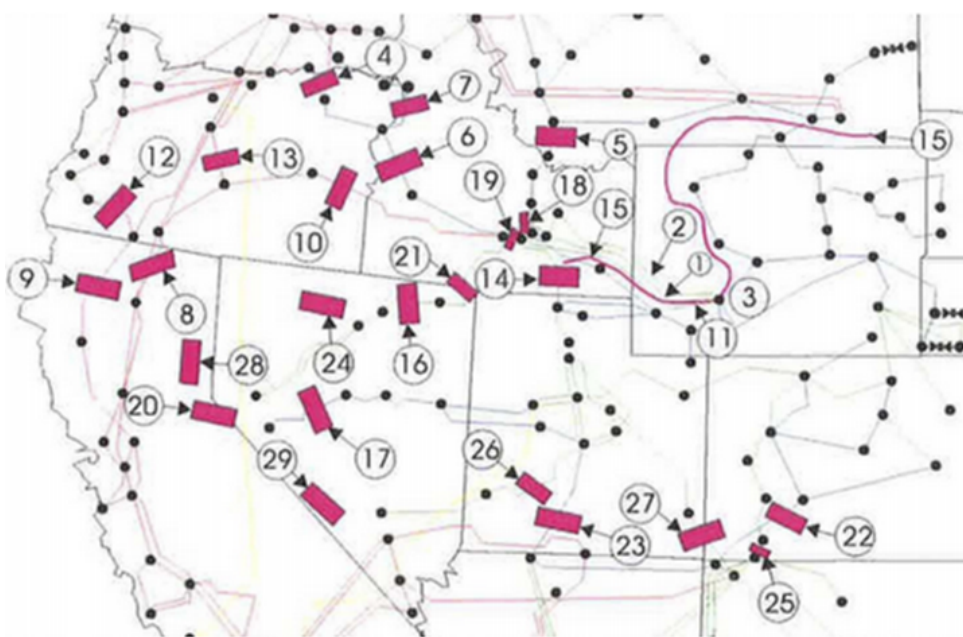


Рис. 1. Схема отключений западной ЭЭС США, 1996 г.

моделирования. С другой стороны, если пользоваться при анализе надежности ЭЭС графом влияний, подобная информация — разрезы графа и смежные элементы — может облегчить его формирование. Основная проблема при получении подобных графов — отсутствие необходимой статистики, которая бы описывала все варианты, в то время как для отказов по разрезам графа и по смежным элементам ЭЭС имеются хоть какие-то данные, в силу частоты развития именно таких сценариев. Кроме того, зная наиболее уязвимые места при выходе из строя какого-либо элемента, появляется возможность максимально уточнять значения соответствующих вероятностей — например, при помощи имитационного моделирования. В последнем случае нет необходимости рассмотрения всех вариантов распространения каскада, а только по разрезам и другим уязвимым местам, что может ускорить процесс имитационного моделирования (т. е. наработки статистики) и повышает его точность.

Также как и при использовании графа зависимостей, будем исходить из предположения, что вероятность выхода из строя элемента на i -м шаге процесса отказов зависит только от того, какой элемент вышел из строя на $(i - 1)$ -м шаге. При этом на первом шаге элементы считаются отказывающими независимо с вероятностями p_i . На каждом шаге отказывает только один элемент, либо не отказывает ни один вообще, в этом случае каскад считается законченным.

В качестве структурной модели ЭЭС будем использовать неориентированный граф $G = (V, E)$, где V — это множество вершин, а E — множество ребер графа G . Пусть для каждого ребра e задано вещественное число $0 \leq p_e \leq 1$, т. е. вероятность присутствия ребра e в графе. Будем интерпретировать это число как вероятность безотказной работы соответствующей линии электропередачи в течение заданного интервала времени. При этом предполагается, что вершины абсолютно надежны, т. е. присутствуют с вероятностью, равной 1.

Вероятность изначального отказа хотя бы одного элемента:

$$P(A) = 1 - \prod_1^M p_i.$$

Вероятность того, что при этом отказал элемент с номером i :

$$P(B_i) = \frac{q_i}{\sum_1^M q_i}, \quad q_i = 1 - p_i.$$

Следовательно, вероятность того, что первым отказал i -й элемент:

$$Q_i = P(A) P(B_i) = \frac{(1 - \prod_1^M p_i) q_i}{\sum_1^M q_i}.$$

Пусть задан граф зависимостей G' .

Зададим дискретное вероятностное пространство $W = (\Omega, P)$.

Ω — пространство элементарных событий (элементарных исходов), образованное множеством всевозможных путей $S_i = (l_1, \dots, l_n)$ без самопересечения в графе зависимостей G' , вероятность такого события будет определяться как

$$P(Q) = Q_e \prod_{n \geq l > 1} r_l,$$

где r_l — вес дуги, e — самая первая вершина в пути.

В качестве показателя надежности рассматривают, как правило, вероятность связности сети. В данной работе также предлагается рассматривать такой показатель как вероятность доступности какого-либо источника питания для каждого потребителя в течение определенного времени. Другой показатель — это вероятность доступности какого-либо источника питания для заданной доли потребителей. Такой показатель может быть более информативным, чем предыдущие, при рассмотрении ЭЭС большой размерности, например — в масштабах страны, или нескольких стран, если соответствующие сети объединены.

Все эти показатели формально определяются через соответствующие критерии успешности события. Например, вероятность связности вершин в графе G — вероятность того, что эти вершины связаны исправными ребрами, т. е. вероятность события, состоящего из всех элементарных событий, в которых вершины связаны, и только из них. Вероятность доступности какого-либо источника питания для каждого потребителя $R(G)$ — вероятность события, состоящего из всех элементарных событий, в которых каждая вершина-потребитель связана с какой-либо вершиной-центром питания, и только из них.

2. Метод факторизации в расчете показателей надежности и кумулятивном уточнении их верхней и нижней границ. Расчет надежности непосредственно по определению приведет к перебору всех реализаций графа, что делает расчет невозможным даже при небольшой размерности. Поэтому для расчета различных показателей надежности используются другие методы, самый распространенный из которых — метод факторизации (ветвления, Мура-Шеннона) [15]. Метод заключается в рекурсивном применении формулы полной вероятности при рассмотрении в качестве альтернативных гипотез наличия, либо отсутствия очередного разрешающего ребра. В нашем же случае главной особенностью будет наличие большего числа таких гипотез — по числу исходящих линий в графе влияний (для каждой линии нужно рассмотреть гипотезу ее отказа), плюс еще одной, по которой каскад прекратился. Далее для графа с удаленным ребром

также осуществляем факторизацию, и так далее, либо до получения варианта, где каскад прекратился (возвращается 1), либо до получения неуспешного графа (возвращается 0). Например, для расчета вероятности доступности какого-либо источника питания для каждого потребителя граф неуспешный, если какой-либо потребитель оказывается отрезанным от всех центров питания.

Введем ряд обозначений: $P(a)$ [Probability] — вероятность текущего события; TE [Target Element] — номер элемента, рассматриваемого на текущем шаге; $Adj(i)$ [Adjacent] — множество смежных с i элементов в графе зависимостей; FE [Failed Edges] — множество отказавших ребер; $P(a, b)$ — вероятность отказа элемента b после отказа a ; R — надежность сети.

Булева функция **Check(FE)** — проверяет граф, представляемый списком отказавших ребер **FE**, на критерий успешности. Ниже приведен пример псевдокода такой функции для наличия пути от каждой вершины-потребителя до одной из заданного наперед списка вершин-источников питания S :

```
function Check(FE)
   $G' := G - FE$ 
  for each  $i$  in  $E'$ 
    for each  $j <> i$  in  $S$ 
      if PathCheck( $i, j$ ) = true then
        return true
      end
    end
  end
  return false
```

Метод факторизации реализуется при помощи процедуры **Branching(FE, P, TE)**, которая вычисляет вероятность того, что, при заданных списке уже отказавших ребер, вероятности отказа и целевом элементе, сеть устояла:

```
procedure Branching(FE, P, TE)
   $R1 := 0$ 
  if Check(FE) = true then
    for each  $i$  in Adj(TE)
      if  $i$  not in FE then
        Branching(FE +  $i$ ,  $P * P'(TE, i)$ ,  $i$ )
         $R1 = R1 + P'(TE, i)$ 
      end
    end
   $R = R + P * (1 - R1)$ 
end
```

Тогда псевдокод общей программы расчета надежности будет иметь следующий вид:

```
procedure main()
   $R = 0$ 
  FE := { }
  for  $TE := 1$  to  $M$ 
    Branching(FE + TE,  $P'(TE)$ , TE)
  end
```

В [16] предлагается другой подход к задаче о надежности сети — ставится задача установить, превосходит ли надежность сети заданное значение (порога), т.е. является ли сеть достаточно надежной? Идея решения данной задачи проста. Каждый раз, когда в процессе ветвления получается подграф, для которого можно вычислить надежность непосредственно, мы уточняем верхнюю и нижнюю границу исходной сети и, если одно из полученных значений пересекает значение порога, принимаем решение о надежности или ненадежности сети. Подход, названный кумулятивным уточнением границ, был далее развит, например, в [17].

Опишем, как осуществлять подобное кумулятивное уточнение границ для рассматриваемых показателей надежности в случае наличия зависимых отказов. Изначально они иницируются как $RL = 0$, $RU = 1$. В теле псевдокода процедуры *Branching* при вызове функции *Check* в случае отрицательного ответа необходимо изменить верхнюю грань: $RU = RU - P$. Нижняя грань уточняется вместе с уточнением точного значения надежности: $RL = RL + P * (1 - R1)$. Используя такой подход, возможно значительно ускорить оптимизацию структуры ЭЭС по критерию надежности. При эвристическом нахождении мест для наиболее надежного размещения центров питания появляется возможность раннего отсеивания неподходящих вариантов.

3. Приближенный расчет на основе метода Монте-Карло. Для оценки различных показателей надежности широко используются методы Монте-Карло [3, 18], способные дать приближенное решение для сколь угодно трудоемких задач.

Опишем общую схему применения методов Монте-Карло к рассматриваемым задачам. Задается характеристическая функция

$$\chi(E) = \begin{cases} 1, & E \text{ — успешное событие,} \\ 0, & \text{— в противном случае.} \end{cases}$$

E — элементарное событие, соответствующие частной реализации сети после отказа определенных элементов, в данном случае — после каскадного отключения ребер. Факт успешности E определяется возможностью соответствующей сети успешно функционировать, исходя из практических соображений. Как правило, это определяется различными характеристиками связности. В данном случае E успешно, если в соответствующей сети остается работоспособный путь от каждого потребителя к хотя бы одному из поставщиков ресурса (центру питания).

Тогда вероятность успешного функционирования сети аппроксимируется как:

$$R = M[\chi] \approx \frac{1}{L} \sum_{i=1}^L \chi_i, \quad (1)$$

где L — количество выбранных событий.

Согласно ЦПТ, с ростом L вычисляемое приближение R становится все ближе к точному значению:

$$P \left\{ \left| R - \frac{1}{L} \sum_{i=1}^L \chi_i \right| \leq 3\sigma \right\} \rightarrow 0,9973. \quad (2)$$

Дисперсия случайной величины равна:

$$D(\chi) = M(\chi^2) - M^2(\chi) = R - R^2.$$

Соответственно, дисперсия оценки надежности (среднего значения по выборке) равна:

$$D(R_L) = \frac{L \cdot D(\chi)}{L^2} = \frac{R(1-R)}{L}.$$

А утроенное среднеквадратичное отклонение, которое мы используем для оценки погрешности ε (2), оценивается сверху как:

$$\varepsilon = 3\sigma = 3\sqrt{\frac{R(1-R)}{L}} \leq 3\sqrt{\frac{1}{4L}} = \frac{3}{2\sqrt{L}}. \quad (3)$$

Неравенство в центре в выражении (3) имеет место, так как $0 \leq R \leq 1$.

Таким образом, для количества итераций статистического моделирования

$$L = \frac{9}{4\varepsilon^2}, \quad (4)$$

неравенство (3) выполняется и получаемая оценка надежности (1) верна с погрешностью ε с вероятностью 0,9973 (2).

4. Численные эксперименты. В качестве тестового примера был взят граф одной электросети во Франции [19] (рис. 2), содержащий 141 вершины и 216 ребер. На рисунке слева синей линией обозначен один из разрезов, по трем ребрам, который отделяет 17 вершин. В случае отказа одной из этих трех линий энергопередачи нагрузка перераспределится на две оставшиеся, что, в свою очередь, с высокой вероятностью может вызвать отказы и в них. В зависимости от конфигурации (мощности станций, текущего потребления и пр.) без электричества могут остаться все потребители как в меньшем подграфе, так и в большем, или в различных частях и те, и другие. В качестве входных данных было взято разбиение графа на 7 подграфов (рис. 2) в предположении, что распространение каскадного отказа идет в группах ребер между каждой парой подграфов.

Для экспериментов использовался ПК с процессором Intel Xeon E5450 2,7 GHz. В качестве показателя надежности рассматривалась вероятность существования пути от каждого узла к источнику. Изначальная надежность каждого ребра полагалась 0,99, выбор же дальнейшего развития отказа или его остановки полагался равновероятным по всем вариантам в соответствии с моделью. При расчетах также учитывалось ограничение d на глубину каскадного отказа. Рост значения этого ограничения приводил к значительному увеличению времени расчета. Так, для $d = 8$ точный расчет надежности занял 0,24 с, для $d = 10$ — 1 с, $d = 12$ — 4 с. При ограничении 16 время расчета составило уже 70 с, полученное значение надежности: 0,787863. Оценка методом Монте-Карло дала близкое значение: 0,786287 за 0,2 с.

Заключение. Основной результат статьи — это модель распространения каскадных (зависимых) отказов элементов в сети по ее разрезам. Развитие подобных сценариев характерно для электроэнергетических сетей, когда первый отказ линии, спровоцированный внешними или внутренними причинами, может приводить практически к мгновенной перегрузке и отказу других линий. В качестве основного объекта для описания сценариев отказов был взят граф зависимостей, информация в котором может быть уточнена и дополнена подобными зависимыми отказами по структурным разрезам сети.

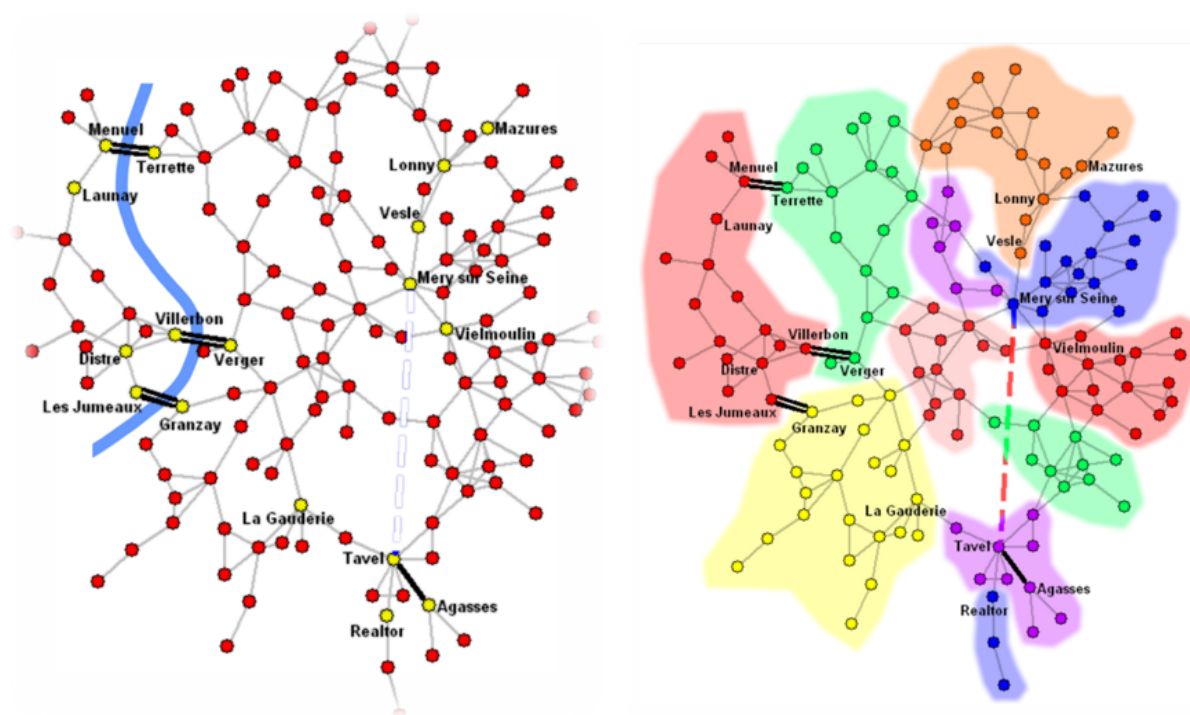


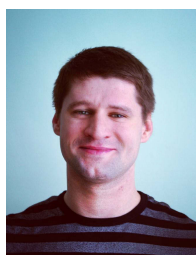
Рис. 2. Пример графа электросети и его структурные компоненты

В рамках данной модели были введены в рассмотрение различные показатели надежности сети. Предложены алгоритм точного их расчета, на основе использования формулы полной вероятности, и алгоритм оценки, на основе метода Монте-Карло. Также предложен метод кумулятивного уточнения границ показателей надежности, позволяющий делать заключение о достаточной надежности/ненадежности сети по отношению к заданному порогу.

Список литературы

1. Манов Н. А., Хохлов М. В., Чукреев Ю. Я., Шумилова Г. П., Успенский М. И., Чукреев М. Ю., Полуботко Д. В., Готман Н. Э., Старцева Т. Б. Методы и модели исследования надежности электроэнергетических систем. Коми научный центр УрО РАН. Сыктывкар, 2010.
2. Шилин А. Н., Сошинов А. Г., Елфимова О. И., Шилин А. А. Оценка надежности воздушных линий электропередачи по топологическим схемам с учетом влияния погодных условий для реализации информационно-измерительной системы. ВолгГТУ, КТИ (филиал) ВолгГТУ. Волгоград, 2017.
3. Krupenev D., Boyarkin D., Iakubovskii D. Improvement in the computational efficiency of a technique for assessing the reliability of electric power systems based on the Monte Carlo method // Reliability Engineering and System Safety. 2020. Vol. 204. ID: 107171. DOI: 10.1016/j.res.s.2020.107171.
4. Методические вопросы исследования надежности больших систем энергетики. Вып. 61. Проблемы исследования и обеспечения надежности либерализованных систем энергетики / Отв. ред. Н. И. Воропай, А. Д. Тевяшев. Иркутск: ИСЭМ СО РАН, 2011.
5. Kaplunovich P. A., Turitsyn K. S. Fast selection of $N - 2$ contingencies for online security assessment // Proc. of 2013 IEEE Power and Energy Society General Meeting (PES), 21–25 July 2013, Vancouver. IEEE Press. P. 1–5. DOI: 10.1109/PESMG.2013.6672792.

6. Colbourn Ch. J. The combinatorics of network reliability. N. Y., Oxford Univ. Press, 1987.
7. Cancela H., El Khadiri M., and Petingi L. Polynomial-time topological reductions that preserve the diameter constrained reliability of a communication network // IEEE Trans. on Reliability. 2011. Vol. 60, N 4. P. 845–851.
8. Родионов А. С., Родионова О. К. Кумулятивные оценки средней вероятности связности пары вершин случайного графа // Проблемы информатики. 2013. № 2. С. 3–12.
9. Родионов А. С., Хапугин С. А. Моделирование групповых отказов в анализе надежности сетей // Проблемы информатики. 2015. № 2. С. 31–43.
10. Chen W., Huang N., Kang R.. A reliability model with the dependent failures for telecommunication network // Proc. of the 8th International Conference on Reliability, Maintainability and Safety, Chengdu, 2009. IEEE Press. P. 1129–1132. DOI: 10.1109/ICRMS.2009.5270059.
11. Barrera J., Cancela H., Eduardo M. Topological optimization of reliable networks under dependent failures // Oper. Res. Let. 2015. 43(2). P. 132–136.
12. Chang L., Wub Z. Performance and reliability of electrical power grids under cascading failures // International Journal of Electrical Power & Energy Systems. Oct. 2011. Vol. 33. N 8. P. 1410–1419.
13. Hines P., Balasubramaniam K., Sanchez E. Cascading failures in power grids // IEEE Potentials. Sept. 2009. Vol. 28. N 5. P. 24–30.
14. Hines P. D. H., Dobson I., Rezaei P. Cascading power outages propagate locally in an influence graph that is not the actual grid topology // IEEE Transactions on Power Systems. March 2017. Vol. 32. N 2. P. 958 — 967.
15. Page L. B., Perry J. E. A practical implementation of the factoring theorem for network reliability // IEEE Trans. on Reliability. 1998. Vol. 37. N 3. P. 259–267.
16. Won J.-M., Karray F. Cumulative update of all-terminal reliability for faster feasibility decision // IEEE Trans. Reliability. 2010. V. 59. N 3. P. 551–562.
17. Rodionov A., Migov D., and Rodionov O. Improvements in the efficiency of cumulative updating of all terminal network reliability // IEEE Trans. Reliability. 2012. V. 61. N 2. P. 460–465.
18. Мигов Д. А., Винс Д. В. Параллельная реализация и имитационное моделирование оценки надежности сети методом Монте-Карло // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2019. Т. 47. С. 66–74. (DOI: 10.17223/19988605/47/8).
19. Crucitti P., Latora V., Marchiori M. Locating critical lines in high-voltage electrical power grids // Fluctuation and Noise Letters. 2005. Vol. 5. N 2. P. 201–208. DOI 10.1142/S0219477505002562.



Мигов Денис Александрович — канд. физ.-мат. наук, старш. науч. сотр. Института вычислительной математики и математической геофизики СО РАН; 630090, Новосибирск; e-mail: mdinka@rav.sccc.ru.

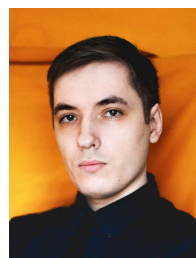
Денис Мигов окончил в 2003 году Механико-математический факультет Новосибирского государственного университета, получив квалификацию „Математик, системный программист“ по специальности „Прикладная математика и информатика“. В 2008 году защитил диссертацию „Расчет вероятности связности случайного графа с применением сечений“ на соискание ученой степени кандидата физико-математических наук по специаль-

ности 05.13.18 — „Математическое моделирование, численные методы и комплексы программ“ в диссертационном совете при Институте вычислительной математики и математической геофизики Сибирского отделения Российской академии наук. В настоящее время является старшим научным сотрудником лаборатории Системного моделирования и оптимизации указанного института. Денис Мигов является дважды лауреатом Именной премии правительства Новосибирской области в 2011 г. и в 2015 г., также неоднократно становился призером различных конференций. Область его научных интересов включает в себя теорию графов, методы анализа надежности сетей, структурную оптимизацию сетей, беспроводные сенсорные сети и параллельные алгоритмы на графах и сетях.

Denis Migov received the diploma of mathematician and programmer in applied mathematics and informatics from the Novosibirsk State University, Novosibirsk, Russia, in 2003. In 2008, he received Ph.D. (Candidate of science) degree in the field of Mathematical modeling, numerical methods, and program complexes from the Institute of Computational Mathematics and Mathematical Geophysics of the Siberian Branch of the Russian Academy of Sciences. He is a senior research fellow at the Laboratory of System modeling and optimization of the named institute. Denis Migov is twice a laureate of the Name Prize of the Government of the Novosibirsk Region in 2011 and in 2015; he also repeatedly became a prize-winner of various conferences. His scientific interests are in graph theory, network reliability analysis, network topology optimization, wireless sensor networks, and parallel algorithms on graphs and networks.

Коротков Антон Николаевич — ведущий программист ООО „Центр развития проектов „Петербургская Недвижимость“, 196066, Санкт-Петербург; e-mail: the.korotkov@gmail.com.

Антон Коротков окончил в 2016 году бакалавриат факультета информационных технологий Новосибирского государственного университета по направлению подготовки 09.03.01 „Информатика и вычислительная техника“. Тема дипломной работы: „Разработка математических методов и программных средств для расчета надежности сетей с зависимыми отказами элементов“. В настоящее время является ведущим программистом ООО „Центр развития проектов „Петербургская Недвижимость“.



Anton Korotkov graduated in 2016 with a bachelor's degree from the Faculty of Information Technologies of Novosibirsk State University in the field of 09.03.01 Informatics and Computer Engineering. Thesis topic: „Development of mathematical methods and software for calculating the reliability of networks with dependent element failures“. Currently, he is a leading programmer at the Center for Project Development Petersburg Real Estate.

Дата поступления — 27.11.2020

COGNITIVE MODEL FOR EVALUATING CUSTOMER EXPERIENCE IN THE STRUCTURE OF INFOCOMMUNICATION LANDSCAPE OF A TELECOM OPERATOR

V. A. Akishin

The Bonch-Bruевич Saint Petersburg State University of Telecommunications,
193232, Saint Petersburg, Russia

DOI: 10.24412/2073-0667-2021-3-34-55

Modern telecommunication market is highly competitive and, as a result, there is a high cost of attracting new customers. In addition, players in the telecommunications market have leveled off in terms of provided services and prices. That is why, telecom operators are looking for new development scenarios and sources of income.

Customer Experience Management (CEM) is one of the promising development direction that can provide a competitive advantage for a telecom operator in today's telecommunications market. The CEM concept focuses on the customer and its (CEM) goals are to build long-term relationships with the customer. Such relationships are based on formation positive customer experience and impressions from interaction with a company. In other words, the CEM concept focuses on working with loyalty and customer retention.

The implementation of the CEM concept assumes that all operational processes of the company should be built taking into account the total customer experience. In other words, all functional blocks of the company (even blocks, which does not work directly with the customer) should focus on the experience and impressions of the company's customers. Such a customer-centric model of managing business of a telecom operator is impossible without measuring customer experience in the context of the main operational processes and business functions. In other words, a telecom operator must understand how certain operational processes (and their KQI/KPI) affect the total customer experience.

To solve the above problems, it is required to develop a decision support system (DSS), which aimed on improving quality of management decision related to customer experience. Such system should be based on a cognitive model for assessing customer experience. The cognitive model should provide an opportunity, on the one hand, to assess the dependence of customer experience from the efficiency of the company's operational processes, and, on the other hand, to simulate optimal scenarios for managing customer experience on the telecom operator's network.

The main task of the study is development of a cognitive model for assessing customer experience in the context of managing operational processes and the infocommunication landscape of a domestic telecom operator. To solve this problem, the author sequentially analyzes the subject area and develops a cognitive model for assessing customer experience in the structure of the operational processes of a domestic telecom operator.

In the first part of the work, the author investigates existing research and approaches in the Customer Experience domain. Special attention is paid to the research of TM Forum. In particular, the thesaurus and the basic concepts that TM Forum uses when working on the topic of Customer

Experience are examined. In addition, TM Forum functional model for assessing the integral customer experience is considered — this model can provide an unified measure of CX for the telecommunications industry.

The second part of the work highlights the development of a model for assessing the integral customer experience based on fuzzy cognitive maps in the specifics of a domestic telecom operator. First of all, a functional representation of the model is formulated - value of the integral customer experience should be formed as a result of aggregation of metrics that affect the customer experience in the context of various dimensions of the company. Further, the issues of the mathematical basis of the model are highlighted, in particular, the basic calculations, which are required for solving research problem. For example, it is highlighted: calculation of integral customer experience in the context of the stages of the customer's lifecycle (based on the metrics of customer experience); calculation of integral customer experience for the entire customer lifecycle; calculation of the force of mutual influence between different factors of the model.

In the final part of the work, the author determines the practical value of the developed model (in particular, static and dynamic analysis of the model) for the tasks of managing the experience of the telecom operator's clients.

Key words: DSS, cognitive model, fuzzy cognitive maps, customer experience, operational processes.

References

1. GB962 Customer Experience Management Solution Suite. [Electron. Res.]: <https://www.tmforum.org/resources/suite/gb962-customer-experience-management-solution-suite-r17-5-0/> (date of access: 02.04.2020).
2. GB962 Customer Experience Management: Introduction and Fundamentals R16.0.1. [Electron. Res.]: <https://www.tmforum.org/resources/best-practice/gb962-customer-experience-management-introduction-and-fundamentals-r16-0-1/> (date of access: 02.04.2020).
3. GB962A Customer Experience Management Lifecycle Metrics R15.0.1. [Electron. Res.]: <https://www.tmforum.org/resources/best-practice/gb962a-customer-experience-management-lifecycle-metrics-r15-0-1/> (date of access: 30.04.2020).
4. RN341 Customer Experience Management Index (CEMI) Release Notes R2.0. [Electron. Res.]: <https://www.tmforum.org/resources/reference/rn341-customer-experience-management-index-cemi-release-notes-r2-0/> (date of access: 01.05.2020).
5. TMF066 Customer Experience Management Index Technical Specification V1.1. [Electron. Res.]: <https://www.tmforum.org/resources/technical-report-best-practice/tmf066-customer-experience-management-index-technical-specification-v1-1/> (date of access: 02.03.2020).
6. TR193 Customer Experience Management Index v1.3. [Electron. Res.]: <https://www.tmforum.org/resources/technical-report-best-practice/tr193-customer-experience-management-index-v1-3/> (date of access: 01.05.2020).
7. Akishin V. A., Kislyakov S. V., Fenomenov M. A. // Funkczionalnaya arkhitektura CEM-kompleksa dlya vnedreniya v IT-landshaft krupnogo operatora svyazi // T-COMM: TELEKOMMUNIKACzII I TRANSPORT. 2016. N 10. S. 12–16.
8. Akishin V. A., Kislyakov S. V., Fenomenov M. A. // Polzovatel'skij opyt v kognitivnoj modeli upravleniya setyu operatora // T-comm. 2016. T. 10, N 10.
9. Goldshtejn A. B. O kognitivnykh kartakh v upravlenii telekommunikacionnym operatorom / A. B. Gol'dshtejn, N. A. Pozharskij, D. A. Likhachev // Informatizacziyaisvyaz'. 2016. N 1.
10. Zagranovskaya A.V. Sistemnyj analiz na osnove nechetkikh kognitivnykh kart. Vestnik Rossijskogo ekonomicheskogo universiteta imeni G. V. Plekhanova. 2018. N 4. P. 152–160.

11. Marenko M., Malczeva M. Primenenie kognitivnogo modelirovaniya dlya analiza problem malogo biznesa // *Izvestiya Irkutskoj gosudarstvennoj e'konomicheskoy akademii*, 2015. T. 25. N 6.
12. Rotshtejn A. P. Nechetkie kognitivnye karty v analize nadezhnosti sistem // *Nadezhnost'*. 2019. N 19(4). P. 24–31.
13. Akishin V., Goldstein A., Goldstein B. Cognitive Models for Access Network Management // Galinina O., Andreev S., Balandin S., Koucheryavy Y. (eds) *Internet of Things, Smart Spaces, and Next Generation Networks and Systems. NEW2AN 2017, ruSMART 2017, NsCC 2017. Lecture Notes in Computer Science*. Springer, Cham. 2017. V. 10531. P. 375–381 (Scopus).
14. Axelrod R. *Structure of Decision. The Cognitive Maps of Political Elites*. Princeton University Press, 1976.
15. Kosko B. Fuzzy systems as universal approximators // *IEEE Transactions on Computers*, November 1994. V. 43. N 11. P. 1329–1333.
16. Gulakov V. K. SPPR na osnove kognitivnogo modelirovaniya „IGLA“ / V. K. Gulakov, D. G. Lagerev, A. G. Podvesovskij. M.: Programmnyeproduktyisistemy, 2007.
17. Groumpos, P. Fuzzy cognitive maps: basic theories and their application to complex systems. / Glykas, M. (ed.) *Fuzzy Cognitive Maps: Advances in Theory, Methodologies, Tools and Applications*, Springer, Berlin, 2010, P. 1–22.
18. Kosko B. Fuzzy Cognitive Maps // *International Journal of Man-Machine Studies*. 1986. N. 24. P. 65–75.

КОГНИТИВНАЯ МОДЕЛЬ ОЦЕНКИ КЛИЕНТСКОГО ОПЫТА В СТРУКТУРЕ ИНФОКОММУНИКАЦИОННОГО ЛАНДШАФТА ОПЕРАТОРА СВЯЗИ

В. А. Акишин

СПбГУТ им. проф. М. А. Бонч-Бруевича,
193232, Санкт-Петербург, Россия

УДК 65.011.56

DOI: 10.24412/2073-0667-2021-3-34-55

В статье освещаются вопросы оценки клиентского опыта в контексте управления операционными процессами и инфокоммуникационным ландшафтом отечественного оператора связи. Актуальность исследования обусловлена потребностью субъектов управления в системе поддержки принятия решений, которая позволит, с одной стороны, оценивать зависимость клиентского опыта от эффективности операционных процессов компании, а, с другой стороны, моделировать оптимальные сценарии управления клиентским опытом на сети оператора связи. Задачей исследования является разработка когнитивной модели оценки клиентского опыта, которая станет основой для системы, ориентированной на поддержку принятия управленческих решений (СППР), направленных на улучшение опыта клиентов оператора связи.

Для решения обозначенной задачи автором последовательно проводятся анализ предметной области и разработка когнитивной модели оценки опыта клиентов в структуре операционных процессов отечественного оператора связи. В рамках исследования предметной области автор проводит анализ совокупности факторов, формирующих клиентский опыт, а также подходов к его оценке в телекоммуникационной отрасли. В контексте разработки когнитивной модели автором предлагается функциональная модель оценки клиентского опыта, а также ее математическая основа на базе аппарата нечетких когнитивных карт.

Ключевые слова: СППР, когнитивная модель, нечеткие когнитивные карты, клиентский опыт, операционные процессы.

Введение. Современный рынок телекоммуникационных услуг характеризуется высокой конкуренцией и, как следствие, высокой стоимостью привлечения новых клиентов. Кроме того, крупные игроки на рынке телекоммуникационных услуг в последние годы практически сравнялись по уровню цен и качеству предоставляемого сервиса. Очевидно, что в подобных условиях трансформируется и клиент оператора связи — современные клиенты становятся все более требовательными к качеству услуг и процессов обслуживания, а также ожидают от компании определенного „предвосхищения“ их потребностей. Качество опыта и впечатлений, которые клиент получает в рамках взаимодействия с оператором, становится ключевым фактором, определяющим то, останется ли клиент с компанией или уйдет к конкуренту.

Концепция Customer Experience Management (CEM) или управление клиентским опытом является одним из наиболее перспективных трендов развития на современном телекоммуникационном рынке. Центральным понятием концепции CEM является

CustomerExperience (CX) — клиентский опыт. Клиентский опыт — это некоторая мера опыта и впечатлений клиента от компании, полученных на протяжении всего цикла взаимодействия клиента с компанией. Внедрение концепции СЕМ предполагает, что все операционные процессы компании должны строиться с учетом максимизации совокупного клиентского опыта. Иными словами, ориентироваться на клиента и его опыт должны не только подразделения, которые напрямую работают с клиентом, но и те функциональные блоки компании, деятельность которых не связана напрямую с клиентом, например, эксплуатационные подразделения оператора связи, которые отвечают за функционирование и эксплуатацию сети.

Очевидно, что подобная клиентоцентричная модель управления структурой бизнеса оператора связи невозможна без измерения клиентского опыта в разрезе основных операционных процессов и бизнес-функций. Иными словами, для внедрения концепции СЕМ оператор связи должен понимать, каким образом те или иные операционные процессы (и их KQI/KPI) влияют на совокупный клиентский опыт. Так, например, в точке принятия управленческих решений о модернизации процессов взаимодействия с клиентом, или разработке новых продуктов, или планировании строительства сети субъект управления должен оперировать информацией о том, каким образом его решения влияют на опыт клиентов — в частности, какие точки контакта с клиентом порождают негативную динамику клиентского опыта, какие сотрудники и подразделения недостаточно клиентоориентированы, какие продукты и сервисы формируют наилучший клиентский опыт, какая мера корреляции существует между основными показателями функционирования сети связи и клиентским опытом и т. д. [1].

Для того чтобы иметь возможность отвечать на подобные вопросы, оператор связи должен решить задачу оценки интегрального клиентского опыта как некоторой совокупной меры опыта и впечатлений клиента от компании. При этом оценка интегрального клиентского опыта должна формироваться на основе вертикальной агрегации данных о клиентском опыте на разных уровнях иерархии компании: от конкретных сотрудников до подразделений и региональных филиалов, от конкретных точек контакта и каналов взаимодействия с клиентом до сквозных процессов продаж и обслуживания. В этом случае субъект управления получает модель, которая, с одной стороны, дает возможность выявлять проблемные области с точки зрения клиентского опыта, а, с другой стороны, позволяет выстраивать основные операционные процессы и взаимодействие с клиентами таким образом, чтобы максимизировать положительный клиентский опыт.

Основная цель представленной работы — построение когнитивной модели оценки клиентского опыта в структуре процессов оператора связи. В контексте работы рассматриваются:

- Основные понятия предметной области СЕМ, необходимые для построения модели.
- Вопросы разработки модели оценки клиентского опыта, в т. ч. ее функциональное и математическое представление. В качестве математической основы для модели используются нечеткие когнитивные карты.
- Потенциальная область применения модели в структуре бизнеса отечественного оператора связи.

1. Предметная область СЕМ и подходы к оценке CX. Наиболее системный тезаурус для описания предметной области СЕМ можно сформировать на основе исследований, представленных организацией TMForum. Так, в 2013–2017 годах TMForum опубликовал ряд исследований, в которых формулируется видение концепции СЕМ в телеком-

муникациях и, в частности, модель оценки клиентского опыта применительно к структуре бизнеса оператора связи [1–6]. Ниже представлены основные понятия, составляющие основу предметной области СЕМ в области телекоммуникаций.

Основой концепции СЕМ является понятие CustomerExperience (CX) — клиентский опыт. Понятие клиентский опыт TMForum формализует как совокупность ощущений и впечатлений, которые получает клиент на протяжении всего цикла взаимодействия с компанией. Понятие CustomerExperienceManagement (СЕМ), в свою очередь, определяется как стратегия компании, позволяющая трансформировать внутренние процессы и образ мышления компании таким образом, чтобы она стала максимально клиентоориентированной. По мнению экспертов TMForum, концепция СЕМ должна формировать модель управления бизнесом таким образом, чтобы потребности клиента и взгляд на проблему „с точки зрения клиента“ стали отправной точкой при организации внутренних операционных процессов и формировании продуктов/сервисов. Понятия CX и СЕМ также рассматриваются экспертами TMF в контексте таких доменов как CustomerExperienceManagement, SocialMedia, ServiceQualityManagement, CustomerInteractionManagement — отсюда можно сделать вывод, что СЕМ является „обобщающим“ понятием, которое так или иначе связано с любым доменом, в котором присутствует понятие Customer [2].

Одной из важнейших составляющих концепции СЕМ является понятие жизненного цикла клиента [7, 2, 3]. Под жизненным циклом понимается совокупность этапов (также встречается термин „фазы“) взаимодействия клиента с компанией, которые проходит клиент последовательно или параллельно, начиная с момента появления у клиента осведомленности о компании и заканчивая точкой прекращения взаимоотношений: „BeAware“ — осведомленность; „Interact“ — начало взаимодействия; „Choose“ — выбор; „Consume“ — потребление; „Manage“ — управление; „Pay“ — оплата; „Renew“ — обновление соглашения на использование сервиса; „Recommend“ — рекомендации; „Leave“ — окончание использования.

Следует отметить, что на каждом из этапов жизненного цикла у клиента формируется определенный опыт от взаимодействия с компанией. Как следствие, совокупный клиентский опыт для всего жизненного цикла клиента может быть представлен как агрегация опыта на различных этапах.

В качестве меры для оценки клиентского опыта эксперты TM Forum вводят такие понятия как CustomerExperienceLifecycleMetrics (далее по тексту метрики CX) и CE Index. CE Index рассчитывается на основе метрик CX. Данные понятия являются основополагающими в данном исследовании, поэтому рассмотрим их подробнее.

Метрики CX — это совокупность влияющих на клиентский опыт показателей, определенных для каждого этапа жизненного цикла клиента [3]. В общем случае TMForum выделяет 3 основных типа CX метрик:

- PerCustomerMetrics — показатели, определяющие опыт конкретного клиента в точках касания с компанией. По сути своей данные показатели определяют „взгляд снаружи на компанию“, иными словами, то, как со стороны клиента выглядят основные интерфейсы и операционные процессы компании. Примером PerCustomerMetrics может быть, например, время ожидания клиентом ответа в контактном центре.

- PerEnterpriseCustomerMetrics — показатели, определяющие опыт конкретных клиентов корпоративного сегмента. По своему функциональному смыслу данный тип метрик идентичен PerCustomerMetrics, однако накладывается специфика взаимодействия с кор-

поративными клиентами, например, специфичные персональные SLA, выделенное сопровождение и т. д.

— FunctionalMetrics — показатели, связанные с внутренними операционными процессами компании, которые оказывают непосредственное влияние на клиентский опыт. Однако, как правило, подобные метрики далеко не всегда можно измерить в контексте конкретного клиента. Ключевое отличие FunctionalMetrics от PerCustomer (Enterprise) Metrics состоит в том, что FunctionalMetrics определяют „взгляд изнутри“ на понятие клиентского опыта и влияющие на него операционные процессы. Примером FunctionalMetrics может быть доля неуспешных (незавершенных) контактов с клиентом.

Исходя из своего определения, типы метрик „PerCustomerMetrics“ и „PerEnterpriseCustomerMetrics“ измеряются в контексте точек контакта или каналов взаимодействия с клиентом, например, web-сайт, офис продаж и обслуживания, IVR, Call-центр, социальные сети и медиапространства и т. д. В свою очередь, FunctionalMetrics рассматриваются в контексте основных функций предприятия, в т. ч. клиентский сервис, клиентская поддержка, сеть связи, биллинг, продажи, маркетинг, управление персоналом, а также безопасность.

Описанные выше метрики CX могут быть сформированы на основе данных из информационного ландшафта компании, результатов исследования мнений клиентов, а также „внешних“ данных, характеризующих состояние рынка или конкретного бренда.

Отдельного внимания заслуживает такой источник данных как исследование мнений клиентов — существенная часть метрик CX формируется именно на основе субъективной оценки, которую клиент дает компании посредством ответа на вопрос. В общем случае эксперты TMForum предлагают:

— Методы для исследования абсолютных значений субъективной клиентской оценки некоторого измерения компании (например, оценка клиентом некоторого продукта, канала взаимодействия, сотрудника и т. д.). Для этого используются вопросы, в которых клиенту предлагается оценить некоторое измерение по шкале от 1 до 6, где „1“ — „полностью не удовлетворен или полностью не согласен“, а „6“ — „полностью удовлетворен или полностью согласен“.

— Методы для исследования субъективной оценки клиентов в контексте сравнения некоторых измерений (например, цен на услуги компании и ее конкурентов). Для этого используются вопросы, постановка которых предполагает сравнение по шкале от 1 до 6, где „1“ — „намного хуже“, а „6“ — „намного лучше“.

Метрики CX и структура их описания являются отправной точкой для измерения индекса клиентского опыта.

Индекс клиентского опыта или CE (CX) Index — это количественная мера, позволяющая определить некоторое значение совокупного опыта клиента, основываясь на измерении показателей функционирования различных бизнес-доменов компании, влияющих на клиентский опыт [4–6]. Ключевыми понятиями, на основе которых рассчитывается CE Index, являются понятия KeyPerformanceIndicator (KPI) и KeyQualityIndicator (KQI), которые в своей основе используют описанную выше модель CX метрик.

KeyPerformanceIndicators — это совокупность показателей, которые в данном контексте позволяют измерять меру эффективности функционирования некоторой бизнес-единицы (например, процесса, услуги, элемента сети), которая, в свою очередь, влияет на клиентский опыт. KeyQualityIndicators представляет собой более обобщенное понятие, которое определяет некоторую меру качества функционирования определенной бизнес-единицы —

в частности, TMForum предлагает рассматривать KQI в разрезе точек контакта и каналов взаимодействия с клиентами. По сути своей, KQI формируются на основе набора KPI — так, например, для оценки качества (KQI) функционирования контактного центра используется сумма оценок по набору KPI, которые определяют эффективность работы операторов, эффективность IVR, загрузку входящего канала и т. д.

В ряде своих исследований эксперты TMForum предлагают функциональную модель для оценки CE Index на основе расчета значений KPI и KQI, которые, в свою очередь, являются производными от CX метрик [4–6]. Функциональная модель расчета CE Index построена на принципе агрегации данных от частного к общему. Модель состоит из трех уровней:

- Уровень 3 (L3) — данный уровень описывает сценарии сбора исходных данных для модели, а также то, каким образом рассчитываются показатели качества KQI и показатели эффективности KPI для операционных процессов, влияющих на формирование клиентского опыта.

- Уровень 2 (L2) — данный уровень описывает то, каким образом рассчитываются показатели качества (KQI) для конкретного канала взаимодействия или точки контакта с клиентом на основе результатов KPI/KQI, полученных на уровне 3. Далее полученные значения качества канала взаимодействия или точки контакта „калибруются“ с помощью субъективных оценок клиентами данных измерений (которые также формализованы в модели CX метрик).

- Уровень 3 (L3) — данный уровень описывает то, каким образом рассчитывается индекс клиентского опыта в контексте определенного этапа жизненного цикла клиента, а также совокупное значение в контексте всего жизненного цикла клиента.

2. Постановка задачи на разработку модели и исходные данные. Исходя из тезисов, сформулированных во вступлении, а также основываясь на анализе модели TMForum, можно сформулировать следующие задачи, которые должна решать разрабатываемая модель оценки интегрального клиентского опыта:

- Определение единой меры для измерения понятия интегрального клиентского опыта, а также показателей, влияющих на значение интегрального клиентского опыта;

- Расчет значения интегрального клиентского опыта (интегральный CX — аналог CE Index) в момент времени в разрезе таких измерений как этап жизненного цикла клиента, точка контакта, сегмент клиента, конкретный клиент

- Анализ взаимовлияния между значением интегрального клиентского опыта и факторов, на него влияющих.

В качестве исходных данных для разработки модели принимается структура данных отечественного оператора связи, включающая в себя совокупность показателей, которые, по мнению экспертов, влияют на клиентский опыт. При этом исходная структура данных адаптируется под условно эталонную структуру модели CX метрик, предложенную TMForum, в частности:

Во-первых, выделяются 2 типа показателей:

- Показатели, связанные с конкретным клиентом и характеризующие опыт конкретного клиента (аналог PerCustomerMetrics). Например, время ожидания клиентом ответа специалиста при обращении в КЦ или количество перерывов связи по конкретной услуге клиента.

- Показатели, характеризующие операционную деятельность компании, не связанные с конкретным клиентом, однако, оказывающие непосредственное влияние на клиентский

опыт (аналог FunctionalMetrics). Например, степень загруженности выездных бригад, проводящих работы у клиента в процессах подключения или технической поддержки.

Во-вторых, показатели группируются в разрезе таких измерений как:

- Точки контакта с клиентом;
- Этапы жизненного цикла клиента (согласно модели, предложенной TM Forum);
- Сегмент клиента.

В целях унификации используемой терминологии описанные выше показатели, приведенные к соответствующей структуре, далее по тексту будут называться „метрики клиентского опыта“ (по аналогии с тезаурусом TMForum). Для проведения исследования были отобраны 52 метрики клиентского опыта, которые были соотнесены со структурой данных из В/OSS среды двух отечественных операторов связи. Ниже приведены некоторые примеры данных метрик:

— Средняя оценка клиентом сотрудников Оператора, с которыми он имел взаимодействие — сформированная на основе опроса степень удовлетворенности клиента сотрудниками, с которыми он взаимодействовал (этап Interact).

— Количество заявок на подключение, не выполненных в запрашиваемый клиентом период — количество заявок на подключение, выполненных во временной интервал, указанный клиентом при оформлении заявки (этап Choose).

— Оценка клиентом качества услуги интернет — сформированная на основе опроса степень удовлетворенности клиента качеством предоставления услуги интернет (этап Consume).

— Количество незаявленных перерывов связи и/или их деградации — количество незаявленных клиентов перерывов в функционировании услуг связи по причине плановых ремонтных работ или групповых повреждений на сети связи (этап Consume).

— Количество клиентских инцидентов, связанных с работоспособностью услуги — количество клиентских инцидентов, связанных с прекращением функционирования или частичной деградации услуг связи (этап Manage).

— Оценка клиентом прозрачности выставляемого компанией счета на оплату услуг — сформированная на основе опроса степень удовлетворенности клиента понятностью и обоснованностью выставляемого счета (этап Pay).

— Доля клиентов, обновивших соглашение — доля клиентов, перезаключивших договор с компанией относительно активных клиентов (этап Renew).

— Доля принятых RealTimeOffers — доля принятых клиентом продуктовых предложений (upsale, crossale), сделанных клиенту в течение его жизненного цикла, начиная с этапа Consumer (этап Renew).

— Доля положительных комментариев в соц. медиа пространствах — доля положительных комментариев в социальных сетях (ВКонтакте, Одноклассники, FB) на официальных страницах Оператора от общего количества (этап Recommend).

Таким образом, входные данные, с которыми будет работать модель оценки клиентского опыта, с одной стороны, основываются на реальных данных отечественного оператора связи, а, с другой стороны, она приближена к условно эталонной модели учета CX метрик, сформулированной TMForum.

3. Функциональное представление модели. В первую очередь, определим функциональный смысл модели. Оценка значения интегрального клиентского опыта должна формироваться как результат агрегации значений метрик, влияющих на клиентский опыт в разрезе различных измерений компании. Опираясь на исследования TMForum [3, 5],

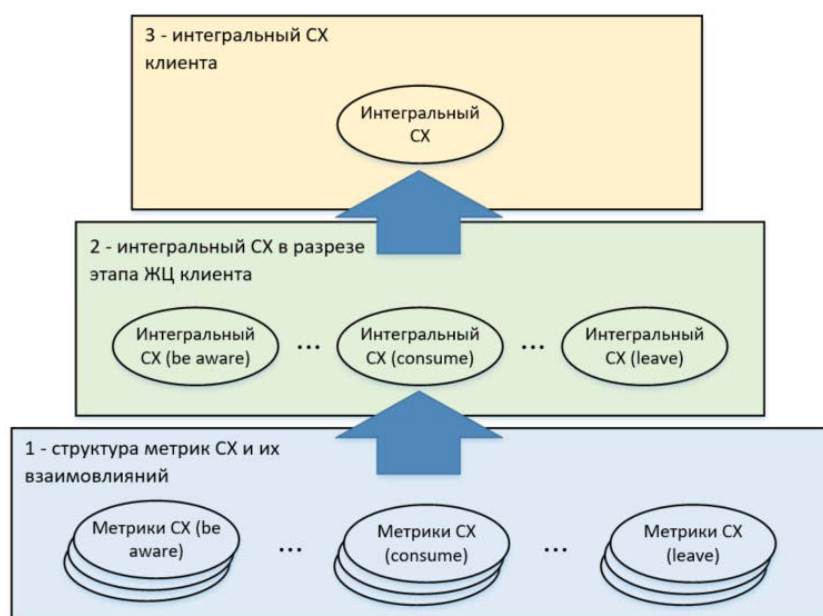


Рис. 1. Функциональная структура модели

функциональный смысл модели можно описать в виде структуры из трех уровней, где каждый из уровней определяет вычисления значений клиентского опыта разного уровня абстракции — от атомарных метрик клиентского опыта до интегральных значений клиентского опыта:

— Уровень 1 — уровень метрик клиентского опыта. Данный уровень определяет исходную структуру и взаимосвязи метрик клиентского опыта, а также инициализацию исходных значений метрик клиентского опыта.

— Уровень 2 — уровень интегрального клиентского опыта в разрезе этапов жизненного цикла клиента. Данный уровень определяет то, каким образом вычисляется значение интегрального клиентского опыта для конкретного этапа жизненного цикла на основе связанных с данным этапом метрик клиентского опыта (уровень 1). В общем случае, вычисления интегрального CX на данном этапе могут формироваться не только в разрезе этапов жизненного цикла клиента, но и в разрезе конкретных точек контакта или каналов взаимодействия с клиентом.

— Уровень 3 — уровень интегрального клиентского опыта в разрезе исследуемого клиента или группы клиентов. Данный уровень определяет то, каким образом вычисляется значение интегрального клиентского опыта на основе его значений в разрезе конкретных этапов жизненного цикла клиента (уровень 2).

В зависимости от исходных данных, которые модель принимает на вход на 3 уровне, конечное значение интегрального CX может рассчитываться либо в контексте конкретного клиента, либо в контексте группы клиентов (например, сегментов B2B/B2C). Схематично функциональная архитектура модели представлена на рис. 1.

Описанная выше функциональная структура достаточно хорошо ложится на математический аппарат нечетких когнитивных карт. Выбор нечетких когнитивных карт в качестве математической основы для данной модели обусловлен исследованиями [8–15]. Таким образом, описанную выше функциональную модель можно представить в виде когнитив-

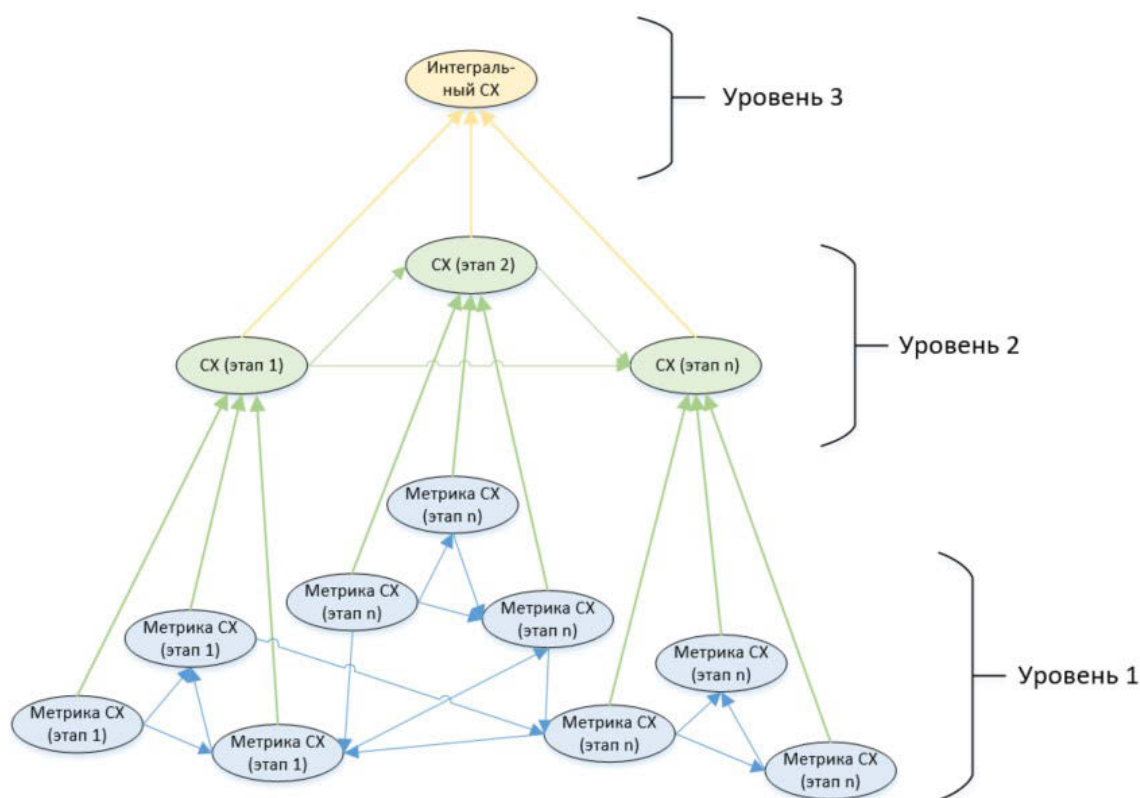


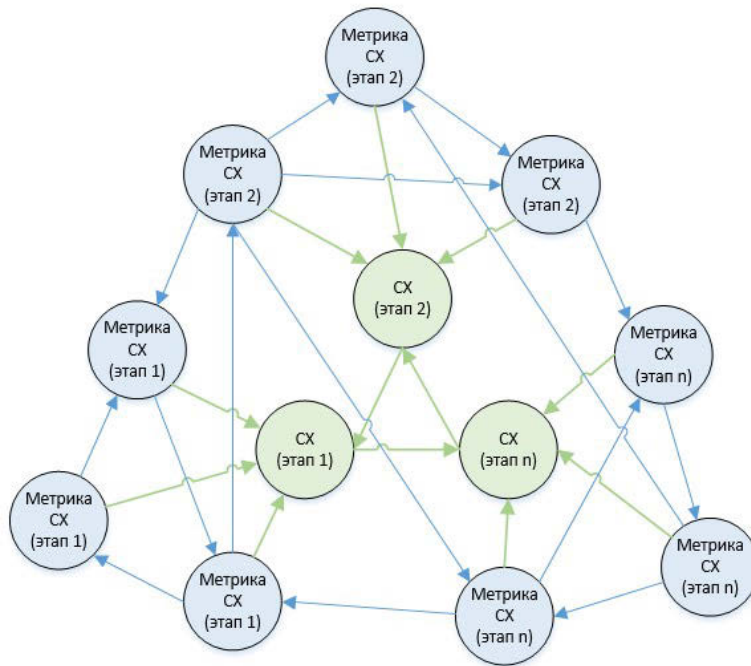
Рис. 2. Представление функциональной структуры модели в виде иерархической когнитивной карты

ной карты иерархической структуры, где каждый уровень иерархии будет соответствовать одному из уровней функциональной структуры модели (рис. 2).

Представленную на рис. 2 иерархическую модель можно декомпозировать на две независимо вычисляемые когнитивные карты.

Первая когнитивная карта G_1 будет описывать модель оценки интегрального CX для этапов жизненного цикла на основе метрик клиентского опыта. Управляющими факторами (концептами) данной когнитивной карты будут являться метрики клиентского опыта, а целевыми факторами будут являться значения интегрального CX в разрезе определенного этапа жизненного цикла клиента. Взаимовлияние факторов модели друг на друга будет характеризоваться некоторым весом и направлением, что позволит нам моделировать важность влияния метрик клиентского опыта друг на друга и на значение интегрального CX для этапа жизненного цикла.

Вторая когнитивная карта G_2 будет описывать модель оценки интегрального CX на основе значений клиентского опыта в разрезе этапов жизненного цикла. Управляющими факторами модели будут являться значения интегрального CX в разрезе этапа жизненного цикла клиента, а целевым фактором — конечное интегральное значение клиентского опыта. Как и в первом случае, взаимовлияние факторов модели друг на друга будет характеризоваться некоторым весом и направлением, что позволит нам моделировать важность влияния управляющих факторов друг на друга и на целевой фактор. Ключевым отличием второй когнитивной карты является наличие у управляющих факторов дополнительного

Рис. 3. Схема первой когнитивной карты G_1

веса, который моделирует сценарий накопления опыта клиента в разрезе этапа жизненного цикла — данный вес в модели будет представлен влиянием фактора на самого себя.

Математическим представлением когнитивных карт G_1 и G_2 будет являться векторный функциональный граф, который можно записать кортежем вида:

$$G = \langle V, E, A, W \rangle,$$

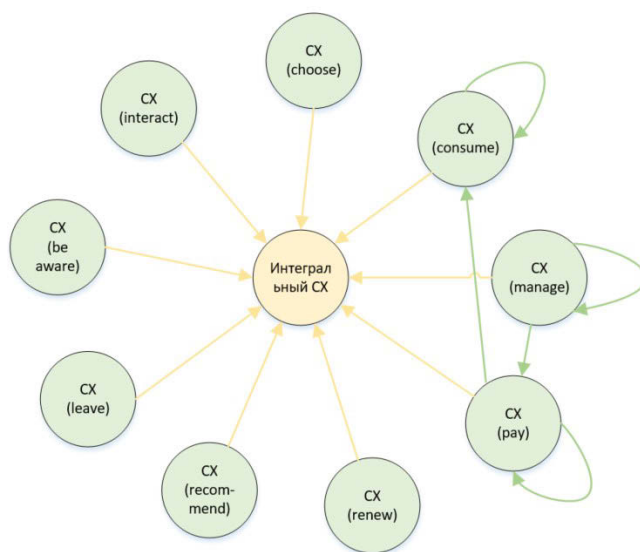
где V — множество вершин графа, моделирующих управляющие и целевые факторы когнитивной модели; E — множество дуг графа между вершинами v_i, v_j ($i = 0, 1, 2, \dots$; $j = 0, 1, 2, \dots$), моделирующих направленность и вес взаимосвязи между факторами модели; A — множество значений вершин графа (по сути, a_i представляет собой вектор из одного значения); W — множество значений степени влияния дуги e_{ij} ;

Далее рассмотрим принципы формирования значений (a_i) факторов когнитивной модели.

Значения факторов, формирующих когнитивное поле модели, могут формироваться на основе:

- Значений метрик, представляющих собой субъективные оценки, которые клиенты дают в качестве обратной связи компании (например, в рамках ответов на опросы);
- Значений метрик, представляющих собой объективные количественные показатели, которые формируются в операционной среде компании или среде взаимодействия с клиентом;
- Предрасчитанных значений интегрального клиентского опыта в разрезе определенных измерений.

Для того, чтобы нормализовать значения всех факторов, а также для того, чтобы упростить их восприятие, введем в нашу модель понятие лингвистической переменной, которая будет определяться кортежем вида:

Рис. 4. Схема второй когнитивной карты G_2

$$\langle a, T(a), U, G, M \rangle,$$

где a — название лингвистической переменной; $T(a)$ — обозначает терм-множество переменной a или, иными словами, множество названий лингвистических значений переменной a , при этом, каждое из таких значений является нечеткой переменной \tilde{a} со значениями из универсального множества U (т.е. которое является областью определения нечетких переменных, относящихся к лингвистической переменной); G представляет собой синтаксическое правило, порождающее название переменной a из множества $T(a)$; M — семантическое правило, которое позволяет определять для каждого нового значения (путем использования функции G) некоторое осмысленное содержание посредством формирования соответствующего нечеткого множества. Значения лингвистической переменной a будут соотноситься с значениями множества U , которое принадлежит области значений от -1 до 1 :

$$U \in [-1; 1].$$

В нашем случае терм-множество $T(a)$ будет состоять из конечного числа термов, для которых определяется конкретный диапазон значений нечеткой переменной множества U . Как следствие, функциональный смысл функции M будет сводиться к выставлению соответствия каждому элементу терм-мноества его смысла. Соответствие значения термина лингвистической переменной a и численных значений множества U отражено в таблице 1.

Таким образом, значение каждого фактора v_i нашей когнитивной модели будет определяться лингвистической переменной a_i — иными словами, значение каждого фактора в модели в момент времени будет представлено в виде значения „Максимальное отрицательное значение“, „Нейтральное значение“, „Максимальное положительное значение“ и т.д., что позволит унифицировано интерпретировать значение каждого из факторов когнитивной модели, несмотря на разнородность исходных данных, формирующих данное значение.

Для того чтобы преобразовать исходные данные для модели (метрики клиентского опыта) в лингвистическую переменную, будет использоваться процедура фазификации,

Таблица 1

Соответствие значения термина лингвистической переменной a
и численных значений множества U

Значения терм-множества для лингвистической переменной	Численное значение множества нечетких переменных U
Очень низкое	$[-1; -0,72]$
Низкое	$(-0,72; -0,44]$
Ниже среднего	$(-0,44; -0,16]$
Среднее	$(-0,16; 0,12]$
Выше среднего	$(0,12; 0,40]$
Высокое	$(0,40; 0,68]$
Очень высокое	$(0,68; 1]$

т.е. по сути приведение значений метрик к „нечеткости“. Это позволит преобразовать конкретные количественные значения метрик в термы лингвистической переменной. Для каждой метрики клиентского опыта, участвующей в моделировании, будет задаваться функция преобразования диапазона значения метрики в термы множества $T(x)$, которая будет определяться на основе экспертных знаний специалистов в предметной области. Например, имеется метрика „Количество клиентских инцидентов за период“, значения которой могут определяться в диапазоне $[0; \infty]$. Эксперты должны установить правила соответствия значений из данного диапазона значениям терм-множества, например: значение 0 соответствует терму „Очень низкое“, диапазон $(0; 1]$ соответствует терму „Низкое“ и т. д.

Далее определим принципы для описания взаимовлияния между факторами в модели.

Взаимовлияние между факторами модели (w_{ij}) характеризуется направленностью и весом влияния.

$$w_{ij} = w(v_i, v_j),$$

где v_i, v_j — факторы когнитивной карты, w — мера интенсивности (веса) влияния между факторами модели. Направленность взаимовлияния в когнитивной модели описывается следующим набором правил:

— Положительное влияние ($w_{ij} > 0$) — определяет влияние фактора v_i на фактор v_j таким образом, что изменение в большую сторону значения фактора v_i влечет за собой изменение в большую сторону значения фактора v_j , и наоборот: изменение в меньшую сторону значения фактора v_i влечет за собой изменение в меньшую сторону значения фактора v_j ;

— Отрицательное влияние ($w_{ij} < 0$) — определяет влияние фактора v_i на фактор v_j таким образом, что изменение в большую сторону значения фактора v_i влечет за собой изменение в меньшую сторону значения фактора v_j , и наоборот: изменение в меньшую сторону значения фактора v_i влечет за собой увеличение значения фактора v_j ;

— Нулевое влияние ($w_{ij} = 0$) — определяет отсутствие взаимовлияния (отношения причинности) между факторами v_i и v_j .

Для описания взаимовлияния w_{ij} между факторами v_i и v_j в нашей когнитивной модели также будет использоваться лингвистическая переменная по аналогии с описанием значений факторов (концептов) модели. В данном случае лингвистическая переменная будет описываться кортежем вида:

Таблица 2

Соответствие значения термина лингвистической переменной w и численных значений множества Z

Значения терм-множества для лингвистической переменной	Численное значение множества Z
Максимальное отрицательное влияние	$[-1; -0,75]$
Сильное отрицательное влияние	$(-0,75; -0,50]$
Отрицательное влияние	$(-0,50; -0,25]$
Слабое отрицательное влияние	$(-0,25; 0)$
Нейтральное влияние	0
Слабое положительное влияние	$(0; 0,25]$
Положительное влияние	$(0,25; 0,50]$
Сильное положительное влияние	$(0,50; 0,75]$
Максимальное положительное влияние	$(0,75; 1]$

$$e = \langle w, S(w), Z, H, R \rangle,$$

где w — название лингвистической переменной, $S(w)$ — обозначает терм-множество переменной y или, иными словами, множество названий лингвистических значений переменной w , при этом каждое из таких значений является нечеткой переменной \tilde{w} со значениями из универсального множества Z (т.е. которое является областью определения нечетких переменных, относящихся к лингвистической переменной). H представляет собой синтаксическое правило, порождающее название переменной w из множества $Z(w)$. R — семантическое правило, которое позволяет определять для каждого нового значения (путем использования функции H) некоторое осмысленное содержание посредством формирования соответствующего нечеткого множества. Данные значения переменной w будут соотноситься с значениями множества Z , которое принадлежит области значений от -1 до 1 :

$$Z \in [-1; 1].$$

Как и в случае с факторами, терм-множество $S(y)$ состоит из конечного числа термов, для которых определяется конкретный диапазон значений нечеткой переменной множества Z , как следствие, функциональный смысл функции R будет сводиться к выставлению соответствия каждому элементу терм-множества его смысла. Соответствие значения термина лингвистической переменной w и численных значений множества Z представлено в табл. 2.

Таким образом, значение каждой взаимосвязи e_i нашей когнитивной модели будет определяться лингвистической переменной w — иными словами, вес и направленность взаимосвязи факторов друг на друга будет описываться терминами („Максимальное отрицательное влияние“, „Нейтральное влияние“, „Максимальное положительное влияние“ и т.д.). Преимущество подобного подхода при оценке взаимовлияний факторов друг на друга будет заключаться в том, что для оценки взаимовлияния экспертами не будут использоваться числовые веса для каждой причинно-следственной связи.

При этом экспертная оценка взаимовлияний будет формироваться с помощью метода парных сравнений (автор Т. Саати), основная идея которого — обработка суждений

эксперта об относительном превосходстве степеней принадлежности различных элементов [16]. В данном случае формируется группа экспертов, где каждому эксперту будут задаваться три последовательных вопроса:

- Если значение фактора v_i увеличивается, то как ведет себя фактор v_j : увеличивается, уменьшается или остается неизменным?;
- Если фактор v_j увеличивается или уменьшается, то какова степень данного изменения в терминах лингвистической переменной?;
- Оцените степень своей уверенности в ответе на второй вопрос по шкале от 1 до 3, где „1“ — „не уверен или нет мнения“, „2“ — „скорее уверен, чем не уверен“, „3“ — „абсолютно уверен“.

Далее искомая величина w_{ij} оценивается как сумма оценок, взвешенная уверенностью экспертов:

$$w_{ij} = \frac{\alpha_1 w_{ij}^1 + \alpha_2 w_{ij}^2 + \dots + \alpha_l w_{ij}^l}{\alpha_1 + \alpha_2 + \dots + \alpha_l},$$

где w_{ij}^p — оценка силы влияния p -ым экспертом, выраженное в виде численного значения термина лингвистической переменной; α_p — степень уверенности p -ого эксперта в своей оценке; l — количество экспертов. Диапазон значений α_p определяется для каждого конкретного исследования, но по умолчанию используются следующие значения α_p : $\alpha_p = 0$, если эксперт „не уверен или нет мнения“, $\alpha_p = 0,5$, если эксперт „скорее уверен, чем не уверен“, $\alpha_p = 1$, если эксперт „абсолютно уверен“.

4. Математическая основа модели. Основываясь на сформированном выше функциональном описании когнитивной модели, рассмотрим основные вычисления, которые потребуются для решения задач, поставленных перед моделью в разделе 3, в т. ч.:

- Вычисление значения интегрального клиентского опыта в разрезе этапов жизненного цикла клиента на основе метрик клиентского опыта;
- Вычисление значения интегрального клиентского опыта для всего жизненного цикла клиента;
- Вычисление силы взаимовлияния факторов модели друг на друга.

Вычисление значения интегрального клиентского опыта в разрезе этапов жизненного цикла клиента на основе метрик клиентского опыта. Рассмотрим первую когнитивную карту G_1 , которая моделирует первый уровень иерархии в функциональной модели. Целевыми факторами данной модели будут являться значения интегрального клиентского опыта в разрезе этапа жизненного цикла клиента, а управляющими факторами будут являться метрики клиентского опыта, рассчитываемые в рамках данного этапа жизненного цикла. Когнитивная карта будет записываться в виде векторного функционального графа, вершины которого будут определяться с помощью значений лингвистической переменной, которые, в свою очередь, являются результатом функции фазификации для исходного значения метрики клиентского опыта.

$$G_1 = \langle V_1, E_1, A_1, W_1 \rangle,$$

где V_1 — множество вершин графа, которые являются факторами когнитивной модели G_1 ; E_1 — множество дуг графа, моделирующих взаимовлияние факторов модели G_1 ; A_1 — множество значений вершин графа v_i ($v_i \in V_1$; $i = 0, 1, 2, \dots$); W_1 — множество значений степени влияния дуги e_{ij} ($e_{ij} \in E_1$; $i = 0, 1, 2, \dots$; $j = 0, 1, 2, \dots$).

Значение вершины v_i запишем в виде численного значения лингвистической переменной a_i . В качестве начальных значений для концептов множества V_1 будет принято значение лингвистической переменной „Среднее“, что соответствует значению 0 в множестве нечетких значений.

Дуги графа описывают влияние факторов друг на друга. Оценка влияния фактора v_i на фактор v_j можно записать в виде матрицы W_1 (для удобства работы запишем не значения термов лингвистической переменной, а соответствующие им значения нечетких переменных множества Z):

		факторы					
		v_1	v_2	v_3	v_j
факторы	v_1	0	-0,25	-0,75	w_{1j}
	v_2	-0,25	0	0,25	w_{2j}
	v_3	1	0,75	0	w_{3j}
	0
	v_i	w_{i1}	w_{i2}	w_{i3}	0
	0

Смоделируем возмущение модели. Предположим, что необходимо вычислить значение целевого фактора v_i в момент времени t , при условии, что на связанный с v_i фактор v_j подается возмущение или, в терминах функциональной модели, изменяется значение некоторой метрики клиентского опыта. Тогда значение фактора v_i будет вычисляться по формуле:

$$a_{v_i}^t = f \left(\sum_{\substack{i=1 \\ i \neq j}}^n a_j^{t-1} w_{ji} + a_i^{t-1} \right),$$

где a_i^t — значение фактора v_i в момент времени t ; a_j^{t-1} — значение концепта v_j в момент времени $t-1$; w_{ji} — вес дуги e_{ji} как параметр, определяющий степень влияния концепта v_j на v_i . Функция f используется в качестве пороговой функции, которая позволяет определить выходное значение $a_{v_i}^t$ в диапазоне $[-1;1]$, который будет соответствовать нечеткому множеству для нашей лингвистической переменной.

В качестве пороговой функции f для когнитивных моделей обычно используют либо сигмовидную функцию, либо функцию гиперболического тангенса. Т. к. область определения функции f предполагает наличие как положительных, так и отрицательных значений, в качестве пороговой функции будет использоваться гиперболический тангенс.

$$f(x) = \tanh(x).$$

Или, выражая через экспоненту,

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}.$$

Таким образом, на данном шаге вычисления моделируется решение задачи оценки значения интегрального клиентского опыта для этапа жизненного цикла на основе метрик клиентского опыта.

Вычисление значения интегрального клиентского опыта для всего жизненного цикла клиента. Интегральный СХ для всего жизненного цикла клиента по сути своей будет являться целевым выходным параметром всей модели, агрегируя значения клиентского опыта на всех этапах жизненного цикла. Для оценки интегрального СХ будем использовать когнитивную карту G_2 . Целевым фактором в данном случае является интегральный клиентский опыт для всего жизненного цикла клиента, а управляющими факторами модели являются значения интегрального клиентского опыта на каждом из этапов жизненного цикла, рассчитанные в качестве целевых факторов когнитивной модели G_1 . Когнитивная карта G_2 будет записываться в виде векторного функционального графа, вершинами которого будут значения лингвистической переменной интегрального клиентского опыта.

$$G_2 = \langle V_2, E_2, A_2, W_2 \rangle,$$

где V_2 — множество вершин графа, которые являются факторами когнитивной модели G_2 ; E_2 — множество дуг графа, моделирующих взаимовлияние факторов модели G_2 ; A_2 — множество значений вершин графа v_i ($v_i \in V_2$; $i = 0, 1, 2, \dots$); W_2 — множество значений степени влияния дуги e_{ij} ($e_{ij} \in E_2$; $i = 0, 1, 2, \dots$; $j = 0, 1, 2, \dots$).

Значение вершины v_i запишем в виде численного значения лингвистической переменной a_i . В качестве начальных значений для концептов множества V_2 будет принято значение, рассчитанное для целевых факторов модели G_1 на предыдущем шаге вычислений.

Оценку влияния фактора v_i на фактор v_j по аналогии с предыдущим шагом вычислений запишем в виде матрицы W_2 . Ключевым отличием в данном случае будет являться наличие дуги e_{ii} с весом w_{ii} , который по сути своей будет определять „значимость“ конкретного фактора в общей модели — в терминах функциональной модели это позволит смоделировать ситуацию, когда предыдущий интегральный клиентский опыт имеет накопительный эффект.

		факторы					
		v_1	v_2	v_3	v_j
факторы	v_1	0,25	-0,25	-0,75	w_{1j}
	v_2	-0,25	0,5	0,25	w_{2j}
	v_3	1	0,75	-0,25	w_{3j}

	v_i	w_{i1}	w_{i2}	w_{i3}	w_{ij}

Смоделируем вычисление целевого фактора когнитивной карты G_2 , т. е. интегрального клиентского опыта в разрезе этапа его жизненного цикла.

$$a_i^t = f \left(\sum_{\substack{j=1 \\ j \neq i}}^n a_j^{t-1} w_{ji} + w_{ii} a_i^{t-1} \right),$$

где a_i^t — значение фактора v_i в момент времени t ; a_j^{t-1} — значение концепта v_j в момент времени $t-1$; w_{ji} — вес дуги e_{ji} как параметр, определяющий степень влияния концепта v_j и v_i ; w_{ii} — вес дуги e_{ii} как параметр, учитывающий предысторию, т. е. вклад значения концепта в момент времени $t-1$ (по сути данный параметр позволяет моделировать „накопление“ опыта); a_i^{t-1} — значение концепта v_i в момент времени $t-1$. Функция f , по аналогии с предыдущим шагом вычислений, представляет собой функцию гиперболического тангенса и используется в качестве пороговой функции, которая позволяет определить выходное значение $a_{v_i}^t$ в диапазоне $[-1; 1]$.

Таким образом, на данном шаге вычисления моделируется решение задачи оценки интегрального клиентского опыта для всего этапа жизненного цикла клиента на основе значений клиентского опыта на каждом из этапов.

Вычисление силы взаимовлияния факторов модели друг на друга. Рассмотрим, каким образом можно оценить влияние фактора v_i на фактор v_j .

Смоделируем кейс для оценки влияния фактора v_i на фактор v_j . В общем случае, в когнитивных картах выделяется косвенный и общий причинный эффект влияния. Рассмотрим некоторый „путь“ из фактора v_i к фактору v_j : $v_i \rightarrow v_{i+1} \rightarrow v_{i+2} \rightarrow \dots \rightarrow v_{j-1} \rightarrow v_j$. Данный путь можно определить упорядоченными индексами факторов: $(i, i+1, i+2, \dots, j-1, j)$. Тогда косвенный эффект влияния фактора v_i на фактор v_j будет определяться через путь $(i, i+1, i+2, \dots, j-1, j)$. Общий эффект влияния фактора v_i на фактор v_j будет определяться множеством путей N , существующих между данными факторами.

Запишем косвенный эффект влияния фактора v_i на фактор v_j :

$$B_n(v_i, v_j) = \min w_{d, d+1},$$

где B_{n1} — влияние фактора v_i на фактор v_j через некоторый путь n из множества путей N ; d и $d+1$ — смежные слева направо индексы факторов, через которые строится путь из фактора v_i к фактору v_j . Операция \min , согласно ее определению в нечеткой когнитивной алгебре, в данном случае будет эквивалента операции умножения [17, 18].

Тогда общий причинный эффект влияния фактора v_i на фактор v_j можно записать следующим образом:

$$C(v_i, v_j) = \max B_n(v_i, v_j),$$

где $C(v_i, v_j)$ — общее влияние фактора v_i на фактор v_j через множество путей N ; $B_n(v_i, v_j)$ — влияние фактора v_i на фактор v_j через путь n_1 из множества путей N . Операция \max , согласно ее определению в нечеткой когнитивной алгебре, будет в данном случае эквивалента операции сложения.

Таким образом, на данном шаге вычисления решается задача оценки степени влияния метрик клиентского опыта и значений интегрального клиентского опыта.

5. Область применения модели. Модель оценки клиентского опыта, построенная на основе нечетких когнитивных карт, дает возможность решать ряд задач, связанных

с поддержкой принятия операционных и стратегических решений в структуре бизнеса оператора связи. Основная ценность модели, как уже было отмечено выше, состоит в возможности в момент времени рассчитывать интегральный клиентский опыт в зависимости от множества разнообразных факторов, на него влияющих. Кроме того, в контексте практических задач в структуре процессов оператора связи определенный интерес также представляют статический и динамический анализ полученных моделей.

Статический анализ модели позволяет оценивать разработанную модель с точки зрения ее адекватности и релевантности практической задаче. Для этого могут использоваться различные методы, в т. ч. оценка степени взаимовлияния факторов друг на друга, оценка влияния системы на конкретный фактор, оценка влияния конкретного фактора на систему, а также консонанс влияния как мера „уверенности“ в выводах, которые позволяет делать модель. С точки зрения практического применения статического анализа модели в процессах управления наибольший интерес представляет анализ взаимовлияния факторов друг на друга. Подобный анализ модели дает возможность субъекту управления выявлять факторы, которые оказывают наибольшее влияние на клиентский опыт, а также определять меру и пропорции влияния. Практический эффект статического анализа над моделью заключается в том, что:

- Субъект управления (например, директор по клиентскому опыту) имеет возможность выявлять неочевидные показатели операционных процессов, которые влияют на клиентский опыт, даже если данные показатели не связаны напрямую с взаимодействием с клиентом;

- Модель позволяет выстроить единую причинно-следственную цепочку между параметрами операционных процессов оператора связи и клиентским опытом, что, в свою очередь, дает возможность выстраивать клиентоцентричную модель управления компанией, когда все операционные и бизнес-подразделения ориентированы на улучшение клиентского опыта.

Задача динамического анализа когнитивной модели, в общем случае, позволяет формировать множество сценариев изменения клиентского опыта в зависимости от разных комбинаций значений управляющих факторов в определенный момент времени. Динамический анализ когнитивной модели может использоваться как в контексте операционных процессов, связанных с взаимодействием с клиентом, так и в рамках стратегического и операционного анализа в целях поддержки принятия управленческих решений. В рамках операционных процессов основная ценность модели заключается в возможности рассчитать в момент времени значение интегрального клиентского опыта для конкретного клиента и на его основе построить персонализированное взаимодействие в точках контакта с клиентом (например, использовать значение клиентского опыта в качестве одного из предикторов для допродажи или сохранения клиента). Для поддержки принятия управленческих решений динамический анализ модели позволяет моделировать совокупность сценариев управления операционными процессами и выявлять те сценарии, которые позволяют, с одной стороны, максимизировать клиентский опыт, а, с другой стороны, выбрать оптимальную модель влияния на параметры операционных процессов оператора связи. Иными словами, динамический анализ модели дает возможность формировать ретроспективу изменения клиентского опыта, рассчитывать текущее значение, а также прогнозировать будущую динамику клиентского опыта.

Заключение. На сегодняшний день элементы концепции СЕМ внедряются у большинства ведущих отечественных операторов связи. При этом, следует отметить, что сама те-

матика СЕМ достаточно слабо разработана в рамках отечественных исследований в области управления телекоммуникациями — об этом свидетельствует относительно небольшое количество публикаций и работ по данному направлению. Как следствие данных тезисов, комплексная разработка предметной области управления клиентским опытом является одним из наиболее актуальных научно-практических направлений для исследования.

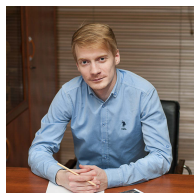
В рамках работы автором представлена когнитивная модель оценки клиентского опыта, разработанная для функционирования в структуре информационной среды отечественного оператора связи. Функциональная структура модели разработана с учетом практик и результатов исследований TMForum в области оценки и измерения клиентского опыта. Математической основой модели являются нечеткие когнитивные карты, зарекомендовавшие себя как эффективный инструмент для решения задач поддержки принятия решений в условиях неопределенности. Практическая ценность модели обусловлена возможностью ее использования, с одной стороны, в качестве инструмента поддержки принятия управленческих решений, а, с другой стороны, как подхода к измерению совокупного опыта и впечатлений клиента на основе данных операционного ландшафта оператора связи.

Дальнейшее исследование тематики измерения клиентского опыта на сети оператора связи будет строиться в двух направлениях: интенсивном и экстенсивном. Интенсивное направление предполагает дальнейшую разработку модели, ее апробацию в практических реалиях, а также внедрение модели оценки клиентского опыта в процессы оператора связи. Экстенсивное направление включает в себя более широкий взгляд на методы и модели оценки клиентского опыта как в области телекоммуникаций, так и в смежных областях.

Список литературы

1. GB962 Customer Experience Management Solution Suite. [Electron. Res.]: <https://www.tmforum.org/resources/suite/gb962-customer-experience-management-solution-suite-r17-5-0/> (дата обращения: 02.04.2020).
2. GB962 Customer Experience Management: Introduction and Fundamentals R16.0.1. [Electron. Res.]: <https://www.tmforum.org/resources/best-practice/gb962-customer-experience-management-introduction-and-fundamentals-r16-0-1/> (дата обращения: 02.04.2020).
3. GB962A Customer Experience Management Lifecycle Metrics R15.0.1. [Electron. Res.]: <https://www.tmforum.org/resources/best-practice/gb962a-customer-experience-management-lifecycle-metrics-r15-0-1/> (дата обращения: 30.04.2020).
4. RN341 Customer Experience Management Index (CEMI) Release Notes R2.0. [Electron. Res.]: <https://www.tmforum.org/resources/reference/rn341-customer-experience-management-index-cemi-release-notes-r2-0/> (дата обращения: 01.05.2020).
5. TMF066 Customer Experience Management Index Technical Specification V1.1. [Electron. Res.]: <https://www.tmforum.org/resources/technical-report-best-practice/tmf066-customer-experience-management-index-technical-specification-v1-1/> (дата обращения: 02.03.2020).
6. TR193 Customer Experience Management Index v1.3. [Electron. Res.]: <https://www.tmforum.org/resources/technical-report-best-practice/tr193-customer-experience-management-index-v1-3/> (дата обращения: 01.05.2020).
7. Акишин В. А., Кисляков С. В., Феноменов М. А. Функциональная архитектура СЕМ-комплекса для внедрения в ИТ-ландшафт крупного оператора связи // Т-КОММ: ТЕЛЕКОММУНИКАЦИИ И ТРАНСПОРТ. 2016. № 10. С. 12–16.
8. Акишин В. А., Кисляков С. В., Феноменов М. А. Пользовательский опыт в когнитивной модели управления сетью оператора // Т-comm. 2016. Т. 10, № 10.

9. Гольдштейн А. Б. О когнитивных картах в управлении телекоммуникационным оператором / А. Б. Гольдштейн, Н. А. Пожарский, Д. А. Лихачев // Информатизация и связь. 2016. № 1.
10. Заграновская А. В. Системный анализ на основе нечетких когнитивных карт // Вестник Российского экономического университета имени Г. В. Плеханова. 2018. № 4. С. 152–160.
11. Маренко М., Мальцева М. Применение когнитивного моделирования для анализа проблем малого бизнеса // Известия Иркутской государственной экономической академии. 2015. Т. 25. № 6.
12. Ротштейн А. П. Нечеткие когнитивные карты в анализе надежности систем // Надежность. 2019. № 19(4). С. 24–31.
13. Akishin V., Goldstein A., Goldstein B. Cognitive Models for Access Network Management // Galinina O., Andreev S., Balandin S., Koucheryavy Y. (eds) Internet of Things, Smart Spaces, and Next Generation Networks and Systems. NEW2AN 2017, ruSMART 2017, NsCC 2017. Lecture Notes in Computer Science. Springer, Cham. 2017. V. 10531. P. 375–381 (Scopus).
14. Axelrod R. Structure of Decision. The Cognitive Maps of Political Elites. Princeton University Press, 1976.
15. Kosko B. Fuzzy systems as universal approximators // IEEE Transactions on Computers, November 1994. V. 43. N 11. P. 1329–1333.
16. Гулаков В. К. СПИР на основе когнитивного моделирования „ИГЛА“ / В. К. Гулаков, Д. Г. Лагереv, А. Г. Подвесовский. М. : Программные продукты и системы, 2007.
17. Groumpos, P. Fuzzy cognitive maps: basic theories and their application to complex systems. / Glykas, M. (ed.) Fuzzy Cognitive Maps: Advances in Theory, Methodologies, Tools and Applications, Springer, Berlin, 2010, P. 1–22.
18. Kosko B. Fuzzy Cognitive Maps // International Journal of Man-Machine Studies. 1986. N. 24. P. 65–75.



Акишин Владимир Андреевич окончил аспирантуру на кафедре инфокоммуникационных систем в СПбГУТ. Основное направление научных исследований: системный анализ

и процессы управления, системы поддержки операционных и бизнес процессов оператор

ов связи, управление клиентским опытом на сети оператора связи.

Akishin Vladimir Andreevich graduated from postgraduate school at the Department of Information and Communication Systems at SPbSUT. His research interests are system analysis and management processes, business and operation support systems, customer experience management on a telecom operator's network.

Дата поступления — 13.08.2020

THE STRUCTURE AND PARAMETERS OF THE UNWEIGHTED CO-AUTHORSHIP NETWORK BASED ON DB REPEC DATA

S. V. Bredikhin, V. M. Lyapunov, N. G. Scherbakova

Institute of Computational Mathematics and Mathematical Geophysics SB RAS,
630090, Novosibirsk, Russia

DOI: 10.24412/2073-0667-2021-3-56-67

In this paper we investigate the structure of scientific collaboration based on co-authorship in publications retrieved from the RePEc database. The main attention is paid to the co-authorship network: nodes correspond to authors, and two authors are considered connected if they are coauthors of at least one publication. The network is represented by the undirected unweighted collaboration graph.

The co-authorship in articles provides a window on patterns of collaboration within the academic community. The study of scientific collaboration networks is carried out in two main directions. Empirical measurements provide detailed characteristics of the network: statistical properties of the distribution of degrees of nodes, global network parameters, nodes centralities (Newman 2001a; 2001b; 2001c). The study of the dynamic properties of real networks and network models makes it possible to identify the structural principles that govern the evolution of networks; dynamic properties, in turn, can explain static ones (Barabasi et al., 2001; Savić et al., 2017).

Co-authorship networks are used to identify research groups and the most important researchers and to predict their scientific success; to classify journals by degree of co-authorship and to make maps of science. Co-authorship networks can be constructed for different components of analysis, such as researchers, institutions, and countries. We use a researcher as the unit of analysis.

Let P be the set of publications under consideration and V — the set of the authors of these publications. We assume that each publication in P has at least two authors. For $i, j \in V$ let us define the binary relation of co-authorship R^{ca} : $i R^{ca} j \equiv (\exists p \in P)$ the author i is one of the authors of p and the author j is one of the authors of p . The co-authorship network is the weighted undirected network $N^{ca} = (V, E, w)$, V — the set of the authors, $E \subseteq V \times V$ — the set of weighted edges, $e = (i, j) \in E$, if $i R^{ca} j$. In the simplest case, $w(e) = 1$, regardless of how often these authors are coauthors and how much each contributed to coauthored publications, that is, the network can be considered unweighted. Then the graph corresponding to the unweighted network is represented by the $(0, 1)$ adjacency matrix $U = (u_{ij})$.

The articles used in this study were retrieved from the RePEc database (REPEC). In order to identify authors uniquely and to infer actual author identity we use the author „profile“ that authors create basing on the Author Service provided by the RePEc database (similar to Google Scholar). So we have $|P| = 91\,113$ publications with more than one author and $|V| = 32\,434$ authors of these publications. In total, these authors have published 364 979 articles. The publications refer to the period from 1954 to 2019.

We presented the base statistics such as papers per author, authors per paper, maximal authors per paper and collaboration indices CI, DC, CC, MCC. We also measured the network parameters

This work was carried out under state contract with ICMMG SB RAS (0251-2021-0005).

such as the size of the largest component, network density, average distance, maximal distance and clustering coefficient. It was found that the distribution of numbers of coauthors for authors follows power law with exponent $\gamma = 1.3$.

The study showed that in the collection of publications under consideration the fraction of coauthored publications is small (25 %) and the prevailing trend is the presence of two coauthors in a publication (77 %). The most authors are indirectly connected to each other - the maximum component includes 90 % of authors. The co-authorship network under consideration is scale-free and shows the „small world“ effect.

Key words: bibliometrics, co-authorship, collaboration indices.

References

1. BEAVER D. D. Reflections on scientific collaboration (and its study): past, present and future // *Scientometrics*. 2001. V. 52. P. 365–377.
2. LAUDEL G. Who do we measure by co-authorship? // *Research Evaluation*. 2002. V. 11. P. 3–15.
3. STEPHAN P. How economics shapes science // Cambridge, Mass.: Harvard University Press, 2012. 384 pp.
4. DING Y. Scientific collaboration and endorsement: Network analysis of coauthorship and citation networks // *J. of Informetrics*. 2011. V. 5. P. 187–203.
5. WALLACE M. L., LARIVIRE V., GINGRAS Y. A SmallWorld of citations? The influence of collaboration networks on citation practices // *PLoS ONE*. 2012. V. 7(3), e33339. [Electron. resource]. URL: <https://doi.org/10.1371/journal.pone.0033339>.
6. MARTIN T., BALL B., KARRER B., NEWMAN M. E. J. Coauthorship and citation patterns in the Physical Review // *Phys. Rev.* 2013. E 88, 012814. [Electron. resource]. URL: <https://doi.org/10.1103/PhysRevE.88.012814>.
7. NEWMAN M. E. J. The structure of scientific collaboration networks // *Proc. of the Nat. Acad. of Sci. of the USA*. 2001. V. 98(2). P. 404–409.
8. NEWMAN M. E. J. Scientific collaboration networks. I. Network construction and fundamental results // *Phys. Rev. E*, 64(1), 016131.
9. NEWMAN M. E. J. Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality // *Phys. Rev. E*, 64(1), 016132.
10. BARABASI A. L., JEONG H., NEDA, Z. RAVASZ E., SCHUBERT A., V. CSEK T. Evolution of the social network of scientific collaborations // *arXiv: cond-mat/0104162* (2001).
11. Savić M., Ivanović M., Radovanović M., Ognjanović Z., Pejović A., Jakšić Kruger T. The structure and evolution of scientific collaboration in Serbian mathematical journals // *Scientometrics*. 2014. V. 101(3). P. 1805–1830. [Electron. resource]. URL: <https://doi.org/10.1007/s11192-014-1295-6>.
12. RePEc. General principles. [Electron. resource]. URL: <http://repec.org/>.
13. YOSHIKANE F., NOZAWA T., TSUJI K. Comparative analysis of co-authorship networks considering authors' roles in collaboration: Differences between the theoretical and application areas // *Scientometrics*. 2006. V. 68, iss. 3. P. 643–655.
14. BOUROS N., SOTIROPOULOS D. N., POURNARAKIS D., & GIAGLIS G. M. Social Network Analysis Within the ICMB Community: Co-Authorship Networks // *The 13th Intern. conf. on Mobile Business, ICMB 2014, London (UK), June 4–5, 2014*. P. 8. [Electron. resource]. URL: <http://aisel.aisnet.org/icmb2014/8>.
15. LAWANI S. M. Quality, collaboration and citations in cancer research: A bibliometric study // *Ph. D. Dissertation*. Florida State Univ., 1980.
16. SUBRAMANYAM K. Bibliometric studies of research collaboration: A review // *J. Inform. Sci.* 1983. V. 6, iss. 1. P. 33–38.

17. AJIFERUKE I., BURELL Q., AND TAGUE J. Collaborative coefficient: A single measure of the degree of collaboration in research // *Scientometrics*. 1988. V. 14, iss. 5–6. P. 421–433.
18. EGGHE L. Theory of collaboration and collaborative measures // *Inform. Proc. and Manag.* 1991. V. 27, iss. 2–3. P. 177–202.
19. SARVANUR K. AND SRIKANTH R. Modified collaboration coefficient: A new measure for quantifying the degree of research collaboration // *Scientometrics*. 2010. V. 84, iss. 2. P. 365–376.
20. DISTEL P. *Teorya grafov*. Novosibirsk: Int-t matematiki, 2002. 336 s.
21. KHARARI F. *Teorya grafov*. M.: Mir. 1973. 297 s.
22. WATTS D.J. Networks, dynamics and the small-world phenomenon // *Amer. J. of Sociol.* 1999. V. 105, iss. 2. P. 493–527.
23. WATTS D.J., STROGATZ S.H. Collective dynamics of 'small-world' networks // *Nature*. 1998. V. 393. P. 440–442.
24. MILGRAM S. The small world problem // *Psychol. Today*. 1967. V. 2. P. 60–67.

СТРУКТУРА И ПАРАМЕТРЫ НЕВЗВЕШЕННОЙ СЕТИ СОАВТОРСТВА НА ОСНОВЕ ДАННЫХ БД PERES

С. В. Бредихин, В. М. Ляпунов, Н. Г. Щербакова

Институт вычислительной математики и математической геофизики СО РАН,
630090, Новосибирск, Россия

УДК 001.12+303.2

DOI: 10.24412/2073-0667-2021-3-56-67

Исследуется структура сотрудничества на основе соавторства в научных публикациях. На основе информации, извлеченной из библиографической базы данных, построена сеть соавторства и измерены ее основные параметры. Показано, что сеть является масштабно-инвариантной и обладает структурой, свойственной „малому миру“. Вычислены индексы сотрудничества, отражающие вовлеченность ученых в совместную деятельность.

Ключевые слова: библиометрия, сеть соавторства, индексы сотрудничества.

Введение. Одним из классов комплексных сетей являются „сети сотрудничества“, в которых акторы связаны между собой благодаря членству в некоторой группе, например, спортсмены, занимающиеся одним видом спорта, актеры кино, участвующие в общем фильме; группы авторов, совместно подготовивших общую научную публикацию (НП) — соавторы.

Если более одного автора разделяет ответственность за НП, содержащую результаты конкретных исследований, то предполагается, что соавторы какое-то время работали сообща. Соавторство является достаточно хорошо документированным типом научного сотрудничества ученых. Общая ответственность за публикацию в солидном научном журнале редко возможна без какой-либо формы сотрудничества, которое, как правило, способствует эффективному решению поставленных в НП проблем, сокращая число заблуждений и ошибок.

Обзоры факторов, способствующих повышению роли, которую играет сотрудничество в научных исследованиях, а также побуждающих к нему мотиваций, можно найти в работах [1, 2], а также в книге [3]. Анализ сетей соавторства позволяет оценить степень взаимопроникновения идей, знаний и технологий и характеризует структуру научного сообщества. При этом могут рассматриваться различные уровни сотрудничества: между странами, в пределах одной страны, среди сотрудников организации, между научными направлениями, внутри одного направления, в рамках журнала или трудов конференции и т. д.

Сети цитирования и сети соавторства представляют разные взгляды на массив НП. Источником информации о НП, как правило, являются библиографические базы данных. При изучении сети цитирования НП являются ее узлами, а ориентированные связи между узлами строятся на основании цитирования. В сетях соавторства узлами являются авторы НП, а неориентированные связи строятся на основании того, что авторы являются

Исследования выполнены в рамках государственного задания ИВМ и МГ СО РАН (0251-2021-0005).

соавторами. Тем не менее, наблюдается взаимное влияние практик сотрудничества и цитирования ([4–6]). Например, результаты анализа параметров сети соавторства позволяют определить, насколько часто авторы цитируют своих соавторов по сравнению с другими цитируемыми авторами. Сети соавторства используются при выявлении групп ученых и наиболее важных направлений их деятельности и для прогнозирования успеха, позволяют классифицировать журналы по степени соавторства, составлять карты науки.

Изучение „сетей сотрудничества“ ведется по двум основным направлениям. Эмпирические измерения предоставляют детализированную характеристику сети: статистические свойства распределения степеней узлов, глобальные параметры сети, центральность [7–9]. Исследование динамических свойств реальных сетей и их моделей позволяет выявить структурные принципы, управляющие эволюцией сетей; динамические свойства, в свою очередь, могут объяснить статические [10, 11].

Данная работа содержит методы и результаты измерения статических параметров сети соавторства, представленной неориентированным невзвешенным графом, построенном на основе информации, извлеченной из библиографической базы данных (БД) RePEc [12]. Соавторство исследуется в рамках одной научной области.

1. Сеть публикаций и сеть соавторства. Обозначим $P = \{p_1, p_2, \dots, p_m\}$ множество НП, содержащихся в БД, $|P| = m$; $V = \{v_1, v_2, \dots, v_n\}$ — множество авторов НП, $|V| = n$.

Сеть публикаций (сеть авторства) $N^{\text{pub}} = (V, P, E^{\text{pub}})$ — это двудольная сеть с узлами, разделенными на два непересекающихся независимых (т.е. никакие два узла не являются смежными) множества V и P , при этом дуга $e = (v_i, p_j) \in E^{\text{pub}}$, если v_i является автором или соавтором НП p_j .

Матрица смежности (*матрица авторства*) $A = (a_{ij})$ графа, соответствующего N^{pub} , является матрицей размера $(n \times m)$, строки соответствуют авторам, а столбцы — публикациям:

$$a_{ij} = \begin{cases} 1, & \text{если } v_i \text{ является автором } p_j, \\ 0 & \text{в противном случае.} \end{cases} \quad (1)$$

Для $v_i, v_j \in V$ определим бинарное отношение соавторства R^{ca} :

$$v_i R^{\text{ca}} v_j \equiv (\exists p \in P), \text{ авторами } p \text{ являются } v_i, \text{ и } v_j. \quad (2)$$

Сетью соавторства будем называть взвешенную неориентированную сеть $N^{\text{ca}} = (V, E, w)$, в которой V — множество авторов, $E \subseteq V \times V$ — множество взвешенных ребер, $e = (v_i, v_j) \in E$, если выполняется $v_i R^{\text{ca}} v_j$. Узлами N^{ca} являются авторы НП, а в качестве ребер выступают пары авторов, являющихся соавторами по крайней мере одной НП. В простейшем случае $w(e) = 1$ независимо от того, насколько часто авторы v_i и v_j являются соавторами и какой вклад внес каждый автор, т.е. в этом случае сеть N^{ca} можно считать невзвешенной.

Тогда граф, соответствующий невзвешенной сети, представляется $(0, 1)$ -матрицей смежности $U = (u_{ij})$,

$$u_{ij} = \begin{cases} 1, & \text{если } v_i R^{\text{ca}} v_j, \\ 0 & \text{в противном случае.} \end{cases} \quad (3)$$

Матрицу U можно представить в виде произведения $A \cdot A^T$ с заменой всех ненулевых элементов на единицу. Здесь A — матрица смежности N^{pub} (1). Заметим, что сеть соавторства не всегда моделируется неориентированным графом. Например, если подчеркивается лидерство первого автора в списке соавторов, то сеть соавторства может быть представлена орграфом [13, 14].

Возможно, первым случаем использования сетевого подхода к понятию соавторства можно считать вычисление так называемых чисел Эрдеша (Erdős numbers) — расстояния в сети N^{ca} от какого-либо ученого, работающего в области математики, до венгерского математика Р. Erdős (1913–1996), автора около 1525 НП. Непосредственные соавторы Эрдеша, их 511, имеют число Эрдеша, равное единице.

Исследование невзвешенных N^{ca} -сетей показало, что типичное среднее расстояние между узлами приблизительно равно шести, что соответствует среднему расстоянию в сети социальных знакомств [9]. Сеть N^{ca} можно отнести к сетям „малого мира“, для которых среднее расстояние $L(N)$ растет логарифмически с ростом числа узлов, $L(N) \propto \log(n)$, а расстояние l между парой произвольно выбранных авторов растет пропорционально $\log(l)$.

2. Исходные данные. Проблема идентификации авторов на основе текста НП для БД значительного размера признана достаточно трудоемкой и зависящей от того, насколько хорошо документирована база данных. Так, например, трудности возникают при распознавании авторов НП, имеющих короткие обозначения ФИО, что приводит к необоснованным совпадениям и ошибочным выводам. Подход к идентификации авторов на основе анализа заголовка статьи, содержащего список авторов, места их работы и страны (с использованием библиотеки *Python difflib library*) можно найти в работе [6]. В нашем случае при отборе авторов и НП была использована информация, содержащаяся в БД. Во избежание путаницы, возникающей из-за различных форматов представления информации об авторе в НП, авторы создают „профили“ на основе сервиса *Author Service*, предоставляемого БД RePEc (подобно *Google Scholar*). В профиле автора указаны НП, размещаемые в БД, автором или соавтором которых он является. На момент 2020.01.31 число зарегистрировавших себя авторов составляет 70 549, среди них присутствуют и не задекларировавшие ни одной НП. Уникальных идентификаторов НП 470 333. Научных публикаций, задекларированных несколько раз — 138 882, эти НП предположительно имеют более одного автора. Дескрипторов НП, имеющих более одного автора и задекларированных тем же числом авторов, сколько указано в дескрипторе, 91 113. Число авторов этих НП 32 434. Всего этими авторами задекларировано 364 979 НП (в том числе и имеющих одного автора). Эти „фильтрованные“ данные взяты за основу при построении сети соавторства. Публикации относятся к периоду с 1954 по 2019 гг. Следует отметить, что дата создания БД RePEc относится к 1997 г.

Введем обозначения: S — множество авторов, имеющих публикации в соавторстве; Pa — множество публикаций, имеющих a авторов; P_+ — множество публикаций, имеющих более одного автора; P — множество всех публикаций, $P = P_1 \cup P_+$; $|P| = \sum_j |P_j|$; q — максимальное число соавторов публикаций из множества P .

Базовая статистическая информация о соавторстве в рассматриваемом множестве НП представлена в табл. 1. Таким образом, среднее число авторов НП множества P_+ равно 2,3, а именно приблизительно 77 % всех НП множества P_+ имеют двух авторов, а примерно 20 % — трех. На остальные группы приходится около 3 % множества всех НП, выполненных в соавторстве. Значение данного параметра зависит, в частности, от рассматриваемой области науки (в данном случае — экономика). Пример для наиболее значимых размеров соавторства выглядит так (число соавторов, число статей): (2, 70 399); (3, 18 073); (4, 2245); (5, 251).

Среднее число НП в соавторстве, приходящихся на каждого автора, составляет 2,8. На рис. 1 приведено распределение продуктивности авторов множества S , опубликовавших

Таблица 1

Базовая информация о соавторстве

Параметр	Значение
Число авторов, имеющих НП в соавторстве	$ S = 32\,434$
Число НП, имеющих более одного автора	$ P_+ = 91\,113$
Число НП множества P , имеющих одного автора, для авторов из A	$ P_1 = 273\,866$
Максимальное число соавторов публикаций из множества P	$q = 17$
Последовательность значений числа соавторов публикаций	$\{2, 3, \dots, 12, 15, 16, 17\}$
Среднее число авторов НП из множества P_+	2,3
Среднее число НП из множества P_+ , приходящихся на одного автора	2,8

НП в соавторстве. По оси абсцисс указано число НП, по оси ординат — доля (в промилле от общего числа) авторов, опубликовавших указанное число НП, шкалы логарифмические. Распределение следует степенному закону $p_x \sim x^{-\gamma}$ с параметрами $x_{min} = 6$, $\gamma = 1,64$ (вычислено с помощью пакета *igraph*), т.е. большинство авторов имеет незначительное число НП, выполненных в соавторстве.

Отметим, что общее число НП растет со временем, в том числе растет и число НП, опубликованных в соавторстве, (период, число НП): (1954–1989, 3964); (1990–1999, 10 781); (2000–2009, 33 288); (2010–2019, 42 941).

3. Эмпирические характеристики.

3.1. *Меры сотрудничества.* Для оценки степени сотрудничества и определения трендов в развитии сотрудничества разработан ряд мер. Рассмотрим некоторые из них, используемые для анализа соавторства в НП, относящихся к одной области науки.

Индекс сотрудничества (*Collaborative Index* [15]) определяется как

$$CI = \frac{\sum_{a=1}^q a \cdot |P_a|}{|P|}. \quad (4)$$

Это среднее число авторов публикаций из множества P . Заметим, что CI не равен нулю, если все НП имеют одного автора.

Степень сотрудничества (*Degree of Collaboration* [16]) определяется как

$$DC = 1 - \frac{|P_1|}{|P|}. \quad (5)$$

Это доля НП, имеющих более одного автора, по отношению ко всему множеству НП этих авторов. Множество значений лежит в интервале $[0, 1]$. Индекс равен нулю, если все НП имеют одного автора, и имеет тем большее значение, чем больше НП подготовлены в соавторстве.

Коэффициент сотрудничества (*Collaborative Coefficient* [17]) определяется как

$$CC = 1 - \frac{\sum_{a=1}^q |P_a|/a}{|P|}. \quad (6)$$

Таблица 2

Индексы сотрудничества для множества P

Мера сотрудничества	Значение
Индекс сотрудничества (4)	$CI = 1,315768$
Степень сотрудничества (5)	$DC = 0,249639$
Коэффициент сотрудничества (6)	$CC = 0,134959$
Модифицированный коэффициент сотрудничества (7)	$MCC = 0,134963$

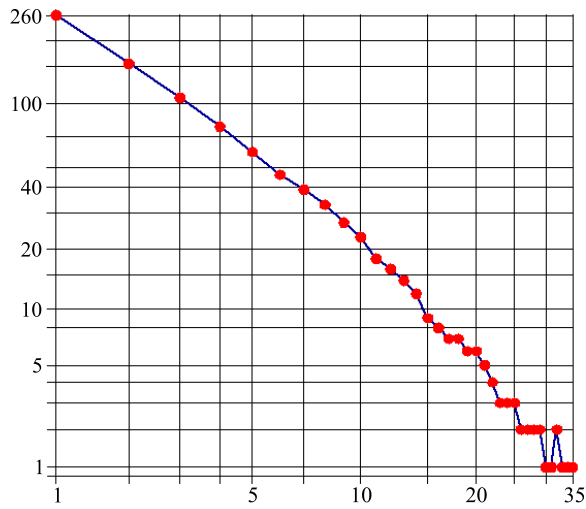
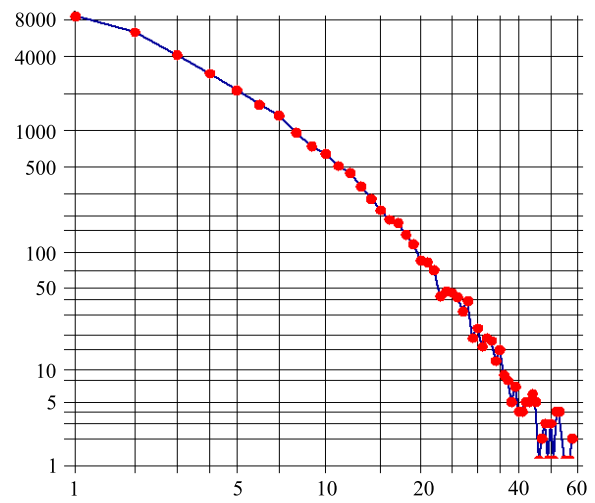
Рис. 1. Продуктивность авторов N^{ca} 

Рис. 2. Распределение степеней узлов сети

Коэффициент равен нулю, если в коллекции P нет публикаций в соавторстве. Дифференцирует уровни соавторства (2, 3 автора и т.д.). Заметим, что CC не равен единице при максимальном соавторстве, когда все авторы являются соавторами любой публикации коллекции.

Модифицированный коэффициент сотрудничества (*Modified Collaboration Coefficient*) [18, 19] определяется как

$$MCC = \frac{|S|}{|S| - 1} CC. \quad (7)$$

Это нормированный коэффициент сотрудничества. Предполагается, что $|S| > 1$. Коэффициент равен единице при максимальном числе соавторов и нулю, если все НП имеют одного автора.

В табл. 2 приведены значения рассмотренных индексов сотрудничества, вычисленных на основе анализа множества публикаций P . Согласно значениям индексов, уровень вовлеченности в совместную работу авторов публикаций множества P не высок.

3.2. *Параметры сети соавторства.* При построении сети соавторства мы рассматриваем множество авторов S и множество публикаций P_+ , каждая из которых имеет более одного автора. В невзвешенной сети соавторства $N^{ca} = (V, E^{ca})$ вершинам множества V соответствуют элементы множества S , ребро $e = (i, j) \in E^{ca}$, если выполняется (2) для $P \equiv P_+$. Размеры сети: $n = |V| = 32\,434$, $m = |E^{ca}| = 73\,969$. Матрица смежности $U = (u_{ij})$

Таблица 3

Параметры, определяющие структуру N^{ca}

Параметр	Значение
Распределение степеней узлов N^{ca} (число соавторов)	$p_x \sim x^{-\gamma}$, $x_{\min} = 4$, $\gamma = 1,3$ (рис. 2)
Среднее значение степеней узлов	$\langle k \rangle = \langle \deg(\cdot) \rangle \sim 4,6$
Число связных компонент	1242
Размеры максимальной связной компоненты $N = (V, E)$	$ V = 29\,270$, $ E = 71\,780$, $\langle k \rangle \sim 4,9$
Вторая по числу узлов связная компонента	25 узлов
Реберная плотность в максимальной компоненте [20]:	$\Delta^U(N^{\text{ca}}) = 0,000168$
$\Delta^U(G) = \frac{2 E }{ V (V - 1)}$	
Радиус [21]: $rad(G) = \min_{x \in V(G)} \max_{y \in V(G)} d(x, y)$	$rad(N^{\text{ca}}) = 13$
Диаметр [21]: $diam(G) = \max_{x \in V(G)} \max_{y \in V(G)} d(x, y)$	$diam(N^{\text{ca}}) = 25$
Характеристическая длина пути (среднее расстояние) [22]:	$L(N^{\text{ca}}) = 6,577475$
$L(G) = \frac{1}{n(n-1)} \sum_{x \in V} \sum_{y \neq x \in V} d(x, y)$	
В случайном графе $L_{\text{rand}} = \ln n / \ln \langle k \rangle$	$L_{\text{rand}} \sim 6,739148$
Локальный коэффициент кластеризации [23]:	$CC(N^{\text{ca}}) = 0,264428$
$CC(i) = \frac{\sum_{j,h} u_{ij}u_{ih}u_{jh}}{\deg(i)(\deg(i) - 1)}$; $CC(N) = \frac{1}{n} \sum_i CC(i)$	
В случайном графе $C_{\text{rand}} = \langle k \rangle / n$	$C_{\text{rand}} \sim 0,000141$

сети N^{ca} организована в соответствии с равенством (3). В табл. 3. представлены основные параметры, характеризующие структуру N^{ca} .

На рис. 2 приведено распределение числа соавторов у авторов множества S . По оси абсцисс указано число соавторов, по оси ординат — число авторов, имеющих указанное число соавторов. Шкалы логарифмические. Распределение следует степенному закону $p_x \sim x^{-\gamma}$ с параметрами $x_{\min} = 4$, $\gamma = 1,3$.

Число узлов в максимальной компоненте N^{ca} составляет примерно 90 % от всего множества узлов. Остальные компоненты имеют не более 25 узлов. Таким образом, большинство авторов к рассматриваемому моменту времени опосредованно связаны друг с другом. Реберная плотность достаточно низкая. Параметр $L(N)$ позиционируется в работе [22] как одна из характеристик сетей „малого мира“ [24], для которых характерно малое значение $L(N)$ (сравнимое со средним расстоянием в случайном графе) при достаточно большом значении n и небольшом среднем значении степеней вершин ($\langle k \rangle \ll n$). Такие сети также характеризуются более высокими, по сравнению со случайными графами, значениями локальных коэффициентов кластеризации. Сравнение значений параметров $L(N^{\text{ca}})$ и $CC(N^{\text{ca}})$ с соответствующими значениями параметров случайного графа позволяет заключить, что рассматриваемую сеть соавторства можно отнести к сетям „малого мира“.

Заключение. Анализ N^{ca} показал, что в рассматриваемой коллекции НП RePEc доля командной работы невелика и составляет приблизительно 25 %. При этом преобладающей тенденцией является наличие двух соавторов (77 % НП). Большинство авторов опосредованно связано друг с другом — максимальная компонента связности включает 90 % авторов. Сеть соавторства является масштабно-инвариантной и относится к сетям „малого мира“. Дальнейшему исследованию подлежит взаимосвязь между соавторством и цитированием, а именно, получают ли НП, опубликованные в соавторстве, больше цитирований, чем опубликованные одним автором. Интерес представляют также параметры взвешенной сети соавторства, например, устойчивость авторских групп.

Список литературы

1. BEAVER D. D. Reflections on scientific collaboration (and its study): past, present and future // *Scientometrics*. 2001. V. 52. P. 365–377.
2. LAUDEL G. Who do we measure by co-authorship? // *Research Evaluation*. 2002. V. 11. P. 3–15.
3. STEPHAN P. How economics shapes science // Cambridge, Mass.: Harvard University Press, 2012. 384 pp.
4. DING Y. Scientific collaboration and endorsement: Network analysis of coauthorship and citation networks // *J. of Informetrics*. 2011. V. 5. P. 187–203.
5. WALLACE M. L., LARIVIRE V., GINGRAS Y. A SmallWorld of citations? The influence of collaboration networks on citation practices // *PLoS ONE*. 2012. V. 7(3), e33339. [Electron. resource]. URL: <https://doi.org/10.1371/journal.pone.0033339>.
6. MARTIN T., BALL B., KARRER B., NEWMAN M. E. J. Coauthorship and citation patterns in the Physical Review // *Phys. Rev.* 2013. E 88, 012814. [Electron. resource]. URL: <https://doi.org/10.1103/PhysRevE.88.012814>.
7. NEWMAN M. E. J. The structure of scientific collaboration networks // *Proc. of the Nat. Acad. of Sci. of the USA*. 2001. V. 98(2). P. 404–409.
8. NEWMAN M. E. J. Scientific collaboration networks. I. Network construction and fundamental results // *Phys. Rev. E*, 64(1), 016131.
9. NEWMAN M. E. J. Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality // *Phys. Rev. E*, 64(1), 016132.
10. BARABASI A. L., JEONG H., NEDA, Z. RAVASZ E., SCHUBERT A., V. CSEK T. Evolution of the social network of scientific collaborations // *arXiv: cond-mat/0104162* (2001).
11. Savić M., Ivanović M., Radovanović M., Ognjanović Z., Pejović A., Jakšić Kruger T. The structure and evolution of scientific collaboration in Serbian mathematical journals // *Scientometrics*. 2014. V. 101(3). P. 1805–1830. [Electron. resource]. URL: <https://doi.org/10.1007/s11192-014-1295-6>.
12. RePEc. General principles. [Electron. resource]. URL: <http://repec.org/>.
13. YOSHIKANE F., NOZAWA T., TSUJI K. Comparative analysis of co-authorship networks considering authors' roles in collaboration: Differences between the theoretical and application areas // *Scientometrics*. 2006. V. 68, iss. 3. P. 643–655.
14. BOUROS N., SOTIROPOULOS D.N., POURNARAKIS D., & GIAGLIS G.M. Social Network Analysis Within the ICMB Community: Co-Authorship Networks // *The 13th Intern. conf. on Mobile Business, ICMB 2014, London (UK), June 4–5, 2014*. P. 8. [Electron. resource]. URL: <http://aisel.aisnet.org/icmb2014/8>.
15. LAWANI S. M. Quality, collaboration and citations in cancer research: A bibliometric study // *Ph.D. Dissertation. Florida State Univ.*, 1980.

16. SUBRAMANYAM K. Bibliometric studies of research collaboration: A review // J. Inform. Sci. 1983. V. 6, iss. 1. P. 33–38.
17. AJIFERUKE I., BURELL Q., AND TAGUE J. Collaborative coefficient: A single measure of the degree of collaboration in research // Scientometrics. 1988. V. 14, iss. 5–6. P. 421–433.
18. EGGHE L. Theory of collaboration and collaborative measures // Inform. Proc. and Manag. 1991. V. 27, iss. 2–3. P. 177–202.
19. SARVANUR K. AND SRIKANTH R. Modified collaboration coefficient: A new measure for quantifying the degree of research collaboration // Scientometrics. 2010. V. 84, iss. 2. P. 365–376.
20. ДИСТЕЛЬ Р. Теория графов. Новосибирск: Ин-т математики, 2002. 336 с.
21. ХАРАРИ Ф. Теория графов. М.: Мир, 1973. 297 с.
22. WATTS D.J. Networks, dynamics and the small-world phenomenon // Amer. J. of Sociol. 1999. V. 105, iss. 2. P. 493–527.
23. WATTS D.J., STROGATZ S.H. Collective dynamics of 'small-world' networks // Nature. 1998. V. 393. P. 440–442.
24. MILGRAM S. The small world problem // Psychol. Today. 1967. V. 2. P. 60–67.



Бредихин Сергей Всеволодович — канд. техн. наук, зав. лабораторией Ин-та вычислительной математики и математической геофизики СО РАН; e-mail: bred@nsc.ru;

Сергей Бредихин окончил механико-математический факультет Новосибирского государственного университета в 1968 г. С 1968 г. — сотрудник Института автоматизации и электрометрии СО РАН. Кандидат технических наук с 1983 г. С 1988 г. — заведующий Лабораторией прикладных систем Института вычислительной математики и математической геофизики СО РАН. Являлся техническим директором проекта „Сеть Интернет Новосибирского Научного Центра“. Лауреат государственной премии по науке и технике 2012 г. В сфере его научных интересов — измерение и анализ сетей распределенных информационных структур. Автор и соавтор более 110 работ и двух монографий: „Методы библиометрии и рынок электронной научной периодики“, „Анализ цитирования в библиометрии“.

Sergey Bredikhin graduated from Novosibirsk State University in 1968 (faculty of Mechanics and Mathematics). In 1968 he became an employee of Institute of Automation and Electrometry SB RAS. In 1983 he received PhD degree in Engineering Science. Since 1988 he is the head of Applied Systems laboratory of Institute of Computational Mathematics and Mathematical Geophysics SB RAS. He was the technical manager

of „Akademgorodok Internet Project“. He is the state prize winner in science and engineering (2012). Sphere of his scientific interests — the measurement and analysis of networks of the distributed information structures. He is the author and co-author of more than 110 works and two monographs: „Metody bibliometrii i rynek elektronnoy nauchnoy periodiki“, „Ansliz tsitirovaniya v bibliometrii“.



Ляпунов Виктор Михайлович — ведущий инженер Ин-та вычислительной математики и математической геофизики СО РАН; e-mail: vic@nsc.ru;

Виктор Ляпунов окончил механико-математический факультет Новосибирского государственного университета в 1978 году. В 1978 г. стал сотрудником Вычислительного Центра СО АН СССР, а с 1990 г. — сотрудником Института систем информатики СО АН СССР. С 2004 г. — ведущий инженер Института вычислительной математики и математической геофизики СО РАН. Занимается вопросами извлечения информации из баз данных и обработкой больших массивов данных. Соавтор более 10 работ в этой области.

Victor Lyapunov graduated from Novosibirsk State University in 1978 (faculty of Mechanics and Mathematics). In 1978, he became an employee of Computing Center of SB AS USSR, since 1990 — an employee of Institute

of Informatics Systems SB RAS. Since 2004 he works as software engineer in Institute of Computational Mathematics and Mathematical Geophysics SB RAS. His current research interests include methods of information extracting from databases and processing of large data sets. He is the co-author of more than 10 works in that area.



Щербакова Наталья Григорьевна — ст. науч. сотр. Ин-та вычислительной математики и математической геофизики СО РАН; e-mail: nata@nsc.ru.

Наталия Щербакова окончила Новосибирский государственный университет по специальности „Математическая лингвистика“ в 1967 г. С 1967 г. работала в Институте математики СО РАН, затем в Институте автоматизации и электрометрии СО РАН в области создания программного обеспечения систем передачи данных. С 2000 г. — сотрудник Института вычислительной математики и математической геофизики СО РАН, где с 2002 г. занимает должность старшего научного сотрудника. Являлась участником проекта „Сеть Интернет

Новосибирского Научного Центра“, занималась вопросами мониторинга и анализа IP-сетей. Автор и соавтор более 40 работ, соавтор монографии „Анализ цитирования в библиометрии“. Текущие интересы лежат в области исследования методов оценки научной деятельности на основе анализа цитирования научной литературы.

Natalia Shcherbakova graduated from Novosibirsk State University in 1967 (mathematical linguistics). Since 1967 she worked at Institute of Mathematics SB RAS, then at Institute of Automation and Electrometry SB RAS in the field of software design for data transmission systems. In 2000 — the employee of Institute of Computational Mathematics and Mathematical Geophysics SB RAS, since 2002 works as senior researcher. She is a member of „Akademgorodok Internet Project“, dealt with software of monitoring and the analysis of IP networks. She is the author and co-author of more than 40 works, the co-author of the monograph „Ansliz tsitirovaniya v bibliometrii“. The current research interests lie in the field of bibliometrics: methods of measuring of scientific activity on the base of citations.

Дата поступления — 16.03.2021

NEW SCIENTIFIC INDEX OF THE AUTHOR'S PUBLISHING ACTIVITY

V. V. Mironov

Ryazan State Radio Engineering University named after Academician V. F. Utkin,
390005, Ryazan, Russia

DOI: 10.24412/2073-0667-2021-3-68-78

A new (three-dimensional) scientometric index is presented — a (weighted) index of the author's publishing (or public) activity, depending on the number of author's publications in various journals. Informatively, the index of the author's publishing activity characterizes the breadth of the author's scientific interests, attention to his work on the part of publishers and (indirectly) the level of scientific significance of his results. A weighted estimate is achieved by accounting for the quartiles of the journals. For the convenience of comparing the publishing activity of scientists, a one-dimensional pseudo-norm (or assessment) of the activity index has been introduced. The new indices are easy to use, based on the Hirsch methodology already familiar to the scientific community, and at the same time, in quality they surpass the Hirsch methodology in analyzing the scientist's contribution to the development of his branch of science.

Key words: scientometric indicator, multidimensional and one-dimensional indices of the author's publishing activity, breadth of scientific interests, balanced and informative indices.

References

1. Grigor'eva E. I. Naukometrija – nauka vazhnaja, interesnaja, poleznaja. 2020. [Electron. resource]: <http://www.isran.ru/> (date of access: 14.01.2021).
2. Mironov V. V. Novye indeksy publikacionnoj aktivnosti // Vestnik RAN. 2020. T. 90. N 10. S. 959–966.
3. Mironov V. V. Obrabotka dannyh o publikacionnoj aktivnosti avtora v sostave avtorskogo kollektiva s uchetom kvartilej zhurnalov // Nauchno-tehnicheskaja informacija. Serija 2. Informacionnye processy i sistemy. 2020. N 11. S. 30–35.
4. Mironov V. V. Informatizacija obrazovanija: dostizhenija i problemy // Informatizacija obrazovanija i nauki. 2017. N 4 (36) . S. 3–18.
5. Hirsch J. E. An index to quantify an individual's scientific research output // Proceedings of the National Academy of Sciences of the United States of America. 2005. Vol. 102. N 46. P. 16569–16572.
6. Google Scholar. [Electron. resource]: <https://scholar.google.com/> (date of access: 15.01.2021).
7. Elibrary.Ru, Nauchnaja Jelektronnaja Biblioteka. [Electron. resource]: <https://elibrary.ru/> (date of access: 15.01.2021).
8. Astrophysics data system (ADS NASA). [Electron. resource]: <http://adsabs.harvard.edu/> (date of access: 17.01.2021).
9. Tagiew R., Ignatov D. I. Behavior mining in h-index ranking game // CEUR Workshop Proceedings. 2017. Vol. 1968. P. 52–61.

-
10. Imaev V. Tehnologii uvelichenija indeksa Hirsha i razvitie imitacionnoj nauki // Komissija RAN po bor'be s lzhenaukoj i fal'sifikaciej nauchnyh issledovanij. V zashhitu nauki. 2016. N 17. S. 38–51.
 11. Demina N. Hirshemaniya i hirshefobiya. Troickij variant — Nauka. [Electron. resource]: <https://trv-science.ru/2016/12/06/khirshemaniya-i-khirshefobiya/> (date of access: 17.01.2021).
 12. Mihajlov O. V. Novaja versija indeksa Hirsha — j-indeks // Vestnik RAN. 2014. T. 84, N 6. S. 532–535.
 13. Egghe L. Theory and practise of the g-index // Scientometrics. 2006. Vol. 69, N 1. P. 131–152.
 14. Kosmulski M. I. A bibliometric index // Forum Akademickie. 2006. Vol. 11. P. 31.
 15. Prathap G. Hirsch-type indices for ranking institutions' scientific research output // Current Science journal. 2006. Vol. 91 (11). P. 1439.
 16. Jones T., Huggett S., Kamalski J. Finding a Way Through the Scientific Literature: Indexes and Measures // World Neurosurgery. 2011. Vol. 76, N 1, 2. P. 36–38.
 17. Holodov A. S. Ob indeksah citirovanija nauchnyh rabot // Vestnik RAN. 2015. T. 85, N 4. C. 310–320.
 18. Mazov N. A, Gureev V. N. Al'ternativnye podhody k ocenke nauchnyh rezul'tatov // Vestnik RAN. 2015. T. 85, N 2 . C. 115–122.
 19. Kuznecov A. V. Dlja nachala nado navesti porjadok v sushhestvujushhej sisteme RINC. Pis'ma v redakciju // Vestnik RAN. 2014. T. 84, N 3. S. 268–269.
 20. Ludo Waltman, Nees Jan van Eck. Robust Evolutionary Algorithm Design for Socio-Economic Simulation: Some Comments // Comput. Econ. 2009. N 33. P. 103–105, DOI 10.1007/s10614-008-9148-0.
 21. Hirsch J. E. h_α : An index to quantify an individual's scientific leadership. 2019. [Electron. resource]: <http://link.springer.com/article/10.1007/s11192-018-2994> (date of access: 18.01.2021).
 22. Bredihin S. V., Shherbakova N. G. Formalizacija indeksa Hirsha. Obzor // Problemy informatiki. 2014. N 2 (23). S. 17–27.
 23. Moiseenko V. V., Rodionov A. S. Vozrastnaja zavisimost' soavtorstva pri provedenii nauchnyh issledovanij v akademicheskom institute // Problemy informatiki. 2017. N 1 (34) . S. 62–73.
 24. Bredihin S. V., Ljapunov V. M., Shherbakova N. G. Dinamika rosta seti citirovanija nauchnyh statej // Problemy informatiki. 2020. N 1. S. 5–20.

НОВЫЙ НАУКОМЕТРИЧЕСКИЙ ИНДЕКС ИЗДАТЕЛЬСКОЙ АКТИВНОСТИ АВТОРА

В. В. Миронов

Рязанский государственный радиотехнический университет имени академика

В. Ф. Уткина,

390005, г. Рязань, Россия

УДК 001.5 + 519.24

DOI: 10.24412/2073-0667-2021-3-68-78

Представлен новый (трехмерный) наукометрический индекс — (взвешенный) индекс издательской (или публичной) активности автора, зависящий от числа публикаций автора в различных журналах. Информативно индекс издательской активности автора характеризует широту научных интересов автора, внимание к его работам со стороны издательств и (косвенно) уровень научной значимости его результатов. Взвешенность оценки достигается путем учета квартилей журналов. Для удобства сравнения издательской активности ученых введена одномерная псевдонорма (или оценки) индекса активности. Новые индексы просты в использовании, основаны на уже привычной для научного сообщества методологии Хирша и одновременно по качеству превосходят методику Хирша в анализе вклада ученого в развитие его отрасли науки.

Ключевые слова: наукометрический показатель, многомерный и одномерный индексы издательской активности автора, широта научных интересов, взвешенность и информативность индексов.

Введение. Методология Хирша. Направление исследования — наукометрия (англ. *Scientometrics*) — изучение самой науки, как таковой, количественными методами; изучение эволюции науки через измерения и статистическую обработку данных. В этом контексте ее можно рассматривать как часть информатики. Информатика (фр. *Informatique*; англ. *Computer science*) — наука о методах и процессах сбора, хранения, обработки, передачи, анализа и оценки информации с применением компьютерных технологий.

Наукометрия, несмотря на противоречия и критику, продолжает жить, развиваться и исследоваться (к примеру, [1]).

Представленная автором статья — попытка отойти от принятых шаблонов и схем в оценке деятельности научных и педагогических работников и, особенно, в оценке перспектив ученого. В этом контексте статья примыкает к работам автора [2–4].

В работах [2, 3] по сути, представлена платформа, на базе которой можно строить различные „модели“ анализа творческой активности автора, в работе [4] анализируются общие вопросы информатизации, в том числе относящиеся к рассматриваемой тематике, не избегая при этом нравственных вопросов и последствий применения информационных технологий.

Так сложилось, что в научном сообществе и в центрах руководства наукой периодически возникает проблема сравнения вклада ученых в представляемую ими отрасль науки для эффективного финансирования работ, кадровых решений и других научно-организационных проблем. Широко известным методом решения подобных задач стал

метод Хирша и его результат — индекс Хирша, который является количественной характеристикой продуктивности ученого (или группы ученых, или научной организации или даже страны в целом), исчисляемый на основе количества публикаций и количества цитирований (в любых изданиях) публикаций этого ученого.

Методология Хирша — это (в рассматриваемом контексте) комплекс или система базисных принципов, методов, методик, способов и средств их реализации для оценки научно-практического вклада ученого в „его“ научную отрасль по данным его публикаций в научных журналах.

Ядром методологии Хирша является индекс Хирша (или h -индекс) — наукометрический числовой показатель [предложенный в 2005 г. американским физиком аргентинского происхождения Хорхе Хиршем (Hirsch, J. E.) из Калифорнийского университета (Сан Диего, США)] для оценки публикационной активности физиков [5].

Индекс Хирша является количественной характеристикой продуктивности ученого (или группы ученых, или научной организации, или даже страны в целом), исчисляемый на основе количества публикаций и количества цитирований (в любых изданиях) публикаций этого ученого. Ввиду простоты вычисления и необходимости хоть как-то количественно оценивать работу ученых со стороны чиновников, индекс Хирша быстро распространился на все виды научной деятельности.

Обратим внимание, что индекс Хирша был предложен и был корректен для ученых, работающих в одной той же предметной области и имеющих близкий научный стаж (от года первой публикации). Ныне это первоначальное требование зачастую нарушается, ибо идея Хирша оказалась и продуктивной (в первом приближении), и простой в вычислении.

Индекс Хирша вычисляется просто: если h статей ученого из их общего числа N_p цитируются h (или более) раз каждая, и каждая из оставшихся $N_p - h$ статей цитируется менее (или ровно) h раз, то h -индекс ученого равен (натуральному) числу h (в противном случае, если нет статей или ссылок на них $h = 0$).

Принято считать, что индекс Хирша оценивает „ядро“ (в интуитивном его понимании) публикационной активности ученого. Индекс Хирша формируется на основе свободных в доступе наукометрических баз публикационных данных через Интернет (например, Google Scholar [6], Elibrary.ru [7], ADS NASA [8]), а также на основе платных баз (Scopus или Web of Science Group).

Очевидно, индекс Хирша не оптимален для своих же целей, его недостатки отмечены уже в основополагающей статье [5]. Главный его недостаток в том, что соотношение h -индексов ученых зачастую не соответствует их вкладу в развитие соответствующей отрасли науки (подразумевается, по определению, чем больше h -индекс, тем больше вклад или „вес“ ученого в развитии его отрасли науки). К прочим недостаткам можно отнести то, что автоматизированный расчет индекса Хирша возможен только строго по одной базе и только по публикациям, проиндексированным в базе, иначе расчет потребует ручного разбора дубликатов, как публикаций, так и цитирований.

Недостатки количественной оценки вклада ученого в науку, предложенной Х. Хиршем, породили *многочисленные* предложения по улучшению этой оценки и на основе методологии Хирша, и на других принципах. См., к примеру, работы [9–20].

Еще почти четыре десятка работ с различными предложениями по модернизации идеи Хирша или „удаленными“ от Хирша нововведениям в наукометрию можно найти на сайте [21].

Все эти нововведения не свободны от серьезных недостатков. Главные из которых — это отход от привычной и уже повсеместно принятой системы Хирша, вычислительная сложность и небольшое количество мало-весомых преимуществ над h -индексом, а также психологическое неприятие дискуссионных модернизаций.

Одновременно неугасающий интерес научной общественности и организаторов науки к теме учета научной и публикационной активности ученых говорит о том, что сама эта тема является, если так можно выразиться, „фрактальной“. В этом контексте отметим работы [22–24].

Значит, наряду с уже признанной характеристикой — индексом цитируемости работ автора (в разных его модификациях) — можно (и на взгляд автора, нужно) предложить альтернативный наукометрический или библиографический показатель, который „с другой стороны“ охарактеризовал бы творчество автора, его активность и популярность.

Такой новый индекс и предлагается в статье — это новый трехмерный индекс издательской активности ученого или I -индекс для оценки внимания издательств и их журналов (с учетом их квартилей — взвешенный индекс) ко вкладу ученого в его научную отрасль. Одновременно с введением псевдонормы I -индекса строится одномерный индекс издательской активности, по значению которого издательствам (или заинтересованным лицам) удобно „ранжировать“ представителей науки.

В данной работе рекомендуется (для определенности) учитывать журналы, входящие в Перечень ВАК, и базы „более высокого уровня“ (Scopus, Web of Science и др.), что, вообще говоря, не принципиально для определения новых индексов, и может быть использована любая база данных о журнальных публикациях.

Простой I -индекс и взвешенный I^* -индекс издательской активности автора.

При оценке издательской активности автора возникает необходимость учитывать статус журналов, в которых опубликованы статьи автора. Многие журналы из перечня ВАК относятся к 3-му или 4-му квартилям, и лишь отдельные — к 1-му или 2-му квартилю. Считается (и не без оснований), что авторы работ с большей ответственностью относятся к публикации своих статей в „серьезных“ журналах из 1-ой или 2-ой квартилей. Одновременно число постоянных авторов, публикующих свои статьи в журналах 1-й и 2-й квартилей, весьма невелико. Учет квартилей журналов при анализе библиометрической информации об ученых компенсирует, в определенной мере, эту диспропорцию, а также и необычайно возросший круг научных работников по многим направлениям науки (один из критериев продолжающейся научной революции), и возросший круг научных журналов 3-й и 4-й квартилей, порою платных (спрос рождает предложение).

Определение. Для данного автора NN расположим все журналы Перечня ВАК и более „высоких“ баз (Scopus, Web of Science др.) списком сверху вниз по мере убывания числа публикаций автора в этих журналах. Для каждого журнала из списка и для каждой публикации в этом фиксированном журнале зарезервируем ячейку таблицы (матрицы), в которую проставим число, равное $k_i = 1/i$, где $i = 1, 2, 3, 4$ — та или иная квартиль журнала (на период опубликования работы автора). Журналы с одинаковым числом публикаций расположим в произвольном порядке. Получим ориентированный список (матрицу) под условным названием „ P -матрица“ размера $N_p \times M_p$, где N_p — число журналов, публиковавших работы автора NN ; M_p — максимальное число публикаций по каждому журналу, равное, очевидно, числу публикаций в верхнем журнале таблицы (по определению размеры $N_p > 0$, $M_p > 0$).

Ячейки (элементы) матрицы P , не заполненные квартилями журналов, заполним (по определению) нулями.

Определение. За „единицу измерения“ индексов принята величина $e = 1 \times 1$, т.е. квадрат со „стороной“ 1, где единица слева — это 1 (один) журнал, который опубликовал некую работу автора 1 (один) раз (единица справа).

По определению „ядро издательской активности“ (или „широта активности“) автора (или его I -индекса) — это квадратная таблица размером $h \times h$, где первое h — это число журналов, считая сверху таблицы, каждый из которых издал не менее (второе) h работ, а из оставшихся журналов нет ни одного, что опубликовал бы более h работ.

„Широту издательской активности“ автора будем характеризовать числом h .

Замечание. В случае, если для первичной оценки учет квартилей журналов не существенен, достаточно положить в „ P -матрице“ все $k_i = 1$. Получим вариант так называемой простой оценки (по определению) издательской активности автора (I -индекс издательской активности). Все остальные определения и рассуждения сохраняют силу.

Из определений легко вытекает, что „широта издательской активности“ автора h методологически (не численно!) соответствует индексу Хирша. *Конец замечания.*

По определению „хвост ядра активности“ или „глубина активности“ — это таблица, у которой „по вертикали“ стоят все журналы (или их порядковые номера), названия которых входят в ядро издательской активности; „по горизонтали“ — это статьи автора, изданные в этих журналах соответственно, которые (статьи) не вошли собственно в ядро.

По определению „подвал ядра активности“ (или „фон активности“) — это таблица, у которой „по вертикали“ стоят все журналы (или их порядковые номера), названия которых не вошли в ядро цитирования; „по горизонтали“ стоят статьи автора, изданные в этих журналах соответственно, при этом каждый журнал из подвала опубликовал ровно h работ автора.

Определение. Рассмотрим журналы и работы автора только из ядра активности. По определению положим

$$S_0 = k_1 n_{1,0} + k_2 n_{2,0} + k_3 n_{3,0} + k_4 n_{4,0}, \quad (1)$$

здесь $k_i = 1/i$ — коэффициент журнала с i -ой квартилью (индекс квартили пробегает значения $i = 1, 2, 3, 4$), где напечатана статья из ядра, а $n_{i,0}$ — число всех таких работ в журналах с квартилью i ;

Рассмотрим журналы и работы автора только из хвоста ядра активности. По определению положим

$$S_1 = k_1 n_{1,1} + k_2 n_{2,1} + k_3 n_{3,1} + k_4 n_{4,1}, \quad (2)$$

где k_i — имеет тот же смысл, что и в (1), а $n_{i,1}$ — число всех таких работ в журналах с квартилью i ;

Рассмотрим журналы и работы автора только из подвала ядра активности. По определению положим

$$S_2 = k_1 n_{1,2} + k_2 n_{2,2} + k_3 n_{3,2} + k_4 n_{4,2}, \quad (3)$$

где k_i — имеет тот же смысл, что и в (1), (2), а $n_{i,2}$ — число всех таких работ в журналах с квартилью i .

По определению положим, что взвешенный I -индекс издательской активности автора есть (упорядоченная тройка натуральных чисел):

$$I^* = (h^*, p^*, q^*), \quad (4)$$

где компоненты (индексы), вычисляемые на основе данных (1)–(3) есть

$$h^* = \left[\sqrt{S_0} \right], \quad (5)$$

$$p^* = \left[\sqrt{S_1} \right], \quad (6)$$

$$q^* = \left[\sqrt{S_2} \right], \quad (7)$$

где в формулах (5)–(7) символ $[\dots]$ — стандартная функция целой части числа.

Замечание. Содержательно смысл составляющих компонент I -индекса состоит в том, что в рассматриваемом случае „реальная“ „ P -матрица“ заменяется „виртуальной“ „ P -матрицей“, в которой все клетки заполнены коэффициентами $k_i = 1$. Получаем некий вариант простой оценки издательской активности автора.

В этой виртуальной ситуации индекс $I^* = (h^*, p^*, q^*)$ предоставляет информацию о том, что в ядре „содержится“ h^* журналов, каждый из которых „опубликовал“ h^* работ автора; в хвосте „остались“ неучтенными еще p^* работ автора, „издаваемые“ как бы теми же самыми журналами, а в подвале „находятся“ другие журналы, опубликовавшие то же h^* работ автора. *Конец замечания.*

Индекс (4) позволяет дать качественную и количественную оценку издательской широты и его „окружения“ в виде глубины и фона. Но тройка чисел (4) психологически воспринимается (неподготовленным пользователем) достаточно трудно. В целях удобства сравнения издательской активности ученых (обозначим ее $A(NN)$ для ученого NN) в целях поддержания финансирования работ или кадрового роста персонала и т. п. введем (по определению) псевдонормированный (одномерный) индекс δ^* (издательская активность ученого NN).

По определению индекс

$$\delta^* = \left[\sqrt{(h^*)^2 + (p^*)^2 + (q^*)^2} \right], \quad (8)$$

где в формуле (8) символ $[\dots]$ имеет тот же смысл, что и в (5)–(7).

По определению для двух авторов NN_1 и NN_2 их издательская активность сравнивается следующим образом:

1) Одномерная или сильная (или псевдонормированная) сравнимость

$$A(NN_1) \gg A(NN_2) \stackrel{def}{\iff} \begin{cases} \delta_1^* > \delta_2^*, \\ \delta_1^* = \delta_2^* \wedge h_1^* = h_2^* > h_2^*, \\ \delta_1^* = \delta_2^* \wedge h_1^* = h_2^* \wedge p_1^* > p_2^*, \\ \delta_1^* = \delta_2^* \wedge h_1^* = h_2^* \wedge p_1^* = p_2^* \wedge q_1^* > q_2^*. \end{cases} \quad (9)$$

В противном случае, по определению, издательские активности авторов (в этом измерении) считаются одинаковыми: $A(NN_1) \equiv A(NN_2)$.

2) Если компоненты вектора $I^* = (h^*, p^*, q^*)$ проранжировать по степени „важности“, полагая (по определению), что h^* „важнее“ p^* , а p^* , в свою очередь, „важнее“ q^* , полагая также, что бинарное отношение „важности“ транзитивно, то можно ввести многомерную или слабую сравнимость по правилу:

$$A(NN_1) > A(NN_2) \stackrel{def}{\Leftrightarrow} \begin{cases} h_1^* > h_2^*, \\ h_1^* = h_2^* \wedge p_1^* > p_2^*, \\ h_1^* = h_2^* \wedge p_1^* = p_2^* \wedge q_1^* > q_2^*. \end{cases} \quad (10)$$

В противном случае, по определению, издательские активности авторов (в этом измерении) считаются одинаковыми: $A(NN_1) = A(NN_2)$.

Замечание. В формулах (9), (10) знак \wedge — (стандартный) символ бинарной операции конъюнкции двух высказываний, более слабый (по определению, во избежание „лишних“ скобок) по сравнению с соседними символами равенства или неравенства.

С целью детализации информации индексы h^* , p^* , q^* , δ^* можно вычислять с точностью до десятых долей. *Конец замечания.*

Алгоритм вычисления I^* -индекса. Приведем алгоритм вычисления I^* -индекса, давая при необходимости пояснения к его шагам.

1. Для данного автора NN расположим все журналы выбранной базы данных (Перечня ВАК и в более „высоких“ базах (Scopus, Web of Science др.)), в которых опубликованы его работы, списком сверху вниз по мере убывания числа публикаций его статей в этих журналах.

Журналы с одинаковым числом публикаций располагаются в произвольном порядке. Получим список под условным названием „ P -список“ автора NN с общим числом журналов N_p .

2. На основе P -списка построим P -матрицу без проставления коэффициентов $k_i = 1/i$ на основе квартилей $i = 1, 2, 3, 4$ журналов. Полученную матрицу назовем и обозначим (для определенности) P_0 -матрицей.

3. На основе P_0 -матрицы составим ядро активности автора, хвост ядра и подвал ядра.

4. Построим P -матрицу, проставив коэффициенты $k_i = 1/i$ на основе квартилей $i = 1, 2, 3, 4$ журналов в P_0 -матрице.

5. По формулам (1), (2), (3) вычислим значения S_0, S_1, S_2 соответственно.

6. По формулам (5), (6), (7) вычислим значения h^*, p^*, q^* соответственно.

7. По формуле (4) вычислим значения (трехмерного) индекса издательской активности I^* .

8. По формуле (4) вычислим значения (трехмерного) индекса издательской активности I^* .

9. По формуле (8) вычислим значения (одномерного) индекса издательской активности δ^* .

При необходимости сравнения активностей двух авторов NN_1 и NN_2 вычислим их активности по приведенному алгоритму и дополним алгоритм еще одним, следующим шагом.

10. По формулам (9) или (10) сравним активности $A(NN_1)$ и $A(NN_2)$ двух авторов.

Некоторые свойства новых индексов. Определение. Будем говорить, что индекс X (где синтаксическая переменная X может принимать значения $X = h^*, p^*, q^*, I^*, \delta^*$) устойчив в области D (где синтаксическая переменная D может принимать значения $D =$ „ядро“, „хвост“, „подвал“), если никакие одна-две статьи в новых или старых журналах из области D „ P -списка“ не могут с большой долей вероятности изменить значения индекса X .

В противном случае индекс X неустойчив в области D .

Будем говорить, что индекс X глобально устойчив, если он устойчив при всех значениях переменной D .

Тезис 2. Индексы h^*, p^*, q^* — устойчивые характеристики в своих областях; индекс δ^* — глобально устойчивая характеристика.

Доказательство. В самом деле, нетрудно посчитать (во избежание ненужных для сущности работы объемных вычислений данный расчет опущен) классические (статистические) вероятности изменения характеристик h^*, p^*, q^*, δ^* при одной и двух новых статьях автора из области D и убедиться в относительно небольших значениях этих вероятностей.

Примеры. Пример 1. Один журнал 1-ой квартили опубликовал 100 работ автора NN_1 , других публикаций у автора нет. Из определений вытекает, что $I^* = (1, 9, 0)$, $\delta^* = 9$.

Автор NN_2 опубликовал 100 работ в 100 разных журналах 1-ой квартили. Из определений вытекает, что $I^* = (1, 0, 9)$, $\delta^* = 9$.

Сравнивая активности авторов, получаем: $A(NN_1) \gg A(NN_2)$ и $A(NN_1) > A(NN_2)$.

Пример 2. Автор NN_3 опубликовал свои работы в 10 журналах, все 1-й квартили, из которых 1-й журнал опубликовал 100 работ, 2-й журнал — 90 работ, 3-й — 80 работ и т.д., последний 10-й журнал опубликовал 10 работ. Из определений вытекает, что $I^* = (10, 21, 0)$, $\delta^* = 23$.

Пример 3. Автор NN_4 опубликовал работы в 30 журналах 1-й квартили, из которых 10 журналов опубликовали по 100 работ, а оставшиеся 20 журналов опубликовали по 10 работ. Из определений вытекает, что $I^* = (10, 30, 10)$, $\delta^* = 33$.

Если в примере 3 все журналы были 3-ей квартили, то из определений вытекает, что $I^* = (5, 17, 5)$, $\delta^* = 18$.

Уже из этого примера видно, что учет квартилей журналов может привести к существенному изменению издательского „статуса“ автора публикаций.

Заключение. Таким образом, автор представил новый наукометрический индекс — индекс издательской (или публичной) активности автора, зависящий от числа публикаций автора в различных журналах.

Информативно индекс издательской активности автора характеризует широту научных интересов автора, внимание к его работам со стороны издательств и (косвенно) уровень научной значимости его результатов.

Новые индексы просты в использовании, основаны на уже привычной для научного сообщества методологии Хирша и одновременно „с другой стороны“ дают более качественное представление о вкладе ученого в развитие его отрасли науки.

Вопрос о связи (или корреляции) публикационной активности автора и его издательской активности представляется весьма интересным и нетривиальным и вполне может составить тему научно-квалификационной работы, например, по специальности 05.13.10 — „Управление в социальных и экономических системах“.

В заключение я хочу выразить слова благодарности моей дочери Кристине Мироновой (канд. техн. наук, ведущему специалисту крупной IT-компании) за помощь в анализе очень большой информации по наукометрической проблематике в Интернете и полезные советы и замечания.

Список литературы

1. Григорьева Е. И. Наукометрия — наука важная, интересная, полезная. 2020. [Электронный ресурс]: <http://www.isran.ru/> (дата обращения: 14.01.2021).

2. Миронов В. В. Новые индексы публикационной активности // Вестник РАН. 2020. Т. 90. № 10. С. 959–966.

3. Миронов В. В. Обработка данных о публикационной активности автора в составе авторского коллектива с учетом квартилей журналов // Научно-техническая информация. Серия 2. Информационные процессы и системы. 2020. № 11. С. 30–35.
4. Миронов В. В. Информатизация образования: достижения и проблемы // Информатизация образования и науки. 2017. № 4 (36). С. 3–18.
5. Hirsch J. E. An index to quantify an individual's scientific research output // Proceedings of the National Academy of Sciences of the United States of America. 2005. Vol. 102. № 46. P. 16569–16572.
6. Google Scholar. [Электронный ресурс]: <https://scholar.google.com/> (дата обращения: 15.01.2021).
7. Elibrary.Ru, Научная Электронная Библиотека. [Электронный ресурс]: <https://elibrary.ru/> (дата обращения: 15.01.2021).
8. Astrophysics data system (ADS NASA). [Электронный ресурс]: <http://adsabs.harvard.edu/> (дата обращения: 17.01.2021).
9. Tagiew R., Ignatov D. I. Behavior mining in h-index ranking game // CEUR Workshop Proceedings. 2017. Vol. 1968. P. 52–61.
10. Имаев В. Технологии увеличения индекса Хирша и развитие имитационной науки // Комиссия РАН по борьбе с лженаукой и фальсификацией научных исследований. В защиту науки. 2016. № 17. С. 38–51.
11. Демина Н. Хиршемания и хиршефобия. Троицкий вариант — Наука. [Электронный ресурс]: <https://trv-science.ru/2016/12/06/khirshemaniya-i-khirshefobiya/> (дата обращения: 17.01.2021).
12. Михайлов О. В. Новая версия индекса Хирша – j-индекс // Вестник РАН. 2014. Т. 84. № 6. С. 532–535.
13. Egghe L. Theory and practise of the g-index // Scientometrics. 2006. Vol. 69. N 1. P. 131–152.
14. Kosmulski M. I. A bibliometric index // Forum Akademickie. 2006. Vol. 11. P. 31.
15. Prathap G. Hirsch-type indices for ranking institutions' scientific research output // Current Science journal. 2006. Vol. 91 (11). P. 1439.
16. Jones T., Huggett S., Kamalski J. Finding a Way Through the Scientific Literature: Indexes and Measures // World Neurosurgery. 2011. Vol. 76. № 1, 2. P. 36–38.
17. Холодов А. С. Об индексах цитирования научных работ // Вестник РАН. 2015. Т. 85, № 4. С. 310–320.
18. Мазов Н. А, Гуреев В. Н. Альтернативные подходы к оценке научных результатов // Вестник РАН. 2015. Т. 85, № 2. С. 115–122.
19. Кузнецов А. В. Для начала надо навести порядок в существующей системе РИНЦ. Письма в редакцию // Вестник РАН. 2014. Т. 84, № 3. С. 268–269.
20. Ludo Waltman, Nees Jan van Eck. Robust Evolutionary Algorithm Design for Socio-Economic Simulation: Some Comments // Comput. Econ. 2009. N 33. P. 103–105, DOI 10.1007/s10614-008-9148-0.
21. Hirsch J. E. h_α : An index to quantify an individual's scientific leadership. 2019. [Электронный ресурс]: <http://link.springer.com>article/10.1007/s11192-018-2994> (дата обращения: 18.01.2021).
22. Бредихин С. В., Щербакова Н. Г. Формализация индекса Хирша. Обзор // Проблемы информатики. 2014. № 2 (23). С. 17–27.
23. Моисеенко В. В., Родионов А. С. Возрастная зависимость соавторства при проведении научных исследований в академическом институте // Проблемы информатики. 2017. № 1 (34). С. 62–73.
24. Бредихин С. В., Ляпунов В. М., Щербакова Н. Г. Динамика роста сети цитирования научных статей // Проблемы информатики. 2020. № 1. С. 5–20.



Миронов Валентин Васильевич — доктор физико-математических наук, профессор по специальности 05.13.01. „Системный анализ, управление и обработка информации“. E-mail: mironov1vv@mail.ru; моб. + 7 (903) 839-40-85.

Основные области научных интересов: высшая алгебра и алгебра логики, дискретная математика, дифференциальная геометрия и дифференциальные уравнения, теория управления и устойчивости, системный анализ, философия.

Основные результаты: автор более 160 научных работ в разных областях математики. Подготовил 1 доктора наук и 4 кандидатов наук, в настоящее время – руководитель работ 2-х докторантов и 2-х аспирантов.

Профессор кафедры высшей математики Рязанского государственного радиотехнического университета (РГРТУ) им. В. Ф. Уткина, директор лаборатории системного анализа, президент Рязанского физико-математического общества, заместитель Генерального директора по Центральному Федеральному округу Международного фонда ученых и инженеров имени академика В. Ф. Уткина.

Поэт, издал несколько сборников стихов и рассказов. Рисует, организовал выставку своих художественных работ и инсталляций.

Mironov Valentin Vasilievich — Doctor of Physics and Mathematics, Professor, specialty 05.13.01. „System analysis, management and information processing“. E-mail: mironov1vv@mail.ru; phone no. + 7 (903) 839-40-85

Main areas of research interests: higher algebra and algebra of logic, discrete mathematics, differential geometry and differential equations, control and stability theory, systems analysis, philosophy.

Main results: author of over 160 scientific papers in various fields of mathematics. Prepared 1 Doctor of Science and 4 Candidates of Science, at present — the head of the work of 2 doctoral students and 2 postgraduates.

Professor of the Department of Higher Mathematics, Ryazan State Radio Engineering University (RSREU) named after V.F. Utkin, Director of the Laboratory of Systems Analysis, President of the Ryazan Physics and Mathematics Society, Deputy Director General for the Central Federal District of the International Fund for Scientists and Engineers named after Academician V.F. Utkin.

Дата поступления — 24.01.2021

Правила представления и подготовки рукописей для публикации в журнале „ПРОБЛЕМЫ ИНФОРМАТИКИ“

Общие требования.

Редакция принимает к рассмотрению статьи в электронном виде (*исходный файл в L^AT_EX* и файл PDF, либо MS Word с приложением оригиналов рисунков в формате тех программ, в которых они были сделаны, отдельными файлами).

Файлы, содержащие текст статьи, иллюстрации и дополнительные материалы, можно пересылать на электронный адрес редакции: problem-info@sscc.ru.

Принимаются файлы, архивированные архиваторами ZIP/7Z или RAR; применение самораспаковывающихся архивов не допускается.

При повторной отправке материалов, а также при внесении в исходный текст дополнений или исправлений необходимо сообщить об этом в редакцию в тексте электронного письма.

Текст статьи с формулами, рисунками, таблицами должен быть подготовлен на стандартном листе формата A4 через 1,5 интервала, размер шрифта 12 pt, все поля по 2 см.

Статьи, содержащие формулы, следует набирать в редакторе L^AT_EX. В остальных случаях допускается использование программы MSWord, шрифт — TimesNewRoman; автоматическая расстановка переносов в документе должна быть отключена.

Направляя статью в редакцию журнала, автор (соавторы) на безвозмездной основе передает(ют) издателю на срок действия авторского права по действующему законодательству РФ исключительное право на использование статьи или отдельной ее части (в случае принятия редколлегией Журнала статьи к опубликованию) на территории всех государств, где авторские права в силу международных договоров Российской Федерации являются охраняемыми, в том числе следующие права: на воспроизведение, на распространение, на публичный показ, на доведение до всеобщего сведения, на перевод на иностранные языки и переработку (и исключительное право на использование переведенного и (или) переработанного произведения вышеуказанными способами), на предоставление всех вышеперечисленных прав другим лицам.

Журнал „Проблемы информатики“ является некоммерческим изданием. Плата с авторов за публикацию статей не взимается.

К статье должны быть приложены:

— **разрешение на публикацию** от экспертного совета организации, в которой выполнена работа (для авторов из России);

— **оригинал рецензии;**

портретные фотографии авторов разрешением не менее 300 dpi.

— Блоки информации и на русском, и на английском языках просьба присылать отдельными файлами:

— **Название** статьи;

— **Инициалы и фамилии** авторов;

— **Места работы** авторов: полное наименование организации, почтовый индекс, город, страна;

— **Код(ы) классификации УДК;**

— **Аннотации**, содержащие краткую постановку задачи и описание метода решения: на русском языке объемом не более 1000 знаков, на английском языке расширенную, объемом от 4000 до 8000 знаков, что соответствует требованиям ВАК и Scopus.

— **Ключевые слова;**

— **Списки используемой литературы** в соответствии с ГОСТ Р 7.0.5—2008 (в английской версии необходимо выполнить **транслитерацию** неанглоязычных элементов списка литературы в соответствии с ГОСТ Р 7.0.34-2014) — составляются по ходу упоминания источников в тексте;

— **Краткие биографии (БИО)** авторов с указанием ключевых научных достижений (включая ученую степень, ученое звание — при наличии; основные области научных интересов и формулировку основных результатов, место работы, занимаемую должность, контактные данные — почтовый адрес с индексом, адрес электронной почты, контактный телефон).

Подготовка статьи.

1. Материал статьи должен быть изложен в следующей последовательности:

1. название статьи на английском языке;
2. инициалы и фамилия автора(ов) на английском языке;
3. место работы автора(ов) (на английском языке): полное наименование организации, индекс, город, страна;
4. англоязычная аннотация;

5. ключевые слова на английском языке;
6. references+транслитерация неанглоязычных элементов списка литературы;
7. название статьи на русском языке;
8. инициалы и фамилии и авторов;
9. место работы авторов: полное наименование организации, почтовый индекс, город, страна;
10. индекс УДК;
11. аннотация на русском языке;
12. ключевые слова (не более 8);
13. текст статьи;
14. список литературы, оформленный в соответствии с требованиями ГОСТ;
15. краткие биографии авторов на английском и русском языках с указанием ключевых научных достижений (ученую степень, ученое звание — при наличии; место работы, занимаемую должность, контактные данные — почтовый адрес с индексом, адрес электронной почты, контактный телефон, основные области научных интересов и формулировка основных результатов).

2. Требования к формулам:

- Нумерация формул сквозная, выносные формулы центрируются, номер выровнен по правому краю.

3. Требования к рисункам:

- Файлы с рисунками присылаются отдельно в формате программ, в которых они были выполнены: в формате MS Excel (для графиков и диаграмм), eps, pdf, png, tiff, bmp или jpeg (с максимальным качеством).

- Рисунки с подрисуночными подписями завершаются в текст статьи.

- Тексты, являющиеся частью рисунка, выполняются шрифтом TimesNewRoman.

- Фотографии должны иметь разрешение не менее 300 dpi.

4. Дополнительные требования:

- В текст статьи необходимо включать ссылки на рисунки и таблицы, а также подрисуночные подписи и заголовки таблиц. Все буквенные обозначения, приведенные на рисунках, необходимо пояснить в основном тексте или в подрисуночных подписях.

- Сокращения слов не допускаются (кроме общепринятых).

- Векторные переменные обозначаются полужирным шрифтом без курсива.

- Таблицы не должны быть громоздкими. Значения физических величин в таблицах, на графиках и в тексте должны указываться в единицах измерения СИ.

- Графики, если их на рисунке несколько, а также отдельные детали на чертежах, узлы и линии на схемах следует обозначать цифрами, набранными курсивом.

- Нумеровать следует только те формулы и уравнения, на которые имеются ссылки в тексте, нумерация сквозная.

- Ссылки на источники в тексте заключаются в квадратные скобки.

- Иностранные источники приводятся на языке оригинала. Ссылки на неопубликованные работы не допускаются.

Все статьи, опубликованные в журнале „Проблемы информатики“, доступны на сайте https://elibrary.ru/title_about.asp?id=30275 и на сайте журнала <http://problem-info.sssc.ru> спустя год после опубликования.

Пример оформления статей можно посмотреть на сайте журнала <http://problem-info.sssc.ru>.