Modeling self-organizing teams in a research environment

Tamara V. Voznesenskaya [a]
E-mail: tvoznesenskaya@hse.ru

Fedor V. Krasnov [b]
E-mail: krasnov.fv@gazprom-neft.ru

Rostislav E. Yavorsky [a]
E-mail: ryavorsky@hse.ru

Polina V. Chesnokova [c]
E-mail: chepolina27@gmail.com

[a] National Research University Higher School of Economics
Address: 20, Myasnitskaya Street, Moscow 101000, Russia

[b] Gazprom Neft PJSC
Address: 75–79 liter D, Moika River Embankment, Saint Petersburg 190000, Russia

[c] LANIT Group
Address: 14 block 1, Murmansky Proezd, Moscow 129075, Russia

Abstract

This paper presents the results of studies aimed at analyzing the effectiveness of a research center. The study focuses on the process of self-organization of project teams (groups of co-authors) for project implementation (writing a scientific article). The initiative to create a team comes from one of its members. The paper describes a formal model, based on a competence approach, which considers the types of tasks to be solved and the necessary skills of the staff. The paper also presents the results of simulation in the AnyLogic environment and problems for further research.

The competency profile of each employee is a vector where each coordinate describes the level of mastery of the corresponding skill. The competency profile of the team is a vector obtained as the result of simple addition of the competency profiles of the participants. The proposed model assumes that each task requires a certain set of competencies and that the list of competencies and the level of experience are the criteria for deciding whether to join the team. The logic of decision making at various stages of team creation is modelled by functions. At each step of the modelling, the next employee is chosen randomly. To calibrate the team member's competency profile, internal data on employee qualifications of the Gazpromneft Research Center was used. The constructed model is the basis for further studies of the process by which project teams are created and function in a scientific environment and for developing a methodology to assess the effectiveness of the work of research teams. It helps to predict the need for personnel with different competencies, plan activities to improve the skills of employees and strengthen communication in the team.

Key words: team information model; simulation modeling; characteristics of research work; scientometrics; organizational hypothesis.

Citation: Voznesenskaya T.V., Krasnov F.V., Yavorsky R.E., Chesnokova P.V. (2019) Modeling self-organizing teams in a research environment. *Business Informatics*, vol. 13, no 2, pp. 7–17. DOI: 10.17323/1998-0663.2019.2.7.17

Introduction

Analysis of the effectiveness of basic research and research and development, as well as studies of factors influencing research effectiveness have been discussed in many publications (e.g., [1–4]). As a rule, these publications consider the research team as a "black box" that outputs research results. Effectiveness analysis only considers results, while the inner structure of the research group is usually not taken into account. The self-organizing team is studied in detail in [5]. Other publications study motivating factors [4] and factors impacting effectiveness [3]. Teamwork modeling and analysis has also been studied in detail since the mid-20th century [6–8]. A formal description of competency profiles is the subject of numerous studies and publications (e.g., [8, 9]).

In the present article, we propose using a mathematical model of team formation based on competencies in order to increase the effectiveness of research management[1]. We employ a simplified model that posits the existence of a fixed set of skills. The competency profile of each team member can be represented as a vector, each of whose coordinates describes the level of the corresponding skill. The vector representing the competency profile of the team is obtained by simply adding the competency profiles of participants. Such a model assumes that a competency is measured by the effectiveness of solving the corresponding type of problem. It is natural to posit that, in the case of teamwork, the effectiveness of team members adds up. A similar vector can be used to represent the problem profile. To prepare and implement the research project, a certain level of effectiveness is needed for each type of problem in order to be able to finish on time. A rigorous mathematical description of this model is given in Section 2.

This article has the following structure. Section 1 gives an informal description of the competency model and the formation of a coauthor team. Section 2 elaborates mathematical definitions and an algorithm for team formation. The computer experiment and its results are described in Section 3. The Conclusion lists open questions and areas for further study.

1. Coauthor team formation

This section gives an informal description of the formation of a coauthor team and the principal assumptions involved. They will be subsequently expressed in mathematical language in Section 2.

1.1. Assumptions and restrictions

The present article studies small self-organizing teams (with up to six members [10]), in

[1] Preliminary results of the study were presented in the graduate work by Polina V. Chesnokova, defended at the HSE Faculty of Computer Science in 2018

which the initiative of team formation is taken by the staff members themselves. This assumption reflects the state of things with regard to most research teams: the administration may try to motivate staff members to submit applications to research conferences or write articles for research journals, yet the final decision is made, as a rule, by the staff members themselves.

It is assumed that the list of competencies and amount of experience are the main criteria for a staff member's decision to join the team.

The model's input data are the set of topics specified in incoming invitations to submit applications to journals and conferences. One or several topics are given for each journal or event. The preparation of an article on the given topic requires a certain set of competencies.

It is assumed that the model shall be used to analyze the work of medium or large teams with a membership of between several dozen and several hundred people over a long period (1–3 years).

1.2. Start of team creation

Team formation begins with the decision taken by the first team member to create a team for drafting an application to a conference or an article for a journal or edited volume. The procedure is as follows. A staff member looks through the list of invitations and assesses his competencies with regard to each topic. If at least one of his competencies matches or surpasses the requirements, he decides to form a team and becomes its first member. At the beginning, the team's competency profile coincides with the competency profile of the first member. Subsequent members join the team on the basis of the requirements of the given topic and the competency profiles of other team members.

1.3. Recruitment of new team members

The second (or subsequent) member learns from one of the team members about the project goal and the current team competency profile. This information circulates among staff members, who are acquainted with each other. We assume that, in comparison to the total project duration, one can neglect the time spent on communication during team formation. The model represents this communication by a graph. Each subsequent team member assesses his competencies from the standpoint of the team's needs for attaining its goal and makes a decision about joining the team. This decision shall be affirmative if at least one of the competencies of this staff member helps the team to attain its goal.

1.4. Finalizing the team

Given the limited time allocated for solving the problem, the time for team formation is limited as well. If a team with the required set of competencies has not been formed by the deadline, the process stops, the team members are relieved of their responsibilities, and they are free to look for other problems.

If the team is formed successfully, then one can assume that its members will be occupied for some time to come and that their joint work will lead to a publication.

2. Model

In this section, we present the formal mathematical definitions of the notions and procedures described above.

2.1. Formal competency model

Let N denote the number of key skills necessary for working in the given subject area and W – the set of staff members of the organization. Then the competency profile of the staff member is the vector

$$\vec{k}(w) = k_1, ..., k_N, \text{ где } w \in W, k_i \in R^+.$$

The competency profile of the team T consisting of M members is a vector of the same dimension N that is equal to the sum of vectors of individual team members:

$$\vec{k}(T) = \sum_{i=1}^{M} \vec{k}(w_i),$$

$$\text{where } T = \{w_1, \ldots, w_M : w_i \in W\}.$$

The i-th component of the vector informally corresponds to the effectiveness of the member or team in solving a problem of the given type.

The profile of the problem p is also an N-dimensional vector:

$$\vec{k}(p) = (k_1, ..., k_N).$$

Here the i-th component of the vector corresponds to the minimal effectiveness of the team assuring that all the problems of the corresponding type will be solved on time and with the required quality.

2.2. Model for making the key decisions

The key functions of team formation are the functions that model the decision-making logic at the different stages of the process:

✦ $\alpha(w, p)$ describes the choice of goal by the first team member: $\alpha(w, p) = 1$ if staff member w considers goal p and takes a positive decision to form a team and $\alpha(w, p) = 0$ otherwise;

✦ $\beta(w, T, p)$ formalizes the decision-making process with regard to the recruitment of the second and subsequent members to the team;

✦ $\gamma(T, p, t)$ models the decision of voluntary team dissolution at time t based on a comparison between the team profile and the problem profile.

For the purposes of the present article, we assume that α, β and γ are functions that depend only on the competency profiles of the team member, the team and the problem, respectively. Functions α, β and γ take values on the set {0, 1}. Thus

$$\alpha(w, p) = \alpha'(\vec{k}(w), \vec{k}(p)),$$

$$\beta(w, T, p) = \beta'(\vec{k}(w), \vec{k}(T)\, \vec{k}(p)),$$

$$\gamma(T, p, t) = \gamma'(\vec{k}(T), \vec{k}(p), t).$$

Furthermore, let K denote the space of possible values of the competency vector. Then the fact that the team formation algorithm in our model depends only on the competency profiles of the team member, team and goal results in the function types α', β' and γ':

$$\alpha' : K^2 \to \{0, 1\},$$

$$\beta' : K^3 \to \{0, 1\},$$

$$\gamma' : K^2 \to \{0, 1\}.$$

These functions can be described with the following logical formulae:

$$\alpha'(x, y) = 1 \Leftrightarrow \exists\, i\, (x_i \geq y_i)$$

$$\beta'(x, y, z) = 1 \Leftrightarrow \exists\, i\, [(x_i > y_i) \wedge (y_i < z_i)]$$

$$\gamma'(x, y, t) = 1 \Leftrightarrow \exists\, i\, (x_i < y_i) \wedge (t > \tau_{\max}).$$

2.3. Team formation

When team formation is complete, a list of open problems P is specified, and a profile $k(p)$ is defined for each problem $p \in P$. The set of staff members W is also specified, and a competency profile $k(w)$ is assigned to each staff member $w \in W$. A communication graph $G \subseteq W \times W$ is also given for the staff members. Another parameter is the time $\tau_{\max}$ during which team formation must be complete.

At each step, the following occurs:

1. Every staff member w_0 that is not on a team and has not received an invitation to join a team considers the list of goals P. If he finds a p_0 for which $\alpha(w_0, p_0) = 1$, then the staff members makes the decision to create the new team T_0 and sends an invitation to all his neighbors on the communication graph G to join the team.

2. If the staff member w_1 is not a member of team T_1, that was created to solve problem p_1 and has not received an invitation to join it, then he accepts the invitation if $\beta(w_1, T_1, p_1) = 1$ and sends an invitation to all his neighbors on the graph G. Otherwise, the invitation is declined.

3. If, for the team T_2, created for solving problem p_2 the condition $\gamma(T_2, p_2) = 1$ holds, then the team sets to work, and all invitations are annulled.

4. If, for the team T_3 created for solving problem p_3, the condition $\gamma(T_3, p_3) = 0$ holds after the fixed time τ_{max}, then the team is dissolved, and all invitations are annulled.

Although α, β and γ are determined, the algorithm allows for a large degree of indeterminacy that stems from the undetermined nature of the objects within the system. In particular, the result depends on the following parameters, which are random in nature:

✦ order in which a free staff member considers the list of problems;

✦ order in which a staff member considers the received invitations;

✦ order in which staff members are chosen for applying a step of the algorithm.

3. Computer experiment

The key parameters of the model that are required for the computer experiment are:

✦ structure of the space K of possible values of the competency vector;

✦ characteristics of the staff communication graph $G \subseteq W \times W$;

✦ distribution of competency profiles among staff members $k(w)$;

✦ characteristics of incoming problems $k(P)$.

The configuration of the above parameters depends on the specifics of the modeled organization. Some of them can be measured or easily calculated by obtaining accessible information from the personnel department or analyzing data from corporate information systems. For other parameters, the choice is made by solving the inverse problem, i.e., selecting values for which modeling results best match the parameters of teams observed in practice [11, 12].

In this article, the input data for modeling the interaction of staff members was taken from telephone logs as well as office parameters such as the schedule of company events and meetings, typical staff itineraries, etc. The information on competencies was taken from an analysis of the publication activities of the organization.

The model described above was run in the AnyLogic environment [13, 14]. The base classes *Competency, Topic, Staff Member, Team, Publication* and the aforementioned process of team formation were run. The computer experiments were made on a high-performance computing cluster consisting of ten computers with network attached storage.

3.1. Research topic

In the field of oil and gas production and treatment, the list of current research topics constantly changes. Changes in research topics in the oil and gas industry were studied in [15]. A list of current topics appeared in a leading geophysics journal in 2018[2], for example. Moreover, the calendar of key theoretical and applied research conferences is known ahead of time[3]. These data were used to configure the parameters of the set of current research problems P: the problem topics were taken from the *Neftyanoye khozyaystvo* ("Oil Industry") journal over the period 2011–2017.

[2] http://fb.eage.org/index/guidelinesforauthors?p=103
[3] http://www.spe.org/events/calendar/

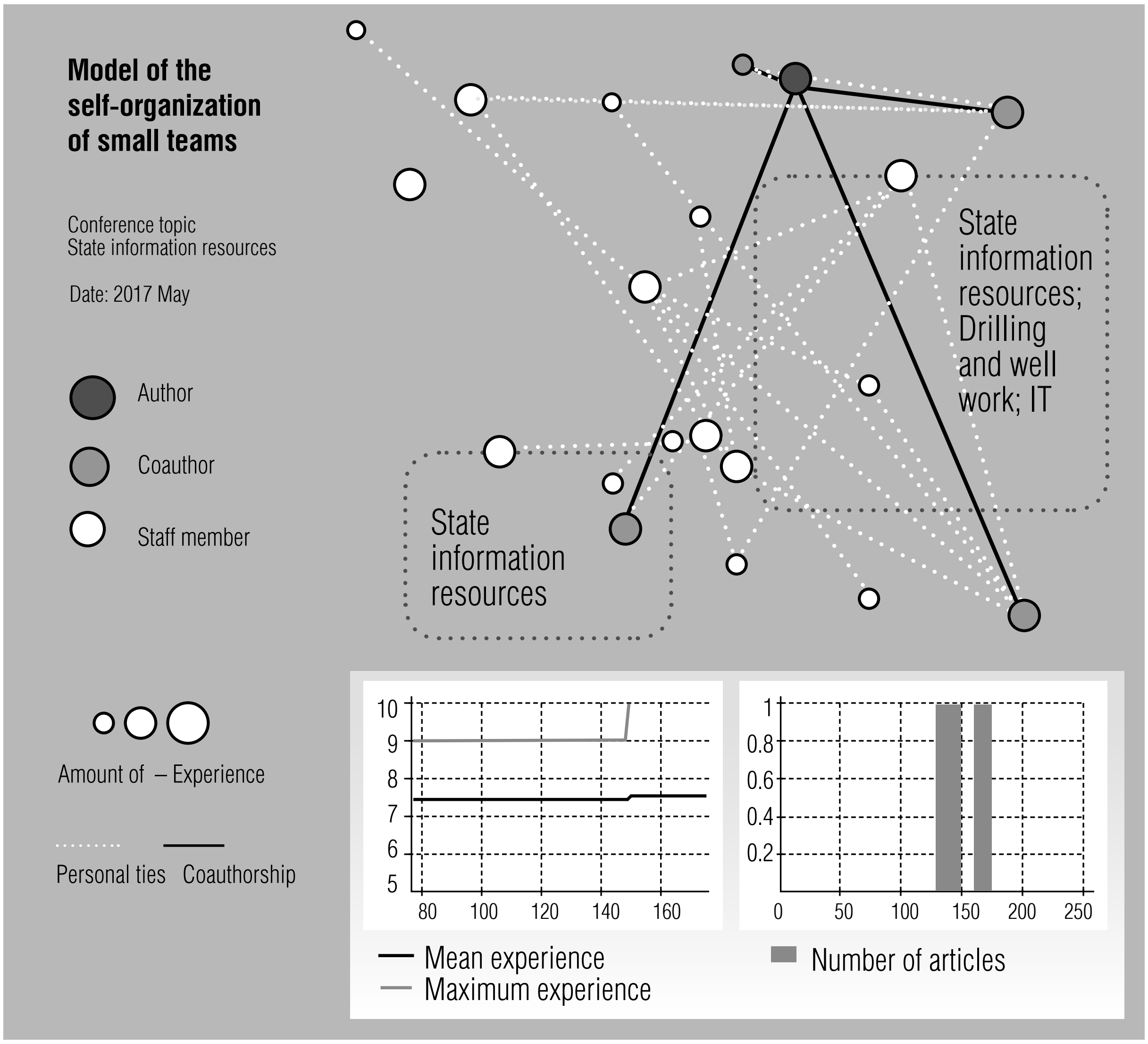

Fig. 1. A step in the simulation process

Actors took decisions on team creation and coauthorship in accordance with the aforementioned algorithm. *Figure 1* shows one of the steps of the simulation process. Some staff members are already working together, as one can see from the coauthorship graph.

3.2. Competency profiles

For the purposes of calibrating staff competency profiles, we used company data on staff qualifications of the GazpromNeft Science and Technology Center (these data are constantly collected and updated). In all, 20 staff members from a single department took part in the experiment. We used the personnel management model given in [16] to keep track of the recruitment and dismissal of staff members.

3.3. Communication graph of an organization

To make a communication graph, one analyzes communication by telephony, SMS, and email, communication at meetings, and

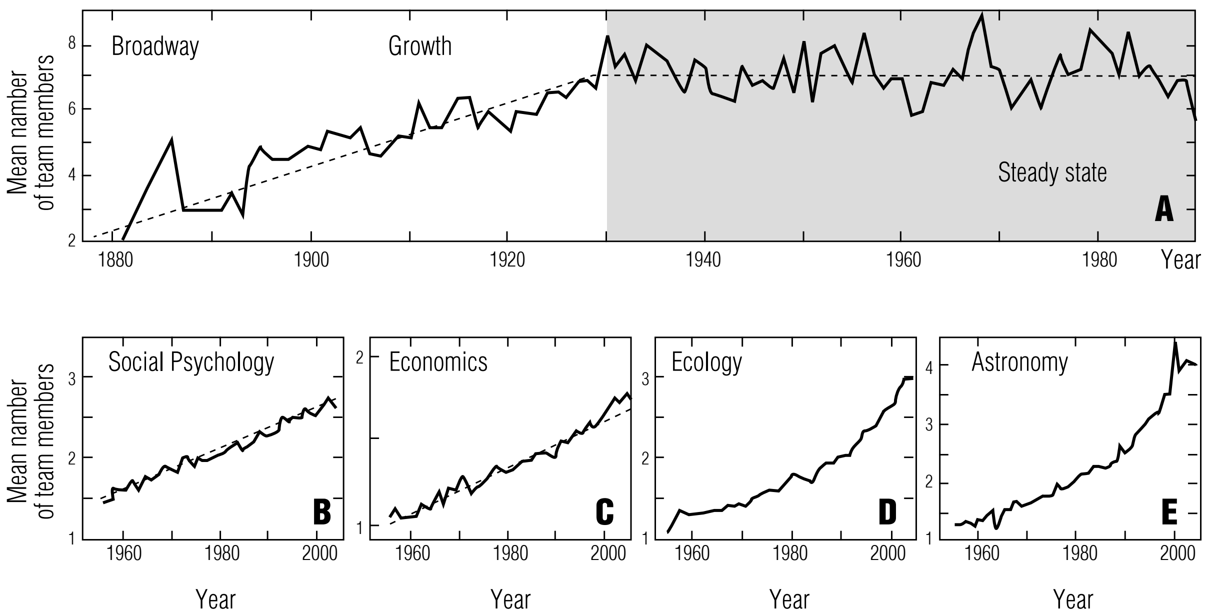

Fig. 2. Mean size of the coauthor team by research field [10]

communication for the purposes of reserving rooms, organizing trips, receiving technical support, etc. [17, 18]. Note that such attributes of communication as calendar entries, meeting proceedings, attached documents and protocols may be digital even if the content of communication channels is non-digital.

For the purposes of the present study, it is important to note that one can draw the communication graph G by analyzing data from the corporate information system, if available.

The principal factor in team formation is personal ties: if the coauthors do not know each other, the probability of them writing a joint article amounts to zero.

In the present experiment, we disregarded many social factors described in the books [19, 20]. For example, personal affinity [21] or respect for rank [22] may change the pattern of staff interaction in the office. The impact of social factors greatly depends on cultural specifics and varies between countries and regions. For the purposes of the present article, we took personal ties as the key element of team formation. In subsequent studies, we may add other socio-cultural factors to this model.

3.4. Coauthor team size

A detailed analysis of team size is given in the article [10], which studies the mean size of coauthor teams in different fields of research (*Figure 2*). Today, the mean number of coauthors of an article ranges from 2 to 6. This data was used to calibrate the model and, more precisely, to refine the topics and competency scale for staff members.

3.5. Results of the computer experiment

Conducting computer experiments based on the above model allowed experts to use the communication graph to identify potential professional ties that could greatly improve the effectiveness of work. These results may be used for organizing company events (seminars and conferences) with a well-defined list of participants to promote maximum effectiveness.

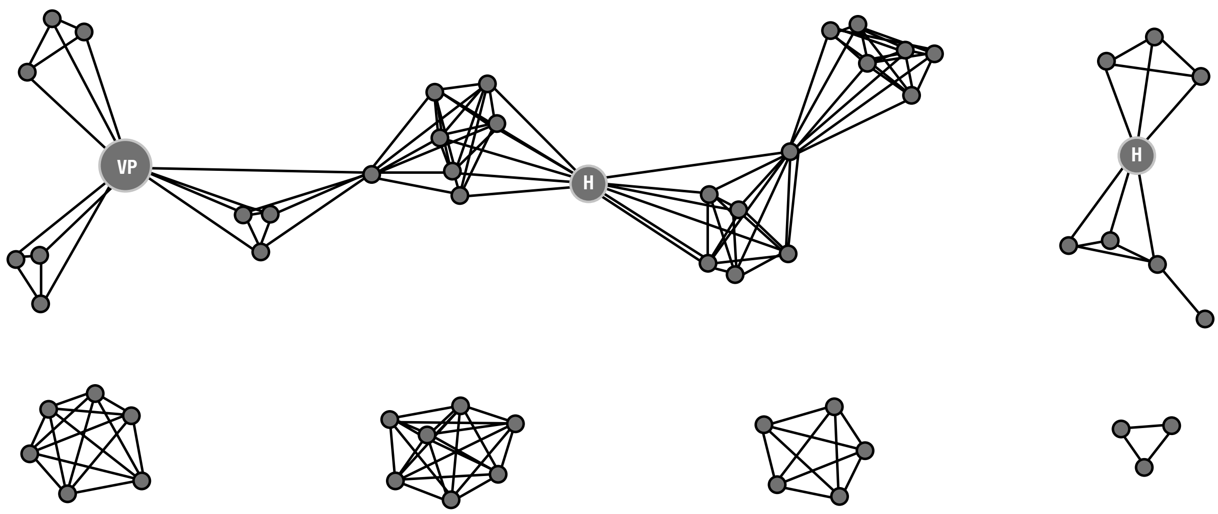

Fig. 3. Coauthorship graph for the Bashneft Company Research Center [11]

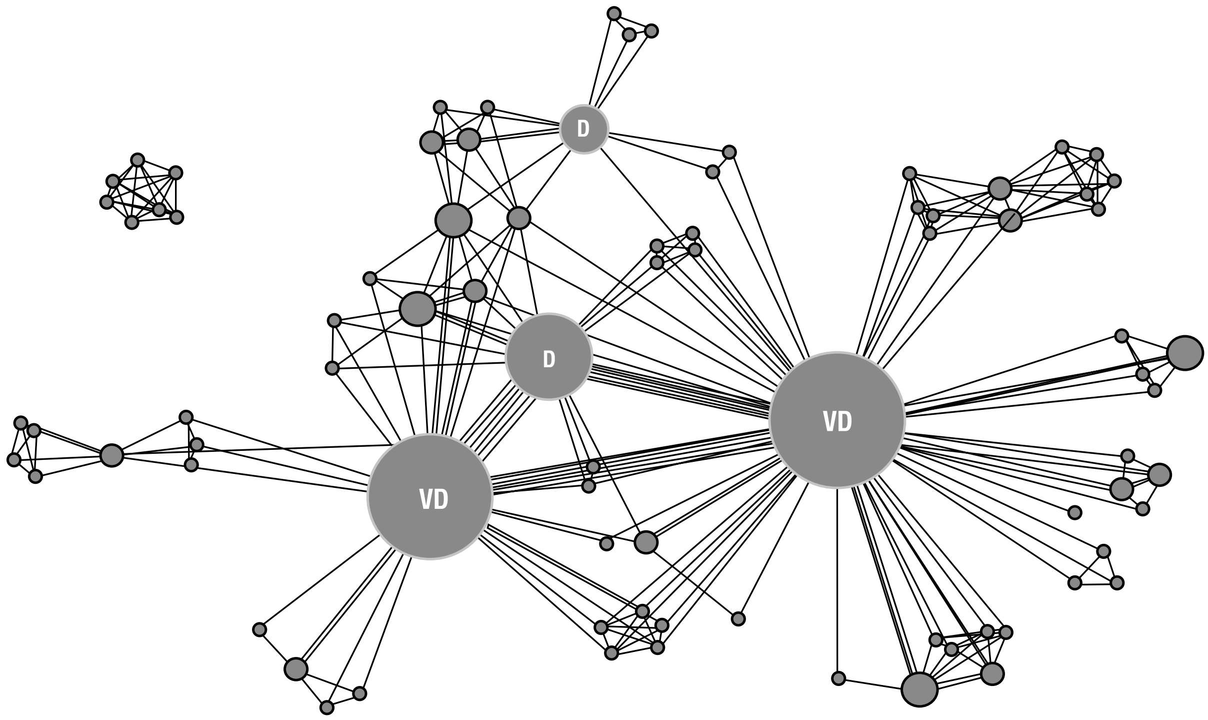

Fig. 4. Coauthorship graph for the Tatneft Company Research Center [11]

Conclusion

The simulation model constructed here can be used to conduct computer experiments for the scenario analysis of different situations relating to possible personnel changes at a research center. In particular, this model can help to analyze the impact of recruiting a new staff member with a given set of competencies, dismissing a staff member, or changing a staff member's competency profile. Another useful feature is the possibility of simulating different scenarios involving changes in the communication graph. Thus, this model can be used by research division managers and directors of personnel departments to answer the following questions:

✦ What will be the impact on a team's effectiveness of the dismissal of a staff member or his transfer to another department?

✦ What is the most effective way of involving a new member in teamwork? With which members should he predominantly interact?

✦ Which links should be inserted into the communication graph of staff members to improve the team's effectiveness?

✦ What is the best way of allocating personnel development resources? The development of which competencies shall have the biggest impact on the team's overall effectiveness?

Earlier studies [11] have shown that the coauthorship graph can vary greatly between sectoral research centers (*Figures 3* and *4*). This is often due to poorly organized communication.

The constructed model can serve as a basis for further research on the formation and management of project teams in a research environment. In particular, we plan to use it to develop a methodology for analyzing the effectiveness of academic research.

It may also be interesting to improve and expand the model in the following ways:

✦ Competence models may be improved with the help of fuzzy logic;

✦ When modeling long-term periods, it is necessary to keep track of the career and professional development of staff members and the corresponding changes in their competency profiles;

✦ Special attention should be paid to synergetic effects, which include, first and foremost, knowledge exchange in a group, the existence of supplementary skills that are not needed to solve the problem, and the emergence of mentorship. Synergetic effects should be taken into account when keeping track of changing team competency profiles;

✦ The functions α, β and γ that describe the process of making key decisions may be improved by taking other individual and group characteristics and the specifics of the task into account;

✦ The team formation algorithm can have a more complex iterative logic that takes agile project management methodologies into account;

✦ The case of unsuccessful project completion should be considered. With regard to research activity, this would mean that the prepared article was not accepted for publication, yet its results can be used for further work. In the present article, we make the assumption that each coauthorship leads to a publication. ■

The authors are aware of these aspects and plan to conduct research in these areas.

References

1. Scherbakov A.I. (1982) *Efficiency of scientific activity in the USSR: Methodological aspect.* Moscow: Economika (in Russian).
2. Ovchinnikov O.A. (2009) On the methodology of evaluation of scientific activity in scientific and educational institutions of the Russian Federation. *Bulletin of Moscow University of Russian Interior Ministry*, no 3, pp. 48–51 (in Russian).
3. Fursov K., Roschina Y., Balmush O. (2016) Determinants of research productivity: An individual-level lens. *Foresight and STI Governance*, vol. 10, no 2, pp. 44–56.

4. Shmatko N., Volkova G. (2017) Service or devotion? Motivation patterns of Russian researchers. *Foresight and STI Governance*, vol. 11, no 2, pp. 54–66.

5. Moe N.B., Dingsøyr T., Dybå T. (2008) Understanding self-organizing teams in agile software development. Proceedings of the *19th Australian Conference on Software Engineering (ASWEC 2008), Perth, Australia, 26–28 March 2008*, pp. 76–85.

6. Bavelas A. (1948) A mathematical model for group structures. *Human Organization*, vol. 7, no. 3, pp. 16–30.

7. Novikov D.A. (2008) *Mathematical models of formation and functioning of teams*. Moscow: Fizmatlit (in Russian).

8. Beilkhanov D.K., Kvyatkovskaya I.Yu. (2014) Using the competence model in the process of team building. *Technical Sciences – From Theory to Practice*, no 30, pp. 7–12 (in Russian).

9. Różewski P., Małachowski B. (2009) Competence management in knowledge-based organisation: Case study based on higher education organization. Proceedings of the *Third International Conference on Knowledge Science, Engineering and Management (KSEM 2009), Vienna, Austria, 2 5–27 November 2009*, pp. 358–369.

10. Guimera R., Uzzi B., Spiro J., Amaral L.A.N. (2005) Team assembly mechanisms determine collaboration network structure and team performance. *Science*, vol. 308, no 5722, pp. 697–702.

11. Krasnov F., Dokuka S., Yavorskiy R. (2016) The structure of organization: The coauthorship network case. Proceedings of the 5*th International Conference on Analysis of Images, Social Networks and Texts (AIST 2016). Yekaterinburg, Russia, 7–9 April 2016*, pp. 100–107.

12. Krasnov F., Dokuka S., Gorshkov I., Yavorskiy R. (2017) Analysis of strong and weak ties in oil & gas professional community. Proceedings of the *International Workshop on Formal Concept Analysis for Knowledge Discovery (FCA4KD 2017). Moscow, Russia, 1 June 2017*, pp. 22–33.

13. Borschev A. (2008) Application of simulation modeling in Russia – State for 2007. *Business Informatics*, no 4, pp. 64–68 (in Russian).

14. Borschev A. (2013) *The big book of simulation modeling: Multimethod modeling with AnyLogic 6.* AnyLogic North America.

15. Krasnov F.V., Ushmayev O.S. (2018) Exploration of hidden areas of research in the oil and gas industry by analyzing the OnePetro library. *International Journal of Open Information Technologies*, vol. 6, no 5, pp. 7–14 (in Russian).

16. Kurchakova N.V., Krasnov F.V. (2017) Modeling of intellectual capital changes in the conditions of increased load on the staff of the research organization. *Naukovedenie*, vol. 9, no 6. Available at: https://naukovedenie.ru/PDF/157TVN617.pdf (accessed 01 September 2018) (in Russian).

17. Krasnov F.V. (2008) A human and communications. *Director of Information Service*, no 11. Available at: https://www.osp.ru/cio/2008/11/5480255/ (accessed 01 September 2018) (in Russian).

18. Krasnov F., Sergeev A. (2008) Segmentation of IT customers on internal market. Proceedings of the *5th Software Engineering Conference (SEC(R) 2008). Moscow, Russia, 21–25 October 2008,* pp. 457–465.

19. Homans G.C. (2017) *The human group*. Routledge.

20. Katz E., Lazarsfeld P.F., Roper E. (2017) P*ersonal influence: The part played by people in the flow of mass communications*. Routledge.

21. Mann R.D. (1959) A review of the relationships between personality and performance in small groups. *Psychological Bulletin*, vol. 56, no 4, pp. 241–270.

22. Hare A.P. (1976) *Handbook of small group research*. N.Y.: Free Press.

About the authors

Tamara V. Voznesenskaya
Cand. Sci. (Phys.-Math.);
Associate Professor, Big Data and Information Retrieval School,
Faculty of Computer Science, National Research University Higher School of Economics,
20, Myasnitskaya Street, Moscow 101000, Russia;
E-mail: tvoznesenskaya@hse.ru

Fedor V. Krasnov
Cand. Sci. (Tech.);
Expert, Science & Technology Centre, Gazprom Neft PJSC, 75–79 liter D, Moika River Embankment,
Saint Petersburg 190000, Russia;
E-mail: krasnov.fv@gazprom-neft.ru

Rostislav E. Yavorsky
Cand. Sci. (Phys.-Math.);
Associate Professor, School of Data Analysis and Artificial Intelligence,
Faculty of Computer Science, National Research University Higher School of Economics,
20, Myasnitskaya Street, Moscow 101000, Russia;
E-mail: ryavorsky@hse.ru

Polina V. Chesnokova
System Analyst, LANIT Group, 14 block 1, Murmansky Proezd, Moscow 129075, Russia;
E-mail: chepolina27@gmail.com