ISSN 2587-814X (print), ISSN 2587-8158 (online)
Russian version: ISSN 1998–0663 (print), ISSN 2587–8166 (online)

№4(42)-2017

BUSINESS INFORMATICS

HSE SCIENTIFIC JOURNAL

CONTENTS

Publisher:
National Research University
Higher School of Economics

Subscription index in the «Rospechat» catalogue – 80870

The journal is published quarterly

The journal is included into the list of peer reviewed scientific editions established by the Supreme Certification Commission of the Ministry of Education and Science of the Russian Federation

Editor-in-Chief:
A. Golosov

Deputy Editor-in-Chief
Y. Koucheryavy

Computer Making-up:
O. Bogdanovich

Website Administration:
I. Khrustaleva

Address:
33, Kirpichnaya Street, Moscow,
105187, Russian Federation

Tel./fax: +7 (495) 771-32-38
http://bijournal.hse.ru
E-mail: bijournal@hse.ru

Circulation:
English version – 300 copies,
Russian version – 300 copies,
online versions in English and Russian – open access

Printed in HSE Printing House
3, Kochnovsky Proezd, Moscow,
Russian Federation

ABOUT THE JOURNAL

Business Informatics is a peer reviewed interdisciplinary academic journal published since 2007 by National Research University Higher School of Economics (HSE), Moscow, Russian Federation. The journal is administered by School of Business Informatics. The journal is published quarterly.

The mission of the journal is to develop business informatics as a new field within both information technologies and management. It provides dissemination of latest technical and methodological developments, promotes new competences and provides a framework for discussion in the field of application of modern IT solutions in business, management and economics.

The journal publishes papers in the areas of, but not limited to:

- ✦ data analysis and intelligence systems
- ✦ information systems and technologies in business
- ✦ mathematical methods and algorithms of business informatics
- ✦ software engineering
- ✦ Internet technologies
- ✦ business processes modeling and analysis
- ✦ standardization, certification, quality, innovations
- ✦ legal aspects of business informatics
- ✦ decision making and business intelligence
- ✦ modeling of social and economic systems
- ✦ information security.

The journal is included into the list of peer reviewed scientific editions established by the Supreme Certification Commission of the Ministry of Education and Science of the Russian Federation.

The journal is included into Russian Science Citation Index (RSCI) database on the Web of Science platform.

International Standard Serial Number (ISSN): 2587-814X (in English), 1998-0663 (in Russian).

Editor-in-Chief: Dr. Alexey Golosov – President of FORS Development Center, Moscow, Russian Federation.

EDITORIAL BOARD

ABOUT THE HIGHER SCHOOL OF ECONOMICS

Consistently ranked as one of Russia's top universities, the Higher School of Economics (HSE) is a leader in Russian education and one of the preeminent economics and social sciences universities in Eastern Europe and Eurasia. Having rapidly grown into a well-renowned research university over two decades, HSE sets itself apart with its international presence and cooperation.

Our faculty, researchers, and students represent over 50 countries, and are dedicated to maintaining the highest academic standards. Our newly adopted structural reforms support both HSE's drive to internationalize and the groundbreaking research of our faculty, researchers, and students.

Now a dynamic university with four campuses, HSE is a leader in combining Russian educational traditions with the best international teaching and research practices. HSE offers outstanding educational programs from secondary school to doctoral studies, with top departments and research centers in a number of international fields.

Since 2013, HSE has been a member of the 5-100 Russian Academic Excellence Project, a highly selective government program aimed at boosting the international competitiveness of Russian universities.

ABOUT THE SCHOOL OF BUSINESS INFORMATICS

The School of Business Informatics is one of the leading divisions of HSE's Faculty of Business and Management. The School offers students diverse courses taught by full-time HSE instructors and invited business practitioners. Students are also given the opportunity to carry out fundamental and applied projects at various academic centers and laboratories.

Within the undergraduate program, students participate each year in different case-competitions (PWC, E&Y, Deloitte, Cisco, Google, CIMA, Microsoft Imagine CUP, IBM Smarter Planet, GMC etc.) and some of them are usually as being best students by IBM, Microsoft, SAP, etc. Students also have an opportunity to participate in exchange programs with the University of Passau, the University of Munster, the University of business and economics in Vienna, the Seoul National University of Science and Technology, the Radbound University Nijmegen and various summer schools (Hong Kong, Israel etc.). Graduates successfully continue their studies in Russia and abroad, start their own businesses and are employed in high-skilled positions in IT companies.

There are four graduate programs provided by the School:

- ✦ Business Informatics (with three specializations – Informational Business Analytics, Information Systems Management and Business Process Modelling and Optimization);
- ✦ E-Business;
- ✦ Information Security Management;
- ✦ Big Data Systems.

The School's activities are aimed at achieving greater integration into the global education and research community. A member of the European Research Center for Information Systems (ERCIS), the School cooperates with leading universities and research institutions around the world through academic exchange programs and participation in international educational and research projects.

Open business model innovation: Literature review and agenda for future research

Tiare-Maria Brasseur
Lecturer, Department of e-Business
University of Vienna
Address: 1, Oskar-Morgenstern-Platz, Vienna, 1090, Austria
E-mail: tiare.brasseur@gmail.com

Andreas Mladenow
Lecturer, Department of e-Business
University of Vienna
Address: 1, Oskar-Morgenstern-Platz, Vienna, 1090, Austria
E-mail: andreas.mladenow@univie.ac.at

Christine Strauss
Professor, Department of e-Business
University of Vienna
Address: 1, Oskar-Morgenstern-Platz, Vienna, 1090, Austria
E-mail: christine.strauss@univie.ac.at

Abstract

In today's fast-paced business environment, firms are constantly pressured to innovate in order to remain competitive. Business model innovation (BMI) has recently attracted increasing attention as a promising approach to achieve competitive advantage in the face of fierce competition. Despite its great potential, however, BMI also entails high degrees of complexity, uncertainty and financial risk. Fueled by the rise of digital technologies, BMI has become increasingly open and collaborative in the recent past.

The aim of this paper is to investigate the role and implications of open and collaborative practices in BMI and to provide a comprehensive review of available literature in this field. Therefore, a systematic review of literature at the intersection of Open Innovation (OI) and BMI has been carried out. Our analysis of the literature identified two major research streams in open business model innovation (OBMI): OBMI trends (customer-driven BMI, BM co-creation, early BM validation, virtual collaboration, design thinking) and OBMI effects. Overall, the findings support a growing trend of collaboration and co-creation in BMI supported by digital or tangible tools, and further reveal that OI has a direct positive effect on BMI success. Analysis of the literature also shows that the field of OBMI is still an under-researched area.

Key words: business model innovation, open business model innovation, digital business modeling, open innovation, collaboration, collaborative infrastructure, co-creation, customer innovation.

Citation: Brasseur T.-M., Mladenow A., Strauss C. (2017) Open business model innovation: Literature review and agenda for future research. *Business Informatics*, no. 4 (42), pp. 7–16. DOI: 10.17323/1998-0663.2017.4.7.16.

Introduction

It was not so long ago that in real world scenarios "innovation" implied higher investments in internal research and development (R&D) than the competitors [1]. More recent and enhanced innovation strategies rely on collaborative approaches with external partners to complement their innovative capabilities. Chesbrough first introduced the notion of open innovation (OI) as a mode of innovation in which firms use external knowledge sources and external markets to accelerate internal innovation [1]. OI has received increasing attention from researchers [1–10] and the media in the recent past. The consensus view among OI scholars is that the value of reusing the same knowledge over and over again is limited and can harm a firm's ability to innovate [1, 2, 7]. Therefore, the OI model encourages firms to open up their processes to the outside world and allow a more participatory and decentralized approach to innovation in contrast to the previously dominant model of "closed innovation" [1]. The notion of OI overlaps with numerous other concepts such as open source, crowdsourcing, co-creation, user innovation and other forms of distributed innovation [11].

OI helps firms to access a wider array of talent and skills, to reduce cost of R&D, to gain additional revenue sources (through licensing, joint ventures, etc.) and to reduce the time-to-market for new products [12–14]. By combining internal and external sources of innovation, firms gain access to new innovations, new customers, new ways of commercialization and additional markets [12, 13]. In practice, firms such as Lego, 3M, P&G, IBM and Dell have demonstrated the power of OI to improve innovation and financial performance.

Besides OI, BMI has emerged as a promising approach to break out of fierce competition, particularly in face of the (higher) imitability of products and processes [15, 16]. It has been argued in the literature that business models (BMs) have become the new basis of competition [17] and probably the most challenging type of innovation. When trying to develop a new BM, firms are confronted with high levels of uncertainty, ambiguity and complexity [18]. A change in the BM can mean multi-billions in profits or it can mean corporate failure [19]. The high risk associated with BMI hence resides in the large scope of business transformation when BMI occurs. At the same time, however, BMI is increasingly recognized as a key driver of super normal profits and the main source of sustainable competitive advantage [1, 20]. In this context, a systematic literature review has been conducted in five leading scientific databases (ACM, EBSCO, IEEE, Springer Link, Wiley) to provide a comprehensive overview of available literature on BMI to identify the major topics and concepts discussed in this field.

1. OBMI trends and effects

Analysis of the literature identified two major new research streams in OBMI: OBMI trends and OBMI effects. In the available literature, various approaches and toolkits have been discussed, ranging from theoretical models, to (open) software solutions for virtual collabo-

ration and playful design thinking methods to advance BMI. Furthermore, five emerging trends in OBMI have been identified. Overall, the literature showed a trend of more collaborative forms of BMI and suggested the use of different approaches and tools to integrate outside partners in OBMI activities, as depicted in *Table 1*.

Envisaging the participation of supply chain members in OBMI activities, all studies involve customers in the BMI process. This is not surprising, as the keywords used for the systematic literature search aimed to identify articles exploring target group involvement, i.e. the participation of customers, in OBMI practices [2]. Moreover, 5 out of 11 publications choose to collaborate with both – customers and suppliers – to advance BMI. Additionally, a single study chose to involve customers, suppliers and a research institution in a collaborative setting of OBMI [21].

Envisaging the intensity of supply chain partner participation in OBMI activities, the International Association for Public Participation [22] has developed a stakeholder participation spectrum to demonstrate the possible types (i.e. levels) of stakeholder participation. Based on this spectrum, two participation approaches dominate the literature on OBMI: "involvement" and "collaboration". The majority of available publications, i.e. 7 out of 11 articles, choose a collaboration approach which allowed supply chain members to actively contribute new BM ideas. On the other hand, two publications pursued an involvement approach, which included adding to, modifying, or recombining knowledge contributed by supply chain partners.

2. BM co-creation

The topic of BM co-creation is most frequently discussed among available literature on OBMI [23–26]. Ebel et al. define BM co-creation as the activity of "developing business models together with customers in a collaborative manner" [24]. In other words, BM co-creation encourages the active participation of supply chain partners in the creation of new BM. Furthermore, co-innovation describes a new innovation paradigm, which incorporates collaboration and co-creation in a platform approach [26]. Accordingly, co-innovation emphasizes on the co-creation of unique value and experience with external stakeholders, rather than merely creating value for them.

The literature analysis we performed revealed that opening up the BMI process to allow BM co-creation enables new business opportunities [27] and enhances the quality of the developed BMs [24]. It has been argued that by seeking external feedback and collaborating with customers to share BMs, firms can test the commercial viability of a new BM before implementing it [23, 24]. Particularly, the role of customers and suppliers as valuable participants in the BMI process has been highlighted in the literature. Ogilvie [25] argues that engaging directly with customers and prospective partners helps to co-create solutions that deliver value to all stakeholders involved. Buur and Gudiksen [21] arrive at the conclusion that the "discussion of business models within the company, and with suppliers and customers is necessary to ensure competitive edge". Similarly, Ebel et al. [24] denote target group involvement as crucial for successful BMI. Looking at co-creation from a customer-perspective, Ogilvie claims that customers embraced the invitation to participate in BM co-creation sessions and were enthusiastic about the idea to co-create a new solution [25].

Furthermore, Chew put forward the view that "service innovation and business model innovation are the source for business growth" [23]. The author introduces the prototype of a new integrated design method, the so-called iSIM (integrated service innova-

Table 1.

Literature analysis

Source	OBMI Trends					OBMI Effects	Toolkits			Additional Information
	BM co–creation	Customer–driven BMI	Early BM validation	Virtual collaboration	Design thinking		Software solution	Tangible tools	Theoretical model	
Berre et al. (2013) [27]	X		X	X			X			• Cloud–based platform for OBMI • Pilot testing of business operations platforms • Social media based innovation community
Buur and Gudiksen (2012) [21]	X	X			X			X		• BM design experiments with "tangible" BMs • Interactive pinball design (gamification) • Customer–side of the BM
Chew (2015) [23]	X	X	X	X			X		X	• Integrated service innovation method (iSIM) model for simultaneous service innovation & BMI • Customer co–created value (user value) • "Backward" BM design (customer as start)
Denicolai et al. (2014) [33]						X				• Effects of external knowledge sourcing on firm growth (as proxy for BMI success) • Annual report analysis (310 European firms)
Ebel et al. (2016) [24]	X	X	X	X			X		X	• Pilot testing of business model development tool with test users • Based on online collaboration literature • Target group / expert / stakeholder integration
Gudiksen (2015) [32]	X				X			X		• Playful BM design experiments (gamification) • Role plays, use of tangible resources • Participatory innovation
Huang et al. (2013) [13]						X			X	• Effects of OI on BMI & organisational inertia • Effects of BMI on firm performance • Study among 141 small and medium enterprises (in Taiwan)
Ogilvie (2015) [25]	X		X		X			X	X	• BM co–creation (customers & industry partners) • Early BM & revenue model validation • 5Bs template, BM prototypes (mock–ups)
Pynnönen et al. (2012) [29]		X	X	X			X		X	• Managing customer–driven BMs • Customer participation in BMI, BM validation • Social media based innovation community
Trimi et al. (2012) [26]		X	X						X	• BM experimentation & BM flexibility • Customer development process (validation) • Lean philosophy (testing of BM hypotheses)
Zolnowski et al. (2014) [28]	X	X							X	• Customer as start of (new) BM design • Customer impact on other BM dimensions • BM canvas (firm / customer / partner perspective)

tion method), for simultaneous service innovation and BM design. The model is based on the assumption that customer value co-creation is central to BMI. Thus, the iSIM model facilitates the engagement of customers as co-producers of value. A co-creation approach in BMI enhances the mutual value proposition alignment between the customer-side and the supply-side business ecosystem [23, 28]. Similarly, Ebel et al. recommend the discussion and refinement of new BM alternatives with customers and suppliers until mutual agreement among all stakeholders is reached [24].

To practice co-creation, it has been proposed that firms need an organizational culture of "open leadership" and "organizational learning" [23]. With the use of rapid BM experimentation, firms can test the commercial viability of new BM concepts before committing capital on the design and implementation of the BM [23]. Ebel et al. point to a certain flexibility in terms of BM resources and capabilities "to ensure that the business models can be adapted to changing market situations" [24]. Zolnowski et al. stress the importance of adopting a "value network logic" instead of focusing on a single value chain, to foster the integration of and interaction with external partners [28]. Furthermore, Chew emphasized a combination of resource integration and configuration capabilities to be able to leverage externally generated knowledge for internal processes, i.e. so-called "absorptive capacity" [23].

3. Customer-driven BMI

Another aspect is the need for a more "customer-centered model" of BMI [26]. Pynnönen et al. argue that "firms do not necessarily know what the value preferences of their customers are" [29], so they should recognize customers as valuable participants in various roles in the BMI process. Zolnowski et al. take the customer as the "starting point" of BMI by positioning the customer at the top of the BM [28]. Visualising all potential interaction points with the customer helps to identify how the customer co-determines and influences other BM dimensions [28]. Chew proposes to first envision what the unique customer experience should be and then backward design the BM and service offering accordingly [23]. In the same way, Pynnönen et al. introduce a four-stage Business Mapping Framework to analyze the fit between the firm's current BM and customer value [29]. Each BM element was weighted according to customer value preferences to identify the core and non-core components of the BM and exclude non-value adding elements [29]. Trimi and Berbegal-Mirabent stress the importance of "customer validation" before creation [26]. They seize on the idea of customer development. The main principle of the customer development process is that a new concept or idea, first needs to be validated by customers to be able to proceed to the creation stage, or otherwise pivots back to the discovery stage. Regarding BMI, customer validation includes the verification of BM elements such as the perceived value of the offering, or the appropriateness of pricing or distribution channels [26]. Finally, Pynnönen et al. point to the importance of continuous involvement of customers in the BMI process to repeatedly align a firm's BM to current and emerging market needs in an iterative manner to obtain the essential real-time information about changing customer preferences [29].

4. Early BM validation

Envisaging the topic of early BM validation, researchers agree that "validation before creation", i.e. seeking external feedback and target group involvement in early stages of BM development, is vital to the success of BMI [24, 26]. Thus, through using an early BM validation approach, firms can test the com-

mercial viability of a new BM concept before making investments beyond planning [23, 25]. Early BM validation helps firms to reduce the high degrees of uncertainty and risk entailed in BMI. Ogilvie argues that the development of early BM validation techniques is the "key to success" [25]. Thus, through early and ongoing interactions with customers, firms can increase their chances of success [26]. BMs should hence be flexible enough to allow quick iterations and trial-and-error learning from early validation attempts. Particularly for early-stage businesses, the development of experimentation capabilities to rapidly test and validate their business hypotheses is crucial [26]. Ogilvie involves the use of visualization and design thinking techniques in BM co-creation sessions with customers and prospective partners. A telematics provider used visual BM prototypes (i.e. posters) "with just enough information to allow customers to understand how the BM might work", but also leaving enough space for comments and ideas by the workshop participants [25]. This innovative approach allowed the firm to validate and reconfigure a new BM concept together with a revenue model at minimal expense [25].

5. Virtual collaboration

Another major OBMI topic identified in the course of the review is "virtual collaboration". Available literature highlights the crucial role of IT tools to support open and collaborative approaches of designing new BMs [24]. In this regard, digital OBMI platforms aim at providing BM development teams with innovation community support for BMI challenges. The principal idea behind virtual collaboration for BMI is thus very similar to a crowdsourcing approach, which makes use of collective capacities to solve complex problems [30, 31]. The digital business modeling tools presented in the literature support a variety of different features such as voting, evaluations, rating, search and access control, community functions (profile pages, interest groups), messaging, commenting and other collaboration features such as file sharing [23, 24, 27, 29].

6. Design thinking

The fifth major topic covered by the literature is design thinking. Design thinking describes different participatory innovation methods involving experimentation, tangible resources and game-like innovation activities to stimulate the creative exploration of innovative solutions [21]. The approach of "tangible business modelling" allowed participants to "express themselves with their hands" [21]. The analyzed literature reported the use of tangible resources such as sticky notes [25] or balls and bricks [32] to facilitate collaboration among (multi-disciplinary) BM co-creation teams. Buur and Gudiksen conducted iterative BM experiments in various interactive settings to trigger new discussions and solution-based thinking. The researchers arrived at the conclusion that "playfully 'thinking with hands and body' enhances 'the quality of conversations' towards an innovative outcome" [21]. Similarly, Gudiksen chose different BM design games for the playful experimentation with new BM ideas [32]. It has been demonstrated that approaching the BM as a design problem advances new perspectives on BMI initiatives. The use of tangible resources and a randomizer (i.e. a dice) helped to spark new ideas and combinations. Eventually, BM design games have two major benefits: (1) they impose clear rules that had to be followed by all participants, and (2) they provide a joyful and engaging atmosphere, which allowed participants to "step out of the real life" ("foolishness") and enjoy the "freedom to play" and stimulated experiential learning [32]. To successfully practice design thinking for BMI, firms need to create a tempo-

rary space for imagination and encourage an experimental, game-oriented culture and the freedom to improvise, play and test BM concepts [21, 32].

7. OBMI effects

Two publications in available literature on OBMI discussed how OI can be effective in creating BMI [13, 33]. Both studies used a quantitative approach to explore the implications of external knowledge sourcing for BMI and firm performance [13] or firm growth "as a proxy for BMI success" [33], respectively. By using survey research methodologies Huang et al. found that OI can be effective in changing "organisational inertia" and hence create BMI [13]. The notion of organizational inertia describes an organization's difficulty in changing its organizational structure. In addition, organizational inertia has a negative influence on both OI and BMI [13]. Therefore, it is vital to overcome organizational inertia to be able to adapt to fast changing market needs [13]. Furthermore, the researchers have claimed that OI has a direct positive influence on both BMI and firm performance.

By using accounting indicators as a proxy for knowledge acquisitions, Denicolai et al. aim at understanding how internal and external knowledge assets can be combined to contribute to a firm's growth as a proxy for BMI success [33]. For this purpose, scholars investigated the interplay between internally developed knowledge (KINT) and externally acquired knowledge assets (EINT) to create value. Their research identifies an inverted U-shaped tension between the costs and benefits of external knowledge sourcing and internal absorptive capacity [33]. Thus, it has been argued that increasing investments in externally generated knowledge assets has a positive effect on a firm's growth, but only up to a certain point – beyond this threshold it has a negative effect on a firm's growth as the organization encounters difficulties in absorbing such high levels of external knowledge [33]. The researchers arrive at the conclusion that there is an "optimal level" of external knowledge, depending on the absorptive capacity of the firm, at which firm growth, and thus BMI success, is maximized.

Conclusion

The analysis of literature we presented shows that the field of OBMI is still an underresearched area, as only 11 publications met the inclusion criteria. It was found that available literature discusses various trends as well as the effects of OBMI practices. Analysis of the available literature reveals that OI has a direct positive effect on BMI success. It was found that involving outside partners, particularly customers, in BMI processes may enhance both the quality of developed BMs as well as a firm's performance – at least up to a certain level. Regarding the "openness" of BMI activities, all 11 publications included involve customers in the process of BMI, while some chose to additionally cooperate with suppliers and research institutions to advance BMI. Furthermore, the majority of publications chose a collaborative BMI setting, which allowed outside partners to actively contribute new BM ideas. Furthermore, five major trends in OBMI have been identified.

First, the trend of BM co-creation has been most frequently discussed, indicating the growing importance of collaboration in BMI. Second, the literature analysis revealed the need for a more customer-centric model of BMI. Third, the findings of this review suggest that establishing capabilities to validate new BM concepts in early stages of development is crucial for successful BMI. Therefore, it has been proposed that firms should seek early target group involvement to be able to achieve BM "validation before creation". Fourth, available literature has presented different methods of "virtual col-

laboration" for BMI. The digital collaboration tools discussed in the literature draw on crowdsourcing methods for BMI challenges and highlight the crucial role of ICT in facilitating OBMI. Fifth, the application of design thinking in BMI represents another key finding of this review. This literature review reveals that approaching the BM as a design problem through the use of tangible and visual BM "mock-ups" in experimental, game-like innovation activities enhances the quality of discussions among participants and thus increases innovation output. It has been argued that using design thinking methods in OBMI helps firms to advance new perspectives on BM challenges and allows them to both validate and collect new BM ideas at minimal expense.

It can be concluded that OBMI represents an important research area, which is still in its infancy. This paper delivers a comprehensive review of knowledge in this emerging research field. Given the insights on the growing trend of collaboration in BMI, further research will focus on exploring increasingly collaborative approaches in BMI. Additionally, future work might shed light on BM co-creation activities regarding e.g. their application in different stages of the innovation process and the direction of information flows. Furthermore, as virtual collaboration tools and design thinking methods are both expected to play an increasing role in facilitating BMI in the future, they represent promising areas for further research. ■

References

1. Chesbrough H. (2010) Business model innovation: Opportunities and barriers. *Long Range Planning*, vol. 43, no. 2–3, pp. 354–363.
2. Brasseur T.-M., Mladenow A., Strauss C. (2017) Business model innovation to support smart manufacturing. Proceedings of *American Conference on Information Systems 2017. Workshop on Smart Manufacturing. Boston, USA, 10–12 August 2017.* Available at: http://aisel.aisnet.org/sigbd2017/9 (accessed 16 October 2017).
3. Enkel E., Gassmann O., Chesbrough H. (2009) Open R&D and open innovation: Exploring the phenomenon. *R&D Management*, vol. 39, no. 4, pp. 311–316.
4. Gassmann O., Enkel E., Chesbrough H. (2010) The future of open innovation. *R&D Management*, vol. 40, no. 3, pp. 213–221.
5. Huizingh E.K. (2011) Open innovation: State of the art and future perspectives. *Technovation*, vol. 31, no. 1, pp. 2–9.
6. Mladenow A., Fröschl K.A. (2011) *Kooperative forschung.* Frankfurt am Main: Lang.
7. Laursen K., Salter A. (2006) Open for innovation: the role of openness in explaining innovation performance among U.K. manufacturing firms. *Strategic Management Journal*, vol. 27, no. 2, pp. 131–150.
8. Lopez-Vega H., Tell F., Vanhaverbeke W. (2016) Where and how to search? Search paths in open innovation. *Research Policy*, vol. 45, no. 1, pp. 125–136.
9. Prugl R., Schreier M. (2006) Learning from leading-edge customers at The Sims: Opening up the innovation process using toolkits. *R&D Management*, vol. 36, no. 3, pp. 237–250.
10. West J., Bogers M. (2014) Leveraging external sources of innovation: a review of research on open innovation. *Journal of Product Innovation Management*, vol. 31, no. 4, pp. 814–831.
11. Mokter H. (2013) Open innovation: so far and a way forward. *World Journal of Science, Technology and Sustainable Development*, vol. 10, no. 1, pp. 30–41.

12. Alexy O., George G. (2013) Category divergence, straddling, and currency: Open innovation and the legitimation of illegitimate categories. *Journal of Management Studies*, vol. 50, no. 2, pp. 173–203.

13. Huang H.-C., Lai M.-C., Lin L.-H., Chen C.-T. (2013) Overcoming organizational inertia to strengthen business model innovation: An open innovation perspective. *Journal of Organizational Change Management*, vol. 26, no. 6, pp. 977–1002.

14. Kleemann F., Voss G.G., Rieder K. (2008) Un(der)paid innovators: The commercial utilization of consumer work through crowdsourcing. *Science, Technology & Innovation Studies*, vol. 4, no. 1, pp. 5–26.

15. Becker A., Mladenow A., Kryvinska N., Strauss C. (2012) Aggregated survey of sustainable business models for agile mobile service delivery platforms. *Journal of Service Science Research*, vol. 4, no. 1, pp. 97–121.

16. Kopetzky R., Günther M., Kryvinska N., Mladenow A., Strauss C., Stummer C. (2013) Strategic management of disruptive technologies: a practical framework in the context of voice services and of computing towards the cloud. *International Journal of Grid and Utility Computing*, vol. 4, no. 1, pp. 47–59.

17. Bursuk I., Mladenow A., Novak N.M., Strauss C. (2016) Online cofounder search in tech start-ups. Proceedings of the *18th International Conference on Information Integration and Web-based Applications & Services (iiWAS 2016). Singapore, 28–30 November 2016*, pp. 482–488.

18. Andries P., Debackere K. (2013) Business model innovation: Propositions on the appropriateness of different learning approaches. *Creativity and Innovation Management*, vol. 22, no. 4, pp. 337–358.

19. Gibson E., Jetter A. (2014) Towards a dynamic process for business model innovation: A review of the state-of-the-art. Proceedings of the *Infrastructure and Service Integration Conference (PICMET 2014). Kanazawa, Japan, 27–31 July 2014*, pp. 1230–1238.

20. Teece D.J. (2010) Business models, business strategy and innovation. *Long Range Planning*, vol. 43, no. 2–3, pp. 172–194.

21. Burr J., Guriksen S. (2012) Interactive pinball business. Proceedings of the *7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design (NordiCHI 2012). Copenhagen, Denmark, 14–17 October 2012*, pp. 129–138.

22. IAP2 (2014) *IAP2's Public Participation Spectrum*. Available at: https://www.iap2.org.au/Tenant/C0000004/00000001/files/IAP2_Public_Participation_Spectrum.pdf (accessed 01 March 2017).

23. Chew E.K. (2016) iSIM: An integrated design method for commercializing service innovation. *Information Systems Frontiers*, vol. 18, no. 3, pp. 457–478.

24. Ebel P., Bretschneider U., Leimeister J.M. (2016) Leveraging virtual business model innovation. *Information Systems Journal*, vol. 26, no. 5, pp. 519–550.

25. Ogilvie T. (2015) How to thrive in the era of collaborative services entrepreneurship. *Research Technology Management*, vol. 58, no. 5, pp. 24–33.

26. Trimi S., Berbegal-Mirabent J. (2012) Business model innovation in entrepreneurship. *International Entrepreneurship and Management Journal*, vol. 8, no. 4, pp. 449–465.

27. Berre A.J., Man H.D., Lew Y., Elvesaeter B., Ursin-Holm B.M. (2013) Open business model, process and service innovation with VDML and ServiceML. Proceedings of *5th International IFIP Working Conference on Enterprise Interoperability (IWEI 2013). Enschede, The Netherlands, 27–28 March 2013*, pp. 127–142.

28. Zolnowski A., Weiss C., Bohmann T. (2014) Representing service business models with the service business model canvas: The case of a mobile payment service in the retail industry. Proceedings of the *47th Hawaii International Conference on System Science (HICSS 2014). Hawaii Big Island, 6–9 January 2014*, pp. 718–727.

29. Pynnönen M., Hallikas J., Ritala P. (2012) Managing customer-driven business model innovation. *International Journal of Innovation Management*, vol. 16, no. 4, pp. 1–13.

30. Mladenow A., Bauer C., Strauss C. (2016) Crowd logistics: the contribution of social crowds in logistics activities. *International Journal of Web Information Systems*, vol. 12, no. 3, pp. 379–396.

31. Ernst C., Mladenow A., Strauss C. (2017) Collaboration and crowdsourcing in emergency management. *International Journal of Pervasive Computing and Communications*, vol. 13, no. 2, pp. 176–193.

32. Gudiksen S. (2015) Business model design games: Rules and procedures to challenge assumptions and elicit surprises. *Creativity and Innovation Management*, vol. 24, no. 2, pp. 307–322.

33. Denicolai S., Ramirez M., Tidd J. (2014) Creating and capturing value from external knowledge: the moderating role of knowledge intensity. *R&D Management*, vol. 44, no. 3, pp. 248–264.

Digital economy: Conceptual architecture of a digital economic sector ecosystem

Yury M. Akatkin
Head of Laboratory of Social-Demographic Statistics
Plekhanov Russian University of Economics
Address: 36, Stremyanny Lane, Moscow, 117997, Russian Federation
E-mail: u.akatkin@semanticpro.org

Oleg E. Karpov
Corresponding Member, Russian Academy of Sciences
General Director of Pirogov National Medical-Surgical Center
Address: 70, Nizhnyaya Pervomaiskaya Street, Moscow, 105203, Russian Federation
E-mail: nmhc@mail.ru

Valery A. Konyavskiy
Head of Information Security Department
Moscow Institute of Physics and Technology
Address: 1A, Kerchenskaya Street, Moscow, 117303, Russian Federation
E-mail: konyavskiy@gospochta.ru

Elena D. Yasinovskaya
Senior Researcher, Laboratory of Social-Demographic Statistics
Plekhanov Russian University of Economics
Address: 36, Stremyanny Lane, Moscow, 117997, Russian Federation
E-mail: elena@semanticpro.org

Abstract

The main objective of digital transformation is to fulfill the needs of a "new digital generation customer" for on-demand delivery, quality and personalization. "Anything as a service" has become the key principle of the digital paradigm. This is about a data-oriented service which relies on sharing information resources (including public ones) and the requirements for interoperability, security and trust. This paper presents the main approaches to digital transformation based on the example of the most innovatively active sectors such as banking and healthcare. We compare the proprietary development of digital services (products) to the building of a digital sector ecosystem aimed at attracting an unlimited number of participants. We defined the purpose of creating an ecosystem that is to provide the population with digital services formed on demand, in real time, in compliance with legislation and regulations, as well as in the context of maximum trust. We emphasize the role of openness for uniting the efforts of the community interested in the development of a digital industry, extension of public-private partnerships and building a competitive environment in order to ensure the rapid growth of available digital services, as well as to improve their quality. Since the knowledge

economy is the basis for the digital economy, the authors consider it especially important to form a semantic core which acts as the carrier of knowledge in a digital sector ecosystem. We confirmed the necessity to implement the semantic core by a brief analysis of modern semantic approaches to standardization of information sharing in the above-mentioned industries, such as FIBO, BIAN (banking), HL7 and UMLS (health). The research carried out allowed the authors to design the conceptual architecture of the ecosystem and to suggest several proposals for digital transformation of an industry. The proposals express the necessity of state support for innovation and providing the conditions for the entry of new digital products based on the following principles: accessibility, timeliness, personalization, adaptability and security.

Key words: digitalization, digital economy, digital service, economic sector, ecosystem, conceptual architecture, interoperability, information sharing, semantic integration, semantic core.

Citation: Akatkin Y.M., Karpov O.E., Konyavskiy V.A., Yasinovskaya E.D. (2017) Digital economy: Conceptual architecture of a digital economic sector ecosystem. *Business Informatics*, no. 4 (42), pp. 17–28. DOI: 10.17323/1998-0663.2017.4.17.28.

Introduction

Millennials, the generation born between 1980 and 2000, have grown up and are now involved in the labor market, which substantially changes the average customer's needs. They expect high quality services (products) on demand as well as to use mobile and other innovative technologies to which they are accustomed. When the experts forecast The Fourth Industrial Revolution [1], the current situation in various economic sectors calls for digital transformation as well as their integration into the digital world and digital economy.

Although experts have generally reached a consensus about the goals of digital transformation, the term "digital economy" stays rather ambiguous. It retains a focus on marketing following the market trends set for the benefit of the main players. During the last few years, various reviews and concepts made by global analysts and ICT leaders on digitization of economy and its particular sectors have filled the information landscape. Based on research materials, the Analytical Center for the Government of the Russian Federation has compiled seven definitions of "digital economy" and invited experts to discuss them [2]. However, we consider it more useful to move forward from the exact terminology to making concrete steps towards the development of digital economic sectors and further improvement of public administration.

The public sector should regard digital transformation within the overall complex of measures which is required for successful involvement in the global digital economy. The suggestions on e-government digitalization which the authors proposed in the paper "Towards the digital government in Russia: Integrative approach" [3] remain relevant and should be considered within this complex of measures. We discuss the following issues: providing interoperability (at all levels including semantic); building open platforms which enable collaboration between domain and IT experts; establishing an expert community involving key scientists.

In this paper, we suggest a concept of digital sector ecosystem architecture in order to define the main ways and methods of digital transformation. We have considered potential participants and basic ecosystem elements, realization features (based on banking and healthcare sectors) and the role of the state in the future growth of digital economy. This approach has

allowed us to work out proposals aimed at providing innovations with state support based on public-private partnership principles and to create optimal conditions for the development of customer-oriented digital services and products.

1. Current state of digital transformation in some economic sectors

Using up-to-date technologies, digitalization reshapes the competitive picture and blurs the boundaries between the players of economic sectors. Digital leaders, such as Google, Apple, Facebook and Amazon, promote the digital paradigm and people define its significance in accordance with the expectation level of "new digital generation clients", who get used to permanent service accessibility as well as to immediate service delivery. Therefore, large IT corporations, equipment and software providers primarily focus their efforts on changing the form in which they offer products and services to the customer.

Since clients quickly get accustomed to a high level of digital services, the market offers solutions which meet their needs in a better way, by increasing the availability, providing high usability and convenience together with personalization of services according to the customers' individual characteristics and interests.

Thus, the experts characterize the digital transformation by the following features:

✦ orientation to customers (a digital service or a digital product is formed in accordance with customer requests and doesn't depend on the producer's offer);

✦ orientation towards mobility and velocity (the principle "here and now": using any device, any communication channel);

✦ orientation on data (the acquisition of new data from existing data, analysis and decision-making).

This leads to the transformation of a business model towards the individualization of services and products. The principle "anything as a service" appears, in which the service is data-oriented. For instance, many people consider the use of "car-sharing" services on par with the purchase of a car, not to mention Uber-like services, which have become a popular success story of the digital economy.

To meet these high customer expectations, companies must accelerate the digitization of their business processes. They should go beyond simply automating existing processes: "They must reinvent the entire business process, including cutting the number of steps required, reducing the number of documents, developing automated decision making, and dealing with regulatory and fraud issues. Operating models, skills, organizational structures, and roles need to be redesigned to match the reinvented processes. Data models should be adjusted and rebuilt to enable better decision-making, performance tracking, and customer insights. Digitization often requires that old wisdom be combined with new skills, for example, by training a merchandising manager to program a pricing algorithm. New roles, such as data scientist and user-experience designer, may be needed" [4].

The banking industry is a significant example of rapid digital development. Deloitte conducted research on the use of new technologies in Russian banks, from which experts chose 11 innovations and divided them into several groups: security ("smart" authentication), analytics (Big Data, personal financial assistants), digital technologies (online wallet, contactless payment, etc.), automation (robots in departments), gamification (games and quests for customers) and P2P-crediting [5].

We should note that leaders in the banking sector pay particular attention to IT development and this causes the establishment of digital factories.

Thus, one of the largest banks in Canada, The Bank of Nova Scotia or Scotiabank recently announced the development of a new Digi-

tal Factory that will pursue, design and deliver digital innovations and solutions for the Bank's customers with the philosophy of "reinventing how banking serves people"[1].

Deutsche Bank now has its own center for developing digital banking products: The Digital Factory in Frankfurt. Roughly 400 software developers, IT specialists and financial experts from 14 nations are working together to develop digital products using state-of-the-art methods [6].

In November 2016, Deutsche Bank (DB) gave software developers from Germany and abroad the opportunity to create digital solutions for bank clients that go far beyond traditional financial services. This approach ensured the creation of an innovation ecosystem which united three innovation labs where the bank cooperates with start-ups from all over the world, the Digital Factory and DB's new research and development center. The bank provides access to its proprietary development environment via an application programming interface (dbAPI) which allows programmers to test their ideas for digital services of the future [7].

The PSD2 (EU Revised Payment Service Directive)[2] also supports the logic of open APIs for the provision of payment services. According to this quite disruptive initiative, European banks are obliged to provide third-party providers with access to their customers' accounts through open APIs. This will enable third parties (e.g. application developers) to build financial services on top of the banks' data and infrastructure. On behalf of the client and without the need to make an agreement with the bank, a third party will be able to make payments and display information about transactions in its own applications tailored to the customers' expectations. This is another step towards open banking, which, on the one hand, brings anxiety into the banking market, and on the other – it seems to be a new, strategically important,prospect [3].

On the other hand, the proprietary approach to the development of digital banking services is quite traditionally applied. For example, in 2011 Russia's Sberbank created the IT company "Sbertech". Currently, about 7000 employees in 16 cities are developing 350 projects for this bank. The largest projects are the "Integrated Front-End System", which is a client service in any place through any channel; the "Platform for Business Development support", a tool for business applications creation and the "Data Factory" for monetization of information about clients and analysis of their behavior [8].

Today Sbertech focuses its projects on the modernization of their own in-house developed banking system. However, the development of open interfaces (Open API) is also said to be one of the future directions. For this reason, Sbertech management points towards the creation of marketplaces or aggregators where "together with their partners they will begin to sell not only financial but also related services such as travel and insurance. This is especially true since the banking ecosystem is growing and there are many developers who want to write their own applications related to banking services" [9]. This shows the desire to follow the principles of openness and to attract a large number of independent IT professionals focused on customer-oriented start-up ideas. This is the approach prescribed by the digital paradigm.

The innovation market in healthcare is one of the most active. Digital transformation in this domain is aimed at identification and prevention of most (80/20) life and health hazards due to timely preliminary diagnostics and health monitoring, primary medical consul-

[1] http://digitalfactory.scotiabank.com/ (accessed 19 September 2017)
[2] https://ec.europa.eu/info/law/payment-services-psd-2-directive-eu-2015-2366_en (accessed 19 September 2017)
[3] https://nordeaopenbanking.com/ (accessed 19 September 2017)

tations and personalized on demand medical services, emergency medical care services initiation and referral to in-depth medical examination in high-tech health care centers [10].

The "Accenture Digital Health Technology Vision 2016" [11] emphasizes that keeping up with changing technology is vital, but it is just as important to evolve the consumer experience, care delivery methods and career development opportunities for the healthcare workforce. It reveals five trends that prove winning in the digital age that depends on people:

1. Intelligent Automation: Do things differently, do different things to new jobs, products and services in healthcare;

2. The Liquid Workforce: Today's digital demands call for highly agile healthcare skillsets;

3. Platform Economy: Ecosystems are the new bedrock of digital healthcare;

4. Predictable Disruption: Digital ecosystems blur healthcare's boundaries in a foreseeable way;

5. Digital Trust: As healthcare data security risks increase, so do opportunities to earn consumer trust [11].

The United Kingdom, which is a digital economy leader, is actively creating an infrastructure to provide innovation solution development by involving a wide range of developers interested in digital field progress.

One example of such projects is the "Digital Catapult"[4], a platform for technical specialists, creative professionals, business and academic representatives who develop new ideas within corporations and promote their products to the British market and abroad. This allows British "digital" companies to implement innovations faster and with less risk. In this way, the entry of new products and services to the market can be accelerated.

Code4Health[5] is another initiative. It is a program of NHS England and NHS Digital to support health and care professionals with digital tools and technologies to provide safe, effective, and high quality telecare services.

Code4Health intends to educate and inform all community members (citizens, patients, guardians, health care employees and digital technology specialists) about tele-healthcare, the option to use digital technologies and related tools, knowledge and skills for collaboration in design and implementation of high-quality digital solutions [12].

The project comprises the following parts:

✧ **Platform**. Code4Health provides a simulated environment in which you can explore the resources you will find in the emerging open digital health and care ecosystem, learn to code, build apps, discover and create content and test your ideas;

✧ **Learning.** Code4Health provides clinicians with an explanation of the technologies they use and with opportunities for coeducation to help in designing new advanced services for rendering medical care;

✧ **Communities.** Code4Health brings together people with shared interests in digital healthcare. Communities can involve people with a shared interest in some specific aspect of digital healthcare or involve people chosen geographically.

Hence, based on the above information, in order to facilitate digital transformation, the efforts in three interrelated areas will be fundamental:

1. The reengineering of business processes, technologies and business cultures to provide customers with a service in any place and at any time in the digital format they are accustomed to;

[4] https://digital.catapult.org.uk/ (accessed 19 September 2017)
[5] https://code4health.org/ (accessed 19 September 2017)

2. Implementation of innovative technologies either in the industry or in the economy in general, to be present in any digital platform using mobile and Internet of Things (IoT) devices;

3. Creation of an organizational form and normative base for attracting innovation and effective cooperation with developers of digital solutions, applications and devices, in the framework of sectoral state regulation, making it possible to reduce administrative barriers to digital transformation.

Therefore, it is also necessary to solve emerging legal problems in a timely manner and change the existing legislation, both in the industry and in the country's digital economy, while maintaining a reasonable degree of state regulation and providing a sufficient level of state support.

The above overview shows that today's digital leaders use different approaches towards the realization of a system for development and implementation of digital services:

1. To create a proprietary system for the development and the implementation of digital services, as well as to establish a digital services development company and to lead independent activity (in the example of Sbertech). Along with obvious benefits, the following disadvantages stand out in this approach:

- ✦ it is the developer (founder) of a proprietary system who defines what clients want; therefore it is difficult to carry out the requirements of the digital economy to focus on the customer's needs;
- ✦ despite increased competition in the IT community, integration with the industry community is complicated, and it is therefore difficult to achieve the proper quality and timeliness of digital services development;
- ✦ substantial investments in the creation of specific products and services which may become outdated before the end of the development process.

2. To create a digital sector ecosystem by consolidating the efforts of digital technology and the application developer community, industry enterprises (for example, in healthcare - medical institutions providing digital healthcare as well as other stakeholders) In our opinion this approach is preferable and is outlined in more detail below.

Current practice shows that digital transformation requires openness and involvement of a wide range of developers who are ready to create innovative digital services. Therefore, the second approach is the most promising. Along with the relevance of this approach for digital transformation, it is important to consider it in more detail.

2. Digital sector ecosystem

Experts and consultants [13, 14] have long said that the digital economy is becoming one of the key contributors to GDP growth. It also has important implications for GDP measurement, productivity and household welfare in the retail sector and in all service sectors of the economy [15].

This year the challenges of the digital economy became one of the main themes at the St. Petersburg Economic Forum. It was emphasized that Russia needs to build up human, intellectual and technological advantages in the digital economy sector, and intends to act in directions which have systemic importance. This necessitates the formation of a principally new, flexible normative base to introduce technology in all areas of life[6]. At the same time, all decisions should consider the information security of the state, businesses and citizens[7].

The following stakeholders should support digital transformation:

- the state – from the standpoint of sectoral regulation, setting norms and rules, providing support measures, as well as encouraging

[6] https://sputniknews.com/business/201706021054244434-russia-economy-putin- spief-2017/ (accessed 19 September 2017)
[7] https://ria.ru/economy/20170602/1495725962.html (accessed 19 September 2017)

industry leaders to migrate to a digital paradigm;

• leading enterprises of economic sectors – within the framework of consolidating the efforts of expert communities, sharing information resources and providing access to them for the development of digital services, standardization and ensuring a sufficient level of trust and security.

• innovative companies – as the main driver for the emergence of new customer-oriented digital services and products.

In conclusion, the principles of openness and the union of the community's efforts are fundamentally important for digitalization, and the formation of digital economic sector ecosystems becomes the basic solution for the digital economy as a whole.

A digital sector ecosystem is an environment that serves innovative development and the expansion of digital services and products, applications and devices in a particular sector of the digital economy.

The purpose of creating an ecosystem is to provide the population with the digital services formed on demand, in real time, in compliance with legislation and regulations, as well as in the context of maximum trust. Such applications will allow consumers' access to services and products without thinking how the industry generally works and how the information systems that support the automation of the sector function.

This ecosystem creates the basis for a public-private partnership in digitization of various economic sectors and provides an opportunity for many third-party developers: the producers of new devices, products, tools and digital services to join the process. Openness for participation, government support and a competitive environment will create the proper conditions for the number of available digital services to grow and for their quality to improve. At the same time, the state will be able to avoid expenses for mass development and the implementation of application services, creating an environment that encourages small and medium business participation, thus supporting Russian providers of services, gadgets and other devices.

While considering different ways to realize this ecosystem, we should take into account that a technology that can solve all industry problems simultaneously does not exist. The challenges of sectoral digital transformation are too monumental and complicated for one company. Therefore, we should not think about separate systems and services but discuss the possibility of using a platform that provides the opportunity for the collaboration of separate systems and organizations from both technological and commercial standpoints.

The economy of knowledge is the basis for all digital economic sectors (sub-sectors). The knowledge keeper in this economy should be a semantic core, the support and development of which is a vitally important and scientifically intensive activity for the digital economy. For example, in the banking sector, where semantic interoperability is important [16], Financial Industry Business Ontology (FIBO) is being implemented[8]. A joint effort by the Object Management Group (OMG) and the Enterprise Data Management (EDM) Council, FIBO is an industry initiative to define financial industry terms, definitions and synonyms using semantic web principles such as RDF/OWL and widely adopted OMG modeling standards such as UML, started in the end of 2015. FIBO is designed and developed according to regulatory requirements, e.g. the Basel Committee on Banking Supervision (BCBS) 239 compliance for effective risk data aggregation and risk reporting is a major driver for FIBO [17].

[8] https://www.edmcouncil.org/financialbusiness (accessed 19 September 2017)

To reduce the costs of integration, increase the flexibility of IT solutions and to optimize banking IT architecture, the Banking Industry Architecture Network (BIAN) was created and is being developed by an international association of banks and suppliers of IT solutions[9]. BIAN is a collaborative not-for-profit ecosystem formed from global leading banks, technology providers, consultants and academics. It can become a good example of a second approach where community members combine their industry expertise to create a revolutionary banking technology framework that standardizes and simplifies core banking architecture, which has typically been convoluted and outdated. Based on service-oriented architecture principles, the comprehensive model provides a future-proof solution for banks that fosters industry collaboration. BIAN plans to use FIBO, considering it to be a crucial, complementary initiative"[10].

The current version (BIAN v.5) includes 7 business lines, 36 business domains, about 300 services in different domains, more than 700 business scenarios and about 2000 standard business operations in these services. Among the participants, there are 27 financial institutions (ABN AMRO Group, Credit Suisse, Societe Generale Group, Deutsche Bank, Unicredit Group, ING, Achmea, Rabobank, UBS, Banko Galicia and others) and 43 software producers (Temenos, Diasoft, Infosys, Sopra Banking Software, TCS Banks, IBM, SAP, Microsoft and others).

In healthcare, interoperability and standardization have been developing since the 1970s within the strategic healthcare initiative started in the USA. These initiatives have led to such significant projects as the Health Level Seven International (HL7) and the Unified Medical Language System (UMLS). HL7 was founded in 1987 to produce a standard for the exchange of data with hospital information systems.

Presently HL7 is not only a set of standards, but it is a not-for-profit, ANSI-accredited standards developing organization dedicated to providing a comprehensive framework and related standards for the exchange, integration, sharing, and retrieval of electronic health information that supports clinical practice and the management, delivery and evaluation of health services[11].

The Unified Medical Language System (UMLS) integrates and distributes key terminology, classification and coding standards, and associated resources to promote the creation of more effective and interoperable biomedical information systems and services, including electronic health records. It is an example of a universal reference book of medical knowledge in the broadest sense possible, created with the use of the most effective methods made in computer knowledge processing. UMLS development started in 1986 at the National Library of Medicine, USA. Such terms as semantic representation and medical knowledge processing are used in UMLS ontologies describing all possible domains making up the medical knowledge universe which are created and constantly developed[12].

UMLS has three tools which developers call Knowledge Sources:

✧ Metathesaurus: Terms and codes from many vocabularies, including CPT®, ICD-10-CM, LOINC®, MeSH®, RxNorm, and SNOMED CT®;

✧ Semantic Network: Broad categories (semantic types) and their relationships (semantic relations);

✧ SPECIALIST Lexicon and Lexical Tools: Natural language processing tools.

Considering the international approaches

[9] https://bian.org/ (accessed 19 September 2017)
[10] http://fibo2017.dataversity.net/sessionPop.cfm?confid=116&proposalid=10193 (accessed 19 September 2017)
[11] http://www.hl7.org/ (accessed 19 September 2017)
[12] https://www.nlm.nih.gov/research/umls/ (accessed 19 September 2017)

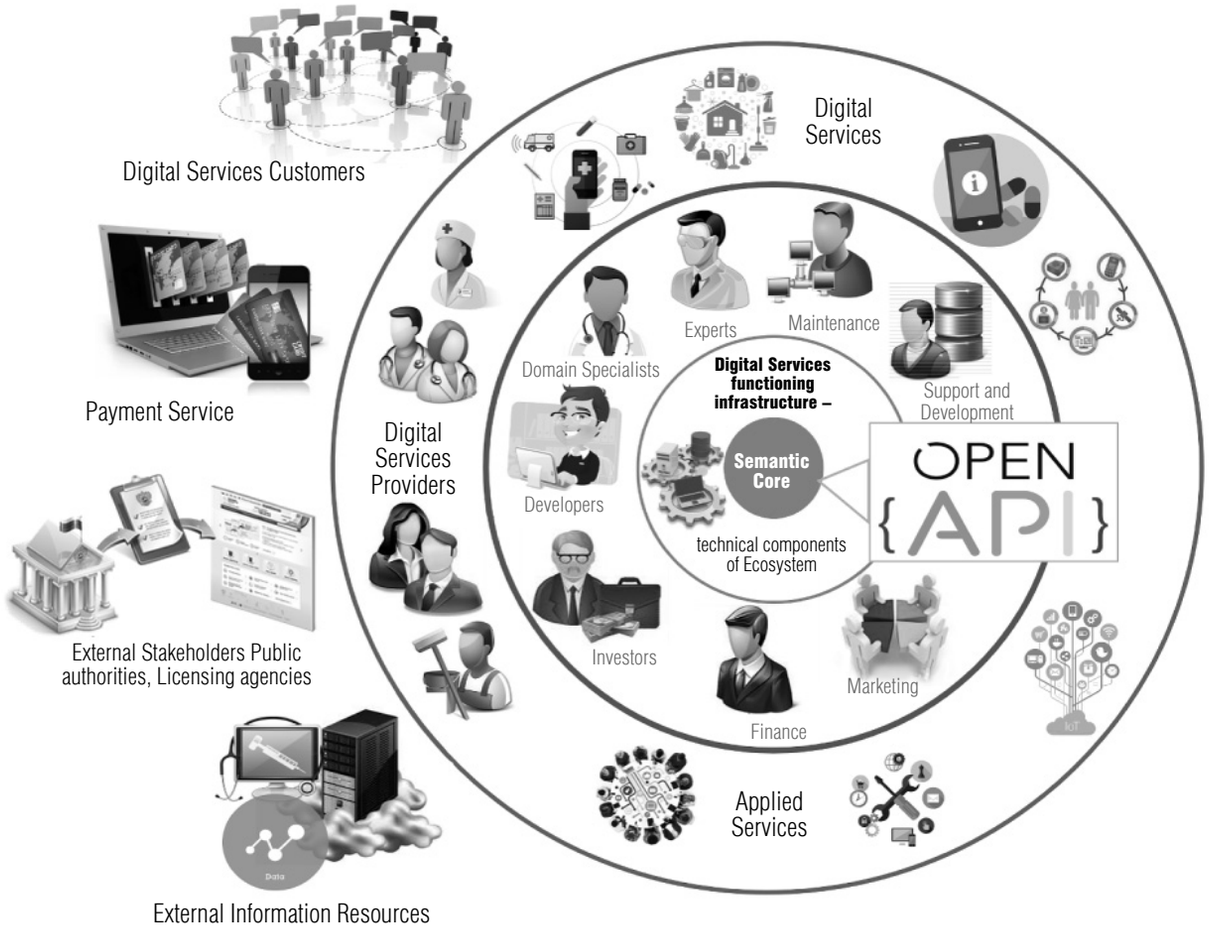

Fig. 1. Conceptual architecture for a digital economic sector ecosystem

to building digital economy platforms in the healthcare and financial sectors, the authors developed a conceptual architecture for a digital economic sector ecosystem.

3. Conceptual architecture of a digital economic sector ecosystem

We represent a conceptual architecture for the digital economic sector ecosystem (*Figure 1*) as a set of levels, interacting with each other in compliance with consistent rules:

1. Semantic core;

2. Digital services functioning infrastructure, which is the technical component of the ecosystem;

3. Digital services for customers and applied services for ecosystem participants, including visual interfaces and the application marketplace, as well as open tools for developers;

4. Personalized gadgets, wearables and IoT devices.

The basis for digital sector ecosystem creation and development can be an industry cluster formed as an independent entity by uniting leading enterprises interested in the promotion of the digital sector (for example, in healthcare it can be National scientific and medical centers), within state regulation, cluster policy realization and other measures of state support [18].

The core of the cluster should be a specially organized entity which will serve as a tool for

the ecosystem's development, support and expansion.

1. For affiliated organizations, the cluster is a project office, center of competence, R&D center which carries out:

- preventive engineering and the approbation of new technologies;
- designing the transformation of existing procedures and business processes, embedding and integration in existing business processes;
- product testing on a platform that simulates digital sector business procedures for the embedding and integration of services;
- development of ecosystem technical components, including OpenAPI and infrastructure services;
- creation of a monitoring system to control the quality of industry digital services through all lifecycle phases.

2. For sector companies, institutions and structures, it is a center for digital transformation and the adaptation of services and products to a digital market, providing:

- digitization of existing processes and the adoption of best solutions within the digital sector;
- legitimization of existing mobile services, cooperation with other enterprises, organizations or institutions for the improvement of service quality;
- simplification of information security procedures.

3. For customers, it is an application access center, supporting:

- extended access mechanisms, concealing the complexity of procedures from clients, requiring no additional hardware facilities and simplifying the access of various customer categories to the services;
- publication of applications developed by ecosystem participants;
- implementation and integration of IoT devices and relevant applications for their everyday use.

4. For development companies, it is a platform for quality control and product expansion, a competence center and business accelerator, facilitating:

- the use of infrastructure applied services and tools for developers;
- the design of applications that can work with OpenAPI and other digital systems and services within the ecosystem;
- tests of applications, publishing and selling them via standard application marketplaces;
- the search for investors and the creation a primary reputation for start-ups;
- participation in the expert community for semantic core development and standardization of the requirements for applications and services.

Conclusion

Hence, for the successful transformation to a digital economy in key economic industries it is advisable to form a state regulation system for the digital sector (sub-sector) as well as:

✦ to provide legitimization to digital services and, if necessary, digital device registration (which is essential in sectors that are highly regulated by public authorities, e.g. in healthcare, education, finance);

✦ to develop a system of state support measures for the digital sector (including the creation of pilot projects, the development of PPPs, the formation of business incubators and stimulation of digital services implementation in industry enterprises;

✦ to create and develop a digital sector ecosystem;

✦ to conceptualize a methodology for the formation of a semantic core as a warehouse of complete and consistent knowledge of the subject area (glossaries, classifiers, thesauri, ontologies, models, standards, interaction schemes,

including international schemes) and to create an expert community:

✦ to provide the filing of the semantic core based on harmonized Russian and international reference books, classifiers, thesauri and ontologies for informational resources, sharing and standardization of interaction between providers and customers of digital services;

✦ to develop an infrastructure for the functioning of digital services, and to ensure the use of information resources existing in the sector;

✦ to create pilot digital services and test them using the tools of the functioning infrastructure, and then to provide their seamless integration with existing systems for subsequent digital transformation of business processes performed by sector enterprises;

✦ to provide the transfer of innovative solutions to sector enterprises and support for domestic start-up companies.

Based on the approaches presented, the authors of this paper prepared detailed suggestions for the digital transformation in healthcare, the fundamental propositions of which were included in the "Digital Healthcare" section of the "Digital Economy" program draft worked out in accordance with "Assignment List for Realization of President's Message to the Federal Assembly" [19].

With the expansion of the adopted program for the "Digital Economy" [20] the implementation of the suggested sectoral strategy for digital transformation will create significant benefits through the rapid development of new, customer-oriented products and solutions, as well as high flexibility and adaptability to changing needs. The effectiveness of the strategy is largely determined by the creation of an ecosystem that provides favorable conditions enabling innovations in developing companies. ■

References

1. Schwab K. (2016) *Chetvertaya promyshlennaya revolyutsiya* [The fourth industrial revolution]. Moscow: Eksmo (in Russian).
2. Mitin V. (2017) *Sem' opredeleniy tsifrovoy ekonomiki* [Seven definitions of digital economy]. Available at: https://www.crn.ru/news/detail.php?ID=116780 (accessed 19 September 2017) (in Russian).
3. Akatkin Y., Yasinovskaya E., Drozhzhinov V., Konyavskiy V. (2016) Towards the digital government in Russia: Integrative approach. Proceedings of *International Conference on Electronic Governance and Open Society: Challenges in Eurasia (EGOSE'16). St. Petersburg, Russia, 22–23 November 2016.* N.Y.: ACM New York, pp. 41–48.
4. Markovitch S., Willmott P. (2014) *Accelerating the digitization of business processes.* Available at: http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/accelerating-the-digitization-of-business-processes (accessed 19 September 2017).
5. Stogney A., Sedov D. (2016) *Banki dlya gikov: eksperty nazvali liderov finansovykh innovatsiy* [Banks for geeks: experts listed the leaders of finance innovations]. Available at: http://money.rbc.ru/news/57bb7ffe9a7947340fa28c1f (accessed 19 September 2017) (in Russian).
6. Pertlweiser M. (2016). *For the banking of the future – Deutsche Bank's digital factory.* Available at: https://www.db.com/newsroom_news/16_09_29_Digital_Factory_Opening_final_Englisch.pdf (accessed 19 September 2017).
7. Deutsche Bank (2016) *Deutsche Bank opens its data store to external software developers* (October 2016). Available at: https://www.db.com/newsroom_news/2016/medien/deutsche-bank-opens-its-data-store-to-external-software-developers-en-11735.htm (accessed 19 September 2017).
8. Kantyshev P. (2016) «Sbertekh» stal pervym po golovam [Sbertech became a headhunting leader]. *Vedomosti*, no. 4173 (3 October 2016) (in Russian).
9. TAdviser (2016) *«Sberbank rasskazal na TAdviser SummIT o revolyutsionnoy transformatsii IT-sistem* [Sberbank told on TAdviser SummIT about revolutionary transformation of IT systems]. Available at: http://tadviser.ru/a/318028 (accessed 19 September 2017) (in Russian).

10. Karpov O., Akatkin Y., Konyavsky V., Mikerin D. (2016) *Tsifrovoe zdravookhranenie v tsifrovom obshchestve* [Digital healthcare in digital society]. Moscow: Delovoy Express (in Russian).
11. Accenture (2016) *Accenture digital health vision 2016.* Available at: https://www.accenture.com/us-en/insight-healthcare-technology-vision-2016 (accessed 19 September 2017).
12. GOV.UK (2017) *Personalised health and care 2020.* Available at: https://www.gov.uk/government/publications/personalised-health-and-care-2020 (accessed 19 September 2017).
13. Oxford Economics (2011) *The new digital economy. How it will transform business*. Available at: http://www.pwc.com/mt/en/publications/assets/the-new-digital-economy.pdf (accessed 19 September 2017).
14. European Parliamentary Research Service Blog (2013) 2*1% GDP productivity growth – Why the digital economy is important!* Available at: https://epthinktank.eu/2013/11/11/21-gdp-productivity-growth-why-the-digital-economy-is-important/ (accessed 19 September 2017).
15. Betancourt R.R. (2017) *Distribution services and the digital economy: Implications for GDP measurement, productivity and household welfare*. Available at: http://economics.fiu.edu/events/2017/seminar-roger-betancourt-2/betancourt2017.pdf (accessed 19 September 2017).
16. Reingold L.A., Volkov A.I., Kopaygorodsky A.N., Pustozerov E.Y. (2016) Semanticheskaya interoperabel'nost' v reshenii finansovykh zadach i sposoby ee izmereniya [Semantic interoperability in solving financial problems and ways of measuring it]. *Applied Informatics*, vol. 11, no. 4 (64), pp. 115–134 (in Russian).
17. Zaino J. (2016) *Banking on FIBO: Financial institutions turn to semantic standard.* DATAVERSITY (October 2016). Available at: http://www.dataversity.net/banking-fibo-financial-institutions-turn-standard-value-compliance/ (accessed 19 September 2017).
18. Skvortsov E.N., Guskova N.D. (2016) *Analiz organizatsionnykh struktur sistem upravleniya innovatsionnymi promyshlennymi klasterami v Rossii* [Analysis of organizational structures of management systems for innovative industrial clusters in Russia]. Economics Portal (14 June 2016). Available at: http://institutiones.com/innovations/2806-analiz-organizacionnyx-struktur.html (accessed 19 September 2017) (in Russian).
19. The President of Russia (2016) *Perechen' porucheniy po realizatsii Poslaniya Prezidenta Federal'nomu Sobraniyu* [Assignment list for realization of the President's Message to the Federal Assembly]. Available at: http://kremlin.ru/acts/assignments/orders/53425 (accessed 19 September 2017) (in Russian).
20. The Government of the Russian Federation (2017) *Programma «Tsifrovaya ekonomika Rossiyskoy Federatsii»* [The Digital Economy of the Russian Federation Program]. Approved by the Government of the Russian Federation in its resolution No. 1632-r, 28 July 2017. Available at: http://ac.gov.ru/files/content/14091/1632-r-pdf.pdf (accessed 19 September 2017) (in Russian).

Improvement of the procurement process using the integrated tender recommendation system

Zinaida K. Avdeeva
Senior Researcher, Trapeznikov Institute of Control Sciences;
Associate Professor, Department of Innovation and Business in Information Technologies
National Research University Higher School of Economics
Address: 65, Profsoyuznaya Street, Moscow, 117997, Russian Federation
E-mail: avdeeva@hse.ru

Alexander A. Utrobin
Bachelor in Business Informatics
National Research University Higher School of Economics
Address: 20, Myasnitskaya Street, Moscow, 101000, Russian Federation
E-mail: alexanderutrobin11@gmail.com

Ilya Y. Lykov
MSc Program Student, Faculty of World Economy and International Affairs
National Research University Higher School of Economics
Address: 20, Myasnitskaya Street, Moscow, 101000, Russian Federation
E-mail: lykov.ilya@gmail.com

Abstract

Currently, tender procedures occupy an important place in the work of the procurement department of any company. Most of the modern recommendation services operate on web-platforms. Implementation of a tendering system in a company can increase the level of maturity of the procurement process and will not require serious changes in the structure of the processes. This article is devoted to the study of the structure of an integrated tender recommendation system. The integrated tender recommendation system is based on procedures prescribed by federal laws, aggregated tenders from different e-trade sites (presenting state and commercial platforms); it offers to its users additional services.

The main purpose of the study is to develop an effective model of an integrated tender recommendation system. In the description part of this article, we present information on peculiarities of the tender procedure in the Russian Federation and modern advisory services are considered. The functional advantages of an integrated system in comparison with the web platform are set out. The structure of the system is designed using several approaches. Using the IDEF0 methodology, a functional model of the system that reflects the work of processes has been developed and described. The operation of the main system and subsystems has been analyzed using the projected diagrams of the DFD methodology. A mathematical model of dynamic filtering of tenders for creating recommendations to users is described. Relying on the basic principles of collaborative filtering and, with the help of appropriate algorithms, the integrated system gives recommendations to users and determines the probability of success in a particular tender. Application of such technology of tenders is possible in companies of different scale. The developed structure of the integrated system and filtering methods for recommendations are based on the basic principles of a new international trend – e-tendering.

Key words: recommendation services, tender, electronic trading platform, integrated system, dynamic filtration.

Citation: Avdeeva Z.K., Utrobin A.A., Lykov I.Y. (2017) Improvement of the procurement process using the integrated tender recommendation system. *Business Informatics*, no. 4 (42), pp. 29–39. DOI: 10.17323/1998-0663.2017.4.29.39.

Introduction

With the development of information technology, most of the world's countries have switched to electronic trading platforms. Tender is an integral part of the governmental and commercial sectors of the economy. There are many electronic trading platforms (ETP) and recommendation web services that collect requirements of different categories and deliver information through queries. Such sites and services also offer additional services.

The efficient operation of an enterprise's purchasing department is directly dependent on automation of the tender process. In "routine" tenders, employees regularly make mistakes during preparation of applications or when assembling a package of documents; they also may miss important changes in the specifications of terms or periods. In order to avoid errors and increase the purchasing department's efficiency, an integrated tendering system should be implemented at the enterprise. As a result, employees will be able to upload documents into the system and conduct all operations related with the tender, from planning to signing contracts with customers. An integrated system of tendering combines the functions of classical ETP and recommendation web services. It also allows users to automate "routine" tenders and proposes the recommended tenders. The system always provides warnings regarding changes in the terms of the application, the timing and composition of documents.

Modern ways of tender processing represent the object of this research, and the functioning of an integrated tender recommendation system is the subject of this study.

This article has the following structure. In the first section, peculiarities of competitions (tenders) in the Russian Federation are considered. In the second section, the work of modern advisory services is analyzed. In the third section, the model of an integrated tender recommendation system is presented, an effective structure of the system is designed, and the functions and subsystems are defined. This section also provides a dynamic filtration system, which is based on certain mathematical methods used in recommendation services. Finally, in the fourth section, the results of the study are presented.

1. Features of electronic tenders in Russia

A tender is a method of competing for a contract among which is compliant with state regulations and, under certain conditions, depends on the internal rules of the trading platform. As a result of tenders, the process of concluding a contract between the customer and the winner is performed. With this type of trading operation, the customer makes a choice from a set of proposals from different companies. As a result, the selected company undertakes an obligation to perform work, service or delivery of some goods to the customer. In Russian legal acts (http://zakupki-tendery.ru/zakonodatelstvo), the concept of "tender" is not officially used; instead we find the term "competition".

Modern web sites designed for holding trades are presented in three fields: state order, municipal order and commercial order. Each area is

related with a certain organizer and the inherent specialized selection criteria. The organizers include government agencies, local governments and business entities. The special criteria are value for money, information about orders performed by the participant earlier, as well as additional conditions provided by the participants in the application. Within these three fields, various forms of the competition are available. The most frequently used tender forms are an open tender, a closed tender, a specialized closed auction, a request for quotations, a purchase from one supplier, and a request for proposals.

Since January 2011, all government customers are required to register and conduct tenders on one of five federal state electronic platforms: Sberbank AST, Rosetorg, OSET, ETP, MMBV "Goszakupki". Since January 2014, the Federal Law of 05.04.2013 No. 44-FZ "On the contract system in the sphere of procurement of goods, works and services to provide for state and municipal needs" is applied. Now, in addition to the tender procedures, analysis of previous trades is conducted in the electronic trading platforms, the statistics on the results of the competitions are posted, and information on planning and orders is provided. The federal law from 18.07.2011 No. 223-FZ "On procurement of goods, works and services by certain types of legal entities" regulates activities of legal entities – state owned corporations, monopolies, autonomous institutions, economic societies, companies of small and medium business. Before the appearance of the law, the activities of legal entities and the organization of competition were regulated through a variety of legal acts.

2. Modern recommendation services for tender search and planning

In order to find potential business partners, companies may use search systems, web aggregators of state trading platforms, industry related and universal sites, free and commercial bidding sites, forums, e-trading platforms of debtors' property sales. With the development of information technologies, most modern ETPs provide for their users certain additional functions of planning and holding tenders. However, at present only commercial systems present comprehensive services for tender search and planning. These services solve a number of problems, such as the huge number of online sites of competitions, specific interface of systems, weak return from search, complexity of monitoring the changes in tender conditions, as well as organizing documents.

Let us consider the results of the survey in large companies in Australia, the UK, China, the USA and Japan, represented in [1]. There are three architectural types of modern electronic trading platforms. The most popular is an architecture based on the principal. In this case, the state acts both as a customer and as an organizer of the system. The second type is an architecture based on an independent third party. A vivid example of such a system is a commercial electronic trading platform. The third type is an architecture based on responsibilities distributed between independent third parties. In such systems, functions are performed by different legal entities. For example, one company deals with the security of the system, another is responsible for the electronic digital signature. At present, the advisory systems for the search and planning of the competition are presented in the last two types of architecture.

3. An integrated tender recommendation system

Improving the functioning of the purchasing department of a company can be carried out with the help of an integrated tendering system. For modern Russia, we usually find that competitions related recommendation services provide their services through a web site and cloud computing. The integrated system has several advantages: minimization of the probability of system failures at critical times, con-

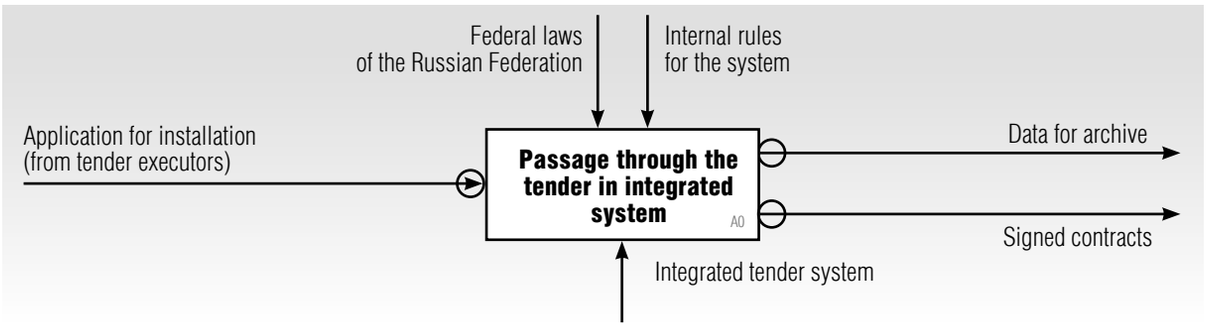

Fig. 1. Context diagram of the tender process in the integrated system

fidentiality of correspondence and control over access to tender information, minimization of financial and legal risks associated with third parties, using an effective system of tender filtering, application of an individual approach to each company registered in the system.

In general, the development of online bidding systems for large companies aims at improving the efficiency of the procurement process through open tenders. For medium-sized businesses, this means an opportunity to expand partnerships and find new business partners. For a small business, it is a convenient search system and an effective system of recommendations which the company can use to determine tenders where it is most likely to win, to view its statistics and to develop a specific strategy for activities of the procurement department.

3.1. The process of passing through the tender in an integrated tender system

The formalization of passing through the tender is constructed on the basis of the IDEF0 methodology. *Figure 1* shows a contextual diagram of a tender passage process within an integrated system, as well as related additional information. Entrance objects are applications from the executors; the controls are the Federal Law of the Russian Federation and internal rules of the system; the mechanism is an integrated tender system; and the output objects are data for the archive and signed contracts.

The main functions provided by the integrated system are registration in the system, manual search, automatic search, rendering additional services, reporting the tender passage and archiving (*Figure 2*).

Let us consider the diagram by its stages. Registration in the system involves working with approved applications received from companies. This function is performed using an integrated system and is based on the company's internal rules. As a result, the function provides information for the file and user's ID. Companies registered in the system can perform a manual search through appropriate inquiries. Automatic search involves recommendations to users, providing information on tenders that are important for the company, as well as automatic submission of applications for regular competitions and setting up an electronic signature in the user's account. The function of providing additional services implies an individual approach to users, if necessary – assistance in drawing up an application for a tender, reminders related with changes in terms and conditions, as well as the preparation of a package of documents. The function of forming a package of documents uses the results of previous functions – the results of manual and automatic searches and additional services. The function of passing through the tender implies working with applications and packages of documents received from suppliers. With the help of the archiving function, data analysis is performed, which is necessary for the automatic search function. In addition, the data to be used for further work are archived.

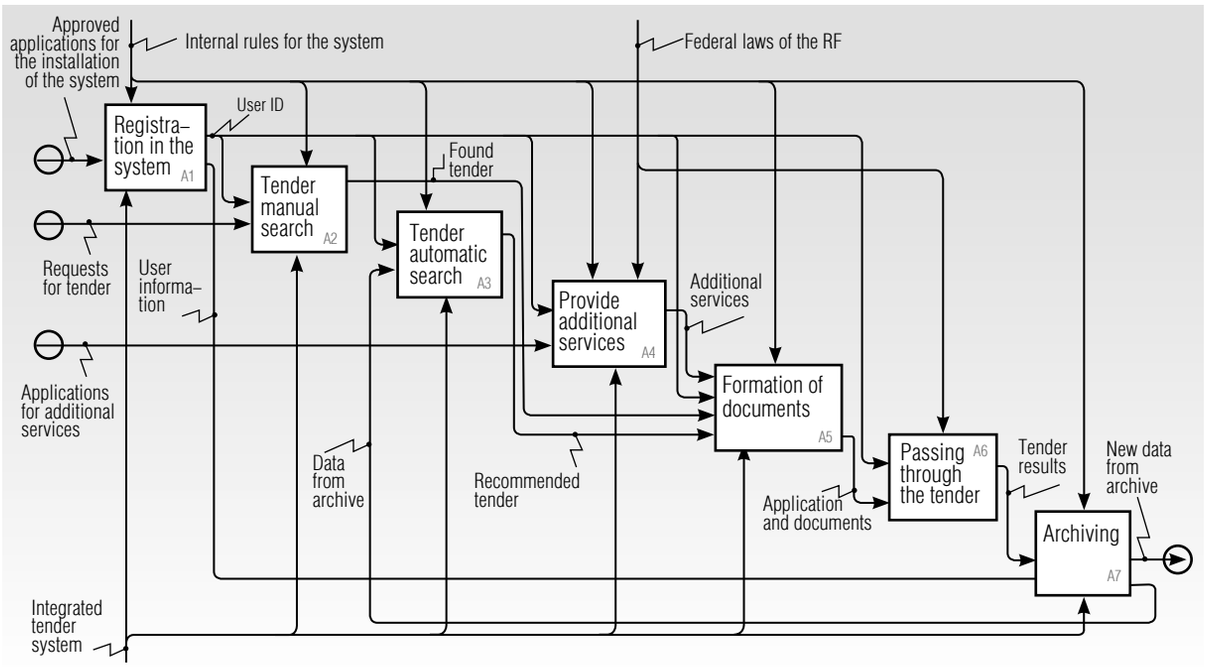

Fig. 2. The first level diagram of the tender process in an integrated system

3.2. Analysis of data flows in an integrated system

An integrated system implies working with executors of future and current competitions, as well as interaction with the management of the system. A relevant context DFD diagram is presented in *Figure 3*. From registered users, the system obtains information about the company, then on the basis of joint work in a personal account certain statistics supporting the recommendation system are generated. From the system, companies receive necessary information on tenders. From the management, the system receives internal rules, and management receives data for the archive.

The context diagram shows all the subsystems of the integrated tendering system (*Figure 4*). The archive subsystem performs storage of information on transactions and customers, as well as their search in the data warehouse. The search subsystem allows users to search and sort tenders. The search may be carried out using the manual method. It is also possible to select tenders according to required criteria among tenders recommended by the system. The subsystem of additional services provides the user with the opportunity to choose necessary services with the help of which the probability of winning the tender may be increased. The subsystem of the tender passage performs interaction of the installed program with state and commercial sites. Through a well-designed interface, users can go directly to the requirements of the integrated system. The dynamic filtering subsystem is looking for tenders recommended for each company on an individual basis, relying on mathematical methods and

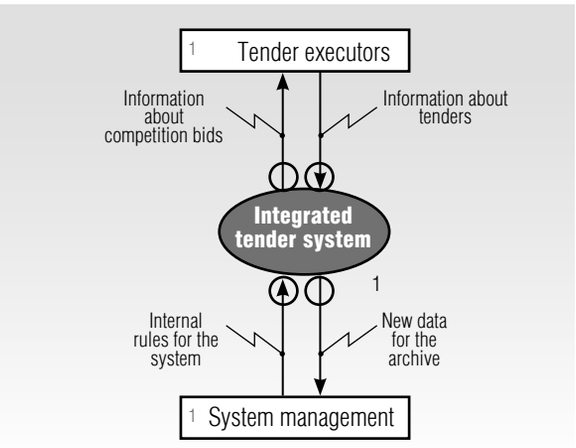

Fig. 3. Context diagram of the integrated tender system

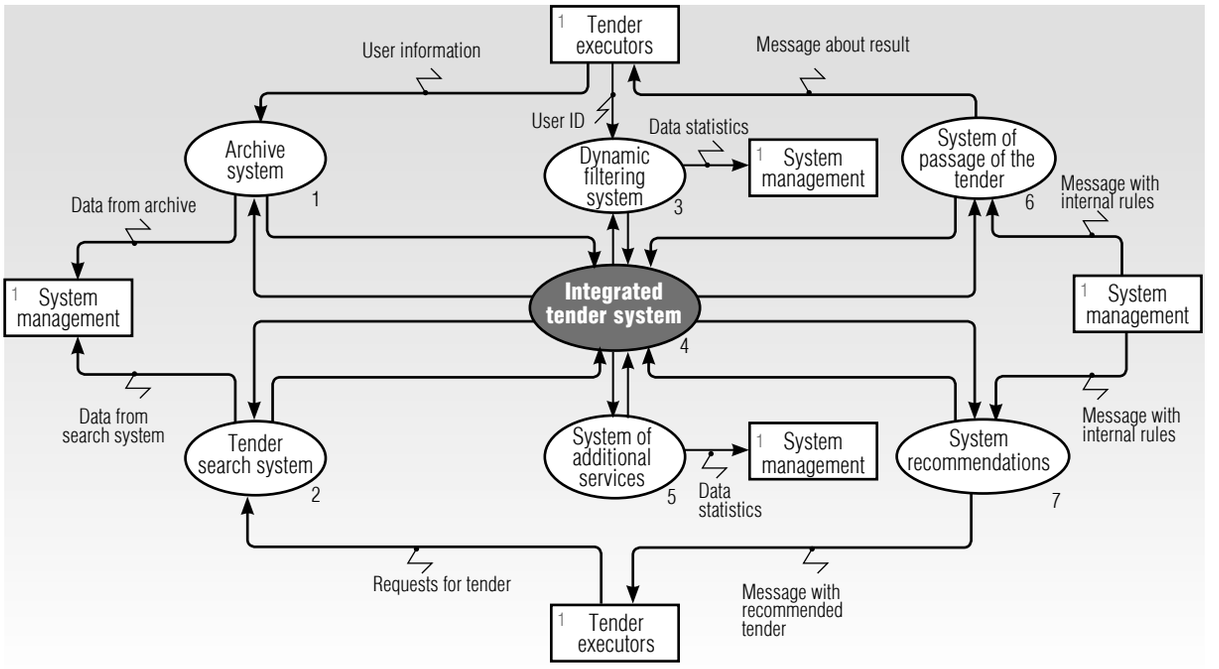

Fig. 4. Zero–level diagram of the integrated tender system

statistics generated in the archive system. The recommendation subsystem provides the user with the results of the work under a dynamic filtering system. If the user decides to conduct an automatic compilation and send regular requests for tenders, the recommendation system performs this work with the help of a pre-defined algorithm and digital signature.

3.3. Dynamic filtering subsystem

The maturity of the company's procurement process is determined by the ability to automate part of the processes, thus saving resources of the enterprise [2]. The electronic format makes it possible to avoid transaction costs arising during the search for a suitable supplier. In addition, the tender format opens up a number of possibilities for contracting companies in their search for customers. However, the process of searching for a customer can be quite time-consuming and requires collection of large amounts of information. At this stage, the likelihood for both sides to make the wrong choice is increased. In this part of the work, a model that allows automatic selection of the tender is proposed. The model is based on the use of the basic principles of an advisory service. The key difference is that the method is applied for selection of a tender in which the contractor can participate. It is based on the principle of collaborative filtering: recommendations are based on the history of the estimates made both by a certain user and other users.

To start the recommendation system, users should assess the tenders in which they participated through this program. This condition is mandatory, as further recommendations will be based on these data. To evaluate different tenders, users should provide a certain set of characteristics. Ratings for each parameter are in the range of from 1 to 5.

Among the characteristics examined are the following:

- ✦ stability of documents status;
- ✦ stability of application conditions;
- ✦ stability of terms;
- ✦ level of feedback with the customer;

✦ degree of the deal's transparency;

✦ correctness of application conditions.

Relying on estimations for each of the listed characteristics, the overall assessment of the tender is formed, calculated by the weighted average formula:

$$\overline{x} = \frac{\sum_{i=1}^{n} w_i \, x_i}{n},$$

where $\overline{x}$ – score assigned by the user for a particular tender;

x_i – score related with the i-th characteristic, $i = 1, ..., n$;

w_i – weight of the i-th characteristic, $i = 1, ..., n$;

n – number of characteristics.

The purpose of the algorithm is to generate a recommendation for each contractor. The recommendation is a tender offer which may be interesting for this executor. The system makes a recommendation, predicting the assessment that the user would put for a particular tender on the basis of its previous estimates. It is important that when logging into the system the user specifies the area of its activity for selection of the appropriate proposal.

The simplest metric that enables us to estimate proximity of two elements is the Euclidean distance between vectors. Obviously, that the smaller this value is, the closer the elements are and thus the more similarities they have. Relying on this certain recommendation, collection may be arranged. However, this method is not the only one. A better choice is possible using calculations that are more complex. Let us consider several possible algorithms for predicting the evaluation.

First, let us consider more advanced filtering methods. There is a set of executors ($u \in \mathrm{U}$) and a set of tenders suitable to the particular executor ($t \in \mathrm{T}$). On the basis of the processed information, the system makes a recommendation e_{ut}.

The first algorithm is associated with clustering of users. It is necessary to introduce the function of similarity $\mathrm{sim}(u, v)$, where u is the selected user of the system, and v is a user compared with the first one. This function shows how similar are scoring histories of the two users.

This function can be set in several ways. For example, it can be characterized as a percentage of tenders in which the two companies participate, or a share of tenders for which the companies have matched estimates. Further, users are divided into clusters, relying on the "like with like" principle: $u \to \mathrm{G}(u)$.

The predicted user's score of a certain tender is calculated as the average of all of the cluster to which the user belongs:

$$\widehat{e_{ut}} = \frac{1}{\left|\mathrm{G}(u)\right|} \sum_{v \in \mathrm{G}(u)} e_{vt},$$

where $\widehat{e_{ut}}$ – predicted score of tender t by user u, $t \in \mathrm{T}$, $u \in \mathrm{G}(u)$;

$|\mathrm{G}(u)|$ – the number of users in a cluster, for which user u belongs;

e_{vt} – score of tender t by user v (another user of the cluster), $t \in \mathrm{T}$, $v \in \mathrm{G}(u)$;

T – a set of tenders;

G(u) – a set of users of the cluster.

Let us consider another "user-based" algorithm which can be used to create recommendations [4]. We give a formula for predicting the evaluation. Here we also use the similarity function that was mentioned earlier:

$$\widehat{e_{ut}} = \overline{e_u} + \frac{\sum_{v \in \mathrm{U}} \mathrm{sim}(u,v)\left(e_{vt} - \overline{e_v}\right)}{\sum_{v \in \mathrm{U}} \mathrm{sim}(u,v)},$$

where $\widehat{e_{ut}}$ – predicted score of tender t by user u, $t \in \mathrm{T}$, $u \in \mathrm{G}(u)$;

$\overline{e_u}$ – average typical score of user u, $u \in \mathrm{G}(u)$;

$\overline{e_v}$ – average score of a similar user v, $v \in \mathrm{G}(u)$;

e_{vt} – actual score of tender t by a similar user v, $t \in \mathrm{T}$, $v \in \mathrm{G}(u)$;

$\mathrm{sim}(u, v)$ – function of similarity of scores of the users u and v, $u \in \mathrm{G}(u)$, $v \in \mathrm{G}(u)$;

T – a set of tenders;

G(u) – a set of users of the cluster.

Another possible algorithm of estimation is based on the principle of singular decomposition [4]:

$$\widehat{e_{ut}}(\beta) = p_u^{\mathrm{T}} q_t,$$
$$\beta = \{p_u; q_t \mid u \in \mathrm{U}, t \in \mathrm{T}\},$$

where $\widehat{e_{ut}}(\beta)$ – predicted score of tender t by user u, $t \in \mathrm{T}$, $u \in \mathrm{G}(u)$;

p_u – possible score of user u, $u \in \mathrm{U}$;

q_t – possible score of tender t, $t \in \mathrm{T}$;

β – a couple of scores (possible score of user p_u and tender q_t);

T – a set of tenders;

U – a set of users.

In this case, p_u and q_t are parameters that should be evaluated. It is necessary to optimize the quality of future predictions:

$$(\widehat{e_{ut}}(\beta) - e_{ut})^2 \to \min, \qquad (1)$$

where $\widehat{e_{ut}}(\beta)$ – predicted score of tender t by user u, $t \in \mathrm{T}$, $u \in \mathrm{G}(u)$;

e_{ut} – actual score of tender t by user u, $t \in \mathrm{T}$, $u \in \mathrm{G}(u)$;

β – a set of possible scores of user u and tender t;

T – a set of tenders;

U – a set of users.

While the system has a number of previous estimates, we can speak about availability of a training sample. Let us rewrite (1), performing regularization:

$$\sum_{u,t} (\widehat{e_{ut}}(\beta) - e_{ut})^2 + \lambda \sum_{\varphi \in \beta} \varphi^2 \to \min, \qquad (2)$$

where $\widehat{e_{ut}}(\beta)$ – predicted score of tender t by user u, $t \in \mathrm{T}$, $u \in \mathrm{G}(u)$;

e_{ut} – actual score of tender t by user u, $t \in \mathrm{T}$, $u \in \mathrm{G}(u)$;

λ – regularization parameter, $\lambda > 0$;

φ – stabilizing functional;

β – a set of possible scores of user u and tender t;

T – a set of tenders;

U – a set of users.

The feature of the regularization parameter λ is that when $\lambda \to 0$ and the availability of errors in source data the solution of a correct task strives for the true solution of an appropriate incorrect task. The aim of the stabilizing functional φ is to provide stability of solution of the task if source data are inaccurate.

Let us rewrite (2) for the purposes of the task under consideration:

$$\sum_{u,t} (p_u^{\mathrm{T}} q_t - e_{ut})^2 + \lambda \left(\sum_u \|p_u\|^2 + \sum_t \|q_t\|^2 \right) \to \min,$$

where β – a set of possible scores of user u and tender t;

p_u – possible score of user u, $u \in \mathrm{U}$;

$\|p_u\|$ – norm of scores of user u, $u \in \mathrm{U}$;

q_t – possible score of tender t, $t \in \mathrm{T}$;

$\|p_u\|$ – norm of scores of tender t, $t \in \mathrm{T}$;

e_{ut} – actual score of tender t by user u, $t \in \mathrm{T}$, $u \in \mathrm{G}(u)$;

λ – regularization parameter, $\lambda > 0$;

T – a set of tenders;

U – a set of users.

Thus, the problem has been reduced to finding the minimum of the functional. This allows us to find the most correct assessments for the predictions.

A separate advantage of the system so developed is not only automatic selection of the tender for the contractor, but also the evaluation of the probability of winning it. To calculate the probability, the system analyzes several criteria, assigning them ratings from 1 to 10. Among the criteria are the following:

✦ the degree of occupancy of the tender's executor;

✦ user's experience in working with the system;

✦ compliance with the business area of the tender;

✦ production costs;

✦ geographic distance (if the user is irrelevant for a particular situation, the system uses the first four criteria for assessing the probability of winning).

The probability is calculated as follows:

$$\mathrm{P}(\mathrm{T}) = \frac{\sum_{i=1}^{n} A_i}{\sum_{i=1}^{n} B_i} \cdot \frac{1}{n}, \qquad (3)$$

where P(T) – probability of winning the tender;

A_i – the score assigned by the system to the tender, relying on the evaluation presented above;

B_i – the maximal score;

$\sum_{i=1}^{n} A_i$ – the sum of the scores assigned to the tender relying on the criteria;

$\sum_{i=1}^{n} B_i$ – the maximal possible sum of the scores;

n – the number of participants in the tender.

In addition to calculating the percentage of probability, the partitioning of the results by categories is provided (*Table 1*).

Table 1.

Categories of probability of winning the tender

Categories	Name of the category	Probability of winning (%)
Category 1	The highest probability of winning	90 – 100
Category 2	High probability of winning	70 – 89
Category 3	Average probability of winning	50 – 69
Category 4	Low probability of winning	30 – 49
Category 5	The lowest probability of winning	0 – 29

Thus, in this part of the paper the following filtering techniques are discussed: estimation of important parameters of the tender and calculating a recommendation by the weighted average; simple metrics through Euclidean distance; the algorithm associated with clustering of users; a "user-based" algorithm; an algorithm based on the singular decomposition of the matrix. The probability of winning the tender is evaluated using the criteria and calculation of estimates using formula (3).

4. The results of the study

As a result of the study, an integrated recommendation tender system has been designed. Due to aggregation and data selection from different ETPs, so-called "false" tenders can be identified and excluded from the set of tenders under consideration (lots with single provider, lots with special constraints, etc.). With the help of automated functions and recommendations, the level of transaction costs related with a tender executor is decreased. A high-quality integrated system includes the ability to store and process huge data sets.

This system can be used by companies of all sizes and all kinds of production activities. With the automation of processes in the purchasing department related to the tender, companies can save money. The effectiveness of the system depends on the elaborated structure of sub-systems, environs of the basic system functions, and on basic algorithms of the filtering system. The development of an integrated system was based on the basic principles of modern trends of e-tendering. The main features of this area are identified in [5]; they include minimization of legal and financial risks, minimization of errors in preparation of applications, and an effective tenders search system.

In the near future, e-tendering systems will be able to reach a higher level of quality. Many foreign authors speak about possible changes in the structure of the tender (for example, in [6]), and also examine the details of such changes. With the development of recommendation service technologies, new integrated systems will appear that are able to automate part of the process in the purchasing department, resulting in reduced costs of making a tender and procurement through tender. They will increase the level of maturity of the procurement process, will increase deals flows, will expand partner bases and will automate "routine" tenders.

4.1. An example of the filtering system in operation

Let us present some calculations illustrating the operation of the filtering system. For example, we consider an algorithm based on the "user-based" method. Suppose that there are

two users *A* and *B*. There is information on how they evaluated tenders in which they took part (*Tables* 2 and 3). It is necessary to predict what score will be assigned by user *A* to the tender 6.

Table 2.

Evaluations of user A

Tender	Score
Tender 1	5
Tender 2	3
Tender 3	4
Tender 4	2
Tender 5	5

Table 3.

Evaluations of user B

Tender	Score
Tender 1	5
Tender 2	3
Tender 4	4
Tender 5	5
Tender 6	1
Tender 8	2

Let us introduce a similarity function as a share of tenders for which the performers have matched estimates. Both customers participated in tenders 1, 2, 4 and 5. However, the evaluations coincided only for tenders 1, 2 and 5. Thus the share of tenders with coinciding estimates for user *A* is 3/5.

Then we apply the formula for evaluation prediction and calculate an evaluation prediction of tender 6 by user A:

$$\widehat{e_{ut}} = \overline{e_u} + \frac{\sum_{v \in U} \mathrm{sim}(u,v)\left(e_{vt} - \overline{e_v}\right)}{\sum_{v \in U} \mathrm{sim}(u,v)},$$

where $\widehat{e_{ut}}$ – predicted score of tender t by user u, $t \in T$, $u \in G(u)$;

$\overline{e_u}$ – average typical score of user u, $u \in G(u)$;

$\overline{e_v}$ – average score of a similar user v, $v \in G(u)$;

e_{vt} – actual score of tender t by a similar user v, $t \in T$, $v \in G(u)$;

$\mathrm{sim}(u, v)$ – function of similarity of scores of the users u and v, $u \in G(u)$, $v \in G(u)$;

T – a set of tenders;

$G(u)$ – a set of users of the cluster.

The average estimate of a typical user ($\overline{e_u}$) we take as the average number of scores set out by the system. If we are afraid to overestimate the assessment, the geometric mean can be taken, since it is always less than the arithmetic mean, except when all the estimates are equal. In this case, it is 3.8, and in the case of using the geometric average it is 3.59. For user *B*, the average ($\overline{e_v}$) is 3.3. Making calculations, we get 3.83. This means that it is likely that user *A* will assign the score 4 to tender 6.

4.2. An example of calculations

To estimate the probability of winning the tender, the system uses a proposed technique assigning to each of the parameters a score from 1 to 10. It is worth paying attention to the fact that different parameters are taken into account by a system in a different way. Once again we turn to the characteristics used in the calculation of probabilities and give some comments on the methodology for assigning the scores:

✦ degree of tender executor occupancy (in this case, the relationship is reversed: the more loaded the executor is, the fewer points he receives for this feature);

✦ user's experience in working with the system (direct relationship: the more experienced a user is, the higher is the score assigned to him);

✦ correspondence of business activity to the subject of the considered tender (direct relationship);

✦ cost of production (in this case, the relationship is reversed: the lower the cost of production from this manufacturer, the greater score it gets. The lower the production cost, the lower price the user may ask, and therefore, it has a better chance of getting the order, which should be reflected in the formula);

✦ geographical distance (if this parameter is irrelevant for a particular situation, the system uses the first four criteria for assessing the probability of winning. There is a direct relationship: the closer the performer is to the customer, the greater the probability of winning).

Assume that there are two firms participating in the tender – *A* and *B*. *Table 4* presents evaluations estimated by the system.

Table 4.

Scores assigned by the system

Parameter	Firm A	Firm B
Occupancy degree	8	2
Experience	7	4
Compliance	10	7
Cost and price	8	8
Geographical distance	10	5

Using formula (3) to define winning probability, we find that the wining probability of firm A is 43%, and the wining probability of firm B is 26 %. It is obvious that company A has a greater chance to win, and with the growth of the number of participants the probability of winning will decrease.

Conclusion

Thus, in this article the efficiency of the integrated tender system has been considered. The current state of the federal legislation of the Russian Federation is analyzed, and specific features of conducting competitions are highlighted. The problems that may be solved using the system for companies of different scale are formulated. The structure of a recommendation service using IDEF0 and DFD methodologies has been designed. The mathematical methods and algorithms that may be applicable within the dynamic tenders filtering subsystem are considered. ■

References

1. Du R., Foo E., Nieto J.G., Boyd C. (2005) Designing secure e-tendering systems. *Trust, privacy and security in digital business* / S. Katsikas, J. López, G. Pernul (eds). Lecture Notes in Computer Science, vol. 3592. Springer, Berlin, Heidelberg, pp. 70–79.

2. KPMG (2011) *Model' zrelosti zakupok. Analiz funktsii zakupok v rossiyskikh kompaniyakh* [Procurement maturity model. Analysis of the procurement function in Russian companies]. Available at: https://home.kpmg.com/ru/ru/home/insights/2011/06/model-of-procurement-maturity.html (accessed 01 September 2017) (in Russian).

3. Sarwar B.M., Karypis G., Konstan J.A., Reidl J. (2001) Item-based collaborative filtering recommendation algorithms. Proceedings of the *10th International World Wide Web Conference. Hong Kong, 1–5 May 2001.* P. 285–295.

4. Jones T.M. (2014) *Rekomendatel'nye sistemy. Chast' 1. Vvedenie v podkhody i algoritmy* [Recommendation systems. Part 1. Introduction to approaches and algorithms]. Available at: https://www.ibm.com/developerworks/ru/library/os-recommender1/ (accessed 15 March 2017) (in Russian).

5. Dawson E.P., Duncan W.D., Christensen S.A., Black P.J., Foo E., Du R., Boyd C.A., Nieto J. (2005) *E-tendering – Security and legal issues: Research report.* Brisbane: CRC Construction Innovation.

6. Lavelle D., Bardon A. (2009) E-tendering in construction: time for a change? *Northumbria Working Paper Series: Interdisciplinary Studies in the Built and Virtual Environment*, no. 2 (2), pp. 104–112.

Specifics of the integration of Business Intelligence and Big Data technologies in the processes of economic analysis

Stanislav Mitrovic
University of Novi Sad, Republic of Serbia;
Doctoral Student of the Faculty of Economics
Lomonosov Moscow State University
Address: 1, Leninskie Gory, Moscow, 119991, Russian Federation
E-mail: Mitrovic.Stanislav@hotmail.com

Abstract

The volume of data used for economic analysis of the activities of organizations is growing every year. Despite the fact that all information required for economic analysis is available from various sources, such data are very often useless for analysis from the point of view of their economic potential.

The purpose of this study is to outline a foundation for integrating Business Intelligence and Big Data into economic analysis processes. The theoretical and methodological basis of this study is provided by scientific research, methodological and practical developments of domestic and foreign authors on the application of IT solutions in economic analysis.

According to the results of the research, modern information technologies, in particular, the Business Intelligence and Big Data systems have considerably changed the possibilities for improving economic analysis and reducing decision-making time. From the methodological point of view, many aspects of integration of BI and Big Data solutions and their implementation in the economic analysis processes in Russia's companies remain insufficiently developed. The foreign market of modern information technologies for business analytics has a longer history and is being developed more rapidly.

The main conclusions of the study indicate that modern organizations operating on a highly competitive market should understand that the accumulation of Big Data does not always lead to the expected business benefits. In this context, the conclusion is that a modern company should not set as its goal to process all the available data in order to improve the quality of its economic analysis. It is more significant to use the entire volume of data for segmentation, which allows effective construction of a large number of models for small clusters, solving specific problems of economic analysis based on the application of modern IT systems.

Key words: economic analysis, modern information technologies, Business Intelligence, Big Data, Data Lake, business analytics, Big Data analytics.

Citation: Mitrovic S. (2017) Specifics of the integration of Business Intelligence and Big Data technologies in the processes of economic analysis. *Business Informatics*, no. 4 (42), pp. 40–46.
DOI: 10.17323/1998-0663.2017.4.40.46.

Introduction

In contrast to the Russian market, the foreign market of modern information technologies, in particular, of Business Intelligence (BI) systems and Big Data used to analyze economic information, has a longer history and is developing more rapidly. According to an IDC forecast (March 2017), worldwide profits from the Big Data and Business Analytics market will reach $150.8 billion by the end of 2017, which is 12.4% more than the previous year. It is expected that commercial purchases of the hardware, software, and services related to Business Analytics of Big Data will show an average Compound Annual Growth Rate (CAGR[1]) of 11.9% until 2020, when the profits will exceed $210 billion [1].

The largest volume of investments into Big Data business analytics technologies is expected in 2017 in the fields of banking, continuous production, federal (central) governmental authorities, as well as in the field of professional services. It is predicted that all these industries combined will spend a total of $72.4 billion by the end of 2017 for the development of Big Data tools. These industries will remain the largest investors into Big Data in the year 2020 as well, when their investments will total $101.5 billion. The highest growth rate of expenditures will be observed in the banking industry (CAGR of 13.3%), health services, insurance, operations with securities, investment management and telecommunications (each with CAGR of 12.8%) [2].

1. Application of Business Intelligence systems and Big Data for the improvement of economic analysis in Russia and worldwide

According to Gartner Research, Big Data business analytics makes it possible to control in-house information in modern banking structures better and to expose signs of fraud more promptly than was possible before, thereby ensuring financial safety. As follows from this study, economic analysis of Big Data helps us to "solve almost all key banking problems more efficiently, including client acquisition, improvement of services, clients' credit score evaluation, and so forth. Moreover, these technologies help banks to comply with the regulations issued by the authorities by increasing the speed and quality of report preparation, making the economic analysis of data deeper and broader, helping to prevent fraud and money laundering" [3].

BDA+BI technologies are mostly used in the banking industry to analyze client databases, including the clients' behavior, preferences and expectations. For example, PNC Bank (USA) performs a behavioral analysis of their clients' activities on different websites (primarily their online accounts in social networks), processes the information about the purchases made by clients and about their lifestyle, and uses these data to offer flexible interest rates. Another example is Commonwealth Bank of Australia (CBA), which analyses all transactions of their clients and complements this research with analysis of the clients' behavior in social networks. The combination and analysis of these data flows have allowed the bank to significantly reduce its credit default rate [4].

There is also positive experience in Russia. For instance, the Ural Bank for Reconstruction and Development launched a project for Big Data processing using BI technologies several years ago. This project used the client database to create personalized credit offers, investment tools, and other services. Within just one year of using the innovative IT solutions, the bank managed to increase its retail loan portfolio by more than 50% [4].

[1] Compound Annual Growth Rate (CAGR) is the average rate of growth of investments over a period greater than one year

Even though insurance companies are among the leaders of the segment and are interested in applying Big Data technologies and Business Intelligence to economic analysis, Bravura Solutions and Financial Services Council agencies conclude that only a small percentage of insurance companies compared to the banking industry has started to move in this direction [5]. According to a poll, 67% of insurance companies believe that they only have limited access to their clients' data. Respondents think that these data are sufficient to personalize the interaction with the clients, but they cannot be considered Big Data and be used to predict the clients' behavior. However, about 30% of foreign insurance companies that participated in the poll already use Big Data and BI technologies for economic analysis to predict their clients' needs and create personalized offers, whereas 23.7% of the companies did not even consider such innovational solutions yet [5].

Active development of the technologies of Big Data analysis in the retail sector – another leader of this segment – goes in different directions. Let us list the possibilities of practical application that can be considered as positive experience for Russian conditions. For instance, in worldwide practice, the brick-and-mortar retail industry uses the economic analysis of Big Data with BI technologies to analyze customers' behavior, plan their trips inside the store, group and place the goods according to customers' preferences, and plan acquisitions. As a result, sales increase. In the online retail sector, the economic analysis of Big Data – even an analysis using the most basic set of features of these technologies – can serve as the basis for the store algorithm itself: the customers receive suggestions based on their previous choices and personal preferences. This information can be obtained, for example, from social networks, prior purchases, and orders from online stores. In all cases, the intelligent economic analysis of Big Data helps to reduce overheads, increase the customers' loyalty, nurture their interest, and involve broader audiences [6].

Experts forecast that under present conditions, investments into analytical software for intelligent analysis of Big Data will continue to grow, reaching more than $70 billion in 2020. The main part of this amount will be spent on processing consumer surveys, tools of reporting and analysis, and means of data storage management [5]. It is worth noting that the world market demand for technologies and services in the field of Big Data and business analytics is stimulated by the increase in data volumes, development of new technologies, and a cultural shift towards making decisions related to information [2]. According to the data published by RAC, such a fast increase in the use of business analytics technology related to Big Data and the development of the new methods in this field can be explained by the fact that this segment achieved a new level of development compared to Business Intelligence technology. This new level allows the companies to use an extended range of applications (with built-in BI functions) and does not require the specialists responsible for economic analysis to have any additional specialized skills, whereas the similar segment of the CRM systems requires a different methodological approach and level of skills of the personnel [5].

Currently, the technologies of Big Data processing are reaching a higher level of development worldwide compared to the practices observed in Russia. According to the latest version of the "Cycle of technologies maturity" [7] published by Gartner, which estimates the level of methodology maturity and implementation, Big Data have already passed "the peak of excessive expectations, which they had in 2016, and moved into the stage of "disillusioning", i.e. the stage of active prototyping and development of the technology. According to expert estimates,

Big Data will reach their final stage of maturity in 5–10 years. Simultaneously, the analysis of data that combines all approaches to Big Data including Business Intelligence technologies has also reached its peak [7].

2. Prospects of interaction of Business Intelligence systems and Big Data in economic analysis

Even though the most investments into the technology of Big Data business analysis are made in the USA, other world players show higher activity in this area than before. According to IDC's data, 45% of the market of hardware, software, and services related to Big Data business analysis belong to European, Middle Eastern, African, and Asian (not counting Japanese) companies [1].

Demand is growing for analytical tools that can perform a complicated forecasting analysis in the field of Big Data. At the same time, the availability and range of standard solutions for Big Data business analysis are increasing. These solutions are already available from such major companies as Microsoft, IBM, Teradata, Oracle and SAP. This promotes a higher quality of products developed in this market.

78% of the companies around the world are certain that the business analysis of Big Data will be profitable and lead to various advantages within the next 1–3 years. Worldwide, 70% of major companies on the market not only use their Big Data for business analytics, but also purchase data from external sources (it is expected that this percentage will reach 100% by 2019) [1]. It is also expected that more and more companies will profit from their data by selling them or by offering some valuable content based on their data. We should note that for Russia, this is still only a distant perspective. However, there is a new type of offers – "data as a service" (from the Google search engine and others). For many companies worldwide, including Russian companies – and small businesses as well – this will make it easier to enter the Big Data market in the near future.

According to a forecast by IDC, investments into Big Data business analysis systems convenient for all users will become a must for all companies by 2018 [1]. It is also expected that the budgets for cheaper financial and organizational cloud solutions in the field of Big Data business analysis will grow three times as fast as those for localized solutions.

Based on the opinion of Gartner's experts, we can conclude that the Big Data and BI markets are at the last stage of a long-term transition from systematic reporting (IT approach) to the means of automated analysis (business approach). As a result, the modern platforms of business analytics and Big Data analysis (BI&BDA) will need to comply with new requirements for availability, flexibility and depth of analytic understanding [7]. This promotes the development of simpler, more flexible, and quicker solutions of business analytics, capable of expanding the business opportunities using a deep understanding of various data sources. This is why cloud business analytics and Big Data storage (for example, in the Amazon cloud) is one of the priority directions in the foreign technologies. This opportunity is important for many companies, especially for small businesses, and may be recommended as a promising trial or working version for Russia's companies. The advantage of this version is that the company does not need to spend a lot of money – which often is the reason why business analytics is not affordable for small businesses – buy or lease server space for the Big Data storage, considering that the company does not need this equipment on a permanent basis.

So-called streaming analytics, which makes real-time data analysis possible, is another promising tool for Big Data business analysis. We should note that unlike the Russian market, the world market tends to divide the field

of Big Data into multiple distinct areas (for example, storage, processing, data filtering, real-time on-demand models, data visualization, deep learning, etc.) that solve narrower specific problems of economic analysis. The development of ready-to-use solutions for small and medium businesses which work as independent applications or as SaaS or BDaS models (Big Data as Service) is closely connected with the aforementioned subdivision [8].

When we talk about the connection between the two innovation technologies (Big Data and BI) in economic analysis, we have to mention that they are not developing in parallel and are not subordinate to each other (as some experts believe). Since 2016, the development of the Big Data business analysis technology is closely integrated with the development of the Business Intelligence technology. This creates a two-level model of economic analysis. The first level is the "traditional" Big Data analytics, where huge volumes of data are processed in the offline mode (not in real time). A new, second level provides the possibility of economic analysis of relatively big volumes of data in real time thanks to the in-memory technologies which are included today in most BI solutions. For today's Russian business, this level opens up new possibilities as a consequence of economic analysis of the data accumulated by a company. This is connected, in particular, with the appearance and application of such phenomenon as "data lakes," which are the next evolutionary step to improve the business analytic processes "on the fly" and so influence events as they occur in real time. Let us briefly characterize this innovative phenomenon from the point of view of its application to economic analysis and give several methodological recommendations in this field for modern Russian companies.

A data lake is a repository that can store massive volumes of raw data. These data should be connected with the company's economic analytical system to perform economic analysis of the company's activities. Data lakes allow one to analyze various incomparable data in their original format. This is a huge advantage for today's organizations since no big expenses are necessary for storage and processing, and the data from the "lake" are accessible for business analytics purposes to any user within the company at any time [9].

Economic data are collected and stored in the "lake" because it is unknown beforehand how they could be used for business development. However, the very popular opinion shared by many organizations that "lakes" could replace traditional data storage facilities is erroneous. Lakes should provide organizations with new analytic possibilities and optimize the expenses for data storage and processing during economic analysis. In this context, many modern companies in the field of Big Data economic analysis face the question of whether their organizational activities and financial expenses for the creation of data lakes will be justified, especially if the company either already has or is creating a data storage for the BI technologies integration to process Big Data. According to Teradata researchers' estimates [9], data lakes can be used for more affordable data storage, and these technologies, when integrated, exert a stronger synergistic effect on economic analysis. The main distinction from the traditional storage is that a "lake" can grow almost indefinitely, keeping the expenses low, and can provide instant access to any data for all users. Moreover, while data in a traditional storage are structured, "lakes" provide cheap storage for all types of data, i.e. data from social networks or various devices, including data in audio and video formats [9].

At the same time, companies should be aware that it may be risky to accumulate "lakes" with the aim of improving the results of economic analysis within the company. For example, company employees tend to store any kind of data in the "lake" without any quality con-

trol. Data security and access control are rarely maintained. Data from the "lake" can be effectively used primarily by business analytics specialists, however, they can be accessed by any business users within the company. Lastly, the main danger is that these "lakes" – if they are not sorted and cleaned periodically – could be transformed with time into "swamps", making the data unusable even for the simplest processes of business analytics [10]. We have to note that the concept of Big Data assumes that no company data should be discarded, especially if they can be stored at low cost. Any data should be initially considered to be "smart data". If some data are discarded right away, there will be no way of knowing if they were useful for business purposes. At the same time, we have to keep in mind that it is rather difficult to determine the potential usefulness of such data. Hence, an organization does not necessarily have to create and keep such a "lake" on its own. Outsourcing is already a common practice, and different companies propose appropriate tools – in particular, the "cloud" tools mentioned above.

Conclusions

Analyzing the problem under consideration, we have to conclude that from the methodological point of view, many aspects of the solutions integration and implementation into the processes of economic analysis in the field of BI and Big Data are still not perfectly developed. One of the main problems is to find algorithms and methodological principles as well as factors and conditions that would allow modern companies to perform economic analysis of Big Data using the capabilities of Business Intelligence with maximum effectiveness for business, as well as the advantages that could give the company a competitive edge. As demonstrated by the analysis performed, the methodological aspects of this question are covered in theoretical and technical business literature much less than the aspects of storage, analytical processing of data, and Big Data management technology.

The main conclusion of this research is that the modern market is very competitive, and companies should understand that it is not necessarily profitable to collect and store Big Data. As with any project aimed at the optimization of business activity, when starting a project for applying Business Intelligence to process Big Data, companies should begin with solving the issue of data storage and systematization, as well as the pertinence of the data for certain business goals. The company data analysts need to ask themselves the following questions: "what should the ideal results look like?", "what is the measure of success?", and "what types of information are more useful for achieving an ideal result?". When answered, these questions may or may not lead to the conclusion that Big Data storage is necessary. As an alternative, a cloud solution may be chosen for the storage and processing of Big Data to help the company make strategic managerial decisions.

It is natural to draw another important conclusion: despite the concept of BI & Big Data, a modern company should not try to process all available data in the hope of increasing the quality of the results of economic analysis. It should be kept in mind that Big Data economic analysis that uses Business Intelligence and other modern IT tools is most effective when it is used to build predictive models. The precision of such models depends on the quality of the data sample. Hence, the real challenge of Big Data integrated with BI in economic analysis is not to utilize all available data to create predictive models, as this will not increase the models' accuracy. The more important thing is to use all available data to segment and cluster the data. This will provide an opportunity to effectively create many models for small clusters while solving practical problems of economic analysis based on modern information technologies. ■

References

1. Tadviser (2017) *Bol'shie dannye (Big Data): mirovoy rynok* [Big Data: world market]. Available at: http://www.tadviser.ru/index.php/Статья:Большие_данные_(Big_Data)_мировой_рынок (accessed 14 August 2017) (in Russian).
2. IDC (2015) *Double-digit growth forecast for the worldwide Big Data and business analytics market through 2020 led by banking and manufacturing investments, according to IDC*. Available at: http://www.idc.com/getdoc.jsp?containerId=prUS41826116 (accessed 14 July 2017).
3. CNews (2014) *Infografika: Chto mogut bol'shie dannye v bankakh?* [Infographics: What Bid Data can do in banks?]. Available at: http://www.cnews.ru/articles/infografika_chto_mogut_bolshie_dannye (accessed 14 July 2017) (in Russian).
4. CNews (2014) *Kak krupneyshie banki ispol'zuyut bol'shie dannye* [How the largest banks use Big Data]. Available at: http://www.cnews.ru/articles/kak_krupnejshie_banki_ispolzuyut_bolshie (accessed 14 July 2017) (in Russian).
5. Tadviser (2017) *Teoriya i praktika Bol'shikh dannykh v otraslyakh* [Theory and practice of Big Data in industries]. Available at: http://www.tadviser.ru/index.php/Статья:Теория_и_практика_Больших_данных_в_отраслях#cite_note-4 (accessed 18 August 2017) (in Russian).
6. CNews (2014) *Infografika: Kak roznitsa ispol'zuet bol'shie dannye* [Infographics: How retailers use Big Data]. Available at: http://www.cnews.ru/articles/infografika_kak_roznitsa_ispolzuet (accessed 13 July 2017) (in Russian).
7. Gartner (2016) *Gartner hype cycles 2016: Major trends and emerging technologies*. Available at: https://www.gartner.com/webinar/3377422 (accessed 18 August 2017).
8. CNews (2016) *Biznes-analitika i bol'shie dannye v Rossii 2016* [Business analytics and Big Data in Russia 2016]. Available at: http://www.cnews.ru/reviews/bi_bigdata_2016 (accessed 14 August 2017) (in Russian).
9. Teradata (2017) *Insights and tools from Big Data and analytics innovators*. Available at: http://bigdata.teradata.com/ (accessed 16 August 2017).
10. CNews (2016) *Kak razvernut' «ozero dannykh»: 5 sovetov* [How to rollout a data lake: 5 recommendations]. Available at: http://www.cnews.ru/articles/kak_razvernut_ozero_dannyh_5_sovetov (accessed 16 August 2017) (in Russian).

Using a web analytics system as the basis for integration with CPA services

Vyacheslav I. Zhukov
Assistant Professor, Department of Innovation and Business in Information Technologies
National Research University Higher School of Economics
Address: 20, Myasnitskaya Street, Moscow, 101000, Russian Federation
E-mail: vzhukov@hse.ru

Mikhail M. Komarov
Associate Professor, Department of Innovation and Business in Information Technologies
National Research University Higher School of Economics
Address: 20, Myasnitskaya Street, Moscow, 101000, Russian Federation
E-mail: mkomarov@hse.ru

Abstract

In modern e-business, there are many ways to attract potential customers to the site both with the help of offline and online methods. Companies usually use several channels to attract customers, differing in the placement of advertising, payment models and other parameters.

One of the most popular online methods is using CPA networks, which allow webmasters to place on their websites links to the advertised website and earn rewards for customers who purchased a service by clicking on the link. CPA networks work on the basis of payment for achieving targeted goals. A targeted goal can occur both online and offline. The most important task is to link the source of attraction (usually certain UTM tags) to the target goal of the client, since remuneration for the CPA network should only occur for orders from customers who are drawn to the CPA network. There are problems fixing operations, linking to source of attraction, storing and providing access to these data.

In this paper, we give a brief overview of various approaches to solving the problem: log analysis, use of marketing pixels and web analytics tools. We have analyzed the benefits and challenges of these methods, which were given to solve the task of fixing the target actions of clients and provided access to the data for the CPA network. Also in this article, we have described a practical case of integration with the CPA network based on the use of end-to-end web analytics. The advantages, disadvantages and limitations of the proposed method are set out in this paper.

Key words: web analytics, Internet marketing, CPA networks, end-to-end web analytics, Google Analytics, CRM.

Citation: Zhukov V.I., Komarov M.M. (2017) Using a web analytics system as the basis for integration with CPA services. *Business Informatics*, no. 4 (42), pp. 47–54.
DOI: 10.17323/1998-0663.2017.4.47.54.

Introduction

Running an online business usually requires dealing with the problem of landing target customers using the global network of Internet. This problem can be solved using both online and offline methods. The online methods suggest advertising in one form or another being placed on external sources linking to the company site. Banners, contextual advertising, display advertising, links in the articles and other approaches are used for that purpose.

To determine the payment for placement of advertisements on an external source, a few models are used: payment for the placement for a defined period, payment for the number of views (cost per view, CPV), or a frequently used scheme of payment for a thousand of impressions (cost per millennium, CPM), payment for clicks (cost per click, CPC), payment for the target actions (cost per action, CPA) [1]. A payment model for the landed customers as a special case of the CPA model also exists. Each of the models has its pros and cons: as for the CPC or CPV the result cannot be ensured in contrast to CPA providing the advertiser pay only for the landed customers who make the payments. Thus, using CPA allows us to significantly reduce the risk of potential financial losses [2]. However, it should be noted that the CPA model demands that significant technical improvements be made by the advertiser that are not required when using CPV or CPM because the website of the advertisement placement usually provides convenient ways of publishing it.

CPA networks operate based on the CPA model. The CPA networks are binding and medium agents between the site owners (often with low traffic) and the business that places the ads on these websites. Webmasters publish the ads on their sites. Users click on these ads, and the webmasters get the payments for the landed customers in case the users make the target actions. The type of action and its price are defined by the advertiser and they are named and offer. The CPA network is in fact an intermediate link between the huge amount of the webmasters, their sites and the companies which want to attract the audience from these sites to their own websites for the purpose of selling goods and services.

Usually the advertiser uses a couple of channels of customers' acquisition in the Internet. Special parameters (codes) should be added to the link to determine the acquisition source. Usually they are the UTM codes [3]. The special identifiers assigned to customers are often used for the referral programs [4].

Before doing the targeted action, the customer client can enter the site via several channels (for example, from contextual ads and then from the search results). To consider this fact during the process of analyzing acquisition source effectiveness, you should define the attribution as a rule of assigning the target action to the acquisition source or the rule of distributing value by the conversion value between the sources. The attribution of the last relevant click [5] is usually used for the CPA networks without taking free traffic sources into account (direct traffic, unpaid search traffic) i.e. the last payment click attribution.

Insofar as the payment is based on the number and types of the target actions, there appears to be a problem of recording these actions on the side of the advertiser and passing firm data about them to the CPA network in order to measure the payment for the webmasters.

1. Problem-solving approaches

A few methods of recording the target actions of the users on the website exist: who got in the site from a certain channel to provide access to them for the CPA network: the web server log analysis service, web beacons activation and using the web analysis record-keeping systems (trackers).

1.1. Web server log analysis service

In this case, all the "mechanics" of the attribution and attaining the target actions are developed on the server as the log analysis service [6] that records all the references to the web pages; submitting the transaction forms and other activities of the users. The referral and the UTM codes used for the user to get into the site are recorded in particular. It also records the customer's IP-address, his identifier and other service information (browser and its version, operating system, time of the reference etc.) [7]. These logs are usually saved in one of the following formats [8]:

- ✦ NCSA Common Log;
- ✦ NCSA Combined Log;
- ✦ NCSA Separate Log;
- ✦ W3C Extended Log.

Further, the service for building the reports on this data including the attribution model logic is developed. If necessary, this data is combined with the other data sources of the company, for example, with the CRM-systems. The data is linked based on the client's identifier that should be the same for the same customer in all the systems. To integrate with the CPA network, a special API is carried out to provide access to the information about the payments of the customers landed from the CPA network (defined based on the UTM codes).

The ability to get absolute accuracy of the calculations and carry out a larger amount of assorted target actions can be seen as the advantage of this approach. In addition, the possibility exists to record the actions done beyond the website, for example, to identify users who paid for the service or the product in cash, using payment terminals or by bank transfer due to the ability of integration with other company's internal data sources.

We should consider the substantial expenses for the development of this service (or the deployment of a ready-made solution) and its further support as the disadvantage. It is worth mentioning the issue of confidence, since the business can provide false information for the CPA network and it is quite difficult to check this.

1.2. Web beacons

This approach is based on the activation of external pixels when certain events take place on the website [9, 10]. Special pixels of the CPA network are placed on the company's site. They are activated if certain target actions have been done. They usually are gif 1x1 px. [11], located on the CPA network's sever. The pixel's activation activates passing the parameters and the cookies of the user. These pixels are activated only for the users attributed based on the UTM codes as those landed from the CPA network. Thus, CPA networks get information about attaining the target actions. Usually this solution is used for relatively simple target actions that take place on the site (viewing articles, clicking certain buttons or submitting a request).

The advantages of this approach are transparency (as the logics of pixel activation can be checked by the CPA network) and unsophisticated deployment in the customer's site.

The disadvantage is poor scalability of the approach: when increasing the number of recorded target actions each of them should be separately marked with a separate pixel with other parameters of activation. It is worth mentioning the problems related to the modernization of the website's interface: it is necessary to keep the domain logic of activating these pixels. There is a problem of the complexity of recording events that can take place beyond the website, for example, when the client pays for the service. Another disadvantage is the partial accuracy of data due to losses related to JavaScript turned off at the user's browser and ad blocking tools.

1.3. Using the goals in the web analytics systems

In this case, we use the goals that are already customized in the web analytics systems (for example, Google Analytics[1], Yandex.Metrika[2], Piwik[3], etc.) and the abilities of these systems to set the conversions attributes by the sources. The CPA network specialists have access to the web analytics trackers with filtering on the traffic sources corresponding to the CPA network (usually it is a specific code in the UTM parameters).

For specialists, there is CPA network access to the trackers of web analytics. Based on this data, reports are generated and the amount of remuneration is determined. The reports are built and the payment value is determined based on this data.

Transparency is also an advantage of this method, since the CPA network always can check the consistency of collecting the goals and attributions. We should also consider the relative cheapness of integration: usually the counters are already installed on the client's site and the goals are configured and are used for analysis tasks within the company.

Loss of data is the disadvantage. Its reasons coincide with the "web beacons" approach.

2. Case description

2.1. Task assignment

Company "A" works in the field of domain sales and providing hosting services. It decided to use a CPA network to increase sales. To solve the assigned task, the "B" CPA network was chosen. Company "A" specified the "paid order" to be the target action. However, since the company sells a wide range of different services with different marginality, the payment ("offer") for the purchase by the customer led by the CPA network, should also vary depending on the type of service. It is also important that a huge amount of payments be done not via the website but via the payment terminals, by bank transfers etc.

The CPA network demanded availability of the API for access to the data on service payments of the users led from this CPA network with passing information about the category of the purchased service; the attribution model – the last payment click. Company "A" did not have such a service and its development would require extensive resources. That brought to nought the beneficial effect of using the CPA network.

Previously, company "A" implemented the end-to-end web analytics approach [3] to carry out more accurate analysis of advertising campaign efficiency and deep analysis of the customer's behavior from the moment of client acquisition to the site until the moment of the real payment, based on the integration of Google Analytics with the internal self-developed billing solution via the Measurement Protocol (MP API) [12, 13]. That allowed them to record the target actions of the users (payments) done beyond the website.

To carry out technical integration into the CPA network with a view to economy of resources, the decision was made to use the existing system of end-to-end analytics but improving it.

2.2. Description of the integration of Google Analytics with the billing

When the customer enters into a purchase order, the Client ID (cid) is transmitted from the Google Analytics system (used by Google Analytics to identify the individual user of the website) as the parameter, and when the customer pays for the order, a special pur-

[1] https://www.google.com/analytics/
[2] https://metrika.yandex.ru/
[3] https://piwik.org/

pose request – the Enhanced Ecommerce[4] request – is sent from the billing to Google Analytics with binding to this customer. This data multiplexing allows tracking the customer's full path from getting into the site to the real payment for a specific service and the profit of selling this service.

The payment hit includes the following information:

- ✦ the number of the customer's order;
- ✦ SKU of the service;
- ✦ category of the service;
- ✦ name of the service;
- ✦ revenue;
- ✦ profit;
- ✦ clientID.

2.3. "Last payment click" emulation

To complete the tasks of the regular web analytics in the company, the Google Analytics system was customized according to the standard attribution model – the last non-direct click. However, it doesn't match up for the integration with the CPA network since, besides the paid traffic, the unpaid traffic (mostly organic search and referral traffic) can dominate and "erase" the paid acquisition source.

This would take place if the user followed the paid source and then, for example, followed the results of the organic search and bought the goods. At that time, he searched for the name of the company.

To implement the "last payment click" model for the integration with the CPA network, there was an additional Javascript assembly unit developed on the front-end side of the website that emulated the "last payment click" logics via the specific user parameter of Google Analytics set on the user level (saved during the whole life time of the clientID). The UTM codes were recorded to these parameters and sent to Google Analytics. The information was available for analysis and building the reports; at the same time the processes and systems already used in the company to analyze customer behavior on the site were not damaged.

2.4. Data access API

The mentioned organization allowed making Google Analytics a general store of the data on the full track of the client's interaction with the site: from the acquisition source until the moment of the real payment. At the same time, since Google Analytics already has the instruments of getting the data via API, there was no necessity to develop the service for this data access.

The individual Google Analytics view was customized for the CPA network. It included the CPA networks of the traffic source that were transmitted via the UTM codes. The information was available through Google Analytics Core API [14].

2.5. Interaction patterns and stages

The integration diagram with the CPA network shown on *figure 1* provides for carrying out the following steps.

Step 1. The customers get into the client's site from the webmasters' resources via the URL with special UTM parameters that contain the information about the network and the webmaster identifier.

Step 2. When the site is being loaded, the information about the UTM codes in the specific parameters on the user level is also transmitted to Google Analytics besides the standard information about the page view (cid etc.).

Step 3. When ordering the service on the site,

[4] https://developers.google.com/

the cid (Client ID, client identifier in Google Analytics service) is sent to billing.

Step 4. After the service payment, billing transmits the hit with the contents of the paid services and the cid attached to the order to Google Analytics via MP. After receiving the payment data, Google Analytics relates it with the specified user (based on cid) whose visit information was recorded at step 2.

Step 5. Using Google Analytics Core API, the CPA network uploads the information about the payments of the services and their categories carried out by the customers who got into the site of company "A" from the resources of the webmasters who work with this CPA network. The subsequent cost clearings between the company and the CPA network are carried out based on this information.

3. Advantages and restrictions of the suggested method

The described method has a variety of advantages and restrictions that should be taken into account when solving the problem of integration with the CPA network.

One of the main advantages of the method is the fact that a small improvement of the infrastructure developed for the analytical tasks allows us to solve the integration problem. This method would allow the following things for companies that use end-to-end analytics in their work:

✦ significantly reduce the time and costs for solving the problem of integration with the CPA network;

✦ to test this channel of solving the market-

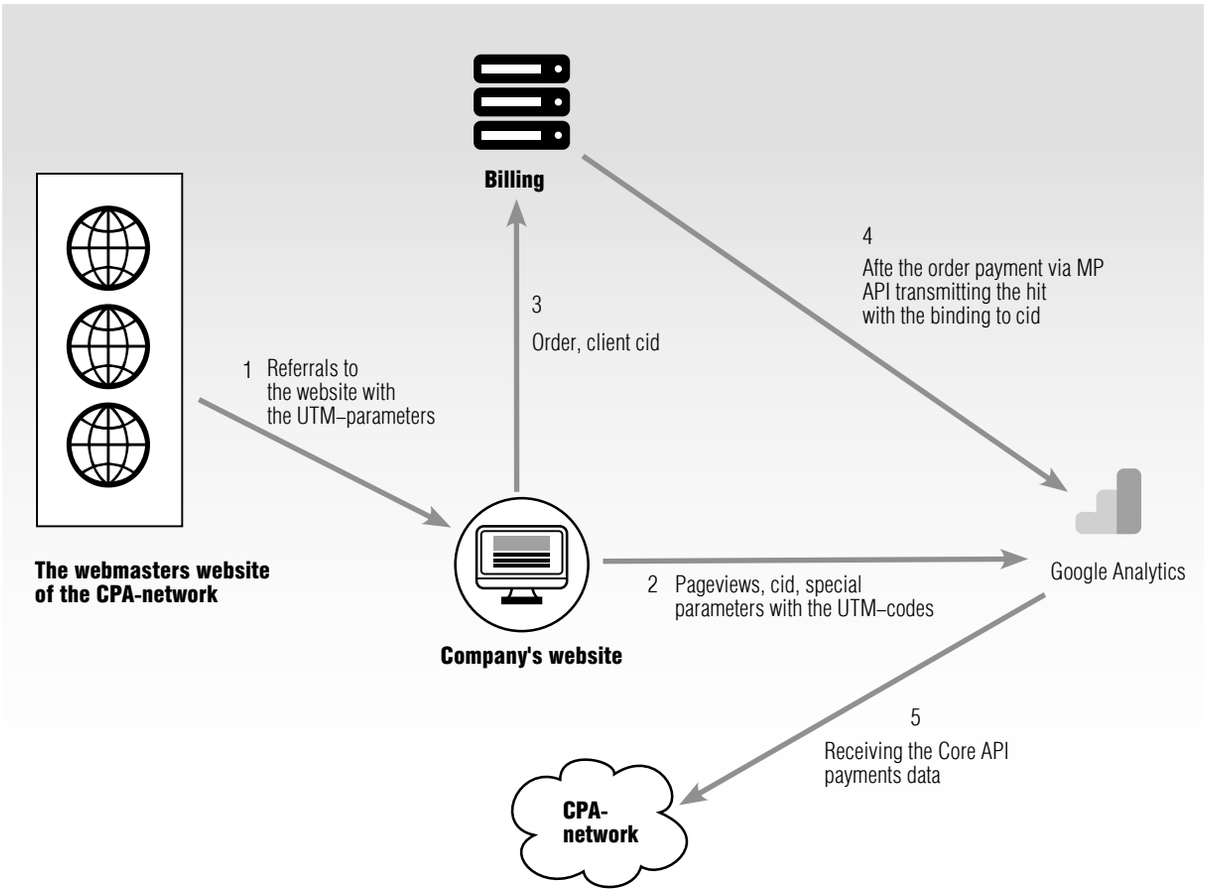

Fig. 1. Integration solution with the CPA network

ing problems of increasing the customer base, growth of profits or launching a new service as quickly as possible;

✦ quickly stop working with a promotion channel in case of its inefficiency;

✦ the ability to take service categories into account when calculating payment for the webmasters.

At the same time, the given method has a couple of technical restrictions that cause partial loss of data.

First of all, Google Analytics can get under the operation of the anonymizers and ad blockers for example, AdBlock[5], that can cause the loss of payment data for the customers attracted from the CPA network who use these kinds of blockers. This problem can be completely solved only by using the previously described approach of log analysis, but due to the high cost of its development and support the common method is to simply increase the webmasters' reward by a certain percentage. This method is more profitable than log analysis system implementation and support for small volumes of the orders.

Secondly, some customers can get to know about the services (be attracted to the site) using one device (e.g., a cellphone) and buy the service using another device. This situation can also cause the loss of payment data in this circuit. To solve this problem, you can use one of the methods to identify the customer from various devices [13] by additionally passing the userID from billing or CRM to Google Analytics. This solution would be successful if the user logged in to the site on each of the devices.

Conclusion

As far as the information about the real service payments of the site's users was bound with their information in the Google Analytics system and this information was available to be received by the CPA network, the integration task was solved. However, the transmitting of the goods categories and their names allowed estimation of the various payments for the webmasters depending on the service paid by the attracted customer.

Due to the built-in functionality of Google Analytics (which includes integration with other company sources of information and the possibility to access this information using API), it became possible to solve two business tasks at once: to bind the visitor's clickstream on the site to the payment data for deeper and more accurate analysis of the customer's behavior (end-to-end web analytics) and provide integration with the CPA network.

This method expands the approach to carry out the integration with the CPA network based on systems of web analytics, upgrading it by the method of end-to-end web analytics. The method allows one to record the actions of the customer beyond the website and at the same time to bind these actions with the visitor's clickstream and the acquisition source. This approach can also be used to record other types of target actions that take place offline, preceded by visiting the website. ■

References

1. Rzemieniak M. (2015) Measuring the effectiveness of online advertising campaigns in the aspect of e-entrepreneurship. *Procedia Computer Science*, no. 65, pp. 980–987.
2. Hu Y., Shin J., Tang Z. (2012) *Performance-based pricing models in online advertising: Cost per click versus cost per action*. Atlanta: Georgia Institute, 2012.

[5] https://adblockplus.org/ru/android

3. Korobkov S.A. (2016) Osnovnye komponenty sistemy skvoznoy marketingovoy Internet-analitiki dlya malogo biznesa [Main components of a system of end-to-end marketing Internet-analytics for small business]. *International Scientific and Research Journal*, no. № 11–1 (53), pp. 48–50 (in Russian).

4. Brown B.C. (2009) *The complete guide to affiliate marketing on the web: How to use and profit from affiliate marketing programs.* Ocala, Florida: Atlantic Publishing Company.

5. Berman R. (2015) *Beyond the last touch: Attribution in online advertising.* Available at: http://ron-berman.com/papers/attribution.pdf (accessed 01 December 2017).

6. Grace L.K.J., Maheswari V., Nagamalai D. (2011) Analysis of web logs and web user in web mining. *International Journal of Network Security and its Applications*, vol. 3, no. 1, pp. 99–110.

7. Goel N., Jha C.K. (2013) Analyzing users behavior from web access logs using automated log analyzer tool. *International Journal of Computer Applications*, vol. 62, no. 2, pp. 29–33.

8. Booth D., Jansen B.J. (2009) A review of methodologies for analyzing websites. *Handbook of research on web log analysis* (eds. B.J. Jansen, A. Spink, I. Taksa). Hershey, PA: IGI Global, pp. 143–164.

9. D., Kaushik A. (2009) Web Analytics 2.0: Empowering customer centricity. *Search Engine Marketing Journal*, vol. 2, no. 1, pp. 5–11. Available at: https://pdfs.semanticscholar.org/f804/b8c0f66b28220d-64060e27892dbe0a7a3baa.pdf (accessed 01 December 2017).

10. Mayer J.R., Mitchell J.C. (2012) Third-party web tracking: Policy and technology. Proceedings of *2012 IEEE Symposium on Security and Privacy (SP 2012). San Francisco, California, 20–23 May 2012*, pp. 413–427.

11. Dwyer C. (2009) Behavioral targeting: A case study of consumer tracking on Levis.com. Proceedings of the *Fifteenth Americas Conference on Information Systems (AMCIS 2009). San Francisco, California, 6–9 August 2009*, pp. 1–10.

12. Waisberg D. (2015) *Google Analytics integrations.* John Wiley & Sons.

13. Weber J. (2015) *Practical Google Analytics and Google Tag Manager for developers.* Apress.

14. Clifton B. (2012) *Advanced web metrics with Google Analytics.* John Wiley & Sons.

Estimating the net premium using additional information about a quantile of the cumulative distribution function

Zhanna N. Zenkova
Associate Professor, Institute of Applied Mathematics and Computer Science
National Research Tomsk State University;
SRO "Association of professional actuaries"
Address: 36, Lenin Avenue, Tomsk, 634050, Russian Federation
E-mail: zhanna.zenkova@mail.tsu.ru

Elizaveta A. Krainova
Doctoral Student, Institute of Applied Mathematics and Computer Science
National Research Tomsk State University
Address: 36, Lenin Avenue, Tomsk, 634050, Russian Federation
E-mail: lanshakoval@gmail.com

Abstract

In this paper, the task of increasing the accuracy of net premium estimations in non-life insurance is considered. Improvements are achieved by involving additional information about a known quantile of loss cumulative distribution function. The additional information is used by projection the empirical cumulative distribution function onto the class of cumulative distribution functions with a certain quantile, and then the modified empirical cumulative distribution function is substituted into the integral that yields the mean value. This allows us to obtain a modified estimation of mean value using additional information about the quantile which is unbiased and its variance is asymptotically less than the variance of the classical sample mean, so that the mean-square error of the modification is also smaller. Therefore, the modified estimation is more accurate than the classical one for a large sample size.

The influence of a quantile value on the variance of the new estimation is studied for uniform, triangular and normal distributions. It is suggested that the minimum of the variance is reached when a known quantile is equal to the median (symmetry center) for symmetrical distribution. Based on Simpson triangular distribution, it was shown that for cases of skewed distributions involving the quantile allows one to decrease the variance more significantly than for symmetrical ones.

The modified estimation of mean value is applied to a real data set for calculation of a net premium. The data contain information about payments for voluntary health insurance of some insurance company. It is demonstrated that the classical method underestimates the net premium, and so it could lead to the company's bankruptcy. After applying the new modified technique, the net premium becomes higher and the bankruptcy risk is reduced as well.

This paper contains practically significant results which make it possible to give important recommendations to an insurance company.

Key words: net premium, sample mean, additional information, cumulative distribution function quantile, modified estimation of mean value, accuracy of estimation, mean-square error, non-life insurance.

Citation: Zenkova Z.N., Krainova E.A. (2017) Estimating the net premium using additional information about a quantile of the cumulative distribution function. *Business Informatics*, no. 4 (42), pp. 55–63. DOI: 10.17323/1998-0663.2017.4.55.63.

Introduction

The insurance business is actively employing the latest scientific developments, especially in the area of modern statistical methods and models. This paper considers the cases of non-life insurance [1] in which the insurance indemnity is only paid upon the occurrence of an insured event, which are essentially non-deterministic. Because of their random character, the insurance companies use the so-called net premium [2] as the foundation for the calculation of insurance premiums. The net premium depends on the mean loss amount and the probability that the insured event will occur. Then the net premium is adjusted by a coefficient greater than one, which is calculated based on the probability of the companyr's bankruptcy set by an actuary. The premium is further incremented so that the company could make a profit and pay the agents' commissions.

To optimize their business, insurance companies always aim at setting policy premiums that, on the one hand, would be attractive to clients, and, on the other hand, accurately account for the random character of the insured events as accurately as possible. For these purposes, different statistical methods and models are used, including methods that involve various additional types of information [3–14].

This paper proposes a new method of improvement of the net premium estimation by considering additional information about the quantile of the loss cumulative distribution function. This results in a more accurate modified estimation of the expected value of the insurer's payouts per insured event. The new method of calculations was tested using real-life data on voluntary health insurance and allowed us to re-evaluate the policy premiums, thus reducing the risk of bankruptcy of the insurance company.

This paper should be of a considerable practical interest because this approach to the calculation of the net premiums has not been used before.

1. Estimating the mean insurance indemnity payment using additional information about the quantile of the cumulative distribution function

The net premium p plays the key role in the calculation of the insurance premium rates in non-life insurance [1], and it is defined as follows [2]:

$$p = z \cdot EX, \quad (1)$$

where X – the payout amount, a random variable with the cumulative distribution function $F(x) = P(X < x)$;

EX – the expected value of the random variable;

z – the probability of occurrence of the insured event.

In insurance practice, neither z nor EX are known exactly. They must be estimated based on the available data on payouts. Very often, the researcher possesses some additional information about the random variable under consideration and its distribution. This information could come from the conditions of an experiment, the researcher's professional experience, etc. It is well known that the use of additional information positively affects the properties of

the modified statistics by increasing the accuracy of the estimation [3–13].

This paper considers the additional information about the known quantile of the cumulative distribution function x_q of the given level q, i.e., we know that

$$F(x_q) = q. \quad (2)$$

Using (2), we obtain the modified estimation of the expected value by substituting [15] the modified empirical distribution function into the functional for EX:

$$\bar{X}^q = \int_{-\infty}^{+\infty} x dF_N^q(x).$$

The modified empirical distribution function $F_N^q(x)$ is obtained as the projection of the ordinary empirical distribution function onto the prior class [9] by the formula:

$$F_N^q(x) = \begin{cases} q \cdot \left(\dfrac{F_N(x)}{F_N(x_q)} \wedge 1 \right) + \\ + (1-q) \cdot \left(\left(\left(\dfrac{F_N(x) - F_N(x_q)}{1 - F_N(x_q)} \right) \vee 0 \right) \wedge 1 \right), \\ if \quad F_N(x_q) \in (0;1); \\ F_N(x), \ if \ F_N(x_q) = 0 \quad or \ F_N(x_q) = 1, \end{cases} \quad (3)$$

where

$$F_N(x) = \frac{1}{N} \sum_{i=1}^{N} H(x - X_{(i)}) \quad (4)$$

is the ordinary empirical distribution function, symbols $\wedge$ and $\vee$ denote the maximum (minimum) of the two values, $H(y) = \{0 : y \le 0; 1 : y > 0\}$ is a Heaviside step function, $X_{(1)} \le X_{(2)} \le ... \le X_{(N)}$ are the order statistics of the independent sample $\{X_1, X_2, ..., X_N\}$, N is a sample size.

Then

$$\bar{X}^q = \frac{q}{N \cdot F_N^q(x_q)} \cdot \sum_{i=1}^{N} X_i \cdot I_{(X_i < x_q)} +$$

$$+ \frac{(1-q)}{N(1 - F_N^q(x_q))} \cdot \sum_{i=1}^{N} X_i \cdot I_{(X_i \ge x_q)} =$$

$$= \frac{q \cdot N}{N \cdot r} \cdot \sum_{i=1}^{r} X_{(i)} + \frac{N \cdot (1-q)}{N \cdot (N-r)} \cdot \sum_{i=r+1}^{N} X_{(i)} =$$

$$= \frac{q}{r} \cdot \sum_{i=1}^{r} X_{(i)} + \frac{1-q}{N-r} \cdot \sum_{i=r+1}^{N} X_{(i)}.$$

Here the random variable r is defined through

$$F_N(x_q) = \frac{r}{N}$$

and it has the binomial distribution $Bi(N, g)$ [15].

After the modification using the known quantile, the estimated expected value is given by the equation

$$\bar{X}^q = \begin{cases} \dfrac{q}{r} \cdot \sum_{i=1}^{N} X_i \cdot I_{(X_i < x_q)} + \dfrac{1-q}{N-r} \cdot \sum_{i=1}^{N} X_i \cdot I_{(X_i \ge x_q)}, \\ \qquad\qquad if \ \ r = \overline{1, N-1}; \\ \bar{X}, \qquad\qquad if \ \ r = 0 \quad or \ N, \end{cases} \quad (5)$$

where

$$\bar{X} = \frac{1}{N} \cdot \sum_{i=1}^{N} X_i - \quad (6)$$

is the classical sample mean. Additionally, for $r = \overline{1, N-1}$ the equation (5) can be defined using the order statistics:

$$\bar{X}^q = \frac{q}{r} \cdot \sum_{i=1}^{r} X_{(i)} + \frac{1-q}{N-r} \cdot \sum_{i=r+1}^{N} X_{(i)}.$$

Let us note that the equation (5) can be considered as a generalization of the asymptotically normal estimation proposed in [8], which dealt with a more general case of the functional estimation using several exactly known expected values; however, the situations when a modification cannot be defined were not taken into account. It is worth noting that the aforementioned estimation was obtained not by substitution, but by projection onto the prior class using the Kullback-Leibler divergence.

Let us find the expected value of the estimate (5):

$$E\bar{X}^q = E\left\{E\left(\bar{X}^q \mid r\right)\right\} = E\left\{E\left\{\frac{q}{r}\cdot\sum_{i=1}^{N} X_i\cdot I_{(X_i<x_q)} + \right.\right.$$

$$\left.\left. + \frac{1-q}{N-r}\cdot\sum_{i=1}^{N} X_i\cdot I_{(X_i\ge x_q)} \mid r=\overline{1,N-1}\right\}\right\} +$$

$$+E\left\{E\left\{\frac{1}{N}\cdot\sum_{i=1}^{N} X_i \mid r=0 \;\; or N\right\}\right\} = \left(1-q^N-(1-q)^N\right)\cdot$$

$$\cdot E\left\{\frac{q}{r}\cdot r\cdot\int_{-\infty}^{x_q} xd\left(\frac{F(x)}{q}\right) + \frac{1-q}{N-r}\cdot(N-r)\cdot\int_{x_q}^{+\infty} xd\left(\frac{F(x)-q}{1-q}\right) \mid r = \right.$$

$$\left. = \overline{1,N-1}\right\} + EX\cdot\left(q^N + (1-q)^N\right) = EX.$$

Therefore, the modified estimation of the expected value (5) is unbiased.

Let us find the variance of the estimation (5) using the equation

$$D\bar{X}^q = D\left(E\left(\bar{X}^q \mid r\right)\right) + E\left(D\left(\bar{X}^q \mid r\right)\right),$$

where $D\left\{E\left(\bar{X}^q \mid r\right)\right\} = D\{EX\} = 0.$

Using independence of X_i, $i = \overline{1,N}$, we obtain

$$D\left\{\frac{q}{r}\cdot\sum_{i=1}^{N} X_i I_{(X_i<x_q)} + \right.$$

$$\left. + \frac{1-q}{N-r}\cdot\sum_{i=1}^{N} X_i I_{(X_i\ge x_q)} \mid r=\overline{1,N-1}\right\} =$$

$$\frac{q^2}{r} D\left\{X_i I_{(X_i<x_q)}\right\} + \frac{(1-q)^2}{N-r} D\left\{X_i I_{(X_i\ge x_q)}\right\}.$$

According to the properties of expectation, $E\left(\frac{1}{r}\right) \le \frac{1}{Er}$ because function $f(r) = \frac{1}{r}$, is convex. Convergence of $E\left(\frac{1}{r}\middle| r>0\right) \xrightarrow[N\to\infty]{} \frac{1}{E\left(r\middle| r>0\right)}$ was proved numerically for $N = \overline{10,100}$. *Figure 1* shows the relative deviation

$$\Delta = \frac{\left|E\left(\frac{1}{r}\middle| r>0\right) - \frac{1}{E\left(r\middle| r>0\right)}\right|}{E\left(\frac{1}{r}\middle| r>0\right)}\cdot 100\%$$

as a function of the sample size N for q = 0.2, 0.5 and 0.8.

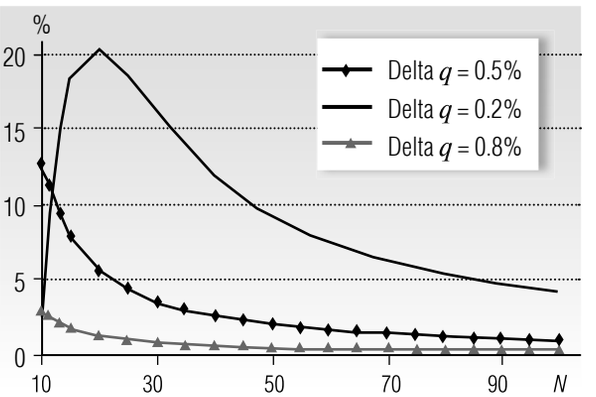

Fig. 1. Deviation Δ for q = 0.2, 0.5, 0.8

Since $q \in (0,1)$, then

$$P\left(r=\overline{1,N-1}\right) = 1-q^N-(1-q)^N \xrightarrow[N\to\infty]{} 1,$$

$$P\left(r=0\right) = q^N \xrightarrow[N\to\infty]{} 0,\; P\left(r=N\right) = \left(1-q\right)^N \xrightarrow[N\to\infty]{} 0.$$

To obtain the asymptotic value of the variance, it is sufficient to consider the case $r = \overline{1,N-1}$, then

$$\sigma_q^2 = \lim_{N\to\infty} ND\bar{X}^q = \frac{q^2}{q} D\left\{X_i\cdot I_{(X_i<x_q)}\right\} +$$

$$+\frac{(1-q)^2}{(1-q)} D\left\{X_i\cdot I_{(X_i\ge x_q)}\right\} = q\cdot D\left\{X_i\cdot I_{(X_i<x_q)}\right\} +$$

$$+(1-q)\cdot D\left\{X_i\cdot I_{(X_i\ge x_q)}\right\} = q\cdot\int_{-\infty}^{x_q} x^2 d\,\frac{F(x)}{q} - q\cdot\left(\int_{-\infty}^{x_q} xd\,\frac{F(x)}{q}\right)^2 +$$

$$+(1-q)\cdot\int_{x_q}^{+\infty} x^2 d\left(\frac{F(x)-q}{1-q}\right) - (1-q)\cdot\left(\int_{x_q}^{+\infty} xd\left(\frac{F(x)-q}{1-q}\right)\right)^2 =$$

$$= \int_{-\infty}^{x_q} x^2 dF(x) - \frac{1}{q}\cdot L^2 + \int_{x_q}^{+\infty} x^2 dF(x) - \frac{1}{1-q}\cdot R^2 =$$

$$= DX + \left(\int_{-\infty}^{+\infty} xdF(x)\right)^2 - \frac{1}{q}\cdot L^2 - \frac{1}{1-q}\cdot R^2.$$

Here $L = \int_{-\infty}^{x_q} xdF(x),\quad R = \int_{x_q}^{+\infty} xdF(x).$

Therefore, the asymptotic variance of the modified estimation of the expected value, normalized to the sample size, is given by the following equation:

$$\sigma_q^2 = DX + \left(EX\right)^2 - \frac{1}{q}\cdot\left(\int_{-\infty}^{x_q} xdF(x)\right)^2 - \frac{1}{1-q}\cdot\left(\int_{x_q}^{+\infty} xdF(x)\right)^2. \quad (7)$$

Since $(EX)^2 = L^2 + 2 \cdot L \cdot R + R^2$, then

$$(EX)^2 - \frac{1}{q} \cdot L^2 - \frac{1}{1-q} \cdot R^2 = -\frac{1-q}{q} \cdot L^2 +$$
$$+ 2 \cdot \sqrt{\frac{1-q}{q}} \cdot \sqrt{\frac{q}{1-q}} \cdot L \cdot R - \frac{q}{1-q} \cdot R^2 =$$
$$= -\left(\sqrt{\frac{1-q}{q}} \cdot L - \sqrt{\frac{q}{1-q}} \cdot R \right)^2 \le 0,$$

which means that the equation (7) can be rewritten as follows:

$$\sigma_q^2 = DX - \left(\sqrt{\frac{1-q}{q}} \cdot \int_{-\infty}^{x_q} x dF(x) - \sqrt{\frac{q}{1-q}} \cdot \int_{x_q}^{+\infty} x dF(x) \right)^2,$$

or, equivalently [8],

$$\sigma_q^2 = \sigma^2 - \frac{1}{q(1-q)} \cdot \left(\int_{-\infty}^{x_q} x dF(x) - q \cdot EX \right)^2, \quad (8)$$

where $\sigma^2 = DX = NDX$.

For the case when the known quantile is the median, we have:

$$\sigma_{0,5}^2 = \sigma^2 - \left(\int_{-\infty}^{x_{0,5}} x dF(x) - \frac{EX}{2} \right)^2.$$

From (8) it is obvious, that $\sigma_q^2 \le \sigma^2$, i.e., when the number of observations is large enough, the use of the additional information about the quantile may reduce the mean-square error, and, consequently, increase the accuracy of the estimate of the expected value.

For the uniform distribution $U_{[0,1]}(x)$, the equation (8) takes the following form:

$$\sigma_q^2 = \frac{1}{12} - \frac{q(1-q)}{4}. \quad (9)$$

It is obvious that the minimum of the variance is reached at $q = 0.5$. *Figure 2* shows the dependence given by the equation (9).

Figure 3 shows σ_q^2 vs q for the normal distribution function $N(10,4)$.

Figure 4 shows the second term of the equation (8), namely

$$d = -\frac{1}{q(1-q)} \cdot \left(\int_{-\infty}^{x_q} x dF(x) - q \cdot EX \right)^2 \quad (10)$$

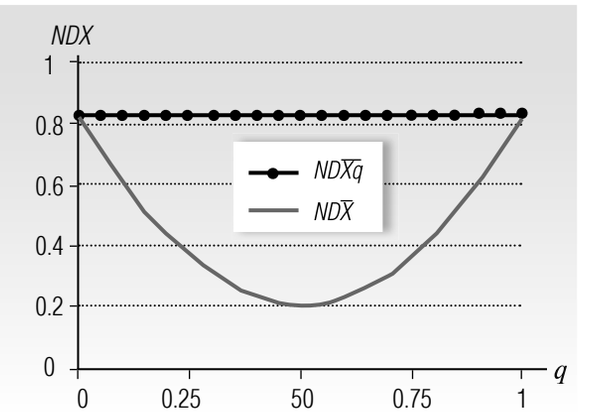

Fig. 2. σ_q^2 vs. q for $U_{[0,1]}(x)$, $\sigma^2 = \frac{1}{12}$

for different values of c for Simpson's triangular distribution over the interval [0,1] where $c \in (0,1)$. The value of d shows how much smaller the final variance σ_q^2 is compared to the variance of the original estimate of the mean $\sigma^2 = NDX$. Note that we did not obtain the best results for the symmetrical case (when $c = 0.5$). The consideration of the quantile turned out to be more important in the cases of a large skewness ($c = 0.9$ and $c = 0.1$). When the skewness is positive ($c < 0.5$), the accuracy increases if we use the information about $q > 0.5$, and vice versa (see *Figure 5* and *Table 1*).

Figure 5 shows the combination of the parameter c and the quantile q that allows for the maximum improvement in the quality of the estimation of the modified mean, meaning that d reaches its minimum value. *Table 1*

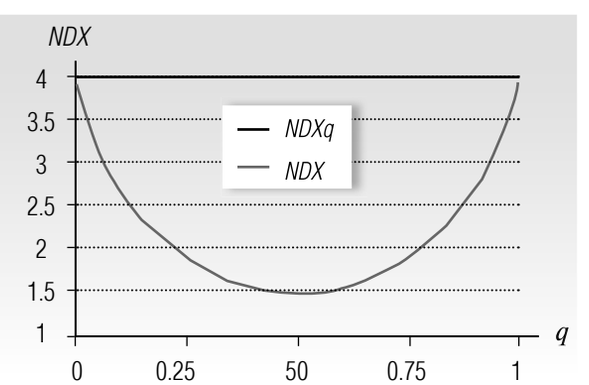

Fig. 3. σ_q^2 vs. q for the normal distribution function $N(10,4)$, $\sigma^2 = 4$

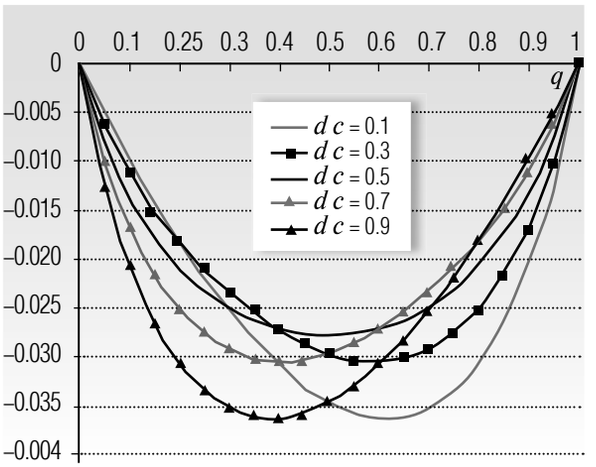

Fig. 4. The dependence of d vs. q for the triangular distribution over [0,1] with the parameter $c \in (0,1)$

shows the numerical values for asymptotic normalized variances σ_q^2 and the minimum values of d for these combinations of c and q.

Therefore, for all distributions we considered, the account of the additional information leads to a significantly smaller variance of the estimations. For symmetric distributions, the variance reached the minimum when the quantile coincided with the median, i.e. with the center of symmetry. Skewed distributions require additional studies; however, judging by the example of the triangular distribution, we can assume that, for a positively skewed distribution, a more accurate estimation of

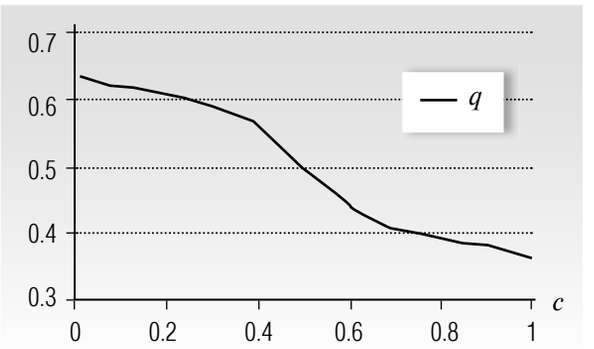

Fig. 5. The combination of the parameter c and the quantile q that allows for the maximum improvement of the quality of the estimation of the modified mean for the triangular distribution over [0,1] with the parameter $c \in (0,1)$

Table 1.

Values of asymptotic normalized variances $\sigma_q^2 = \lim_{N\to\infty} ND\overline{X}^q$, variances $\sigma^2 = ND\overline{X}$ and $d = \sigma_q^2 - \sigma^2$ for optimal combinations of c and q giving the minimum d values

c	q	σ^2	σ_q^2	d
0.01	0.6355	0.055006	0.015375	–0.03963
0.1	0.616	0.050555	0.014224	–0.03633
0.2	0.609	0.046667	0.013541	–0.0331
0.3	0.5935	0.043889	0.013383	–0.03051
0.4	0.564	0.042222	0.013619	–0.0286
0.5	0.500	0.041667	0.013889	–0.02778
0.6	0.436	0.042222	0.013619	–0.0286
0.7	0.4065	0.043889	0.013383	–0.03051
0.8	0.3910	0.046667	0.013541	–0.03313
0.9	0.384	0.050555	0.014224	–0.03633
0.99	0.3645	0.055006	0.015375	–0.03963

the expected value may be obtained using the information about a quantile that is greater than the median, and vice versa.

It is worth mentioning that for uniform distribution, the estimation of the cumulative distribution function involving symmetry has a smaller variance compared to the estimation that takes the median into account, as shown in [7].

As a result, the modified estimation of the net premium (2) with the account for the information about the quantile can be calculated as follows:

$$p^q = z \cdot \overline{X}^q, \qquad (11)$$

This formula yields a more accurate result compared to the classical method of estimation that uses the sample mean because in this case, the mean-square error $\overline{X}^q$ is smaller.

2. Estimation of the net premium using a quantile for voluntary health insurance

The proposed method for the net premium estimation (11) was applied to real-life data on insurance indemnity payments made on voluntary health insurance policies. For privacy reasons, the actual values are scaled, and the name of the insurance company is concealed. There have been $N = 239$ insured events over the considered period. *Table 2* presents the sample $X = \{X_1, X_2, ..., X_N\}$, nonrepeating payments Y_i, n_i – the number of repetitions of Y_i in the sample, $N = \sum_{i=1}^{k} n_i$, and k – the number of nonrepeating payments.

When the net premiums are calculated by the traditional method, the sample mean $\overline{X} = 504.73$ cmu/unit, the probability of an insured event was estimated as the ratio of the number of insured events to the number of policies, and it turned out that $z = 0.035$. The final net premium is $p = 17.67$.

However, from his long-term experience in the health insurance business, the company's actuary knows that in 90% of the cases the indemnity payments do not exceed 750 cmu/unit, i.e., we know the quantile $x_q = x_{0.9}$ of the $q = 0.9$ level for the random value considered here. Using (5), we obtain $\overline{X}^q = 516.23$ cmu/unit. As a result, by accounting for the quantile, we find the modified net premium of $p^q = 18.07$, which is 2.26% greater than p.

Table 2.

Scaled data on indemnity payments for insured events of voluntary health insurance in conventional monetary units (cmu) per unit

i	Y_i cmu/unit	n_i	i	Y_i cmu/unit	n_i	i	Y_i cmu/unit	n_i	i	Y_i cmu/unit	n_i
1	16.9	2	16	131.3	2	31	437.5	1	46	1343.8	1
2	20.6	1	17	137.5	1	32	468.8	5	47	1500.0	1
3	25.0	1	18	156.3	24	33	487.5	1	48	1562.5	2
4	28.1	1	19	162.5	1	34	500.0	3	49	1662.5	1
5	31.3	8	20	175.0	1	35	531.3	1	50	1687.5	1
6	37.5	9	21	187.5	18	36	562.5	2	51	2020.0	1
7	46.9	1	22	200	4	37	600.0	1	52	2500.0	1
8	50.0	1	23	210.0	1	38	625.0	8	53	3125.0	2
9	56.3	11	24	218.8	2	39	739.8	1	54	3750.0	2
10	62.5	19	25	243.8	2	40	750.0	2	55	4687.5	1
11	87.5	1	26	250.0	5	41	781.25	1	56	5000.0	1
12	93.8	20	27	281.3	4	42	937.5	2	57	5625.0	1
13	100.0	5	28	312.5	19	43	1000.0	1	58	8312.5	1
14	112.5	2	29	375.0	10	44	1125.0	2	59	8437.5	1
15	125.0	11	30	406.25	1	45	1250.0	2	60	9375.0	1

Therefore, knowledge of the quantile leads to a more accurate evaluation of the net premium. The premium calculated previously was underestimated, meaning an increased risk of bankruptcy for the insurance company. Since the modified estimate is more accurate and provided a large enough sample size ($N = 239$), the company's actuary was urged to recalculate the health insurance premiums.

Conclusion

This paper presents a new approach to the calculation of net premiums in the non-life insurance business. The approach uses additional information about the quantile of the cumulative distribution function to calculate the expected value of loss. The modified mean estimation is unbiased and more accurate than the classical sample mean, because its asymptotic mean-square error is smaller. Therefore, the new estimate of the net premium is more accurate for a sufficiently large number of observations.

The paper also studies the effect of the quantile on the asymptotic normalized variance of the modified mean in the cases of uniform, triangular and normal distributions. It is shown that in the case of a symmetric distribution, maximum accuracy is achieved by using the median for the adjustment calculation. As follows from the consideration of triangular distribution, positively skewed distributions will benefit from employing quantile levels greater than 0.5, and vice versa.

The proposed method was tested using real-life data on voluntary health insurance indemnity payments. The use of additional information allowed us to obtain a more accurate value for the net premium. The company's actuary was urged to recalculate the insurance premium rates to reduce the risk of the company's bankruptcy. ■

References

1. Mironkina Y.N., Sorokin A.S. (2011) *Osnovy aktuarnykh raschetov* [Fundamentals of actuarial calculations]. Moscow: Center EAOI (in Russian).
2. Falin G.I., Falin A.I. (2004) *Teoriya riska dlya aktuariev v zadachakh* [Theory of risk for actuaries in tasks]. Moscow: Mir, Nauchny Mir (in Russian).
3. Abu-Dayyeh W.A., Ahmed M.S., Ahmed R.A., Muttlak H.A. (2003) Some estimators of a finite population mean using auxiliary information. *Applied Mathematics and Computation*, no. 139, pp. 287–298.
4. Haq A., Shabbir J. (2014) An improved estimator of finite population mean when using two auxiliary attributes. *Applied Mathematics and Computation*, no. 241, pp. 14–24.
5. Singh H.P., Tailor R. (2005) Estimation of finite population mean with known coefficient of variation of an auxiliary character. *Statistica,* vol. LXV, no. 3, pp. 301–313.
6. Tarima S., Pavlov D. (2006) Using auxiliary information in statistical function estimation. ESAIM: *Probability and Statistics*, no. 10, pp. 11–23.
7. Dmitriev Y.G. (1976) O svoystvakh otsenok funktsii raspredeleniya i funktsionalov pri dopolnitel'noy apriornoy informatsii [On the features of distribution function and functionals estimations in case of additional a priori information]. *Mathematical Statistics and Its Applications*, no. 4, pp. 63–76 (in Russian).
8. Dmitriev Y.G., Tarasenko P.F. (1992) Ispol'zovanie apriornoy informatsii v statisticheskoy obrabotke eksperimental'nykh dannykh [Using a priori information in statistical processing of experimental data]. *Izvestiya vuzov. Physics*, no. 9, pp. 136–142 (in Russian).
9. Dmitriev Y.G., Ustinov Y.K. (1988) *Statisticheskoe otsenivanie raspredeleniy veroyatnostey s ispol'zovaniem dopolnitel'noy informatsii* [Statistical estimation of probability distributions using additional information]. Tomsk: TSU (in Russian).

10. Zhurko E.S., Zenkova Z.N. (2017) Vliyanie apriornoy informatsii na rezul'taty metoda tsenoobrazovaniya na tovar-novinku PSM [The influence of a priori information on the results of the PSM new product pricing method]. Proceedings of the *III International Scientific and Practical Conference "Actual Problems and Perspectives of State Statistics Development in Modern Conditions". Saratov, 5–7 December 2016.* Saratov: Saratovstat, vol. 2, pp. 66–68 (in Russian).
11. Zhurko E.S., Zenkova Z.N. (2016) Modifikatsiya metoda tsenoobrazovaniya PSM s uchetom kvantilya zadannogo urovnya [Modification of the PSM pricing method considering quantile of the specified level]. Proceedings of the *International Scientific and Practical Conference "Information Technologies of Siberia". Kemerovo, 10 November 2016.* Kemerovo: KuzSTU, pp. 134–136 (in Russian).
12. Zenkova Z.N., Makeeva O.B. (2015) Ispol'zovanie informatsii o kvantile pri analize oborachivaemosti oborotnykh sredstv [Using quantile information in current assets turnover analysis]. Proceedings of the *III All-Russian Youth Conference "Mathematical Support and Software for Informational, Technical and Economical Systems". Tomsk, 22–23 May 2015.* Tomsk: TSU, pp. 82–87 (in Russian).
13. Zenkova Z.N., Muravleva M.A. (2014) Asimptoticheskaya nesmeshchennost' otsenki funktsii raspredeleniya po odnokratno intervalom tsenzurirovannym dannym s privlecheniem informatsii o simmetrii [Asymptotic unbiasedness of distribution function estimation based on single interval-censored data using information about symmetry]. Proceedings of the *X All-Russian Conference "New Information Technologies in Complex Structures Research". Katun', 9–11 June 2014.* Tomsk: TSU, pp. 114–115 (in Russian).
14. Zenkova Z.N., Krakovetskaya I.V. (2013) Neparametricheskaya otsenka Ternbulla dlya interval'no-tsenzurirovannykh dannykh v marketingovom issledovanii sprosa na bioenergeticheskie napitki [Nonparametric Turnbull estimator for interval-censored data in the marketing research of the demand of bio-energy driks]. *Tomsk State University Journal of Control and Computer Science*, no. 3(24), pp. 64–69 (in Russian).
15. Borovkov A.A. (1997) *Matematicheskaya statistika* [Mathematical Statistics]. Novosibirsk: Nauka; Institute of Mathematics (in Russian).

Combined method to detect communities in graphs of interacting objects[1]

Sofiya Y. Lobanova
BSc Program Student
Tikhonov Moscow Institute of Electronics and Mathematics
National Research University Higher School of Economics
Address: 20, Myasnitskaya Street, Moscow, 101000, Russian Federation
E-mail: s.lobanova@usabilitylab.net

Alexander A. Chepovskiy
Associate Professor,School of Applied Mathematics
Tikhonov Moscow Institute of Electronics and Mathematics
National Research University Higher School of Economics
Address: 20, Myasnitskaya Street, Moscow, 101000, Russian Federation
E-mail: aachepovsky@hse.ru

Abstract

In this paper, we propose and implement a method for detecting intersecting and nested communities in graphs of interacting objects of different natures. For this, two classical algorithms are taken: a hierarchical agglomerate and one based on the search for k-cliques. The combined algorithm presented is based on their consistent application. In addition, parametric options are developed that are responsible for actions with communities whose sizes are smaller than the given k, and also with single vertices. Varying these parameters allows us to take into account differences in the topology of the original graph and thus to correct the algorithm.

The testing was carried out on real data, including on a group of graphs of a social network, and the qualitative content of the resulting partition was investigated. To assess the differences between the integrated method and the classical algorithms of community detections, a common measure of similarity was used. As a result, it is clearly shown that the resulting partitions are significantly different. We found that for the approach proposed in the article the index of the numerical characteristic of the partitioning accuracy, modularity, can be lower than the corresponding value for other approaches. At the same time, the result of an integrated method is often more informative due to intersections and nested community structure.

A visualization of the partition obtained for one of the examples by an integrated method at the first and last steps is presented. Along with the successfully found set of parameters of the integrated method for small communities and cut off vertices in the case of social networks, some shortcomings of the proposed model are noted. Proposals are made to develop this approach by using a set of parametric algorithms.

Key words: graph, graph analysis, community detection, graph structure, social network analysis, big data.

Citation: Lobanova S.Y., Chepovskiy A.A. (2017) Combined method to detect communities in graphs of interacting objects. *Business Informatics*, no. 4 (42), pp. 64–73. DOI: 10.17323/1998-0663.2017.4.64.73.

[1] This work was supported by the Russian Foundation for Basic Research (projects No. 16-29-09546 and No. 16-07-00641)

Introduction

Today, a wide variety of real-world systems of many fields can be represented as complex networks or graphs, namely, sets of vertices and edges with non-trivial topological internals. Examples of such networks run from technology [1] to biological fields [2] and social [3] studies.

The notion of community is quite general and associated with object classification. Depending on the context, it may refer to the module, the class, the group, the cluster, etc. This notion is of great value for solving the community detection task for the Internet [4]. The issue is topical for the implementation of on-the-edge search engines, data filtration and further automated sorting.

Typical networks to study are popular social network services. Then the user accounts are taken for the nodes, and the connections between them are determined by the formal notions of "friendship", "subscription", etc., so that we obtain a graph of interacting objects. Such networks have the structure of implicit communities – groups of people united by some additional feature, for example: classmates, subscribers to a number of public persons or thematic publics, criminal elements.

The problem of community detection in graphs of interacting objects is of great value to network analysis. In the general case, this task is NP-complete, so the paper proposes to use a method with a number of terminal options since it is required to deal with graphs from various fields [5, 6]. These options are designed to possibly refine the resulting community structure and are based on the initial graph topology and corresponding field.

1. Community detection methods

The well-known GN algorithm proposed by Girvan and Newman in 2002 [7, 8] is among the pioneers of divisive algorithms. Initially all the graph vertices share one community. Each step of the GN algorithm starts with the edge betweenness [9] calculation for every edge in the graph [8] and then it removes those with the highest score. The algorithm repeats this procedure, and it splits the network into disconnected parts, which in their turn go through the same process. The loop breaks when there are no more edges in the graph or when the split reaches modularity's [10] maxima.

Modularity is a numerical characteristic of the correctness of a partition. Currently, this is the most popular measure of the quality of the partition, and it is used directly or indirectly in the study of complex networks. As a rule, a closer to real partitioning has larger values of modularity.

Modularity Q is defined on an existing community structure of a graph:

$$Q = \frac{1}{2m}\sum_{i,j}\left[A_{ij} - \frac{k_i k_j}{2m}\right]\delta\left(c_i, c_j\right), \qquad (1)$$

where m the number of edges;

A_{ij} – the adjacency matrix element;

k_i – the degree of the vertex i;

c_i – the index of the class to which the vertex belongs;

$$\delta\left(c_i, c_j\right) = \begin{cases} 0, & c_i = c_j; \\ 1, & c_i \neq c_j. \end{cases}$$

The edge betweenness for a given edge is the fraction of the shortest paths passing this edge among all shortest paths:

$$c_B(e) = \sum_{\substack{s \neq t \\ s,v \in V}} \frac{\sigma_{st}(e)}{\sigma_{st}}, \qquad (2)$$

where e – the given edge;

V – the set of graph vertices;

σ_{st} – the number of the shortest paths between the vertices s and t;

$\sigma_{st}(e)$ – the number of the shortest paths between the vertices s and t passing through the edge e.

The calculation of edge betweenness $C_B(e)$ is computationally expensive, as was mentioned by Girvan and Newman [1, 8]. So in the case of studying large networks, it is better to apply different techniques [1].

The fast hierarchical approach was proposed by Blondel et al. in 2008 [11] for the general case of weighted graphs. The algorithm is agglomerative, so it combines current vertices to new supervertices (communities) step by step. The increase of modularity for such split must be maximal as well. At the end of each iteration, a new hierarchy level is reached. The loop breaks if there is no increase of modularity. The method is limited more by the processing memory than the time.

The resulting community structure may not make sense since the approach is based on the locals of vertices. That is why it is not always obvious if some intermediate split meets the reality. Moreover, the outcomes depend on the vertices order at the first step [12].

The aforementioned approaches mean that communities are not intersecting with each other. Furthermore, in real-world systems vertices are rarely belong to single community [13]. As a consequence, the resulting partition may not accurately describe the actual network.

Further on, for CPM we denote one of the well-known methods. In their work, Palla et al. [14, 15] presented a novel approach for detecting the overlapping community structure in complex networks.

In this method, finding the structure of overlapping communities is related to the problem of finding k-cliques in a network (complete subgraphs of size k). Given the clique size k, this algorithm first finds all k-cliques within the network. Then it associates two k-cliques if they share k - 1 nodes (*Figure 1*). Each community is made over a maximum set of adjacent k-cliques [15].

To measure the quality of such partitions, there are many modifications of classical modularity. In this paper, we consider the variant proposed in [13]:

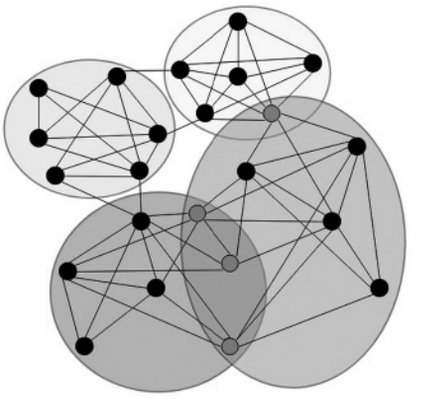

Fig. 1. Determined overlapping communities with the CPM algorithm for 4–cliques

$$Q_{ov} = \frac{1}{2m}\sum_{c\in C}\sum_{i,j\in c}\left[A_{ij} - \frac{k_i k_j}{2m}\right]\frac{1}{O_i O_j}, \qquad (3)$$

where $C = \{c_1, c_2, ..., c_{|C|}\}$ – the set of overlapping communities;

Q_i – the number of communities to which the vertex i belongs.

2. Description of the Combined method

In this paper, we consider a combination of the hierarchical Blondel algorithm with the CPM method of network partition into overlapping communities. By combination we mean a consistent application of the two methods. The idea is that at first our network is divided into large non-intersecting communities, and then we find smaller intersecting communities within them. Hereafter, we call this algorithm the "Combined method" (CM).

However, if Blondel's method identifies a finite community that does not split into parts in terms of a real network, then it is no longer useful to run CPM on it. Therefore, we apply the CPM algorithm to each separate c community identified by the Blondel method only if the percentage of vertices with high connection centrality does not exceed a predetermined value

$$\frac{\sum_{v \in c} \alpha(v)}{n_c} \le \beta(n_c), \alpha(v) = \begin{cases} c_B(v), c_B(v) \ge \tilde{\alpha} \\ 0, \quad c_B(v) < \tilde{\alpha} \end{cases}, \quad (4)$$

where n_C – number of vertices in c community; $c_B(v)$ – connection centrality of the vertex v; $\tilde{\alpha}$ and $\beta(n_c)$ – algorithm parameters, $0 \le \tilde{\alpha} < 1$; $0 < \beta \le 1$.

The relatively large connection centrality value characterizes the vertex as possibly connecting different communities. If the vertex is removed, the graph can be divided into several components. A similar consideration is used in the Girvan–Newman algorithm [1, 8]. The presence of a sufficient number of such peaks in the community may indicate that it has not been finalized completely, and further fragmentation by the Blondel algorithm is required. Therefore, the condition of applying CPM depends both on the value of connection centrality and on the percentage of vertices which have this value higher than some threshold $\tilde{\alpha}$. The threshold values for $c_B(v)$ and the vertex percentage are determined by the input parameters of algorithm $\tilde{\alpha}$ and β, and thus they are able to influence the final partition.

As a result of using CPM, vertices that are not included in cliques are cut off, and there is a problem of assigning them to the community. The solution is a system of two new parameters t_{opt} and m_{opt} of the Combined method. Depending on their values, not only the cut off vertices are redistributed, but also small communities. This allows us to partially influence the resulting structure, and in some cases – to avoid combining graph sheets or small communities into large, meaningless communities. This characteristic of the partition, which can be called "picking up junk", is intrinsic to many modularity optimization methods, including the Blondel algorithm [11, 16, 17].

To accurately define and illustrate the operation of the parameters t_{opt} and m_{opt}, let us consider $C_R = \{C_{R_1}, C_{R_2}, \ldots, C_{R_{rc}}\}$ the set of communities identified by the Blondel algorithm, where rc is their number. For convenience, we assume that if the individual community $C_{R_i} \in C_R$ does not satisfy the condition of applying the CPM, then the set $C_{CPM_i} = C_{R_i}$ duplicates it. Otherwise, C_{CPM_i} is defined as the set of communities obtained as a result of the CPM inside C_{CPM_i}:

$$C_{CPM_i} = \{C_{CPM_{i_1}}, C_{CPM_{i_2}}, \ldots, C_{CPM_{i_{mc_i}}}\}, \quad (5)$$

where mc_i number of communities in $C_{R_i} \in C_R$ that were identified by CPM algorithm.

Let us fix the result of CPM partition as C_f and denote the community structure after redistribution with the parameter i_{opt} as C_f^i. The parameters i_{opt} and m_{opt} are required when the algorithm starts, along with the size of the clique k and the limitations $\tilde{\alpha}$ and β. The community structure C_f is first transformed according to the value of the parameter i_{opt}, and then – according the value of the parameter m_{opt}.

The first parameter i_{opt} is responsible for the distribution of the cut off vertices of the CPM application and may have the following values (*Table 1*).

Table 1.

Description of the operation of i_{opt} parameter

None	Cut off vertices that do not belong to any community
i2c	Each cut off vertex is assigned to an individual (single) community within the community in which it was before applying CPM
i2r	Each cut off vertex is assigned to an individual (single) community that does not intersect with others
ia2c	A group of all cut off vertices that were in the same community prior to CPM application forms a community within it
ia2r	All the cut off vertices form one community that does not intersect with others

The second parameter m_{opt} works with small communities. The community is considered small if the number of vertices in it does not exceed the size of the clique k – the input parameter of the algorithm. This restriction is chosen because k is the minimum community size that can be indicated using the CPM method. The m_{opt} parameter can have the following values (*Table 2*).

Table 2.

Description of the operation of m_{opt} parameter

None	The structure of communities does not change
mdel	Small communities disband, but if the vertex belongs to other communities, then it remains in them. Otherwise, it will not belong to any community
m2c	All small communities that are in a common community (i.e., enclosed in it) are united into one common community that does not have enclosed communities
m2r	All small communities form one common community that does not intersect with others, even if the vertices belonged to other communities

Each of the transformations is done once, so when new small communities (possibly for *m2r*) appear, they will not be merged with the new common community. The resulting structure of C_f^* communities network is formed from C_f in accordance with the principles shown in *Tables 1* and *2*.

Separately, we should consider a combination of parameters with the values "*i2c / None*". In this case, when a small community is identified, it often turns out that all its members are also identified as individual communities, and large communities are divided into one larger part and several individual parts. To avoid a bunch of communities in the final structure, the authors use a modification: in the case described for small communities, individual communities disband, so only the smallest community remains. For a large community, it remains itself and the individual communities remain within it (that is, we remove the part of the internal partition except of those individual communities). *Figure 2* shows a possible result of the CM operation within one of the communities obtained in the first step. No similar result can be obtained by the separate application of Blondel and CPM algorithms. Meanwhile, in a number of cases in the graph analysis such a partition would be extremely informative.

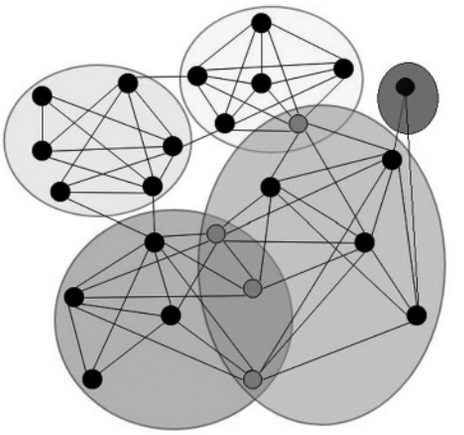

Fig. 2. Possible final partition within one of the communities obtained in the first step of CM

3. Tests

The results of the CM operation on different types of graphs were evaluated and compared with the separate results of the Blondel and CPM algorithms. In addition, comparison was made with the results of some classical algorithms: GN, quick greedy optimization of the modularity FG [18] and the CONGA algorithm (Cluster Overlap Newman-Girvan Optimized algorithm) [19]. We considered dependence of the CM results on the parameters k, $\tilde{\alpha}$, β, i_{opt}, m_{opt}. The size of the clique k is considered in the following range: $k \in [3, 10]$. For each partition, the value of Q_{ov} is calculated, and for the partitions obtained by CM, their similarity measure ω to other partitions must be calculated.

Collins and Dent [20, 21] proposed the ω (Omega Index) for estimating the similarity of two partitions, based on the number of pairs of vertices adjacent in the same number of communities. Let C_1 and C_2 be different partitions, K_1, K_2 – corresponding number of communities. Let $t_j(C)$ be the set of pairs of vertices occurring together in the C partition exactly j times. Then ω is defined as follows:

$$\omega(C',C'') = \frac{\omega_u(C',C'') - \omega_e(C',C'')}{1 - \omega_e(C',C'')}, \quad (6)$$

where $\omega_u(C',C'')$ – observed value,

$\omega_e(C',C'')$ – expected value:

$$\omega_u(C',C'') = \frac{1}{m}\sum_{j=0}^{\min(K_1,K_2)} \left| t_j(C') \bigcap t_j(C'') \right|,$$

$$\omega_e(C',C'') = \frac{1}{m^2}\sum_{j=0}^{\min(K_1,K_2)} \left| t_j(C') \right| \cdot \left| t_j(C'') \right|.$$

Value of $\omega(C',C'') = 1$ in the case of absolute equality of two partitions.

The authors studied the results of applying CM and selected classical algorithms on more than 70 graphs of interacting objects of various nature, such as: social networks, scientists citation networks, interactions of proteins, infrastructure networks. In a number of cases, for graphs from one group and with similar topological parameters, the most optimal values of the correcting parameters for CM have been identified. The software realization is written in Python 3.5, standard and open libraries, including classical algorithms, were used to work with graphs. Semi-automatic test blocks are implemented to find the best values for Q_{ov} parameter values. Visualization of the partitions obtained for the graphs is provided by third-party software.

Let us consider the results on ego-graphs obtained from the social network VKontakte (*Table 3*) in more detail. The ego-graphs in this case are graphs formed by a width search starting from a single vertex with a depth of 1 and adding all the edges between the vertices previously obtained. At the same time, as a starting point we took vertices with a degree in the range from about 100 to 500. This made it possible to obtain not only graphs with different topological properties, but ones which are also convenient for qualitative analysis of their sizes. The main objectives of this work are to study the content of the resulting partitions and to find the most suitable values of CM parameters.

Table 3.

Social network graphs for testing

	n	***m***	***avk***	***dst***	***avntr***	***avclu***	***maxcl***	***bet***	***clo***
G_1	105	792	15.1	0.1451	38.04	0.712	15	44.5	0.5434
G_2	120	1226	20.4	0.1717	52.39	0.668	14	49.3	0.5507
G_3	158	1352	17.1	0.1090	44.41	0.767	17	69.9	0.5309
G_4	158	1596	20.2	0.1287	63.80	0.638	21	68.4	0.5374
G_5	187	1873	20.0	0.1077	51.11	0.674	16	83.0	0.5308
G_6	229	1726	15.1	0.0661	26.67	0.633	14	106.5	0.5184
G_7	439	4143	18.9	0.0431	41.47	0.602	16	209.6	0.5117
G_8	461	4299	18.7	0.0405	49.36	0.665	17	220.7	0.5110

Notation: ***n*** – number of vertices; ***m*** – number of edges; ***avk*** – average degree of vertices; ***dst*** – connection density; ***avntr*** – average number of triangles; ***avclu*** – average cluster coefficient; ***maxcl*** – size of a maximal clique; ***bet*** – averaged betweenness centrality; ***clo*** – averaged closeness.

Table 4.

Test results

	$s(CM)$	Q_{ov}^{KA}	$s(BL)$	Q_{ov}^{BL}	$s(CP)$	Q_{ov}^{CP}	$s(GN)$	Q_{ov}^{GN}	$s(FG)$	Q_{ov}^{FG}	$s(CN)$	Q_{ov}^{CN}
G_1	38	0.213	4	0.320	4	0.167	17	0.250	3	0.315	3	0.245
G_2	21	0.367	4	0.385	4	– 0.013	20	0.329	4	0.371	4	0.136
G_3	28	0.650	4	0.614	5	0.390	15	0.581	5	0.581	5	0.494
G_4	39	0.327	4	0.342	3	0.016	47	0.175	3	0.313	4	0.100
G_5	36	0.048	5	0.369	3	– 0.008	27	0.250	5	0.325	110	0.011
G_6	51	0.224	5	0.465	4	– 0.029	22	0.408	6	0.393	4	0.273
G_7	90	0.501	7	0.572	8	– 0.149	54	0.536	8	0.509	5	0.304
G_8	99	0.349	7	0.569	11	–0.272	53	0.514	6	0.552	8	0.195

Notation: $s(CM)$ – number of communities identified by CM; $s(BL)$ – number of communities identified by the Blondel algorithm; $s(CP)$ – number of communities identified by CPM; $s(GN)$ – number of communities identified by GN; $s(FG)$ – number of communities identified by FG; $s(CN)$ – number of communities identified by CONGA; Each of Q_{ov}^{A} – value Q_{ov} for the partition, obtained by A algorithm.

In most cases, CM identified, as expected, a greater number of communities on these graphs (*Table 4*). For this group of graphs representing ego-graphs of the same social network, the most effective parameter option for the content of the partitions was the *i*2*c*/*m*2*r* and $k = 4$. A qualitative expert analysis of the composition of the groups singled out by the algorithms examined, especially of small communities of several vertices, showed the greatest efficiency especially for CM. For example, the Blondel, CPM and FG algorithms, as expected, left no large communities. And the smallest communities (including 1 element) identified by the GN algorithm for each of the graphs were not all meaningfully justified, and in some cases, taking into account the edges of the graph, they should be combined with the others.

It should be noted that the value of Q_{ov}^{KA} was not always higher than of other similar algorithms due to the fact that CM calculates another local extremum of the given quality measure, which is partially confirmed by the calculated values of the similarity measure ω of these partitions (*Table 5*). As you can see from the tables, not only the number of the communities obtained by CM and the other community algorithms differs greatly, but so do the sets of vertices combined into groups. After all, due to the peculiarities of the considered classical algorithms, some types of partitions cannot be obtained from their application, even theoretically. For example, the situation in *Figure 2* for CPM cannot be obtained. All combined, this shows that the partition obtained as a result of CM was essentially different from the others.

The partition obtained within the CM is hierarchical, with intersecting and nested communities, which gives an opportunity for more complete understanding of the structure of the investigated graph and better corresponds to the actual system of interaction of objects of some nature, in particular, users of social networks. As an illustration, let us consider the example for G_3 in more detail.

At the first step, the CM identifies four communities in the graph G_3: $C_{R_1}, C_{R_2}, C_{R_3}, C_{R_4}$, with the sizes 44, 41, 37, 36, respectively (*Figure 3*). This graph is very revealing, because the communities identified by the Blondel algorithm are of the similar size. Here C_{R_1} is a typical

Table 5.

The values of ω of the results of partitioning the graphs by CM and other algorithms

	$\omega(BL)$	$\omega(CPM)$	$\omega(GN)$	$\omega(FG)$	$\omega(CONGA)$
G_1	0.0811	–0.0478	0.0209	0.0729	0.0226
G_2	–0.0125	0.0139	0.0152	–0.0087	0.0655
G_3	0.0160	–0.0166	–0.0017	0.0043	0.0149
G_4	0.0459	–0.0077	–0.0098	0.0047	–0.0147
G_5	0.0651	–0.0058	0.0232	0.0133	–0.0055
G_6	0.0915	0.0003	0.0004	0.0159	0.0088
G_7	0.0452	0.0009	0.0061	0.0020	0.0066
G_8	0.0152	0.0137	–0.0070	0.0096	0.0079

Notation: $\omega(BL)$ – for partitioning a graph with CM and the Blondel algorithm; $\omega(CPM)$ – for partitioning a graph with CM and CPM algorithm; $\omega(GN)$ – for partitioning a graph with CM and GN algorithm; $\omega(FG)$ – for partitioning a graph with CM and FG algorithm; $\omega(CONGA)$ – for partitioning a graph with CM and CONGA algorithm.

"junk" community containing the original vertex of the ego-graph, leaf vertices and small communities. The other three communities are more meaningful, they are user groups, most of which have no connections with representatives of the other two communities.

Then, after applying the following CM steps on the graph G_3, the four communities are transformed into 28 as follows. Within the three meaningful dense communities, minimal changes occur. Thus, the algorithm found one vertex less connected to the others inside the C_{R_2}, for which, according to the *i2c/None* options, a separate community inside C_{R_2} was selected, and the remaining 40 vertices were put in the second internal community for C_{R_2}. The community C_{R_3} remained completely unchanged, and for C_{R_4} three vertices were found. They were separated into their individual internal communities, like in the C_{R_2} case, meanwhile creating a large remainder – a community of 33 vertices .

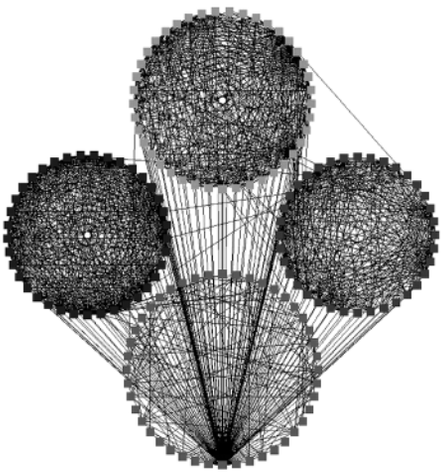

Fig. 3. The four communities identified by the Blondel algorithm in the graph G_3

The main changes occurred at the second step of the CM within C_{R_1}. Thus, in this community 4 others were identified. They intersected along the original vertex of the ego-graph. In addition, similar to the situation with the other two upper-level communities, 14 individual communities within the original community were identified (*Figure 4*).

In other ego-graph G_4 with the same number of vertices, but with a different structure (*Table 3*), 39 communities were identified after applying CM to it. And initially, at the first step, 4 large groups were also identified (here

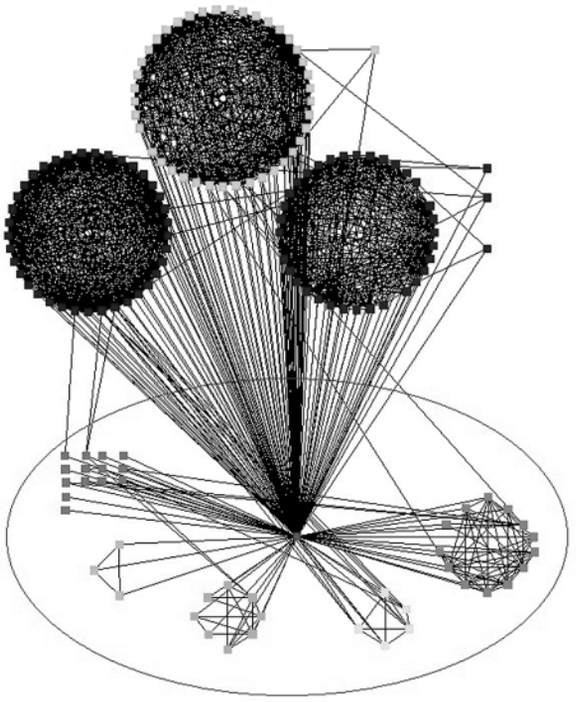

Fig. 4. Communities obtained by CM

$|C_{R_1}| = 50$, $|C_{R_2}| = 44$, $|C_{R_3}| = 44$, $|C_{R_4}| = 20$) inside of which already at the second step a search for intersecting nested communities is made. In this case, C_{R_2} and C_{R_3} did not change, but within the C_{R_1} and C_{R_4} 28 and 7 nested communities with intersections of different sizes were allocated, respectively.

It should be noted that according to the users of social networks whose ego-graphs were analyzed, other partitions obtained by the compared methods carried less information in themselves. However the Q_{ov} score was higher than for CM in a number of cases (*Table 4*). In particular, the results of algorithms that are components of CM were recognized as less informative and, for example, for GN the partitioning was partially superfluous.

Conclusion

A new algorithm, called by the authors the "Combined method" (CM), has been invented and presented to effectively identificate partitions of implicit communities in the graphs of interacting objects of various natures. We apply CM and well-known classical algorithms on the set of graphs from different industries which were taken from open sources. We studied the graph partitions into communities so obtained. Comparison of the modularity value Q_{ov} on the resultant partitions and the similarity measure ω between the result of CM and the remaining algorithms showed that in most cases there are significant differences in the resulting sets of communities. In addition, for a set of graphs from the social network, a meaningful analysis of the partitions obtained was carried out, and it showed that the result of CM with the options selected on this set more accurately identifies groups of interactions between users.

The proposed algorithm was run on graphs with different parameters. During the testing, it was found that the use of the parameters of the restrictions $\tilde{\alpha}$ and β on these graphs does not improve the quality of the resulting partition, since a result different from the results of applying the Blondel and CM algorithm (with no constraints) was achieved only in the case $\tilde{\alpha} > 0$ and $\beta = 0$. Further investigation was carried out at the values of the parameters $\tilde{\alpha} = 0$ and $\beta = 1$. Thus, the CPM method was run on all the Blondel communities generated by the algorithm.

The research showed that the most suitable values of additional parameters for social networks are *i2c/None* within the CM designations. At the same time, it cannot be stated that for a graph of interacting objects of a different nature this combination will be more effective than the others. Thus, the authors consider further studies on the design and implementation of parametric algorithms for the partition of communities. We also consider a possible use of automated tools to select suitable options and parameter values based on topological features and the original nature of graphs. ■

References

1. Guo Y., Xu K. (2007) An improved algorithm for finding community structure in networks with an application to IPv6 backbone network. *Frontiers of Computer Science in China*, vol. 1, pp. 459–467.
2. Wu F., Chen L., Wang J., Alhajj R. (2014) Biomolecular networks and human diseases. *BioMed Research International*, vol. 2014, pp. 1–2.
3. White H.C., Boorman S.A., Breiger R.L. (1976) Social structure from multiple networks. I. Blockmodels of roles and positions. *American Journal of Sociology*, vol. 81, no. 4, pp. 730–780.
4. Broder A.Z., Kumar S.R., Maghoul F., Raghavan P., Rajagopalan S., Stata R., Tomkins A., Wiener J. (2000) Graph structure in the Web. *Computer Networks*, vol. 33, no. 1–6, pp. 309–320.
5. Ahuja M.S., Singh J., Neha (2016) Practical applications of community detection. *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 6, no. 4, pp. 412–415.
6. Pennacchioli D., Coscia M., Pedreschi D. (2014) Overlap versus partition: Marketing classification and customer profiling in complex networks of products. Proceedings of 2014 *IEEE 30th International Conference on Data Engineering Workshops (ICDEW 2014). Chicago, 31 March – 4 April 2014*, pp. 103–110.
7. Newman M., Girvan M. (2004) Finding and evaluating community structure in networks. *Physical Review E*, vol. 69, no. 2, pp. 1–15.
8. Girvan M., Newman M. (2002) Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, vol. 99, no. 12, pp. 7821–7826.
9. Newman M. (2010) Networks: An introduction. Oxford: Oxford University Press.
10. Newman M. (2006) Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, vol. 103, no. 23, pp. 8577–8582.
11. Blondel V.D., Guillaume J.L., Lambiotte R., Lefebvre E. (2008) Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, no. 10, P10008.
12. Fortunato S. (2010) Community detection in graphs. *Physics Reports*, vol. 486, no. 3–5, pp. 75–174.
13. Shen H.W., Cheng X.Q., Cai K., Hu M.B. (2009) Detect overlapping and hierarchical community structure in networks. *Physica A: Statistical Mechanics and its Applications*, vol. 388, no. 8, pp. 1706–1712.
14. Palla G., Derényi I., Farkas I., Vicsek T. (2005) Uncovering the overlapping community structure of complex networks in nature and society. *Nature*, vol. 435, no. 7043, pp. 814–818.
15. Palla G., Ábel D., Farkas I.J., Pollner P., Derényi I., Vicsek T. (2009) k-clique percolation and clustering. *Handbook of Large-scale Random Networks*, Springer, ch. 9, pp. 369–408.
16. Fortunato S., Barthelemy M. (2006) Resolution limit in community detection // *Proceedings of the National Academy of Sciences*, vol. 104, no. 1, pp. 36–41.
17. Orlov A.O., Chepovskiy A.A. (2016) Osobennosti algoritma Blondelya pri vyyavlenii soobshchestv v grafe sotsial'noy seti [Features of Blondel algorithm to reveal communities in social networks]. Proceedings of *International Scientific Conference of Moscow Institute of Physics and Techniques and Institute of Physical and Technical Information (SCVRT1516*). Moscow, Protvino: Institute of Physical and Technical Information, pp. 124–129 (in Russian).
18. Clauset A., Newman M., Moore C. (2004) Finding community structure in very large networks. *Physical Review*, E 70. 066111.
19. Gregory S. (2007) An algorithm to find overlapping community structure in networks. Proceedings of *Knowledge Discovery in Databases: PKDD 2007: 11th European Conference on Principles and Practice of Knowledge Discovery in Databases, Warsaw, Poland, 17–21 September 2007.* Berlin, Heidelberg: Springer, 2007, Vol. 4702. P. 91–102.
20. Collins L.M., Dent C.W. (1988) Omega: A general formulation of the Rand index of cluster recovery suitable for non-disjoint solutions. *Multivariate Behavioral Research,* vol. 23, no. 2, pp. 231–242.
21. Gregory S. (2011) Fuzzy overlapping communities in networks. *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2011, no. 02, pp. 1–18.

Modeling and optimization of business processes and process systems under conditions of uncertainty

Alexander G. Madera
Professor, Department of Mathematics of the Faculty of Economics
National Research University Higher School of Economics
Address: 20, Myasnitskaya Street, Moscow, 101000, Russian Federation
E-mail: amadera@hse.ru

Abstract

This paper is devoted to mathematical modeling and optimization of business processes and process systems under conditions of uncertainty. At present, modeling of business processes is mainly descriptive, which does not allow quantitative modeling and optimization in the design of processes and process systems. In addition, the existing methods of decision-making in business processes are based on the assumption that the decisive factors are deterministic. Despite uncertainty of the real processes caused by the uncertainty of future costs of resources, the market environment, economy, finances, etc., the factors of an uncertain future are either not taken into account, or are believed to be the same as those observed currently.

In this paper, a stochastic interval mathematical optimization model is developed. This model allows us to simulate in a quantitative way the business processes and process systems in which they take place, taking into account the uncertainties of the future state of the economy, finances, market environment, costs of resources, as well as future realization of chances and risks related to the productive, supporting, and service processes. The criterion for optimality of the model is the maximization of the smallest deviation of the projected chances and risks, which makes it possible to make the best decision in the case that the most unfavorable conditions for the business process occur in the future. The criterion of optimality adopted in the mathematical model takes into account not only the uncertainty of the future state of the economy, finance, and market environment, but also the psychology of decision-making and the subjective nature of judgments and estimates. We present a concept and method for estimating the inductive (logical, subjective) probabilities of the occurrence of uncertain projected business process factors.

The models and methods developed in the paper make it possible to carry out mathematical modeling and optimization of business processes in a variety of activities without restrictions on the complexity of the structural model of the business process, the qualitative and quantitative composition of the connections in the process systems. On their basis, a software package for the quantitative design of business processes and process systems under conditions of uncertainty can be developed.

Key words: business process, process system, conditions of uncertainty, mathematical model, optimization, interval stochastic model, probability, optimization criterion, chances, risks.

Citation: Madera A.G. (2017) Modeling and optimization of business processes and process systems under conditions of uncertainty. *Business Informatics*, no. 4 (42), pp. 74–82.
DOI: 10.17323/1998-0663.2017.4.74.82.

Introduction

Quantitative design of business processes and corresponding process systems requires appropriate quantitative methods of modeling and optimization. However, modeling of processes and process systems in all fields of activity (manufacture, service, maintenance, support, supply, logistics, supply chains) is nowadays mainly qualitative and descriptive. The models are presented using either verbal, textual, graphical, or other kinds of descriptions or notations that form a stream of works, resources, and information. In the existing literature on business processes, their modeling is assumed to encompass the so-called regulation, documenting, and supporting document flow, and the optimization is understood as re-engineering, which includes a series of actions aimed at coordinating and amending the processes [1, 2]. However, all types of organizational instructions or actions based on some qualitative criteria of an optimum cannot be, in fact, regarded as optimization. Optimization should be based on mathematical analysis and the search for an extremum of some quantitative measure of optimality. Using qualitative analysis alone, it is impossible to ensure that an optimal solution is found.

In spite of the high demand for the development of mathematical methods of modeling and optimization of business processes and process systems, the design of such processes and systems is still mainly performed using qualitative approaches. This conclusion follows from the clearly insufficient number of published works on mathematical methods of modeling and process optimization (as pointed out, e.g., in [3]), as well as from the dire lack of case studies on real companies that practice quantitative design, modeling and optimization of their business processes [4, 5]. We note that while the publications [3, 6–14] present some mathematical models of business processes, these *ad hoc* results are only applicable to the particular cases considered in those papers. Moreover, the majority of publications consider business processes under fully determined conditions only, i.e., on the assumption that all the underlying factors of the business processes are known exactly and unambiguously.

In practice, however, all the factors that determine the business processes and process systems are usually uncertain. Among those factors are the future state of the economy, finances, market demand and supply, future prices of the resources and energy, investment amounts, demand for new products, chances, risks, and future financial stability of the organizational units comprising the process system. In reality, the parameters of both the present and the future necessarily involve uncertainty. Therefore we need to incorporate uncertainty conditions in the methods of modeling and business process optimization in order to adequately describe reality.

In this paper, we develop a mathematical model for the optimization of business processes and process systems under uncertainty conditions using stochastic interval factors obeying the uniform probability distribution. The models and methods of modeling and optimization developed in this paper can serve

as a foundation for the qualitative design of processes and process systems, development of the relevant software, and can be employed in practical applications in the areas of manufacture, support, supply, maintenance, customer service, logistics, supply chains among others.

1. Structural and mathematical models of the process elements

A structural model of a process and a process system includes the set of interacting organizational units that are modeled by the process elements and connections between them. Within each process element, various activities are performed that are aimed at the same common goal of creating the ultimate end-product possessing value for the consumer. A process element receives a flow of resources (industry factors) at the input. These resources are transformed into either intermediate products or end-products at the output of the element.

According to [2], there are the following types of business processes and corresponding elements (*Figure 1*):

✦ productive, or main, processes and elements (*Figure 1a*) – processes that increase both the value and cost of the end-product or intermediate product, by an amount equal to the cost of the product's manufacture. Within a productive element, the input flows of the material factors, energy, and labor are transformed into the output product (intermediate product, semi-finished product, unfinished production, the end-product of the overall process);

✦ supporting processes and elements (*Figure 1b*) – processes in which the product is neither produced or changed either qualitatively or quantitatively, value is not added, in contrast with the added value due to the energy and labor used in the production activities. Within a supporting element, the amounts of the production factors and products at the input and output of the element are equal to each other. Examples of supporting activities include loading, unloading, delivery of the factors and products, storage, documentation paperwork, and so on;

✦ service processes and elements (*Figure 1c*) – processes that increase both the cost and the value for the consumer for which they are prepared to pay. Similar to a productive element, a service element adds value to the product due to the spending of production factors, energy, and labor in the course of performing the services. At the same time, the amounts of the product at the input and output of the service element are equal, as they are in a supporting element. The service works may include any processing of the end-product done at the consumer's request and for which the consumer is prepared to

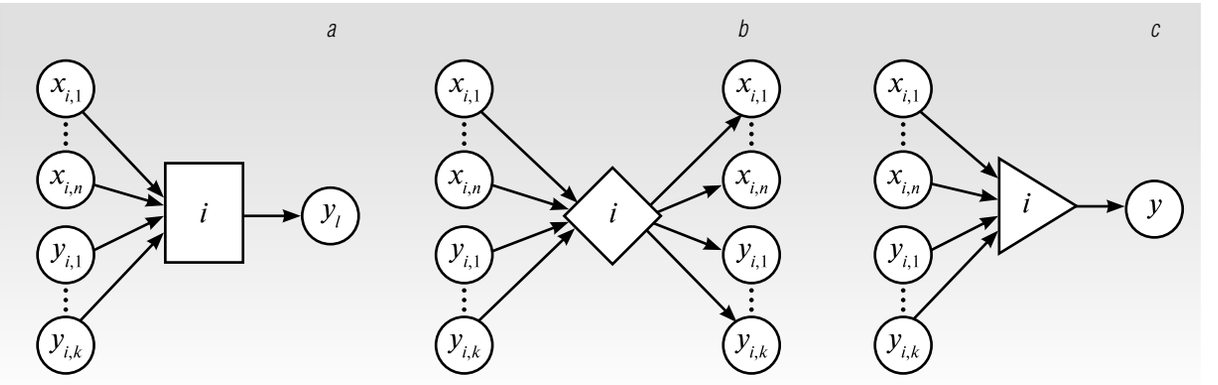

Fig. 1. Structural models for the (*a*) productive, (*b*) supporting, and (*c*) service elements

pay: custom packaging, sorting, labeling, improving the appearance, direct delivery to the consumer, etc.

The input flows of each element include both the exogenous flows of resource purchased outside the process (set $\boldsymbol{X}$), and the endogenous flows of products produced within prior elements of the same process (set $\boldsymbol{Y}$) that are fed, for the production purposes, into other elements of the same process. Exogenous and endogenous flows in the sets $\boldsymbol{X}$ and $\boldsymbol{Y}$ are numbered independently and continuously throughout the entire structure of the process system.

Let us consider the structural model of a process system of a sample business process (*Figure 2*). The flows of exogenic and endogenic resources are shown by circles with resource amounts marked inside them (x for exogenic, y for endogenic resources); the incoming resources and/or products are shown by arrows; the production elements are shown by rectangles (*Figure 1a*), the supporting elements are shown by diamonds (*Figure 1b*), and the service elements are shown by triangles (*Figure 1c*). The structural model in *Figure 2* contains two supporting elements (1 and 6), four production elements (2, 3, 4, 5 and 7), and a single service element (8). The exogenic resources are purchased in the amounts x_1, x_2, x_3, x_4, x_5 outside the process; the endogenic products are produced in the amounts y_1, y_2, y_3, y_4 within the process elements 2, 3, 4 and 5. The end-product of the entire business process is the output of the production element 7; it is produced in the amount y_5 and it serves as the input for the service element 8, in which service works are performed but $y_6 = y_5$.

The works to create the end-product in the amount y_5 are performed the following way. Resource x_1 and part x_2 of the resource x_2'(such that $x_2 = x_2' + x_2''$) are fed into the warehouse element 1, and from there to the output of the production elements 3 and 4. Part x_2'' of the resource x_2, and resource x_3 are directly fed into the production element 2, where they are used to create a product in the amount y_1. A part of this product, y_1' (such that $y_1 = y_1' + y_1''$), enters the production element 4, and the remaining part y_1'' is fed into the production element 5. Coming out, the production elements 3 and 4, the products y_1 and y_3 together with x_5 serve as the input for the production element 7, where they are transformed. The product y_4 after the production element 5 enters the warehouse together with the resource x_4. After that, they are fed as the input into the production element 7, where, together with the products y_2, y_3, and the resource x_5, they are transformed into the end-product of the business process y_5. The latter undergoes service actions in the service element 8, where it acquires properties demanded and paid for by the consumer.

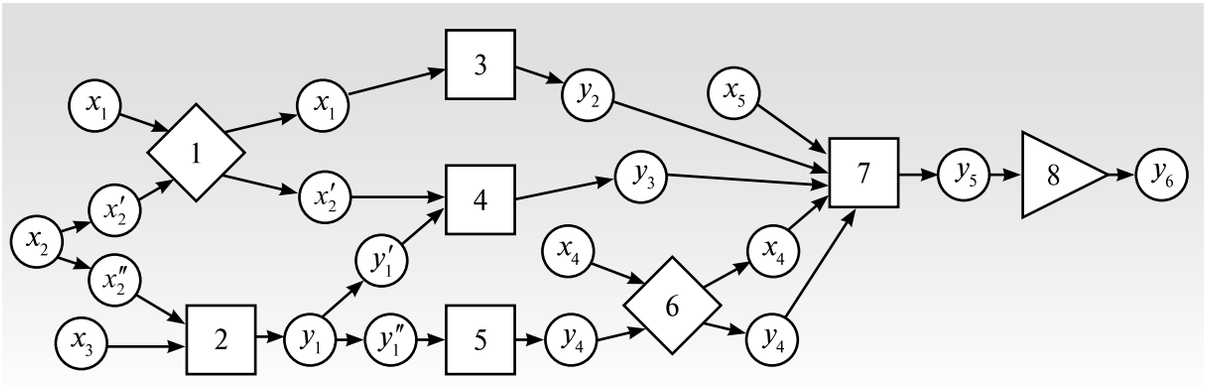

Fig. 2. A structural model of the business process and process system
(1, 6 – supporting elements, 2, 3, 4, 5, 7 – production elements, 8 – service element)

The mathematical model of the *i*-th production element (*Figure 1a*), which is fed the ordered sets of exogenic, $\boldsymbol{X}_i$, and endogenic, $\boldsymbol{Y}_i$, flows of resources $\{x_{i,1}, x_{i,2}, \ldots, x_{i,n}\} \in \boldsymbol{X}_i \subset \boldsymbol{X}$ and $\{y_{i,1}, y_{i,2}, \ldots, y_{i,k}\} \in \boldsymbol{Y}_i \subset \boldsymbol{Y}$, is a multifactor production function [15, 16], which can be represented in the following natural form:

$$y_l = f_i\left(x_{i,1}, x_{i,2}, \ldots, x_{i,n}; y_{i,1}, y_{i,2}, \ldots, y_{i,k}\right),$$

where $y_l \in \boldsymbol{Y}$ – the amount of the product that was produced within the *i*-th element.

In contrast to the production elements, where production factors are transformed into a qualitatively new product at the output, the supporting elements (*Figure 1b*) perform actions to support the production process, whereas within the service elements (*Figure 1c*), service works are carried out upon the created end-product. The production factors are not spent within the supporting or service elements during various actions; however, the energy, labor, and means of labor are spent during the works upon the end-product, but this does not change its amount. While the supporting processes do not add value to the produced products, the service processes, on the contrary, add value to the end-product by ensuring the properties requested by the consumer and providing services paid for by the consumer.

Therefore, the mathematical models of the supporting elements can be represented by identity production functions of the production factors' transformation into themselves,

$$f_i\left(x_{i,1}, x_{i,2}, \ldots, x_{i,n}; y_{i,1}, y_{i,2}, \ldots, y_{i,k}\right) = x_{ij}, j = 1, 2, \ldots, n,$$

$$f_i\left(x_{i,1}, x_{i,2}, \ldots, x_{i,n}; y_{i,1}, y_{i,2}, \ldots, y_{i,k}\right) = y_{ij}, j = 1, 2, \ldots, k,$$

The mathematical models of the service elements are then represented by production functions that transform the production factors and the end-product in the amount of *y* into a product of the same amount, $f\left(x_{i,1}, x_{i,2}, \ldots, x_{i,n}; y\right) = y$ [15, 16].

2. A mathematical model of a business process and process system under conditions of uncertainty

To adequately describe real business processes, mathematical models should take into account the uncertainty of the future state of economy, finances, market demand prices for production factors and energy, volumes of investments, chances, risks, and other factors of uncertainty (see Section 1), as well as the psychological factors that affect the decisions, opinions, and estimates [17–21]. Uncertainty is modeled by the stochastic interval factors that are uniformly distributed inside the intervals whose boundaries are determined by inductive (logical, or subjective) probability [16, 20–25]. To set up the optimization model, we need to define the quantitative optimization criterion and constraints on the range of acceptable solutions, $\boldsymbol{D}$ [26].

Optimization criterion for the mathematical model. The uncertain events that might be crucial in the future could be divided into two types: events that could be favorable for the business process (chances) – high profits, goal achievement, obtaining the planned results – and unfavorable events (risks) such as losses, insufficient profit, failures, bankruptcies, etc. As shown in [16, 24], to ensure the best solutions, the optimization criterion should maximize the chances and minimize the risks. The complex criterion "chances-risks" *Ch&R* [16, 24] satisfies these conditions and is defined as

$$Ch\&R = \beta_{Ch} Ch - \beta_R \left| R \right|,$$

where *Ch* and *R* – the total projected chances and risks relevant to the process under consideration;

$\beta_{Ch}, \beta_R \in [0,1]$ – the coefficients of relative importance of the chances and risks.

The values of the total chances $Ch \in \boldsymbol{M}$ and risks $R \in \boldsymbol{M}$ (where $\boldsymbol{M}$ set includes *L* chances and *K* risks) are determined [16, 24]

as sums of the products of the quantitative measures of chances (income, profit, cash) $M_{Ch,k},(k=1,2,\ldots,L)$ by risks (losses, defaults, damages) $M_{R,k},(k=1,2,\ldots,K)$ by the measures of their possible realization, or probabilities, $P_{Ch,k}$ and $P_{R,k}$, i.e.

$$Ch=\sum_{k=1}^{L}M_{Ch,k}P_{Ch,k},\quad R=\sum_{k=1}^{K}M_{R,k}P_{R,k}.$$

Here the qualitative measures of the k-th chance or risk are stochastic intervals, varying within the intervals $M_{Ch,k}\in\left[\underline{M}_{Ch,k},\overline{M}_{Ch,k}\right]$, $M_{Ch,k}\in\left[\underline{M}_{R,k},\overline{M}_{R,k}\right]$, where $\underline{M}_{Ch,k},\overline{M}_{Ch,k}$ and $\underline{M}_{R,k},\overline{M}_{R,k}$ are the lower and upper bounds of the intervals of the quantitative measures of the chances and risks, which are calculated according to the rules of the interval arithmetic [27].

To obtain the best solution relevant for the business process that will maximize the chances and minimize the risks at the same time, the optimization criterion "chances-risks" $Ch\&R$ must be maximized. This means we have to find a solution under uncertainty conditions such that we ensure the maximum possible difference between the lower bound for chances $\underline{Ch}$ and the upper bound for risks $\overline{R}$ [16, 24] for any $\omega\in\Omega$ (where ω is an elementary event, and Ω is the space of elementary events). We obtain the following optimization criterion for business processes:

$$\max\left\{Ch\&R\right\}=\max\left\{\beta_{Ch}\underline{Ch}-\beta_{R}\left|\overline{R}\right.\right\},\qquad(1)$$

where $\underline{Ch}=\sum_{k=1}^{L}\underline{M}_{Ch,k}P_{Ch,k}$, $\overline{R}=\sum_{k=1}^{K}\overline{M}_{R,k}P_{R,k}$.

Constraints of the mathematical optimization model. Constraints of the mathematical model are determined by the budget and set the conditions for the value of the total expenses for the process activities. These expenses should not exceed the finances available for the business process. Let us denote the amounts of the purchased exogenic production factors by $\left\{x\right\}=\left\{x_1,x_2,\ldots,x_m\right\}\in\boldsymbol{D}$, $x_i\geq 0$, and their costs by $c_{Ch,i}$ and $c_{R,i}$, (i = 1, 2, ...m). Considering that the costs of the production factors are uncertain and given by stochastic intervals, and that they vary within the intervals $c_{Ch,i}(\omega)\in\left[\underline{c}_{Ch,i},\overline{c}_{Ch,i}\right]$ and $c_{R,i}(\omega)\in\left[\underline{c}_{R,i},\overline{c}_{R,i}\right]$, $\overline{c}_{Ch,i}\leq\underline{c}_{R,i}$ with the lower and upper bounds $\underline{c}_{Ch,i},\overline{c}_{Ch,i}$ and $\underline{c}_{R,i},\overline{c}_{R,i}$, we obtain the expressions for the total expenses for the purchase of the production factors for each $\omega\in\Omega$:

$$C_{Ch}(\omega)=\sum_{i=1}^{m}c_{Ch,i}(\omega)x_i,\ C_R(\omega)=\sum_{i=1}^{m}c_{R,i}(\omega)x_i.$$

Since both the expenses and the funding volumes of the business process are stochastically uncertain, we should consider the budget constraint in a probabilistic sense, namely through the probabilities of satisfying the budget constraints (inequalities), which should be lower than certain threshold values p_{Ch} and p_R set by the owner of the business process [28] and the business analysts. Therefore, the constraints for the chances and risks can be written down as follows:

$$P\left\{C_{Ch}(\omega)\leq I_{Ch}(\omega)\right\}\geq p_{Ch},\ P\left\{C_R(\omega)\leq I_R(\omega)\right\}\geq p_R,\qquad(2)$$

where $I_{Ch}(\omega)\in\left[\underline{I}_{Ch},\overline{I}_{Ch}\right]$ and $I_R(\omega)\in\left[\underline{I}_{R},\overline{I}_{R}\right]$, $\overline{I}_{Ch}\leq\underline{I}_{R}$ – the stochastic interval volumes of the funding of the business process;

$P\{\cdot\}$ – the inductive (subjective, or logical) probabilities.

The stochastic interval mathematical optimization model. The stochastic interval mathematical optimization model of a business process with uncertain parameters is set up as follows: we need to find the amounts of the production factors $\{x\}=\{x_1, x_2, \ldots x_m,\}\in\boldsymbol{D}$, $x_i\geq 0$, i = 1, 2, ... m, that provide the maximum value of the criterion "chances-risks" (1) and satisfy the probabilistic constraints for the chances and risks, i.e.:

$$\max_{x}\left\{Ch\&R\right\}=\max_{x}\left\{\beta_{Ch}\underline{Ch}-\beta_{R}\left|\overline{R}\right.\right\},$$

$$P\left\{C_{Ch}(\omega)\leq I_{Ch}(\omega)\right\}\geq p_{Ch},\quad P\left\{C_R(\omega)\leq I_R(\omega)\right\}\geq p_R.$$

To find the optimal solution for the given optimization model under the conditions of the stochastic interval uncertainty, we need to rewrite the probabilistic constraints (2) as the corresponding deterministic equivalents [29].

The equivalent deterministic mathematical optimization model. Let us note that the stochastic interval costs of the production factors $c_{Ch,i}(\omega)$, $c_{R,i}(\omega)$ and the funding volumes $I_{Ch}(\omega)$, $I_R(\omega)$, $\omega \in \Omega$ are uniformly distributed over respective intervals. By Lyapunov's central limit theorem [30], their sums, which are equal to the total expenses, are normally distributed with the expected values

$$m(C_{Ch}) = \sum_{i=1}^{m} m(c_{Ch,i})x_i, \;\; m(C_R) = \sum_{i=1}^{m} m(c_{R,i})x_i$$

and dispersions

$$\sigma^2(C_{Ch}) = \sum_{i=1}^{m} \sigma^2(c_{Ch,i})x_i^2, \;\; \sigma^2(C_R) = \sum_{i=1}^{m} \sigma^2(c_{R,i})x_i^2,$$

where

$$m(c_{Ch,i}) = \left(\underline{c}_{Ch,i} + \overline{c}_{Ch,i}\right)/2 \;, m(c_{R,i}) = \left(\underline{c}_{R,i} + \overline{c}_{R,i}\right)/2,$$

$$\sigma^2(c_{Ch,i}) = \left(\underline{c}_{Ch,i} - \overline{c}_{Ch,i}\right)^2/12, \;\; \sigma^2(c_{R,i}) = \left(\underline{c}_{R,i} - \overline{c}_{R,i}\right)^2/12.$$

By the theorem [29], we rewrite the probabilistic inequalities in the constraints (2) in a deterministic form and thus obtain the final equivalent deterministic mathematical optimization model:

$$\max_{\{x\}} \{Ch \& R\} = \max_{\{x\}} \left\{\beta_{Ch}\underline{Ch} - \beta_R \left| \overline{R} \right.\right\},$$

$$m(C_{Ch}) + k_{Ch}\sqrt{\sigma^2(C_{Ch}) + \sigma^2(I_{Ch})} \le m(I_{Ch}),$$

$$m(C_R) + k_R\sqrt{\sigma^2(C_R) + \sigma^2(I_R)} \le m(I_R),$$

$$x_i \ge 0, \;\; i = 1, 2, \ldots, m,$$

where k_{Ch}, k_R are the quantiles of the standard normal distributions;

$m(I_{Ch})$, $m(I_R)$ and $\sigma^2(I_{Ch})$, $\sigma^2(I_R)$ are expected values and dispersions for the stochastic interval volumes of funding:

$$m(I_{Ch}) = \left(\underline{I}_{Ch} + \overline{I}_{Ch}\right)/2, \; m(I_R) = \left(\underline{I}_R + \overline{I}_R\right)/2,$$

$$\sigma^2(I_{Ch}) = \left(\underline{I}_{Ch} - \overline{I}_{Ch}\right)^2/12, \; \sigma^2(I_R) = \left(\underline{I}_R - \overline{I}_R\right)^2/12.$$

The deterministic mathematical optimization model so obtained belongs to the class of general problems of nonlinear programming. Its optimal solution can be obtained using specialized software packages, such as NEOS, AMPL, MATLAB, MathOptimizier, GALAHAD, and MS Excel. From the optimal solution $\{x^*\} = \{x_1^*, x_2^*, \ldots, x_m^*\}$ obtained numerically, we can find the optimal values for all the factors of the business process and process system: the amount and cost of the end-product, the amounts and costs of the intermediate products obtained as the outputs of each element, incomes and profits under different projected future conditions, etc.

3. Method for estimating the inductive probabilities of projected events

Probabilities of future events (the state of economy, finances, risks, chances, etc.) should be estimated using the concept of inductive probability, because the projected events are unique, and the concepts of the classical or the statistical probabilities are not applicable. Unique events are not the random objects that classical probability theory deals with, because such events do not satisfy the necessary conditions of being indefinitely repeatable, uniform, and possessing a stable frequency in a large number of trials under identical conditions [20–24].

The concept and method for estimating the inductive probabilities $P(A)$ for the events $\{A\} = \{A_1, A_2, \ldots A_n\}$ to occur are developed in [21] and lead to the projections that are considerably more accurate. The method is based both on the statistical data on the projections of the relevant events over past periods of time and on new information about the present trends. The data for the past periods of time characterize

the errors made by the person who performed the projections of the events and estimated their probabilities. The data derived from the history of the events' development and information about the current trends can be used to calculate the corrections. These two types of data allow one to derive two matrices: the matrix of projecting errors made by the given person, $\mathcal{L}$, and the matrix of corrected projections, $\mathcal{M}$, calculated by the use of current information. As shown in [21], the unknown vector $P(A)$ describing the probability that the projected events will occur is the eigenvector of the complete projection error matrix $\mathcal{K} = \mathcal{M} \cdot \mathcal{L}$ corresponding to the eigenvalue 1, i.e. $P(A) = \mathcal{K} \cdot P(A)$.

Conclusion

The stochastic interval mathematical optimization model proposed in this paper allows for the modeling of optimal business processes and optimal structures and parameters of the process systems under conditions of uncertainty reflecting the state of the future economy, finances, market environment, demand for new products, costs of the production factors, sale price of the end-product, investments volumes, future chances and risks, and so on. The structural model of a production system may include the productive, supporting, and service elements that cover the entire range of activities performed during various business processes that transform the input factors into the output products. The mathematical models and methods developed in this paper facilitate the mathematical modeling and optimization of various business processes and process systems without limitations regarding the complexity, qualitative and quantitative composition of the process elements, or the complexity of the structure of the process system. They can serve as a foundation for the development of software packages for the quantitative design of business processes and complex process systems under conditions of uncertainty. ■

References

1. Robson M., Ullah P. (1996) *A practical guide to business process re-engineering.* Hempshire: Gower.
2. Harrington H.J., Esseling K., van Nimwegen H. (1997) *Business process improvement workbook: Documentation, analysis, design, and management of business process improvement.* N.Y.: McGraw-Hill.
3. Vergidis K., Tiwari A., Majeed B. (2008) Business process analysis and optimization: Beyond reengineering. *IEEE Transactions on Systems, Man, and Cybernetics. Part C (Applications and Reviews)*, vol. 38, no. 1, pp. 69–82.
4. Stock J.R., Lambert D.M. (2001) *Strategic logistics management.* Boston: McGraw-Hill.
5. Shapiro J.F. (2001) *Modeling the supply chain. Duxbury*, CA: Thomson Learning.
6. van der Aalst W.M.P. (1998) The application of Petri nets to workflow management. *Journal of Circuits, Systems and Computers*, vol. 8, no. 1, pp. 21–66.
7. Hofacker I., Vetschera R. (2001) Algorithmical approaches to business process design. *Computers & Operations Research*, no. 28, pp. 1253–1275.
8. Koubarakis M., Plexousakis D. (2002) A formal framework for business process modelling and design. *Information Systems*, vol. 27, no. 5, pp. 299–319.
9. Valiris G., Glykas M. (2004) Business analysis metrics for business process redesign. *Business Process Management Journal*, vol. 10, no. 4, pp. 445–480.
10. Seuring S. (2013) A review of modeling approaches for sustainable supply chain management. *Decision Support Systems*, vol. 54, no. 4, pp. 1513–1520.
11. Wibig M. (2013) Dynamic programming and genetic algorithm for business processes optimization. *International Journal of Intelligent Systems and Applications*, vol. 5, no. 1, pp. 44–51.

12. Sawicki P., Sawicka H. (2014) Logistics process improvement using simulation and stochastic multiple criteria decision aiding. *Procedia – Social and Behavioral Sciences*, no. 111, pp. 213–223.
13. Zhang Q., Shah N., Wassick J., Helling R., van Egerschot P. (2014) Sustainable supply chain optimisation: An industrial case study. *Computers & Industrial Engineering*, no. 74, pp. 68–83.
14. Ahmadikatouli A., Motameni H. (2015) Enrichment of object oriented Petri net and object Z aiming at business process optimization. *International Journal of Advanced Computer Science and Applications*, vol. 6, no. 7, pp. 13–19.
15. Kleiner G.B. (1986) *Proizvodstvennye funktsii: Teoriya, metody, primenenie* [Production functions: Theory, methods, application]. Moscow: Finance and Statistics (in Russian).
16. Madera A.G. (2015) Matematicheskoe modelirovanie i optimizatsiya biznes-protsessov na osnove kompleksnogo kriteriya "shansy – riski" [Mathematical modeling and optimization of business processes based on an integrated criterion of chances and risks]. *Russian Management Journal*, vol. 13, no. 4, pp. 51–68 (in Russian).
17. Kozielecki J. (1979) *Psikhologicheskaya teoriya resheniy* [Psychological decision theory]. Moscow: Progress (in Russian).
18. Plous S. (1993) *The psychology of judgment and decision making.* N.Y.: McGraw-Hill, Wesleyan University.
19. Werth L. (2004) *Psychologie für die wirtschaft. Grundlagen und anwendungen.* Berlin: Spectrum Akademischer Verlag.
20. Madera A.G. (2014) Interval'no stokhasticheskaya neopredelennost' otsenok v mnogokriterial'nykh zadachakh prinyatiya resheniy [Interval stochastic uncertainty of estimates in multiple criteria decision making problems]. *Artificial Intelligence and Decision Making*, no. 3, pp. 105–115 (in Russian).
21. Madera A.G. (2016) Metod opredeleniya veroyatnostey prognoziruemykh sobytiy pri prinyatii resheniy [Method for of probabilities of the projected events when making decisions]. *Artificial Intelligence and Decision Making*, no. 2. pp. 38–45 (in Russian).
22. Carnap R. (1971) *Logical foundations of probability.* Chicago: The University of Chicago Press.
23. de Finetty B. (1974) Bayesinism: Its unifying role for both the foundations and applications of statistics. *International Statistical Review*, no. 42, pp. 117–130.
24. Madera A.G. (2014) *Riski i shansy: neopredelennost', prognozirovanie i otsenka* [Risks and chances: Uncertainty, forecasting and estimation]. Moscow: Krasand (in Russian).
25. Madera A.G. (2016) Estimating the probability of forecasted events. *International Journal of Accounting and Economics Studies*, vol. 4, no. 1, pp. 76–80.
26. Madera A.G. (2009) *Modelirovanie i prinyatie resheniy v menedzhmente* [Modeling and decision making in management]. Moscow: LKI (in Russian).
27. Alefeld G., Herzberger J. (1983) *Introduction to interval computations.* N.Y.: Academic Press.
28. ISO/IEC 15288:2002. *System engineering – System life cycle processes.*
29. Liu B. (2002) *Theory and practice of uncertain programming.* Heidelberg: Physica-Verlag.
30. Feller W. (1970) *An introduction to probability theory and its applications.* N.Y.: John Wiley.

XIX April International Academic Conference on Economic and Social Development

In Moscow on April 10-13, 2018, the National Research University Higher School of Economics (HSE), with the support of the World Bank, will be holding the 19th April International Academic Conference on Economic and Social Development. The Conference's Program Committee is chaired by Professor Evgeny Yasin, HSE Academic Supervisor.

The Conference offers a diverse agenda concerning social and economic development in the Russian Federation.

The plenary sessions and roundtable discussions will focus on the following topics:

- ✦ Economic Stability and Growth;
- ✦ Inequality and Inclusive Development;
- ✦ Perspectives and Drivers of Globalization and Regionalization;
- ✦ Technological Future: Towards a Smart Society;
- ✦ The Arctic: Challenges of the 21st Century.

AUTHORS GUIDELINES

Articles should be topical and original, should outline tasks (issues), describe key results of the author's research and appropriate conclusions.

Manuscripts are submitted via e-mail: bijournal@hse.ru.

MANUSCRIPT REQUIREMENTS

TEXT FILES should be submitted in electronic form, as a MS Word document (version 2003 or higher).

LENGTH. Articles should be between 20 and 25 thousand characters (incl. spaces).

FONT, SPACING, MARGINS. The text should be in Times New Roman 12 pt, 1.5 spaced, fit to the width, margins: left – 25 mm, all other – 15 mm.

TITLE of the article should be submitted in native language and English.

AUTHORS' DETAILS are presented in native language and English. The details include:

- Full name of each author
- Position, rank, academic degree of each author
- Affiliation of each author, at the time the research was completed
- Full postal address of each affiliation (incl. postcode / ZIP)
- E-mail address of each author.

ABSTRACT are presented in native language and English.

- The abstract should be between 200 and 300 words.
- The abstract should be informative (no general words), original, relevant (reflects your paper's key content and research findings); structured (follows the logics of results' presentation in the paper)
- The recommended structure: purpose (mandatory), design / methodology / approach (mandatory), findings (mandatory), research limitations / implications (if applicable), practical implications (if applicable), originality / value (mandatory).
- It is appropriate to describe the research methods/ methodology if they are original or of interest for this particular research. For papers concerned with experimental work the data sources and data procession technique should be described.
- The results should be described as precisely and informatively as possible. Include your key theoretical and experimental results, factual information, revealed interconnections and patterns. Give special priority in the abstract to new results and long-term impact data, important discoveries and verified findings that contradict previous theories as well as data that you think have practical value.
- Conclusions may be associated with recommendations, estimates, suggestions, hypotheses described in the paper.
- Information contained in the title should not be duplicated in the abstract. Authors should try to avoid unnecessary introductory phrases (e.g. «the author of the paper considers...»).
- Authors should use the language typical of research and technical documents to compile your abstract and avoid complex grammatical constructions.
- The text of the abstract should include key words of the paper.

KEYWORDS are presented in native language and English. The number of key words / words combinations are from 6 to 10 (separated by semicolons).

FORMULAE should be prepared using Math Type or MS Equation tool.

FIGURES should be of high quality, black and white, legible and numbered consecutively with Arabic numerals. All figures (charts, diagrams, etc.) should be submitted in electronic form (photo images – in TIF, PSD or JPEG formats, minimum resolution 300 dpi). Appropriate references in the text are required.

REFERENCES should be presented in Harvard style and carefully checked for completeness, accuracy and consistency.

The publication is free of charge.