

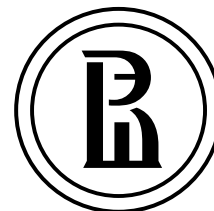
ISSN 2587-814X (print), ISSN 2587-8158 (online)

Russian version: ISSN 1998-0663 (print), ISSN 2587-8166 (online)

№3(41)-2017

BUSINESS INFORMATICS

HSE SCIENTIFIC JOURNAL



CONTENTS

Modeling of social and economic systems

*Tatiana A. Gavrilova, Artem I. Alsufyev,
Elvira Y. Grinberg*

Knowledge visualization: Critique
of the St. Gallen School and an analysis
of contemporary trends.....7

Elena D. Kopnova, Lilia A. Rodionova

Modeling the influence of foreign investments
on food security based on panel
cointegration models20

Information systems and technologies in business

Eduard A. Babkin, Nikita O. Ponomarev

Analysis of the consistency of enterprise
architecture models using formal
verification methods30

Ruslan A. Dolzhenko

Approaches to geomarketing via office
locations serving individuals and legal entities
of a commercial bank.....41

Pavel V. Malyzhenkov, Marina I. Ivanova

An architectural approach
to IT–business alignment56

Information security

Marina V. Tumbinskaya

Process of distribution of undesirable
information in social networks65

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. Koucheryav

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

© National Research University
Higher School of Economics

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

EDITOR-IN-CHIEF

Alexey Golosov –

President of FORS Development Center, Russian Federation

DEPUTY EDITOR-IN-CHIEF

Yevgeni Koucheryavy –

Professor, Department of Electronics and Communication Engineering, Tampere University of Technology, Finland

EDITORIAL BOARD

Habib Abdulrab –

Professor, Mathematical and Software Engineering Department, National Institute of Applied Sciences – Institut national des sciences appliquées de Rouen (INSA de Rouen), Rouen, France

Sergey Avdoshin –

Professor, Head of School of Software Engineering, National Research University Higher School of Economics, Russian Federation

Andranik Akopov –

Professor, Department of Business Analytics, National Research University Higher School of Economics, Russian Federation

Fuad Aleskerov –

Professor, Head of Department of Mathematics, National Research University Higher School of Economics, Russian Federation

Anton Afanasyev –

Leading Researcher, Laboratory of Social Modeling, Central Economics and Mathematics Institute, Russian Academy of Science, Russian Federation

Eduard Babkin –

Professor, Department of Information Systems and Technologies, National Research University Higher School of Economics, Russian Federation

Alex Bayer –

Head of KAFAN FX Information Services, New York, USA

Alexander Baranov –

Deputy Head of Central Scientific and Research Computing Center, Federal Tax Service of Russia, Russian Federation

Jorg Becker –

Vice-Rector, Professor, Director of European Research Center for Information Systems (ERCIS), University of Munster, Germany

Vladimir Belov –

Professor, Department of Computational and Applied Mathematics, Ryazan State Radio Engineering University, Russian Federation

Andrey Gribov –

Director General, CyberPlat Company, Russian Federation

Alexander Gromov –

Professor, Head of Department of Modeling and Business Process Optimization, National Research University Higher School of Economics, Russian Federation

Vladimir Gurvich –

Invited Professor and Researcher, Rutgers Center for Operations Research, Rutgers, The State University of New Jersey, USA

Laurence Jacobs –

Professor, Medical School, University of Zurich, Switzerland

Isif Diskin –

Academic Supervisor, Chairmen of Scientific and Expert Council, Russian Public Opinion Research Center (VCIOM); Member of the Council, The Russian Public Chamber; Russian Federation

Kurt Sandkuhl –

Professor, Head of Department of Business Information Systems, University of Rostock, Germany

Nikolay Ilyin –

Deputy Head, Administration of Special Communication, Federal Security Guard, Russian Federation

Dmitry Isaev –

Associate Professor, Department of Business Analytics, National Research University Higher School of Economics, Russian Federation

Valery Kalyagin –

Professor, Head of Department of Applied Mathematics and Informatics, National Research University Higher School of Economics, Russian Federation

Maria Kamennova –

Director General, BPM Logic, Russian Federation

Tatiana Kravchenko –

Professor, Head of Department of Business Analytics, National Research University Higher School of Economics, Russian Federation

Sergey Kuznetsov –

Professor, Head of School of Data Analysis and Artificial Intelligence, National Research University Higher School of Economics, Russian Federation

Mikhail Lugachev –

Professor, Head of Department of Economic Informatics, Lomonosov Moscow State University, Russian Federation

Svetlana Maltseva –

Professor, Head of Department of Innovation and Business in Information Technologies, National Research University Higher School of Economics, Russian Federation

Peter Major –

Vice-chairman, Radiocommunication Advisory Group of International Telecommunication Union (ITU), vice-chairman of the UN Commission on Science and Technology for Development (CSTD), Geneva, Switzerland

Boris Mirkin –

Professor, School of Data Analysis and Artificial Intelligence, National Research University Higher School of Economics, Russian Federation

Vadim Mottl –

Professor, Department of Information Security Management, Tula State University, Russian Federation

Dmitry Palchunov –

Head of Department of General Informatics, Novosibirsk State University, Russian Federation

Panagote (Panos) Pardalos –

Distinguished Professor and University of Florida Research Foundation Professor, Director of Center for Applied Optimization, Department of Industrial and Systems Engineering, University of Florida, USA

Albert Silantsev –

Professor, Department of Information Business Systems, National University of Science and Technology «MISIS», Russian Federation

Victor Taratukhin –

Managing Director, European Research Center for Information Systems (ERCIS), University of Munster, Germany

Mikhail Ulyanov –

Professor, School of Software Engineering, National Research University Higher School of Economics, Russian Federation

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 Radboud 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.

Knowledge visualization: Critique of the St. Gallen School and an analysis of contemporary trends¹

Tatiana A. Gavrilova

*Professor, Head of Department of Information Technologies in Management
Saint Petersburg University
Address: 3, Volkhovsky Pereulok, St. Petersburg, 199004, Russian Federation
E-mail: gavrilova@gsom.pu.ru*

Artem I. Alsufyev

*Assistant Professor, Department of Organizational Behavior and Personnel Management
Saint Petersburg University
Address: 3, Volkhovsky Pereulok, St. Petersburg, 199004, Russian Federation
E-mail: alsufyev@gsom.pu.ru*

Elvira Y. Grinberg

*Doctoral Student, Department of Information Technologies in Management
Saint Petersburg University
Address: 3, Volkhovsky Pereulok, St. Petersburg, 199004, Russian Federation
E-mail: st057454@student.spbu.ru*

Abstract

The purpose of this article is the analysis of leading European research in the field of knowledge visualization from the point of view of the accumulated theoretical base, practice of application, problems, and trends.

The need for digital business transformation for survival in the era of high-speed, mobile intelligent applications and big data has become apparent. However, understanding and interpretation of information can be performed only by humans. Modern managers cope with information “explosion” through visualization. Visualization helps them to understand, to compress and to demonstrate the ocean of numbers, words, and ideas. The number of works devoted to the theme of visualization is growing every year. There are numerous studies on the visualization of networks and relationships, and visualization of communication with a consumer. Fewer articles have been devoted to the visualization of knowledge in the implementation of business practices. At the same time, scientists are examining

¹ This work was supported by the Russian Science Foundation (grant No. 15-18-30048)

one specific area of application of visualization and only a few contribute to the theory of the subject and study it in a versatile manner. The latter include the works of researchers from the University of St. Gallen (Switzerland), which we call in this article the St. Gallen School.

We propose systematization of the following basic stages of research formation of the aforementioned School: 1) the preliminary stage, 2) the stage of empirical data accumulation, and 3) the stage of theory development. The School's contribution to the theory and practice of management was analyzed. Its contribution to theory includes the classification of visualization techniques, a description of visualization use in business, the development of the boundary objects theory, as well as a detailed description of experimental studies. Contribution to business practices means implementation of educational projects and the development of new visual models. The fragmented nature of research is identified: theoretical work is focused on how several visual models influence the implementation of certain business practices; empirical work often describes consulting projects, but do not provide an understanding of how to apply visualization techniques when there is no researcher-consultant.

Based on our analysis of the literature, we demonstrate that the major trend in information processing is focus on knowledge representation based on data, not data as such. The challenging areas related to applied research methods are highlighted as follows: lack of consistency, and lack of distinction between the concepts of "data visualization" and "knowledge visualization". Thus, there is a need to distinguish visualization of knowledge in a separate area of study.

Key words: knowledge visualization, knowledge communication, knowledge management, business trend, infographics, business research.

Citation: Gavrilova T.A., Alsufyev A.I., Grinberg E.Y. (2017) Knowledge visualization: Critique of the St. Gallen School and an analysis of contemporary trends. *Business Informatics*, no. 3 (41), pp. 7–19. DOI: 10.17323/1998-0663.2017.3.7.19.

Introduction

The necessity to transform business for survival in a rapidly changing world of informational overloads and ultra high speeds, mobile intellectual apps and big data has become evident. However, the comprehension and interpretation of information remains a human task. The interest in infographics and visualization assisting rapid understanding is no coincidence. Precisely through visualization, modern managers handle the information "explosion". Visualization helps grasp and comprehend, condense and clearly present an ocean of numbers, words, and ideas. Modern

knowledge management is inconceivable without extensive use of diagrams, graphics and schemas.

From among numerous modern works on information visualization, the authors selected works by researchers from St. Gallen University (Switzerland), dubbing them the St. Gallen School. These works focus on knowledge and data visualization in business. Notably, the leader of the St. Gallen School and head of the Institute for Media and Communication Management, Professor Martin Eppler, began his scientific journey from his dissertation on information overloads.

Visualization is also “social glue”, ensuring communication between different individuals and groups [1]. This means that visual methods work as a medium that sets general frameworks for communication.

This article analyses the works of the St. Gallen School according to the chosen topic, indexed in the Scopus and Web of Science databases. The analysis was conducted with regard to accumulated theoretical knowledge, use, issues, and modern trends in the area of knowledge visualization. Of the presented publications, works from the last 10 years were closely examined, including conference materials with a high citation index and journal articles. The reviewed publications’ extensive contributions to science and business practice makes it possible not only to trace the formative phases of St. Gallen School visualization, but also to draw conclusions about issues and development trends of this scientific area in general.

1. The main characteristics of the works of the St. Gallen School

Notwithstanding the multidimensionality of the works of the St. Gallen School, two main characteristics may be identified (*Figure 1*):

1) the breadth of the research area, which focuses on the study of the specificity of visual method application in business practice;

2) the orientation of research objectives towards knowledge visualization theory construction, which emerges in the authors’ commitment to the expansion of the theoretical base and the classification of known visualization methods.

The first characteristic is the largest, and is distinguished by breadth of coverage and practical orientation. The second characteristic sets the trends in research development in this area.

1.1. Area of research

As was already noted, the majority of the St. Gallen School works are dedicated to visualization in management. The authors show how visualization helps executives solve fundamental business problems, such as business model development, strategizing, multi-level communication, idea generation, team collaboration improvement, increasing knowledge exchange between people and companies, risk management, company core competency analysis, et al. (*Table 1*).

A few empirical studies of knowledge visualization methods are presented in *Table 2*. In particular, articles [7–11, 14–16, 21, 22] present the results of a number of empirical investigations focused on the study of the specificity of using a separate visualization method and its influence on a specific business objective (in

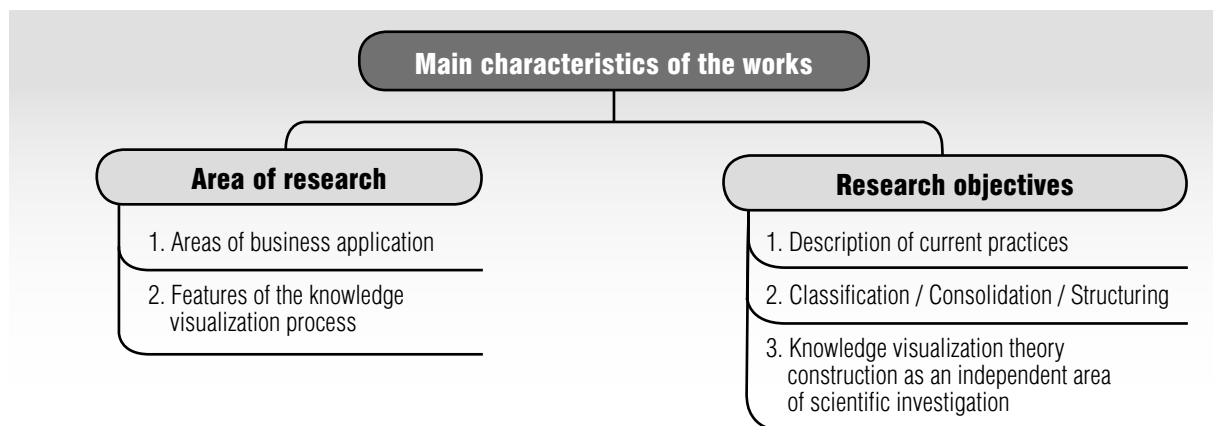


Fig. 1. The main characteristics of the works of the St. Gallen School in the area of knowledge visualization

Table 1.

List of analyzed papers

Ref.	M. Eppler's co-authors	Year	Title
PRELIMINARY PHASE			
[2]	R. Lengler	2007	Towards a periodic table of visualization methods of management
[3]	R.A. Burkhard, et.al.	2007	Visualization Summit 2007: Ten research goals for 2010
[4]	–	2007	Toward a visual turn in collaboration analysis?
[5]	R.A. Burkhard	2007	Visual representations in knowledge management: Framework and cases
EMPIRICAL DATA ACCUMULATION PHASE			
[6]	R.A. Pfister	2012	The benefits of sketching for knowledge management
[7]	F. Hoffmann	2012	Does method matter? An experiment on collaborative business model idea generation in teams
[8]	S. Bresciani, M. Tan	2011	Augmenting communication with visualization: Effects on emotional and cognitive response
[9]	F. Hoffmann, S. Bresciani	2011	New business models through collaborative idea generation
[10]	A. Comi	2011	Assessing the impact of visual facilitation on inter-organizational collaboration: An experimental study
[11]	N. Bischof	2011	Caring for clarity in knowledge communication
[12]	S. Bresciani	2009	The benefits of synchronous collaborative information visualization: Evidence from an experimental evaluation
[13]	M. Aeschmann	2009	A systematic framework for risk visualization in risk management and communication
[14]	K.W. Platts	2009	Visual strategizing. The systematic use of visualization in the strategic-planning process
THEORY CONSTRUCTION PHASE			
[15]	L. McGrath, S. Bresciani	2016	We walk the line: Icons provisional appearances on virtual whiteboards trigger elaborative dialogue and creativity
[16]	E. Alexander, S. Bresciani	2015	Understanding the impact of visual representation restrictiveness on experience sharing: An experimental assessment
[17]	E. Alexander, S. Bresciani	2015	Knowledge scaffolding visualizations: A guiding framework
[18]	S. Kernbach, S. Bresciani	2015	Slip-sliding-away: A review of the literature on the constraining qualities of PowerPoint
[19]	S. Bresciani	2015	The pitfalls of visual representations: A review and classification of common errors made while designing and interpreting visualizations
[20]	A. Comi	2014	Diagnosing capabilities in family firms: An overview of visual research methods and suggestions for future applications
[21]	R.A. Pfister	2014	Beyond projection: Using collaborative visualization to conduct qualitative interviews
[22]	R.A. Pfister, N. Bischof	2014	Best of both worlds: Hybrid knowledge visualization in police crime fighting and military operations
[23]	S. Bresciani	2013	Visualization in management: From communication to collaboration. A response to Zhang

Table 2, works [7–11, 14] pertain to the empirical data accumulation phase, while works [15, 16, 21, 22] to the theory construction phase). Around half of the specified works are dedicated to subjects such as communication and collaboration. They can also be supplemented through decision-making, work with consumers, new product development, etc. Together

with dozens of existing visualization methods, we obtain an extensive field for study. However, this approach is fragmentary and does not give an understanding of the system of interrelations inside this informational field as a whole. The majority of researchers in the area of visualization have unfortunately chosen precisely this approach.

Table 2.

Empirical research of knowledge visualization methods

Ref.	Type of research	Visual method	Business objective	Research method	Result
[7]	Quantitative	Templates and other boundary objects	Creativity and team cooperation	Experiment, 3 teams of managers	Templates are capable of directing teamwork. The use of physical objects and sketches did not yield significant results
[8]	Quantitative	Interactive visualization	Communication and organization	Experiment, a group of professionals	Communication can be directed and improved through interactive visualization, which entails brief and apt expression of ideas
[9]	Quantitative	Business models	Team development of a new business model	Experiment, 3 teams of managers	The use of a digital model of a business template significantly increases perceived collaboration, but reduces creativity
[10]	Quantitative	Interactive visualization, posters	Collaboration between teams of different companies	Experiment, 96 participants	Visual facilitation raises performance and satisfaction from collaboration
[11]	Quantitative	PowerPoint presentation	Knowledge communication at university	Survey, 145 respondents	The CLEAR formula has been developed, along with a corresponding list of control questions
[12]	Quantitative	Interactive visualization	Work with knowledge in teams	Experiment, 131 managers	Interactive visualization exerts a statistically significant positive influence on knowledge exchange and involvement in work, but respondents are unaware of this
[14]	Qualitative	Group of methods	Company strategizing	Action research, 5 companies	Visualization is a process that simplifies strategizing. But it can also have negative consequences if it is not administered properly
[15]	Quantitative	Sketches, icons	Communication using IT	Experiment, 37 pairs of managers	The incompleteness of imagery stimulates social cooperation
[16]	Quantitative	Grid, matrix	Communication in small groups	Experiment, 96 professionals	The average level of visual restrictiveness stimulates experience exchange
[21]	Qualitative	Visualization of organizational competencies	Qualitative interviewing	Individual and group interviews	The integration of design procedures with facilitations using visualization improves mutual understanding and makes it possible to elicit more profound knowledge from the interviewed party
[22]	Qualitative	Knowledge boards	Knowledge exchange in extreme situations	Observations, focus groups, interviews	The combination of virtual and physical visualization is a useful strategy for corporate knowledge management.

Another important aspect is the organizational level the analysis is conducted on. The authors focus on teamwork, but also study personal, organizational and interorganizational cooperation levels through visualization. Additionally, the authors describe the process of development and use of visualization, conduct experiments with the goal of revealing the most effective approaches for specific situations, and devote attention to the degrees of visualization maturity.

The researchers conduct an unbiased analysis of all aspects of visual method use, paying attention to issues that arise during work with visualization. These issues are generally subdivided into three groups: cognitive, emotional and social.

The first group is the most extensive and is often the only one in the other authors' studies. This is why St. Gallen School researchers pay special attention and give a more detailed description of emotional and social issues in the development and perception of visualization. For example, during the development of a certain diagram, the opinions of certain group participants may carry greater weight owing to their status, while experts possessing critical knowledge do not have the opportunity to be heard. Therefore, visualization is considered, among other things, as a process driven by a facilitator. In this case, the level of social tension during group work and the objectivity of the result, as well as the effectiveness of the visual model, depend on the skill of the facilitator.

In general, the authors strive to describe a wide range of visualization methods and related features, however, its boundaries may be so indistinct that it is difficult to identify precise recommendations for work with a certain method or a group of methods. The issues described correlate with specific visualization methods in only one article [24]. In this way, the list of issues without an indication of "where they live" creates a false sensation of the failure of visualization as an effective tool for business. However, the causes

of errors and ways of eliminating them are generally not addressed in the articles.

It should be separately noted that a few works spark theoretical interest, focusing on the study of the combination of different visualization procedures.

1.2. Main research objectives

In earlier works, the authors primarily answer the question "what?": "what is visualization?", "what is associated with it?", "what can be visualized?". They examine a broad range of visualization methods and provide their detailed definitions with striking examples, which makes it possible to treat the St. Gallen School works as a kind of dictionary of visual methods. In this regard, the authors established a clear tradition of discussing visualization methods, specifically their direct combination with illustrations. This may be a stylized image of a specific diagram, or a real-life example from business practice. Such an approach makes it possible to avoid discrepancies in term comprehension and is more than appropriate for this area of research. At the same time, early approaches to classification did not involve a division between data and knowledge visualization.

St. Gallen School researchers extend the theoretical basis of knowledge visualization through close examination of the organizational abilities of family firms [20], new business model development [9], strategic planning [14], and interviewing [21]. They consider visualization to be a form of boundary object. Boundary objects are "plastic enough to adapt to local needs and constraints of the several parties employing them, yet robust enough to maintain a common identity across sites" [26]. This ethnographic understanding of the term "boundary object" was reevaluated in the work [1] illustrating that all engineering schemas are boundary objects. The School authors continue work in this area and consider all visual knowledge models as boundary objects. For example,

an article [9] reveals that the use of a digital business model template significantly increases perceptions of collaboration, but reduces creativity (other examples are provided in *Table 2*). In this way, the authors expand the concept of boundary objects and elaborate it with examples, as well as formulating a number of new research objectives in this area.

2. The formative phases of the St. Gallen School

Based on an analysis of the works of Martin Eppler, three phases of the formation and development of the St. Gallen School can be

identified: preliminary, empirical data accumulation and theory construction phases.

2.1. Preliminary phase

This phase, which lasted from 2002 to 2008, is characterized by several key features. The first is the predominance of theoretical articles of a synthesizing nature. In a number of works, attempts are made at the classification of visualization methods. Complexity level and solvable issues are employed (*Figure 2*) as the basis for classification. The authors also theoretically developed and applied their own typology of complexity levels of work with visualization:

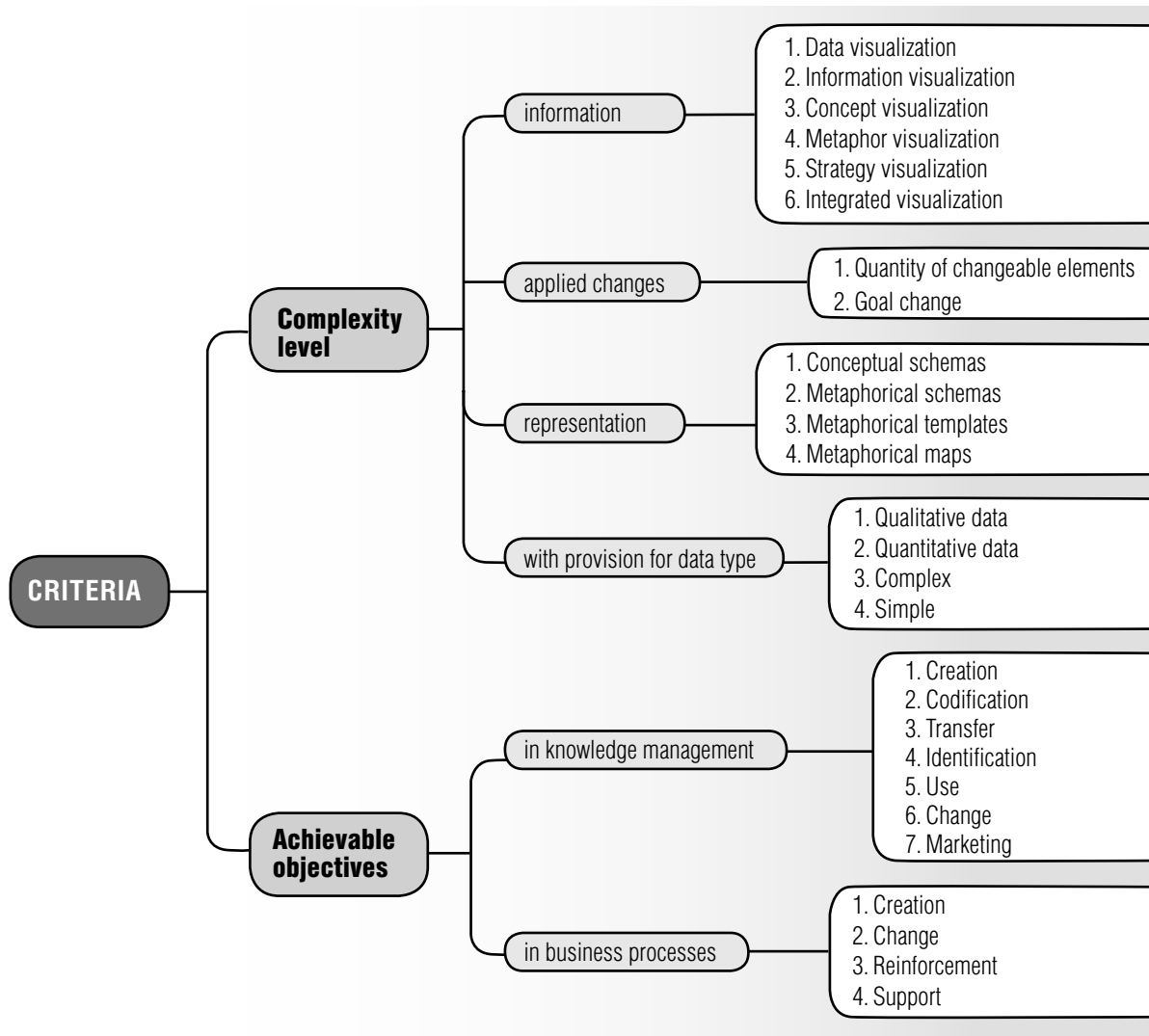


Fig. 2. Criteria for visual method classification

cognitive, emotional, and social. This classification is found in both the preliminary phase and in later works [14, 11, 19].

The second feature is the predominance of the number of conference proceedings over the number of publications in journals. In these proceedings, the researchers are fairly inconsistent in visual model structuring: a form of the periodic table is used [2], answers to the questions “what – why – how?” [5] and “who – what – where – why?” [26]. This inconsistency can be seen in later stages: various kinds of tables are used [16, 27], along with matrixes [28], which have a different structure and cannot be combined. By contrast, at this stage works appear which propose using visualization tools to complement one another. A striking example of this approach is an article [24], in which the author proposes applying, in an integrated manner, concept maps, mind maps, conceptual diagrams and visual metaphors, introducing them sequentially at different phases of the manager’s work.

2.2. Empirical data accumulation phase

From 2009 to 2013 the works of the St. Gallen School are characterized by a predominance of empirical works. The main research method at this phase is the experiment. The main research question that the authors raise is the influence of specific kinds of visualization on certain types of managerial activity (*Figure 3*). Specifically, the authors address how the implementation of visualization affects strategizing and business models, risk management and communication.

The general findings indicate the advantages of visualization: ease of communications and increased work performance of managers or teams of managers. However, the rigid frameworks that it can impose reduce creativity and have a number of negative consequences in social and emotional spheres.

The empirical research contains examples of questionnaires and descriptions of experiments that can be repeated in a different context or as a supplement to achieve other managerial objectives. The features of empirical research methods are characteristic for this scientific area as a whole. Among eight empirical articles in this phase, only one is based on a non-interventive research method. The extensive use of experimental research methods is a shortcoming of the St. Gallen School, as there is no theoretical basis for these experiments.

The article [11] is based on a survey of students and teachers on the particularities of using PowerPoint as a knowledge exchange tool. In this case, the use of closed questions was possible because PowerPoint is a popular tool that all respondents were familiar with.

2.3. Theory construction phase

In publications of 2013 and later the authors, having accumulated empirical material, return to the theoretical questions that were raised at the preliminary phase. However, if in this case the opportunities to combine different visualization methods, such as concept maps, mind maps, and argument maps [24] were investigated, now opportunities to combine different model construction procedures are studied.

In particular, the article [22] is dedicated to the combination of virtual and real visualization. The case description of visualization knowledge in police force crime fighting is notable, but is the only example of a descriptive qualitative research method to date. The authors showed how visual methods are applied to real life in situations when it is necessary to make responsible decisions within constrained timeframes, which is especially relevant for the modern business environment. Moreover, this article examines features of the visual model development procedure. The resulting model is further used for facilitating communications during qualitative research.

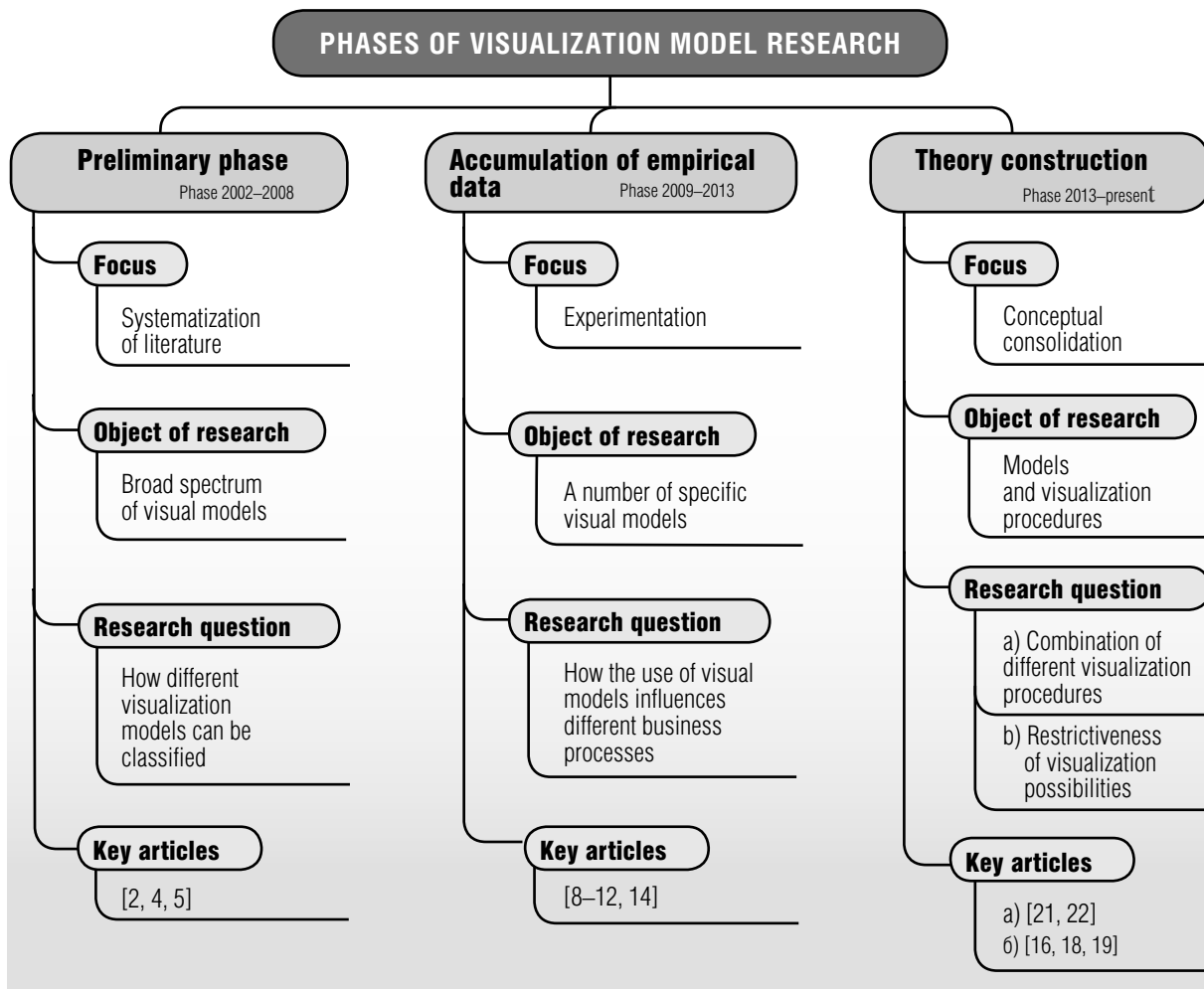


Fig. 3. Research phases of St. Gallen School visual models

3. The current state of knowledge visualization research

3.1. Issues and trends in the works of the St. Gallen School

In the works of Martin Eppler and his co-authors, the difference between knowledge and data visualization is identified as the visualization of qualitative and quantitative data. The conceptualization of the difference between data and knowledge visualization is valuable with regard to the development of a knowledge visualization theory. The data visualization theory is well developed, and the quantity of accumulated knowledge exceeds thousands of works related to dozens of areas. At the same time, knowledge

visualization theory is a new subject that necessitates a different approach to research.

Researchers at the St. Gallen Institute for Media and Communication Management strive to raise the overall level of visual literacy and therefore they are creating a web portal of the same name, www.visual-literacy.org, aimed at eliminating this gap. Additionally, the center’s staff regularly holds seminars and master classes for the business community. One of the main trends in the work of the St. Gallen School is its practical orientation. The authors give multiple detailed recommendations and examples for practitioners. Problems arising from perception of visual forms are closely examined in the work [19] and illustrated in a number of works [8, 11, 18].

3.2. The current state of research

As knowledge visualization is a new and rapidly developing area of research, at its initial phase there is not yet theoretical maturity or a precise classification of terms, and there is still no systematic approach to research.

A classification of visualization methods was the main focus of the works of the St. Gallen School in the preliminary phase. However, after the creation of an entire range of classifications, the authors set aside the subject for five years and returned to it in earnest only after accumulating extensive work experience with commercial companies and government organizations. *Figure 4* presents different approaches that the authors undertook in order to unite visualization methods into specific groups and identify links between these groups. An entire range of factors may be noted as the cause of the failures of the first classification attempts. First of all, at the time the articles were written, terms that designated different types of diagrams, graphs, matrices, and other graphics were largely undefined. For example, the term “knowledge map” in certain works is defined as a kind of analogue of a locality map, making it possible to determine where certain knowledge can be found while other authors use this term to designate the totality of diagrams displaying knowledge. Second, neither in the St. Gallen School works, nor in those of other authors studying visualization, can we witness the use of a powerful knowledge-structuring tool such as ontology. Ontology or specification conceptualization [29] is a hierarchical model of a subject field possessing a sound mathematical and programmatic foundation. Without ontology, it is fairly difficult to construct any kind of classification, particularly if the number of elements at the lower level of the hierarchy exceeds one hundred [2]. As regards the rules of ontology construction, it is apparent that division into 4-7 groups is insufficient; it is necessary to add one or two more hierarchical levels.

Returning to the issue of the division of research boundaries in data and knowledge visualization, we propose constructing separate ontologies for both the former and the latter. In this way, the number of elements will be reduced, while the classification criteria will become more lucid.

Research of visual methods in the current phase of development theory and practice are frequently focused on the features of a specific visual method. However, the number of visualization methods is not only great at present, but continues to rise. In keeping with the logic proposed, following the appearance of each new visualization method, researchers must explore the possibilities of using it in different business spheres, as well as the particularities of combining it with other methods. Such work appears to be excessively labor-intensive and to lead to the blurring of the subject of research instead of focusing on and forming a synthesized theory. It has evidently become imperative to redefine the structure of research in the area of study under discussion.

3.3. Features of knowledge visualization research

Currently, the majority of empirical research in the area of knowledge visualization is based on interventive methods, among which the experiment predominates. A similar strategy during visualization research is observed by many authors [30, 31]. It is likely that this is due to the fact that quantitative research entails the provision of a set of closed questions, which is only feasible if the respondent is well acquainted with the terminology of the subject area. As was previously mentioned, a uniform classification of visualization methods has not formed in scientific literature. Moreover, practitioners may work with visualization without realizing precisely which method they are employing and which business objectives this method can be applied to. Thus, the research methods used in this phase of the area of knowledge visualization make it impossible

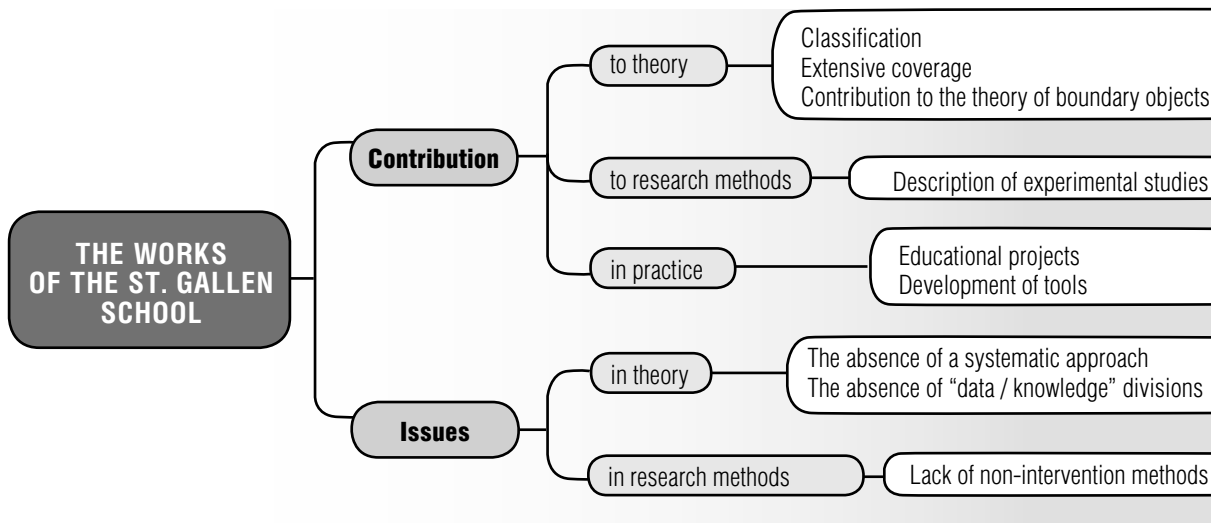


Fig. 4. Issues and trends in the study of visualization methods conducted by St. Gallen School

to reveal and observe the phenomenon in question in all of its manifestations.

In this case, the most appropriate method is the observation method (without interference), which was employed in only one study [32]. However, the study was conducted not in the business environment, but using the work of a police station as an example. As a result, the scientific community currently possesses an extremely restricted set of data on what is actually happening in visualization practice in business. This situation raises the issue of the importance and prospects of further study of this field using descriptive methods with the aim of identifying prevailing approaches to visualization in business [33–35].

Conclusion

The aim of this article was to analyze the leading European research in the area of knowledge visualization with the goal of identifying modern trends for their effective use in scientific and practical activity.

The article identifies the main phases of research formation of the so-called St. Gallen School: the preliminary, empirical data accumulation and theory construction phases.

Problematic areas are also identified, associ-

ated with applied research methods, the absence of a systematic structure and the inadequate differentiation of the concepts of “data visualization” and “knowledge visualization”.

As a whole, the works of the St. Gallen School made a significant contribution to the study of the matter of visualization and identified promising areas of research. Its contribution to theory includes classification of visualization methods, description of the use of visualization in business, development of a theory of boundary objects, as well as a detailed description of experimental research. Its contribution to business practice lies in its initiation of educational projects and development of new visual models.

Thus, there is a need to formulate a general theory of knowledge visualization in business, to create special tracks and seminars at international conferences of varying levels (for example, the Academy of Management, IFKAD, VISUAL, IEEE Pacific Visualization Symposium, et al.), as well as to formulate corresponding internet platforms. To effectively discuss the issues surrounding the creation of a general framework for methods, metrics and theoretical boundaries necessary for the development of the theoretical area, the development of a single ontology of visual methods seems particularly significant [36].

Among development trends in this area, the efforts of researchers to enlarge the scope of knowledge visualization in use by different business practices have been noted. In information processing (for example, when working with Big Data), the use of visualization as a clear, condensed description of data makes it possible to transition to more profound levels of interpretation. Currently, researchers note both the presence of diverse visualization software even for graphs with thousands of peaks and millions of connections [37] and the impossibility of depicting business processes using some uniform notation [38]. Thus, the crucial role in the processing of available data into knowledge, making it possible to take managerial decisions, is played by manager's expertise in synthesizing,

interpreting and systematizing information. The need to allocate knowledge visualization to a separate area of research is becoming apparent. It is worth taking into account that methods and approaches that were applied in the study of data visualization may be only partially applied to this situation. Because the field of study has not reached the necessary level of maturity, the use of quantitative methods or a deductive approach is currently underproductive. It is necessary, first and foremost, to develop an integral theoretical basis for further research, and then formulate empirically supported hypotheses.

This review makes it possible to identify new opportunities for research and focuses attention on the necessity of a systematization of existing knowledge. ■

References

1. Henderson K. (1991) Flexible sketches and inflexible data bases: Visual communication, conscription devices, and boundary objects in design engineering. *Science, Technology, & Human Values*, vol. 16, no. 4, pp. 448–473.
2. Lengler R., Eppler M.J. (2007) Towards a periodic table of visualization methods for management. *Proceedings of 2007 IASTED International Conference on Graphics and Visualization in Engineering (GVE 2007)*, Clearwater, Florida, USA, 3–5 January 2007, pp. 83–88.
3. Burkhard R.A., et.al. (2007) Visualization Summit 2007: Ten research goals for 2010. *Information Visualization*, no. 6, pp. 169–188.
4. Eppler M.J. (2007) Toward a visual turn in collaboration analysis? *Building Research & Information*, vol. 35, no. 5, pp. 584–587.
5. Eppler M.J., Burkhard R.A. (2007) Visual representations in knowledge management: framework and cases. *Journal of Knowledge Management*, vol. 11, no. 4, pp. 112–122.
6. Pfister R.A., Eppler M.J. (2012) The benefits of sketching for knowledge management. *Journal of Knowledge Management*, vol. 16, no. 2, pp. 372–382.
7. Eppler M.J., Hoffmann F. (2012) Does method matter? An experiment on collaborative business model idea generation in teams. *Innovation: Management, Policy & Practice*, vol. 14, no. 3, pp. 388–403.
8. Bresciani S., Tan M., Eppler M.J. (2011) Augmenting communication with visualization: Effects on emotional and cognitive response. *Proceedings of IADIS ICT, Society and Human Beings 2011 (ICT 2011)*, Rome, Italy, 20–26 July 2011, pp. 109–121.
9. Eppler M.J., Hoffmann F., Bresciani S. (2011) New business models through collaborative idea generation. *International Journal of Innovation Management*, vol. 15, no. 6, pp. 1323–1341.
10. Comi A., Eppler M.J. (2011) Assessing the impact of visual facilitation on inter-organizational collaboration: An experimental study. *Journal of Universal Computer Science*, vol. 17, no. 10, pp. 1434–1454.
11. Bischof N., Eppler M.J. (2011) Caring for clarity in knowledge communication. *Journal of Universal Computer Science*, vol. 17, no. 10, pp. 1455–1473.
12. Bresciani S., Eppler M.J. (2009) The benefits of synchronous collaborative information visualization: Evidence from an experimental evaluation. *IEEE Transactions on Visualization and Computer Graphics*, vol. 15, no. 6, pp. 1073–1080.
13. Eppler M.J., Aeschmann M. (2009) A systematic framework for risk visualization in risk management and communication. *Risk Management*, vol. 11, no. 2, pp. 67–89.
14. Eppler M.J., Platts K.W. (2009) Visual strategizing. The systematic use of visualization in the strategic-planning process. *Long Range Planning*, vol. 42, no. 1, pp. 42–74.

15. McGrath L., Bresciani S., Eppler M.J. (2016) We walk the line: Icons provisional appearances on virtual whiteboards trigger elaborative dialogue and creativity. *Computers in Human Behavior*, no. 63, pp. 717–726.
16. Alexander E., Bresciani S., Eppler M.J. (2015) Understanding the impact of visual representation restrictiveness on experience sharing: An experimental assessment. *Journal of Visual Languages and Computing*, no. 31, pp. 30–46.
17. Alexander E., Bresciani S., Eppler M.J. (2015) Knowledge scaffolding visualizations: A guiding framework. *Knowledge Management & E-Learning: An International Journal*, vol. 7, no. 2, pp. 179–198.
18. Kernbach S., Bresciani S., Eppler M.J. (2015) Slip-sliding-away: A review of the literature on the constraining qualities of PowerPoint. *Business and Professional Communication Quarterly*, vol. 78, no. 3, pp. 292–313.
19. Bresciani S., Eppler M.J. (2015) The pitfalls of visual representations: A review and classification of common errors made while designing and interpreting visualizations. *SAGE Open*, vol. 5, no. 4, pp. 1–14.
20. Comi A., Eppler M.J. (2014) Diagnosing capabilities in family firms: An overview of visual research methods and suggestions for future applications. *Journal of Family Business Strategy*, vol. 5, no. 1, pp. 41–51.
21. Comi A., Bischof N., Eppler M.J. (2014) Beyond projection: using collaborative visualization to conduct qualitative interviews. *Qualitative Research in Organizations and Management: An International Journal*, vol. 9, no. 2, pp. 110–133.
22. Eppler M.J., Pfister R.A. (2014) Best of both worlds: Hybrid knowledge visualization in police crime fighting and military operations. *Journal of Knowledge Management*, vol. 18, no. 4, pp. 824–840.
23. Eppler M.J., Bresciani S. (2013) Visualization in management: From communication to collaboration. A response to Zhang. *Journal of Visual Languages and Computing*, no. 24, pp. 146–149.
24. Eppler M.J. (2006) A comparison between concept maps, mind maps, conceptual diagrams, and visual metaphors as complementary tools for knowledge construction and sharing. *Information Visualization*, vol. 5, no. 3, pp. 202–210.
25. Star S.L., Griesemer J.R. (1989) Institutional ecology, translations, and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907–39. *Social Studies of Science*, vol. 19, no. 3, pp. 387–420.
26. Eppler M.J., Burkhard R.A. (2004) *Knowledge visualization: Towards a new discipline and its fields of application*. Working Paper of NetAcademy of Knowledge Media. St. Gallen.
27. Eppler M.J., Hoffmann F. (2011) Challenges and visual solutions for strategic business model innovation. *Strategies and communications for innovations: An integrative management view for companies and networks* (eds. M. Hulsmann, N. Pfeffermann). Berlin Heidelberg: Springer, pp. 25–36.
28. Eppler M.J., Pfister R.A. (2010) Drawing conclusions: Supporting decision making through collaborative graphic annotations. Proceedings of the *IEEE 14th International Conference on Information Visualisation (IV 10)*, London, 26–29 July 2010, pp. 369–374.
29. Gruber T.R. (1993) A translation approach to portable ontologies. *Knowledge Acquisition*, vol. 5, no. 2, pp. 199–220.
30. Zhao M., Hoefler S., Dahl D.W. (2009) The role of imagination-focused visualization on new product evaluation. *Journal of Marketing Research*, vol. 46, no. 1, pp. 46–55.
31. Cheema A., Bagchi R. (2011) The effect of goal visualization on goal pursuit: Implications for consumers and managers. *Journal of Marketing*, vol. 75, no. 2, pp. 109–123.
32. Eppler M.J., Oste H.F., Bresciani S. (2013) An experimental evaluation on the impact of visual facilitation modes on idea generation in teams. Proceedings of the *IEEE 17th International Conference on Information Visualisation (IV 13)*, London, 15–18 July 2013, pp. 339–344.
33. Gavrilova T., Alsufyev A., Yanson A.-S. (2014) Modern notation of business models: Visual trend. *Foresight—Russia*, vol. 8, no. 2, pp. 56–70.
34. Yau N. (2013) *Iskusstvo vizualizatsii v biznese: Kak predstavit' slozhnuyu informatsiyu prostymi obrazami* [Art of visualization in business: How to present complex information using simple images]. Moscow: Mann, Ivanov and Ferber (in Russian).
35. Cybulski J.L., Keller S., Nguyen L., Saundage D. (2015) Creative problem solving in digital space using visual analytics. *Computers in Human Behavior*, no. 42, pp. 20–35.
36. Gavrilova T.A., Kudryavtsev D.V., Leshcheva I.A., Pavlov Ya.Yu. (2013) Ob odnom metode klassifikatsii vizual'nykh modeley [On a method of visual models classification]. *Business Informatics*, no. 4 (26), pp. 21–34 (in Russian).
37. Kolomeychenko M.I., Chepovskiy A.M. (2014) Vizualizatsiya i analiz grafov bol'shikh razmerov [Huge graph visualization and analysis]. *Business Informatics*, no. (30), pp. 7–16 (in Russian).
38. Fiodorov I.G. (2016) Overcoming expressiveness deficit of business process modeling languages. *Business Informatics*, no. 3 (37), pp. 62–71.

Modeling the influence of foreign investments on food security based on panel cointegration models

Elena D. Kopnova

*Associate Professor, Department of Statistics and Data Analysis
National Research University Higher School of Economics
Address: 20, Myasnitskaya Street, Moscow, 101000, Russian Federation
E-mail: ekopnova@hse.ru*

Lilia A. Rodionova

*Associate Professor, Department of Statistics and Data Analysis
National Research University Higher School of Economics
Address: 20, Myasnitskaya Street, Moscow, 101000, Russian Federation
E-mail: lrodionova@hse.ru*

Abstract

The paper studies the problems of food security as the basis for the stability of the country's economic development. Despite the economic growth of most countries in the world, the problems of hunger continue to be acute in developing countries. This undermines the food security of these countries and can pose a threat to world peace as a whole. Numerous studies show that foreign investment plays an important role in the creation of food security, but this effect is ambiguous. The most common methods of analysis in recent researches are descriptive analysis and panel data models.

In this article, our approach to analyzing the impact of foreign investment is based on panel cointegration models; the interpretation of the results uses the impulse response function based on an error correction model. Countries of North Africa are considered as an example of the implementation of this approach. Data from the World Bank and the Food and Agriculture Organization for 1991–2014 for seven countries (Algeria, Egypt, Libya, Morocco, Sudan, Tunisia, Western Sahara) were selected for analysis. In this paper, we have also tested the methodology for selecting the food security indicator based on analysis of the matrix of cointegration relations.

The results show that foreign investment has a significant long-term impact on food security, but in the short term no effect was detected. The proposed modeling methodology can be extended to any region of the world to monitor and evaluate the effectiveness of the current economic policy to combat hunger and poverty.

Key words: modeling of economic systems, panel cointegration, food security, foreign investment, Pedroni panel tests, panel tests of unit root.

Citation: Kopnova E.D., Rodionova L.A. (2017) Modeling the influence of foreign investments on food security based on panel cointegration models. *Business Informatics*, no. 3 (41), pp. 20–29. DOI: 10.17323/1998-0663.2017.3.20.29.

Introduction

One of the important goals of humanity is sustainable development of society. This is difficult to imagine without progress in the eradication of hunger and malnutrition [1]. Despite the fact that enough food is produced in the world to feed everyone, the Food and Agriculture Organization (FAO) estimates that in 2015 about 792.5 million people continued to suffer from chronic hunger¹. The vast majority of the hungry lived in developing countries. According to the FAO, food security is defined by the conditions “when all people, at all times, have physical, social and economic access to sufficient, safe and nutritious food that meets their dietary needs and food preferences for an active and healthy life” [2]. Today, food security is not only one of the basic human needs and basic human rights, but it plays an important role in a country’s economic development and the preservation of world peace [3]. International organizations under FAO auspices have adopted an agreement on ensuring food security in developing countries to achieve Zero Hunger by 2030. Investments in agricultural development amounting on average to US\$ 265 billion a year are one of the necessary measures [4]. Such investment flows require constant monitoring of the situation and the evaluation of possible risks and the effectiveness of the policy based on thorough consideration of general long-term indicators for food security and socio-economic development. Thus, the need to monitor the current situation of food security in developing countries, the development of a methodology for monitoring economic policy and obtaining quantitative estimates of the long-term effects of investment determine the relevance of the selected topic of this study. Note that the preliminary results of this study were

published in the series of “Economics” preprints of the National Research University Higher School of Economics [5].

1. The impact of investment on food security: the theoretical background

In the literature on food security, different concepts to explain the determinants of food security are highlighted: modernization, economic dependency, urban bias, neo-Malthusian population pressure, ecological evolutionary processes and militarism among others [6]. Let us dwell on some of them.

In **modernization theory** internal sources of economic development play an important role. Domestic investment and the development of education are a reliable basis for industrialization and the cultural transformation of society, which leads to economic growth and the creation of stronger institutions that, in turn, contribute to improving the overall welfare of the population. Investing in human capital and increasing the level of education creates a workforce with modern motivation, skills and a high degree of adaptation to modern technologies, which is an important component of economic growth [7]. International trade and investment stimulate economic development, promote the spread of new technologies and methods, and improve social security. Foreign studies also show that economic development, industrialization, education and urbanization can improve the quality of life of the population, and economic growth has a positive effect on the food supply in the country and life expectancy in general. According to modernization theory, international capital flows and technology contribute to economic growth, increase the incomes of all population groups and, as a result, reduce the risks of hunger in the country [8, 9]. In the theory

¹ Statistical data are available at FAO website: <http://faostat3.fao.org/download/I/IC/E>

of modernization, an important role in the formation of food security is also assigned to political institutions. Strong political institutions, and the democratization of society, contributes to economic and social development, inhibits corruption, encourages the growth of public investment in nutrition, education and health, provides political stability, improves the quality of life, all of which significantly affects food security [10].

In **economic dependency theory**, international trade and foreign investment have a negative impact on social security. Studies have shown that export dependence reduces economic growth and the quality of life, cuts total food supply, and contributes to child mortality [11]. Foreign direct investment (FDI) has a negative impact on food security for several reasons. Multinational corporations repatriate most of their profits and prevent the formation of domestic firms, which leads to a reduction of domestic economic growth. Foreign investment is largely focused on advanced technology and the creation of a smaller number of jobs, although highly paid. From this come inequality of income and poverty. Foreign investment is mainly concentrated in export industries and has weak links with the internal market, which also has a negative impact on domestic production. In the short term, FDI can create economic growth, but the long-term effects of the penetration of foreign capital are negative [12].

These concepts of food security are reflected in the research related to Africa [13]. Study of the FDI effect in African countries shows that countries with well-developed financial markets benefit significantly, and the problem of excessive dependence of the economy on FDI exists. In our study, the effect of FDI will be evaluated by the example of the countries of North Africa, where some progress has

already been achieved in the field of combating hunger.

2. Review of existing methods for modeling food security factors

In the last decade, problems of food security have been actively explored and significant results have been obtained, especially by the FAO. Only in a few works, however, is panel data analysis used as an econometric tool for the full sample of developing countries [9, 12, 14]. In most works, the authors confine themselves to descriptive data analysis. If regional comparisons are carried out, for example, in African countries in conditions of limited sampling, and modeling methods using panel time series are used, then a number of problems in estimating models with fixed and random effects arise [15, 16]. Nonstationarity of time-series is one such problem. In our work, we will try to develop a methodology for modeling the determinants of food security based on panel cointegration [17]. This methodology makes it possible to take into account the problems of nonstationarity in modeling the determinants of food security using panel time series. This approach will be described in more detail in the “Proposed research methodology” section.

3. Problems of measuring food security

For a comprehensive assessment of food security in countries around the world since 1996, the FAO uses a set of 30 indicators² covering various aspects of the problem and classifying food security into four dimensions – availability, access, utilization, stability. The indicators of availability describe the value of food production and productivity, level of stocks, adequacy of average dietary energy supply, etc. Access

² A list of all food safety indicators, definitions and detailed calculation methodology are available on the FAO website: <http://www.fao.org/economic/ess/ess-fs/ess-fadata/en/#.V2Qip-pv7IV>

covers the economic opportunity to purchase food and transport it (percent of paved roads over total roads, road density, GDP per capita, etc.). Utilization evaluates food consumption in terms of nutritional standards for calories, protein and trace elements. Stability characterizes the situation with the adequacy of food in the country in different periods (cereal import dependency ratio, domestic food price volatility, etc.).

Such a wide range of indicators is necessary in order to assess food security in various aspects. Food security is not just enough food. For example, despite the declining share of hungry people in Africa and the problem of the lack of energy value of nutrition, a number of problems related to the quality of the diet remain. In sub-Saharan Africa, the phenomenon of “hidden hunger” is observed when, due to lack or inadequate intake of nutrients, various manifestations of malnutrition (anemia among pregnant women and children under 5 years of age, vitamin A deficiency, children under 5 years of age who are underweight, etc.) arise.

Note that empirical studies on food security use different indicators to measure food security. Among all the ones proposed, it can distinguish the average energy value of the diet, the percentage of newborns weighing less than 2500 grams at birth, the proportion of foods consumed containing starch [19], daily protein intake [6, 12], the proportion of undernourished [14], etc. It can be concluded that among researchers there is no single opinion and substantive justification for answering the question which of the 30 indicators should be chosen as a measure of food security. In our study, we propose a methodology for selecting a food security indicator based on an analysis of the matrix of cointegration links, described in the “Proposed research methodology” section.

4. Information base of the research

The countries of North Africa were considered as an example of the implementation of the proposed methodology. Data from the FAO³ and the World Bank⁴ for seven countries of North Africa (Algeria, Egypt, Libya, Morocco, Sudan, Tunisia, Western Sahara) covering the period 1991–2014 were used in the work. For analysis of the food security of the countries, a number of indicators were chosen that characterize the considered problem from different angles: food security indicators, characteristics of economic development and resources, international trade indicators, indicators of investment in agriculture and indicators of the financial system. A more detailed description of variables, descriptive statistics of indicators are given in the pre-print [5]. Note that the problem of skipping data for some indicators and countries (especially for Western Sahara and Sudan) exists.

The change over time of some indicators that characterize food security in the countries studied in different aspects is presented in *Figure 1*. The amount of food consumed, expressed by average dietary energy supply adequacy, exceeds 100% of the daily ration. This suggests that the population of North Africa is on average not starving. The prevalence of anemia among pregnant women, indicating the quality of food consumed, declined significantly during the study period and was less than 35% for the countries under consideration after 2011. It should be noted that a rather large variation of cereal import dependency ratio exists (which has a negative impact on food security): Libya (more than 90%) demonstrates the highest figures, Egypt is the lowest (less than 40%). In general, a positive trend is observed both in quantity and as a diet of the population. This shows that measures to improve food security

³ <http://faostat3.fao.org/download/I/IC/E>

⁴ <http://data.worldbank.org/indicator>

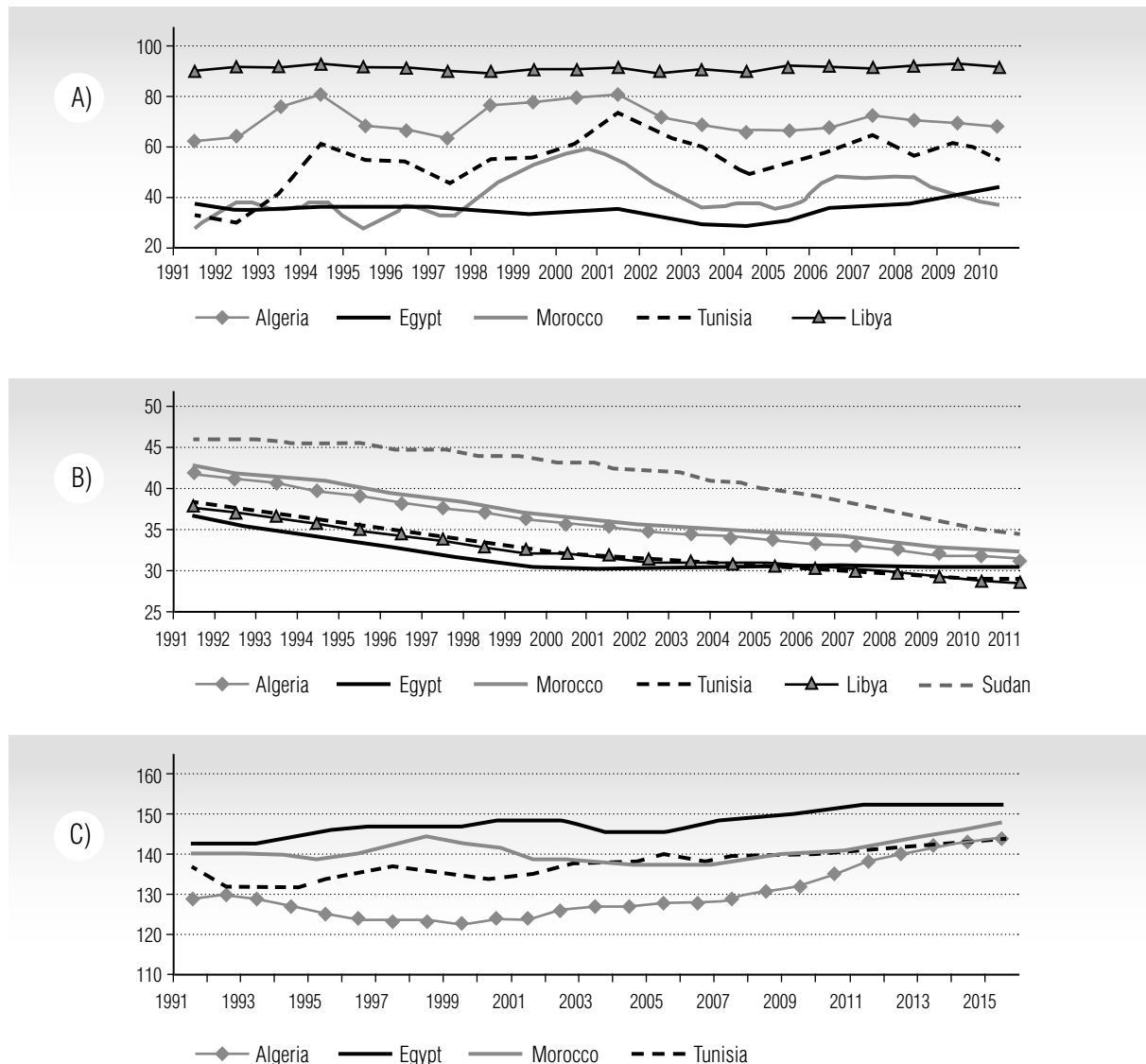


Fig. 1. Evolution of some food security indicators (calculated based on FAO statistical data)

- A) Cereal import dependency ratio, %
- B) Prevalence of anemia among pregnant women, %
- C) Average Dietary Energy Supply Adequacy, %

in the countries of North Africa are effective and this experience is useful for study.

5. Proposed research methodology

The specificity of the available data for Africa (a lot of omissions, a limited sample of countries for the selected region, nonstationarity of time series) do not allow us to evaluate the clas-

sical panel data models. In this paper, to test the theoretical assumptions about the effect of foreign investment, we will use an approach based on panel cointegration [17] and implemented in this study in several stages.

Stage 1. Stationarity of processes. To study the long-term relationships between the indicators being analyzed, it is first necessary to check the time series for stationarity, taking

into account the panel structure of the data. For each of the indicators, panel unit-root tests were carried out at 10% significance level. We considered five tests: the Levin-Lin-Chu unit-root test (LLC), the Breitung unit-root test, the Im-Pesaran-Shin unit-root test (IPS), Fisher-type unit-root tests based on augmented Dickey-Fuller tests (ADF-Fisher) and based on Phillips-Perron tests (PP-Fisher), all three modifications (with a constant, trend, constant and trend) [16, 19, 20]. All tests have as the null hypothesis that all the panels contain a unit root. The results of testing for one of the characteristics of food security – cereal import dependency ratio (CIDR) in the countries of North Africa are given as an example. The statistics and the corresponding probabilities (P-value) are given in *Table 1*⁵.

These results allow us to conclude that the CIDR is the realization of a nonstationary process, integrated of the first order (I(1)). A similar testing approach was used for the other 40 indicators of the social, economic and financial spheres of countries. Among the characteristics of food security, 10 out of 30 indicators also turned out to be a realization of I(1) and were taken for further analysis. These results also

indicate that the processes (different aspects of food security, economic characteristics and foreign investment) are in development, which should contribute to progress in improving the lives of the poor. Nonstationarity of processes is also necessary for further research of long-term relationships between indicators.

Stage 2. Panel cointegration test. To further analyze long-term relationships, it is necessary to determine the cointegration of processes. For time series that were a realization of I(1), Pedroni cointegration tests were applied [21]. Seven different tests with a null hypothesis of no cointegration in non-stationary panels were performed. Test statistics were grouped into two categories: group-mean statistics that average the results of individual country test statistics and panel statistics that pool the statistics along the within-dimension [15]. In *Table 2*, the results of the Pedroni cointegration tests for CIDR and FDI (inflow) are given: the test statistics and their corresponding probabilities (P-value). The null hypothesis of no cointegration is rejected at 10% significance level. The decision about the cointegration of processes was accepted if the hypothesis was rejected for 3 or more statistics out of 7.

Table 1.

**Panel unit-root tests
for cereal import dependency ratio (CIDR)**

Indicators		CIDR		ΔCIDR	
Type of statistics	Test	Statistics	P-value	Statistics	P-value
Panel statistics	LLC	-1.02	0.16	-3.02	0.001
Group-mean statistics	IPS	-2.94	0.002	-5.17	0.000
	ADF-Fisher	24.71	0.006	44.08	0.000
	PP-Fisher	18.47	0.048	45.06	0.000

⁵ The decision about the stationarity was accepted, if the hypothesis of a unit root was rejected at 10% level for all modifications for three or more statistics. For conflicting testing results, priority was given to group average statistics in favor of the reliability of the conclusion

Table 2.

Pedroni cointegration tests for cereal import dependency ratio (CIDR) and inflow of FDI

Variables		CIDR and FDI	
Test statistics		Statistics	P-value
Panel statistics	v-St	2.08	0.02
	rho-St	-1.31	0.09
	PP-St	-2.17	0.02
	ADF-St	-3.13	0.001
Group-mean statistics	rho-St	-0.14	0.44
	PP-St	-1.36	0.08
	ADF-St	-1.61	0.05

Based on the test results, it can be concluded that the CIDR and FDI processes analyzed are cointegrated and further analysis of long-term relationships is possible for them.

Stage 3. Selecting a food security indicator. Based on Pedroni cointegration tests, the cointegration relationships between the other indicators were analyzed. Based on analysis of the test statistics, the matrix of cointegration links was compiled in which the maximum number of statistics indicating cases of rejection of the null hypothesis of no cointegration at the 10% significance level is given [5]. The calculations so obtained made it possible to reveal the distribution of mutual influences of the indicators analyzed and to determine those of them that accumulate the greatest number of connections and can be used as integral indicators in further analysis. The indicator that accumulates the largest number of links can be viewed as a proxy of food security. Cereal import dependency ratio (CIDR) was selected as such a characteristic in our study. It has 33 links to other indicators from 40 analyzed. Access to improved water sources also accumulates a large number of cointegration links (31 out of 40).

Stage 4. The long-run relationships. The expected long-term relations were analyzed for cointegrated variables. Each cointegrating relation contained three endogenous variables and a constant.

$$y_{1it} = \alpha_i + \beta_1 y_{2it} + \beta_2 y_{3it} + \varepsilon_{it}, \quad (1)$$

where $y_{1it}, y_{2it}, y_{3it}$ – the values of the indicators for country i in year t ;

ε_{it} – the error of cointegration;

α_i – coefficients corresponding to individual effects;

β_k – components of cointegrated vectors.

Multivariate cointegration was also tested by the Fisher-Johansen test.

The evaluation of the parameters for cointegrating equations was carried out using Dynamic Ordinary Least Squares with automatic selection of the auxiliary lags number of y_{1it} applying Schwarz criterion. Standard errors were calculated using the New-West formula adjusted for heteroscedasticity and autocorrelation. The estimation results of equation (1) for CIDR and FDI (inflow and outflow) are presented in Table 3. Analysis of the regression residuals of the model (1) was carried out using

panel unit root tests, which indicate about cointegration between the investigated variables and the suitability of the model.

Stage 5. The short-run relationships. To estimate the short-term relationships, error correction model based on panel data was used in the form of the equation:

$$\Delta y_{1it} = \alpha_i^{(\Delta)} + \beta_1^{(\Delta)} \Delta y_{1it-1} + \beta_2^{(\Delta)} \Delta y_{2it-1} + \beta_3^{(\Delta)} y_{3it-1} + \delta_1^{(\Delta)} ECM_{it-1} + \varepsilon_{it}^{(\Delta)}, \quad (2)$$

where $\Delta y_{jit} = y_{jit} - y_{jit-1}$ $j = 1, 2, 3$ – increments of the related indicators;

$\alpha_i^{(\Delta)}$ – coefficients characterizing the individual effects;

ECM_{it-1} – the adjustment mechanism equilibrium, which are the residuals of the cointegrating equations (1);

$\delta_1^{(\Delta)}$ – coefficient which characterizes the recovery rate of the equilibrium state for y_{1it} ;

$\beta_k^{(\Delta)}$ – regression coefficients;

$\varepsilon_{it}^{(\Delta)}$ – regression error.

An estimation of the short-term relation parameters was carried out according to a method of Panel Least Squares, followed by

residuals diagnostics. The residuals of the model were checked for autocorrelation with the Wooldridge test for serial correlation in panel data. The estimation results of the model (2) are also given in *Table 3*.

Note that the coefficient at ECM, characterizing the correction to long-term equilibrium, is significant at the level of 0.01 and has the correct sign. It shows that the correction of CIDR in its deviation from the equilibrium trajectory with potentially possible short-term shocks occurs about a year and two months later (1/0.7059).

6. Results and prospects of the study

Estimation results of the models (1)–(2) and significant estimates of model parameters confirm the expected long-term and short-term relationships. At 5% significance level we can assume that there is a direct long-term relationship between CIDR and FDI inflows. The relationship between CIDR and FDI outflow was not detected. With the increase in FDI inflows in the country, CIDR increases in the long

Table 3.

The results of evaluation of long- and short-term relation between CIDR and FDI

Cointegrating equation (1)		Error correction model (2)	
Variable	Coefficient	Variable	Coefficient
FDI inflows, FDI	0.004**	$\Delta CIDR(t-1)$	0.2902***
		$\Delta FDI(t-1)$	0.0005
FDI outflows, FDO	-0.001	$\Delta FDO(t-1)$	0.0001
		$ECM(t-1)$	-0.7059***
N=87; R ² =0.96; DW=1.741#; J-B-test $\chi^2=0.144##$		N=87; R ² =0.28; DW=1.81; J-B-test $\chi^2=0.21$	

** Indicate the p-value is significant at 5% (p<0.05);

*** Indicate the p-value is significant at 1% (p<0.01).

Indicate the value of Durbin-Watson statistics for testing the model residuals for first-order autocorrelation.

Indicate the value of Jarque-Bera statistics for testing the model residuals for normality.

term. This effect can be regarded as negative, as food security deteriorates and theoretical backgrounds for the concept of economic dependence are confirmed. The long-term elasticity coefficient was 0.088. So, an increase in FDI by 1% is accompanied by an increase in CIDR of 0.088%. Speaking of short-term effects, the results show that the effect of FDI in the short term are not significant (model (2)).

The significant and positive estimate at the lags of the absolute growth of $\Delta CIDR(t-1)$ in equations (2) indicates the inertial variability of CIDR and this should be taken into account in forecasting. This fact, in addition to a low coefficient of determination in error correction model ($R^2 = 0.28$), points to the need for further analysis in identifying the factors of CIDR dynamics.

Based on the estimates of model parameters for the food security indicator impulse response functions were calculated (Figure 2). This allows an assessment of the impact of a single error change (shock) at a rate of one standard deviation and characterizes the mechanisms of short-term adjustments to the equilibrium state.

CIDR initially increases in response to shocks of FDI inflows, but after 4 years there is a loss of reaction (to zero). After 5–6 years an indicator reduction is observed which in the context of food security is a positive effect because it slightly decreases CIDR. The opposite reaction is observed in response to shocks of FDI outflows. In this case, the CIDR increases gradually in 6–7 years by an average of 1.6% and then remains at the new, higher level. These results confirm the long duration of the analysed processes. The detection of the

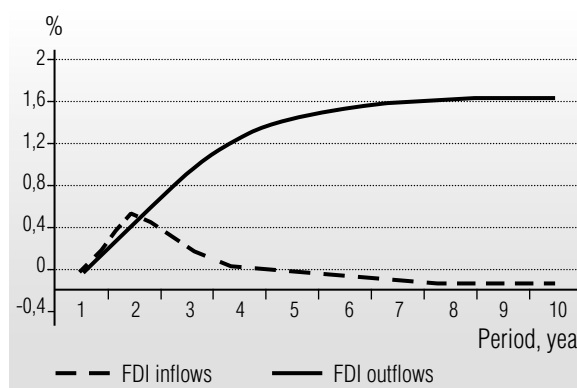


Fig. 2. Responses of CIDR on the impulse FDI

long-term effects of FDI indicates that investments will only be seen as significant over the long term. This fact also points to the necessity of improving the investment policy and to raise the role of domestic assets.

Conclusion

The proposed approach for a justified choice of the food security indicator based on the matrix of cointegration links, the method of estimating the effect of foreign investment based panel cointegration model and the calculation of impulse response functions, provides an opportunity to quantify the effect of foreign investment on food security in the long term. It is impossible to expect instant effects in the modeling of such complex socio-economic processes as food security. This study can also contribute to the development of a methodology for monitoring economic policies in African countries. It will create conditions for food security, taking into account long-term financing indicators, and thus have some value in light of the formulated millennium goals to achieve Zero Hunger. ■

References

1. FAO (2015) *FAO and the 17 sustainable development goals, 2015*. Available at: <http://www.fao.org/3/a-i4997e.pdf> (accessed 20 March 2017).
2. FAO (2015) *The state of food insecurity in the world, 2015*. Available at: <http://www.fao.org/3/a-i4646e/index.html> (accessed 20 March 2017).

3. Jenkins J.C., Scanlan S.J., Peterson L. (2007) Military famine, human rights, and child hunger: A cross-national analysis, 1990–2000. *Journal of Conflict Resolution*, vol. 51, no. 6, pp. 823–847.
4. FAO (2015) *Achieving zero hunger: The critical role of investments in social protection and agriculture, 2015*. Available at: <http://www.fao.org/publications/card/ru/c/91014696-3723-4df5-b729-2b4e55b22e8f/> (accessed 20 March 2017).
5. Kopnova E., Rodionova L. (2017) *An analysis of the economic determinants of food security in North Africa*. Moscow: HSE. Series WP BRP “Basic research program”, no. 166/EC/2017.
6. Jenkins J.C., Scanlan S.J. (2001) Food security in less developed countries, 1970–1990. *American Sociological Review*, no. 66, pp. 718–744.
7. Barro R.J. (1991) Economic growth in a cross section of countries. *Quarterly Journal of Economics*, no. 106, pp. 407–443.
8. Firebaugh G., Beck F. (1994) Does growth benefit the masses? Growth, dependence, and welfare in the Third World. *American Sociological Review*, no. 59, pp. 631–653.
9. Scanlan S.J. (2004) Women, food security, and development in less-industrialized societies: Contributions and challenges for the new century. *World Development*, no. 32, pp. 1807–1829.
10. Safaei J. (2006) Is democracy good for health? *International Journal of Health Services*, no. 36, pp. 767–786.
11. Wickrama K.A., Mulford C.L. (1996) Political democracy, economic development, disarticulation, and social well-being in developing countries. *The Sociological Quarterly*, no. 37, pp. 375–390.
12. Mihalache-O’Keef A., Li Q. (2011) Modernization vs. dependency revisited: Effects of foreign direct investment on food security in less developed countries. *International Studies Quarterly*, vol. 55, no. 1, pp. 71–93.
13. Skoet J., Stamoulis K., Deuss A. (2004) *Investing in agriculture for growth and food security in the ACP countries*. WP No. 04-22 FAO – ESA.
14. Austin K.F., McKinney L.A. (2012) Disease, war, hunger, and deprivation: A cross-national investigation of the determinants of life expectancy in less-developed and sub-Saharan African nations. *Sociological Perspectives*, no. 55, pp. 421–447.
15. Phillips P.C.B., Moon H.R. (2000) Nonstationary panel data analysis: An overview of some recent developments. *Econometric Reviews*, no. 19, pp. 263–286.
16. Im K.S., Pesaran M.H., Shin Y. (2003) Testing for unit roots in heterogeneous panels. *Journal of Econometrics*, no. 115, pp. 53–74.
17. Hamilton J.D. (1994) *Time series analysis*. Princeton: Princeton University Press.
18. Poleman T.D. (1997) Recent trends in food availability and nutritional wellbeing. *Population and Environment*, no. 19, pp. 145–165.
19. Breitung J., Das S. (2005) Panel unit root tests under cross-sectional dependence. *Statistica Neerlandica*, no. 59, pp. 414–433.
20. Choi I. (2001) Unit root tests for panel data. *Journal of International Money and Finance*, no. 20, pp. 249–272.
21. Pedroni P. (2004) Panel cointegration: Asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis. *Econometric Theory*, no. 20, pp. 597–625.

Analysis of the consistency of enterprise architecture models using formal verification methods¹

E.A. Babkin

*Professor, Department of Information Systems and Technology
National Research University Higher School of Economics
Address: 25/12, Bolshaya Pecherskaya Street, Nizhny Novgorod, 603155, Russian Federation
E-mail: eababkin@hse.ru*

N.O. Ponomarev

*Student, Business Informatics MSc Program
National Research University Higher School of Economics;
Software Engineer, Intel Corporation
Address: 25/12, Bolshaya Pecherskaya Street, Nizhny Novgorod, 603155, Russian Federation
E-mail: nik4nikita@gmail.com*

Abstract

Enterprise architecture design is a complex process which makes it possible to synchronize the capabilities and needs of business and information technologies (IT). It can be achieved by clarifying the understanding and formalization of the business processes and the interaction of the elements of the system through their formal description. The large number of interacting business processes and enterprise architecture entities raises the question of verifying their correctness. Therefore, it is necessary to formalize the requirements for architecture and be able to automatically verify them.

In this paper, we propose a method for detecting logical contradictions in enterprise architecture models based on a model checking approach adopted in the context of business modeling. As an enterprise architecture description language, we use the modern open and independent ArchiMate standard. Developed by The Open Group, the standard provides a general specification for business processes, organizational structures, information flows, IT-systems and the technical infrastructure description of the enterprise. As a verifier, the language and tools of the MIT Alloy Analyzer system were chosen; they facilitate analysis of model constraints in terms of relational logic by automatically generating structures that satisfy the requirements of a logical model.

In this paper, we propose to simplify and automate the process of specification and verification of enterprise architecture domain models using Archi - the visual editor for ArchiMate models. We have developed the editor plug-in which translates the enterprise architecture models into the language of the MIT Alloy Analyzer system and uses the meta-model of the ArchiMate specification as the basis for constructing specific domain models. The proposed method and software solutions have been tested using the ArciSurance case and their enterprise architecture model.

¹ The research was carried out with financial support of Russian Fund of Basic Research No. 16-06-00184 A "Development and investigation models of online-discussion based on materials of political news"

Key words: enterprise architecture, ArchiMate, Alloy Analyzer, verification, consistency analysis, formal methods.

Citation: Babkin E.A., Ponomarev N.O. (2017) Analysis of the consistency of enterprise architecture models using formal verification methods. *Business Informatics*, no. 3 (41), pp. 30–40.
DOI: 10.17323/1998-0663.2017.3.30.40.

Introduction

Gartner defines the term “enterprise architecture” as a discipline for proactive and comprehensive response to a destructive force by identifying, analyzing, and making changes in the desired direction of vision and business results [1]. Enterprise architecture helps managers to find the best strategies for organizational development with respect to the information systems on which it strongly depends nowadays. However, in the process of transformation and gradual complication of the architecture, a serious problem is the lack of any practical opportunity to carry out “manual testing” of the model according to previously formulated requirements. Such testing is known as verification [2, 3].

In this paper, a general approach of formal verification is developed using the principles of model checking applied to the enterprise architecture. In this case, the required properties of the model are expressed by formulas of a certain dialect of formal logic, and the consistency checks are reduced to an exhaustive analysis of the entire space of its states [4]. In comparison with other approaches, this method has two significant advantages. It can be fully automated, and its use does not require that the business analyst should possess special knowledge in the field of mathematical logic and the theory of proofs of theorems. The paper proposes a new method for detecting logical contradictions in enterprise architecture models based on formal requirements. The formal verification of the architecture, as well as its business processes, should provide an opportunity to build a more reliable architecture.

The object of the study is the well-known case of the Open Group – the architecture of

the insurance company ArchiSurance [5]. The subject of the study is the logical consistency of the main elements of the architecture, business processes and services of this enterprise. As the language of the enterprise architecture description, a modern, open and free language – ArchiMate is used in the work [6]. This standard provides a general specification for the description, construction and operation of business processes, organizational structures, information flows, IT systems and the technical infrastructure of the enterprise.

As a verification tool, the language and logic of the Alloy Analyzer system (<http://alloy.mit.edu/alloy/>) [7] is selected. The tool allows you to analyze the limitations of the model in terms of relational logic by automatically generating structures that meet the requirements of the logical model.

The authors believe that the application of formal methods has the greatest effect in the case of close integration with the software environment for modeling the enterprise architecture. The developed integrated software solution in this case allows you to automate the formalization process of the architecture model and its verification. Following this principle, within the framework of this research, it is necessary to develop a meta-model of the ArchiMate specification on the basis of which a user model will be built in a certain modeling environment.

In this article, the results are described as follows. Section 1 outlines the main problems of building an enterprise architecture and the specifics of the ArchiMate language. Section 2 shows the principle of model verification. Section 3 is devoted to the analysis of available verification tools, here the details of the

language used and the logic of the MIT Alloy Analyzer system are described in more detail. Section 4 presents the main results of the creation of the ArchiMate meta-model, and Section 5 describes the proposed general algorithm for converting the model from ArchiMate to the MIT Alloy Analyzer language construct. Section 6 describes the process of creating a formal domain model. Section 7 contains the results of the verification of the enterprise architecture view example based on the proposed algorithm. Finally, section 8 focuses on the details of integrating the presented approach into the Archi modeling tool. The Conclusion sums up the results and determines paths for further research.

1. Enterprise architecture and the problems of its construction

In practice, the enterprise architecture is usually developed because some stakeholders have certain doubts about the functioning of business and IT systems. The role of the architect of the enterprise is to eliminate these problems by defining and specifying the requirements of the stakeholders, developing architectural presentations that show how the problems will be solved taking into account the requirements and trade-offs that need to be coordinated with the potentially conflicting interests of the parties and, in the end, synchronize the business opportunities and needs and IT [6, 8]. The solution to this problem is achieved by clarifying the understanding and formalization of the description of business processes and the interaction of the elements of the system through their formal description.

As a language for describing the enterprise architecture, this work uses the modern, open and independent language of ArchiMate 2.1. Its applicability and popularity is confirmed by a large number of certified organizations in the world (4,314) (<http://www.togaf.info/archimate-visualmap.html>), and the number of participants in the annual ArchiMate Forum reaches

121 (http://reports.opengroup.org/membership_report_archimate_forum.pdf). There are such famous organizations among them as Boeing, Dell, IBM, Philips and many others. This language allows a business analyst to represent the architecture in the form of a set of views that, depending on the needs, can only include elements on one level or can show vertical relationships between levels.

The levels include:

- ◆ a business level that offers products and services to external customers;
- ◆ the level of the application that supports the business layer with application services implemented by software applications;
- ◆ a technological level that provides the infrastructure services necessary to support software systems.

Aspects:

- ◆ the aspect of the active structure is various components that reflect actual behavior, i.e. “subjects” of activity;
- ◆ the aspect of behavior represents processes, functions, events performed by subjects;
- ◆ The aspect of a passive structure represents objects (physical or informational) on which behavior is performed.

2. Construction and verification of models

The presence of a large number of interacting business processes and enterprise architecture entities in the ArchiMate models raises the task of verifying their correctness. This task can be performed using the model checking method. This is a method of verifying that a given logical formula is satisfied on a given formal model of the system, that is, it takes a true value [9].

The method includes several stages: modeling, specification and verification. The first task is to bring the projected architecture model to a formal form that is acceptable for the tools of verification of program models. At the stage of the specification, it is necessary to formulate the properties that the designed model of the

enterprise architecture must have in the language of formal logic.

Verification of the model can show whether the projected system corresponds to a given formal specification, but to determine whether the given specification covers all the properties of the system is not possible. The verification phase, ideally, should be done automatically, but in practice, human intervention is most often needed. In the case of a negative verification result, a counter-example will be generated that will allow the user to track where the error occurred and fix it. It is also possible that the formal form of the system or its requirements have been described incorrectly. The result of verification should also help to identify these problems [10].

3. Comparative analysis of verification tools

In the course of the work, the most popular languages of modeling and analysis of abstractions were considered: B [11], OCL [12], VDM [13], Z [14] and Alloy [7]. All of them are able to describe any complex structure in a concise and abstract form, and each has an active community of users and researchers.

Z. Language Z was first developed in 1977 by Jean-Raymond Abrial at Oxford University and is based on logic and set theory. One of the advantages of Z is that it has a rich mathematical notation, making it an expressive language. A clear style of computational notation makes it possible to maintain many different idioms. However, the automatic proof of the theorems in Z is bounded. It is automated only to a certain extent, and complex proofs often require guidance from an experienced user.

OCL. Object constraint language (OCL) is the UML constraint language developed at IBM and ObjecTime Limited and added to the UML in 1997, which was originally developed as the annotation language for UML class diagrams. OCL is based on the logic of predicates of the

first order, but uses a syntax similar to programming languages, and is very closely related to the UML syntax. OCL allows you to mix declarative elements and elements of operations. However, this language is too implementation-oriented and therefore not suitable for conceptual modeling.

VDM. VDM (Vienna Development Method) is a set of methods for the development of computer systems. It originated from IBM's lab in the mid-1970s and was developed by Cliff Jones and Deans Björner. However, all existing tools do not provide a fully automatic analysis in the style of the model check. VDM supports an object-oriented paradigm and parallelism. Although this is one of the first formal methods of developing IT systems, it has been improved, standardized and is still widely used in the industry.

B. Language B was developed by Jean-Raymond Abrial, one of the creators of Z. It includes the language and method of obtaining implementations from abstract models by stage-by-stage processing. The specification language, Abstract Machine Notation (AMN), reflects its essence in the title: the system for the language is considered (as in VDM and Z) as a finite state machine with operations over the global state. Starting with a very abstract automaton, the details are added one layer at a time until an automaton is obtained that can be translated directly into the code. However, compared to the same Z, B is a lower-level language, similar to an abstract programming language, and is more focused on refinement of code, rather than on system specifications.

Alloy. MIT Alloy Analyzer, developed at the Massachusetts Institute of Technology under the leadership of Daniel Jackson, allows us to identify and detect contradictions in the projected models of systems.

Alloy is a language of structural modeling based on first-order logic for expressing complex structural constraints and behavior. The Alloy language derives from the Z specification language and Tarski's relational calculi, treating relationships as the main unit of analysis, and

uses the relational composition as a powerful operator for combining various structured data.

MIT Alloy Analyzer is a tool for analyzing models written on Alloy. It supports two types of automatic analysis: searching for an entity that satisfies the constraint, and searching for a counter-example for the given model judgments.

To limit the size of the search, the scope function is used; this fixes the number of entities and counter-examples analyzed by this command. Alloy analyzes the limitations of the model and selects the structures that satisfy them. Structures are displayed graphically, and their appearance can be adjusted manually.

During the assessment, the language Alloy was chosen. It is a full-fledged structured declarative modeling language that can express all sorts of complex structural constraints, reflect the logic of model behavior and conduct automatic analysis.

4. Meta-modeling

To find the contradictions in the enterprise architecture models using the MIT Alloy Analyzer tool, we must create a meta-model of the ArchiMate specification, on the basis of which specific domain models will be constructed later. A meta-model is a kind of language model that captures its basic properties, language concepts and semantics [4].

The core of the developed meta-model contains several key entities of the ArchiMate 2.1 speci-

cation common to all levels of representations. From them, all other elements on each level of the language are inherited and expanded by their properties. The keyword “abstract” of the Alloy language emphasizes the signature property that this entity does not include elements outside its extensions. In the proposed meta-mode architecture, all the signatures of the formal model of the system extend the essence of Element.

Graphically, the corresponding fragment of the upper level meta-model ArchiMate is shown in *Figure 1*.

Having defined all the necessary basic entities, we expand the hierarchy of the meta-model towards the business level, which reflects all the entities and links from the ArchiMate 2.1 specification².

Further, by the same principle, we set the meta-model of the application level in the Alloy language from the ArchiMate 2.1 specification³. In this case, the functionality of some types is similar, for example, ApplicationFunction and ApplicationInteraction. In order not to duplicate the links and code of the model, we select a base class with all common types of communication and add its functionality for each type.

The last level of the meta-model is the technology level. It is described by analogy with previous levels in accordance with the specification ArchiMate 2.1 in the language Alloy⁴. Thus, the meta-model architecture has a hierarchical

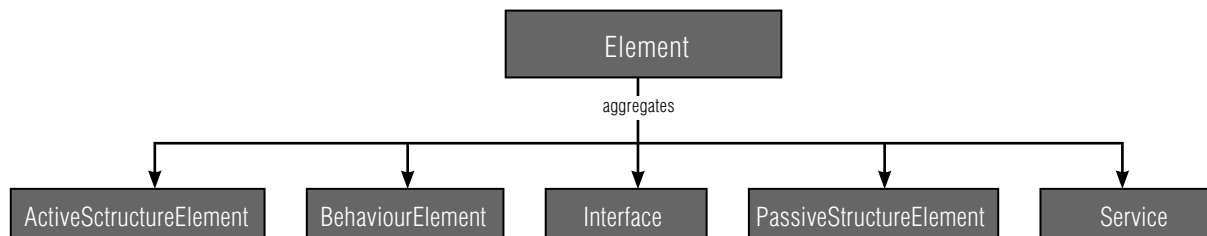


Fig. 1. Graphical representation of the top-level meta-model ArchiMate

² Graphical representation of the meta-model is available at: <https://goo.gl/ZLr4Qv>

³ Graphical representation of the meta-model is available at: <https://goo.gl/J9TTHX>

⁴ Graphical representation of the meta-model is available at: <https://goo.gl/Yo44Dw>

structure: there is an upper common layer from which three ArchiMate level modules are inherited. These, in turn, will be used by the entities of the user case.

5. Model transformation algorithm

Modeling of the entities of the subject domain is completely based on the entities of the created meta-model [4, 10]. Therefore, before describing the formal architecture model, you need to import the module of the developed ArchiMate specification meta-model with the Alloy command “*open ArchimateMetaModel*”.

The algorithm for converting from ArchiMate to Alloy is proposed as follows. Each entity of the view model is transformed into a signature with an extension corresponding to the type of the entity. In the signature name, spaces are replaced with a lower slash, and uppercase letters are replaced with the corresponding lowercase letters. Thus, a unique name is created for the signature. Connections that come from the essence are transformed into limitations (facts) of the signature. If the entity does not have any relationships that are described in the signature of the meta-model from which it is inherited, then the constraint for this relationship is set to “*none*”, that is, it forbids creating a counterexample with this connection when verifying (*Table 1*).

Table 1.

Displaying entities ArchiMate in Alloy

ArchiMate	Alloy
Model elements	The signature (sig “name”) with an extension (extends) the entity type (ex. BusinessFunction)
Communication	The list of restrictions (fact) for a particular signature.

6. Creating a formal model of the subject field

As a verification case, consider a classic example designed to illustrate the use of the ArchiMate modeling language in the context of the TOGAF structure. It describes the company’s basic architecture, as well as a number of change scenarios [5].

The case concerns the insurance company ArchiSurance, which was formed as a result of the merger of three previously independent companies:

- ◆ Home & Away (homeowners and travel insurance);
- ◆ PRO-FIT (auto insurance);
- ◆ Legally Yours (insurance of legal expenses).

At present, the company consists of three divisions with the same names and headquarters as their independent predecessors.

ArchiSurance was created to use the synergy effect between the three organizations to control their costs, maintain customer satisfaction and invest in new technologies. The new company offers all insurance products of the three companies that were merged, and its organizational structure looks like this (*Figure 2*).

To completely cover all aspects of the enterprise architecture on one diagram – the task is not simple and voluminous. More often, the user is only interested in a particular aspect of architecture. Therefore, in ArchiMate there is a concept of “representation”. This is a kind of point of view that allows you to work flexibly in a common architecture, focusing on important aspects, both individual and in a bunch of different levels.

In this paper, we will consider the multi-layered and most general representation of the ArchiSurance architecture, divided into three levels according to the ArchiMate specification⁵.

⁵ Graphical representation of the meta-model is available at: <https://goo.gl/XdvNL>

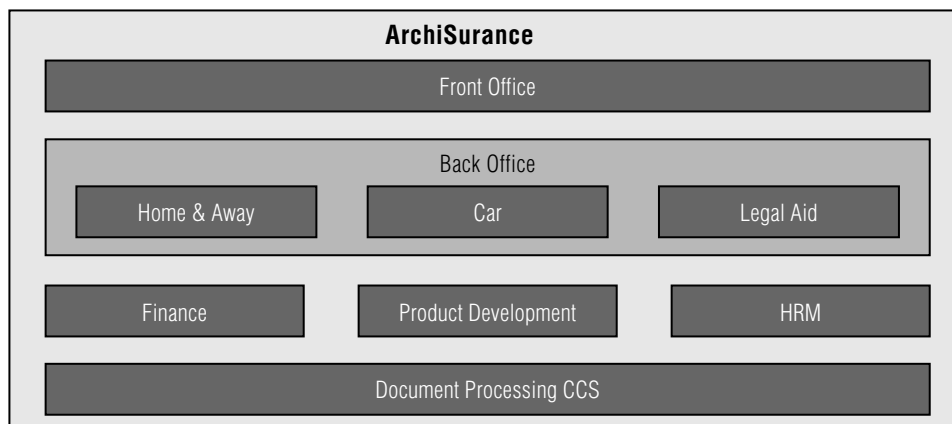


Fig. 2. The organizational structure of the company ArchiSurance

In accordance with the transformation algorithm described earlier, we obtain a model in the language of Alloy for subsequent formal analysis. The model consists of 67 signatures and, accordingly, of the same number of facts. Every fact has from zero to six constraints on connections.

7. Determination of the formal requirements of domain and verification in the MIT Alloy Analyzer

At the architecture design stage, the properties that the future enterprise architecture model should have are formulated. One of these properties is the following statement: “Are there any application components or their functionality that use data, but do not have access to the technology services for their extraction?”

To conduct verification, it is necessary to describe this statement in the language of Alloy’s formal logic, and to determine the scope of the search. It is necessary that the search for a solution be carried out with all the presentation entities, so the query is supplemented with a list of all the elements of the model with the keyword “exactly”.

When starting the search for a counter-example for a given judgment of the model, a significant role is played by scaffolding. Often quantified formulas can be reduced to equivalent

formulas without the use of quantifiers. This abbreviation is called scolemization and is based on the introduction of one or more Skolem constants or functions that fix the limitation of the quantitative formula by their values.

In this example, the MIT Alloy Analyzer finds a counter example for a given restriction and assigns the *app_comp* Skolem relation of the ApplicationComponent type to the name “\$DataAccess_app_comp” for the existing “Policy Data Management” entity, and the *app_funct* of the ApplicationFunction type is named “\$DataAccess_app_funct” for the existing “Create Policy” entity. Thus, in our example, two entities are found that do not meet the specified formal requirements (Figures 3 and 4).

After a negative verification result, there are two possible solutions to the problem: changing the current architecture or formal requirements. If the formal requirements are determined by the stakeholders and, therefore, are not subject to adjustment, then the architecture of the enterprise needs to be reconstructed until verification yields a positive result.

8. Developed integration approach for Archi

One of the main goals of this work is the development of a software tool that would automate part of the verification process of

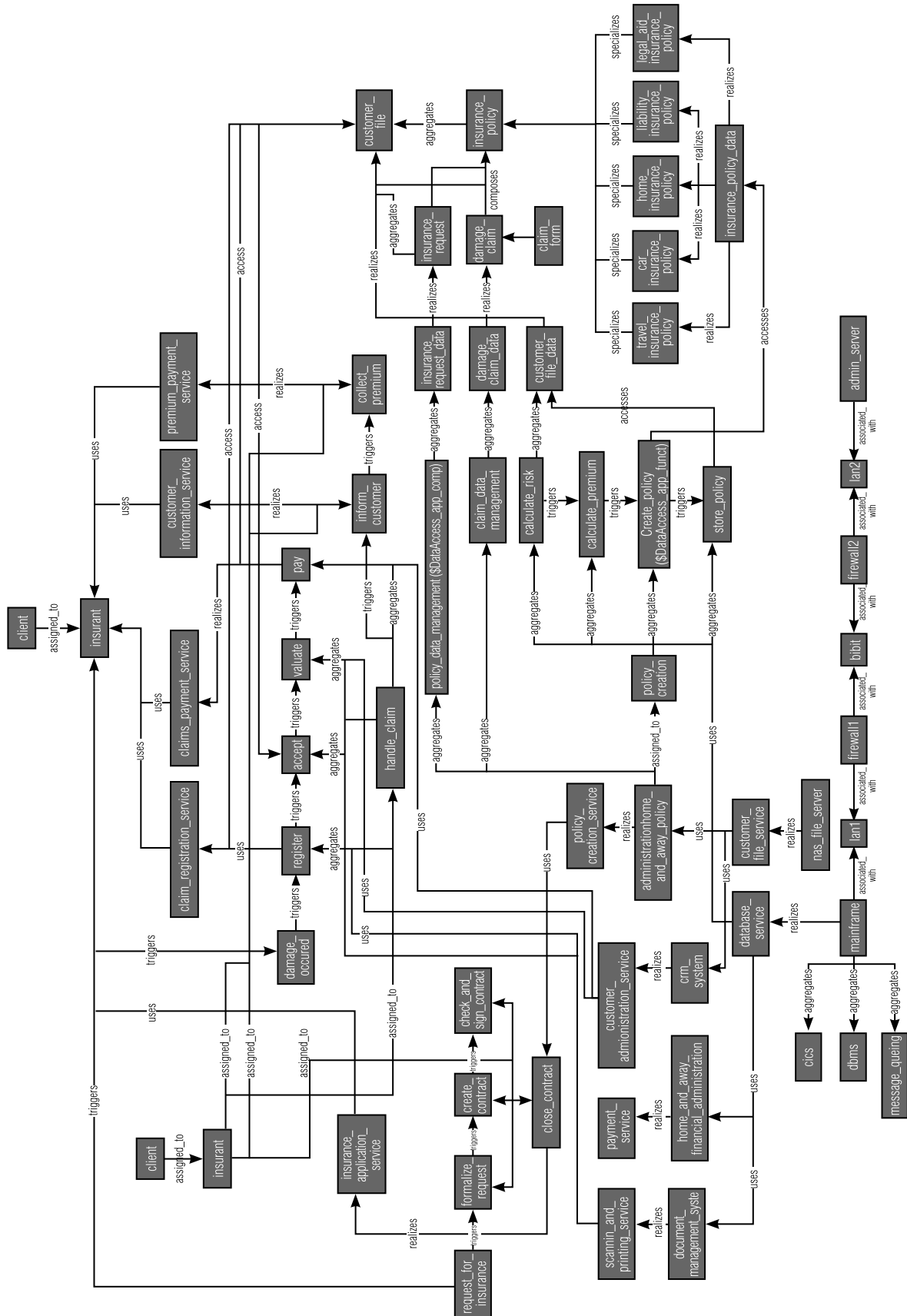


Fig. 3. ArchiSurecase: The result of the verification of the presentation

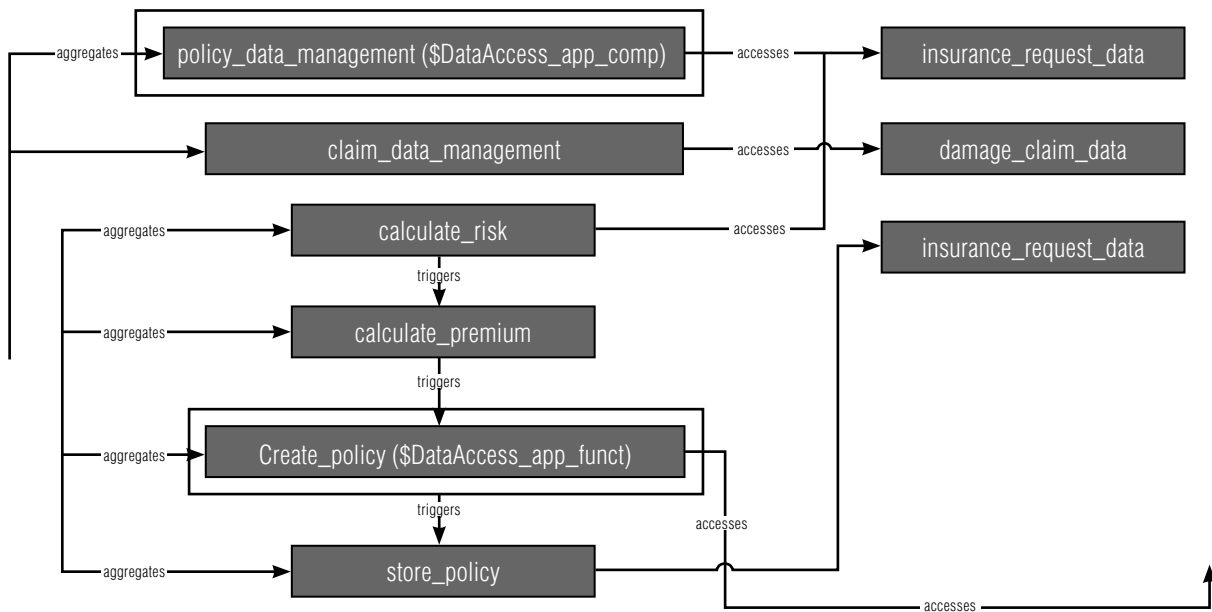


Fig. 4. ArchiSurance case: Found entities that contradict the given condition (in the box)

ArchiMate models and make the approach of testing models in the field of business modeling more accessible. As a software platform, the popular Archi editor has been selected. This is a free cross-platform open source tool for visual modeling and design of ArchiMate models developed on the Eclipse EMF platform and extensible with custom plug-ins.

In this paper, a plug-in was developed for the visual editor Archi which supports the automatic transformation of the enterprise architecture view into the Alloy language (Figure 5). The plugin is written in Java. The design of the solution contains three main classes: the logic of parsing and model transformation, the implementation of the user interface and

auxiliary methods for converting strings. The transformation logic is encapsulated in the Alloy Exporter class, and the public transform-Model method takes an object that implements IArchiMateModel interface and creates a temporary CaseExample.als file into which the converted model is written. The input interface for Alloy commands is written using the Swing library. The amount of code due to the re-use of Alloy libraries is 546 lines (<https://github.com/nik-ponomarev/archimate2alloy>).

After designing a multilevel architecture of the ArchiSurance enterprise, the dialog of the verification window in Archi is opened by selecting “File → Export → Model to Alloy Format” from the context menu.

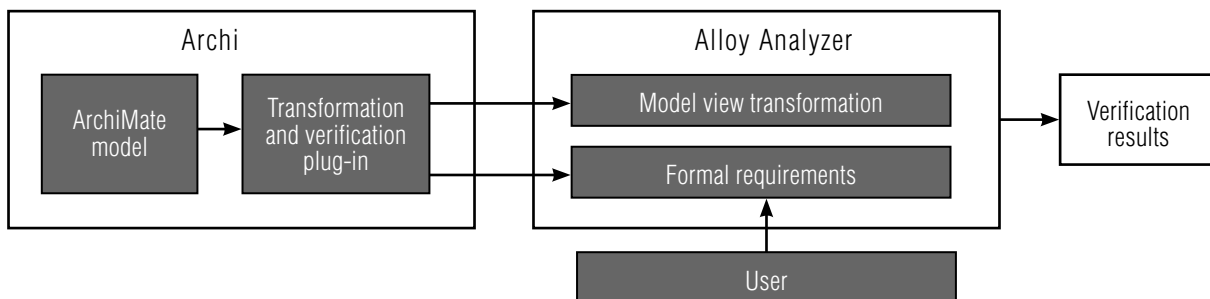


Fig. 5. Scheme of verification toolkit

Next, there appears an interface for entering formal requirements for the architecture in the language Alloy, as well as commands for launching verification, where you need to determine the coverage of the search. At this stage of implementation, it is possible to check only one statement per launch.

The “*Use full scope*” option means that all the entities of the model must participate in the verification, each in a single instance. If the user needs to set his own type of coverage, then there is the option “*Custom scope*”. The “*Find solution*” button starts the verification mechanism.

In the case of a negative verification result, in the presence of skolemized structures, entities that do not satisfy the query condition will be painted red. In the absence of these variables, all elements of the model will be painted in red, which will be generated by Alloy Analyzer as a counter-example. If no counter-examples are found, then the original model satisfies formal requirements. In this case, the corresponding window with the message “*No example/counter-example found*” will be displayed.

In our example, the plugin found two application-level components that did not satisfy the constraint condition, and colored them in red. These are Policy Data Management and Create Policy components that use client data in the current architecture, but do not have access to their receipt from database services of the technological level. In this regard, it is necessary to revise the logic of the current architecture.

The proposed approach to model testing allows to identify errors that are not just related to the incorrect display of the specification, namely errors at the level of the architecture logic of the enterprise architecture and the behavior of business processes. It should be noted that it is not possible to carry out such verification manually with the increase in the size of models, which again confirms the practical value of this tool.

Conclusion

In this paper, the problem of automatic verification of enterprise architecture models in ArhiMate language is investigated. As a tool for the model checking method, the relational logic tool and the MIT Alloy Analyzer modeling system were used.

The idea of the work is to develop a solution that is tightly integrated with the actively used in practice Archi modeling tool to extract all the elements from the architecture view, and then completely automatically create a formal model in the Alloy language and conduct verification through MIT Alloy Analyzer. Verification is based on the projected meta-mode of the ArchiMate specification in the Alloy language, which describes the main entities, the relationships between them, and other language restrictions. Specification of formal requirements in the form of facts, assertions, predicate and functions can be entered by the user in a separate menu, specifying options for restricting the search. The work also discusses the methods for transforming the model and the requirements description rules used in the implemented software.

Finally, the applied nature of the method is presented in the ArchiSurance case – verification of its generalized multi-layered representation of the enterprise architecture. In addition, the work describes the mechanism for displaying verification results in the Archi modeling tool, which allows you to better display the principles of the enterprise’s functioning.

Based on the work done, we can conclude that the mechanism of logical validation for the architecture models of the company ArchiMate is applicable. This method is most important for models with a large number of elements and connections between them, given that manual verification is not possible. The introduction of a formal description of the enterprise architecture, its business processes, and requirements

will allow us to build quality management systems of the company, to solve the problem of building an effective management structure, to optimize processes based on key indicators.

In the future, this work will be improved in the direction of implementation of verification on

models with the help of LTL/CTL logics. Further development of the suggested approaches will allow us to use formal analysis methods for a broad class of the models, including the models of communication of autonomous agents in the framework of DEMO methodology. ■

References

1. James G.A., Handler R.A., Lapkin A., Gall N. (2005) *Gartner enterprise architecture framework: Evolution 2005*. October 25, 2005. Gartner ID: G00130855.
2. Clarke E., Grumberg O., Peled D. (1999) *Model checking*. MIT Press.
3. Korotkov A. (2013) *Arkhitektura predpriyatiya. Kak zastavit' IT rabotat' na vashu kompaniyu* [Enterprise architecture. How to force IT to work in favor of your company]. Available at: http://andrey-korotkov.ru/wp-content/uploads/2013/02/andrey-korotkov.ru_Enterprise_architecture.pdf (accessed 16 March 2017) (in Russian).
4. Clark T., Sammut P., Willans J. (2008) *Applied meta-modeling: A foundation for language driven development*. Ceteva.
5. Jonkers H., Band I., Quartel D. (2012) *The ArchiSurance case study. White paper*. The Open Group.
6. The Open Group (2013) *Archimate 2.1 Specification. Open Group Standard*. Zaltbommel: Van Haren Publishing. Available at: <https://www.vanharen.net/Samplefiles/9789401800037SMPL.pdf> (accessed 16 March 2017).
7. Jackson D. (2006) *Software abstractions: Logic, language and analysis*. MIT Press.
8. Kudryavtsev D.V., Arzumanyan M.Y., Grigoriev L.Y. (2014) *Tekhnologii biznes-inzhiniringa* [Technologies of business engineering]. St. Petersburg, Polytechnic University (in Russian).
9. Karpov Y.G. (2009) *Model checking. Verifikatsiya parallel'nykh i raspredelennykh programmnykh sistem* [Model checking. Verification of parallel and distributed program systems]. St. Petersburg, BHV-Petersburg (in Russian).
10. Szwed P. (2015) Verification of ArchiMate behavioral elements by model checking. Proceedings of the *14th IFIP International Conference on Computer Information Systems and Industrial Management (CISIM 2015)*. Warsaw, Poland, 24-26 September 2015, pp. 132–144.
11. Lano K. (2012) *The B language and method: A guide to practical formal development*. Springer Science & Business Media.
12. Warmer J.B., Kleppe A.G. (1998) *The object constraint language: Precise modeling with UML*. Addison-Wesley.
13. Fitzgerald J.S., Larsen P.G., Verhoef M. (2008) Vienna development method. *Wiley encyclopedia of computer science and engineering*. John Wiley & Sons.
14. Spivey J.M., Abrial J.R. (1992) *The Z notation*. Hemel Hempstead: Prentice Hall.

Approaches to geomarketing via office locations serving individuals and legal entities of a commercial bank

Ruslan A. Dolzhenko

*Professor, Department of Labor Economics and Personnel Management
Ural State University of Economics;*

*Associate Professor, Department of Personnel Management and Socio-Economic Relations
Altai State University*

Address: 62/45, 8 Marta / Narodnoy Voli Street, Ekaterinburg, 620144, Russian Federation

E-mail: snurk17@gmail.com

Abstract

In this article we examine the possibility of using geomarketing as a tool for modeling the spatial location of a bank's offices serving individuals and legal entities. Geomarketing is a marketing concept of modern production management based on geo-information technologies which involves the use of spatially localized information to support decision-making. This concept is particularly timely for commercial banks which present their services to customers in an extensive network of service offices, because for the majority of banking services the determining factor in the consumer's choice of bank for a given service will be the convenience of the location of its office.

The article describes a sequence of typical steps in geomarketing simulations and makes recommendations for the implementation of each selected stage. Methodical approaches to the implementation of geomarketing modeling for an optimal network of the bank's service offices are highlighted. Problems in the organization of the network of service offices of the bank which can be solved with the help of geomarketing design are studied. Recommendations are presented for evaluating the capacity of a commercial bank's office service network depending on their specialization in serving individuals or legal entities. An algorithm is presented for calculating the service office network capacity for individuals together with respective formulas. An approach to evaluating the capacity of the office network servicing legal entities is described. Recommendations are offered for implementing the evaluation of planned location of offices servicing the bank's clients, as well as for determining the habitats of their preferred location to facilitate the most effective coverage of the local market while taking into account the return on investment in offices.

Key words: bank, sales office, geomarketing, geomarketing modeling, rational location of sales offices, placement rating.

Citation: Dolzhenko R.A. (2017) Approaches to geomarketing via office locations serving individuals and legal entities of a commercial bank. *Business Informatics*, no. 3 (41), pp. 41–55.

DOI: 10.17323/1998-0663.2017.3.41.55.

Introduction

Under conditions of a developing market economy and increasing competition, domestic companies have been compelled to review their approach to determining where to locate their offices. Points of sale and service offices must be placed not in some accidentally arrived at locations but with precision, where the client flow and convenience for visitors will be optimal. Companies must evaluate not only current possibilities for rational location of offices but also the dynamics of change over the coming years. Success will accompany only those organizations servicing clients which take into account the geographic location of their offices and customers. This is why it is highly relevant to do research using modeling for rational placement of a company's offices in the markets where it has a presence.

The majority of organizations take a superficial approach to using the potential of modeling their geographic representation in territorial markets and, as a rule, use data from various geographic information systems (GIS), with the help of which they identify and visualize the location of offices on a map of the local market. But displaying the organization's points of presence on the city map does not by itself give any supplemental information which could qualitatively improve the effectiveness of the network's operations. You have to carry out detailed research of local markets, identify data about customers, highlight the factors which can influence their travel around the city. This is why it is relevant to use in business geo-information technologies and, in particular, geomarketing. This field is extremely important for commercial banks, especially those which have an extensive network of offices.

In practice, we see that a bank's success with customers is determined more and more by the intensity of the company's advertising and by successful location of sales offices. Of course, the innovative culture of the financial organi-

zation is also an important factor, but its significance is not so great. For example, the majority of domestic customers – individuals – are not acquainted with the possibilities of mobile banking. For such consumers of banking services, the determining factor when choosing a bank to service accounts will be convenient locations of its offices [1]. That is precisely why geomarketing modeling of the network of sales offices of a credit organization is extremely topical at present. This study is devoted to the given subject.

Before we review the possibilities for network modeling of office locations in a commercial bank, we will define the basic concepts of geomarketing.

1. The essence of geomarketing and the possibilities of using it in practice within a credit organization

From the structure of the word “geomarketing”, one can already conclude that it arose at the intersection of two sciences: marketing and geo-computer science. It is the field of marketing analysis using methods of geo-computer science [2]. Thus, geomarketing was founded on the use of GIS when analyzing a market and making decisions about the activity of the organization within it. There are two points of view on the definition of the given concept: one broad, the other narrow. In the first case, geomarketing can be defined as an approach to collection of data located in space when making marketing decisions. This approach is widespread in Europe. For example, in France it is supported by the National Council on Geographic Information (Conseil national de l'information géographique). The narrow interpretation of the “geomarketing” concept is more widespread in the USA, where it is taken to mean a computer system of programs, processes and materials which allow you to collect, manage, manipulate, analyze, model and select spatial data for complex management of a marketing system in the organiza-

tion. The given approach is used by the Federal Inter-Agency Committee for Coordination of Digital Cartography.

In Russian language scientific works, the narrow approach is more common. For example, in an article by I.G. Zhurkin and V.Y. Tsvetkov we find the following definition of geomarketing: it is the “process of planning, decision-making, price formation, promotion and sale of ideas, goods and services (including information and political services) by means of exchange satisfying the objective of separate individuals and organizations with the help of geo-information technologies” [3].

Use of geomarketing in business practice enables the organization to get answers to three key questions:

- ◆ How many sales offices does the organization need to efficiently serve its current and future client base?
- ◆ How should these offices be configured on the map of the city in order to take in the biggest client flow?

◆ How should the offices be situated relative to one another for maximum financial result?

Essentially, use of geomarketing when modeling the rational location of the organizations’s network of points of presence on the local market enables the firm to create a model of the most effective location of offices on the city map, to compare that with the existing network of offices and to take actions to improve the network till it reaches the target state.

The technology of geomarketing modeling of the organization’s network of offices is applied, as a rule, in three stages which are represented schematically in *Figure 1*.

Evaluation of the potential of sales office locations on the local market has key importance in geomarketing modeling.

Geomarketing evaluation is the method of assessing the location of an office with the help of which we determine the likelihood of high client flow in the bank’s sales office. The special feature of this method is that for each cri-

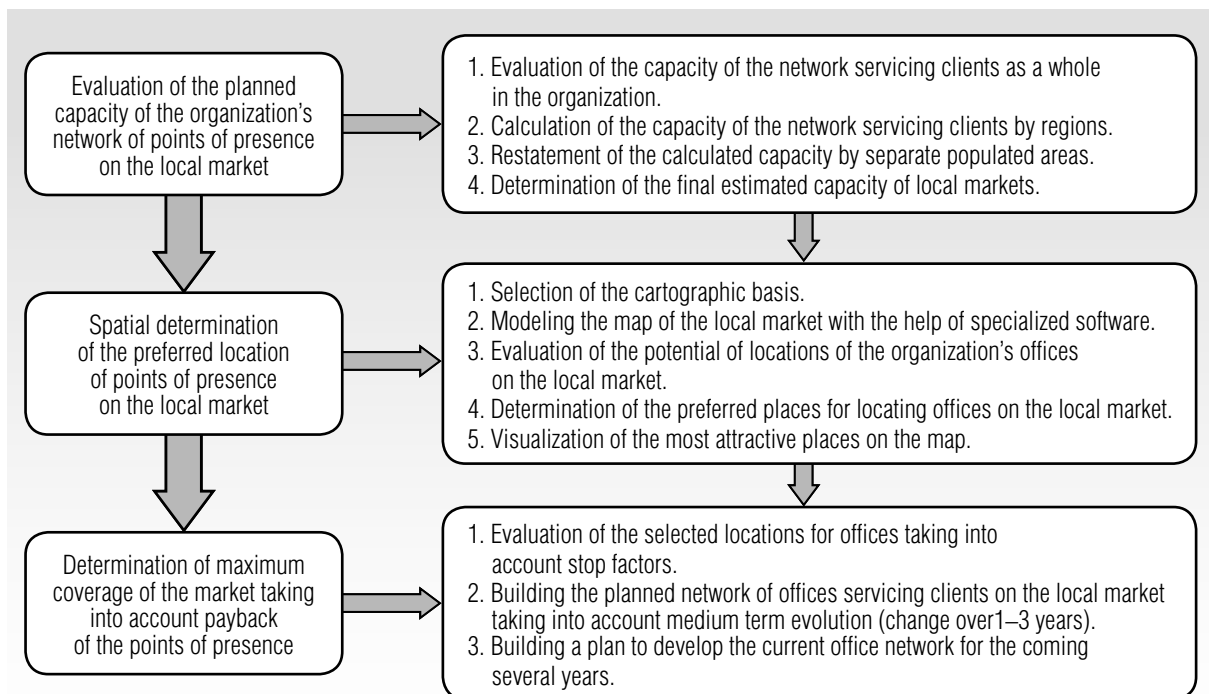


Fig. 1. Stages of geomarketing modeling of rational locations for the organization's office network

terion in the system of evaluating the location a specific score is assigned.

Experience shows that the effectiveness of a geomarketing evaluation largely depends on the extent to which the following three groups of factors are taken into account when modeling.

1. Correspondence to the basic criteria of the local market. In the modeling process, one must take into account the density of population. In addition, one must direct attention to the sources of additional flow of clients such as buildings and places attracting the interest of the population, for example, metro stations, restaurants, entertainment centers, etc.

2. Correspondence to the criteria of location of points of sale. In particular, one has to take into account the location of public transport stops, the transport lines around the planned point of sales (the greater the number of such lines, the higher the evaluation of the location will be). An important criterion when modeling is the density of pedestrian flows in front of the office. A high level shows that in the given place there is no obstacle to pedestrians getting around between parts of the local market. For example, the presence of a pedestrian crossing (whether underground or above ground) increases the flow of people. Moreover, it allows clients to freely walk up to the office from across the street, which is important especially when there is dense automobile traffic.

3. The premises of sales offices, independently of their location on the local market, must correspond to certain criteria: it is desirable for the location of a sales point to be on the ground floor; one must take into account the presence of windows on the facade of the building; the most desirable location of the premises is on the first line of buildings; and, finally, if there is no possibility to arrange a separate entrance, that can significantly lower the quality of the premises.

Such are the general factors connected with geomarketing modeling. There is no doubt

about the prospects for its use in practice within domestic organizations, especially for those which sell products and services to the population [4]. Further on we will review possibilities for using geomarketing tools applicable to a network of sales offices of a commercial bank. We will note that the possibilities for use of this technology for credit organizations are being actively studied by foreign and domestic researchers. Place as an element of the marketing mix is attracting the attention of many researchers. Therefore, the given field of research applicable to retail stores was studied for a long time, and as a result a large number of scientific works accumulated in this field. But what is the situation with respect to research into geomarketing modeling in the field of optimization of the branch network of bank offices?

The first publication on the given topic in foreign scientific literature dates back to 1978 [5]. It is indicative that it reviewed the possibilities of computer-aided mapping (CAM) for building the bank's branch network. On the basis of research into the branch network of a major American bank, the authors of the work proposed an approach to making marketing decisions with help from the mapping of geographic space and determination of the potential locations for branches of the financial organization. Despite the fact that almost 40 years have gone by since its publication, the relevance of the given subject has not declined.

Researchers in the given field offer various approaches to modeling the behavior of consumers, on the basis of which, in their opinion, one can take decisions about the location of offices in the space of populated areas. Approaches to modeling the behavior of consumers were analyzed in detail in the work of S.G. Kiselgof [6]. In this study, it was shown that the behavior of consumers, including consumers of bank services, can be determined by several key factors which researchers use as their basis. For example, they include the level

of education of the population, the number of organizations on the territory, the size of offices, the distance from the nearest commercial center, the location of the office with respect to the daily itinerary of consumers, and others [7].

All of these factors to a greater or lesser degree have their impact on the attractiveness of bank branches. Consequently, by doing modeling, one can determine the expected number of clients in the bank's points of presence and use the given information for taking decisions about the fate of offices in the future. In particular, the given approach is used in the work of L.D. Huff [8].

More and more often, foreign researchers agree that increasing the speed of moving consumers around the city develops further the importance of such an indicator as "client flow of consumers". Their preferences can be ignored in conditions when choice of the place of service is determined by convenience of location, speed of service, as well as throughput capacity of the office [9–12]. It is noteworthy that in domestic banking practice they began to pay attention to this only in the last decade. Thus, Sberbank, as the provider of management innovations in the banking sector, for purposes of resolving the problem of client flows, set for itself the task in 2012 to reduce the time of serving clients in its offices on average to 10 minutes (the Sberbank project "Ocheredey. Net"). For consumers, what counts more now is not the distance between points in the city but the speed with which it can be overcome.

In the domestic scientific milieu, the given topic has been examined in a number of works by F.T. Aleskerov and co-authors [13–15], from the position of the possibilities of mathematically modeling the rational location of bank offices. For example, approaches to modeling behavior in various circumstances are described in the work of S.G. Kiselgof and F.T. Aleskerov [15]. In works by F.T. Aleskerov

and V.Y. Belousova [13, 14], an analysis was made of preferred location territories for the bank at the macro-level. In particular, they determined the most attractive regions of the RF for opening new offices.

This work is a continuation of the scientific work of our colleagues in the given field. The accent is placed on practical aspects of using the technology of geomarketing which can be used by any credit organization when expanding or transforming its branch network.

2. Methodological approaches to geomarketing modeling of a bank's network of sales offices

The use of geomarketing in bank activity makes it possible to solve a number of essential methodological problems which credit organizations face. The key problems and tools for solving them are presented in *Table 1*.

Geomarketing modeling of a bank's network of sales offices includes three main stages.

1. Determination of the capacity of the network of offices. At the given stage, one calculates the necessary number of work places for servicing the client base of the region for each year with a planning horizon of five years. The final result of this stage of geomarketing is an evaluation of the approximate number of offices. A macro-economic model for analysis of capacity will be described below.

2. Determination of the habitats of preferred office locations. The city is divided into local markets, each of which is characterized by the total capacity of sources of additional client flows and the geomarketing scoring system. The result of this stage is the determination of the assessment of locations by number of transactions and volume of sales in the bank's points of sale.

3. Determination of the most effective coverage of the local market taking into account payback of the sales office. Depending on the score, one determines the necessary through-

Table 1.

Problems with organizing a network of bank sales offices and the tools for their solution with the help of geomarketing modeling

Basic problems	Tools for solving them
No forecast for capacity of the network	Macro-model and model of coverage of the local market make it possible to calculate the target capacity of the network for several years ahead
Non-rational selection of the location of sales offices	In the maps, one must consider only the most significant geomarketing factors
Non-rational selection of formats and sizes of sales offices	In the model of coverage of the local market, one must take into account not only large format sales offices but also basic formats. And the latter must gradually be replaced (according to need) by self-service offices
No approach to locating self-service devices and offices	
'Expert' evaluation of payback of sales offices	Model of office payback mathematically ties down the forecast of growth of key business indicators to the evaluation score

put capacity of the sales office for each zone of rational location. At the conclusion of this final stage of geomarketing, one can take decisions on opening/closing/reformatting/moving the bank's sales office.

Let us examine the content of these three stages in greater detail.

3. Determination of the capacity of the office network (stage 1)

This stage comprises three steps. First, it is necessary to calculate the capacity of the network of sales offices as a whole by RF subject (territorial subdivision). This can be done on the basis of macro-economic forecasts and the bank's target benchmarks. The next step is related to restating the calculated capacity from the level of the subject of the federation to the level of a city. It is determined proportionally to the number of the economically active members of the population in the city. Finally, at the last step the forecast capacity in the context of the city is distributed between local markets proportionately to the assessment score of the market and to current distribution of the load on existing sales offices. The target capacity of the local market is determined by its summary evaluation

score, or it can be calculated by two indicators: the number of service offices (taking into account the priority of formats) and the total number of workers who are required according to norms for service of the forecast client base over the coming five years. The given target capacity can be redistributed between shares of the local market across segments, starting from the specifics of the current distribution of clients, as well as the tendencies of the city's future development.

Let us look at an example of the assessment of the capacity of a network of offices for servicing bank clients. Given that the model for calculation differs according to the specialization of the office (servicing individuals or legal entities), we will examine the approach to evaluating capacity for each of these areas.

The evaluation of the capacity of a bank's office network for servicing legal entities can be done starting from the target number of clients across regions and segments (large, medium, small, micro-business), speed of growth of the client base by segments and norms for the number of clients per sales manager and the number of clients per cashier. The algorithm for calculating the target capacity is given in *Figure 2*.

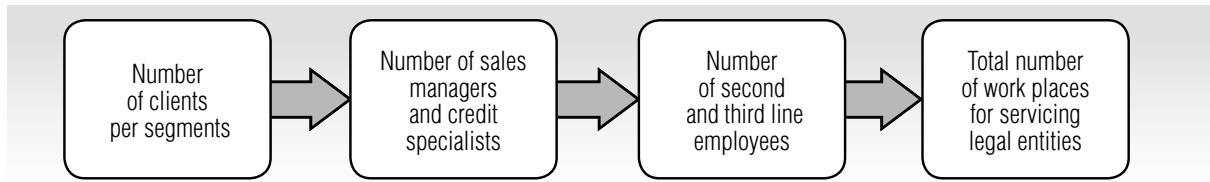


Fig. 2. Algorithm for calculating the capacity of a bank's office network servicing legal entities

The basic result of modeling is the calculation of the necessary capacities of the network by business segment taking into account change in the target shares of the market of the region. Moreover, for legal entities we calculate the target number of work places for workers of various categories (sales managers, credit inspectors, as well as second and third line staff such as cashiers and office managers).

The forecast number of a bank's legal entity clients can be determined by looking at the forecast rates of growth of the client base by segment (large, medium, small, micro-business). To calculate the forecast number of clients, we use the following formula:

$$PNC = ANC \cdot (1 + FGR), \quad (1)$$

where *PNC* is the predicted number of the bank customers from among legal entities;

ANC is the actual number of the bank's clients from among legal entities;

FGR is the forecast growth rate of the bank's client base.

Next it is necessary to calculate the forecast number of work places for cashiers, credit inspectors, sales managers by segment, as well as office managers. This number can be arrived at on the basis of the normative number of legal entity clients per sales manager, the normative number of credit contracts per credit inspector and the normative number of legal entity clients per cashier. As a rule, for each office there should be one office manager; for a large office, one may provide for deputy office manager positions.

Finally, the calculation of a target number of offices is made based on the forecast number

of legal entity clients. In the given instance, the bank, given the accepted approaches to the format of offices (micro-, macro-, specialized, all-purpose), can decide precisely which formats for offices and in which quantities they are needed. For example, in Sberbank, there is an operative norm according to which one bank office should service from 700 to 3000 legal entity clients.

Thus, the macro-model described makes it possible to determine the rational state of a network of offices servicing legal entities, taking into account the planned rates of growth of the client base and the share of service represented by remote banking.

The evaluation of a bank's network of offices servicing individuals should be done using other principles. The main elements for calculating the macro-model are shown in *Table 2*.

The algorithm to calculate indicators for modeling the planned capacity of a bank's office network is shown in *Figure 3*.

Let us look at how within the model described we can calculate the basic indicators shown in *Table 2* and *Figure 3*.

The calculation of the number of accounts per product in an *RF* subject (constituent territory) per member of the economically active population (that portion of the country's population which already uses or wants to and potentially can use bank products and services; for purposes of modeling, one should place in this category individuals aged 18 to 65) is made on the basis of the current population numbers in the *RF* subject, the portion of the population living in cities with 50,000 or more and the share of the population aged 18–65.

Table 2.

Basic elements for calculating a macro-model to evaluate the capacity of a bank’s office network servicing individuals

Calculation indicator	Calculation algorithm
Total transaction volume by city with a population greater than 50,000	Calculated on the basis of the target market share for deposits and consumer credits, as well as forecast numbers of the economically active part of the population in the RF subject (territorial unit)
Volume of transactions performed in an office with physical presence of a cashier	Calculated on the basis of the total transaction volume less the share of remote banking services
Number of work places for cashiers in cities with a population greater than 50,000	Calculated on the basis of the normative work time of one cashier work place and forecast volume of transactions carried out in the office with the physical presence of the cashier
Number of offices by format (large / basic) by city with population greater than 50,000	Calculated on the basis of the developed typology of formats depending on the number of cashier work places
Number of self-service devices (ATMs) per city with population greater than 50,000	Calculated on the basis of the normative load of self-service devices by ATM and information-payment terminals given the forecast volume of transactions and target share of remote banking services
Number of sales managers for banking products	Calculated on the basis of the number of sales and norms of productivity of sales managers for banking products

Thus, the total number of accounts by product within the bounds of the *RF* subject will be correlated with the number of the economically active population, as a result of which one can calculate the number of accounts of products per individual (the share of penetration of bank products within the given *RF* subject). Goskomstat is a source of information for the calculation. The formula for calculating the given indicator looks as follows:

$$NA = \frac{TNA}{P \cdot PUP \cdot PP}, \quad (2)$$

where *NA* is the number of accounts of products per person;

TNA is the total number of accounts by product;

P is the population of the *RF* subject;

PUP is the proportion of urban population living in cities greater than 50,000 (this level of screening is used in order to exclude from the modeling small towns which according to the code book of rules “Urban development. Planning and residential construction of urban and

rural communities” are considered in Russia to be towns with populations less than 50,000 and other types of populated areas);

PP is population proportion of those aged 18–65 (the age of potential bank clients).

The calculation of the target number for the bank’s product accounts can be made on the basis of the forecast total number of accounts within the subject and the bank’s target market share. Using data on the current population of cities with assumptions on growth/loss of population, one can build a forecast of the number of persons in the economically active population. In turn, the forecast total number of accounts by product within the *RF* subject is calculated by year as the product of the number of economically active persons and the number of product accounts per person. Their correlation gives us the target number of bank products:

$$TNA = \frac{EAP \cdot NA}{TMS}, \quad (3)$$

where *TNA* is the target number of accounts of bank products;

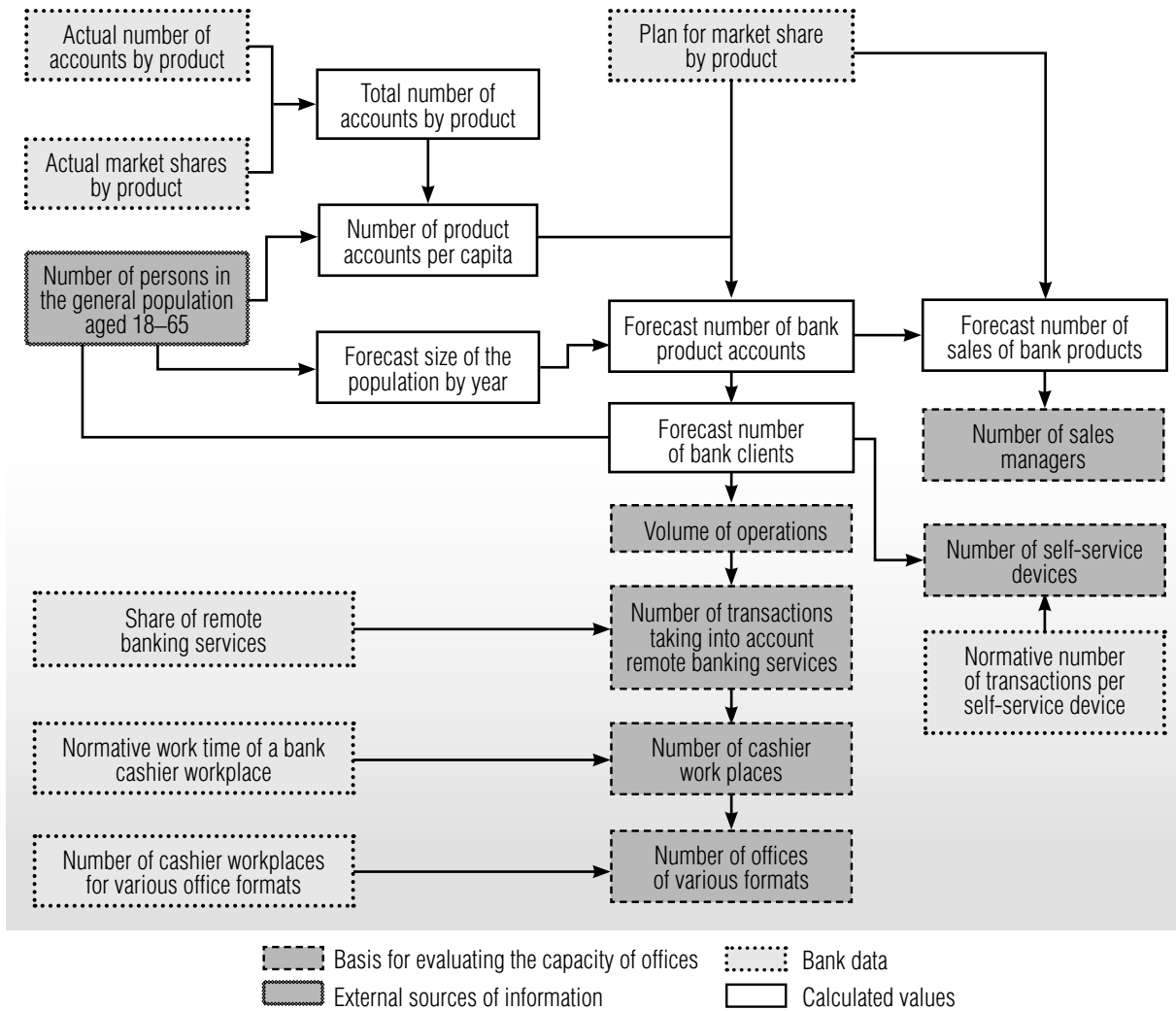


Fig. 3. Algorithm to calculate the capacity of the network servicing individuals

EAP is the economically active population;
NA is the number of product accounts per person;
TMS is the target market share of the bank for deposits within the *RF* subject.
 The calculation of the target number of bank clients is done by dividing the target number of bank deposit accounts by the number of bank deposit accounts per person. The total number of clients is calculated via the cross-selling ratio, which is an input parameter of the model. The formula for calculating the given indicator looks as follows:

$$TNC = \max \left[\left(\frac{TNA_d + TNA_{hl} + TNA_{cc}}{NA_d} \right) \cdot \frac{TNA_d}{CSR}, \frac{TNA_d}{NA_d} \right], \quad (4)$$

where *TNC* is the target number of bank clients;
TNA_d is the target number of deposit accounts;
NA_d is the number of deposit accounts per person;
TNA_{hl} is the target number of housing loan accounts;

TNA_{cc} is the target number of customer credit accounts;

CSR is the cross-selling ratio.

The calculation of the target number of work places for cashiers is made on the basis of the annual operating time fund per work place (assuming a work week of 40 hours and taking into account the given change in the ratio of the cashier's useful work time):

$$TN_{wpc} = \frac{OTF - OTF_{rbs}}{OTF_c}, \quad (5)$$

where TN_{wpc} is the target number of work places for cashiers;

OTF is the target annual operating time fund;

RBS is the share of transactions going through channels of remote banking service (RBS);

OTF_{rbs} is the target annual operating time fund of remote banking service;

OTF_c is the annual operating time fund of one work place for cashiers.

It should be noted that the target annual operating time fund can be divided by channels of remote banking service (RBS) and bank offices in accordance with the target share of remote banking service which has been adopted within the bank for the future.

The calculation of the target number of bank sales offices must be done on the assumption that the bank's office network can be represented by various formats. The credit organization can use its own criteria to determine the necessary number of offices of various format types. For example, one can assume that there should be from 5 to 10 cashiers per office.

The number of self-service devices can be calculated based on the target structure of transactions by channels of remote banking services (RBS) and the normative throughput capacity of the self-service devices of various types. The calculation of the target number can be made in the following way:

$$TN_{atm} = \frac{TAT_{rbs} \cdot S_{atm}}{NT_{atm}}, \quad (6)$$

$$TN_{pt} = \frac{TAT_{rbs} \cdot S_{pt}}{NT_{ps}}, \quad (7)$$

where TN_{atm} is the target number of ATMs;

TN_{pt} is the target number of payment terminals;

TAT_{rbs} is the target annual transactions of remote banking service;

S_{atm} is the share of ATMs in the total volume of transactions carried out by RBS channels;

NT_{atm} is the normative throughput of ATMs;

S_{pt} is the share of payment terminals in the total number of transactions carried out using RBS channels;

NT_{pt} is the normative throughput of payment terminals.

The calculation of the planned number of sales managers for banking products in an office can be made using the following formula:

$$N_{sm} = \sum_m \frac{N_{sn} \cdot S_{ps} \cdot C_{ns}}{C_l \cdot N_{ws} \cdot NS}, \quad (8)$$

where N_{sm} is the number of sales managers;

N_{sn} is the average monthly number of product sales made in a sales office;

S_{ps} is the share of product sales made by sales managers;

C_{ns} is the coefficient of normative sales 'n' per sales manager per day;

C_l is the coefficient of correction taking into account the incompleteness of volume of sales operations given the de facto load. The use of the correction coefficient is necessary in order to take into account auxiliary time, breaks required by the technology and organization of labor, the time for relaxation and personal needs, etc. The value of the coefficient can be determined in detail by a photo of the working time of the sales manager;

N_{ws} is the number of work shifts of the sales manager per month;

NS is the normative of sales per day per sales manager.

4. Determination of the habitats of preferred location of offices (stage 2)

Before proceeding to the determination of the location of points of sale on the local market, it is necessary to perform preliminary work on modeling the habitats of the preferred office locations. For this, one can use specialized software such as ArcGIS¹. Before starting the modeling, one must determine the boundaries of the city's local markets within which the possibilities of locating sales offices will be evaluated.

A local market is characterized by uniformity, compactness and geographic contiguity. For determination of the boundaries of local markets for bank services, one can take into account the following criteria (*Table 3*).

In the next step of the given stage, we superimpose the strata of data on the city map. One after another, we set down the following strata of data: location of sales office, population figures (reflecting the density of population by the intensity of filling in this stratum), key generators of customer flow, public transport stops, major shopping centers, the places of existing and promising new legal entity clients (if they are serviced by the sales offices). A source of the data for analysis of the distribution of legal entities in the city is the place of their registration. Such information can be downloaded from the statistics base RUSLANA² (or from any similar database). The potential boundaries of the local market for legal-entity clients of the bank are those obstacles that are hard to overcome such as mainly rivers and railroads with overloaded bridges and crossings. If such obstacles significantly lengthen the route and/or increase the travel time due to traffic jams, then they must be shown on the map as bound-

Table 3.

Criteria for determining the boundaries of local markets for banking services in a city

Basic criteria	Meaning
Uniformity of type of construction	Type of district (outlying residential, city center, mixed, industrial zone, private sector)
No obstacles to free movement of pedestrians	Examples of boundaries: park/forest zones, highways, railways, waterways, industrial zones, city border
No division of the local market into parts by broad bands of zero population	Width of a band is greater than 1 km

aries separating markets from one another, so that each market services its own client base.

Possible suppliers of information to fill in the respective strata of data are shown in *Table 4*.

After the strata are set down on the city map, one has to calculate the rasters of densities according to the evaluation maps. A raster is a structure which is a grid of variously colored points which display zones of activity or density. With the help of the rasters, one can understand where there is a cluster of interesting buildings and activities. One can also perform a verification of the accuracy of the data – an analysis of the coincidence of the cluster of densities with the real situation in the city.

For clusters of cities with varying populations, we establish a differentiated minimal assessment score: for cities with low population, the minimal score is significantly lower than for large cities.

The potential of a place for a bank office must be evaluated through sales and transaction

¹ ArcGIS – a family of geo-information software products from the American company ESRI. They are applied for land cadasters, in the tasks of land development, to account for real estate, systems of utilities, geodesics and subsoil use, as well as in other fields

² RUSLANA – a database containing information about more than 850,000 Russian companies, as well as around 365,000 companies in the Ukraine

Table 4.

**Possible suppliers of information to fill in data
on the city map in the process of geomarketing modeling**

Data	Presumptive supplier
Sales offices	Information of the respective bank subdivision
Population figures	State statistics bureaus
Generators of additional customer flow (shops, restaurants, entertainment centers)	Navteq ³
Public transport stops	2GIS ⁴
Major shopping centers, shopping and entertainment complexes, major food retailers (hypermarkets)	
Chain-store retailers (supermarkets)	
Data about the registration addresses of legal entities (in case the sales offices service legal entities)	AMADEUS ⁵ / database RUSLANA

Table 5.

**Criteria for the potential of an office location
in terms of sales and servicing transactions**

Factor	Criteria	Description	Sales	Operations
Population	Population numbers within a radius of 250 m	For each 100 people living in a radius of 250 m., the sales point get a score of 1.5. The stipulated maximum score for this criterion is 100	10%	55%
Sources of client flow	Number of sources of client flow (metro stations, food stores, restaurants) in a radius of 250 m	For each unit source of client flow a score of 1.25 is assigned. Maximum score - 100	35%	10%
Transport stops	Number of public transport stops in a radius of 250 m.	The sum of public transport lines is calculated for each stop in a radius of 250 m. and for each unit of this sum the sales point is assigned a score of 3. Maximum score – 100	15%	15%
Pedestrian flow	Pedestrian flow in front of the entrance to the building of the sales office	For every 10 people passing by the sales office every 20 minutes on both sides of the street, the sales point is assigned a score of 1. Maximum score – 100	40%	20%

³ Navteq (abbreviation of the English name “Navigation Technologies”; NYSE: NVT) is an American company, world leader in production of digital maps and other data for geo-information systems used in numerous software packages, including navigation systems. In 2012, the company underwent rebranding and it now operates under the HERE brand name

⁴ 2GIS is a Russian IT company which produces electronic handbooks bearing the same name with city maps. For each organization in the handbook, they provide the address, telephone number, working hours, internet address and location of the entrance to the building. In addition, the books may contain information specific to the organization’s type of activity

⁵ AMADEUS is a pan-European database (more than 9 million European companies). For each company in the system, there is a block of general information including contacts, industry affiliation, number of workers and other information, financial accounting for several years, ownership structure and subsidiaries (moreover, this structure is multi-level). The distinguishing feature of this database is how it structures the information

servicing (with the help of the map of geomarketing evaluation of sales offices). The criteria their description and weighting of sales and operations are shown in *Table 5* (weighting shown in the table relates to offices specializing in servicing individuals; their correlation can be corrected based on bank policy and the structure of the organization’s portfolio of retail services).

To find the zones of highest activity, the rasters of density are assembled using coefficients from the evaluation map. Data for calculating the coefficients are shown in *Table 6* (100% for each evaluation is obtained by adding the value of the pedestrian flow). To obtain the coefficient, one must multiply the assigned score by the necessary percent.

Table 6.

Data for building rasters of density to analyze the potential volumes of sales and operations

Criteria	Sales	Operations
Population (score of 1.5 per 100 people)	10%	55%
Transport stops (score 3 per line)	15%	15%
Generators of client flow (score 1.25 per unit)	35%	10%

As we see from *Table 6*, the coefficients will be different for evaluation of sales and operations. The bank can correct them at its discretion based on its specialization (servicing the population or active sales). The result is two rasters for analysis of potential volumes of sales and operations. After determining the preferred locations of sales offices, the most attractive places on the map are visualized with the help of the map of geomarketing evaluation of the network of sales offices.

On the map of the local market, we place the raster of assessment scores (without taking into account pedestrian flows) separately by criteria of evaluation of the sales potential of transac-

tion servicing and we determine the aggregate of places with the highest assessment scores for the given city. At the same time, depending on the result of the ‘addition’ of two strata, the following situations can arise:

1. Places where there is a coincidence of high scores of potential sales and transaction load (on condition that the distance between geometric centers of the places (habitats) do not exceed 500 meters) are viewed as potential for locating large size sales offices (expanded / flagman or large basic);
2. Places with a high score for sales with relatively low score for transaction load must be seen as potential for locating mini-sales offices;
3. With an inverse correlation between scores, the place can appear to be potentially attractive for locating a basic sales office.

The summary roster obtained in the preceding step is reclassified by 20 quantiles and is color coded. The color range for the raster of sales and transactions is differentiated. As a result, we get a set of habitats ranged by activity and covering around 30% of the city. These habitats are the preferred locations of sales offices. Initially we examine habitats with predominance of the highest evaluation score (5%), then preference is given to habitats with less activity.

5. Determination of the rational coverage of the local market taking into account payback of the sales offices (stage 3)

After the prospective zones of the organization’s presence are determined on the map, it is necessary to perform a search for suitable premises and create the target network of offices. Then comes a search for suitable premises (taking into account the stop factors and the results of payback calculations), which in the future will become elements of the target network of offices for servicing clients in the city. The list of stop factors is shown in *Table 7*.

Table 7.

Stop factors influencing the selection of locations for bank offices

Stop factors	Size of office servicing clients	
	Large	Typical
Location on the first line of houses next to a thoroughfare/street	Hard factor	Hard factor
Location on ground floor of a building	Hard factor	Hard factor
No facade windows	Hard factor	Hard factor
Height of ceilings in the premises	3 m.	2.7 m.
Time in transit to nearest office servicing bank clients	20 minutes	60 minutes
No parking for clients / possibilities to arrange it	Hard factor	Soft factor
No possibility to arrange a separate entrance to the area servicing clients	Hard factor	Soft factor

As the table shows, all stop factors can be divided into two groups:

- ◆ ‘hard’ stop factors: locating offices in premises corresponding to at least one of these characteristics is not permissible;
- ◆ ‘soft’ stop factors: locating offices in premises corresponding to such characteristics is allowed, but on condition of justifying such a necessity by the business needs of the organization.

Each place examined for possibly locating a sales office in the local market should be checked for the existence of stop factors. Responsible executives must go out to the site and evaluate the possibility of locating a sales office there. Out of the remaining premises after evaluation of stop factors, one or two are selected, those with the greatest pedestrian flow on both sides of the street immediately in front of the entrance. The intensity of the flow is measured directly, and the value obtained is added to the assessment score of the place. If the calculated full assessment score turns out to be lower than the passing score for all premises examined, then the place is excluded from the list of potential locations for situating a sales office. From the selected places, one builds the

ideal network of sales offices for the coming 3–4 years.

The capacity calculated for the local market by number of work places is next distributed among all the places proportionally to the respective assessment scores. The number of work places for specialists at each office location determines the refined format and overall throughput capacity of the office. On the basis of comparison of current and ideal networks, one builds the plan for development over the coming several years.

Such are the general recommendations for performing modeling for the rational layout of a network of sales offices for a commercial bank. Use of the described methodological approaches to geomarketing modeling enables the commercial bank to solve the task of laying out its offices. At the same time, outside the framework of the project of modeling such an office network there remain questions that are directly linked with the model for managing them. Among these, for example, are the questions on the necessary quantity of office managers, the distribution of responsibility for decision making on closing/opening part of the subdivision, filling out the staffing schedule, recruiting employees, etc.

Conclusion

In this way, the practical use of geomarketing modeling enables any bank to achieve rational spatial locations of their offices on local markets where they are present. Applicable to the banking sphere, in a short term perspective geomarketing enables the bank to determine the most preferable layout of offices servicing clients in cities from the point of view of efficiency. In the long term perspective, a credit organization obtains a tool for developing and implement-

ing a plan to develop its branch network in the coming years, all of which in the final analysis makes it possible to significantly raise the effectiveness of management decision making as a whole in the bank. The use of tools of geomarketing modeling enables the bank's senior management to effectively use the capacities of its branch network, to react in a timely way to changes in client flows in the city, to achieve greatest effect from managing offices in cities while following its strategic plans. ■

References

1. Soenen L.A. (1974) Locating bank branches. *Industrial Marketing Management*, no. 3 (21), pp. 211–228.
2. Tsvetkov V.Y. (2002) *Geomarketing: prikladnye zadachi i metody* [Geomarketing: Applied tasks and methods]. Moscow: Finance and Statistics (in Russian).
3. Zhurkin I.G., Tsvetkov V.Y. (1998) GIS i geomarketing [GIS and geomarketing]. *Izvestiya Vuzov. Geodesy and aerophotography*, no. 3, pp. 146–150 (in Russian).
4. Bakaeva V.V., Kurushina N.V. (2011) Podkhody k optimizatsii razmeshcheniya predpriyatiy roznichnoy trgovli [Approaches to optimization of retail enterprises location]. *Russian Entrepreneurship*, no 6–2, pp. 160–166 (in Russian).
5. Bell R.R., Zabriskie N.B. (1978) Assisting marketing decisions by computer mapping: A branch banking application. *Journal of Marketing Research*, no. 15 (1), pp. 122–128.
6. Kiselgof S.G. (2011) Modeli restrukturizatsii filial'noy seti kommercheskogo banka [Models of restructuring of a commercial bank's branch network]. Proceedings of *XI International Academic Conference on Economic and Social Development. Moscow, 6–8 April 2010*. Vol. 1 (ed. E.G. Yasin). Moscow: HSE, pp. 594–605 (in Russian).
7. Boufounou P.V. (1995) Evaluating bank branch location and performance: A case study. *European Journal of Operational Research*, no. 87, pp. 389–402.
8. Huff L.D. (1964) Defining and estimating a trading area. *Journal of Marketing*, no. 28, pp. 34–38.
9. Berman O. (1997) Deterministic flow-demand location problems. *Journal of the Operational Research Society*, vol. 48, no. 1, pp. 75–81.
10. Craig S.C., Ghosh A. (1983) Formulating retail location strategy in a changing environment. *Journal of Marketing*, vol. 47, no. 3, pp. 56–68.
11. Ioannou G., Mavri M. (2007) Performance-net: Decision support system for reconfiguring a bank's branch network. *International Journal of Management Science*, no. 35 (2), pp. 190–201.
12. Min H.A. (1989) Model-based decision support system for locating banks. *Information and Management*, no. 17, pp. 207–215.
13. Aleskerov F.T., Belousova V.Y. (2007) Effektivnoe razvitie filial'noy seti kommercheskogo banka [Efficient development of a commercial bank's branch network]. *Management in a Credit Institution*, no. 6, pp. 23–34 (in Russian).
14. Aleskerov F.T., Belousova V.Y. (2007) Otsenka privlekatel'nosti regionov RF dlya razvitiya filial'noy seti kommercheskogo banka [Evaluating attractiveness of RF regions for development of a commercial bank's branch network]. *Banking Business*, no. 8, pp. 54–57 (in Russian).
15. Kiselgof S.G., Aleskerov F.T. (2009) Razmeshchenie otdeleniy banka. Obzor problemy [Location of a bank's branches. Overview of the problem]. *Business Informatics*, no. 1 (07), pp. 59–69 (in Russian).

An architectural approach to IT–business alignment

Pavel V. Malyzhenkov

*Associate Professor, Department of Information Systems and Technologies
National Research University Higher School of Economics
Address: 25/12, Bolshaya Pecherskaya Street, Nizhniy Novgorod, 603155, Russian Federation
E-mail: pmalyzhenkov@hse.ru*

Marina I. Ivanova

*Student, Business Informatics MSc Program
National Research University Higher School of Economics
Address: 25/12, Bolshaya Pecherskaya Street, Nizhniy Novgorod, 603155, Russian Federation
E-mail: miivanova_1@edu.hse.ru*

Abstract

Information technologies have evolved from their traditional back-office role to a strategic resource role that can not only support but also shape business strategies. Still, different analytical sources indicate that only a small number of projects complete on-time and on-budget, leading to initially specified goals and results. The main problem is the creation and formalization of an IT–business alignment mechanism which for over a decade has been ranked as a top-priority management concern and is widely covered in theoretical literature. However, the field is dominated by conceptual studies, while there is little research on practical ways to achieve the alignment. Moreover, most of the existing research focuses on the alignment assessment using questionnaire methods based on the subjective judgement of IT and business executives. From this point of view, Enterprise Architecture development as a methodological approach to the mutually aligned business and IT architectures' design, represents a suitable tool for solving this problem. However, most of the existing EA approaches do not distinguish between different IT–business alignment perspectives. This paper attempts to provide practical guidance for IT–business alignment as well as strategic guidance for EA development by integrating the traditional Strategic Alignment Model and the TOGAF framework.

Key words: IT–business alignment, Enterprise Architecture, SAM, TOGAF, Alloy Analyzer.

Citation: Malyzhenkov P.V., Ivanova M.I. (2017) An architectural approach to IT–business alignment. *Business Informatics*, no. 3 (41), pp. 56–64.
DOI: 10.17323/1998-0663.2017.3.56.64.

Introduction

In the context of today’s dynamic, highly competitive business environment, it is crucial for a company’s survival that it acquire a high level of strategic flexibility, which, in turn, requires agile organizational structure and processes, and, therefore, flexible underlying information system architecture. As research into global IT trends shows [1], there has been a consistent increase in IT budgets across various companies for the last four years (*Figure 1*).

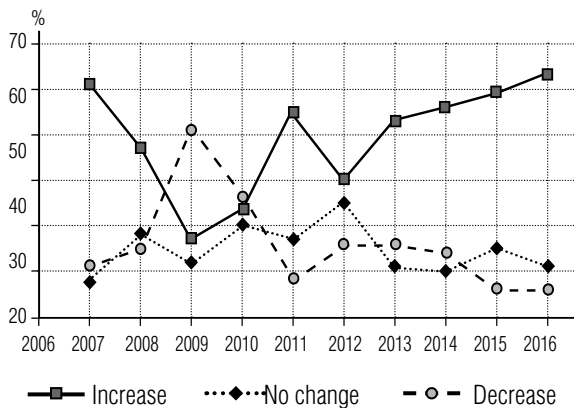


Fig. 1. IT budgets (based on the data provided in [1])

But as the IT project budgets increase, their success rates do not rise. Research into IT project delivery results conducted by the Standish Group (<http://www.standishgroup.com/>) in 2015 on 50,000 IT projects around the world indicates that only 30% of projects reviewed completed on-time and on-budget, with all features and functions as initially specified. Meanwhile, 52% of the projects examined completed over-budget, over the time estimate, and with fewer features and functions than originally specified. Overall, across all companies, the average cost overrun was 189% of the original cost estimate, and the average time overrun was 222%, which, obviously, does not contribute to return on IT investments. Moreover, 19% of projects were cancelled at some point during the development cycle.

Thus, IT, or as it sometimes called digital, strategy is not just about automating existing processes, or making cost savings, or managing complex systems using information technologies. As the Harvey Nash/KPMG CIO Survey states, “the real value of digital strategy lies in how you stitch digital technologies together to create competitive advantage and business growth” [2]. The real problem stands in the grade of coherence between IT and business strategies.

In fact, the strategic alignment of business and information systems has consistently been reported as one of the key concerns of business and IT managers across different industries. From this point of view, it represents an important issue in the field of organizational modeling of enterprises. According to research on international IT management trends [1], the alignment issue has held its position in the top-three key concerns of IT managers since 2000, along with business agility, productivity and cost reduction (*Figure 2*).

As one can see, the problem of IT cost reduction has never risen above fourth place, confirming the fact that business leaders do realize and admit the possible contribution of IT

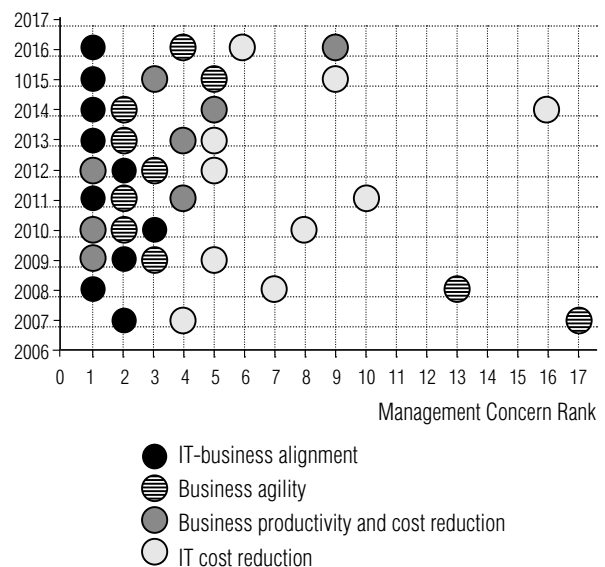


Fig. 2. Top IT management concerns (based on the data provided in [1])

to the organization's value chain and are ready to spend resources on IT projects, but only if there is a positive return on IT investments, which directly depends on the degree of coherence between IT and business within an organization.

The trend is caused by benefits alignment may provide, as well as risks entailed by misalignment. IT–business alignment enables an organization to enhance its flexibility and maximize return on IT investments, which, in turn, leads to increased profitability and sustainable competitive advantage. Thus, failure in leveraging IT may have a considerable negative effect on a firm's performance and viability [3–7].

This paper offers an approach to IT–business alignment based on the TOGAF framework and the Strategic Alignment Model. It is organized as follows: Section 1 summarizes the theoretical background relevant for the approach proposed; Section 2 presents the linkage between the main components of SAM and TOGAF frameworks; in the final section, conclusions are drawn and areas for future research are identified.

1. Theoretical background

An extensive body of research has been conducted on the nature of the IT–business alignment, criteria for its evaluation and approaches to address the issue. IT–business alignment can be determined as “the extent to which the IS strategy supports and is supported by the business strategy” [8] or as “the degree to which the IT mission, objectives and plans support and are supported by business mission, objectives and plans” [9].

However, many researchers consider alignment to be not a static state that can be measured at a single point in time but rather a continuous process of adjustment of business and IT domains [6, 8, 10, 11].

One of the most fundamental and well-recognized alignment frameworks is the Strategic

Alignment Model (SAM) [6, 12]. *Figure 3* is a schematic representation of the SAM illustrating an integration of the business domain, consisting of business strategy and organizational infrastructure and processes, and the IT domain represented by IT strategy and IS infrastructure and processes. The authors distinguish between two types of domain integration:

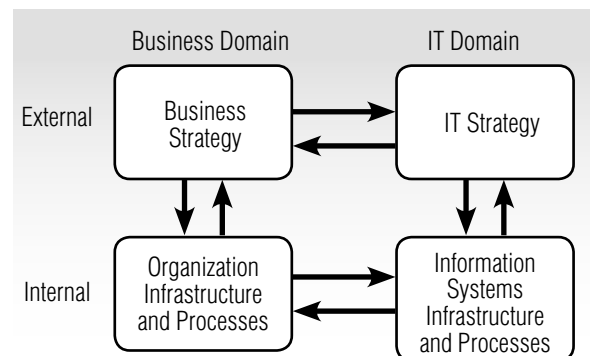


Fig. 3. Strategic Alignment Model (adapted from [12])

1. Strategic integration (of external business and IT domains): the link between business and IT strategies reflecting the capability to leverage IT strategy to both shape and support business strategy;

2. Operational integration (of internal business and IT domains): the link between organizational infrastructure and processes and IS infrastructure and processes reflecting coherence between internal customer requirements and expectations and the delivery capability within the IS function.

In order to ensure the right balance between the choices made across all four domains, it is vital to review multivariate, cross-domain relationships. SAM distinguishes between four dominant cross-domain relationships called alignment perspectives (*Figure 4*) based on the premise that strategic alignment can only be attained when three of the four domains are in alignment. Therefore, changes in one domain affect at least two of the three remaining domains.

The four alignment perspectives may be divided into two groups, based on what kind of a strategy (business or IT) is considered as a driving force.

1. Perspective One: Strategy execution. Business strategy is the driver of both organization and information systems design choices. Strategy is formulated by the top management and executed by the IS management;

2. Perspective Two: Technology transformation. Business strategy drives the development of supporting IT strategy and corresponding IS infrastructure and processes. Business managers seek to identify the best possible IT competencies to support the business strategy. IT managers are responsible for efficient design and implementation of the IS architecture consistent with the IT strategy chosen;

3. Perspective Three: Competitive potential. IT strategy is the driving force of new prod-

ucts and services fostering business strategy and organizational infrastructure and processes modifications. IS management identifies and interprets trends existing in the IT environment that may be considered as opportunities to gain competitive advantage or as a threat to the company's market position. Business executives are able to articulate how to leverage emerging IT capabilities to transform business strategy;

4. Perspective Four: Service level. IT strategy drives the development of IS infrastructure and processes with corresponding implications for the organizational infrastructure and processes. IT executives seek to provide the best possible service to the internal client by developing and implementing the appropriate basis for IS architecture redesign. Business managers are responsible for IT resource allocation and project prioritization.

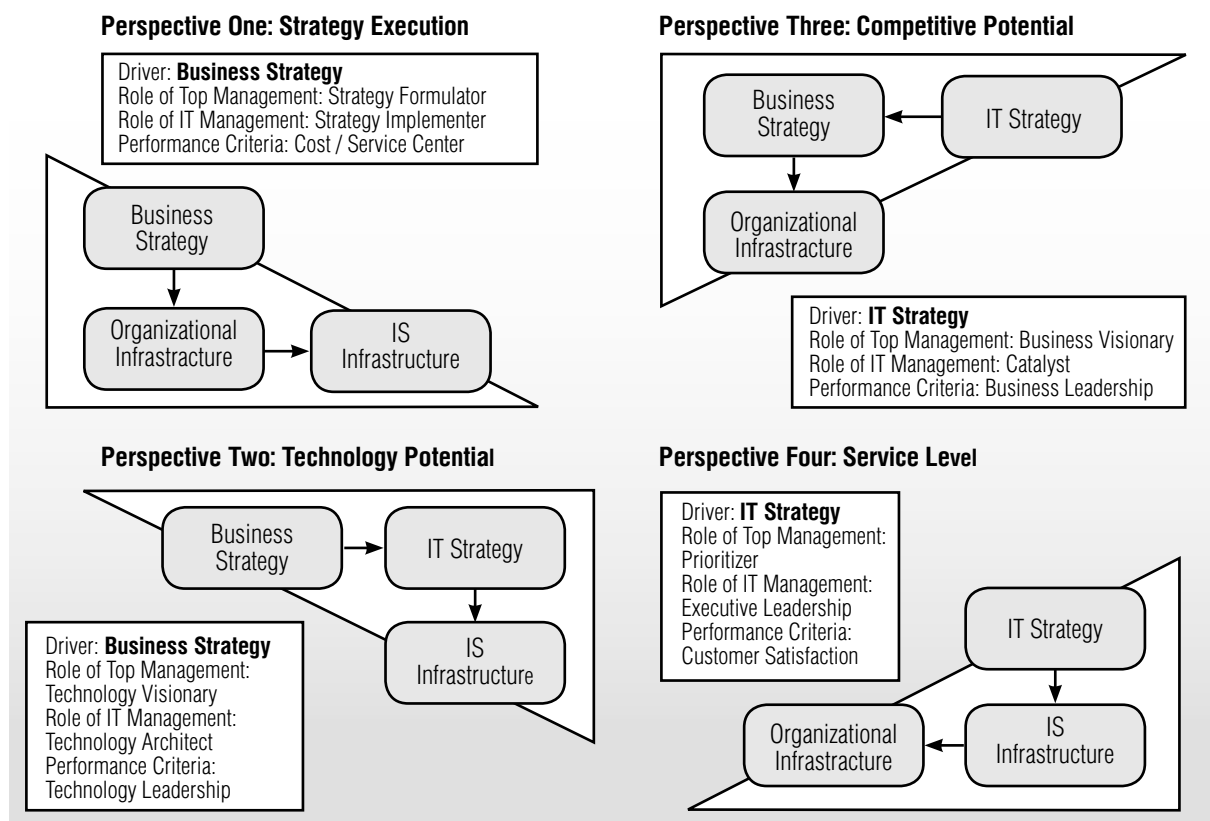


Fig. 4. Alignment perspectives (adapted from [12])

After alternative strategic choices within the four dominant alignment perspectives have been analyzed and evaluated, one or more perspectives should be selected and adopted as the driving force of organizational transformation towards strategic alignment.

The original Strategic Alignment Model is purely conceptual. Therefore, model extensions were later proposed. Thus, in [8] the original SAM was reviewed in a more practical perspective, identifying the major enablers and inhibitors of IT–business alignment, but the model itself was not elaborated. The SAM was also expanded with additional domain components related to information and knowledge management [13] and then combined with the Integrated Architecture Framework in order to enhance practical applicability of the alignment concept. Still, no tool was offered for misalignment detection or capture if alignment processes needed to be established within each of the four alignment perspectives.

This need may be filled by the concept of Enterprise Architecture, which is defined as a set of models and definitions describing the structure of an enterprise, its subsystems and the relationships between them, terminology to employ and guiding principles for design and future evolution [14]. EA development is a continuous iterative process which may be approached using EA frameworks including tools, techniques, process model, artefact descriptions and guidance for EA design.

An integration of the EA framework with the traditional SAM may contribute to the solution of the problems mentioned. We believe that although different in scope, they may complement each other.

Thus, as an attempt to fill the gap in practical guidance for IT–business alignment using SAM as well as in strategic guidance for EA development using EA framework, we propose to combine Henderson and Venkatraman’s SAM with The Open Group Architecture Framework (TOGAF) [15].

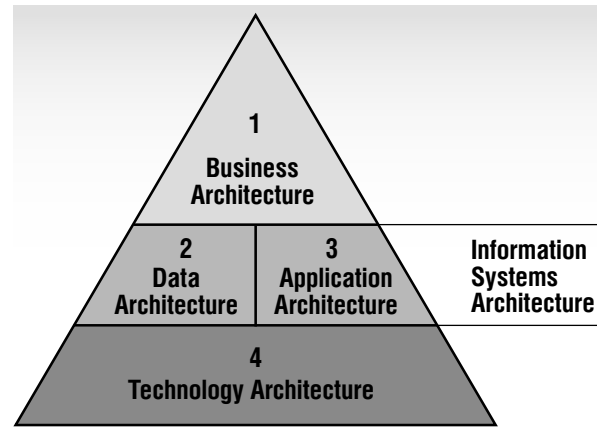


Fig. 5. Architecture domains according to TOGAF [15]

The framework supports four architecture domains that are commonly accepted as subsets of an overall Enterprise Architecture (*Figure 5*):

1. Business Architecture (business strategy, organization structure and processes, business governance);
2. Data Architecture (structure of organization’s data assets and data management resources);
3. Application Architecture (application portfolio);
4. Technology Architecture (software and hardware capabilities including IT infrastructure, networks, communications, standards, etc.).

In this framework, the Data Architecture and Application Architecture domains form the Information Systems Architecture domain.

TOGAF is based on an iterative process model called Architecture Development Method (ADM), consisting of different phases, each provided with its own objectives, approaches, inputs, steps (activities) and outputs [15]:

- ◆ Preliminary Phase. Preparation and initiation activities required for Enterprise Architecture design including customization of TOGAF and definition of the Architecture Principles.
- ◆ Phase A: Architecture Vision. Development of a high-level vision of the capabilities and business value to be delivered as a result of the proposed Enterprise Architecture.

◆ Phase B: Business Architecture. Development of a Business Architecture supporting the approved Architecture Vision; namely, description of the product / service strategy, organizational, process, information and geographical aspects of the business environment.

◆ Phase C: Information Systems Architectures. Development of the Information Systems (Data and Applications) Architecture supporting the agreed Business Architecture and Architecture Vision.

◆ Phase D: Technology Architecture. Development of the Technology Architecture supporting the chosen Information Systems (Data and Application) Architectures as well as Business Architecture and Architecture Vision.

◆ Phase E: Opportunities and Solutions. Identification of projects, programs and / or portfolios that effectively deliver the Target Architectures identified in previous phases.

◆ Phase F: Migration Planning. Planning the transition from the Baseline to the Target Architectures by finalizing a detailed Implementation and Migration Plan.

◆ Phase G: Implementation Governance. Development of the implementation architectural oversight.

◆ Phase H: Architecture Change Management. Establishment of architecture change management procedures.

It is also important to note that the framework has strong documentation support: each of the ADM phases is provided with a set of templates (catalogs, matrices, diagrams, deliverables). For example, the first phase may take advantage of the “Business Principles, Goals, Drivers” and “Architecture Vision” templates; the second may employ “Architecture Definition” and “Architecture Requirements Specification” templates and so on [15].

From the IT–business alignment point of view, TOGAF presents the same disadvantages as most EA design methodologies: it follows the predefined scheme supposing “top-down”

EA design, from business strategy and structure to information systems supporting infrastructure. It does not take into consideration different alignment perspectives while different situations of IT–business misalignment require different approaches to EA design.

In this way, TOGAF and SAM integration may contribute to the problems mentioned above; being different in the scope these methodologies complement each other:

◆ EA design process may incorporate diverse alignment perspectives when guided by SAM;

◆ SAM may be operationalized using methodologies, tools and techniques provided by the EA framework.

2. An integration of the Strategic Alignment Model and the TOGAF framework

Table 1 demonstrates the link between the main SAM components (four integration domains) and TOGAF methodology (EA domains).

Table 1.

Correspondence of SAM and TOGAF domains

Strategic Alignment Model domains	TOGAF domains
External and internal business domains	Business Architecture
Internal IT domain	Application, Data and Technology Architecture
External IT domain	no clear match

As one can see, SAM’s external IT domain does not seem to have a clear match because TOGAF does not explicitly determine the IT strategy or its essential components such as IT vision, goals and objectives, justification of IT investments etc. However, it is reasonable to assume that the IT strategy is formulated and

implemented (which may constitute a broad and fertile field of future research) as part of the overall TOGAF's Information Systems Architecture domain.

Our proposal consists in the following: each TOGAF architecture domain is covered by some ADM phases. A, B, C (Data), C (Application) and D are used to develop baseline and target Business / Data / Application / Technology Architectures and analyze the gap between them. Thus, ADM phases A–D may be used to detect the misalignment between SAM's business and IT domains and identify the target aligned architectures. Then, the next ADM phase E allows us to identify ways of eliminating misalignment by identifying projects, programs and/or portfolios that effectively deliver the target aligned architectures. Finally, ADM phases F–H guide the alignment implementation (the transition from the baseline to the target architectures).

Hence, the proposed framework may be used to measure the alignment by identifying inter-relationships and establishing correspondence between artifacts delivered by ADM phases in different SAM's integration domains. *Figure 6* illustrates the application of SAM's "Strategy Execution" alignment perspective to the

TOGAF framework for misalignment detection.

We can affirm that external and internal business domains are considered aligned if every goal and objective identified by the ADM phase A's "Driver / Goal / Objective Catalog" is covered with some services in the ADM phase B's "Goal / Objective / Service Diagram". Consequently, internal business and IT domains are considered aligned if these services are covered by some applications defined by the ADM phase C's "Application / Function Matrix". This, in turn, should operate the data from the ADM phase C's "Application / Data Matrix" and be based on the technologies identified by the ADM phase D's "Application / Technology Matrix". If, for example, "Driver / Goal / Objective Catalog" contains some goals which are not covered by the "Goal / Objective / Service Diagram", then some new business processes should be introduced in order to implement the business strategy.

Moreover, the proposed framework may be used to create new strategies, structures and processes already aligned across business and IT domains. *Figure 7* illustrates the application of SAM's "Competitive Potential" alignment perspective to the TOGAF framework for cre-

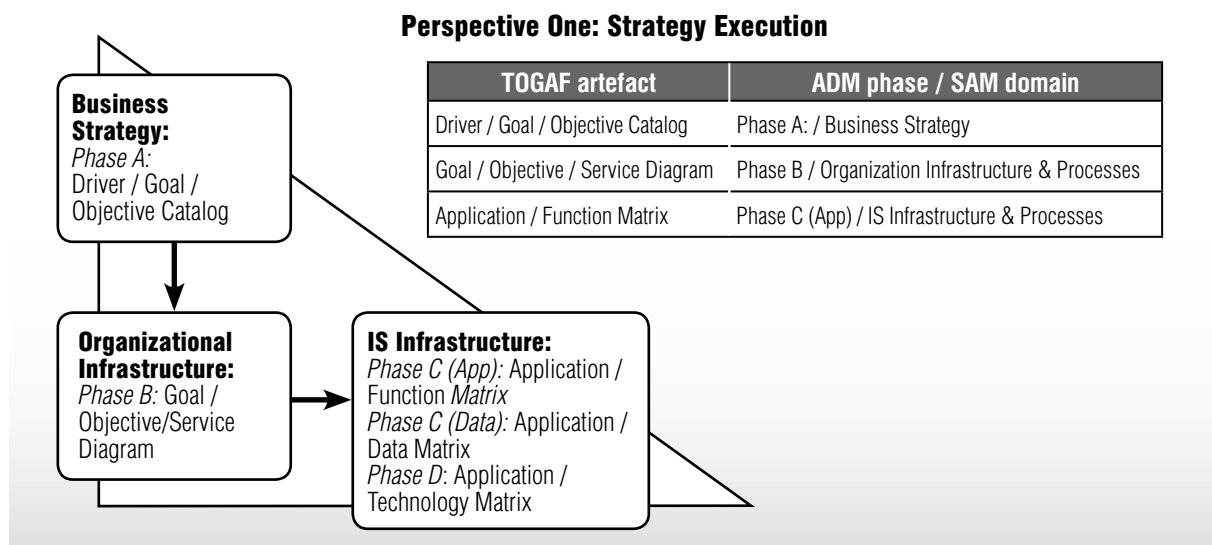


Fig. 6. SAM "Strategy Execution" alignment perspective and TOGAF

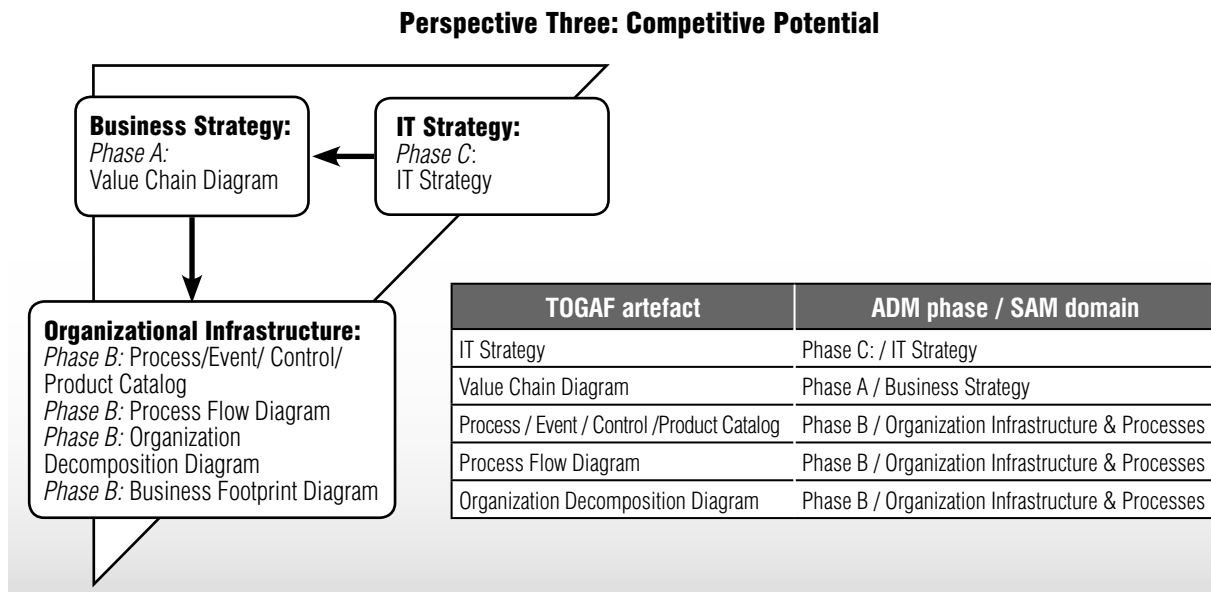


Fig. 7. SAM “Competitive Potential” alignment perspective and TOGAF

ating new business capabilities by leveraging IT opportunities. In this case, if, for example, some production operation is automated, then there will be process changes which should be reflected in the ADM phase B’s “Process / Event / Control / Product Catalog” and “Process Flow Diagram”. And the organizational structure may change: the department manually performing the operation which is to be automated may be disbanded; this should be reflected by the ADM phase B’s “Organization Decomposition Diagram”.

The formalization of our proposal may be realized by means of Alloy language (<http://alloy.mit.edu/alloy/>) which presents the following main characteristics:

- ◆ a structural modelling language, based on first-order logic, for expressing complex structural constraints and behavior of relational models;
- ◆ a constraint solver that provides fully automatic simulation and checking of relational models.

An Alloy model represents a collection of constraints that describes (implicitly) a set of structures, for example, all the possible inter-

dependencies and interrelationships between the layers of Enterprise Architecture. Alloy’s tool, the Alloy Analyzer, is a constraint solver which may be used both to explore the model by generating sample structures, and to check properties of the model by generating counterexamples.

In the context of the approach proposed, the Alloy language and tool may be used to model and analyze a system consisting of artifacts (system structures), delivered by the ADM phases of TOGAF within diverse alignment perspectives, and their interrelationships and interdependencies (system constraints).

Conclusion

In this paper, we establish formal criteria for IT–business alignment by integrating the Strategic Alignment Model with TOGAF. The approach presented, unlike some previous studies (e.g. [16]), allows us to consider the alignment of IT and business components of EA within different alignment perspectives where the driving force is either the business strategy affecting organizational and IT design choices, or the IT strategy fostering business

strategy and organizational changes. Further research activities follow but are not limited to the following branches:

1. specification of each ADM phases' inputs and outputs for each of the SAM's alignment perspectives;
2. formalization of the contents of the IT strategy document;

3. formalization of the alignment assessment model for the resulting EA evaluation;

4. practical application of the proposed framework, taking into account such business factors as industry sector, organizational size and type of strategic positioning. ■

References

1. Global Institute of IT Management (2016) *The changing role of the CIO and top IT executives*. Available at: <https://www.globaliim.com/app/download/967701357/2016+CIO+IT+Trends.pdf> (accessed 05 April 2017).
2. Harvey Nash / KPMG (2015) *Harvey Nash / KPMG CIO Survey 2015 in association with KPMG. Into an age of disruption*. Available at: https://www.harveynash.com/group/mediacentre/Harvey_Nash_CIO_Survey_2015.pdf (accessed 05 April 2017).
3. Byrd A., Lewis B.R., Bryan R.W. (2006) The leveraging influence of strategic alignment on IT investment: An empirical examination. *Information and Management*, vol. 43, no. 3, pp. 308–321.
4. Chan Y.E., Sabherwal R., Thatcher J.B. (2006) Antecedents and outcomes of strategic IS alignment: An empirical investigation. *IEEE Transaction on Engineering Management*, vol. 53, no. 1, pp. 27–47.
5. Gerow J.E., Grover V., Thatcher J.B., Roth P.L. (2014) Looking toward the future of IT–business strategic alignment through the past: A meta-analysis. *MIS Quarterly*, vol. 38, no. 4, pp. 1059–1085.
6. Henderson J.C., Venkatraman N. (1993) Strategic alignment: Leveraging information technology for transforming organizations. *IBM Systems Journal*, vol. 38, no. 1, pp. 4–16.
7. Kearns G.S., Sabherwal R. (2007) Strategic alignment between business and information technology: A knowledge-based view of behaviors, outcome, and consequences. *Journal of Management Information Systems*, vol. 23, no. 3, pp. 129–162.
8. Luftman J.N., Lewis P.R., Oldach S.H. (1993) Transforming the enterprise: The alignment of business and information technology strategies. *IBM Systems Journal*, vol. 32, no. 1, pp. 198–221.
9. Reich B.H., Benbasat I. (1996) Measuring the linkage between business and information technology objectives. *MIS Quarterly*, vol. 20, no. 1, pp. 55–81.
10. Baets W. (1992) Aligning information systems with business strategy. *Journal of Strategic Information Systems*, vol. 1, no. 4, pp. 205–213.
11. Broadbent M., Weill P. (1993) Improving business and information strategy alignment: Learning from the banking industry. *IBM Systems Journal*, vol. 32, no. 1, pp. 162–179.
12. Venkatraman N., Henderson J.C., Oldach S. (1993) Continuous strategic alignment: Exploiting information technology capabilities for competitive success. *European Management Journal*, vol. 11, no. 2, pp. 139–149.
13. Maes R., Rijssenbrij D., Truijens O., Goedvolk H. (2000) *Redefining business–IT alignment through a unified framework*. Primevera Working Paper Series. Amsterdam: University of Amsterdam.
14. Alaeddini M., Asgari H., Gharibi A., Rashidi R.M. (2016) Leveraging business–IT alignment through Enterprise Architecture – an empirical study to estimate the extents. *Information Technology and Management*, vol. 18, no. 1, pp. 55–82.
15. The Open Group (2011) *TOGAF® Version 9.1, an Open Group Standard*. Available at: <http://pubs.opengroup.org/architecture/togaf9-doc/arch/> (accessed 05/04/2017).
16. Kurniawan N.B., Suhardi (2013) Enterprise Architecture design for ensuring strategic business IT alignment (Integrating SAMM with TOGAF 9.1). Proceedings of the 2013 *Joint International Conference on Rural Information and Communication Technology and Electric-Vehicle Technology* (rICT and ICEV-T 2013). Bandung, Indonesia, 26–28 November 2013, pp. 186–192.

Process of distribution of undesirable information in social networks

Marina V. Tumbinskaya

*Associate Professor, Department of Information Protection Systems
Kazan National Research Technical University named after A.N. Tupolev
Address: 10, Karl Marx Street, Kazan, 420111, Russian Federation
E-mail: tumbinskaya@inbox.ru*

Abstract

Currently, users of online social networks increasingly use them to promote business, distribute advertisements for goods and services, engage in leisure, hobbies, personal communication and information exchange. Thus, social networks have become an open source of information for malicious users. Hackers use various ways to implement attacks, one of which is the spread of unsolicited (targeted) information. Successful distribution of unsolicited information entails the implementation of an attack scenario and achievement of the hacker's aim. In this regard, hackers have an interest in involving so-called social networking community leaders (users who have a high level of trust and influence among a large number of community users), who are able to successfully implement part of the attack scenario of the attacker.

This article presents the results of the study in three situations: the user/potential hacker's dissemination of targeted information on the social network, receipt of targeted information by users of the social network, and counteraction and prevention of the dissemination of targeted information on the social network. Experimental data are described and their analysis is presented.

A method of protection from targeted information disseminated on social networks is identified, allowing for an increase in the level of protection of social network users' personal data and personal information and ensuring the reliability of information.

The results of the research will help prevent threats to information security, counteract attacks by hackers, who often use methods of competitive intelligence and social engineering through the use of countermeasures, develop a model of protection against targeted information and implement specialized software for its integration into social networks.

Key words: online social network, targeted information, unsolicited information, hacker, information security.

Citation: Tumbinskaya M.V. (2017) Process of distribution of undesirable information in social networks. *Business Informatics*, no. 3 (41), pp. 65–76. DOI: 10.17323/1998-0663.2017.3.65.76.

Introduction

Nowadays, everyone is an Internet user, and online social networks (OSNs) are actively developing. In literature, “microblogging” is used as a synonym for “social networks”. Social networks are characterized by the ease of carrying out business promotion, the dissemination of advertisements for goods and services, leisure activities, hobbies, personal interaction and information exchange, thus serving as an open source of information for hackers. As a rule, in order to achieve their aims on social networks, hackers employ fraudulent tactics, which is confirmed by studies [1, 2]. Study [3] considers different forms of fraud on major social networks (Facebook, WhatsApp, Twitter, etc.), as well as methods and ways of combating them. One way hackers access confidential information is through the dissemination of targeted information on social networks based on methods of user manipulation [4, 5], and methods of social engineering. The concept of targeted information was formed from the concept of “targeted advertising”. There is no precise definition of the concept of “targeted information”. For this reason, the author of this article considers targeted information to be unsolicited information sent to a certain user or target audience in order to achieve the aim of the sender (for example, sales of goods and services, or, in the context of information security: receipt of confidential information, for example, personal data, usernames, passwords, etc.) using social resources. Research dedicated to targeted advertising on social networks are presented in study [6]. Issues of information dissemination in microblogging systems are addressed in study [7], while the effective dissemination of information on social networks are covered in study [8]. At the same time, the effectiveness of information dissemination is understood to be the level of correspondence between the results of information dissemination and the goal of information dissemination.

For their purposes, hackers may use social networking leaders for recruitment or involvement in terrorist groups [9, 10]. Often, leaders have a high level of trust among a large group of social network or community users, or are the founders (administrators) of communities [11, 12].

The scientific novelty of the study lies in the receipt of experimental data, which help to uncover the parameters of a potential hacker on social networks, underpinning methods of protection from targeted information, as well as forming recommendations for social network users to prevent information security incidents.

1. Examples of hacker implementation of cyberattacks using methods of social engineering

Let us consider examples of hacker implementation of cyberattacks using methods of social engineering. *Figure 1* presents the malicious software introduction process to the office computer of a social network user, while *Figure 2* displays a diagram of the process of corporate network server infection through malicious software. *Figure 3* is a UML diagram of the process hackers employ to use personal data in order to transfer funds.

2. Social information processing and the influence of factors on situations of dissemination of targeted information on social networks

A sample of this study features 2499 users of the Twitter, Facebook and VKontakte social networks who are moderators (administrators) of Russian user communities (the majority aged 17–30). All users participated in a test survey on situations of dissemination of unsolicited information on social networks and resistance to the dissemination of targeted information. Social network users participate in numerous situations linked to the dissemination of unsolicited information, both as victims and as potential hackers. For this reason, it is possible to study their decision-making process and factors in high-risk sit-

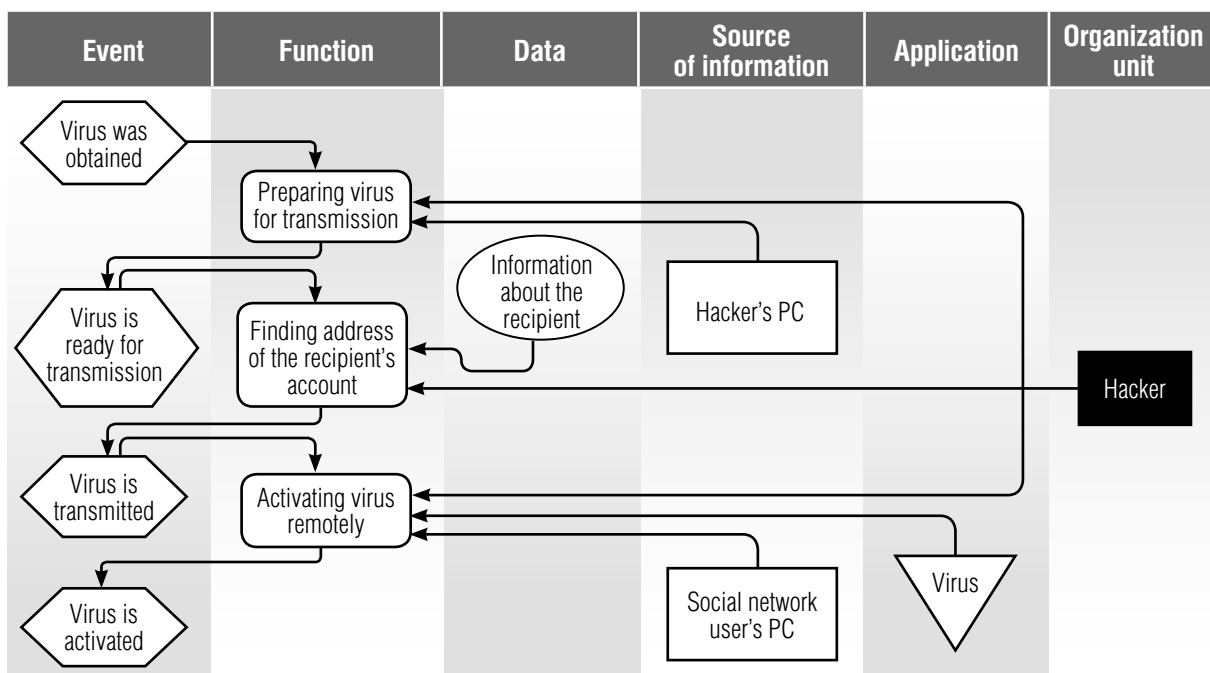


Fig. 1. Diagram of the malicious software introduction process to the office computer of a social network user

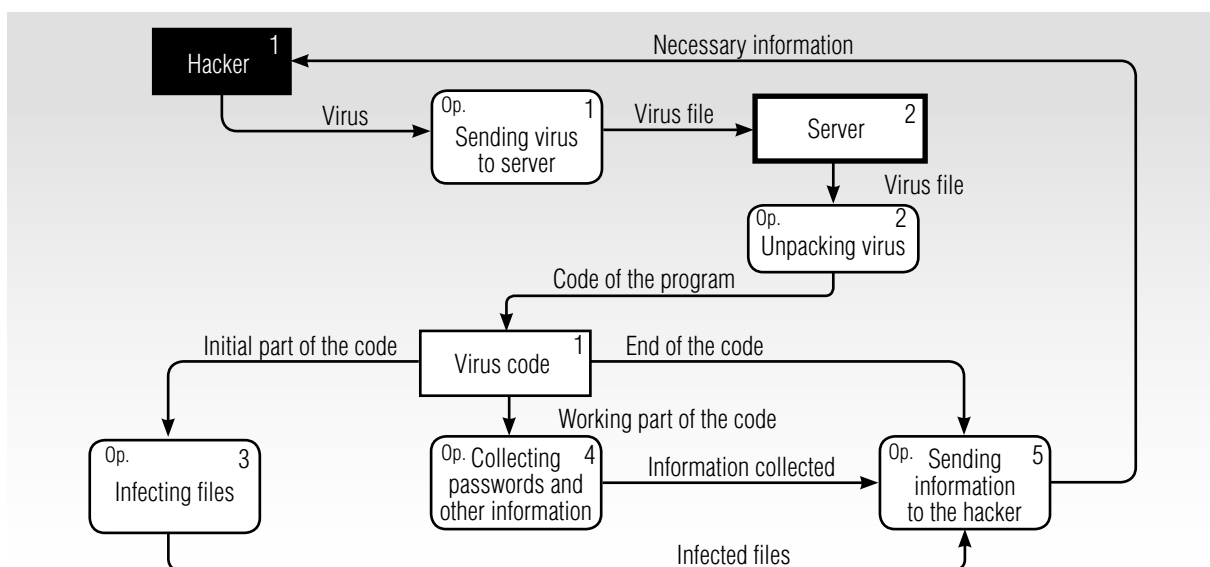


Fig. 2. DFD diagram of the process of corporate network server infection through malicious software

uations of unsolicited information dissemination on social networks.

In the study, all test surveys were anonymous and held over a six-month period in 2016–2017. A single user test survey lasted approximately one hour. The survey was held using test forms, and the results of the survey were processed using the

Statistica 10.0 statistical software package. All respondents gave written consent and voluntarily agreed to participation in the survey.

The research included studies of the influence of social information processing, along with situational and personal parameters, on increased likelihood of unsolicited information dissemi-

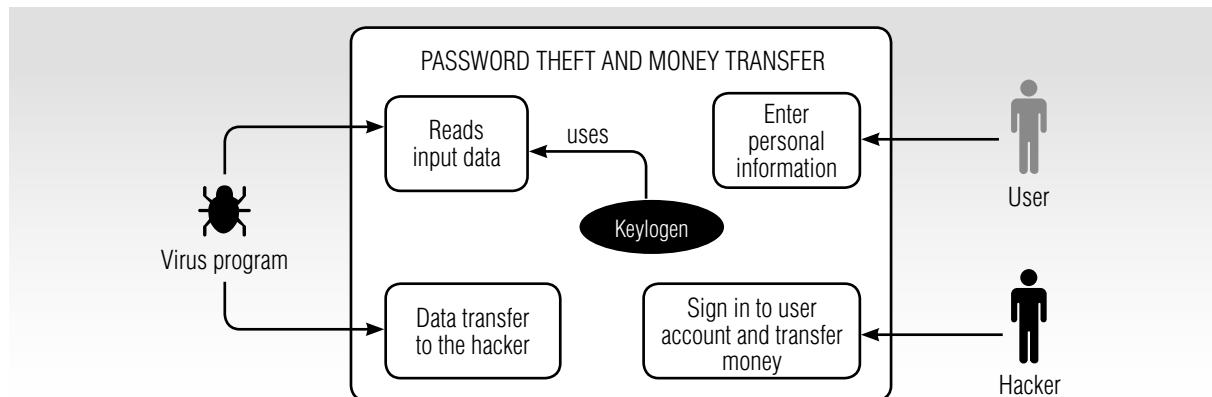


Fig. 3. UML diagram of the process hackers employ to use personal data to transfer funds

nation. To accomplish this, information was collected from respondents about the situations involving receipt and dissemination of unsolicited information that they participated in, as well as about management of those situations.

The situation of receiving targeted information is defined as potential hackers' compulsory transmission of a data message through social networks and microblogging systems to the user (potential victim) to achieve their aim. The situation of dissemination of unsolicited information involves the mass transfer by potential hackers of data messages to social network users to achieve their aim. The resistance to unsolicited information dissemination is a situation in which information dissemination, perceived by the user as possible, did not occur for any reason, for example, as a result of the blocking of a suspicious account sending spam.

The values of the test survey parameters are presented on a binary scale. All parameters take the values "0" or "1", which assists in exposing the links between them. In accordance with the social information processing theory (SIP), we are analyzing the decision-making process of hackers in situations of targeted information dissemination. SIP is a social cognitive approach based on the assumption that humans "enter social situations with a set of biologically limited possibilities and with a database of their past experience". *Table 1* shows the statistical data of a sample from 2499 respondents (with the words of the respondents).

The average age of respondents was 22 years. Nearly 75% of respondents were male. Over half of all respondents possess higher education degrees (66%). The majority of respondents indicated belonging to the lower class (70.1%), as many of them are students whose source of financial support are stipends and temporary work. The remaining respondents consider themselves part of the middle class. In 26% of cases, these are master's and graduate students who have the opportunity to work fully and pursue an academic career. Respondents' marital status statistics likewise demonstrate that during graduate level studies, students are unmarried (69%) or have a civil partner (24%), while only 7% are married.

During research, respondents reported of over 20,000 unsolicited messages received from different social network users. During the period of time under analysis, 33.4% of users received 4 – 10 messages containing unsolicited information, and only 11.7% of respondents noted that they had not received any such messages. In nearly 40% of cases, the sender of messages containing unsolicited information were unknown users, and in 30% of cases, messages were sent from fake accounts. Such messages came least often from friends (5%) and administrators (moderators) of different social network communities (5%). This statistic is due to the fact that friends rarely subject one other to that kind of dissemination, while community administrators (moderators) value their reputation.

Table 1.

Descriptive statistics of a sample of social network users

Label	Variable	Frequency	%
Gender			
n_1	Male	1874	74.99
n_2	Female	625	25.01
Age			
n_3	17–20 y.o.	450	18.01
n_4	20–24 y.o.	950	38.02
n_5	24–27 y.o.	774	30.97
n_6	27–30 y.o.	200	8.00
n_7	over 30 y.o.	125	5.00
Education			
n_8	Secondary	575	23.01
n_9	Basic vocational	275	11.00
n_{10}	Higher, baccalaureate	1049	41.98
n_{11}	Higher, specialist	200	8.00
n_{12}	Master's degree	200	8.00
n_{13}	Graduate (post-master's) degree	200	8.00
Marital status			
n_{14}	Unmarried	1725	69.03
n_{15}	I have a civil partner	599	23.97
n_{16}	Married	175	7.00
Financial situation			
n_{17}	Lower class	1774	70.99
n_{18}	Middle class	650	26.01
n_{19}	Upper class	75	3.00
Level of knowledge in IT field			
n_{20}	Low	100	4.00
n_{21}	Medium	2025	81.03
n_{22}	High	374	14.97
Social network membership			
n_{23}	Twitter	525	21.00
n_{24}	Facebook	550	22.00
n_{25}	Vkontakte	1424	57.00
Membership of social network community groups			
n_{26}	Hobbies, leisure	548	15.81
n_{27}	Education	575	16.58
n_{28}	Religion	577	16.65
n_{29}	Dating	630	18.17
n_{30}	Problem, disaster	552	15.92
n_{31}	Business	585	16.87
Number of followers on social network			
n_{32}	<50	999	39.98
n_{33}	50–100	625	25.01
n_{34}	100–200	375	15.01
n_{35}	200–500	375	15.01
n_{36}	>500	125	5.00
Number of friends on social network			
n_{37}	<50	125	5.00
n_{38}	50–100	500	20.01
n_{39}	100–200	1000	40.02
n_{40}	200–500	500	20.01
n_{41}	>500	374	14.97

On the content of unsolicited messages, respondents note that all proposed response options of the test survey occur: nearly 18% of instances were links to phishing sites. Values for the remaining instances ranged from 15.5% to 16.8%. These were malicious software, recruitment into terrorist groups, engagement in sus-

picious communities, spam, and even advertising for goods and services. 85.8% of respondents noted that there was not one cyberattack on their social network accounts, which is most likely due to the limited time period of the study (6 months). 79.8% of respondents believe contacting technical support services to be pointless.

Social network users very often ask one another for assistance in disseminating information, such as a call for help, etc. According to the statistics, the majority of respondents received such messages fewer than five times (39.2%) or did not receive any at all (13.5%). In agreeing to disseminate such messages, many respondents have more than one aim, such as financial gain (33.5%) or do it with a view to self-affirmation (25.7%). 72.6% of respondents noted that they achieved their aims through the dissemination of unsolicited information.

The dissemination of targeted information can be prevented through filtration of the data messages of social network users. 60% of respondents noted, that the number of key phrases (words) on the message filtration database ranges from 5 to 10. In addition, it is important to consider the semantics of key phrases (words) for message filtration.

The results of the study show that potential hackers can use different methods of unsolicited information dissemination depending on their aims. The simplest and fastest way to disseminate unsolicited information is the coercion or engagement of administrators (moderators) of social network communities, as they most often possess high levels of trust among users. In such cases, hackers' have high chances of achieving their aims.

Descriptive statistics (over 6 months) of a sample from 2499 users about possible situations of targeted information dissemination on social networks are shown in *Table 2*.

As a result of experimental research, three situations were considered:

1. Situations of dissemination of targeted information on social networks by users/potential hackers;
2. Situations of receipt of targeted information by social network users;
3. Situations of counteracting and preventing the dissemination of targeted information on social networks.

During studies of situation No. 1 (dissemination of targeted information on social networks by users/potential hackers), the following parameter correlation was identified.

Table 2.

Descriptive statistics of possible situations of targeted information dissemination on social networks

Values	Variable	Frequency	%
Number of recipients of unsolicited messages			
n_{41}	did not receive	293	11.72
n_{41}	under 3 times	732	29.29
n_{41}	4–10 times	835	33.41
n_{41}	11–15 times	328	13.13
n_{41}	16–20 times	210	8.40
n_{41}	over 20 times	101	4.04
Who was a sender of unsolicited messages on social networks			
n_{48}	Social network community users	500	20.01
n_{49}	Social network moderator (administrator)	125	5.00
n_{50}	Fake account	750	30.01
n_{51}	Friend	125	5.00
n_{52}	Unknown user	999	39.98
Content of unsolicited messages			
n_{53}	Link to a malicious code	388	15.53
n_{54}	Link to a phishing site	449	17.96
n_{55}	Engagement in terrorist groups	407	16.29
n_{56}	Engagement in suspicious groups	415	16.6
n_{57}	Spam	422	16.89
n_{58}	Advertisement of goods, services	418	16.73
Number of successfully-implemented cyberattacks on your account			
n_{59}	none	2145	85.83
n_{60}	under 3 times	353	14.13
n_{61}	4–10 times	1	0.04
n_{62}	11–15 times	0	0.00
n_{63}	over 15 times	0	0.00

Values	Variable	Frequency	%
Number of requests to technical support services			
n_{64}	did not contact	1994	79.79
n_{65}	under 5 times	266	10.64
n_{66}	5–20 times	204	8.16
n_{67}	20–30 times	35	1.40
n_{68}	30–50 times	0	0.00
n_{69}	over 50 times	0	0.00
Number of requests to block certain users made to social network moderators (administrators)			
n_{70}	did not contact	1637	65.51
n_{71}	under 5 times	676	27.05
n_{72}	5–20 times	142	5.68
n_{73}	20–30 times	0	0.00
n_{74}	30–50 times	44	1.76
n_{75}	over 50 times	0	0.00
Number of proposals to disseminate unsolicited information to users from your community you received as social network community moderator (administrator)			
n_{76}	none were received	337	13.49
n_{77}	under 5 times	980	39.22
n_{78}	5–20 times	690	27.61
n_{79}	20–30 times	152	6.08
n_{80}	30–50 times	171	6.84
n_{81}	over 50 times	169	6.76
What was your aim in agreeing to disseminate unsolicited data messages to your community users			
n_{82}	Financial gain	1791	33.48
n_{83}	Self-affirmation	1374	25.69
n_{84}	Revenge against social network community	77	1.44
n_{85}	Revenge against employer	207	3.87
n_{86}	Competitive intelligence	174	3.25
n_{87}	Extremism	88	1.65

Values	Variable	Frequency	%
n_{88}	Hooliganism	143	2.67
n_{89}	Recruitment into terrorist groups	39	0.73
n_{90}	Engagement in hacker groups	205	3.83
n_{91}	Research, interest	1251	23.39
Did you achieve your aim through dissemination of unsolicited information, having agreed to distribute data messages			
n_{92}	yes	1814	72.59
n_{93}	none	685	27.41
How many times did you contact technical support services to request blocking user accounts that disseminated unsolicited information			
n_{94}	did not contact	250	10.00
n_{95}	under 5 times	874	34.97
n_{96}	5–20 times	1000	40.02
n_{97}	20–30 times	250	10.00
n_{98}	over 30 times	125	5.00
How many key phrases / words are there in the community database (where you are moderator) for message filtration			
n_{99}	under 5	75.	3.00
n_{100}	5–10	1500	60.03
n_{101}	10–15	249	9.96
n_{102}	15–20	500	20.01
n_{103}	over 20	175	7.00
Number of users in your community			
n_{104}	≥ 50	204	8.16
n_{105}	50–150	587	23.49
n_{106}	150–300	969	38.77
n_{107}	300–500	153	6.12
n_{108}	500–1000	382	15.30
n_{109}	over 1000	204	8.16

In over 34% of instances, users of each social network (Twitter, Facebook, VKontakte) received unsolicited messages 4–10 times dur-

ing the period under analysis. In 48% of cases, senders were Twitter social network community users, in 40% of cases they were fake Face-

book social network accounts, while in 34% of cases they were fake accounts on the VKontakte social network. The dissemination of unsolicited information most frequently takes place on the Twitter social network (44%). The analysis suggests that users of Twitter social network communities are inclined to disseminate targeted information. It may be assumed that the typical sender of unsolicited information on social networks is male ($n_1 \geq 60\%$), aged 20–27 years ($n_4 \geq 40\%$, $n_5 \geq 30\%$), has an undergraduate degree ($n_{10} \geq 35\%$), unmarried ($n_{14} \geq 65\%$), with a medium level of knowledge in the IT field ($n_{21} \geq 80\%$), and using the Twitter social network.

In 50% of cases, users of “problem, disaster” social network communities received data messages under three times, while in 41% of cases users of “dating” communities received messages 4–10 times. In the “education” community group, users receive unsolicited data messages from unknown users or fake accounts, which indicates hackers’ choices of such a group of communities for the dissemination of information. In more than 50% of cases, dissemination is carried out by either “problem, disaster” community users (55%), thereby seeking vulnerable users for engagement in suspicious or terrorist groups, or fake accounts (50%) in “business” communities. The analysis suggests that the typical sender of unsolicited information is male ($x_1 \geq 61\%$), aged 20–27 years, has an undergraduate degree ($x_{10} \geq 33\%$), is unmarried ($x_{14} \geq 65\%$), with a medium level of knowledge in the IT field ($x_{21} \geq 50\%$), and a member of “problem, disaster”, “dating”, or “business” community groups.

In 31–60% of cases, social network users received unsolicited messages 4–10 times during the period under analysis (n_{44}), and, in 32–46% of cases, under 3 times. Most often (50%) senders are unknown users with over 1000 friends on social networks, in 46% of cases these are fake accounts with 200–500 friends on social networks, and in 40% under 50 friends. The analysis suggests that users with less than 50 friends on a

social network are inclined to the dissemination of unsolicited information. It may be assumed that the typical sender of unsolicited information on social networks is male ($n_1 \geq 62\%$), aged 20–27 years ($n_4 \geq 36\%$, $n_5 = 62\%$), with a basic vocational ($n_9 = 40\%$) or undergraduate degree ($n_{10} \geq 41\%$), unmarried ($n_{14} \geq 50\%$), primarily with a medium level of knowledge in the IT field ($n_{21} \geq 40\%$), and possessing under 50 friends on social networks.

During studies of situation No. 2 (receipt of targeted information by social network users), the following parameter correlation was identified.

Most frequently (44%), Twitter social network users receive unsolicited messages from community users (48%) with the message information content: 2% – engagement in suspicious groups, 26% spam, 26% advertisement of goods, services. In 35% of cases, unsolicited messages are received by Facebook social network users from fake accounts (40%) with the message information content: 35% spam, 35% advertisement of goods, services. In 34% of cases, unsolicited messages are received by VKontakte social network users from fake accounts with offers advertising goods, services (31%). It may be assumed that fake accounts on Facebook and VKontakte social networks send spam and advertisements, while Twitter social network users additionally send information on engagement in suspicious groups.

In the “hobbies, leisure” and “dating” community groups of Twitter and VKontakte social networks, no more than 15% of cases of attempts to implement cyberattacks were recorded under three times. In the “hobbies, leisure” community group of the Facebook social network, hackers attempted to implement cyberattacks more than 15 times in 10% of cases. It may be assumed that potential perpetrators for the implementation of cyberattacks most frequently choose users belonging to the “hobbies, leisure” community group of Facebook social network, “problem, disaster” of Twitter social network, and “dating”

and “business” of VKontakte social network.

During studies of situation No. 3 (counteracting and preventing the dissemination of targeted information on social networks), the following parameter correlation was identified.

Users of the social networks under analysis do not contact moderators (administrators) to request blocking certain community users ($\geq 52\%$), however, in 41% of cases, Twitter social network users contacted moderators (administrators) under 5 times. Social network users belonging to community groups $n_{26}-n_{31}$, most often do not contact moderators (administrators) to request blocking certain community users, or contact them rarely. This means that users do not pay sufficient attention to the informational security policies of social networks.

Most frequently (5–20 times) in 30% of cases, proposals to send messages with unsolicited content are sent to Twitter social network moderators with communities ranging from 150 to 300 users. Less often (under 5 times) in 49% of cases, proposals to send messages with unsolicited content are sent to VKontakte social network moderators with communities ranging from 150 to 300 users, and, in 45% of cases to Facebook network moderators with communities from 150 to 300 users. It may be assumed that, in order to achieve their aims, potential hackers choose to disseminate unsolicited information through moderators (administrators) of communities with 150–300 users. On the Twitter social network, the most vulnerable communities proved to be “hobbies, leisure”, and “religion”, on the Facebook network these were “hobbies, leisure”, and on the VKontakte network “religion” and “dating”.

41% of users of the social networks under analysis contact technical support services with requests to block user accounts that disseminate unsolicited information. Users of the “dating” communities of Facebook and VKontakte social networks most frequently make requests, as well as the “education” and “religion” communities of the Twitter social network.

Most often on social networks (and specifically in the “hobbies, leisure” community groups of the Facebook social network, “religion” on the Twitter social network, and “dating” on the VKontakte network) in more than 44% of cases, 5–10 key phrases are used for message filtration.

Interpretation of the research results demonstrates that potential hackers, disseminating unsolicited information containing spam and advertisements for goods and services, are: male, aged 20–27 years, with higher education, unmarried, with a medium level of knowledge in the IT field, possessing under 50 friends on social networks and concealing their data with a fake account. Most vulnerable are users belonging to “hobbies, leisure”, “problem, disaster”, “dating”, and “business” community groups. Users rarely contact community moderators (administrators) or technical support groups in cases of the emergence of suspicious users. Most vulnerable are users/moderators (administrators) of community groups ranging from 150–300 users. The number of key phrases for the filtration of unsolicited messages (5–10) is insufficient to ensure social network information security.

3. Method for protection from the dissemination of targeted information on virtual social networks

Based on investigation of the situations of dissemination of targeted information on virtual social networks, a method for protection (*Figure 4*) is proposed, which is constituted by the steps of the following sequence:

1. Classification of social network users;
2. Protection of social network leaders;
3. Improvement of the rules of user message filtration;
4. Development of recommendations for protection from the dissemination of targeted information on social networks.

Leaders of social networks are understood to be users possessing a high level of trust and influence among a large number of community users

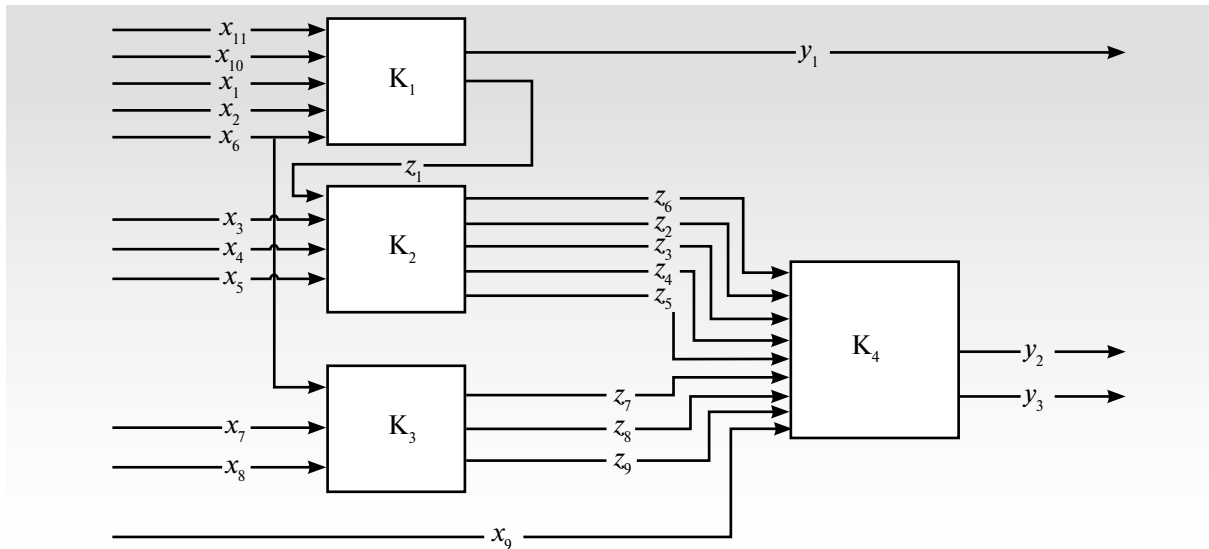


Fig. 4. Block diagram of the protection method from targeted information

capable of successfully implementing part of the actions of hackers' computer attacks. As a rule, leaders are moderators (administrators) of social network communities. This method may be formally presented in the following manner:

◆ $K = \{k_1, k_2, k_3, k_4\}$ – a set of functional blocks of the method, where k_1 – classification of social network users; k_2 – protection of social network leaders; k_3 – improvement of the rules of user message filtration; k_4 – development of recommendations for protection from the dissemination of targeted information on social networks;

◆ $X = \{x_i | i = \overline{1, n}\}$ – a set of input parameters, where x_1 – hacker types; x_2 – criteria of potential hacker classification; x_3 – antivirus software; x_4 – social network user/leader parameters; x_5 – parameters characterizing social network user/leader behavior; x_6 – a set of user messages; x_7 – criteria of user message information evaluation; x_8 – classification rules of user data messages; x_9 – rules of recommendation development for protection from targeted information; x_{10} – a set of social network users;

◆ $Z = \{z_\varphi | \varphi = \overline{1, s}\}$ – a set of internal parameters of the method, where z_1 – a list of social network leaders; z_2 – information messages on the necessity of compliance with security measures;

z_3 – authentication using technical means of communication; z_4 – social network user/leader profile; z_5 – social network user/leader database; z_6 – decision-making on account blocking; z_7 – targeted information message database; z_8 – expected messages of social network users; z_9 – unsolicited messages of social network users;

◆ $Y = \{y_j | j = \overline{1, m}\}$ – a set of output parameters of the method, where y_1 – a list of blocked users; y_2 – information messages to social network users on possible attack implementation; y_3 – recommendations about the adoption of information security measures on social networks.

The “Classification of social network users” functional block (K_1) includes:

- 1) user classification based on hacker types and identification of suspicious users/potential hackers;
- 2) potential hacker classification based on criteria/activity (action) level with respect to social network users during a set period t_1 ;
- 3) decision-making on the blocking of users based on pp. 1 and 2 of this functional block;
- 4) classification of social network users based on “social network user/leader” types.

The “Social network leader protection” functional block (K_2) includes:

1) education and warning of network leaders: introduction of social network leader education measures on the fundamentals of information security (leader accounts are critical resources; access would allow hackers to disseminate targeted information to a large number of users) by sending information messages containing reminders about the necessity of complying with information security measure;

2) implementation of technical security measures: authentication using smart phones (telephones), the use of antivirus software, authentication using hardware, automatic password checks in compliance with information security recommendations;

3) analysis of leader behavior on social networks: development of user profiles (identification of user parameters and their threshold values), creation of databases of user activity, upgrade of user activity databases, classification of user behavior on social networks, development of models of dynamic changes in user profiles and an algorithm of abnormal user behavior identification. If user behavior on networks is abnormal, then a notification is sent informing the user of suspicious activity and subsequently blocking the account.

The “Improvement of the rules of user message filtration” functional block (K_3) is broken down into the following steps:

1) formation of a database of user messages containing targeted information disseminated on social networks, based on analyses of blocked user data;

2) development of criteria for the evaluation of user message information;

3) formation of a rule base of user message information classification;

4) database detailing of users messages containing targeted information, and their classification as expected or unsolicited based on evaluation criteria;

5) improvement of the classification rule base;

6) development of social network user message filtration model.

The “Development of recommendations for protection from the dissemination of targeted information on social network” functional block (K_4) is broken down into the following steps:

1) formation of a rule base of recommendation development for protection from targeted information;

2) social network user notification about the possible implementation of an attack (the likelihood of implementation);

3) development of recommendations on the adoption of measures to ensure information security.

We see prospects of further research of the issue of protection from targeted information in detailed study of the method and development of models of protection from targeted information on its basis. A model of protection from targeted information on social networks permits the implementation of specialized software for its integration in the major social networks, which will help users increase the security of personal information use on social networks and avoid falling prey to hackers. It is anticipated that the specialized software will be a program module (application) permitting:

✧ filtration of the personal messages and messages/records (posts) of social network community users based on the message filtration model;

✧ automatic blocking of users who send unsolicited information based on hacker types and rule bases on user blocking;

✧ provide recommendations to social network administrators (moderators) on possible threats of hacker attack implementation and the adoption of countermeasures to prevent cyberattacks on social networks.

Conclusion

The method of protection from targeted information on social networks proposed in the study facilitates the prevention of threats to information

security, prevention of hacker attempts to implement social engineering attacks, development of a model of protection from targeted information and, in the future, implement specialized software for its integration into virtual social network systems. This all will allow for external monitoring of events on social networks, as well as implementation of a search for vulnerabilities in the mechanisms of instant message exchange in order to lower the possibility of hacker attack implementation and to protect the personal information of social network users. The results of the study permit the development of recommendations for social network users to prevent incidents:

- ◆ apply and regularly update antivirus protection tools;
- ◆ update account passwords at least once a month;
- ◆ be more attentive to the information content of social network user messages, as links to mali-

cious software may be concealed in the guise of advertising links;

- ◆ be selective in opening messages in the “hobbies, leisure”, “problem, disaster”, “dating”, and “business” community groups;
- ◆ comply with social network security policies;
- ◆ contact community moderators (administrators) in cases of emergence of suspicious users;
- ◆ contact technical support groups in cases of emergence of suspicious users;
- ◆ in cases of moderation (administration) of community groups ranging from 150–300 users, check the content of messages sent to users;
- ◆ increase the number of key phrases for unsolicited message filtration.

The results of the study allow for the application of an actively evolving network approach to the study of informal communities on a new level, thereby receiving interesting and illustrative results. ■

References

1. Bradbury D. (2012) Spreading fear on Facebook. *Network Security*, no. 10, pp. 15–17.
2. Kim H.J. (2012) Online social media networking and assessing its security risks. *International Journal of Security and Its Applications*, vol. 6, no. 3, pp. 11–18.
3. Savage D., Zhang X., Yu X., Chou P., Wang Q. (2014) Anomaly detection in online social networks. *Social Networks*, no. 39, pp. 62–70.
4. Krombholz K., Hobel H., Huber M., Weippl E. (2015) Advanced social engineering attacks. *Journal of Information Security and Applications*, no. 22, pp. 113–122.
5. Richard G.B., William B.B., Lewis C. (2012) Flying under the radar: Social engineering. *International Journal of Accounting and Information Management*, vol. 20, no. 4, pp. 335–347.
6. Johnson J.P. (2013) Targeted advertising and advertising avoidance. *RAND Journal of Economics*, vol. 44, no. 1, pp. 128–144.
7. Wang L., Wang M., Guo X., Qin X. (2016) Microblog sentiment orientation detection using user interactive relationship. *Journal of Electrical and Computer Engineering*, vol. 2016, pp. 167–181.
8. Khalilov D. (2013) *Marketing v sotsial'nykh setyakh* [Marketing in social networks]. Moscow: Mann, Ivanov and Ferber (in Russian).
9. *Doktrina informatsionnoy bezopasnosti Rossiyskoy Federatsii* [Doctrine of information security of the Russian Federation]. Available at: <https://rg.ru/2016/12/06/doktrina-infobezobasnost-site-dok.html> (accessed 18 May 2017) (in Russian).
10. Klein G.R. (2015) Ideology isn't everything: Transnational terrorism, recruitment incentives, and attack casualties. *Terrorism and Political Violence*, pp. 868–887. Available at: <http://www.tandfonline.com/doi/full/10.1080/09546553.2014.961635> (accessed 20 January 2017).
11. Murzin F.A., Batura T.V., Proskuryakov A.V. (2015) Programmnyy kompleks dlya analiza dannykh iz sotsial'nykh setey [Software package from social networks data analysis]. *Programmnye produkty i sistemy*, no 4, pp. 188–197 (in Russian).
12. Nazarov A.N., Galushkin A.I., Sychev A.K. (2016) Risk-modeli i kriterii informatsionnogo protivoborstva v sotsial'nykh setyakh [Risk models and information confrontation criteria in social networks]. *T-Comm: Telecommunications and Transport*, vol. 10, no. 7, pp. 81–86 (in Russian).

Russian-French Workshop in Big Data and Applications 2017

On October 12-13, 2017, the Higher School of Economics and the Mines-Télécom Institute (IMT), supported by the Embassy of France in Russia and with the participation of Hub French Tech Moscow, organize the 2nd Russian-French workshop on Big Data technologies and their application.

The Russian-French scientific workshop “Big Data and Applications” will be organized for a second time. The first workshop of Russian and French participants from leading research centers, universities, government agencies and IT companies, led by mutual interest in the subject of the “Big Data”, took place on December 2, 2016. The 1st workshop demonstrated opportunities of cooperation between researchers and business from the two countries in the area of Big Data and became the starting point for fruitful scientific collaboration.

The main objective of the second Russian-French workshop is to strengthen the scientific and educational cooperation between organizations of both countries working in the field of Big Data. It will also lead to exchange views on using Big Data technologies in various fields and applications of science, business and public administration, all of which will definitely lead to joint projects, in particular, within the framework of the European program “Horizon 2020”. During the workshop, a special session will be organized on opportunities within the French CNT “Information and Communication Technologies”.

International Student Conference “E-business. Internet Project Management. Innovations”

On March 13–15, 2018 in Moscow, the School of Business Informatics of the National Research University Higher School of Economics will be holding the International Student Scientific-Practical Conference “E-commerce. Internet Project Management. Innovations”.

Conference attendees have the great opportunity to present their first scientific works and discuss them with well-known experts. The best works will be published in the conference proceedings.

The conference will focus on the following topics:

- ◆ E-commerce;
- ◆ Experience in organizing business on the Internet;
- ◆ Games, virtual reality;
- ◆ ICT for business and public administration;
- ◆ ICT for medicine;
- ◆ Innovation in IT;
- ◆ Internet economics;
- ◆ Internet project management;
- ◆ Marketing and advertising on the Internet;
- ◆ Network community;
- ◆ Internet of things;
- ◆ Research on the Internet;
- ◆ Legal framework of e-business.

XIX April International Academic Conference on Economic and Social Development

On April 10-13, 2018 in Moscow, 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: bjournal@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.