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

BUSINESS INFORMATICS

HSE SCIENTIFIC JOURNAL

№3(45)-2018

CONTENTS

Publisher:
National Research University
Higher School of Economics

Subscription index in the Rospechat catalog – 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 Russian Federation

Editor-in-Chief:
A. Golosov

Deputy Editor-in-Chief
S. Maltseva
Y. Koucheryavy

Computer Making-up:
O. Bogdanovich

Website Administration:
I. Khrustaleva

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

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

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

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

ABOUT THE JOURNAL

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

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

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

✦ data analysis and intelligence systems

✦ information systems and technologies in business

✦ mathematical methods and algorithms of business informatics

✦ software engineering

✦ Internet technologies

✦ business processes modeling and analysis

✦ standardization, certification, quality, innovations

✦ legal aspects of business informatics

✦ decision making and business intelligence

✦ modeling of social and economic systems

✦ information security.

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

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

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

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

EDITORIAL BOARD

ABOUT THE HIGHER SCHOOL OF ECONOMICS

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

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

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

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

ABOUT THE SCHOOL OF BUSINESS INFORMATICS

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

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

There are four graduate programs provided by the School:

- Business Informatics
- 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.

Values-directed enterprise engineering[1]

Evgeny Z. Zinder
Chairman of the Board, NCO "Foundation for System Engineering, Standardization and Project Management Support" (FOSTAS Foundation)
Address: Mailbox 563, Moscow, 107061, Russia
E-mail: ezinder@fostas.ru

Abstract

This article presents an approach to enterprise engineering (EE) enhancement under modern conditions by including in EE possibilities and capabilities of systematic decision making support based on the values analysis of an enterprise and all the engaged parties – stakeholders of different categories. The particularity of the approach consists of combining broad understanding of values, aiming at real business tasks at all the levels of enterprise management, and including digital transformations risk management in these tasks.

The approach includes both "analog" and "digital" values models, and also employs heterogeneous values models. Values are regarded as a social and economic, corporate and personal phenomenon. Subjectivity, situatedness, relativity, and changeability of values are taken into consideration. The article suggests conjoint consideration of various stakeholders' (interested parties') heterogeneous values, among others, of stakeholders affected by digital transformations. The enterprise itself as an economic agent, its owners, employees, enterprise-partners, clients/customers, communities, social organizations, and public agencies might be among the stakeholders considered One of the key requirements is values comparability and the ability to determine the gap between different stakeholders' values. Other requirements define the constitution of values models and characteristics of models elements.

One of the essential results of this research is the developed system of requirements. Besides, the structure of an open multidimensional enterprise values space is suggested as well as the structure of partial and integrated values models placed in this space is presented. The framework for tracing relations between the enterprise components and the stakeholders' values is proposed, and the possibility of introducing functions assessing conformity and nonconformity of different stakeholders' values on the integrated values model of the enterprise is demonstrated. The presented system of requirements and models enables building and applying values models which was tested in completed projects. The approach allows enabling sustainable enterprises development under digital economy conditions at the fundamental level by values-directed management. Moreover, , it might prove useful for expanding various enterprise transformation methods, including their existing value-directed and value-centered variants.

Key words: enterprise engineering; digital transformation; values; values model; risk management; values space; values-directed enterprise engineering; digital values model; analog values model.

[1] This work was supported by the Russian Foundation for Basic Research (project No. 16-07-01062 "Development of methods and resources of enterprise engineering based on smart technologies")

Citation: Zinder E.Z. (2018) Values-directed enterprise engineering. *Business Informatics*, no. 3 (45), pp. 7–19. DOI: 10.17323/1998-0663.2018.3.7.19.

Introduction

Discussions about the "digital economy" concept and recommendations for "digital transformations" (DT) [1] go on along with reports of hacking systems formerly claimed invulnerable, of limitations of artificial intelligence services, and of business models banned judicially [2]. This turbulence forces us to search for approaches to preservation of valuable features of enterprises under changing conditions, and in DT realization in particular. This publication offers a way of enterprise engineering (EE) aimed at sustainability of enterprise development and enabling actualization of various DT methods. This refers, in particular, to well-known methods such as value-oriented (for example, [3; 4]) and value-centered (for example, [5; 6]). The proposed approach builds upon broad understanding of values in contrast to its submission to one particular theory and aims at solving a wide range of business tasks including tasks of DT and sustainable development, as well as tasks of risk management in transformation of enterprises and society in a turbulent digital economy environment. Metamodels for "analog" and "digital" values models are suggested for practical support of this approach; possible directions and examples of their application are provided. Further on, this approach will be called "Values-directed enterprise engineering." This name is believed to be relevant to the vitally important, all-pervading role of values in socio-economic, organizational, managerial, production, and even bio-information [7; 8] processes.

The rest of the paper is structured as follows. Section one presents basics and the purpose of this paper and formulates the main requirement R1 for values modeling. Section two describes the features of values as a phenomenon and their use in the enterprise business tasks, establishes the necessity of more sense bearing values models than those employed in many works on EE, suggests a working definition of values for the given context, and formulates corresponding requirements R2–R5. Section three defines the requirements R6–R9 for singled out features and elements of the general values model structure corresponding to requirements R1–R9, multidimensional Open Enterprise Values Space (OEVS), the rules of formalizing an Open Integrated Values Model (OIVM) of enterprise and its stakeholders. The Conclusion discusses some options for using the proposed rules and models and indicates directions for further research and development.

Note: Stakeholder in this article is understood as an interested party of any type connected with an enterprise being analyzed and its digital transformation, including the enterprise itself as an economic agent.

1. Beginnings of this study and the primary requirement for values models

Mentioned in the Introduction, turbulence of the digital economy prompted us about a year ago to publish our analysis of the situation with respect to the development of conceptions and main EE methods [9]. This showed that the existing EE paradigm retains its efficiency in the face of radically changed organizational structures, communications, and working processes. An assumptive reason is that the basic enterprise features like mission and responsibility for producing some goods or service persist. At the same time, new abilities of DT require more careful tracing of the main guiding lines of enterprise development and paying more attention to possible threats described in the reports [1; 10; 11] as well as to still unknown threats. The need for concrete methods to satisfy these requirements induced additional studies and served as a prerequisite to this publication. Formerly, we described additional studies and their main recommendations for the following areas:

✦ continuous strategic and tactic enterprise governance [12];

✦ knowledge management in dynamically changing network business processes and cooperation and co-creation processes [13];

✦ risk management in enterprise DT projects [14][2].

Each of these studies revealed that values of individuals, their communities, and organizations play a significant or even a leading role as development guiding lines and decision making criteria in various EE key tasks, as well as in day-to-day enterprise activity. Meanwhile, values should be interpreted both broadly and in detail, including humanitarian and economic aspects, because working with the guiding lines requires comprehensive analysis of all essential aspects of an enterprise constitution and activity.

For this reason and according to the restriction of values models mentioned in sub-section 2.1 and their application in known formal models descriptions for EE, the author analyzed values as a phenomenon, examined areas of application of values descriptions at enterprises, formulated a set of requirements to values models associated with an enterprise, and defined a constructive approach to formalized modeling of values and their systems. The primary requirement for values models is:

R1. *Values models must be applicable for comprehensive definition and control of enterprise development guiding lines as well as emerging risk management, including guiding lines and risk management in DT projects and functioning of transformed enterprises.*

2. Values phenomenon analysis and requirements for values modeling

2.1. Value as a phenomenon and the basis of values models formalization

The analysis and description of human values has a long history [15]. We take this history into account and proceed to present tasks combining different views on values as a social, corporate, and personal phenomenon.

Work with values in a broad sense while forming corporate culture has long been put into practice and stated at the level of course books [16]. However, significant progress in structuring, comparing, and formalizing values models can be noticed when working with value as a cost or as a feature well expressible in a monetary form [17; 18], which is not enough for satisfying R1.

The significance of using values as a reference point in EE is recognized by many authors [3–6; 19]. However, in the known attempts of values formalization in EE, interpretations significantly constraining the real phenomenon are employed. For instance, papers [3; 19] inherently consider values within the scope of only one EE theory but not from the point of view of exploring values in real life of enterprises with the subsequent implanting of results of the exploration in EE theory. Besides, this value model appears extremely simplified and is primarily used for reflecting exchanges

[2] The paper [14], in particular, is based on the author's studies 2012–2018 where the analysis of risks and risk management methods in applying new technologies were conjoined with systematic discussions of these risks and methods in various projects and professional groups. In particular, in 2017, such risks and methods were demonstrated and discussed on the video channel "Cyber-tech" https://www.youtube.com/watch?v=6KTG9brsiNU&feature=youtu.be (accessed 05 July 2018) and at the round table "Risk management in the digital age" in Plovdiv, Bulgaria (23 November 2017)

between actors. The approach in [4] takes into consideration various points of view on values, including their nonmonetary and emotional categories to a greater extent and suggests a value metamodel taking into account actors' values of different types. However, in the context of the requirement R1, this model also appears too restricted. In particular, only two types of stakeholders are considered (the service's supplier and client), and their interconnection reflects only their interaction in a co-creation style. Values models in these cases are not expected to be appropriate for solving a wide range of business tasks in which values of public and humanitarian categories play an essential role.

Comprehensive analysis and treatment of social and humanitarian values appears more complicated. Due to this, specific methods and approaches are developed, such as, for instance, "management by values" [20], expanding complex management schemes of different scales by including humanitarian values, up to the highest spiritual values [21]. Suggestions appear simultaneously to transform the values complex of various types into some economic entity reflecting an enterprise sustainability [5], as it is supposed that businesses could be motivated to make meaningful changes only in such a case. At the same time, the use of non-formalized "analog" values models for different stakeholders' categories and their application in various business tasks, including those different from the tasks of values exchange, in particular, in risk management tasks [6], remains actual and topical.

On the whole, there is no generally accepted, sufficiently complete, and at least partially formalized metamodel of values, and comparing values is very difficult due to a number of contradictory approaches to the interpretation of the very concept of value.

Significantly, a huge number of non-formalized models have been accumulated within axiology, including models treating various aspects of values [22; 23]. These models are important, as practically any aspect might be essential for solving a specific business task. However, attempts at direct use of values in their conventional forms face the serious problem of the diversity of values manifestations. In concrete situations, any feature of a material or ideal entity can be recognized as a value. Values can be divided into values-tools (means) and values-goals [16], and be ordered differently according to priorities, and such evaluations are quite subjective and changeable. As a result, the most significant feature of the existing situation is acknowledgment of subjectivity, situatedness, relativity, and changeability of values of different interconnected parties – the enterprise stakeholders. In particular, we point out the extremely high rate of change in the picture of the world and the fact that these changes are crucial in the values variability. For instance, it was shown in [24] that, for new generations this factor leads to inverting of value systems relative to those that were considered conventional not long ago.

As mentioned in the New Philosophic Encyclopedia [22], there is methodological chaos even in the definition of the value concept and in the treatment of value relations. It seems that this chaos only means that unitary values models structured and ordered each in its own way, as was done earlier, are not suitable for reaching the complex goals mentioned above. Pragmatic ways of building and applying values models should work in real "chaotic" conditions, while remaining understandable to different people and engaged in the context of enterprise "digital life," and all this leads to the requirement R2:

R2. *It is necessary to combine forming and using values models of both the "analog" type, that is, oriented to traditional ("manual") ways of using them in enterprises, and of the "digital" type, that is, formalized for enabling their com-*

puter processing in a computerized ("digital") information environment of enterprises. In this case, "analog" and "digital" models of the same single value should not contradict each other in terms of their semantics.

2.2. About the role of common highest values and about a working definition of values

Recognizing the existence and common interpretation of some highest values by various stakeholders provides the foundation for solving many practical problems, including reaching a compromise when other common interpretations of values are missing, and forming a working definition of a value given below:

"Value is understood as the boon to an individual, a group, enterprise, and/or a society in the sense of moving towards highest values or in terms of protecting them from a shift in the opposite direction."

We discussed this and similar interpretations with different audiences. We had very important conversations with Dharia Rahi GuruMata, Head of the International Department in the ashram Atma Kutir[3]. The boon was discussed in two incarnations: as a high ethical category and as progress towards this category in everyday life, including activities of enterprises. The subjective nature of the evaluation of the boon was taken into account, and accordingly, the needs and ways of performing expert assessments of the availability of the boon and its relative scope were discussed.

The author fully recognizes that interpretations of highest values of different stakeholders can be dramatically contradicting each other. However, our times show real possibilities for a strong rapprochement between the positions of various ethical and spiritual teachings which can be caused by the awareness of the increase in common threats in all aspects (ecological, technological, and other threats). The report [11] indicates that today, and in the nearest future, universal highest values are earnestly needed to focus on the boon for all living beings and the world as a whole. This leads to the requirement R3:

R3. *The values modeling approach must provide for the possibility of encompassing highest values relevant as the common highest values at this historical stage.*

2.3. Aiming values models at business tasks solving

Details of values models and characteristics of these details are derived from understanding the needs of the models use for a wide range of practical business tasks executed in the area of EE and DT. Requirements induced by two sample groups of business tasks of enterprise transformation and functioning were explored in this study. Later they were generalized for expansion and use in other tasks and formulated as the requirements R4 and R5.

The first example includes typical business tasks of different scales:

✦ formation of the enterprise development strategy, including comparison of enterprise DT directions;

✦ search for the enterprise's future products (services) on the basis of forecasts of prospective values of customers and promising technologies;

✦ testing job applicants for the concordance of their values-tools and those necessary for completing certain jobs.

The second example refers to using values

[3] Information about Atma Kutir Ashram (Garval Himalaya, India) is available at http://atmakutir.blogspot.ru

models in the sphere of risk management which was considered essential in [6]. The approach to risk management for the DT area suggested in [14] is based on:

✦ architectural analysis of the enterprise and finding connections between the enterprise components and the values these components really support (so-called traceability to real rather than declared values);

✦ comparing this values set to stakeholders' (including clients and enterprise employees) value systems and finding correspondence and possible contradictions.

In both examples, comparing values of different stakeholders and, in some cases, reaching a compromise based on values convergence are required for solving the mentioned tasks. For this purpose, it is necessary to use comparable values models of different stakeholders, in particular, in order to evaluate the concordance of values inherent for stakeholders planning DT and other stakeholders' values. This is reflected in the requirements R4 and R5:

R4. *The set of an enterprise values models must present a set of potentially heterogeneous values for all kinds of stakeholders, including those affected by DT, among those, an enterprise as an agent of the economy, its owners, employees, partners, clients, social organizations, and public authorities.*

R5. *While executing an enterprise tasks, it is reasonable to employ a single space of dimensionality for all possibly heterogeneous values models of different stakeholders' types corresponding to a wide range of elicited business tasks. This space should support the definition of values comparability and the ability to compare them basing on greater–less relation.*

Let us call the dimensionality space mentioned in the requirement R5 the "open enterprise values space (OEVS)".

3. Requirements for values models structure and elements

3.1. Traceability of the components of an enterprise to values models

It is necessary to compare factual values created by the enterprise components, their transformations, and an enterprise as a whole with the enterprise stakeholders' values. For this purpose, in turn, it is necessary to ensure traceability[4] of relations from the components of the enterprise architecture (including its behavior) to the models of the enterprise and its stakeholders' values which are included in an open integrated values model (OIVM). This leads to the requirements R6 and R7.

R6. *An integrated values model of an enterprise must facilitate traceability from any enterprise system component via enterprise goals (or even skipping goals) to the values of all the stakeholders.*

R7. *Comparable values models must have measurable elements of the same type to compare different stakeholders' values and to support execution of EE and DT tasks. The measures of these elements must be determined by the same or compatible metrics. Comparison of these elements must be done for checking equal or not equal and greater or less relations. The exact differences must be calculated when possible.*

For the tracing mentioned in R6, we use the expanded five-model DT scheme proposed in [9]. *Table 1* shows the conception of the relationships between the proposed models (line "1") of partial (column "1") and integrated (column "2") values on the one hand and architectural models in the five-model scheme of the functional architecture on the other hand, which includes services or platforms for business functions (line "3") and their business purposes in terms of supported goals (line "2").

[4] Traceability and tracing are understood similarly to ISO/IEC 15288 "Systems and software engineering. System life cycle processes"

Table 1.

Framework of tracing from components of a five-model scheme of an enterprise transformation to values models

	1. For single party or component of enterprise	2. For integrated representation of enterprise
1. Values (see Section 2)	Model of value v_{qn} for stakeholder s_q	OIVM = $\{v_{qn}\}$
2. Goals (see [9], Fig. 1)	Business goals and goal indicators achieved through a specific system or DT component	Integrated business goals indicators system; their links with OIVM and OICA
3. Components of DT (see [9])	Model c_t for a component of enterprise or DT	OICA = $\{c_t;\ c_t \times c_p \mid p \neq t\}$

In this table and hereinafter:

v_{qn} is the specific value for the stakeholder s_q from the set S of all the stakeholders of the enterprise;

OIVM is Open Integrated Values Model which includes all meaningful v_{qn} related to the enterprise;

c_t and c_p are the high level architectural models of the specific enterprise components;

OICA is Open Integrated Component Architecture which integrates all the enterprise components c_t (primarily the business functions and digital services and platforms supporting them) and their relationships.

3.2. About the heterogeneity of values of different stakeholders

Values subjectivity and situatedness leads to the possibility of incomparability of two values of the same kind and named equally by two stakeholders due to these values' different contents, for example, if the stakeholders are from different social environments or just belong to different generations.

Example: Different forms of interest for the same kind of values in a digital society. The common values category "Prestige" (at level 4 "Respect" according to [25]) can be interesting in the form "Prestige as the *h*-index of the citation", and (for other or same conditions and/or demographic groups) — in the form "Prestige as a number of *Likes* and *Followers* in social networks". These two "Prestige" values are formally incomparable in these two forms of interest as their meanings are defined in different domains of definition.

Notes: 1) Both forms of interest can be used by the same stakeholder as two different specific values of one kind; 2) Domain of definition is referred to as "domain" further in this article.

Another example. Loyalty to one's patron and social group (layer) is an essential value in the environment called "Absolute monarchy with the patrons' hierarchy," but it might not be recognized as a value or might even have a negative value in the environment called "Liberal democracy with a market economy."

These examples demonstrate that to be comparable two values must be defined on the same domain determined by the set of the same value categories. In the given example, they must belong to the values of the same social-economic environment, the same needs category, and the same form of interest. All this is reflected in the requirements R8:

R8. *Values models have to include the elements showing that the values are defined on the specific domain; and accordingly, OEVS has to include the dimensions determining the categories of values, their environments and forms which identify these domains. These dimensions have to enable checking that the various stakeholders' values pairs belong to the same domain and, thus, enable their comparability and homogeneity in this sense.*

3.3. Metrics of values meanings

Comparable values models defined on the same domain of values can differ due to the realization level for the specific values kind and form of interest. For instance, they can differ in the *h*-index of the citation, information safety index, in the level of the company's (or its employee's) responsibility, and the like.

One suppositive example is a generalized value-tool "Ability to do something" and the realization levels of this ability. For example, "1" = "no ability", "2" = "limited ability", "3" = "ability in most situations", "4" = "the highest ability." Such metrics are common in completing qualitative assessments of different objects' states. With the consideration of R7 and R8, this is expressed in R9:

R9. *For every domain of the values of the same kind and the same form of interest, it is necessary to design the metric of the realization level for specific values reflecting, when possible, ordered values meanings from the point of view of a stakeholder as a person defining and expressing these values.*

Notes:

1) As a metric, a scale of qualitative levels, a physical indicator of the level, a composite index, or another parameter can be used together with the rules of measuring a realization level of values in this metric and with the rules for comparing different values by this metric.

2) In many cases, it is the difference of realization levels of two values of the same kind and form of interest that leads to the values conflict and accordingly to the conflict of the enterprise components or decisions supporting this value kind or form of interest.

4. Instantiation and formalization of the values model and values space

The aforesaid requirements and rules for describing values models along with the rules introduced below in subsection 4.1 can be used in "as is" style as metamodels for forming textual and minimally structured "analog" values models. Meanwhile, the formal description of the values space OEVS following the scheme provided in subsections 4.1 and 4.2 establishes the metamodel of the specific values formalized "digital" models, as in this case values models are considered to be the points in OEVS and have to be formed according to its dimensions and other defined requirements.

4.1. Partial values models as the basis for OEVS dimensions

In order to meet the requirement R8, it is possible to establish OEVS dimensions by categories sets which allow anchoring domains of values in practice. A possible variant of such a categories-determinants set for establishing the dimensions is provided below:

✦ the categories (levels) of stakeholders' needs;

✦ the categories (levels) of external conditions of stakeholders as their social-economic environments;

✦ the categories of stakeholders' behavior in the external environment;

✦ values kinds of stakeholders and the levels of their realization according to the requirements R7 and R9;

✦ forms of interest for values of the specific kind.

The kind of the values is traditionally prescribed by its conventional short name (such as, for example, health, honesty, responsibility, and work quality) as well as by its meaty description. The forms of interest for specific value representations depend significantly on the values kind, and these forms might also depend on the domains in other OEVS dimensions.

The OEVS variant suggested below is based on the categories-determinants explicitly or tacitly suggested in the partial models [25–27] and their extensions. These models were selected for the OEVS variant presented due to their wide enough spreading over considerable time and due to the possibility of their extension and adaptation.

In order to fulfill the R8 requirement for the simultaneous coverage of various stakeholders' values, OEVS must also have the dimension $S = \{s_q\}$, where s_q denotes the coordinate of the particular stakeholder in the OEVS space.

The Maslow model usage. For the A. Maslow model [25], the relativity and variability of the order of the "need/value" levels arising due to different priorities in the satisfaction of different needs for different actors are recognized. The extension [28] which includes additional levels of needs/values: cognitive, aesthetic, values of self-realization and transcendental values, including altruistic values is taken into consideration. Altruistic values can be accepted as a kind of highest values mentioned in R4. This model is used in OEVS as the basis of the dimension $M = \{m_i\}$ where m_i is the level in the "extended Maslow pyramid."

For the specific business task, the dimension M can be detailed by defining sub-levels for m_i. Besides, it is possible to introduce one or another relation of order for particular conditions and tasks (for example, the Maslow's original order or the inverted order similar to [24]).

The Graves model usage. The C. Graves model (the "open theory") [26] is applied in both original (simplified in our view) and expanded modes. In a simplified version, the model is used in full accordance with its original scheme as an ordered sequence of development levels of the socio-economic environment joined with the type of the stakeholders' behavior defined as rational for this environment. In order to comply with the described business tasks, the levels of the environment with Graves codes DQ, ER, FS are primarily used. Such a model is used in OEVS as the basis of the dimension $G = \{g_e\}$, where g_e is the development level of the environment and of the stakeholders' behavior in it. In the expanded mode, it is taken into account that the environment and the stakeholder's behavior may not correspond to each other in the manner defined in [26] (an individual can choose an irrational in the view of [26] type of behavior). Taking this into account, such a mode of the partial model induces two dimensions of the OEVS:

✦ dimension $G = \{g_e\}$, where g_e is the level of development of the medium,

✦ dimension $B = \{b_h\}$, where b_h is the level of the behavior of the stakeholder.

The Schwartz model usage. For each value in OEVS, it is necessary to provide the connection of the levels of m_i, g_e, and b_h defined above with a description of the value of a specific kind. The mentioned value description is extended by the actual form of interest for the specific value representation and by measurable levels of the value realization in this form. This extended description, combined with the scales of values realization levels, is used as the basis for the aggregated dimension SH for description of OEVS values.

In the EOVS version presented, the S. Schwartz model ("the theory of basic values") [27] expanded and supplemented with the above-mentioned method is the basis for the aggregated dimension $SH = \{sh_{ijfk}\}$, where sh_{ijfk} is the combination of the elements:

✦ identifying the conventional name of the value kind (for example, "working responsibility") and its generalized description;

✦ the form of interest (in this example, it is the fixation of the fact that a stakeholder holds the characteristic of "working responsibility" expressed in one of the word descriptions predetermined for this kind and form);

✦ the level number from the scale of implementation levels for the values of this kind and form with the meaty levels interpretation (in this example, it is the level number in the scale of work responsibility with word description of characteristics of this level responsibility).

4.2. Values models and an open enterprise values space. An open integrated model of values and values metrics

Versions of values space OEVS and values models for the whole enterprise OIVM based on the version of values models structuring and formalizing described in subsection 4.1 are presented below.

Open enterprise values space OEVS and values models. OEVS is formed by the set of points v – potential values with coordinates:

✦ in the simplified version, $v = (s_q, m_i, g_e, sh_{ijfk})$, which provides a way for visual illustrations making;

✦ in the extended five-dimensional version, $v = (s_q, m_i, g_e, b_h, sh_{ijfk})$, with a potential increase in the number of dimensions up to six or seven in the case of splitting SH dimension.

Here every value v_{qn}, where $n = \{1, ..., N_q\}$, N_q is the number of different values of the stakeholder s_q from the set of stakeholders S, is represented by a point v with above-mentioned coordinates.

Open integrated values model of an enterprise OIVM. The value set of the stakeholder s_q from the set of stakeholders S is formed as his or her value system $V_q = \{v_{qn}\}$, where $n = \{1, ..., N_q\}$, N_q is the number of different values of this stakeholder.

The Open Integrated Values Model (OIVM) of an enterprise is defined as the set of essential values models of the stakeholders of the enterprise and its DT (including customers as part of the enterprise ecosystem).

That is, OIVM = $\{V_q\}$, where $q = \{1, ..., Q\}$, Q is the cardinality of the set of stakeholders S, and all the V_q are defined in the same OEVS. This provides an opportunity to match the values of different stakeholders and to determine their conformities or contradictions to each other.

Values metrics. Realization levels of certain values can be expressed in natural units of measurement (for example, "time spent on receiving a service," "the number of 'likes,'" etc.). For other values, the level of realization is expressed as a qualitative assessment, for example, the ordination number of the conditional realization level with the semantic description of the value realized at this level (see the example in subsection 3.3).

Development of such metrics is performed with each extension of the open scheme of OEVS and the values models in it, among others, for emerging new forms of interest in the SH dimension. This makes it possible to form and accumulate knowledge about values meanings so that during the EE running it is possible to gradually increase the automation level of intelligent analysis, among others for enterprises DT planning.

Conclusion

Derived models and functions for completing business tasks. For completing specific business tasks of an enterprise development, in particular, for evaluating its DT acceptability, it is useful to create and apply derived and adapted models on the basis of integrated OIVM and OICA (*Table 1*). For risk management in DT, for example, it is important to analyze the compatibility of pairwise combinations of the values which can contradict each other. In particular, this is done for evaluating combinations of the enterprise values as values supported (or disappearing because of a component exclusion) by specific functional components c_t on the one hand, and values of enterprise employees, customers, and regulating agencies' values on the other hand, which is described in [14].

In order to formalize such an assessment, it is recommendable to define function $F(c_t, v_{q_1 n_1}, v_{q_2 n_2})$ versions which give, for example, the assessment "conforms" (+), "does not conform" (–), "mixed assessment" (+/–), or "incomparable" (0) for each functional DT component depending on the compatibility assessment $v_{q_1 n_1}$ and $v_{q_2 n_2}$ for the concrete realization of the component c_t from OICA.

Defining these assessments as the meanings of the function F at the initial stage of applying OIVM models requires the "manual" assessment by experts in different subject domains and completing continuous extensions of the SH dimension by descriptions of new kinds, forms of interest, and metrics for the specific values. Such extensions are considered to be knowledge accumulation for further raising the level of function F calculation automation, EE and risk management enhancement.

Development of relations of order in OIES dimensions. It is recommended to develop the metrics scales of values realization levels so that they establish relations of order in OIES dimensions based on ordering the levels of such a scale. It is not always possible to determine such an order in all situations, or only for limited conditions of the specific task, but this enables measuring a relative distance from one value to another or to the reference value. The author and his colleagues have experience in such projects [29; 30] successfully employed in practice [31; 32].

Example: the scales of realization levels of values-tools as IT-specialists' abilities [30] such as "responsibility for the job," "ability to work autonomously," and others were developed for representation of abilities realization levels which are the values both for an individual employee and an enterprise as a whole.

General conclusion and further studies. The system of requirements R1–R9 is regarded as the primary result of this study. Implementing these requirements in the suggested OEVS and OIVM structures is only one possible version. Support of openness of the models included in the approach "values-directed enterprise engineering" provides the basis for enriching different EE methods to a certain degree based on the values in different interpretations.

At the same time, the presented approach in its primary fragments (values models structures, the principle of defining order relations in OEVS, and metrics structure of values realization levels) have already been employed in practice, for example, in the project for assessing IT-specialists' competencies as values, which confirms the feasibility of the suggested approach.

Analysis of the values models structures and experience of their use demonstrates that OIVM formalizing should be combined with accumulating knowledge about representations of actual values forms and their realization levels. This, in particular, will enable developing more enhanced systems of intellectual program agents supporting meaningful decisions in the areas of multi-agent enterprises engineering and continuous risky situations monitoring. On the whole, using knowledge about values in solving an extending range of business tasks and the suggested approach "Values-directed enterprise engineering" will facilitate enterprises sustainable development under digital economy conditions. This development is becoming more attainable due to the coverage of all business areas and enterprise components with development methods directed by values at the fundamental level and in the dynamic regime. ■

References

1. IBRD (2016) *World development report 2016: Digital dividends*. Washington: International Bank for Reconstruction and Development.
2. Alderman L. (2017) Uber dealt setback after European Court rules it is a taxi service. *The New York Times*, 20 December 2017.
3. Pombinho J., Aveiro D., Tribolet J. (2012) Towards value-oriented enterprise engineering — Relativity in service system networks. *Knowledge and Technologies in Innovative Information Systems. Proceedings of the 7th Mediterranean Conference on Information Systems (MCIS 2012). Guimaraes, Portugal, 8—10 September 2012. Lecture Notes in Business Information Processing*, vol. 129, pp. 113—124.
4. Golnam A., Viswanathan V., Moser C.I., Ritala P., Wegmann A. (2014) Designing value-oriented service systems by value map. Proceedings of the *International Symposium on Business Modeling and Software Design (BMSD 2013). Noordwijkerhout, The Netherlands, 8—10 July 2013. Lecture Notes in Business Information Processing*, vol. 173, pp. 150—173.
5. Gilding P., Hogarth M., Reed D. (2002) *Single bottom line sustainability. How a value centered approach to corporate sustainability can pay off for shareholders and society*. Sydney, Australia: Ecos Corporation.
6. Ramamoorti S., Watson M.W., Zabel M. (2008) Engineering value into enterprise risk management. *Internal Auditor*, October 2008, pp. 53—59.
7. DeLozier S., Rhodes M.G. (2015) The impact of value-directed remembering on the own-race bias. *Acta Psychologica*, no. 154, pp. 62—68.
8. Castel A.D., Benjamin A.S., Craik F.I.M., Watkins M.J. (2002) The effects of aging on selectivity and control in short-term recall. *Memory & Cognition*, vol. 30, no. 7, pp. 1078—1085.
9. Zinder E.Z. (2016) Expanding enterprise engineering paradigm. *Business Informatics*, no. 4 (38), pp. 7—18.
10. World Economic Forum (2018) *The global risks report 2018. 13th edition*. Geneva, Switzerland: World Economic Forum.
11. Weizsacker E.U., Wijkman A. (2018) *Come on! Capitalism, short-termism, population and the destruction of the planet. A report to the Club of Rome*. NY: Springer.
12. Zinder E.Z. (2017) Upravlenie balansom strategicheskogo i takticheskogo v realizatsii tsifrovykh predpriyatii i elektronnykh pravitel'stv [Managing the balance of strategy and tactics in digital enterprises and electronic governments implementation]. *Information Society*, no. 2, pp. 9—22 (in Russian).
13. Zinder E.Z., Yunatova I.G. (2017) Digital economy and knowledge barriers: Their origin and dealing with them. Proceedings of *Digital Transformation and Global Society: Second International Conference (DTGS 2017). St. Petersburg, Russia, 21—23 June 2017*, pp. 445—463.
14. Zinder E.Z. (2018) Strategicheskie "tsifrovye" riski predpriyatii i obshchestva. Dvizhenie k nepreryvnomu risk-menedzhmentu [Strategic "digital" risks of enterprises and the society: Movement to continuous risk management]. Proceedings of the *XXI Russian Scientific Conference on Enterprise Engineering and Knowledge Management (EE&KM—2018). Moscow, 26—28 April 2018*. Moscow: REU, vol. 1, pp. 276—281 (in Russian).
15. Humanitarian Technologies (2018) *Aksiologiya* [Axiology]. Available at: http://gtmarket.ru/concepts/6894 (accessed 06 July 2018) (in Russian).
16. Solomanidina T.O. (2007) *Organizatsionnaya kul'tura kompanii* [Organizational culture of a company]. Moscow: INFRA-M (in Russian).
17. Olsen E. (2002) Rethinking value-based management. *Handbook of business strategy*, pp. 286—301. Available at: https://www.bcg.com/documents/file14651.pdf (accessed 06 July 2018).
18. Tang D. (2017) What is value based management (VBM)? *Flevy Blog*. 23 March 2017. Available at: http://flevy.com/blog/what-is-value-based-management-vbm/ (accessed 06 July 2018).
19. Pombinho J.P.M. (2015) *Value-oriented enterprise transformation: Design and engineering of value networks*. PhD Thesis. Lisboa: Universidade de Lisboa.
20. Jaackson K. (2010) Management by values: Are some values better than others? *Journal of Management Development*, vol. 29, no. 9, pp. 795—806.

21. Dolan S.L., Altman Y. (2012) Managing by values: The leadership spirituality connection. *People & Strategy Journal*, vol. 35, no. 4, pp. 20–26.
22. Institute of Philosophy RAS (2018) *Aksiologiya* [Axiology]. Available at: https://iphlib.ru/greenstone3/library/collection/newphilenc/document/HASH0147b7e8f087b539ec51af47 (accessed 06 July 2018) (in Russian).
23. Humanitarian Technologies (2018) *Tsennost'* [Value]. Available at: http://gtmarket.ru/concepts/6895 (accessed 06 July 2018) (in Russian).
24. Sitkevich N.V. (2011) *Osobennosti transformatsii nravstvennykh tsennostei v usloviyakh informatsionnogo obshchestva: etiko-filosofskii analiz* [Particularities of transformation of moral values in the conditions of information society: ethical and philosophical analysis]. PhD Thesis. Novomoskovsk: D. Mendeleev University of Chemical Technology of Russia (in Russian).
25. Maslow A.H. (1962) *Toward a psychology of being.* Princeton, USA: Van Nostrand.
26. Graves C.W. (1970) Levels of existence: An open system theory of values. *Journal of Humanistic Psychology*, vol. 10, no. 2, pp. 131–155.
27. Schwartz S.H. (2012) An overview of the Schwartz theory of basic values. *Online Readings in Psychology and Culture*, vol. 2, no. 1. Available at: https://scholarworks.gvsu.edu/orpc/vol2/iss1/11/ (accessed 05 February 2018).
28. McLeod S. (2018) *Maslow's hierarchy of needs.* Available at: www.simplypsychology.org/maslow.html (accessed 05 February 2018).
29. Zinder E.Z., Telnov Yu.F., Yunatova I.G. (2011) Metodika postroeniya modeli kompetentsii na osnove professional'nykh standartov v oblasti IKT dlya sozdaniya programm dopolnitel'nogo professional'nogo obrazovaniya [The methodology of competence model design based on professional standards in ICT area for creation of the additional professional educational programs]. *Economics, Statistics and Informatics. Vestnik UMO*, no. 6, pp. 112–118 (in Russian).
30. Zinder E.Z., Yunatova I.G. (2012) Sistema metrik i otsenochnykh shkal znanii, umenii i sposobnostei v normalizovannoi modeli kompetentsii primenitel'no k sfere IKT [The system of metrices and evaluative scales of knowledge, skills and abilities in the normalized model of competences applied to the ICT area]. Proceedings of the *XV Conference on Business Processes Reengineering and Knowledge Management Systems (BPR&KMS–2012). Moscow, 23–25 April 2012.* Moscow: MESI, vol. 1, pp. 85–89 (in Russian).
31. Guzik S.V., Zinder E.Z., Yunatova I.G. (2013) Novaya paradigma inzhiniringa predpriyatiya i upravlenie sootvetstviem mezhdu rabochimi protsessami i kompetentnost'yu ikh ispolnitelei [A new paradigm of enterprise engineering and management of compliance between working processes and competency of their executers]. Proceedings of the *XVI Conference on Enterprise Engineering and Knowledge Management (EE&KM–2013). Moscow, 25–26 April 2013.* Moscow: MESI, pp. 90–100 (in Russian).
32. Zinder E.Z., Yunatova I.G. (2013) Conceptual framework, models, and methods of knowledge acquisition and management for competency management in various areas. Proceedings of the *4th International Conference on Knowledge Engineering and the Semantic Web (KESW 2013). St. Petersburg, Russia, 7–9 October 2013* (eds. P. Klinov, D. Mouromtsev). Berlin, Heidelberg: Springer, vol. 394, pp. 228–241.

Exact time-efficient combined algorithm for solving the asymmetric traveling salesman problem[1]

Galina N. Zhukova
Associate Professor, School of Software Engineering
National Research University Higher School of Economics
Address: 20, Myasnitskaya Street, Moscow, 101000, Russia
E-mail: galinanzhukova@gmail.com

Mikhail V. Ulyanov
Leading Researcher, Laboratory of Scheduling Theory and Discrete Optimization
V.A. Trapeznikov Institute of Control Sciences, Russian Academy of Sciences;
Professor, Department of Algorithmic Languages
Lomonosov Moscow State University
Address: 65, Profsoyuznaya Street, Moscow, 117997, Russia
E-mail: muljanov@mail.ru

Mikhail I. Fomichev
Graduate of System and Software Engineering MSs Program
National Research University Higher School of Economics
Address: 20, Myasnitskaya Street, Moscow, 101000, Russia
E-mail: mikhail.fomichev94@gmail.com

Abstract

For practical, important tasks in the fields of economics and logistics, as well as in a number of technical applications, it becomes necessary to solve the traveling salesman problem (TSP). Quite often, the features of these problems lead to the traveling salesman problem in asymmetric formulation (asymmetric traveling salesman problem, ATSP). Moreover, in some practical applications it is desirable to obtain an exact solution. One of the known exact algorithms for solving the ATSP is an algorithm that implements the well-known branch and bound method. The known experimental estimates of its complexity on the average are exponential. However, this does not mean that for small dimensions of the problem (currently, no more than 70–75), the expected time for solving the individual problem is unacceptably high. The need to reduce the time for solving individual problems dictated by practice is associated with the use of various modifications of this algorithm, of which a modification that involves storing truncated matrices in the search decision tree is one of the most effective. In this article, the authors rely on this modification. Other possible improvements in the time efficiency of the software implementation of the branch and bound method are related, among other things, to obtaining the initial approximation by heuristic algorithms. As a result, we get a combined algorithm, in which, at the first stage, some heuristics works to obtain the initial solution, from which the branch and bound method starts. This idea has been discussed for a long time, but the problem is that to reduce time, such a heuristic algorithm is needed that delivers a solution close to optimal which will be found quite fast. One of the possible solutions to this problem is the subject of this article.

[1] This work was supported by the Russian Foundation for Basic Research (project No. 16-07-00160 "Forecasting of time characteristics of efficient implementations of the branch and bound method for the traveling salesman problem relying on characteristics of random matrixes and identification of generated times distribution")

The subject of the research in this article is the choice of the best heuristic algorithm which, when applied, leads to an increase in temporal efficiency in combination with the algorithm of the branch and bound method, and an experimental study of its software implementation in order to obtain an average time for solving individual problems. On the basis of the results obtained, recommendations are given on the limiting dimensions of the problem that allow for an acceptable solution time, something which is of interest in the practical application of this combined algorithm in the tasks of business informatics and logistics.

Key words: travelling salesman problem; branch and bound method; combined algorithm; time efficiency; experimental research.

Citation: Zhukova G.N., Ulyanov M.V., Fomichev M.I. (2018) Exact time-efficient combined algorithm for solving the asymmetric traveling salesman problem. *Business Informatics*, no. 3 (45), pp. 20–28. DOI: 10.17323/1998-0663.2018.3.20.28.

Introduction

The traveling salesman problem (TSP) can be formulated as follows: the salesman is to find the cheapest tour between n cities, visiting each city once and only once (except origin) and returning to the city of origin. We call a tour the sequence of all cities which the salesman is to visit. Instead of 'the travel cost,' one can use distance, time or other indices. The cost of travel is known for all pairs of cities. In the mathematical field of graph theory, the TSP is defined in the following way: what is a minimal Hamiltonian cycle in a complete weighted graph. This graph is represented by an adjacency matrix $C = (c_{ij})$, which is called the cost matrix. The diagonal elements of the adjacency matrix are infinitely large numbers, because the graph of the road net does not contain self-loops. In the general case, the graph is directed. The Hamiltonian cycle of a graph we call a tour. The solution of the TSP is a tour with a minimum sum of arc weight. In such a formulation, the Traveling Salesman Problem belongs to the class of NP-hard problems. That is why it relates to the intractable problems of combinatorial optimization.

The TSP admits a variety of practical important interpretations. For example: scheduling the operation of equipment with reconfigurations, optimization of crane operations, sequencing of burning slots in the manufacture of chips [1]. The cabling of computer networks, predicting protein function and representation of black and white images by a continuous line without intersections can also be reduced to the TSP [2].

Presented in 1963 [3], the first exact algorithm for solving TSP was based on the branch and bound (B&B) method [4]. A detailed description of the earlier works of the Traveling Salesman Problem can be found in [5]. The present methods and approaches to the TSP are described in [6]. The issues of increasing the accuracy of the lower estimate of the cost of the tour can be found in article [7]. We do not give a description of the classical branch and bound algorithm for the TSP which can be found, for example, in [3; 8; 9].

Since a number of practical problems in the field of business informatics, logistics and economic optimization are reduced to the TSP, there are an abundance of heuristic methods for solving it. But this does not mean that there is no need for exact solutions to the problem. Consequently, one may ask:

✦ what is the biggest dimension value of the problem that can be solved exactly within an acceptable time?

✦ how can we increase this dimension value by the use of a modified exact algorithm with better time efficiency?

Since the 1960s, researchers have used a heuristic algorithm for finding optimal or near-optimal solutions (tours) for the individual TSP and then a start branch and bound algorithm with the initial solution. It is intuitively clear that such an approach should reduce the number of vertices of the generated decision tree and therefore the time of finding the exact solution. In this case, we get a combined algorithm which contains a heuristic algorithm and B&B. The present article is devoted to the description and the computational analysis of the combined algorithm with the Linn–Kernigan method as a heuristic part.

1. Definition of the task

We need to find a heuristic algorithm for the asymmetric traveling salesman problem that can help to reduce the total time of solving TSP by starting the branch and bound algorithm with the initial solution, hereinafter called the precomputed tour.

We introduce the following notations:

n – dimension of the problem (number of vertices of the complete graph);

$C = (c_{ij})$ – cost matrix of the individual asymmetric traveling salesman problem that is an adjacency matrix of a directed graph without loops;

T_e – cost of a precomputed tour which is found by a heuristic algorithm;

$t_{BB}(C, T_e)$ – the running time of the software implementation of the branch and bound algorithm (in a certain hardware configuration) for the individual TSP, which is defined by cost matrix C with a precomputed tour of cost T_e;

$t_E(C)$ – the running time of a heuristic algorithm;

$N_0(C)$ – the number of vertices of the search decision tree generated by the classical branch and bound algorithm (without a precomputed tour);

$N_1(C)$ – the number of vertices of the search decision tree generated by the branch and bound algorithm with the precomputed optimal tour;

$\overline{t}_{BB0}(n)$ – sample average running time for software implementation of the classical branch and bound algorithm without a precomputed tour when the sample consists of individual problems with the same dimension n;

$\overline{t}_{BB1}(n)$ – sample average running time for software implementation of the branch and bound algorithm with a precomputed tour when the sample consists of individual problems with the same dimension n;

$\overline{t}_E(n)$ – sample average running time for software implementation of the heuristic algorithm.

Our goal is to find such a heuristic algorithm that delivers the initial tour that decreases the running time of the B&B so that

$$t_{BB}(C, T_e) + t_E(C) < t_{BB}(C, \infty) \qquad (1)$$

for most cost matrices C.

Since the branch and bound algorithm is highly sensitive to the features of individual problems, condition (1) does not mean the time reduction for every individual task. Thereafter, we apply the sample average:

$$\overline{t}_{BB1}(n) + \overline{t}_E(n) < \overline{t}_{BB0}(n). \qquad (2)$$

It is also of interest to find the threshold value N of the dimension of the problem such that a combined algorithm is more effective than the classical B&B when the TSP dimension n is more than N. In order to predict the time efficiency, it is also necessary to approximate a dependence of the average running time on the dimension of the TSP.

2. How to reduce the running time of the branch and bound algorithm

In order to decrease the running time of algorithms that implement the idea of the branch and bound method for solving the traveling salesman problem, various approaches are proposed. Some of them use storing reduced matrices in the nodes of the search decision tree [10]. There are also several combinations of exact and heuristic algorithms [11–13].

The paper [10] presents an experimental study of the impact of additional memory allocation for the storage of truncated cost matrices in the nodes of the search decision tree in the range of TSP of dimension from 25 to 45. *Table 1* presents a forecast based on the experimental data [10]. This means that the software implementation of the algorithm with the storing matrices can be effectively used for finding the exact solution of the TSP of a dimension that is not more than 70 at an available RAM of 16 GB with the expected average calculation time of modern personal computers on the order of magnitude of one minute.

In developing a combination of the branch and bound algorithm and the heuristic algorithms that deliver a precomputed tour, we will further use an implementation that does not include the storage of truncated cost matrices at the vertices of the search decision tree.

The authors of the study [12] sought to level out the shortcomings of the branch and bound algorithm with heuristic algorithms and techniques for parallelizing program flows. However, they also pointed out that they clearly understood the fact that the results obtained are not reliable and had to be treated with great care.

3. The influence of the quality of a precomputed tour on the number of generated vertices

In the classical branch and bound algorithm for solving the TSP, there is no opportunity to begin cutting out the subtrees of the search decision tree until a tour is found. This is because the classical algorithm does not assume the presence of any tour at the time of the initial launch. In view of the particular nature of the problem, the decision tree can seriously grow until the first tour is found. Using a precomputed tour (which is known before the B&B algorithm is launched) can reduce the size of the search decision tree. Under the condition that a precomputed tour is close to the optimal one, a combined approach allows for less time to be spent on finding an optimal tour. That is because there is no need to create and visit unfavorable vertices of the search decision tree. The amount of additional memory required also decreases, since there is no need to store unfavorable vertices of the search tree.

Thus, the combined approach aims at reducing both the running time and the amount of memory required. However, the implementa-

Table 1.

Resource characteristics forecast

Dimension	Prediction of optimal tour calculation time without additional memory	Prediction of optimal tour calculation time with additional memory	Time ratio prediction	Prediction of the average amount of additional memory required
45	1 s	0.2 s	5	30.71 MB
54	7 s	1 s	7	172.3 MB
70	11.7 min	1 min	11.7	12.47 GB
80	2.5 h	10 min	15	136.37 GB
88	19.6 h	1 h	19.6	924.26 GB
102	29.5 days	1 day	29.5	25.69 TB

tion of the combined algorithm causes some problems. Obviously, the closer a precomputed tour is to an exact solution, the less time it takes to find the optimal solution by B&B. On the other hand, a precomputed tour must be found quickly. This means that the running time of the classical branch and bound algorithm should at least be no more than the total running time of the heuristic algorithm and the branch and bound algorithm with the precomputed tour. In other words, the use of a precomputed tour should be justified and rational.

According to [14], if a precomputed tour is optimal, then the search decision tree size is reduced by approximately 40%. The average number of generated vertices of the search decision tree for problems (of different dimensions), obtained experimentally in [14] is presented in *Table 2.* The characteristic $\eta = \left(1 - N_1 / N_0\right)$ shows the decrease of the size of the search decision tree by percentage if the precomputed tour for the branch and bound algorithm is optimal.

However, the sensitivity of the Branch and Bound algorithm to the quality of a precomputed tour is sufficiently high: if a precomputed tour is more than 5% larger than the optimal one, then, as was shown in [14], the search decision tree cannot be significantly reduced.

Table 2.

The influence of the precomputed tour on the number of generated vertices

n	η
35	38%
40	38%
45	39%

4. Heuristic algorithms for calculating a precomputed tour

Heuristic algorithms are algorithms which do not necessarily deliver an exact solution. However, the solutions which can be found by these algorithms are usually quite close to an optimal one; moreover, they are available in a "reasonable" time [15]. Unlike exact algorithms, heuristic algorithms are usually fairly simple to implement, and they work faster. All heuristic algorithms can be divided into the following three types:

- greedy [16];
- swarm intelligence [17–19];
- improvement [20, 21].

In addition, there are many different algorithms that are designed to solve particular cases of the TSP (for example, the metric traveling salesman problem [22]).

Preliminary analysis of the literature sources [14, 23, 24] allows us to conclude that representatives of greedy and swarm intelligence heuristic algorithms cannot provide a sufficiently high-quality solution for the time required in our formulation of the problem. Therefore, the use of the initial tours obtained by these algorithms is not reasonable. The situation is different with a lot of heuristic algorithms that improve solutions.

In 1973, S. Lin and B. Kernigan presented an effective heuristic algorithm for the traveling salesman problem (the Lin–Kernighan algorithm) [25]. It is based on the idea of an iterative improvement of a randomly found tour. As shown by experimental results, this algorithm often even finds globally optimal solutions. At the same time, the complexity of the algorithm is approximately $O\left(n^{2,2}\right)$ [25].

Later, in 2000, K. Helsgaun proposed a modified implementation of the algorithm (the Lin–Kernighan–Helsgaun algorithm) [26]. This algorithm often provides an optimal solution in an acceptable time, even for problems of large dimension.

The algorithms considered are designed to solve the symmetric traveling salesman problem. However, using the improved method described in paper [27], any asymmetric traveling salesman problem (dimension n) can be reduced to the symmetric traveling salesman problem (dimension $2n$). Unfortunately, this transformation also affects the time of solving the asymmetric traveling salesman problem by the Lin–Kernighan–Helsgaun algorithm.

The main idea of the Lin–Kernighan–Helsgaun algorithm is to transform a feasible solution by the replacement of some set of its arcs to another, which delivers a better feasible solution. The process goes on until there is an existing set to replace it with. All details of the algorithm are given in [26].

The efficiency of the Lin–Kernigan–Helsgaun algorithm is achieved, first of all, due to the effective strategy of the search sets of arcs described above. The search is based on the restriction of 5-opt replacements inside the set of possible candidates [26].

The author of [28] attempted to develop an algorithm that provides an optimal or very close to optimal solution relatively quickly. However, he did not pay attention to the complexity of the implementation, and as a result, the last software implementation presented by the author consisted of about 10,000 lines of source code [28].

5. Combined algorithm and experimental results

To analyze the combination of the branch and bound algorithm with a heuristic Lin–Kernigan–Helsgaun algorithm, an experimental study of asymmetric traveling salesman problems of dimension 35, 37, 40, 43, and 45 was carried out. The sample of TSP's of the same dimension consisted of 100,000 individual problems (for each dimension value).

The experiments were carried out on a personal computer with the following characteristics:

- processor: Intel i7 8700K 4700 MHz;
- RAM: Corsair Vengeance LPX CMK 32GX4M2B3466C16R DDR4 3466 MHz 32 GB;
- motherboard: ASRock Fatal1ty Z370 Gaming K6;
- operating system: Arch with kernel version 4.14.13-1-ARCH.

To minimize operating system noise, background processes (for example, firewalls) unnecessary for the research were disabled, and the distribution did not contain the components of the graphical interface (we used command line interface instead). Moreover, swapping was disabled, so that the speed of the HDD did not affect the running time of the implementation of the algorithm.

The algorithms was implemented in C++ and compiled into an executable program using the compiler gcc 7.2.1 20171224.

The combined algorithm was implemented in C++ and compiled into an executable program using the compiler gcc 7.2.1 20171224.

We introduce the notation $\overline{t}_{LKH}(n)$, which means the average running time of the software implementation of the Lin–Kernighan–Helsgaun algorithm where n denotes the dimension of TSP.

Average, minimal and maximal running time of the B&B algorithm implementation without a precomputed tour are presented in the first part of *Table 3* ($\overline{t}_{BB0}(n)$, $\breve{t}_{BB0}(n)$, $\hat{t}_{BB0}(n)$ correspondingly, the sample size is 100,000).

The same characteristics of the implementation of the combined algorithm with a precomputed tour calculated by the Lin–Kernighan–Helsgaun algorithm are denoted as $\overline{t}_{BB1}(n)+\overline{t}_{LKH}$, $\breve{t}_{BB1}(n)+\breve{t}_{LKH}$, $\hat{t}_{BB1}(n)+\hat{t}_{LKH}$. These figures are presented in the second part of *Table 3* (the sample size is 100,000).

The experimental results obtained do not show such a significant reduction of the running time of the implementation of B&B combined with Lin–Kernighan–Helsgaun algorithm as can be expected based on *Table 2*. The reduction of the number of generated vertices of the decision tree by 38% does not provide a similar decrease of the running time of the

Table 3.

Experimental results

n	$\overline{t}_{BB0}(n)$ μs	$\breve{t}_{BB0}(n)$ μs	$\hat{t}_{BB0}(n)$ μs	$\overline{t}_{BB1}(n)+\overline{t}_{LKH}$ μs	$\breve{t}_{BB1}(n)+\breve{t}_{LKH}$ μs	$\hat{t}_{BB1}(n)+\hat{t}_{LKH}$ μs
35	78 009	312	4 866 390	84 080	3 629	4 200 069
37	128 072	357	63 857 740	133 791	3 691	39 814 733
40	261 729	435	34 998 093	264 468	3 840	26 352 357
43	539 085	504	66 511 234	531 160	5 127	58 208 012
45	859 599	578	123 629 945	831 424	6 232	78 427 649

implementation of B&B. The reason is that calculation of the lower bounds of the vertices of the decision tree takes time, but when the estimates are greater than the value of the precomputed tour no new vertices are created. In this case, the running time is spent without the creation of a new vertex of the search decision tree.

Approximation of the obtained experimental results by the method of least squares can be represented as

$$\overline{t}_{BB0}(n) = 17.783 \cdot e^{0.2399n},\ R^2 = 0.9999, \quad (3)$$

$$\overline{t}_{BB1}(n) + \overline{t}_{LKH}(n) = 27.552 \cdot e^{0.2293n},\ R^2 = 0.9999. \quad (4)$$

The abscissa of the intersection point of these exponents is close to the integer point $n = 43$. Taking into account (3) and (4), we concluded that when the dimension of the TSP is more than $n = 43$, the combination of the branch and bound algorithm with the Lin–Kernigan–Helsgaun algorithm delivers an exact solution faster (on average) than the classical B&B. In other words, the sample average running time of B&B with the Lin–Kernigan–Helsgaun algorithm is less than the sample average running time of B&B without an initial tour. The formula (4) provides the values of the dimension of TSP which can be solved using a combined algorithm in time $t_{\max}$ or less. In this case, the largest dimension of the TSP is the solution of the equation

$$27.552 \cdot e^{0.2293n} = t_{\max} \quad (5)$$

For example, when $t_{\max} = 6 \cdot 10^8$ microseconds (i.e. 5 minutes), a combined algorithm can deliver an exact solution of TSP of dimension up to $n = 73$. However, we should not forget that the method is very sensitive to the features of the individual problems, and the amount of running time of some TSP of dimension 73 can be significantly larger than $t_{\max} = 6 \cdot 10^8$ microseconds. Moreover, the requirement for additional memory can be significant.

The authors are aware that the accuracy of the extrapolation obtained is not high, and we have a fairly rough approximation. In general, we hope that when the dimension of the TSP is more than 70–75, the combined algorithm can solve the TSP significantly faster (on average), but we do not state that it is true for all individual tasks.

Conclusion

Thus, on the basis of the experimental research we conducted, the following conclusions can be made:

✦ according to the calculated trend, when the dimension of the TSP is more than $n = 43$, the combined algorithm of B&B and the Lin–Kernigan–Helsgaun algorithm works faster than the classical branch and bound algorithm for the asymmetric traveling salesman problem (on average);

✦ the branch and bound algorithm with a precomputed tour does not provide a solution to the traveling salesman problem in polynomial time, however the use of an initial tour reduces the coefficient at n in the exponent in (3) by 4.4%, which is also a significant result;

✦ it is not efficient to use greedy or swarm intelligence heuristic algorithms for finding an initial tour, because the total time of the operation of the branch and bound algorithm and those heuristic algorithms is much more than the classical branch and bound algorithm (without a precomputed tour).

The authors see the further development of the study in a more detailed analysis of the distributions of running time of software implementations of the algorithms. ■

References

1. Borodin V.V., Loveckiy S.E., Melamed I.I., Plotinskiy U.M. (1980) Ehksperimentalnoe issledovanie ehffektivnosti ehvristicheskih algoritmov resheniya zadachi kommivoyazhera [Experimental research of the effectiveness of heuristic algorithms for solving the traveling salesman problem]. *Automation and Remote Control*, no. 11, pp. 76–84 (in Russian).
2. Kolesnikov A.V., Kirikov I.A., Listopad S.V., Rumovskaya S.B., Domanitsky A.A. (2011) *Reshenie slozhnyh zadach kommivoyazhera metodami funkcionalnyh gibridnyh intellektualnyh system* [Solution of complex traveling salesman problems using the methods of functional hybrid intelligent systems]. Moscow: Institute of Informatics Problems (in Russian).
3. Little J.D.C., Murty K.G., Sweeney D.W., Karel C. (1963) An algorithm for the traveling salesman problem. *Operations Research*, no. 11, pp. 972–989.
4. Land A.H., Doig A.G. (1960) An automatic method of solving discrete programming problems. *Econometrica*, vol. 28, no. 3, pp. 497–520.
5. Slominski L. (1982) Probabilistic analysis of combinatorial algorithms: A bibliography with selected annotations. *Computing*, no. 28, pp. 257–267.
6. Matai R., Mittal M.L., Singh S. (2010) Travelling salesman problem: An overview of applications, formulations and solution approaches. *Travelling salesman problem: Theory and applications* (ed. D. Davendra). Rijeka: Intech Open Access Publisher, pp. 1–24.
7. Toriello A. (2014) Optimal toll design: A lower bound framework for the asymmetric traveling salesman problem. *Mathematical Programming*, vol. 144, no. 1–2, pp. 247–264.
8. Goodman S.E., Hedetniemi S.T. (1981) *Vvedenie v razrabotku i analiz algoritmov* [Introduction to the design and analysis of algorithms]. Moscow: Mir (in Russian).
9. Korte B., Vygen J. (2007) *Combinatorial optimization: Theory and algorithms*. Springer.
10. Ulyanov M.V., Fomichev M.I. (2015) Resource characteristics of ways to organize a decision tree in the branch-and-bound method for the traveling salesmen problem. *Business Informatics*, no. 4 (34), pp. 38-46.
11. Oliver I., Smith D., Holland J. (1987) A study of permutation crossover operators on the traveling salesman problem. Proceedings of the *2nd International Conference on Genetic Algorithms. Cambridge, MA, USA, 28–31 July 1987*, pp. 224–230.
12. Cotta C., Aldana J., Nebro A., Troya J. (1995) Hybridizing genetic algorithms with branch and bound techniques for the resolution of the TSP. *Artificial neural nets and genetic algorithms 2* (eds. D. Pearson, N. Steele, R. Albrecht). Wien, New York: Springer-Verlag, pp. 277–280.
13. Goldberg D., Lingle R.J. (1985) Alleles, loci, and the travelling salesman problem. Proceedings of the *First International Conference on Genetic Algorithms and Their Applications. Pittsburgh, 24–26 July 1985*,

pp. 154–159.

14. Fomichev M.I. (2017) Sravnitelnyi analiz metaehvristicheskih algoritmov resheniya nesimmetrichnoj zadachi kommivoyazhera [Comparative analyses of metaheuristic algorithms of solving asymmetric traveling salesman problem]. *Control Systems and Information Technologies*, no. 3 (69), pp. 88–92 (in Russian).
15. Panteleev A.V. (2009) *Metaevristicheskie algorithmi poiska globalnogo ekstremuma* [Metaheuristic algorithms for search global extremum]. Moscow: MAI-Print (in Russian).
16. Applegate D.L., Bixby R.E., Chvatal V., Cook W.J. (2006) *The traveling salesman problem: A computational study.* Princeton, NJ: Princeton University Press.
17. Colorni A., Dorigo M., Maniezzo V. (1991) Distributed optimization by ant colonies. Proceedings of the *First European Conference on Artificial Life (ECAL 91). Paris, France, 11–13 December 1991*, pp. 134–142.
18. Dorigo M., Gambardella L.M. (1997) Ant colony system: A cooperative learning approach to the traveling salesman problem. *IEEE Transactions of Evolutionary Computation*, vol. 1, no. 1, pp. 53–66.
19. Bonavear E., Dorigo M. (1999) *Swarm intelligence: From natural to artificial systems.* Oxford, UK: Oxford University Press.
20. Gamboa D., Rego C., Glover F. (2005) Data structures and ejection chains for solving large scale traveling salesman problems. *European Journal of Operational Research*, vol. 160, no. 1, pp. 154–171.
21. Kaplan H., Lewenstein M., Shafrir N., Sviridenko M. (2005) Approximation algorithms for asymmetric TSP by decomposing directed regular multigraphs. *Journal of the ACM*, vol. 52, no. 4, pp. 602–626.
22. Mömke T., Svensson O. (2016) Removing and adding edges for the traveling salesman problem. *Journal of the ACM*, vol. 63, no. 1, article no. 2.
23. Cormen T.H., Leiserson C.E., Rivest R.L., Stein C. (2009) *Introduction to algorithms.* Cambridge, MA: MIT Press.
24. Stutzle T., Hoos H.H. (2000) MAX-MIN ant system. *Future Generation Computer Systems*, no. 16. pp. 889–914.
25. Lin S., Kernighan B.W. (1973) An effective heuristic algorithm for the traveling-salesman problem. *Operations Research*, vol. 21, no. 2, pp. 498–516.
26. Helsgaun K. (2000) An effective implementation of the Lin–Kernighan traveling salesman heuristic. *European Journal of Operational Research*, vol. 126, no. 1, pp. 106–130.
27. Jonker R., Volgenant T. (1983) Transforming asymmetric into symmetric traveling salesman problems. *Operations Research Letters*, no. 2, pp. 161–163.
28. Helsgaun K. (2017) *An extension of the Lin–Kernighan–Helsgaun TSP solver for constrained traveling salesman and vehicle routing problems. Technical Report.* Roskilde: Roskilde University.

A heuristic algorithm for generating the numerical terms of a linguistic variable

Elena N. Chujkova
Associate Professor, Department of Computer Systems and Information Security
Don State Technical University
Address: 1, Gagarin Square, Rostov-on-Don, 344000, Russia
E-mail: elenchu@mail.ru

Vasilij V. Galushka
Associate Professor, Department of Computer Systems and Information Security
Don State Technical University
Address: 1, Gagarin Square, Rostov-on-Don, 344000, Russia
E-mail: galushkavv@yandex.ru

Abstract

In this paper we describe an easy-to-implement algorithm for automated generation of the linguistic variable term membership functions to allow for information search in a relational database based on qualitative criteria by means of the SQL query language.

The proposed algorithm makes it possible to calculate the parameters of the triangular and trapezoid membership functions taking into account the distribution of the variable of interest stored in the database. The algorithm defines the intervals covered by the term bases, so that each interval contains about the same number of values. Upper bounds of the defined intervals are used to calculate the parameters of membership functions. The parameters of the membership functions generated with this algorithm can be easily calculated with the limited computational means of the SQL language.

We review the algorithm realizations for the generation of 3 and 5 terms of a linguistic variable based on a sample from a database containing 100 or 500 different values.

The membership functions obtained through the algorithm have the required properties of orderliness, completeness, consistency and normality. They do not require further approximation. Unlike the known methods, the algorithm does not require significant computing resources, the use of specialized software, settings configuring, or a training set formation.

The algorithm implementation creates opportunities to support fuzzy search queries in relational databases using the means of the SQL language, as limited as they are. Thus, the system's level of intelligence would be increased, and the user would be provided with the means of search query formulation in a natural language. The linguistic variable terms generated using our algorithm can be used within the framework of a fuzzy rule-based knowledge base of an information system, as well as to perform fuzzy inference.

Key words: relational database; SQL language; fuzzy logic; linguistic variable; fuzzy set; membership function.

Citation: Chujkova E.N., Galushka V.V. (2018) A heuristic algorithm for generating the numerical terms of a linguistic variable. *Business Informatics*, no. 3 (45), pp. 29–38.
DOI: 10.17323/1998-0663.2018.3.29.38.

Introduction

At the core of any information system lies a database that stores the information processed by the system. One of the functions of the system is searching the database for any requested data. Currently, relational databases that support the search for information using SQL language are the most widely used. Composing a search request in SQL requires the user to set definite ranges for the values of the data being requested, something which is often impossible due the lack of such information. Moreover, the user, especially a novice, is better accustomed to using qualitative search criteria that could be expressed verbally. This kind of user-system interaction is also more convenient for the user.

In an implementation of information search in a database based on qualitative criteria, the tools of fuzzy logic can be used. They involve the use of linguistic variables and the generation of corresponding base term sets (fuzzy sets) that provide a quantitative interpretation for the qualitative search criteria [1–5].

Performing a fuzzy search query in a relational database requires converting it first into a command of the standard SQL language. Paper [6] describes such a conversion procedure that is implemented with the help of the graphical interface of the information system. Within this procedure, the values of the membership functions of the fuzzy sets used are calculated on-the-fly in the course of performing an SQL query. Linguistic variable terms are represented by parametric fuzzy numbers with trapezoid and triangular membership functions. Their parameters can be easily calculated using the commands of the SQL language, which is not intended for complex calculations.

This paper [6] uses the assumption about a uniform distribution of the data sampled from the database to find the parameters of the terms. This assumption limits the applicability of the proposed method and might lead to unacceptable results when it is not valid. In particular, if the distribution of the experimental data is nonuniform, it might happen that no values at all from the database would be in the base of some terms. To avoid such issues, a calculation of the terms' parameters must take the experimental data distribution into account. The term bases must contain approximately equal numbers of values available in the database.

Traditionally, the membership functions of the linguistic variable terms are generated based on expert information, which makes the procedure difficult to automate. For this reason, many researchers have made considerable effort towards the development of algorithmic methods of membership function generation. The papers [7; 8] provide an overview of the existing methods of automatic generation of membership functions. Among them are the methods of inductive logical inference [9], fuzzy c-means clustering methods [10–13], neural networks [14; 15], histograms [16], methods based on fuzzy entropy and other special measures [17; 18], genetic algorithms [19–21], ant colony system algorithms [22], heuristic algorithms [23], and particle swarm optimization [24]. The characteristics of the aforementioned methods are presented in *Table 1* [8].

All the methods listed above have been developed for particular applications of classification and are used to assign experimental data to fuzzy sets. However, they are not guaranteed to produce uniform distribution with an approximately equal number of elements in each set.

Paper [25] proposed a method of partitioning a source sample of values into intervals that contain equal numbers of values. However, using this method one might see identical values being assigned to different intervals, which defeats the purpose of classification.

The membership functions generated using the above methods do not always meet the necessary conditions of being ordered, complete, consistent, or normalized [2; 4], and still

Table 1.

Characteristics of the methods of automatic generation of membership functions

Methods	Characteristics			
	Source of information	Area of applicability	Shape of the membership functions	Number of fuzzy sets
Neural networks	dataset	classification	arbitrary	fixed
Histograms	dataset	image recognition, classification	Gaussian	arbitrary
Fuzzy *c*–means clustering	dataset	classification	triangular, trapezoid	fixed
Genetic algorithms	dataset and expert estimates	controllers	triangular, trapezoid, Gaussian	fixed
Ant colony system algorithms	dataset	data analysis, controllers	arbitrary	arbitrary
Particle swarm methods	dataset and expert estimates	controllers, image processing	Gaussian, triangular S–shaped	fixed
Other methods	dataset	classification, controllers	triangular, Gaussian	2–9

require further approximation. For example, the Gaussian membership functions generated by the particle swarm optimization are subnormal [8].

The existing algorithmic methods of generating membership functions typically have a high computational complexity. Using neural networks makes it necessary to provide a training set. The genetic algorithms, ant colony system algorithms, and particle swarm optimization methods require an objective function to be specified. They are also notorious for their large time to convergence and the possibility of their convergence to a local optimum. The implementation of most methods listed above requires some specialized software.

In this paper, we propose a simple-to-implement method for automatic generation of the membership functions for linguistic variable terms. The method takes into account the distribution of values of the parameter at hand stored in a database, which makes it possible to assign an approximately equal number of values to each term base. In addition, this method can be used to generate an arbitrary number of membership functions that possess the required properties.

1. Problem formulation

Tables of a relational database store numerical values of objects' characteristics used by an information system that must be evaluated and selected based on some qualitative fuzzy criteria (for example, "low," "medium," "high"). To perform a fuzzy search query, one has to provide linguistic variables corresponding to the properties being evaluated. To enable the automatic generation of a linguistic variable, we need an algorithm that would allow us to calculate the number and parameters of the membership functions of the linguistic variable terms. The choice of the number of values (terms) of the linguistic variable is up to the user. The basic scale of a linguistic variable is specified based on a set of all values of the property under investigation obtained from the database and is given by the interval $U=[u_{\min}, u_{\max}]$, where $u_{\min}$ is the minimum value of that property and $u_{\max}$ is the maximum value of that property. Each term base must contain an approximately equal number of values of the property being evaluated that are available in the database. The set of values and their frequencies can be obtained from the database using an SQL query as a sorted (ascending) matrix $H_{k\times 2}$, where the number of rows k is the number of different measurements (values of the parameter at hand in the database), the matrix elements h_{i1} are the values of the i-th measurement, and h_{i2} are the frequencies of the i-th measurement, $i = 1, ..., k$. The membership functions we are seeking to generate should be defined in such a way that their values could be easily calculated by means of the SQL language, which falls short of the computational efficiency of a universal programming language. The membership functions must meet certain requirements [2; 4].

We will represent the terms of a linguistic variable by parametric fuzzy numbers with the most common trapezoid and triangular membership functions. We will use the trapezoid membership functions for the minimum and maximum terms, and the triangular membership functions for the rest of the terms. The above choice of the membership function types was influenced by our goal to ensure the possibility of calculation on-the-fly while running an SQL query, considering the limited capabilities of SQL language. The calculation of the values of the membership functions is performed according to the following formula:

$$\mu(x)=\begin{cases}0,(x<a)\vee(x>d)\\1,b\le x\le c\\\dfrac{x-a}{b-a},(a\le x<b)\wedge(a<b)\\\dfrac{d-x}{d-c},(c<x\le d)\wedge(c<d)\end{cases},\qquad(1)$$

where $a\le b\le c\le d$ are the parameters of the membership function $\mu(x)$.

Let us split the base of the term into two intervals. The total number of such intervals is $m = l + 1$, where l is the number of terms. *Figure 1* shows the shape of the membership functions of linguistic variable terms. Their parameters and the upper bounds of the values of the intervals are provided. The following notation is used:

T_i – the i-th term;

a_i, b_i, c_i, d_i – the parameters of the i-th term, $i = 1, ..., l$;

p_j – the upper bound of the values of the j-th interval, $j = 1, ..., m$.

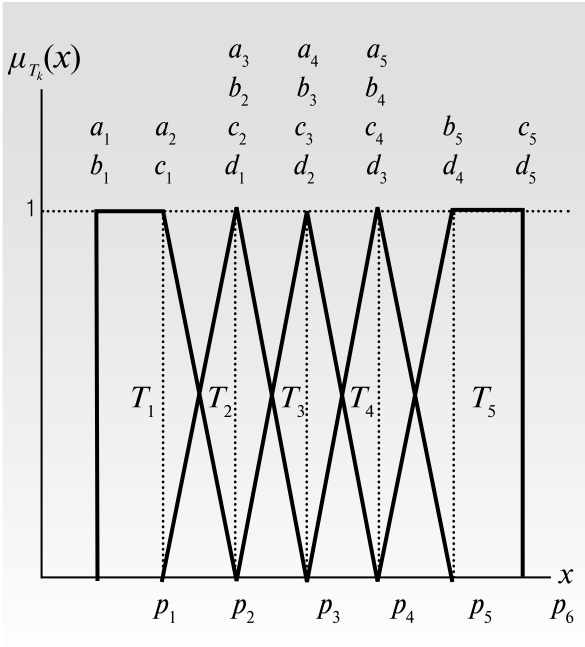

Fig. 1. Membership functions of linguistic variable terms

We need to define the intervals covered by the term bases so that each interval would contain an approximately equal number of values of the property at hand, and moreover, so that identical values would lie within the same interval. The upper bounds p_j of the intervals being defined are used to calculate the parameters a_j, b_j, c_j, d_j of the terms membership functions and are given by the quantiles of the uniform discretization of the source sample.

Based on the aforementioned matrix H, the number n of values that must lie within each interval, assuming a uniform distribution, is given by the equation

$$n = \sum_{i=1}^{k} h_{i2} / m,$$

where $\sum_{i=1}^{k} h_{i2}$ is the total number of measurements (the sample size).

Let us define the vector $\vec{z}$, its components being the numbers of measurements within each interval. Then the j-th component of the vector $\vec{z}$ is the number z_j of measurements within the j-th interval ($j = 1, ..., m$):

$$z_j = \sum_{i \in G} h_{i2},$$

where G is the set of numbers i of the rows of matrix H that contain the values of the frequencies h_{i2} included in z_j.

Let us denote by i^* the largest element of G: $i^* = \max_{i \in G} i$. Then, to calculate the quantiles p_j ($j = 1, ..., l$), we only need to choose p_j to be such elements $h_{i^*1}(i^* = 1,...,k-1)$ of matrix H that the function $E(\vec{z}) = \max_{j=1,...,l+1} |z_j - n|$ is minimized. Alternatively, another equivalent criterion of optimality could be used:

$$E(\vec{z}) = \sum_{j=1}^{l+1} |z_j - n| \to \min.$$

The value p_{l+1} coincides with the largest value in the sample: $p_{l+1} = h_{k1}$.

The optimization is performed under the following conditions:

$$\sum_{j=1}^{l+1} z_j = \sum_{i=1}^{k} h_{i2};$$

for $\forall j$ ($j = 1, ..., l+1$) $z_j = \sum_{i \in G} h_{i2}$, $i_{g+1} = i_g + 1$ $(g = i_{\min},...,i_{\max} - 1)$, where $i_{\min} = \min_{i \in G} i$, $i_{\max} = \max_{i \in G} i$;

for $\forall j_1, j_2, j_1 \neq j_2$, for which $z_{j1} = \sum_{i \in G_1} h_{i2}$, $z_{j2} = \sum_{i \in G_2} h_{i2}$, $G_1 \cap G_2 = \varnothing$ are defined.

To obtain non-trivial and nondegenerate solutions, the number l of the terms must obey the following condition: $2 \le l \le k$. The number of measurements contained in each of m intervals cannot be smaller than the maximum frequency of the measurements, therefore $m \le \sum_{i=1}^{k} h_{i2} / \max_{i=1,...,k} h_{i2}$. This means that the number l of the terms must meet the following condition:

$$2 \le l \le \sum_{i=1}^{k} h_{i2} / \max_{i=1,...,k} h_{i2} - 1 \qquad (2)$$

The parameters a_j, b_j, c_j, d_j of the terms' membership functions are found as follows:

$$a_j = \begin{cases} h_{11}, j = 1 \\ p_{j-1}, 1 < j \le l \end{cases}; \quad b_j = \begin{cases} h_{11}, j = 1 \\ p_j, 1 < j \le l \end{cases};$$

$$c_j = \begin{cases} p_j, 1 \le j < l \\ p_{j+1}, j = l \end{cases}; \quad d_j = p_{j+1}, 1 \le j \le l.$$

2. Algorithm for generation of membership functions

Let us find the actual number z_j of the measurements that get assigned to the j-th interval ($j = 1, ..., m$) and the parameters a_j, b_j, c_j, d_j of the linguistic variable terms according to the following algorithm:

Step 0. Verify that the conditions (2) of the existence of non-trivial and nondegenerate solutions are satisfied. If the conditions (2) are satisfied, then go to Step 1.

Step 1. Initialize i ($i := 0$).

1.1. For each j ($j = 1, ..., m - 1$), find $z1_j$ as follows

1.2.1. Initialize $z1_j$ ($z1_j := 0$).

1.2.2. While $z1_j < n$, consequentially accumulate the values h_{i2} in $z1_j$:

$$i := i + 1$$
$$z1_j := z1_j + h_{i2}.$$

1.2.3. If $z1_j \ge n$, then find R and Q:

$$R := n - (z1_j - h_{i2})$$
$$Q := z1_j - n.$$

Here R is the difference between the required number n of measurements within the interval and the actual number $z1_j$ of measurements within the j-th interval, not counting the i-th measurement; Q is the difference between the actual number $z1_j$ of measurements within the j-th interval (counting the i-th measurement) and the required number n of measurements. The value of R shows how much the number of measurements within the j-th interval is smaller than the required number (the shortage), while the value of Q shows how much the number of measurements within the j-th interval is greater than the required number n (the surplus).

1.2.4. If the shortage without the account for the last measurement is smaller than the surplus ($R < Q$), then remove from $z1_j$ the last measurement added to it h_{i2}:

$$z1_j := z1_j - h_{i2}$$
$$i := i - 1.$$

1.2.5. Define the upper bound $p1_j$ of the j-th interval as follows: $p1_j := h_{i1}$.

1.3. Find the number $z1_m$ of measurements that are assigned to the last m-th interval:

1.3.1. Initialize $z1_m$ ($z1_m := 0$).

1.3.2. Add to $z1_m$ all the remaining measurement: for all i ($i = i + 1, ..., k$)

$$z1_m := z1_m + h_{i2}.$$

1.3.3. Define the upper bound $p1_m$ of the m-th interval as follows: $p1_m := h_{k1}$.

Step 2. Find the total deflection δ_1 of the number of measurements from the required n:

$$\delta_1 = \sum_{j=1}^{m} | z1_j - n |.$$

Step 3. Repeat the calculations z_j by processing the rows of matrix H in reverse order beginning with the last row:

3.1. Initialize i ($i := k$).

3.2. Define the upper bound $p2_m$ of the m-th interval as follows: $p2_m := h_{k1}$.

3.3. For each j ($j = m, ..., 2$), find $z2_j$ as follows:

3.3.1. Initialize $z2_j$ ($z2_j := 0$).

3.3.2. While $z2_j < n$, consequentially accumulate the values h_{i2} in $z2_j$:

$$i := i - 1.$$
$$z2_j := z2_j + h_{i2}.$$

3.3.3. If $z2_j \geq n$, then find R and Q:

$$R := n - (z2_j - h_{i2})$$
$$Q := z2_j - n.$$

3.3.4. If $R < Q$, then remove from $z2_j$ the value h_{i2}:

$$z2_j := z2_j - h_{i2}$$
$$i := i + 1.$$

3.3.5. Define the upper bound $p2_{j-1}$ of the previous ($j - 1$)-th interval as follows:

$$p2_{j-1} := h_{(i-1)1} - h_{i2}.$$

3.4. Find the number $z2_l$ of measurements assigned to the first interval:

3.4.1. Initialize $z2_j$ ($z2_j := 0$).

3.4.2. Add to $z2_l$ all the remaining measurement: for all i ($i = i - 1, ..., l$)

$$z2_l := z2_l + h_{i2}.$$

Step 4. Find the total deflection δ_2 of the number of measurements from the required n:

$$\delta_2 = \sum_{j=1}^{m} | z2_j - n |.$$

Step 5. Find the values of the parameters a_j, b_j, c_j, d_j of the linguistic variable terms using the elements of vector p corresponding to the distribution z of measurements for which the least total deflection δ is achieved (if $\delta_1 < \delta_2$, then use vector $p1$ corresponding to the distribution $z1$ of measurements, otherwise use $p2$ corresponding to $z2$) as follows:

$$a_1 := h_{11}$$
$$b_1 := h_{11}.$$

Let $\delta_1 < \delta_2$.

Then for each j ($j = 1, ..., m - 1$) do:

if $j =< m - 1$, then $c_j := p1_j$, $a_{j+1} := p1_j$;

if $j > 1$, then $d_{j-1} := p1_j$, $b_j := p1_j$;

$$c_{m-1} := p1_m;$$
$$d_{m-1} := p1_m.$$

3. Analysis of the results produced by the algorithm

We consider the realizations of the above algorithm used to generate 3 and 5 terms of a linguistic variable based on samples from a database that contained 100 or 500 different values.

For a sample of 100 values and 3 terms, the number of intervals $m = 4$; the number of values per interval $n = 100/4 = 25$; the number of distinct measurements (the number of rows of matrix H of the values and their frequencies) $k = 20$. The source data (the elements of matrix H) along with the results of Steps 1-4 of the algorithm are presented in *Table 2.*

Table 2 demonstrates that the total deflection of the number of measurements from the required n for the downward processing of matrix H is greater than that for the reverse processing ($\delta_1 > \delta_2$), therefore the parameters of the membership functions are defined using vector $p2$. The values of the parameters of membership functions calculated at Step 5 of the algorithm are presented in *Table 3.*

Table 2.

Source data and the results of steps 1-4 of the algorithm

<table>
<tr><th rowspan="2">Row number of matrix $H(i)$</th><th colspan="2">Elements of matrix H</th><th rowspan="2">Interval number (j)</th><th rowspan="2">Number of measurements in the j-th interval $(z1_j)$</th><th rowspan="2">Upper bound of the j-th interval $(p1_j)$</th><th rowspan="2">Variance (δ_1)</th><th rowspan="2">Number of measurements in the j-th interval $(z2_j)$</th><th rowspan="2">Upper bound of the j-th interval $(p2_j)$</th><th rowspan="2">Variance (δ_2)</th></tr>
<tr><th>Value of the i-th measurement (h_{i1})</th><th>Frequency of the i-th measurement (h_{i2})</th></tr>
<tr><td>1</td><td>354</td><td>5</td><td rowspan="6">1</td><td rowspan="6">28</td><td rowspan="6">718</td><td rowspan="20">12</td><td rowspan="5">21</td><td rowspan="5">718</td><td rowspan="20">10</td></tr>
<tr><td>2</td><td>616</td><td>5</td></tr>
<tr><td>3</td><td>662</td><td>5</td></tr>
<tr><td>4</td><td>695</td><td>4</td></tr>
<tr><td>5</td><td>699</td><td>2</td></tr>
<tr><td>6</td><td>718</td><td>7</td><td rowspan="6">27</td><td rowspan="6">770</td></tr>
<tr><td>7</td><td>741</td><td>2</td><td rowspan="6">2</td><td rowspan="6">26</td><td rowspan="6">784</td></tr>
<tr><td>8</td><td>745</td><td>5</td></tr>
<tr><td>9</td><td>758</td><td>2</td></tr>
<tr><td>10</td><td>764</td><td>2</td></tr>
<tr><td>11</td><td>770</td><td>9</td></tr>
<tr><td>12</td><td>784</td><td>6</td><td rowspan="4">24</td><td rowspan="4">791</td></tr>
<tr><td>13</td><td>785</td><td>2</td><td rowspan="4">3</td><td rowspan="4">27</td><td rowspan="4">800</td></tr>
<tr><td>14</td><td>790</td><td>14</td></tr>
<tr><td>15</td><td>791</td><td>2</td></tr>
<tr><td>16</td><td>800</td><td>9</td><td rowspan="5">28</td><td rowspan="5">855</td></tr>
<tr><td>17</td><td>810</td><td>11</td><td rowspan="4">4</td><td rowspan="4">19</td><td rowspan="4">855</td></tr>
<tr><td>18</td><td>813</td><td>5</td></tr>
<tr><td>19</td><td>814</td><td>2</td></tr>
<tr><td>20</td><td>855</td><td>1</td></tr>
</table>

Table 3.

Parameters of membership functions calculated by the algorithm

Membership function number (j)	Parameters of the j-th membership function			
	a_j	b_j	c_j	d_j
1	354	354	718	770
2	718	770	770	791
3	770	791	855	855

Plots of the membership functions for 3 and 5 terms obtained through the algorithm are shown in *Figures 2* and *3*, respectively. The experimental data sampled from the database are presented by histograms in the figures.

The shape of the membership functions calculated without the account for the distribution of experimental data for the same data samples are shown in *Figure 4.*

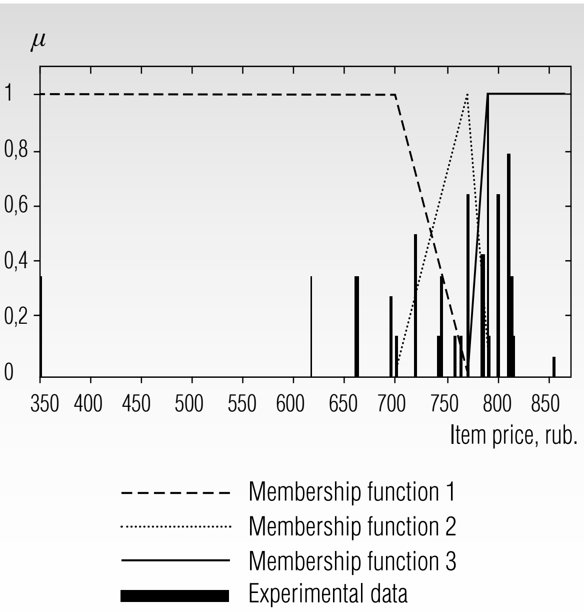

Fig. 2. Membership functions for 3 terms of the linguistic variable "Item price" for a sample of 100 price values

As seen from *Figures 2* and *3*, the proposed algorithm generates the membership functions for the linguistic variable terms that meet all the essential requirements [2; 4]. In particular, it is ensured that each term base contains an approximately equal number of values. By contrast, the membership functions generated without the account for the experimental data distribution (see *Figure 4*) do not possess the required properties and require further approximation.

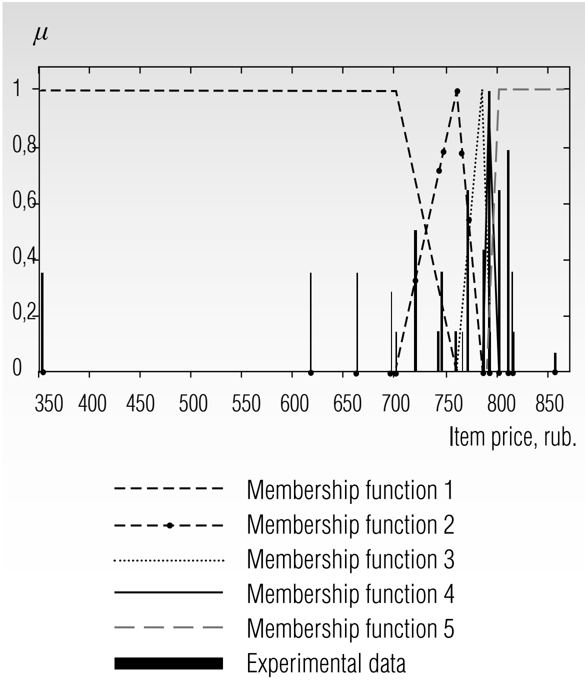

Fig. 3. Membership functions for 5 terms of the linguistic variable "Item price" for a sample of 100 price values

Conclusion

The proposed algorithm makes it possible to automate the procedure of generating the membership functions for the linguistic variable terms. The parameters of such membership functions can be stored in a database and later used to submit fuzzy queries in SQL language. In contrast to the existing methods of automatic generation of the membership functions, our algorithm is simple to implement,

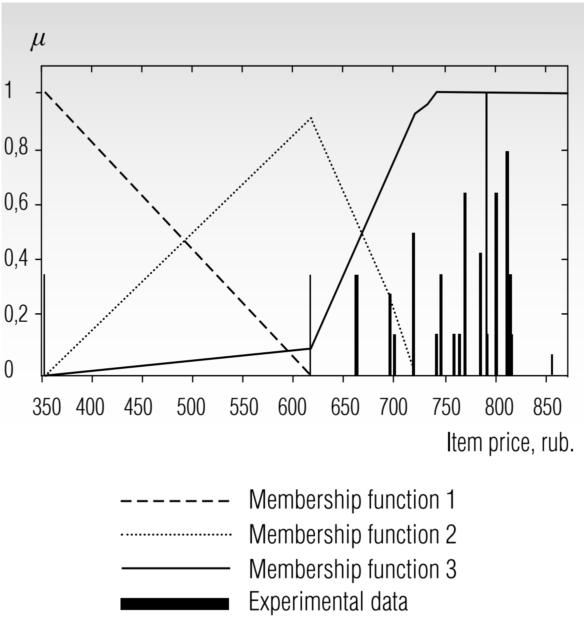

Fig. 4. Membership functions for 3 terms of the linguistic variable "Item price" generated without the account for the experimental data distribution for a sample of 100 price values

does not require considerable computational resources or specialized software, is free of adjustable parameters, does not require a training set, and enables the generation of any user-specified number of triangular and trapezoid membership functions. The membership functions generated with the help of this algorithm take into account the experimental data distribution. Additionally, the respective term bases of the linguistic variable contain approximately equal numbers of values selected from the database. The membership functions obtained through this algorithm possess all the required properties [2; 4] and do not necessitate further approximation. At the same time, the algorithm was developed for the solution of the particular problem stated above, and cannot be claimed to be universal.

The algorithm implementation creates opportunities to support fuzzy search queries in relational databases using the means of SQL language, as limited as they are. Thus, the system's level of intelligence would be increased, and the user would be provided with the means of search query formulation in a natural language. The membership functions of the linguistic variable terms generated using our algorithm can be used within the framework of a fuzzy rule-based knowledge base of an information system, as well as to perform fuzzy inference. ■

References

1. Shtovba S. (2001) *Vvedenie v teoriyu nechetkih mnozhestv i nechetkuyu logiku* [Introduction to the theory of fuzzy sets and fuzzy logic]. Available at: http://matlab.exponenta.ru/fuzzylogic/book1 (accessed 14 October 2013) (in Russian).
2. Korneev V.V., Gareev A.F., Vasjutin S.V., Rajh V.V. (2000) *Bazy dannyh. Intellektual'naya obrabotka informatsii* [Databases. Intelligent processing of information]. Moscow: Knowledge (in Russian).
3. Zadeh L. (1965) Fuzzy sets. *Information and Control*, vol. 8, no. 3, pp. 338–353.
4. Zadeh L. (1975) The concept of a linguistic variable and its application to approximate reasoning – II. *Information Sciences*, vol. 8, no. 4, pp. 301–357.
5. Andrejchikov A.V. (2006) *Intellektual'nye informatsionnye sistemy* [Intelligent information systems]. Moscow: Finance and Statistics (in Russian).
6. Chujkova E.N. (2014) Realizatsiya nechetkogo vybora oborudovaniya v sisteme proektirovaniya informatsionnoy seti [Realization of the fuzzy selection of equipment in the information network design system]. *Vestnik of Don State Technical University*, no. 3, pp. 164–171 (in Russian).
7. Medasani S., Kim J., Krishnapuram R. (1998) An overview of membership function generation techniques for pattern recognition. *International Journal of Approximate Reasoning*, no. 19, pp. 391-417.
8. Schwaab A., Nassar S., Filho P. (2015) Automatic methods for generation of type-1 and interval type-2 fuzzy membership functions. *Journal of Computer Sciences*, no. 11, pp. 976–987.

9. Kim C., Russell B. (1993) Automatic generation of membership function and fuzzy rule using inductive reasoning. Proceedings of the *3rd International Conference on Industrial Fuzzy Control and Intelligent Systems (IFIS '93), Houston, Texas, 1–3 December 1993,* pp. 93–96.
10. Chen M., Wang S. (1999) Fuzzy clustering analysis for optimizing membership functions. *Fuzzy Sets and Systems*, no. 103, pp. 239–254.
11. Liao T., Celmins A., Hammell R. (2001) A fuzzy c-means variant for the generation of fuzzy term sets. *Fuzzy Sets and Systems*, vol. 135, no. 2, pp. 241–257.
12. Lopes P., Camargo H. (2012) Automatic labeling by means of semi-supervised fuzzy clustering as a boosting mechanism in the generation of fuzzy rules. Proceedings of the *13th IEEE International Conference on Information Reuse and Integration (IRI). Las Vegas, USA, 8–10 August 2012*, pp. 279–286.
13. Jamsandekar S., Mudholkar R. (2014) Fuzzy classification system by self generated membership function using clustering technique. *BVICAM's International Journal of Information Technology,* vol. 6, no. 1,
14. Wu S., Er M., Gao Y. (2001) A fast approach for automatic generation of fuzzy rules by generalized dynamic fuzzy neural networks. *IEEE Transactions on Fuzzy Systems*, vol. 9, no. 4, pp. 578–594.
15. Castellano G., Castiello C., Fanelli A., Mencar C. (2005) Knowledge discovery by a neuro-fuzzy modeling framework. *Fuzzy Sets and Systems*, no. 149, pp. 187–207.
16. Refaey M. (2016) Automatic generation of membership functions and rules in a fuzzy logic system. Proceedings of the *5th International Conference on Informatics and Applications (ICIA 2016). Takamatsu, Japan, 14–16 November 2016*, pp. 117–122.
17. Cheng H., Chen J. (1997) Automatically determine the membership function based on the maximum entropy principle. *Information Sciences, v*ol. 96, no. 3–4, pp. 163–182.
18. Nieradka G., Butkiewicz B. (2007) A method for automatic membership function estimation based on fuzzy measures. Proceedings of the *International Fuzzy Systems Association World Congress (IFSA). Cancun, Mexico, 18–21 June 2007*, pp. 451–460.
19. Homaifar A., McCormick E. (1995) Simultaneous design of membership functions and rule sets for fuzzy controllers using genetic algorithms. *IEEE Transactions on Fuzzy Systems*, vol. 3, no. 2, pp. 129–139.
20. Shimojima K., Fukuda T., Hasegawa Y. (1995) Self-tuning fuzzy modeling with adaptive membership function, rules, and hierarchical structure based on genetic algorithm. *Fuzzy Sets and Systems*, vol. 71, no. 3, pp. 295–309.
21. Kaya M., Alhajj R. (2004) Integrating multi-objective genetic algorithms into clustering for fuzzy association rules mining. Proceedings of the *4th IEEE International Conference on Data Mining (ICDM '04). Brighton, UK, 1–4 November 2004*, pp. 431–434.
22. Hong T., Tung Y., Wang S., Wu M., Wu Y. (2009) An ACS-based framework for fuzzy data mining. *Expert Systems with Applications,* no. 36, pp. 11844–11852.
23. Ishibuchi H., Nozaki K., Tanaka H. (1993) Efficient fuzzy partition of pattern space for classification problems. *Fuzzy Sets and Systems,* vol. 59, no. 3, pp. 295–304.
24. Permana K., Zaiton S. (2010) Fuzzy membership function generation using particle swarm optimization. *International Journal of Open Problems in Computer Science and Mathematics*, vol. 3, no. 1, pp. 27–41.
25. Piatetsky-Shapiro G., Connell C. (1984) Accurate estimation of the number of tuples satisfying a condition. *ACM SIGMOD Record*, vol. 19, no. 2, pp. 256–276.

The impact of knowledge management capabilities and processes on SME performance

Naief G. Azyabi
Assistant Professor
Department of Management Information Systems, College of Business Administration
Jazan University
Address: Jazan, 45142, Saudi Arabia
E-mail: ngazyabi@jcba.edu.sa

Abstract

Knowledge management (KM) is a practice where knowledge is captured, distributed and utilized effectively, leading to enhanced productivity and performance of an organization. The prime objective of this study is to examine the influence of KM processes and capabilities on the performance of small and medium-sized enterprises (SMEs) in Saudi Arabia. KM capabilities comprise people, IT, the organizational structure and the organizational culture, which are measured in this research by T-shaped skills, IT support, the level of centralization, and learning. The other dimension is KM processes, which consist of accessing, generating, embedding, representing, facilitating, using, measuring and transferring knowledge. Moreover, KM performance is measured via two factors: the organization's financial performance and customer satisfaction.

The research reviews previous literature related to the KM components (processes, capabilities and performance) to develop the research model and a number of hypotheses to evaluate the research problem. The data is collected through a questionnaire-based survey completed by a total of 126 managers working in different sectors of Saudi SMEs. With the help of a number of statistical tests, the research study found that that the KM capabilities, IT support, learning culture, decentralized structure and the people of the organization contribute to the success of KM practices or processes, validating the theoretical model. The results also show that KM processes, including accessing, generating, measuring, transferring, use, embedding, representing and facilitating, are positively associated with the performance of SMEs in Saudi Arabia.

Key words: knowledge management; IT support; Saudi Arabia; organizational culture; learning; performance.

Citation: Azyabi N.G. (2018) The impact of knowledge management capabilities and processes on SME performance. *Business Informatics*, no. 3 (45), pp. 39–52. DOI: 10.17323/1998-0663.2018.3.39.52.

Introduction

The concept of knowledge management (KM) originated in 1990s and emphasized management of knowledge and information holistically in an organization. The process where knowledge is captured, distributed and utilized effectively is designated as KM [1]. It is an integrated approach to recognize, store, assess, retrieve and share/disseminate all the information assets of an organization. The information assets consist of procedures, policies, documents, databases, and uncaptured experience and expertise possessed by individual workers [2].

The main purpose of knowledge management (KM) in organizations is to enhance productivity and performance through acquiring, using or applying knowledge, converting it into useful forms, and embedding it by systematic and intentional methods in the organization's routine. To understand the concept of KM, it is essential to know the organization's innovation process, where individuals explore creative problem-solving methods. The new marketplace's dynamic nature has originated an incentive or a competitive need to reconcile and consolidate knowledge assets to create sustainable value. Many companies around the world are introducing extensive KM practices to achieve competitive sustainability [3].

The relationship between KM performance, processes and capabilities has been examined by most of the studies such as [4] and [5]. Some studies emphasized the linkage between KM capabilities and processes, while others focused on the association between organizational performance and capabilities [6; 7]. However, very few empirical studies have been conducted with respect to an integrative framework related to KM. Lee and Choi [4] based on relevant theories, stressed the integrative approach of the variables related to KM and presented a framework which includes organizational performance, intermediate outcomes, KM processes and enablers. It is important to identify and assess various factors within an organization that are essential for performance measurement of KM with a balanced view. This provides a better understanding of success and failure and KM.

This research study investigates the structural relationships among different KM value chain factors: KM capabilities, KM processes and small and medium enterprises' (SMEs) performance. For this purpose, the research study is divided into two phases. The first phase consists of reviewing the literature on KM performance, processes, and capabilities to evaluate the core KM value chain constructs, and suggests the KM's integrated framework. The second phase conducts a survey among SMEs in Saudi Arabia in order to examine this framework.

Similar research has been performed by Lee and Lee [7] to investigate the structural relationships between KM capabilities, KM processes and organizational performance. However, the current study investigates these relationships in different contexts. It is well known that the organization's size has an effect on its operations and decisions, thus it is argued that the results that have been found in large organizations could be different from SMEs. Moreover, little research has been conducted in the Saudi context, which has a different culture and regulations that can influence organizational activities and practices. This research contributes to the KM literature in the Saudi SME context by answering the following question: "to what extent do KM processes and capabilities have an influence on SME performance in Saudi Arabia?"

1. Review of the literature

This section presents a review of existing literature on the research components: KM capabilities, KM processes and organizational performance.

1.1. KM capabilities

For effective competition, organizations should leverage their existing knowledge and create new knowledge. Based on the organization's ability to manage their existing and new knowledge, their position in the market can be determined. The task requires organizations to develop the ability to use their previous knowledge, all of which facilitates recognition, assimilation and application of new information to create new capabilities and knowledge [8]. For effective KM, previous research studies have suggested that the KM capabilities are organizational resources or preconditions [3; 4; 9–11].

According to Krogh, Nonaka and Aben [12], the infrastructure of KM can be defined as the mechanism of an organization to create knowledge intentionally and constantly. Their study explained five factors of KM infrastructure, which consist of (a) human resources; (b) employees' relationships; (c) the organizational structure; (d) conversation between employees; and (e) the will to generate knowledge. Quinn [8] stated that in order to utilize the organizational knowledge assets, organizations must perform many activities such as developing a systematic organizational structure, developing technological capabilities and employees' abilities.

Gray [11] has found that the alignment between the organizational KM practices, knowledge creation, knowledge storage and retrieval can positively influence organizational performance. According to Gray [11], KM practices consist of (a) construction of a knowledge repository; (b) formal training; (c) talk rooms of research and development of employees related to the present tasks; (d) communities of practices; and (e) informal knowledge fairs.

An empirical exploration for the KM model that views KM from the capabilities perspective is performed by Gold, Malhotra, and Segars [3]. This model suggests that there are some essential preconditions for KM effectiveness such as a knowledge infrastructure (organizational culture, structure and technology) and knowledge processes (knowledge acquisition, transformation, application, and conservation). Lee and Choi [4] discussed KM capabilities and processes, and explained that the management of knowledge within an organization consists of knowledge processes and the enablers which support and maintain these processes. According to them, the KM enablers (capabilities) consist of IT support, people and the organizational structure and culture.

1.2. KM processes

KM processes are addressed by a number of works in the literature and research which classified the concept of KM into a number of processes [3; 4; 13–15]. The researchers have identified various key processes such as knowledge acquisition, knowledge transfer, knowledge creation, knowledge integration, knowledge exploitation and knowledge capturing [16–19].

For instance, Alavi and Leidner [20] discussed KM processes as creation, transfer, storage and application. Gold, Malhotra, and Segars [3] grouped/assembled a number of capabilities into the four main processes: knowledge acquisition, knowledge conversion, knowledge application and knowledge protection.

In another study, Lee and Choi [4] explored only knowledge creation, utilizing the Nonaka and Takeuchi [21] SECI model (socialization, externalization, combination and internalization). Another classification for KM processes was introduced by Ruggles [22] as:

✦ generating new knowledge and using external valuable knowledge;

✦ using incentives and developing a culture to facilitate knowledge growth, and documenting knowledge through software, databases and documents;

✦ routinizing the accessible knowledge and embedding it within organization norms and procedures;

✦ disseminating the existing knowledge and determining its impact and value on the organization.

1.3. Organizational performance

Among the management activities, performance measurement is considered to be one of the most important areas. The measurement of performance becomes the basis of achievement and establishment of strategy within the organization, because it evaluates how successful the organization is in achieving its strategic targets and communicating its vision to its stakeholders. The conventional tools of performance measurement, which mainly include financial reporting, enable organizations to compare their performance with others. However, these financial indicators are not the only indicators that can measure the organization's performance. Intangible assets such as knowledge play a vital role, apart from tangible assets in determining the growth and worth of organizations. Therefore, there were many attempts to measure organizational performance based on their tangible and intangible assets [4; 6; 23–27].

In a research study conducted by Sveiby [28] to assess organizational performance, an intangible asset monitor (IAM) was established. This IAM is used to measure intangible assets' performance, which include market, structural, and human capital. These performance indicators presented by the intangible asset monitor are simple and plain, and it classifies intellectual capital by external and internal structure and employee capabilities. It further provides three performance indicators: stability, efficiency and innovation/growth.

One of the most known performance measurements at strategic level tools is called the balanced scorecard (BSC). This was developed by Kaplan and Norton [24]. The BSC measures organizational performance based on four main perspectives: financial perspective, customer perspective, internal processes perspective, and innovation and learning perspective. Previous research studies have measured organizational performance with respect to market share, business size, innovativeness, growth rate, profitability and success from a subjective point of view, in comparison with the key competitors, to consider both operational and financial issues.

Arora [23] evaluated the purposes of KM, which include enhancement of employees' jobs through extended cooperation, innovation or creation of new knowledge. To support overall KM, Arora proposes communities of practice activations and construction of a knowledge repository. Gooijer [25] also discussed the importance of KM and proposed a methodology to measure organizational performance by introducing KM balanced scorecards. According to Gooijer, KM is a practice that enhances integration, collaboration and cooperation among employees.

2. Research model

This section discusses the main variables based on our review of the literature and identifies the major factors related to KM capabilities and processes. It further presents our research hypotheses.

2.1. Variables

2.1.1. KM capabilities

The KM capabilities incorporated in the study model include information technology (IT), organizational culture, organizational structure and people. The research study proposes that KM is influenced by IT and its capabilities. In this era of increasing innovation, the use of IT plays a crucial role for the success of any organization. IT is employed widely to establish networks or channels for people to connect and reuse the codified knowledge. IT also plays a significant role in creation of new knowledge by facilitating conversation and enabling sharing, using and storage of knowledge [29]. This study therefore focuses on the capability of IT in the form of ***IT support***, which is an essential part of the KM function.

For successful KM, the organizational culture also plays an important role. The organizational culture determines the organization's values and norms, and it is often considered as the most significant factor in KM success. The culture of sharing knowledge within an organization is beneficial for long-term success, which is associated with organizational learning. Thus, this research study focuses on the organization's ***learning*** as the dimension for organizational culture.

The third important factor is the organizational structure, which may inhibit or encourage the KM function. The impact of the organizational structure on KM is widely recognized [30]. In this research study, the structure of an organization is measured by the degree of ***centralization***, which is a key structural factor and signifies the concentration of decision-making and control within the organization.

People are the key factor in an organization to create and share knowledge, and they are therefore crucial to be managed effectively. Organizations acquire competences and knowledge by recruiting new members/people with the required and unique skills desired by the organization. Particularly, the employees ***T-shaped skills*** are considered as part of the core capabilities. In the individual specialist, these skills may enable synergistic conversation within the organization.

2.1.2. KM processes

Previous studies have acknowledged the processes and capabilities of KM as the antecedents for organizational performance. However, literature has also identified the organization's capabilities as knowledge process preconditions. Therefore, it is essential to understand the impact of an organization's capabilities on its KM processes. Considering KM processes as essential in the organization is based on the process-based view.

This research study adopts eight processes of KM to examine their role as proposed by Ruggles [22]: knowledge generation (***generating***), accessing external knowledge (***accessing***), facilitation of knowledge growth (***facilitating***), knowledge documentation (***representing***), knowledge embedment in processes and routine (***embedding***), knowledge use in decision making (***usage***), knowledge transfer among organization (***transferring***) and measurement of knowledge impact and value (***measuring***).

2.1.3. SME performance

The behavior of employees and managers is strongly affected by measuring organizational performance. From the KM perspective, the performance of an organization can be measured by a number of methods, which can categorized into balanced scorecard, intangible benefits, intellectual capital and financial measures. According to Kaplan and Norton [24], as compared to the tangible or intangible measurement approaches, and the intellectual capital approach, the balanced scorecard method is more beneficial and useful, since it provides the cause and effect associations between organization strategies and knowledge components.

For measuring organizational performance, this research study has adopted a modified method of the balanced scorecard which focuses on measuring customer satisfaction and financial outcomes, following the same approach of Lee and Lee [7].

The diagram below (*Figure 1*) presents the research model. The KM capabilities comprise people, IT, organizational structure and organizational culture, which are represented by IT support, centralization, T-shaped skills and learning. KM processes according to this model consist of accessing, generating, embedding, representing, facilitating, usage, measuring and transferring knowledge. Moreover, KM performance is divided into the organization's ***financial performance*** and ***customer satisfaction***.

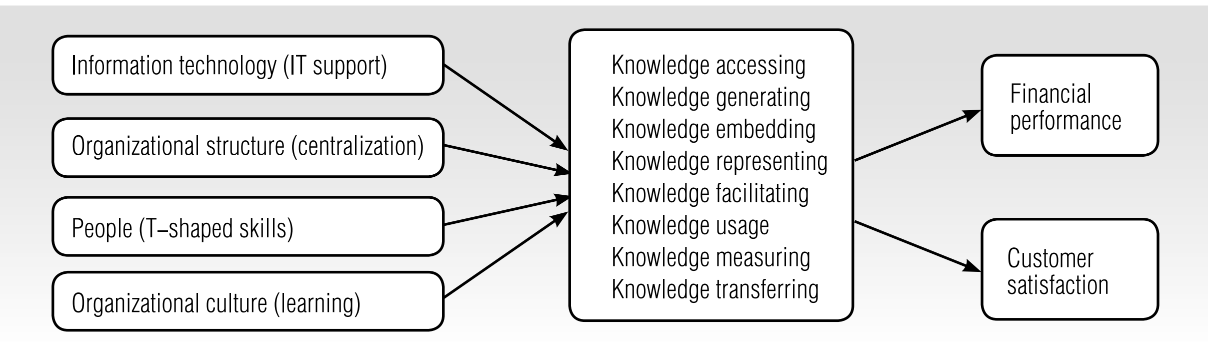

Fig. 1. The research model

2.2. Hypotheses

The research hypotheses are derived from the review of the literature review based on the theoretical statements on KM. The following hypotheses relate to information technology support, learning, centralization, T-shaped skills, and performance of the organization.

2.2.1. T-shaped skills

According to Leonard-Barton [31], the T-shaped skills are both broad and deep, as indicated by the shape of the letter 'T' with its horizontal and vertical parts. As discussed before, the research study has considered the T-shaped skills to measure the KM capabilities' variable – people. The processors of the T-shaped skills can explore particular knowledge domains. For creating new knowledge, the employees/people having T-shaped skills are considered very useful, because of their ability to integrate different knowledge assets.

The people with T-shaped skills can combine both practical and theoretical knowledge, and have the ability to observe how the interaction among different branches of knowledge takes place. Therefore, across several functional areas, they have the ability to expand their competence and create new knowledge. Hence the related hypothesis for the T-shaped skills is:

Hypothesis 1: *There is a positive association between KM processes and the T-shaped skills of the members of an organization.*

2.2.2. Centralization

In this research study, the structure of the organization is measured by the degree of centralization. Centralization in an organization hinders frequent sharing of ideas and interdepartmental communication, due to consumption of time in the communication channels, and results in discontinuousness and distortion of ideas [32].

In contrast, an organization with a decentralized organizational structure facilitates knowledge sharing, where employees tend to participate more in the process of knowledge building. In this kind of working environment, less emphasis is required on the work rules, and the knowledge processes need more flexibility. Therefore, in an organizational structure, increased flexibility may contribute to activated KM practices. The hypothesis for the organizational structure would be:

Hypothesis 2: *There is a negative association between KM processes and the degree of centralization.*

2.2.3. IT support

This explains the degree to which the use or implementation of information technology in an organization supports KM. A number of research works have found an important association between IT use and efficient KM processes, i.e. IT is an essential component of KM [3]. IT enables organizations to collect, exchange and store knowledge rapidly and on a huge scale. Furthermore, the fragmented flows of knowledge are integrated with a well-developed technology. The integration can eliminate communication barriers among various departments. All forms of knowledge processes are supported by IT; these include generation, transferring, usage, facilitating, etc. Therefore, the hypothesis for IT support is:

Hypothesis 3: *There is a positive association between KM processes and IT support.*

2.2.4. Learning

Acquisition of new knowledge is often observed learning by the members of an organization who are willing and able to apply the knowledge in the decision-making process. A learning culture should be developed and various learning means should be provided such as mentoring, training and education within organizations for efficient knowledge processes. Learning not only boosts the efficiency of KM processes but also contributes to the performance and success of an organization. Therefore, the hypothesis for learning culture is:

Hypothesis 4: *There is a positive association between KM processes and learning.*

2.2.5. Performance of the organization

The performance of an organization, in this is research study, is measured by the balanced scorecard's financial and customer perspectives as compared to key competitors. The organizational change goals typically include various aspects of performance of an organization including innovation, improvement, survival or effectiveness. The organizational performance can be regarded as the knowledge processes' output that encourages and improves these aspects. This means that the improvement in knowledge processes contributes to its performance. Therefore, the hypotheses are:

Hypothesis 5: *There is a positive association between financial performance and KM processes;*

Hypothesis 6: *There is a positive association between customer performance and KM processes.*

3. Methodology

The research approach of the study is a quantitative research approach, which tends to explore the research issues by using statistical models, functions and techniques. In this study, the impact of KM processes and capabilities on KM performance is explained with a 'correlational research design' and a 'descriptive design'. Bordens [33] explains that in a correlational design, the variables are measured and the association between them is defined. The descriptive research design here adopts a questionnaire based survey.

The questionnaire was filled out by a total of 126 SME managers working in different sectors; therefore, the data collection source is a primary, where data is directly collected from the respondents.

For all variables/constructs, multiple-item measures were developed. To enhance the confidence of accurate assessment and consistent measurement of the variables or constructs of interest, multiple-item measure was used. They are also used in the measurement process to enhance the validity and reliability of the measures. Moreover, the constructs related to KM processes, capabilities and performance were measured using a 5-point Likert scale.

The questionnaire for the research study was adopted from Lee and Lee [7]. It consists of a total of 34 items related to KM capabilities, processes and performance. The capabilities of KM comprise employees' IT support, the learning organization culture, centralization of the structure and T-shaped skills as shown in *Table 1.*

Table 1.

KM capabilities items

Variables	Items
Information technology (IT support)	Information sharing via Intranet
	Knowledge map for knowledge source
	Use of customer relationship management (CRM)
	Use of data warehouse
	IT support for information acquisition
Organizational culture (learning)	Clubbing and community gatherings
	Contents of job training
	Encouragement to attend seminars, etc.
	Informal individual development
	Formal training programs
Organizational structure (centralization)	Making decisions without approval
	Supervisor's permission to act
	Need to refer to others
	Making own decisions
	Acting without supervisor's consent
People (T–shaped skills)	Knowing core knowledge
	Employees expert in their tasks
	Employees can explain their task
	Employees having accurate know–how

The processes of KM consist of accessing, generating, using, inserting, representing, facilitating, measuring and transferring knowledge that are eight items as represented in *Table 2.*

Table 2.

KM processes items

Variables	Items
KM processes	Accessing valuable knowledge
	Using knowledge that is accessible in decisions
	Embedding knowledge in processes
	Representing knowledge in documents, etc.
	Facilitating growth of knowledge
	Generating new knowledge
	Determining the knowledge assets' value
	Transferring existing knowledge

The financial and customer performance of the SMEs was measured using the KM scorecard to enable the research study to examine the performance of the organization based on a balanced scorecard. The study used cognitive measures for measuring the financial and customer performances, because it is difficult to relate metric financial data (e.g. ROI, ROA, profits) with the KM initiatives. The performance variables and their items are presented in *Table 3* as compared with key competitors.

Table 3.

Performance items

Variables	Items
Financial performance	Greater economic value added
	Greater net profit
	Greater market share
	Greater return on investment
Customer performance	More customer retention
	More customer acquisition
	Greater customer satisfaction

4. Empirical analysis

4.1. Descriptive analysis

In the descriptive analysis, the related industries of the respondents are illustrated in *Table 4.*

Table 4.

Industry analysis

Industry	Frequency	Percent
Wholesale and retail	32	25.4
Consulting and business service	25	19.8
Real estate	21	16.7
Service industry	15	11.9
Information and communication	12	9.5
Construction and engineering	9	7.1
Banking and insurance	9	7.1
Petrochemicals	3	2.4
Total	**126**	**100.0**

According to the table, the highest percentage of respondents is from the wholesale and retail industry, which is equal to 25.4 percent, followed by the consulting and business service and the real estate sector, with a percentage of 19.8 and 16.7 respectively. The data shows that the respondents from the petrochemical industry are the lowest, i.e. only 2.4 percent.

Table 5 shows the percentage of departments to which the participants belong. Most of the respondents are from the Accounting and Finance Department that is 57.1 percent, which is followed by Personnel Management and Training, and Production Departments (10.3 percent and 9.5 percent respectively). The other departments participating in the study include Marketing, Management Information System, Research and Development, General Affairs, Planning and others.

4.2. Reliability analysis

The research study tested the basic assumption related to the items to measure the variables – IT support, learning organization, centralization, T-shaped skills, knowledge processes, financial performance and customer performance. It is also important to test reliability or consistency of the test/survey. For determining reliability, Cronbach's alpha test was applied on the extracted factors (*Table 6*).

Table 5.

Participants by departments

Departments	Frequency	Percent
Accounting and Finance	72	57.1
Personal Management and Training	13	10.3
Production	12	9.5
Marketing	7	5.6
Others	7	5.6
Management Information System	5	4.0
Research and Development	4	3.2
General Affairs	3	2.4
Planning	3	2.4
Total	**126**	**100.0**

According to *Table 6*, the value of Cronbach's alpha for each construct is higher than the assumed threshold of 0.6. Therefore, the internal consistency is acceptable.

4.3. Validity assessment

To test the validity of the items adopted by the research study, construct validity, content validity and criteria-related validity was performed.

4.3.1. Content validity

In this type of validity analysis, we observe to what extent the specific intended content domain is reflected by the measurement. In the research study, the degree of correspondence between the observed items and the conceptual definitions (financial performance, customer performance, knowledge processes, IT support, learning, centralization, and T-shaped skills) is high. The content validity in the research study is recognized by the extensive analysis of previous KM practices.

4.3.2. Construct validity

In this form of validity, the agreement between a specific measuring procedure or device and a theoretical concept is determined. The significance of construct validity is high in a theoretical research. For a research study, three steps should be followed to determine the construct validity. In the first step, it is important to specify the theoretical relationships, which is followed by an evaluation of the empirical relationships

Table 6.

Reliability analysis

Variables		Numbers of items	Cronbach's alpha
KM processes	Accessing, generating, representing, facilitating, measuring, transferring, using, embedding	8	0.921
KM capabilities	IT support	5	0.901
	Learning organization	5	0.926
	Centralization	5	0.897
	T–shaped skills	4	0.851
KM performance	Financial performance	4	0.931
	Customer performance	3	0.891

between the concepts' measures. Lastly, interpretation of the empirical evidence in crucial to clarify the construct validity. The construct validity in the research study is tested using confirmatory factor analysis.

As discussed before, the items used and developed in the research model relating to the different constructs were adopted from a previous research model and theoretical evidence. The confirmatory factor analysis, in the construct development from theory-driven approach, delivers the appropriate method to evaluate the consistency of the structural equation model and the efficacy of measurement. Therefore, it is expected that the associated factors with the developed scales will be uniquely measured and the system of relationships will be represented by a system of factors. In this way, measurement efficacy is provided and the likelihood of confusion is reduced in structural equation modelling.

To examine the measurement strength between the constructs and the associated items, the study estimated three types of measurement models. The first model of measurement evaluates the relationships system among the KM capabilities' measures — IT support, learning, centralization and T-shaped skills (Table 7). The measures of KM process and the system of relationship among them are examined by the second measurement model (*Table 8*). KM performance measures (financial and customer performance) and the system of relationships are examined by the third measurement model (*Table 9*). The following statistical metrics were used: chi-square, *p*-value, root mean square residual (RMR), confirmatory fit index (CFI), non-normed fit index (NNFI), normed fit index (NFI), adjusted goodness of fit index (AGF), goodness of fit index (GF).

According to Table 7, the chi-square values of the four constructs of KM capabilities are significant with the *p*-values less than 0.05 threshold, indicating a good model fit. The non-normed fit index (NNFI), normed fit index (NFI), adjusted goodness of fit index (AGF) and goodness of fit index (GF) are all very high, which suggests the goodness of model fit. This shows that a significant amount of variation is captured by each scale in these KM capabilities' latent dimensions.

According to *Table 8*, the chi-square value is equal to 112.29, which is significant with the *p*-values less than 0.05 thresholds. The non-normed fit index (NNFI), normed fit index (NFI), adjusted goodness of fit index (AGF), and goodness of fit index (GF) are high, suggesting the goodness of model fit.

In *Table 9*, the third measurement models are presented related to the two KM performance measures. According to the statistics, chi-square values of the two variables are significant (0.05 significance level). The NNFI, NFI, AGF and GF indices are high and deducing good model fit. Therefore, in this section, the construct validity was verified using the confirmatory factor analysis.

Table 7.

KM capabilities measurement model

	IT support	Learning culture	Centralization	T-shaped skills
Chi-square	18.91	29.11	25.09	14.119
***P*-value**	0.002	0.000	0.000	0.010
RMR	0.029	0.049	0.029	0.033
CFI	0.913	0.982	0.970	0.969
NNFI	0.982	0.972	0.897	0.952
NFI	0.982	0.945	0.971	0.966
AGF	0.948	0.838	0.901	0.913
GF	0.955	0.898	0.972	0.928

Table 8.

KM processes measurement model

	Knowledge processes
Chi-square	112.29
***P*-value**	0.000
RMR	0.061
CFI	0.911
NNFI	0.812
NFI	0.902
AGF	0.851
GF	0.882

Table 9.

KM performance measurement model

	Financial performance	Customer performance
Chi-square	7.912	4.123
***P*-value**	0.019	0.036
RMR	0.021	0.029
CFI	0.962	0.945
NNFI	0.971	0.967
NFI	0.989	0.978
AG	0.912	0.901
GF	0.978	0.992

4.3.3. Criteria-related validity

In this form of validity, we have the degree to which future performance is predicted by the performance in an assessment i.e. the predictive capability. For instance, the degree to which KM performance is estimated accurately by the KM processes indicates the criteria-related validity.

In this research study, correlated analysis was conducted with the summated scale variables (KM capabilities – IT support, learning, centralization and T-shaped skills, KM processes, and the customer and financial performance). The summated scale's purpose is to raise the representative nature of constructs and to reduce measurement error. In this analysis, the higher value of mean suggests agreement on the constructs' definition. Table 10 illustrates the constructs' correlation analysis using a summated scale.

4.4. Factor analysis

The research study conducted an exploratory factor analysis, using the principle components method for the seven variables (IT support, learning organization, centralization, T-shaped skills, knowledge processes, financial performance and customer performance) to reduce and summarize the number of items.

Table 10.

Correlation coefficients matrix with mean and standard deviation

Variables	Mean	Standard deviation	T-shaped skills	Centralized structure	Learning culture	IT support	KM processes	Customer performance	Financial performance
T-shaped skills	0.32	0.42	1.000	0.254	0.306	0.236	0.289	0.181	0.082
Centralized structure	0.63	0.87	0.254	1.000	0.409	0.184	0.451	0.077	0.057
Learning culture	0.12	0.58	0.306	0.409	1.000	0.344	0.701	0.125	0.037
IT support	0.24	0.46	0.236	0.184	0.344	1.000	0.598	0.056	0.282
KM processes	0.07	0.48	0.289	0.451	0.701	0.598	1.000	0.137	0.106
Customer performance	0.12	0.55	0.181	0.077	0.125	0.056	0.137	1.000	0.258
Financial performance	0.00	0.73	0.082	0.057	0.037	0.282	0.106	0.258	1.000

The tests (*Table 11*) are performance to test the appropriateness of the data for factor analysis. The KMO value equals 0.771, which measures the sampling adequacy. Since the value is above 0.6 (the assumed minimum criteria), the sample is considered to be adequate for the principle components method. The Bartlett's test is showing a sig. value less than 0.05, indicating the appropriateness of running factor analysis.

Table 11.

KMO and Bartlett's test to measure appropriateness

KMO measure of sampling adequacy		0.771
Bartlett's test of sphericity	Approx. chi–square	7353.154
	df	561
	Sig.	0.000

Table 12 shows the results from the rotated component matrix, which is extracted using the principle component method. The correlated individual items related to the seven variables placed in their specific components based on the extraction. The KM process items, the KM capabilities and the performance related items are summarized and grouped into the seven components.

4.5. Structural equation model

In the formation of the research model, it was assumed that the capabilities of KM impact the processes, which may then influence the KM performance. As conceptualized in the previous chapters, the information technology capabilities, structure, culture, and people predict and influence the knowledge processes within an organization, whereas, the knowledge processes' distinct causal paths predict the performance (financial and customer perspectives).

Table 13 shows that the model's chi-square value is 890.12, which is highly significant. The values of NNFI, NFI, adjusted goodness of fit (AGF) and goodness of fit (GF) indices also meet recommended levels.

Table 12.

Rotated component matrix

	Components						
	1	2	3	4	5	6	7
P1	**0.982**	0.056	–0.033	–0.064	0.047	–0.079	0.014
P2	**0.953**	0.020	–0.021	–0.083	0.023	–0.057	0.016
P3	**0.961**	0.053	–0.030	–0.103	0.056	–0.037	–0.009
P4	**0.923**	0.008	–0.001	–0.062	0.072	0.009	0.042
P5	**0.952**	0.073	–0.023	–0.084	0.039	–0.073	–0.020
P6	**0.952**	0.103	–0.042	–0.046	–0.019	–0.018	0.021
P7	**0.949**	0.120	–0.040	–0.066	–0.005	–0.049	0.015
P8	**0.883**	0.001	–0.056	–0.045	0.028	–0.089	0.161
C1	0.071	**0.977**	0.113	–0.082	0.018	–0.034	0.075
C2	0.079	**0.966**	0.104	–0.082	0.013	–0.011	0.085
C3	0.088	**0.947**	0.110	–0.114	0.048	–0.074	0.046
C4	0.010	**0.955**	0.089	–0.113	–0.027	–0.051	0.033
C5	0.075	**0.968**	0.112	–0.089	0.024	–0.037	0.057
L1	–0.015	0.107	**0.968**	–0.021	–0.011	0.121	–0.001
L2	–0.015	0.113	**0.971**	–0.047	–0.014	0.109	–0.002
L3	–0.015	0.098	**0.965**	–0.021	0.005	0.121	0.004
L4	–0.072	0.119	**0.932**	–0.048	–0.043	0.064	–0.014
L5	–0.055	0.078	**0.948**	–0.084	–0.049	0.072	0.022
T1	–0.115	–0.110	–0.042	**0.974**	–0.012	–0.035	–0.043
T2	–0.112	–0.089	–0.046	**0.961**	0.006	–0.034	–0.068
T3	–0.097	–0.122	–0.057	**0.955**	0.002	0.028	–0.021
T4	–0.095	–0.125	–0.065	**0.956**	–0.035	–0.031	0.012
S1	0.325	–0.134	0.121	0.069	**0.635**	–0.047	0.036
S2	0.046	0.001	–0.033	–0.038	**0.955**	0.065	0.047
S3	0.031	0.033	–0.023	–0.022	**0.950**	0.018	0.011
S4	0.036	0.024	–0.065	0.001	**0.883**	0.077	0.040
S5	0.018	0.070	–0.038	–0.015	**0.944**	0.006	0.017
FP1	–0.083	–0.021	0.106	0.051	0.055	**0.919**	–0.102
FP2	–0.117	–0.134	0.089	–0.070	0.034	**0.895**	–0.068
FP3	–0.099	–0.028	0.101	–0.034	–0.002	**0.951**	–0.062
FP4	–0.014	–0.009	0.152	–0.017	0.060	**0.893**	–0.119
CP1	0.092	0.129	0.033	–0.045	0.069	–0.136	**0.940**
CP2	0.050	0.113	–0.003	–0.024	0.022	–0.131	**0.917**
CP3	0.033	0.014	–0.016	–0.038	0.037	–0.064	**0.887**

Table 13.

The hypothesized model fitness

Chi–square = 890.12
P–value = 0.000
RMR = 0.066
CFI = 0.912
NNFI = 0.891
NFI = 0.854
AGF = 0.789
GF = 0.801

The estimated model's path coefficients support the theorized associations in magnitude and direction (*Table 14*). This implies that KM capabilities (IT support, learning culture, decentralized structure and the people of the organization) contribute to the success of KM practices or processes. Meanwhile, successful KM processes enhance the performance of the SMEs in Saudi Arabia.

Table 14.

Hypothesis test results

Hypotheses/Relationships	*t*-value	Path coefficients
T–shaped skills and KM process	4.712	0.355
Centralization and KM process	–3.799	–0.208
IT support and KM process	4.839	0.368
Learning culture and KM process	5.015	0.349
KM process and financial performance	9.320	0.745
KM process and customer performance	4.991	0.469
Financial and customer performance	3.201	0.291

It is important to explain that the theoretical perspective developed in previous sections is consistent with the relationships' mathematical manifestation. A more precise aspect is the contribution of these results and the resultant associations. Although the chi-square values (model fits) are observed to be moderate in strength, the fit measures should be balanced with the complexity of the model.

The item loadings strength, directional path consistency, and the theory match imply that the proposed model of the research study provides valid insight into the KM process, capabilities and the organizational performance in Saudi Arabia.

Conclusion

The objective of this research study was to examine the influence of KM processes and capabilities on the performance of SMEs in Saudi Arabia. The study for this purpose reviewed previous literature on the subject and adopted a quantitative research approach, with an explanatory purpose to examine the research problem. The study has focused on the analysis and discussion of capabilities related to KM, to examine its impact on the facilitation of knowledge process leading towards greater competiveness of an organization.

The research conducted a questionnaire-based survey of 126 respondents related to different sectors. With the help of a number of statistical tests, the research study found that that the KM capabilities (IT support, learning culture, decentralized structure and the people of the organization) contribute to the success of KM practices or processes, validating the theoretical model. The results also show that the KM processes, including accessing, generating, measuring, transferring, usage, embedding, representing and facilitating, are positively associated with the performance of SMEs in Saudi Arabia.

The research study provides strong evidence regarding the impact and association of KM processes and capabilities with performance on organizations (SMEs). However, there are a few limitations related to the study, which include the cross-sectional design of the study overlooking time-lag effects. It is recommended that future researchers perform a longitudinal study for further robust results. The research was focused on some small and medium firms in Saudi Arabia, so it is difficult to generalize the results over the whole population. ■

References

1. Davenport T.H. (1994) Saving IT's soul: Human-centered information management. *Harvard Business Review*, vol. 72, no. 2, pp. 119–131.
2. Duhon B. (1998) It's all in our heads. *Inform*, vol. 12, no. 8, pp. 8–13.
3. Gold A.H., Malhotra A., Segars A.H. (2001). Knowledge management: An organizational capabilities perspective. *Journal of Management Information Systems*, vol. 18, no. 1, pp. 185–214.
4. Lee H., Choi B. (2003) Knowledge management enablers, processes, and organizational performance: An integrative view and empirical examination. *Journal of Management Information Systems*, vol. 20, no. 1, pp. 179–228.
5. Ho C.T. (2009) The relationship between knowledge management enablers and performance. *Industrial Management & Data Systems*, vol. 109, no. 1, pp. 98–117.
6. Mills A.M., Smith T.A. (2011) Knowledge management and organizational performance: a decomposed view. *Journal of Knowledge Management*, vol. 15, no. 1, pp. 156–171.
7. Lee Y.C., Lee S.K. (2007) Capabilities, processes, and performance of knowledge management: A structural approach. *Human Factors and Ergonomics in Manufacturing & Service Industries*, vol. 17, no. 1, pp. 21–41.
8. Quinn J.B. (1999) Strategic outsourcing: leveraging knowledge capabilities. *Sloan Management Review*, vol. 40, no. 4, pp. 9–21.
9. Zack M., McKeen J., Singh S. (2009) Knowledge management and organizational performance: An exploratory analysis. *Journal of Knowledge Management*, vol. 13, no. 6, pp. 392–409.
10. Bhatt G.D. (2001) Knowledge management in organizations: examining the interaction between technologies, techniques, and people. *Journal of Knowledge Management*, vol. 5, no. 1, pp. 68–75.
11. Gray P.H. (2001) A problem-solving perspective on knowledge management practices. *Decision Support Systems*, vol. 31, no. 1, pp. 87–102.
12. Krogh G., Nonaka I., Aben M. (2001) Making the most of your company's knowledge: A strategic framework. *Long Range Planning*, vol. 34, no. 4, pp. 421–439.
13. Lee C., Yang J. (2000) Knowledge value chain. *Journal of Management Development*, vol. 19, no. 9, pp. 783–794.
14. Sabherwal R., Becerra-Fernandez I. (2003) An empirical study of the effect of knowledge management processes at individual, group, and organizational levels. *Decision Sciences*, vol. 34, no. 2, pp. 225–260.
15. Rollett H. (2012) *Knowledge management: Processes and technologies*. Springer Science+Business Media.
16. Teece D.J. (2015) Intangible assets and a theory of heterogeneous firms. *Intangibles, market failure and innovation performance*. Springer International Publishing, pp. 217–239.
17. Skyrme D.J., Amidon D.M. (1998) New measures of success. *Journal of Business Strategy*, vol. 19, no. 1, pp. 20–24.
18. Dalkir K. (2013) *Knowledge management in theory and practice*. Routledge.
19. Hislop D. (2013) *Knowledge management in organizations: A critical introduction*. Oxford University Press.
20. Alavi M., Leidner D.E. (2001) Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly*, vol. 25, no. 1, pp. 107–136.
21. Nonaka I, Takeuchi H. (1995) *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. New York: Oxford University Press.
22. Ruggles R. (1998) The state of the notion. *California Management Review*, vol. 40, no. 3, pp. 80–89.
23. Arora R. (2002) Implementing KM – a balanced score card approach. *Journal of Knowledge Management*, vol. 6, no. 3, pp. 240–249.
24. Kaplan R.S., Norton D.P. (2000) Having trouble with your strategy? Then map it. *Harvard Business Review*, vol. 78, no. 5, pp. 167–176.
25. Gooijer J. (2000) Designing a knowledge management performance framework. *Journal of Knowledge Management*, vol. 4, no. 4, pp. 303–310.
26. Fugate B.S., Stank T.P., Mentzer J.T. (2009) Linking improved knowledge management to operational and organizational performance. *Journal of Operations Management*, vol. 27, no. 3, pp. 247–264.
27. Ho L.A. (2008) What affects organizational performance? The linking of learning and knowledge management. *Industrial Management & Data Systems*, vol. 108, no. 9, pp. 1234–1254.
28. Sveiby K.E. (1997) *The new organizational wealth: Managing and measuring knowledge assets*. San Francisco, CA: Berrett-Koehler.
29. Sher P.J., Lee V.C. (2004) Information technology as a facilitator for enhancing dynamic capabilities through knowledge management. *Information & Management*, vol. 41, no. 8, pp. 933–945.
30. Alawamleh H.S., Kloub M.A. (2013) Impact of organizational structure on knowledge management in the Jordanian insurance companies: From the perspective of the supervisory leadership. *International Journal of Business and Social Science*, vol. 4, no. 11, pp. 82–95.
31. Leonard-Barton D. (1998) *Wellsprings of knowledge: Building and sustaining the sources of innovation*. Boston: Harvard Business School Press.
32. Liao C., Chuang S.H., To P.L. (2011) How knowledge management mediates the relationship between environment and organizational structure. *Journal of Business Research*, vol. 64, no. 7, pp. 728–736.
33. Bordens K. (2006) *Research Design & Methods*. New Delhi: Tata McGraw-Hill.

Determination of the trading discount based on market data and cadastral value

Mikhail B. Laskin
Senior Researcher
St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences;
Associate Professor, Faculty of Economics
St. Petersburg University
Address: 7–9, Universitetskaya Emb., St. Petersburg, 199034, Russia
E-mail: laskinmb@yahoo.com

Abstract

The introduction of the institution of cadastral value in the Russian Federation opens up new opportunities in real estate valuation. In this regard, the new focus for appraisers is statistical analysis of multidimensional empirical distributions that were not previously available, because the real estate market does not have pairwise and multidimensional observations concentrated in unified databases. Data of interest to analysts is usually concentrated in different sources from different owners and pertains to different objects. The goal of combining them can be solved by comparing such data with the data of cadastral records, namely the cadastral number as a unique identifier of the object. Since the cadastral value corresponds to each cadastral number, it is possible to compare the cadastral value with important indicators such as the market price of the offer, the transaction price, the rental rate, the annual price indices, the capitalization rate, the discount rate, the trading discount and many other indicators, the formation of which involves more than two random variables. The construction of the model involves the principle of following the prices formed by pair comparisons to geometric Brownian motion, and hence the formation of lognormal population. As it turned out, as a result of large-scale cadastral work carried out in the Russian Federation in 2014, the cadastral value is also subject to lognormal distribution of prices (in each class of objects). For the market value (as the most probable price of the transaction under conditions of perfect competition), this leads to functional dependences from the cadastral value of the power type. Similarly, many other indicators will also be subject to dependences in the form of power functions. Obviously, having a function depending from the various indicators of cadastral value across the set of values, you can set the relationship between the various indicators, which was impossible before the introduction of the institution of cadastral value. This article proposes a method of calculating the trading discount when appraising real estate based on analysis of market statistics and databases of cadastral records. An analytical formula of the dependence of trading discounts from the offer price is proposed. The method allows us to set the level of the discount not only for objects included in an advertising database, but also for any object that has undergone cadastral registration.

Key words: trading discount; cadastral value; market value; stochastic pricing model; lognormal distribution of prices.

Citation: Laskin M.B. (2018) Determination of the trading discount based on market data and cadastral value. *Business Informatics*, no. 3 (45), pp. 53–61. DOI: 10.17323/1998-0663.2018.3.53.61.

Introduction

In the theory and practice of real estate appraising, one of the unresolved issues is determination of trading discounts. The Russian real estate market is structured in such manner that there is a lot known regarding offers to sell and very limited access (for the appraiser) to information about completed transactions with the actual sales prices. For this reason, it is very inconvenient to find statistics reflecting the offer price and the transaction price for one and the same object of real estate. In practice, appraisers often use "expert" surveys and average out the data obtained. In rising markets, the errors made are quickly compensated by dynamic changes in prices. In present-day conditions, such an approach often does not reflect the real picture.

With the introduction of the institution of cadastral inventory, new possibilities arose to establish the analytic dependence between the offer price and the transaction price, consequently to determine the trading discount.

1. The model

The model is based on the assumption that that bid prices and purchase prices follow lognormal distribution. Cadastral values were determined in 2014 as market values, and also followed lognormal distribution. Such a form of distribution of prices in real estate was pointed out not only by Aitchinson and Brown (the researchers of the University of Cambridge) back in 1963 [1], but also by authors of relatively recent publications [2–4]. Rusakov [5; 6] proved that prices formed by sequential comparisons under conditions of perfect competition aim at formation of lognormal distribution of the population. Market value is understood to mean the most probable price at which the object can be disposed of on the open market under conditions of perfect competition. This understanding fully corresponds to the definition of market value given in the Russian Federal Low No. 135 "About valuation activity" [7] and to foreign appraisal standards [8–11]. For one-dimensional random variables, this leads to an understanding of market value as the modal value of the respective random variable, and in cases of two-dimensional or multidimensional random variables – for the study of the most probable values of multi-dimensional random variables, for example, when studying adjustments [12], the analysis of the auction results [13], the determination of the capitalization ratio corresponding to the current market conditions [14] and other tasks in which the need to review the system of random variables arises.

Since collecting statistics containing the bid prices and purchase prices for each object is inconvenient, one can collect the statistics by the pairs "cadastral value – bid price" and "cadastral value – purchase price" then show the analytical dependences and finally show the dependence relationship of the transaction price on the offer price.

For example, if we suppose that the random variables V_{cv} and V_{op} (V_{cv} – cadastral value, V_{op} – offer price) follow lognormal distribution, then given a fixed cadastral value $V_{cp} = v$ the most probable price MV_{op} (market value appraised using offer prices without consideration of trading discounts) equals:

$$MV_{op} = \mathrm{Mode}(V_{op} \mid V_{cv} = v) =$$
$$\exp(\mu_1 + \rho_1 \frac{\sigma_1}{\sigma_2}(\ln(v) - \mu_2) - \sigma_1^2(1 - \rho_1^2)), \qquad (1)$$

where μ_1, σ_1 – the parameters of lognormal distribution for the cadastral value;

μ_2, σ_2 – the parameters of lognormal distribution for the bid prices;

ρ_1 – the correlation coefficient.

The straightforward finding of this formula is given in the works [14; 15]. We note that the formula (1) describes the dependence of market value without considering trading discounts from the cadastral value in the form of a power function:

$$MV_{op} = \mathrm{Mode}(V_{op} \mid V_{cv} = v) = A_1 v^{B_1}, \qquad (2)$$

where $A_1 = \exp(\mu_1 - \rho_1 \frac{\sigma_1}{\sigma_2}\mu_2 - \sigma_1^2(1-\rho_1^2))$;

$B_1 = \rho_1 \frac{\sigma_1}{\sigma_2}$.

A similar formula can be written as well for appraising market value using purchase prices (MV_{pp}):

$$MV_{pp} = \mathrm{Mode}(V_{op} \mid V_{cv} = v) = A_2 v^{B_2}, \qquad (3)$$

where $A_2 = \exp(\mu_1 - \rho_2 \frac{\sigma_1}{\sigma_3}\mu_3 - \sigma_1^2(1-\rho_2^2))$;

$B_2 = \rho_2 \frac{\sigma_1}{\sigma_3}$

(here μ_3, σ_3 is the parameters of lognormal distribution of purchase prices, ρ_2 is the correlation coefficient of the pair "cadastral value – purchase price").

Then, expressing v through MV_{op} (from the formula (2)) and substituting in the formula (3), we obtain the price dependence of the most likely transaction price MV_{pp} from the most probable bid price MV_{op}:

$$MV_{pp} = A_2 \left(\frac{MV_{op}}{A_1} \right)^{\frac{B_2}{B_1}}. \qquad (4)$$

It is obvious that this dependence also is a power function.

Let us introduce the coefficient K – the ratio of the most probable purchase price to the most probable bid price. From formula (4) we see the power dependence of the coefficient K from the most probable bid price:

$$K = A_2 A_1^{-\frac{B_2}{B_1}} (MV_{op})^{\frac{B_2}{B_1} - 1}. \qquad (5)$$

From formula (5), it is apparent that the trading discount is not constant and given the defined parameters of the laws of distributions of bid prices, purchase prices and cadastral values can significantly change for different offer prices.

2. Calculation example

For our calculations, we have selected data from the Real Estate Bulletin No. 1758 (December 2016)[1], data from the cadastral records of objects in the residential housing stock of St. Petersburg published in [16], on the basis of the report [17] and data from one of the St. Petersburg real estate agencies which presented on a confidential basis its data on the prices of 294 actual transactions in 2017. The information presented by the agency relates only to necessary numerical indicators and does not contain any personal data. Adjustments for time during 2017 were not made inasmuch as there were no significant changes on the market. Unfortunately, databases of an advertising nature such as the "Real Estate Bulletin" do not contain data from the cadastral records (above all, the cadastral number, as the unique identifier of the object and of its cadastral value). The community of appraisers should consider the possibility of creating databases containing not only the market information and description of the price-forming factors but also data from the cadastral records, because the cadastral number is the unique identifier of all real estate objects. If all databases contained the cadastral number (and data from the cadastral records), then it would be possible to integrate the data from various databases into one and examine the multi-dimensional random variables. For the moment, we have to apply other technique to identify the objects. In particular, for data taken from the "Real Estate Bulletin," we succeeded in identifying (establishing the cadas-

[1] Portal of the Bulletin of St. Petersburg Real Estate" (www.bn.ru) dated 8 December 2016, No. 1758

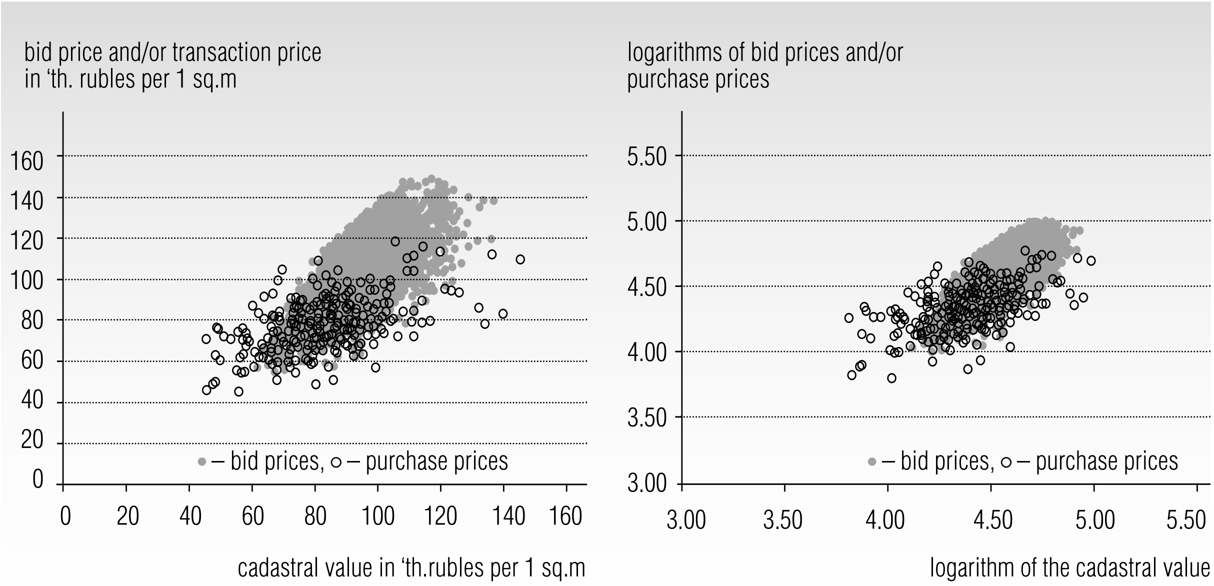

Fig. 1. Scatter–plot chart of the pairs "cadastral value – bid price" and "cadastral value – purchase price" for secondary real estate in St. Petersburg in 2017, for the "mass market" segment in natural prices (on the left) and on the logarithmical plane (on the right)

tral number and cadastral value) of more than 2300 objects. In *Figure 1* we presented the scatter-plot charts of the pairs "cadastral value – bid price," "cadastral value – purchase price" for secondary real estate in St. Petersburg in 2017, for the mass-market segment (the premium segment was not examined as it is a segment requiring separate study).

Empirical distributions of bid prices, very often, do not allow us satisfactorily to approximate right away lognormal distribution. This is explained, above all, by the fact that the market data are a mixture of distributions of prices for objects of various classes. The task of separating out the mixtures belongs to the class of incorrect tasks and cannot be solved by mathematical methods only. Fortunately, in the evaluation of a property, to separate out the mixtures you always can use a set of price-forming factors, part of which is reflected in the market databases. In the given case, a satisfactory approximation was obtained after removing from the sample objects relating to the premium segment and removing an insignificant number of outliers with abnormally elevated expectations of sellers (the remaining size of the sample – 2331 objects). The approximating surfaces are presented in *Figure 2.*

The approximating surfaces have been tested by the circular KS test [14; 15]. The results are presented in *Figure 3.*

In the test for the pair "cadastral value – bid price," the minimal value *p*-value comes to 0.055; and for the pair "cadastral value – purchase price" – the *p*-value equals 0.33. The values obtained provide a basis for not rejecting the hypothesis of combined lognormal distribution.

When modeling, we obtained the following sample values of the parameters of combined lognormal laws of distribution. For the pair "cadastral value – bid price," μ_1=4.53, σ_1=0.126, μ_2=4.56, σ_2 =0.178, ρ_1 =0.756. By substituting the indicated values in the formula (2) we get:

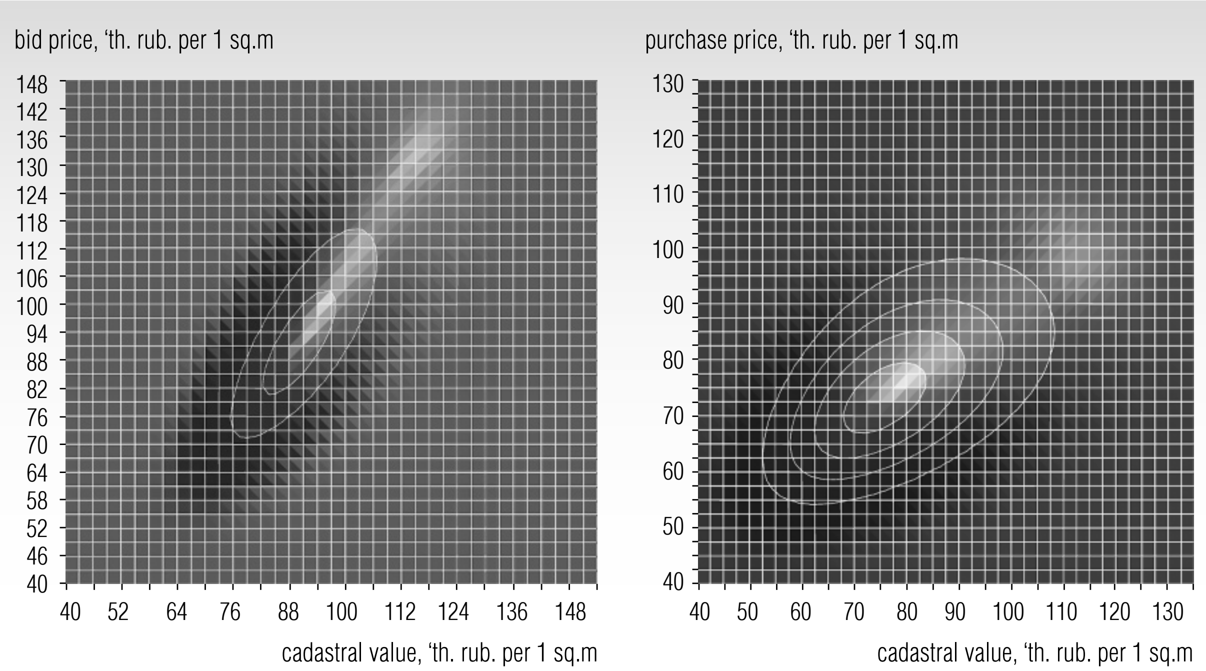

Fig. 2. Approximating surfaces of combined lognormal dstribution of the pairs "cadastral value – bid price" (left) and "cadastral value – purchase price" (right)

$$A_1 = 0.758$$
$$B_1 = 1.064$$
$$MV_{op} = 0.758v^{1.064}. \quad (6)$$

For the pair "cadastral value — purchase price" $\mu_1 = 4.384$, $\sigma_1 = 0.207$, $\mu_3 = 4.337$, $\sigma_3 = 0.17$, $\rho_2 = 0.52$. Substituting the indicated values in the formula (3) we get:

$$A_1 = 11.72$$
$$B_1 = 0.422$$
$$MV_{op} = 11.72v^{0.422}. \quad (7)$$

Both dependences are presented in *Figure 4.*

The dependences obtained for valuations of market value come from different statistics: for market data — on the basis of announcements of sales and for sales data, real estate agencies. These dependences are built on analysis of data from various samples. The objects from these samples are various and are in no way connected to one another except for cadastral values.

Objects having the same cadastral value are not necessarily placed on sale and sold for the same price. However, if when determining the cadastral value the following principle is observed — "cadastral value is determined as the market value obtained by the methods of mass evaluation on the date of appraisal" — then the appraiser has the right, when constructing his valuation, to proceed from the assumption that two such objects have roughly the same market value. Furthermore, since the cadastral value is established for the purposes of taxation, the principle of social justice presupposes that for objects with identical market value the owners should pay the same taxes. That is precisely why the task of periodic review of cadastral value arises.

In our case, cadastral value also acts as the instrument for highlighting the dependence between the bid price and the purchase price (or between the bid price and the multiplier coefficient determining the trading discount). By substituting the values of the coefficients (6) and (7) in the formulas (4) and (5), we get:

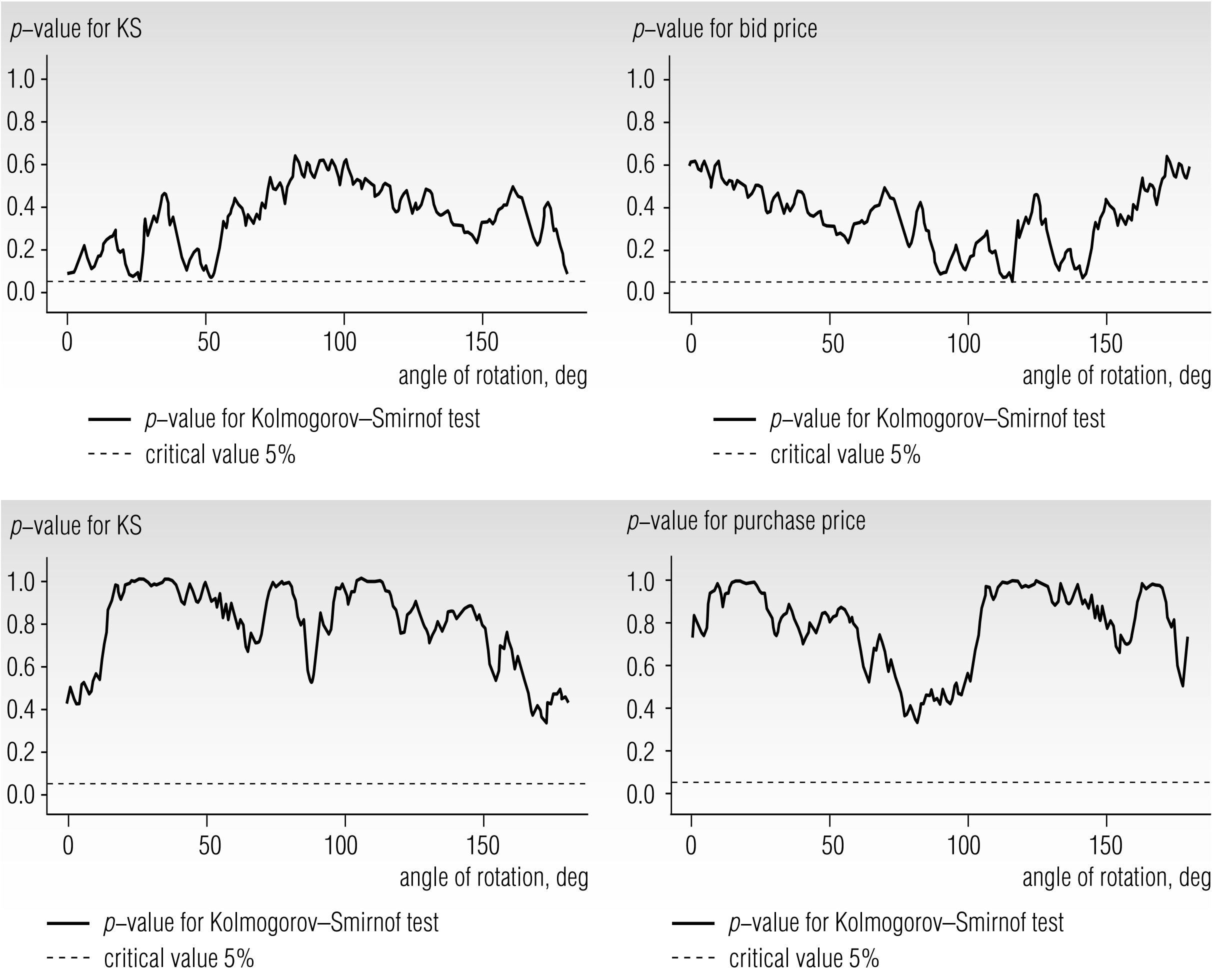

Fig. 3. Results of the circular KS test for the pair “cadastral value – bid price” (upper row) and for the pair “cadastral value – purchase price” (lower row)

$$MV_{pp} = 13.08\ MV_{op}^{\ 0.396}$$

$$K = 13.08\ MV_{op}^{\ -0.604}$$

Both dependences are presented in *Figure 5.*

The dependence indicated in the right chart of *Figure 5* de facto describes the valuation of the trading discount in the sector of secondary real estate in St. Petersburg in 2017. As we see, the valuation of the trading discount is not constant. For the sector of secondary residential real estate in the mass-market segment, at a bid price of 140,000 rubles/sq.m. it can amount to 35%. In addition, this same line shows that a bid price below 70,000 rubles/sq,m. should be applied with caution. For example, a bid price of 60,000 rubles/sq.m. lags behind the most probable sale price by 10% (approximately 66,000 rubles/sq.m.). However, this does not mean that the seller, having received such information, must immediately set a price as much higher as possible. Most sellers proceed from rational considerations, weighing the time for showing on the market with their own need for cash and with their wish to sell the object more expensively. Basically such information is needed by the community of appraisers to make a justified selection of trading discount when appraising objects of real estate.

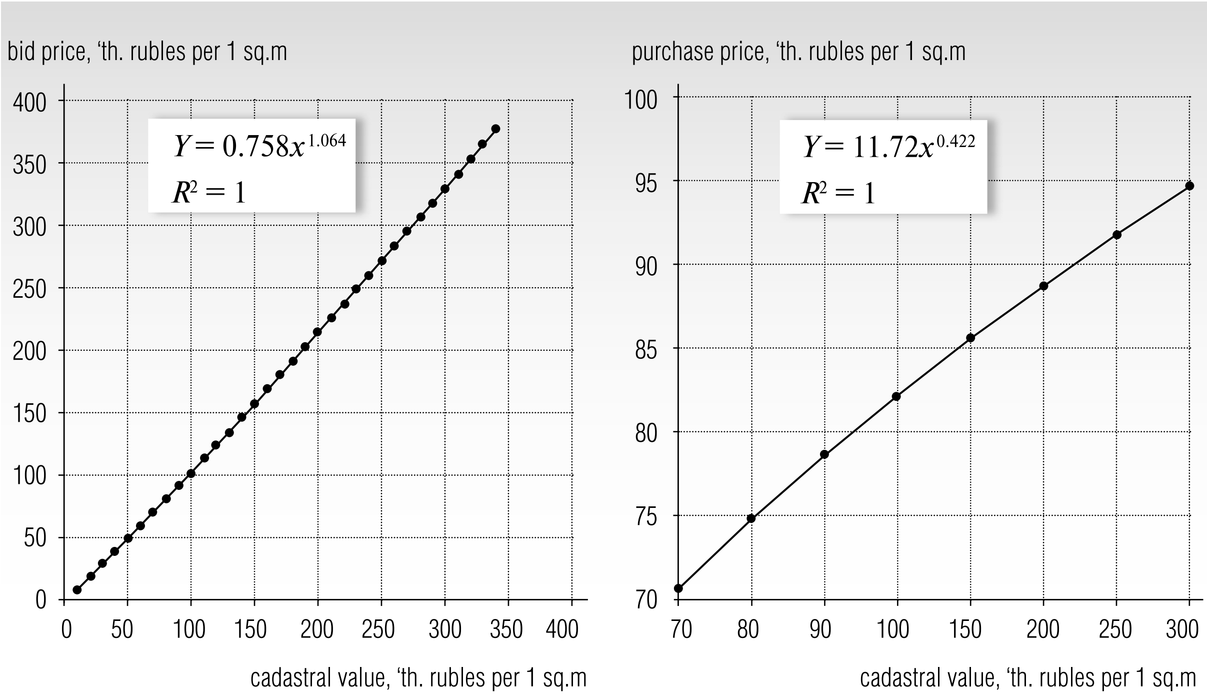

Fig. 4. The dependences of market value determined via bid prices from cadastral value (left) and market value obtained via purchase prices from cadastral value (right)

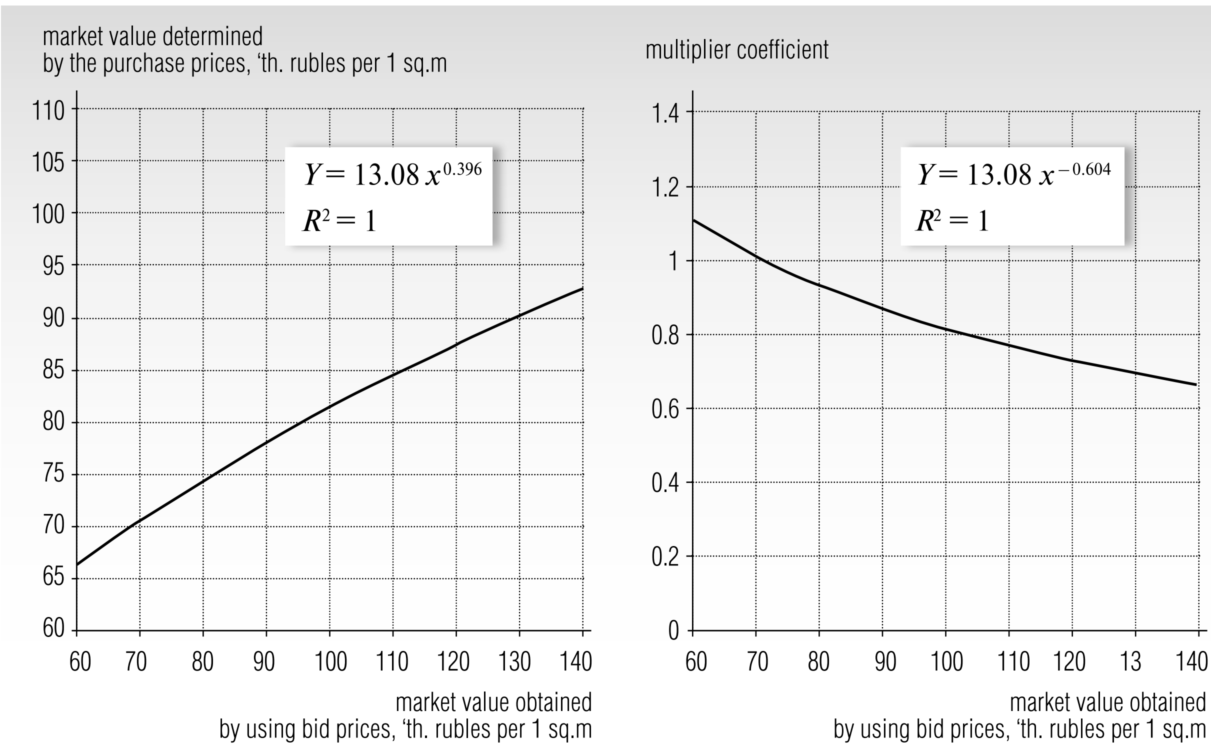

Fig. 5. The dependence of valuation of market value determined by the purchase prices from the valuation of market value obtained by using bid prices (left) and the dependence of the multiplier coefficient from the bid price (right)

Conclusion

The creation of the institution of cadastral value in the Russian Federation opens new, previously inaccessible methods of analyzing the real estate market and evaluating market figures, in particular, determination of the trading discount.

Cadastral value determined as the market value on the date of appraisal prior to the date of its regular review is an important analytical tool, and not just the basis for calculating the property tax.

The proposed approach (juxtaposition of bid prices and purchase prices with the cadastral value determined by methods of mass evaluation as market value on the date of appraisal) can be applied also when adjusting cadastral value. We examine this question in detail in the recently published article [18], which we were prompted to write following the lively discussion in the community of appraisers over the quality of determination of cadastral value. A detailed analysis of the methodological problems of determining cadastral value can be found in the article of Professor Korostelev [19]. ■

References

1. Aitchinson J., Brown J.A.C. (1963) *The lognormal distribution with special references to its uses in economics.* Cambridge, Cambridge University Press.
2. Ciurlia P., Gheno A. (2009) A model for pricing real estate derivatives with stochastic interest rates. *Mathematical and Computer Modeling*, no. 50, pp. 233–247.
3. Ohnishi T., Mizuno T., Shimizu C., Watanabe T. (2011) *On the evolution of the house price distribution.* Columbia Business School. Center of Japanese Economy and Business. Working Paper Series, no. 296.
4. Nikulina T.V., Ponomareva O.A., Popentzova S.V. (2015) Logarifmicheski normal'noe raspredelenie tsen na zhilye ob'ekty nedvizhimosti elitnogo klassa i ekonom-klassa [Lognormal distribution of prices for residential real estate of elite class and economy class]. *Week of science of SPbSPU. Materials of the scientific forum with international participation* (ed. O.V. Kalinina, S.V. Shirokova). Saint Petersburg, pp. 435–437 (in Russian).
5. Rusakov O.V., Laskin M.B., Jaksumbaeva O.I. (2016) Stokhasticheskaya model' tsenoobrazovaniya na rynke nedvizhimosti: formirovanie lognormal'noy general'noy sovokupnosti [Stochastic pricing model for the real estate market. Formation of lognormal general population]. *Economics, Statistics and Informatics. UMO Bulletin*, no. 5, pp. 116–127 (in Russian).
6. Rusakov O.V., Laskin M.B., Jaksumbaeva O.I. (2016) Pricing in the real estate market as a stochastic limit. Log Normal approximation. *International Journal of Mathematical Models and Methods in Applied Sciences*, vol. 10, pp. 229–236.
7. The Federal Law of 29.07.1998 No. 135-FZ (edition of 29.07.2017) *"Ob otsenochnoy deyatel'nosti v Rossiyskoy Federatsii"* ["About valuation activity in the Russian Federation"]. Available at: http://www.consultant.ru/document/cons_doc_LAW_19586/ (accessed 01 June 2018) (in Russian).
8. The European Group of Valuers' Associations (2012) *European valuation standards 2012.* Available at: https://ru.scribd.com/document/254068172/Valuation-Standards-2012-for-European-Countries (accessed 01 June 2018).
9. International Valuation Standard Council (2013) *International valuation standards 2013. Framework and requirements.* Available at: http://www.valuersinstitute.com.au/docs/professional_practice/International%20Valuation%20Standards%202013.pdf (accessed 01 June 2018).
10. Royal Institution of Chartered Surveyors (2014) *RICS valuation professional standard 2014.* Available at: http://hoadd.noordhoff.nl/sites/7605/_assets/7605d24.pdf (accessed 01 June 2018).
11. The Appraisal Foundation (2018) *Uniform standards of professional appraisal practice. 2018–2019 edition.* Available at: http://www.uspap.org/ (accessed 01 June 2018).
12. Laskin M.B. (2017) Korrektirovka rynochnoy stoimosti po tsenoobrazuyushchemu faktoru "ploshchad' ob'ekta" [Market value adjustment for the pricing factor "square"]. *Property Relations in the Russian Federation,* no. 8 (191), pp. 86–99 (in Russian).

13. Laskin M.B. (2018) Statisticheskiy analiz rezul'tatov torgov. Interval startovoy tseny [Statistical analysis of the auction results. Starting price interval]. *Property Relations in the Russian Federation*, no. 1 (196), pp. 19–29 (in Russian).

14. Rusakov O.V., Laskin M.B., Jaksumbaeva O.I. (2016) Opredelenie koeffitsienta kapitalizatsii po statisticheskim dannym [Determination of capitalization coefficient using statistical data]. *Economics, Statistics and Informatics. UMO Bulletin*, no. 1, pp. 14–22 (in Russian).

15. Rusakov O.V., Laskin M.B., Jaksumbaeva O.I. (2016) Otsenka pokazateley rynka nedvizhimosti po statisticheskim dannym na osnove mnogomernogo logarifmicheski normal'nogo zakona [Estimation of the real estate market indexes according to statistical data and based on multidimensional log-normal distribution]. *HSE Economic Journal*, vol. 20, no. 2, pp. 268–284 (in Russian).

16. St. Petersburg Committee of Property Relations (2015) *On approval of the results of the cadastral value of the premises with the area less than 3,000 sq m on the territory of St. Petersburg.* The order of the St. Petersburg Committee of Property Relations, 27 August 2015, no. 59-p.

17. Government of St. Petersburg (2012) *The report on determination of cadastral value of real estate objects (except for the parcels of land) located in the territory of St. Petersburg*, vol. 2, section 2.3.

18. Laskin M.B., Gadasina L.V. (2018) Kak opredelit' kadastrovuyu stoimost' [How to determine cadastral value]. *Property Relations in the Russian Federation*, no. 3, pp. 42–53 (in Russian).

19. Korostelev S.P. (2017) *O "edinoy federal'noy metodologii" kadastrovoy otsenki nedvizhimosti i zemli* [On the "uniform Federal methodology" of cadastral valuation of real estate and land]. Available at: http://www.labrate.ru/articles/2017-1_korostelev.pdf (accessed 28 May 2017) (in Russian).

The impact of the Internet of Things technologies on economy

Mariya S. Tokareva
Expert, Department of Digital Economy Research
Institute for Statistical Studies and Economics of Knowledge
National Research University Higher School of Economics
Address: 20, Myasnitskaya Street, Moscow, 101000, Russia
E-mail: mtokareva@hse.ru

Konstantin O. Vishnevskiy
Head of Department of Digital Economy Research
Institute for Statistical Studies and Economics of Knowledge
National Research University Higher School of Economics
Address: 20, Myasnitskaya Street, Moscow, 101000, Russia
E-mail: kvishnevsky@hse.ru

Ludmila P. Chikhun
Associate Professor, Faculty of Economics
Lomonosov Moscow State University
Address: 1, build. 46, Leninskie Gory, Moscow, 119991, Russia
E-mail: chihun@mail.ru

Abstract

This study presents the analysis of one of the major contemporary transformational forces – the Internet of Things (IoT), which significantly influences the future development of all spheres of life. The purpose of the research is to identify the potential economic effects of IoT implementation in different markets. To achieve this goal, the following tasks are consistently solved in the study: identification and classification of the main IoT applications markets; detection, assessment and analysis of the economic effects of the IoT in the selected segments within the proposed classification; formation of future directions of IoT development. Based on the combination of such methodological approaches as technology life cycle and technology adoption life cycle, perspectives of the IoT development are set out. The technology life cycle is viewed through the prism of the methodology of the research company Gartner (the Gartner Hype Cycle for Emerging Technologies), based on establishing a consensus among a wide range of assessments of leading experts in the field of information and communication technologies. Comparison of the two methods and expert assessments allows us to conclude that, according to the methodology of technology adoption life cycle, the Internet of Things is of interest only for a group of "early adopters."

Key words: Internet of Things; information and telecommunication technologies; technology life cycle; technology adoption life cycle; digitalization; smart things; globalization.

Citation: Tokareva M.S., Vishnevskiy K.O., Chikhun L.P. (2018) The impact of the Internet of Things technologies on economy. *Business Informatics*, no. 3 (45), pp. 62–78.
DOI: 10.17323/1998-0663.2018.3.62.78.

Introduction

On the threshold to the fourth industrial revolution, we see a fusion of technology and the blurring of distinctions between physical, digital and biological spheres. Similar to previous revolutions, the explosive growth of innovative technologies should give impetus to the development of various spheres of public life. The Internet of Things (IoT) has become a large business and technology area with ever-growing market potential. The IoT technology market is expected to grow from $176 Billion in 2016 to $639.74 Billion by 2022, at a CAGR (Compound Annual Growth Rate) of 25.1% between 2017 and 2022 [1].

Largely due to the spread and penetration of information and communication technologies (ICT), the speed of technological innovation in all spheres of life is increasing exponentially today. In the coming 10—15 years, the next radical leap that will be associated with the introduction of the Internet of Things is expected [2]. As various household devices (for example, robotic complexes on digital production) equipped with sensors and connected to the Internet begin "to communicate" among themselves without human intervention, the main sectors of the economy are transformed (industrial production, transport, medicine, agricultural industry, etc.), and the model of interaction between people and machines will be reversed. The economic effect of the Internet of Things is estimated to grow from $2.7 trln to $6.2 trln annually till 2025 [3]. By 2030, the contribution of the Industrial Internet of Things to the global economy can be about $14 trln [4]. The number of connected devices in OECD countries will increase from 1 bln in 2016 to 14 bln by 2022 [5].

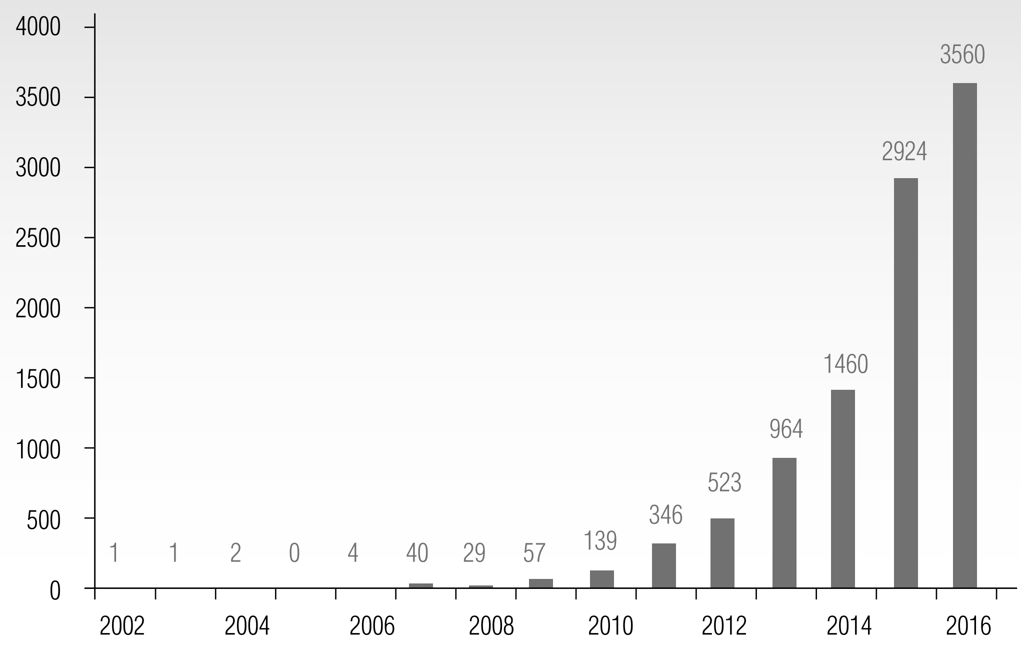

Fig. 1. Publications devoted to the IoT (based on Web on Science)

1. Review of literature

Nowadays there is a growing interest in the Internet of Things technologies as evidenced by the growth of scientific publications in this field of research (*Figure 1*).

The first reference to the Internet of Things can be attributed to Schoenberger and Upbin [6], who described the impact of the Internet of Things technologies on the economy and society. Interest in this issue is not surprising, because in 2002 Amazon launched the first cloud web service, marking the start of cloud computing and providing impetus to the development of IoT technologies. Further, in 2004 Gershenfeld et al. (Massachusetts Institute of Technology) published an article devoted to a wide range of aspects of IoT development [7]. The authors mention the benefits of the objects' ability to connect to data networks: optimization of the configuration of lights and switches at home, reduction of the cost and complexity of building construction, assisting with home health care. Nevertheless, they stressed the importance of solving the problem of standardization [7–9]. It is worth noting that nowadays this remains one of the key challenges for IoT development which can be overcome by the interaction of public authorities, as well as by the formation of consortiums of the Internet of Things.

The evolution of IoT requires the development of a bridging point (i.e., gateway) between the sensor infrastructure network and the Internet. Rahmani et al. [10] consider that it is necessary to develop the concept of Fog Computing in IoT systems by formation of a Geo-distributed intermediary layer of intelligence between sensor nodes and the Cloud. Also Cavalcante et al. [8] claim that Cloud Computing has been advocated as a promising approach to tackle some of the existing challenges in IoT which prevent exploiting its full potential and promoting tangible benefits to society, environment, economy and individual citizens.

One more challenge associated with IoT is noted by Bello et al. [11]. They explore the challenges of integration and interoperability of device to device communication technologies by focusing on network layer functions such as addressing, routing, mobility, security and resource optimization. The limitations of the current TCP/IP architecture of the Internet of Things environment require the development of a low-power wide-area network (LPWAN) which coincides with the fact that the IoT technologies are on the peak of inflated expectations according the Gartner' technologies' life cycle [12].

The development of IoT faces a great challenge – ensuring cybersecurity. Thereby Weber and Studer [13] offer applicable international regulations of a legal cybersecurity environment in the Internet of Things context and alternative approaches to addressing the security issues in IoT. Development of the Internet of Things in the future is directly related to standardization of this field. Park et al. [9] notice that collaboration between standards' bodies to provide interoperability is a key to the success of IoT, because nowadays IoT devices use different standards and are deployed in a large numbers in different domains.

Despite all the difficulties of IoT development, the Internet of Things is an emerging technology demanded in all sectors of the economy that can drive production to a new level, improve the quality of life and more. So the purpose of the research is to identify the economic effects of IoT implementation in different sectors. To achieve this goal, the following research questions are consistently solved in the study: to identify and classify the main IoT applications markets; to assess the economic effects of using the Internet of Things in the selected segments. Based on the combination of such methodological approaches as technology life cycle and technologies adoption life cycle, a scenario of IoT development is set out.

The working paper structure consists of the following parts: a description of the methodology of the study, assessments of the impact of the IoT on the economy, a forecast of the development of the Internet of Things on the basis of theories in the field of technology life cycles, the conclusion and the list of references.

2. Methodology

As the Internet of Things is an emerging technology, it seems relevant to define the segments of the IoT market in the absence of a generally accepted classification. Therefore, the first stage of the study is devoted to the identification of the IoT segments and perfection of existing classifications. The author's classification is used to reveal the impact of the Internet of Things in each segment.

Prediction of the development path of IoT technologies is based on the identification of the level of readiness of IoT technologies through the construction of a timeline. The timeline is a common tool for scientific and technological forecasting and is widely used in Foresight and roadmaps. Similar studies were conducted by Utterback and Abernathy [14], who proposed the dynamic model of process and product innovation presenting the close relationship between innovation (the end product), the innovation process and the company's strategy. The dynamic model of innovation is a combination of the product life cycle model, the process life cycle model and various competitive strategies. The authors identified three phases. Each differently affects individual companies, the market and the resources required to create innovation. Gartner's methodology was chosen as a theoretical basis because of the availability of expert assessments of the emerging technologies' development. Based on a joint analysis of Gartner's Hype Cycle for Emerging Technologies and the technology adoption life cycle, prospects for IoT development were brought into focus.

Gartner's methodology for visualization, the "emerging technology hype cycle," is widely used both in corporate and public environments for more precise decisions concerning the future development of technologies, speed of their adoption, mutual influence and competition of various technologies. The vertical axis on the graph illustrating the Gartner's hype cycle of emerging technologies denotes evolving expectations. Until 2009, the ordinate axis had a different name – visibility, but later that was replaced by "expectations" as a more suitable characteristic. The latter approach includes the attitude of both potential and actual users to the innovative technology, as well as investment decisions regarding its further development [15].

Gartner's methodology includes five key stages: scenarios, surveys, pattern recognition, "stalking horse" and verification. At the first stage it is necessary to establish research objectives using a survey method exploring tough situations or outcomes that other firms might not want to use. By analyzing data from multiple sources such as analytical reports, scientific articles, monographs, statistic collections and others, emerging patterns on markets are identified [16]. Further, there is a need to modify assumptions and update the scenarios. At the fourth stage, a "stalking horse" is a position released into the analyst community to be examined and debated from various viewpoints. Finally, findings are validated against multiple internal and external sources.

Concerning the influence of technology on the economy, the technology adoption life cycle has presented even greater interest [17]. The logic of the life cycle of the new technology adoption is based on the fact that technology is perceived by any community in stages in keeping with the psychological and social portraits of various segments of this community. According to this approach, the population is divided into five groups up to the degree of new technology adoption: innovators, early adopters, early majority, late majority, laggards.

By combining the two approaches outlined above, it is possible to determine which group of consumers should be oriented by the companies developing marketing strategy for promotion of the products from the IoT market. The method was described by several scholars [18; 19].

This approach allowed the authors to determine the portrait of the consumer of products and services created with the Internet of Things technologies. Thus, using this approach, the authors were able to identify a group of consumers for which decisions of the Internet of Things are of the greatest interest at present and a potential group of consumers in the short-term period. The combination of theoretical approaches allowed the authors to transfer the results of the extensive expert poll of Gartner to the field of another theoretical concept (technology adoption life cycle, for which such large-scale expert surveys are not conducted) and to analyze the phenomenon of the Internet of Things through the given prism.

3. Internet of Things market

Based on the analysis of different practices (Gartner, IDC, McKinsey, Forrester, IoT Analytics, Rostelecom, etc.), an author's classification of the IoT market was made. The criterion of this classification is the actors of the market: business, consumers, government. The consumer segment of the Internet of Things includes product solutions used by consumers in everyday life based on the technologies of IoT. The Industrial Internet of Things covers various branches of the economy where the use of IoT technologies can radically change business processes, increase productivity and operational efficiency. Using IoT in the government segment is aimed at improving public safety including re-offense prediction. The information and communication technologies are distinguished by their interdisciplinary nature and relevance in all the afore-mentioned segments, including cloud computing, big data, IoT platform technologies and sensor technologies. The author's classification of the IoT market reveals potential prospective segments of the implementation of IoT technologies attracting a wider range of consumers.

Given the ubiquity of digitalization and automation, when the organizational units of the enterprise are integrated into a single network many production and administration processes can be managed online using cloud computing and IoT systems. According to Verizon, the Internet of Things market in the business segment (Industrial Internet of Things, IIoT) will reach 5.4 billion devices by 2020 [20; 21].

The **Industrial Internet of Things** transforms business processes, increases the effectiveness of the whole value chain, ultimately leading to the formation of new business models and markets. The main directions of IoT technologies application are effective supply management, freight and asset management, machine diagnostics, machine telemetry, inventory control, industrial automation control, real-time monitoring of equipment, etc. The greatest effect from the spread of the Internet of Things in Russia will be observed in the following sectors: health (74% of Russian respondents – representatives of companies from various sectors of the economy); manufacturing (56%); the energy sector (32%) according to the survey conducted by Accenture [4]. *Table 1* presents estimates of economic effects from the implementation of the IIoT in various sectors of the

Table 1.

Economic effects of the Industrial Internet of Things

B2B Industrial IoT (IIoT)			
Manufacturing	**Logistics**	**Energy**	**Mining**
• 2,5–5% saving in operational costs, including maintenance and input efficiencies; • 30% improvement of time–to–market; • 40% reduction in planning and equipment costs; • 10–25% increase in manufacturing productivity; • 40% reduction in factory equipment maintenance costs; • 50% reduction in equipment downtime; • 5% reduction in capital equipment investment costs; • 20–50% reduction in factory inventory carrying costs	• 30% reduction in labor costs; • 30% reduction in the time of order processing; • 12% reduction in the repair expenses; • 30% reduction in overall maintenance costs; • 70% reduction in the downtime	• 2–4% reduction in demand peaks in the grid	• 5–10% saving in op–erating costs from produc–tivity gains; • $3.7 trln economy in global mining operating costs up to 2025
Agriculture	**Transport**	**Construc-tion**	**Finance**
• 25% reduction in cost of vehicle damage from collision avoidance and increased security; • $1.2–1.3 trln in agricultural production (wheat, maize, rice, soybeans, barley); • 20–40% adoption of advanced irrigation systems and precision farming; • 10–20% increase in yields from precision application of fertilizer and irrigation	• $28 mln in value over 10 years with smart buses; • $53 mln in value over 10 years with smart parking; • 20–25% savings in fuel costs; • 79% reduction of crashes; • 40% reduction in vehicle wait time; • 26% reduction in travel time	• 20% reduction in total life–cycle costs of a project	• 30% reduction in building management system deployment and operating expense in financial organizations

economy: manufacturing, logistics, energy, mining, agriculture, transport, construction, finance (based on [22–26]).

Using IIoT technologies provides the opportunity to bring manufacturing to a qualitatively new level. As a result, it becomes possible to integrate flexible production systems as well as digital control systems into manufacturing, all of which makes it easier to control production, accelerate it and significantly increase its flexibility. In this regard, it is necessary to introduce promising technologies in a timely manner and, due to the high cost, the effective use of funds becomes a key issue. It is also important to pay enough attention to the competence of employees and promote their skills [27].

Maximum automation of routine manufacturing operations (advanced manufacturing) through the implementation of IIoT technologies will help to link the activities of various divisions (including suppliers, logistics, marketing and even involve consumers in the production process by obtaining product information from them in real time) through the creation of a single information field. That will help to increase the flexibility of production, the speed of the product's release to market, as well as the optimization cost. Development of IoT technologies in Russia is based on strong competencies of the specialists in the field of algorithms and software. This provides good opportunities for equal participation of Russian organizations in international consortiums and projects related with Internet of Things [28].

There are plentiful examples of the implementation of the Industrial Internet of Things at the nation level, for example, Smart metering in Germany [29] and also at the enterprise level – the ERP system for creating a smart factory in B&R Industrial Automation in Austria and etc. [30].

The **Internet of Things in a consumer segment** contributes to the improvement of the quality of life due to automation of many routine household operations, release of free time, and granting new opportunities previously inaccessible to consumers (*Table 2*, based on [25; 26; 31–43]). Self-driving cars are one prospective application of IoT technologies. According to the forecasts of Hitachi, by 2020 90% of new cars in Europe will be connected to the Internet (versus 30% in 2016). Complementary technologies for self-driving cars are the intellectual transport systems created for these vehicles. Since 2011, in Moscow there is an intelligent transportation system under construction, including management of the technical means of regulating and organizing traffic, controlling parking, photo-video fixation of violations of traffic regulations, monitoring of traffic flow parameters, control and regulation of traffic organization [44].

The technologies of the "smart" home are aimed at providing the most comfortable accommodation, safety and resource-savings. By 2022, the "smart" home market is expected to reach $121.7 bln, CAGR – 14.1% [45]. The "smart" home market includes lighting control, security & access control, HVAC control, entertainment & controls, home healthcare, and the smart kitchen.

The technologies of the Internet of Things are in demand in the sphere of utilities, and their use facilitates the reduction of costs to the general population for electric power, water and heating. Such saving of resources has a positive effect on the environment due to reduced consumption of natural resources including non-renewables. By 2050, a 55% increase in demand for water resources is projected. Furthermore the demand for energy resources will increase for 37% by 2040 [46]. In this regard, the need for resource-saving and energy efficient technologies becomes more acute every year.

Table 2.

Economic effects of IoT in the consumer segment

<table>
<tr><th colspan="3">Consumer segment of Internet of Things (B2C)</th></tr>
<tr><th>Smart city</th><th>Wearable smart devices, healthcare</th><th>Smart home</th></tr>
<tr><td>• 10–20% reduction in average travel time through control of traffic and congestion;
• 10–20% reduction in water consumption and leaks with smart meters and demand control;
• 10–20% reduction in cost of waste handling;
• 60% energy savings by moving to smart street lighting;
• 20% reduction in water losses;
• 30% reduction in street crime</td><td>• 10–20% cost reduction in chronic disease treatment through remote health monitoring;
• 80–100% reduction in drug counterfeiting;
• 0.5–1 hour time saved per day by nurses;
• the value of improved health of chronic disease patients through remote monitoring could be as much as $1.1 trln per year in 2025;
• 25% reduction in costs from clinical and operations inefficiencies (or about $100 bln per year);
• 50% reduction in mean time required to repair connected devices;
• $2,000 reduction in service costs for each problem resolved remotely;
• 20% fewer technician dispatches worldwide</td><td>• 30–50% reduction in water usage;
• $50–90 reduction per water bill;
• 88% reduction in power consumption;
• 96% of construction waste is recycled;
• 30–50% reduction per energy bill (extra $198 per year or $16.5 per month)</td></tr>
<tr><th>Trade and repair</th><th>Sports, leisure, entertainment</th><th>Education</th></tr>
<tr><td>• 1.5–7% increased sales;
• 99.5% inventory accuracy;
• 50% reduction in out–of–stocks;</td><td>• 25% reduction in operational expenses</td><td>• 60% of energy can be saved from IoT connected smart devices in schools/colleges</td></tr>
</table>

One more benefit of applying IoT technologies in the utilities sector is time saving for paperwork (filing receipts, collection of information about monthly expenditure of water and electric power). Release from repeating routine transactions is possible due to the establishment of “smart” counters adjusted to the water, electric power and heat meters via

IoT connections capable of sending notification about the payment to smartphones [44].

The additional possibility of power consumption control relates to "smart" home appliances (washing and dishwashers, dryers, etc.) using built-in sensors and applications of the Internet of Things that can turn them on automatically during the period of lowest daily tariffs. In addition, they can notify the owner about undesirable starts in case of the maximum rate applied.

Numerous "smart" home systems' applications require particular attention and research as they have significant gaps and shortcomings. These issues are related to security systems, safety, energy consumption, marketing and device interoperability. Furthermore, the use of IoT gadgets and IT infrastructure deserves special mention. It is necessary to conduct more research on practical use of new-generation built-in sensors that are equipped with electrical devices as noted in the works of Alaa et al. [47].

IoT technologies in the urban environment are aimed at the development of innovative solutions for infrastructure, energy saving, construction and organization of public space. Thus, connection of the installed sensors on vehicles and roads makes possible real time traffic control. In addition, the development of similar city infrastructure increases the probability of detection of the stolen cars. In the future, traffic information about vehicles will be generated by intellectual transport systems. It will be the means of urban transport system's operation and will lead to the reduction of difficulties with traffic on roads, to decreases in CO_2 emissions, etc.

The strategic direction called "Smart Urban Mobility" was developed within the framework of Singapore's Smart Nation initiative. Implementation of the projects is aimed at improving the system of public transport by introducing digital technologies and the use of data generated by various devices [48]. In Singapore, the development of the Internet of Things is stimulated by providing software development companies with access to a system of sensors deployed throughout the city. Thus, data, IoT platforms and APIs are becoming open and enabling creation of innovative products and services within the ecosystem of the Internet of Things.

The technology of the Internet of Things in healthcare can exert an enormous effect not just on the quality of medical services but also on the capabilities of modern medicine. Use of sensors can detect disease at early stages. The diagnosis of sharp respiratory infections by means of "pocket" biosensors will lead to the reduction of costly treatment and reduced losses from disability. IoT technologies in healthcare aim at forming steady demand for a new quality of life. Medical biotechnology and personalized medicine services may significantly increase the life expectancy of the population, as well as achieve significant progress in treating cancer, cardiovascular and infectious diseases [46].

The Internet of Things opens new opportunities in retail and finance activities. Monitoring information on movements of goods using "smart" packaging with an RFID tag (Radio Frequency IDentification) will let producers develop more efficient marketing strategies (develop various loyalty programs, individual offers for clients, increase the average sales value).

The Internet of Things also impacts on the scope of penetration of financial services due to the use of contactless payments made possible by the technology of wireless high-frequency communication with small radius of action (to 10 cm) – NFC (Near Field Communication) supported by smartphones. Moreover, in banking it is also possible to deepen the customization and personalization of services based on analysis of data consolidated in the development of the IoT technologies.

Media and the entertainment are changing as a result of the development of IoT technologies. Improvement in the quality of telecom services opens the possibility to use media with high resolution at any time and in any circumstances. Virtual reality accessible through mobile devices and applications affects almost all social activities. Virtual reality is becoming not just the popular way of social interaction but also one of the forms of entertainment. Virtual and augmented reality technologies assume immersion of the user inside content directly working with his or her emotions. The viewer will watch a movie and becomes a kind of participant in the story. There is the possibility to act inside the movie, select the view or hero on behalf of the narrator or in the manner he or she wants to act. Virtual and augmented reality technologies are most widely applied in the consumer segment (54%), while the share of government and business segment accounts for only 46% [49].

New directions of IoT concept development in the Social Internet of Things (SIoT) represent an integration of IoT and social networks. The synthesis of two these phenomena provides users with an opportunity to access direct "smart" things (without an intermediary). Thus, the IoT become social. Moreover, in SIoT smart devices are able to establish social links and communicate with each other, just as people do in a social network. This is made possible by the Search Social Internet of Things model (the SSIoT model), where the search engine serves as the medium between two elements of integration. Correspondingly, the SSIoT influences significantly SIoT capabilities and specifications while at the same time it contributes to the speed-up of malicious code propagation which is emphasized in the study of Fu et al. [50].

The new paradigm of the Social Internet of Things (SIoT) has great potential for IoT applications and various networking services which increase the efficiency, speed of development and spread of this innovation. The types and the characteristics of the social relations which can be installed by technologies in the SIoT have been identified: OOR relationships (Ownership Object Relationship), SOR relationships (Social Object Relationship) and C-WOR relationships (Co-work object relationship). In accordance with different types of relationships, the Social Internet of things will have a different architecture, mobility and scalability [51].

IoT smart objects provide people with a perfect opportunity to transfer the reality into the virtual dimension in online protected mode and create augmented reality. In the IoT, smart devices are able not only to be part of people's social network but they are capable of creating their own "social" network. This state of affairs can lead to development of new complex effective services and apps for people. Correspondingly, in the framework of IoT a new concept of the Web of Things has appeared. This paradigm relates to the ability to automatically post various information and data on social network sites. The problem lies in the management of numerous devices and web application architecture [52].

One of the most effective systems for providing public services electronically in the **government segment of Internet of Things** is the Estonian X-Road platform through which citizens can perform almost any operation [53]. Cybersecurity of the X-Road is based on blockchain technology (KSI blockchain), which allows the user to track any interaction with the system [54]. Such convergence of the Internet of Things and blockchain technologies will improve the security of IoT-devices and also increase the scalability of the IoT-system and the speed of interaction of devices included in it. Nowadays there are plans to attract small and medium-sized enterprises to use the platform as well as to expand its borders by involving B2B and B2C segments [55].

For faster penetration of the Internet of Things concept, it is necessary to strengthen the activity in various fields. First of all, this means improving the ICT infrastructure as well as strengthening public support for the introduction of new technologies. For example, in the UK IoTUK Boost launched a program aimed at deployment of IoT networks that can be used by small and medium-sized enterprises [56].A successful example of the deployment of the Internet of Things infrastructure is Estonia, where the international IoT operator Sigfox has created a national IoT network, access to which is planned to be provided for the entire population of Estonia.

Due to digitalization, companies need to revise the formats of partnerships, expand access to information and provide a new level of decentralization of the working environment. For more effective interaction between counterparts, new types of organizational structures will be required at the enterprise level, including those based on platform technologies. Moreover, it will be necessary to develop and implement new standards for interoperability and cybersecurity [21; 57].

Despite effective, dynamic technological progress and outcomes, an IoT vision has not been generated yet. It will be quite difficult for realization, since it is necessary to develop the technological infrastructure and social Internet culture. But due to this, it is indispensable to create and establish criteria and unique standards for the IoT. Consequently, there are not only technical challenges but political, economic, human and social problems. Within the escalating number of users, the variety of mechanisms and usage scenarios will increase as well. It is necessary to consider the evolution of the Internet of Things in connection with the interoperability of human activities and electronic devices. The current state of affairs involves solving a number of problems and conducting social, economic and technological research [58].

The development of the Internet of Things is inextricably linked with the need to increase consumer confidence in new, innovative products and services, based on confidentiality and security of data generated from various devices [59].

A significant initiative in this direction is the activity of the ITU. In 2012, the International Telecommunication Union issued its Recommendation ITU-T Y.2060 with the object in mind to lead standardization in the area of IoT. It explains the definition and the ambit of the IoT, determinates key characteristics, features and high-level direction and objectives for IoT. A reference model IoT is described. Ecosystem and business models are presented as well. Thus, the implementation of these standards has led to the unification of requirements, interoperability of smart objects, the expansion of the range and an increase in the scope of technology and solving the problem of heterogeneity in computing systems [60].

In order to solve IoT problems, various international organizations are developing recommendations and standards in this field. Indeed, the CSA Mobile Working Group issued an IoT Initiative – Security Guidance. The manual in this document was created in such a way as to provide useful guidance in different sectors. It was achieved by researching data and architectures in different industries and screening the security controls that would able to support any one industrial sector. In order to avoid duplication, efforts were made to promote and harmonize working groups of the various sectors [59].

The common world practice is to unite efforts aimed at stimulating innovation by forming global, voluntary standards developed by regulatory bodies or industry consortiums necessary to ensure interoperability and the growth of the IoT ecosystem. Among the key characteristics of consortiums, we can single out the following: prototyping and increasing the scale of production as the main line of business; the network principle of interaction; self-repayable (after the end of budget financing) and an indefinite type of activity based on the cooperation of businesses (including small and medium-sized enterprises), scientific and educational organizations. In March 2014, the Industrial Internet Consortium was founded, with Intel, IBM, Cisco, GE and AT&T among its participants. The main objective of the IIC is to unite organizations to promote and accelerate the growth of the Industrial Internet of Things by identifying, collecting and sharing best practices[1].

There is also a consortium in the field in Russia – the Russian Association of the Industrial Internet[2]. The majority of its members is represented by the information technology sector (55.6%), but representatives of such sectors as telecommunication services, aerospace, electrical equipment and commercial services are also available (11.1% each). This indicates importance of Internet of Things technologies for various sectors of the economy.

Analysis of the members of the Industrial Internet Consortium visually demonstrates those sectors of the world economy that are already showing an interest in solutions based on IoT technologies. In addition to IT companies, scientific organizations, research and consulting companies also are participating in the development and standardization of the Internet of Things. Unlike the foreign consortiums under analysis, Russian scientific organizations are not participants of the RAII. In this regard, as measures to develop the Internet of Things in Russia, authors would like to recommend that scientific organizations take a proactive stance and get involved in the work on the development and standardization of the Internet of Things within the framework of the Russian consortium, as well as companies from other sectors such as energy, finance and machinery.

4. Prospects for IoT development

Regarding the supply side of the IoT market, there is the very important question to be answered: how the suppliers on the IoT market can improve their performance? The correlation between Gartner's 2015 Hype Cycle for Emerging Technologies and the technology adoption life cycle was investigated to determine the target group of population with most likely potential as IoT market consumers.

Comparison of the two afore-mentioned methods [18; 19] shows that IoT solutions are attractive to early adopters who understand the advantages of the new technology, evaluate them and link potential benefits of the IoT with their interests. Their decision to buy something will be based on the degree of this commonality. Because early adopters do not rely on authorities for purchasing decisions, they use their own intuition and foresight. So they a play key role in opening any new digital technologies market segment. According to the technology adoption life cycle, there is a challenging goal

[1] https://www.iiconsortium.org/members.htm2016, No. 1758
[2] http://iotunion.ru/en/uchastniki?limitstart=0

to overcome the “chasm” stage and attract the “early majority” group interested in IoT technologies.

Conclusion

The rate and efficiency of IoT development within the coming 15 years will to a large extent depend on further distribution of fixed and mobile broadband communication and reduction in the cost of connected devices [5; 28]. Moreover, development of the IoT potential is possible thanks to the development and use of processing technologies and big data analysis generated from various sensors and other devices. Skills for data analysis are the key asset for the future. In this regard, there is the possibility to increase social inequality: some part of the population will take advantage of the new opportunities opened by the Internet of Things improving their quality of life while for others these benefits are unavailable [46].

For further research, it seems interesting to compile socio-psychological portraits of consumers for each segment of the IoT market directly affecting the economy. Identification of these groups of consumers will be important for marketing departments of companies, for the Internet of Things providers, as well as for government.

Concluding the research, it is worth mentioning that some people do not notice how the technologies of the IoT are becoming part of their daily lives. When buying a new apartment, consumers often already are acquiring IoT solutions aimed at saving electricity, water and heating. It is necessary to inform the population about the new opportunities opening up for Internet of Things users in order to stimulate the demand for IoT solutions.

It should be stressed that the Internet of Things is one of the key technologies of the fourth industrial revolution, similar to the case of previous industrial revolutions and so it has an impact on the labor market. The latter aspect could be a prospective field for further research. On the one hand, there is growing demand for highly qualified specialists particularly in the field of cloud computing and big data. But on the other hand, the spread of the Internet of Things leads to a reduction in the demand for low skilled labor. In order to settle the social problems which arise, it is necessary to carry out programs to improve the skills of employees both at the governmental level and at the micro (enterprise) level. Therefore, it is necessary to expand the intake of university applicants in ICT technologies, to open new programs within the bachelor’s and master’s programs aimed at training highly qualified specialists in the field of the Internet of Things, cloud computing, big data, distributed computing, etc. For example, in the USA more than half a million students with a specialty in STEM (science, technology, engineering and mathematics) graduate from educational institutions each year and in China – 4.7 mln students [61]. In addition, at the micro level companies need to improve the skills of employees in order to remain competitive on modern dynamically changing markets. ■

References

1. Markets and Markets (2017) *Internet of Things technology market by node component (processor, sensor, connectivity IC, memory device, and logic device), network infrastructure, software solution, platform, service, end-use application, and geography – Global forecast to 2022.* Available at: https://www.marketsandmarkets.com/Market-Reports/iot-application-technology-market-258239167.html (accessed 30 January 2018).

2. ISSEK HSE (2016) Industrial'nyy Internet veshchey: Revolyutsionnye izmeneniya v promyshlennosti [Industrial Internet of Things: Revolutionary changes in industry]. *Global Technological Trends*, no. 10. Available at: https://issek.hse.ru/data/2016/11/15/1110389295/%D0%98%D0%9A%D0%A2_10_2016.pdf (accessed 08 December 2017) (in Russian).

3. Manyika J., Lund S., Bughin J., Woetzel J., Stamenov K., Dhingra D. (2016) *Digital globalization: The new era of global flows. Report, McKinsey Global Institute, February 2016.* Available at: www.mckinsey.com/business-functions/digital-mckinsey/our-insights/digital-globalization-the-new-era-of-global-flows (accessed 25 December 2017).

4. Accenture (2015) *Winning the Industrial Internet of Things. How to accelerate the journey to productivity and growth*. Available at: https://www.accenture.com/us-en/~/media/Accenture/Conversion-Assets/DotCom/Documents/Global/PDF/Digital_1/Accenture-Industrial-Internet-of-Things-Positioning-Paper-Report-2015.pdf (accessed 11 January 2018).

5. OECD (2015) *OECD digital economy outlook 2015.* Paris: OECD Publishing. Available at: http://dx.doi.org/10.1787/9789264232440-en (accessed 10 December 2017).

6. Schoenberger C., Upbin B. (2002) The internet of things. *Forbes*, 18 March 2002. Available at: http://www.fbcoverup.com/docs/library/2002-03-18-The-internet-of-things-by-Chana-Schoenberger-and-Bruce-Upbin-FORBES-Mar-18-2002.pdf (accessed 10 December 2017).

7. Gershenfeld N., Krikorian R., Cohen D. (2004) The Internet of Things. The principles that gave rise to the Internet are now leading to a new kind of network of everyday devices, an "Internet-0". *Scientific American*, no. 10, pp. 76–81. Available at: http://fab.cba.mit.edu/classes/S62.12/docs/Cohen_Internet.pdf (accessed 25 December 2017).

8. Cavalcante E., Pereira J., Alves M.P., Maia P., Moura R., Batista T., Delicato F.C., Pires P.F. (2016) On the interplay of Internet of Things and Cloud Computing: A systematic mapping study. *Computer Communications*, vol. 89–90, pp. 17–33.

9. Park H., Kim H., Joo H., Song J. (2016) Recent advancements in the Internet-of-Things related standards: A oneM2M perspective. *ICT Express*, vol. 2, no. 3, pp. 126–129.

10. Rahmani A.M., Gia T.N., Negash B., Anzanpour A., Azimi I., Jiang M., Liljeberg P. (2017) Exploiting smart e-Health gateways at the edge of healthcare Internet-of-Things: A fog computing approach. *Future Generation Computer Systems*, vol. 78, no. 2, pp. 641–658.

11. Bello O., Zeadally S., Badra M. (2017) Network layer inter-operation of Device-to-Device communication technologies in Internet of Things (IoT). *Ad Hoc Networks*, no. 57, pp. 52–62.

12. Gartner (2015) *Gartner's 2015 Hype Cycle for emerging technologies identifies the computing innovations that organizations should monitor*. Available at: http://www.gartner.com/newsroom/id/3114217 (accessed 10 December 2017).

13. Weber R.H., Studer E. (2016) Cybersecurity in the Internet of Things: Legal aspects. *Computer Law & Security Review*, vol. 32, no. 5, pp. 715–728.

14. Utterback J.M., Abernathy W.J. (1975) A dynamic model of process and product innovation. *Omega*, vol. 3, no. 6, pp. 639–656.

15. Fenn J., Raskino M. (2011) *Understanding Gartner's Hype Cycles, 2011.* Available at: https://www.gartner.com/doc/1748018/understanding-gartners-hype-cycles- (accessed 09 December 2017).

16. Gartner (2017) *Technology-related insights for your critical business decisions. Gartner Research Methodologies*. Available at: http://www.gartner.com/imagesrv/research/methodologies/gartner_methodologies.pdf (accessed 10 December 2017).

17. Beal G.M., Bohlen J.M. (1957) The diffusion process. *Iowa State College. Special report No. 18.* Available at: https://lib.dr.iastate.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=1015&context=specialreports (accessed 19 December 2017).

18. Banerjee U. (2012) *How experts explain technology adoption cycle*. WordPress.com. Technology Trend Analysis. Available at: https://setandbma.wordpress.com/2012/05/28/technology-adoption-shift/ (accessed 20 December 2017).

19. Rao S.V., Cohen M.G. (2014) Staying ahead of the curve. *Cardiovascular Revascularization Medicine,* vol. 15, no. 4, pp. 193–194.
20. Cisco (2016) *Internet of Things*. Available at: http://www.cisco.com/c/dam/en/us/products/collateral/se/internet-of-things/at-a-glance-c45-731471.pdf (accessed 18 December 2017).
21. Verizon (2016) *State of the market. Internet of Things 2016.* Available at: https://www.verizon.com/about/sites/default/files/state-of-the-internet-of-things-market-report-2016.pdf (accessed 25 December 2017).
22. Accenture (2016) *Smart production. Finding a way forward: how manufacturers can make the most of the Industrial Internet of Things.* Available at: https://www.accenture.com/t20160119T041002Z__w__/us-en/_acnmedia/PDF-5/Accenture-804893-Smart-Production-POV-Final.pdf (accessed 25 December 2017).
23. CIMdata (2002) *Disruptive technologies: The Benefits of Digital Manufacturing*. Available at: https://www.cimdata.com/newsletter/2009/25/01/25.01.01.htm (accessed 24 October 2017).
24. Ezell S. (2016) *IoT and smart manufacturing*. Available at: http://www2.itif.org/2016-ezell-iot-smart-manufacturing.pdf?_ga=1.261819661.1089858538.1464487061 (accessed 17 December 2017).
25. Manyika J., Chui M., Bughin J., Dobbs R., Bisson P., Marrs A. (2013) *Disruptive technologies: Advances that will transform life, business, and the global economy*. Available at: https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/disruptive-technologies (accessed 25 December 2017).
26. Robinson A. (2016) *How the IoT platform fixes problems in today's logistics industry*. Available at: https://cerasis.com/2016/05/02/iot-platform/ (accessed 09 December 2017).
27. 451 Research (2015) *2016 Trends in the Internet of Things*. Available at: https://451research.com/search?keyword=%222016+Trends+in+the+Internet+of+things%22 (accessed 09 December 2017).
28. PwC (2016) *The Industrial Internet of Things. Why it demands not only new technology — but also a new operational blueprint for your business*. Available at: https://www.pwc.com/hu/en/kiadvanyok/assets/pdf/industrial-internet-of-things.pdf (accessed 25 December 2017).
29. Metering & Smart Energy International (2017) *Reforming the German energy sector*. Available at: https://www.metering.com/magazine-article/reforming-german-energy-sector/ (accessed 24 April 2018).
30. Industrial Internet Consortium (2016) *Smart Factory: Industry 4.0 @ B&R*. Available at: http://www.iiconsortium.org/pdf/BR_Case_Study.pdf (accessed 24 April 2018).
31. My Alarm Center (2012) *Home automation safety, security, savings*. Available at: http://www.loversiq.com/daut/as/m/v/vashos-green-roofed-and-energy-efficient-rd-house-blends-into-a-home-automation-infographic-efficiency-design-sustainable_green-building-home-designs_home-decor_linon-home-decor-decorators-collection-_797x5636.jpg (accessed 10 December 2017).
32. Michelsen C. (2012) *Smart home test shows 88% reduction in household power consumption*. Available at: https://cleantechnica.com/2012/08/07/smart-home-test-shows-88-reduction-in-household-power-consumption/ (accessed 09 December 2017).
33. General Electric Company (2012) *The many faces of IoT (Internet of Things) in healthcare*. Available at: https://image.slidesharecdn.com/themanyfacesofiotinhealthcareslideshare-141112104650-conversion-gate02/95/the-many-faces-of-iot-internet-of-things-in-healthcare-11-638.jpg?cb=1416371687 (accessed 17 December 2017).
34. Aoun C. (2013) *The smart city cornerstone: Urban efficiency*. Available at: https://www.ogcio.gov.hk/en/news/consultations/d21_submission_2013/doc/079_SchneiderElectric_(Annex).pdf (accessed 22.12.2017).
35. Morgan Stanley (2014) *The Internet of Things is now, connecting the real economy*. Available at: https://www.i-scoop.eu/internet-of-things-guide/internet-things-retail-industry// (accessed 09.12.2017).
36. Honda (2015) *Honda smart home US*. Available at: http://thumbnails-visually.netdna-ssl.com/honda-smart-home-us_53333797b44ff_w1500.jpg (accessed 10 December 2017).
37. Chamberlin B. (2016) *Healthcare Internet of Things: 18 trends to watch in 2016*. Available at: https://ibmcai.com/2016/03/01/healthcare-internet-of-things-18-trends-to-watch-in-2016/ (accessed 25 December 2017).

38. Hewlett Packard (2016) *Smart cities and the Internet of Things*. Available at: https://h20195.www2.hpe.com/V2/getpdf.aspx/4AA6-5129ENW.pdf (accessed 10 December 2017).

39. Sellebraten M. (2016) *Smart cities could cut costs by 30% with open IoT standards*. Available at: https://enterpriseiotinsights.com/20160505/smart-cities/smart-cities-cut-costs-tag29 (accessed 25 December 2017).

40. Thompson J. (2016) *Baseball stadium uses IoT to reduce operational costs by 25 percent*. Available at: https://www.engineering.com/IOT/ArticleID/13095/Baseball-Stadium-Uses-IoT-to-Reduce-Operational-Costs-by-25-Percent.aspx (accessed 25 December 2017).

41. Varian Medical Systems (2016) *Varian Medical Systems case study*. Available at: https://www.ptc.com/en/case-studies/varian-medical-systems (accessed 27 December 2017).

42. Eyebridge (2017) *Internet of Things – Taking digital world by storm*. Available at: http://www.eyebridge.in/blog/internet-of-things-taking-digital-world-by-storm/ (accessed 18 December 2017).

43. HomeSelfe (2017) *How to save water using smart home technology*. Available at: http://www.homeselfe.com/save-water-using-smart-home-technology (accessed 09 December 2017).

44. Rostelecom (2015) Mir, v kotorom vse vklyucheno [A world where everything is included]. *Rostelecom PRO*, no. 3 (18), pp. 38–41 (in Russian).

45. Markets and Markets (2016) *Smart home market by product (lighting control (timer, daylight sensor, occupancy sensor), security & access control, HVAC, entertainment, home healthcare and smart kitchen), software & service (behavioral, proactive), and geography – Global forecast to 2022*. Available at: http://www.marketsandmarkets.com/Market-Reports/smart-homes-and-assisted-living-advanced-technologie-and-global-market-121.html (accessed 10 December 2017).

46. OECD (2016) *OECD science, technology and innovation outlook 2016*. Paris: OECD Publishing. Available at: https://www.oecd-ilibrary.org/science-and-technology/oecd-science-technology-and-innovation-outlook-2016_sti_in_outlook-2016-en (accessed 10 December 2017).

47. Alaa M., Zaidan A.A., Zaidan B.B., Talal M., Kiah M.L.M. (2017) A review of smart home applications based on Internet of Things. *Journal of Network and Computer Applications*, no. 97, pp. 48–65.

48. Smart Nation and Digital Government Office (2018) *Smart nation. Strategic national projects*. Available at: https://www.smartnation.sg/about/Smart-Nation (accessed 15 May 2018).

49. Goldman Sachs (2016) Virtual & augmented reality: Understanding the race for the next computing platform. *Equity Research,* 13 January 2016. Available at: http://www.goldmansachs.com/our-thinking/pages/technology-driving-innovation-folder/virtual-and-augmented-reality/report.pdf (accessed 25 December 2017).

50. Fu C., Peng C., Liu X.-Y., Yang L.T., Yang J., Han L. (2018) Search engine: The social relationship driving power of Internet of Things // *Future Generation Computer Systems (in press)*. Available at: https://www.sciencedirect.com/science/article/pii/S0167739X17307884?via%3Dihub (accessed 26 April 2018).

51. Atzori L., Iera A., Morabito G., Nitti M. (2012) The Social Internet of Things (SIoT) – When social networks meet the Internet of Things: Concept, architecture and network characterization. *Computer Networks*, vol. 56, no. 16, pp. 3594–3608.

52. Atzori L., Carboni D., Iera A. (2014) Smart things in the social loop: Paradigms, technologies, and potentials. *Ad Hoc Networks*, no. 18, pp. 121–132.

53. e-Estonia (2018) *X-Road*. Available at: https://e-estonia.com/solutions/interoperability-services/x-road/ (accessed 23 April 2018).

54. e-Estonia (2018) *KSI blockchain*. Available at: https://e-estonia.com/solutions/security-and-safety/ksi-blockchain/ (accessed 23 April 2018).

55. The World Bank Group (2017) *Internet of Things. The new government to business platform. A review of opportunities, practices, and challenges*. Available at: http://documents.worldbank.org/curated/en/610081509689089303/pdf/120876-REVISED-WP-PUBLIC-Internet-of-Things-Report.pdf (accessed 24 April 2018).

56. IoTUK (2016) *IoTUK Boost 2017. Accelerating Internet of Things innovation with access to LPWAN testbeds across the UK with IoTUK boost*. Available at: https://iotuk.org.uk/iotuk-boost-2017/ (accessed 24 April 2018).

57. Manyika J., Chui M., Bisson P., Woetzel J., Dobbs R., Bughin J., Aharon D. (2015) *Unlocking the potential of the Internet of Things*. Available at: http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/the-internet-of-things-the-value-of-digitizing-the-physical-world (accessed 25 December 2017).

58. Shin D. (2014) A socio-technical framework for Internet-of-Things design: A human-centered design for the Internet of Things. *Telematics and Informatics,* vol. 31, no. 4, pp. 519–531.

59. Cloud Security Alliance (2015) *Security guidance for early adopters of the Internet of Things (IoT)*. Available at: https://downloads.cloudsecurityalliance.org/whitepapers/Security_Guidance_for_Early_Adopters_of_the_Internet_of_Things.pdf (accessed 24 April 2018).

60. ITU (2012) *Recommendation ITU-T Y.2060. Overview of the Internet of Things*. Available at: https://www.itu.int/rec/T-REC-Y.2060-201206-I (accessed 26 April 2018).

61. RIA (2017) *Kvantovaya gonka vooruzheniy: Kitay poshel na obgon SShA* [Quantum arms race. China to overtake USA]. Available at: https://ria.ru/analytics/20170511/1494076882.html (accessed 25 December 2017) (in Russian).

IEEE CBI

21st IEEE International Conference on Business Informatics (IEEE CBI 2019)
July 15–17, 2019, Moscow, Russia

http://www.cbi2019.moscow/
http://www.cbi-series.org

The IEEE Conference Series on Business Informatics is the leading international forum for state-of-the-art research in Business Informatics. The 21st IEEE CBI 2019, held in huge, old and interesting city Moscow, calls for submissions in the multidisciplinary field of Business Informatics, and welcomes a multitude of theoretical and practical perspectives and mind-sets on today's challenges of the digital transformation.

Call for Papers

The IEEE CBI series encourages a broad understanding of Business Informatics research, and intends to further its many different facets, theoretical foundations and experiential body of knowledge. In doing so, the CBI series has proven to be a fertile ground for research with high impact, and a hub for multidisciplinary research with contributions from Management Science, Organization Science, Economics, Information Systems, Computer Science, and Informatics.

In line with this, the CBI series aims to provide a forum for researchers and practitioners from various fields and backgrounds that contribute to the construction, use and maintenance of information systems and their organizational use contexts. A dedicated goal of the CBI series of conferences is to bring together researchers from different fields and disciplines to stimulate discussion, synergy and collaboration. Accordingly, CBI conferences use a format that enables in depth discussions among researchers during the conference. The benefits of such a crossdisciplinary conception are contrasted by a challenge: Authors who submit a paper take the risk to be assessed by standards that are different from those they are used to in their own research communities. The conference organisation accounts for this challenge by involving renowned track chairs and programme committee members from different fields and disciplines.

List of Tracks

- Enterprise Modelling, Engineering and Architecture
- Business Process Management
- Information Systems Engineering
- Artificial Intelligence for Business
- Information Management
- Business Analytics and Business Data Engineering
- Industry 4.0 (Industry Applications)
- Business Innovations and Digital Transformation
- Data-Driven Business Applications
- General Topics

Submission

https://edas.info/newPaper.php?c=25431

Manuscripts must be in English and are restricted to **10 pages in IEEE 2-coloumn template (A4)**.

In the submission form, select the conference track that best fits with your manuscript's topic, or choose "General Topics" if not covered by the tracks.

Review process is **double-blind**; your names and affiliations must not be explicitly listed in the manuscript. However, you should keep all cited references in the submission's bibliography intact.

Further information for submission:

https://cbi2019.moscow/
Accepted conference content will be submitted for inclusion **into IEEE Xplore as well as other Abstracting and Indexing (A&I) databases.**
Selected papers will be invited to submit an extended version to a special issue of the **EMISA Journal** (subject to separate review).

Important Dates

Full paper submission
→ March 1, 2019

Notification of paper acceptance
→ May 10, 2019

Camera-ready paper submission
→ June 1, 2019

AUTHORS GUIDELINES

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

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

MANUSCRIPT REQUIREMENTS

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

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

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

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

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

✦ Full name of each author

✦ Position, rank, academic degree of each author

✦ Affiliation of each author, at the time the research was completed

✦ Full postal address of each affiliation (incl. postcode / ZIP)

✦ E-mail address of each author.

ABSTRACT are presented in native language and English.

✦ The abstract should be between 200 and 300 words.

✦ The abstract should be informative (no general words), original, relevant (reflects your paper's key content and research findings); structured (follows the logics of results' presentation in the paper)

✦ The recommended structure: purpose (mandatory), design / methodology / approach (mandatory), findings (mandatory), research limitations / implications (if applicable), practical implications (if applicable), originality / value (mandatory).

✦ It is appropriate to describe the research methods/ methodology if they are original or of interest for this particular research. For papers concerned with experimental work the data sources and data procession technique should be described.

✦ The results should be described as precisely and informatively as possible. Include your key theoretical and experimental results, factual information, revealed interconnections and patterns. Give special priority in the abstract to new results and long-term impact data, important discoveries and verified findings that contradict previous theories as well as data that you think have practical value.

✦ Conclusions may be associated with recommendations, estimates, suggestions, hypotheses described in the paper.

✦ Information contained in the title should not be duplicated in the abstract. Authors should try to avoid unnecessary introductory phrases (e.g. «the author of the paper considers...»).

✦ Authors should use the language typical of research and technical documents to compile your abstract and avoid complex grammatical constructions.

✦ The text of the abstract should include key words of the paper.

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

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

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

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

The publication is free of charge.