

INFORMATION **SCIENCE**
&
COMPUTING

International Book Series

Number 10

**Intelligent Support
of
Decision Making**

Supplement to
International Journal "Information Technologies and Knowledge" Volume 3 / 2009

ITHEA
SOFIA, 2009

Krassimir Markov, Alexey Voloshyn, Krassimira Ivanova, Iliia Mitov (ed.)

Intelligent Support of Decision Making

International Book Series "INFORMATION SCIENCE & COMPUTING", Number 10

Supplement to the International Journal "INFORMATION TECHNOLOGIES & KNOWLEDGE" Volume 3 / 2009

Institute of Information Theories and Applications FOI ITHEA

Sofia, Bulgaria, 2009

This issue contains a collection of papers in the field of research and development of information technologies to be implemented for automatization of the support of the decision making as well as in the field of business intelligence.

Papers are peer reviewed and are selected from the several International Conferences, which were part of the Joint International Events of Informatics "ITA 2009" – summer session, Varna, Bulgaria.

International Book Series "INFORMATION SCIENCE & COMPUTING", Number 9
Supplement to the International Journal "INFORMATION TECHNOLOGIES & KNOWLEDGE" Volume 3, 2009

Edited by **Institute of Information Theories and Applications FOI ITHEA**, Bulgaria,
in collaboration with

- **V.M.Glushkov Institute of Cybernetics of NAS**, Ukraine,
- **Institute of Mathematics and Informatics, BAS**, Bulgaria,
- **Institute of Information Technologies, BAS**, Bulgaria.

Publisher: Institute of Information Theories and Applications FOI ITHEA, Sofia, 1000, P.O.B. 775, Bulgaria.
Издател: Институт по информационни теории и приложения ФОИ ИТЕА, София, 1000, п.к. 775, България
www.ithea.org, www.foibg.com, e-mail: info@foibg.com

General Sponsor: **Consortium FOI Bulgaria** (www.foibg.com).

Printed in Bulgaria

Copyright © 2009 All rights reserved

© 2009 Institute of Information Theories and Applications FOI ITHEA - Publisher

© 2009 Krassimir Markov, Alexey Voloshyn, Krassimira Ivanova, Iliia Mitov – Editors

© 2009 For all authors in the issue.

ISSN 1313-0455 (printed)

ISSN 1313-048X (online)

ISSN 1313-0501 (CD/DVD)

PREFACE

The scope of the International Book Series "Information Science and Computing" (**IBS ISC**) covers the area of Informatics and Computer Science. It is aimed to support growing collaboration between scientists from all over the world. IBS ISC is official publisher of the works of the members of the ITHEA International Scientific Society.

The official languages of the IBS ISC are English and Russian.

IBS ISC welcomes scientific papers and books connected with any information theory or its application.

IBS ISC rules for preparing the manuscripts are compulsory.

The rules for the papers and books for IBS ISC are given on www.foibg.com/ibsisc .

The camera-ready copies of the papers should be received by ITHEA Submission System <http://ita.ithea.org> .

The camera-ready copies of the books should be received by e-mail: info@foibg.com .

Responsibility for papers and books published in IBS ISC belongs to authors.

This issue contains a collection of papers in the field of research and development of information technologies to be implemented for automatization of the support of the decision making as well as in the field of business intelligence.

Papers are peer reviewed and are selected from the several International Conferences, which were part of the Joint International Events of Informatics "ITA 2009" – summer session, Varna, Bulgaria.

ITA 2009 has been organized by

ITHEA International Scientific Society

in collaboration with:

- Institute of Information Theories and Applications FOI ITHEA
- Taras Shevchenko National University of Kiev (Ukraine)
- International Journal "Information Theories and Applications"
- International Journal "Information Technologies and Knowledge"
- Association of Developers and Users of Intelligent Systems (Ukraine)
- Association for Development of the Information Society (Bulgaria)
- V.M.Glushkov Institute of Cybernetics of National Academy of Sciences of Ukraine
- Institute of Mathematics and Informatics, BAS (Bulgaria)
- Institute of Information Technologies, BAS (Bulgaria)
- Institute of Mathematics of SD RAN (Russia)
- Dorodnicyn Computing Centre of the Russian Academy of Sciences
- Universidad Politecnica de Madrid (Spain)
- BenGurion University (Israel)
- Rzeszow University of Technology (Poland)
- University of Calgary (Canada)
- University of Hasselt (Belgium)
- Kharkiv National University of Radio Electronics (Ukraine)
- Astrakhan State Technical University (Russia)
- Varna Free University "Chernorizets Hrabar" (Bulgaria)
- National Laboratory of Computer Virology, BAS (Bulgaria)
- Uzhgorod National University (Ukraine)

The main ITA 2009 events were:

KDS	XVth International Conference "Knowledge - Dialogue – Solution"
i.Tech	Seventh International Conference "Information Research and Applications"
MeL	Fourth International Conference "Modern (e-) Learning"
INFOS	Second International Conference "Intelligent Information and Engineering Systems"
CFDM	International Conference "Classification, Forecasting, Data Mining"
GIT	Seventh International Workshop on General Information Theory
ISSI	Third International Summer School on Informatics

More information about ITA 2009 International Conferences is given at the www.ithea.org .

The great success of ITHEA International Journals, International Book Series and International Conferences belongs to the whole of the ITHEA International Scientific Society.

We express our thanks to all authors, editors and collaborators who had developed and supported the International Book Series "Information Science and Computing".

General Sponsor of IBS ISC is the **Consortium FOI Bulgaria** (www.foibg.com).

Sofia, June 2009

Kr. Markov, Al. Voloshyn, Kr. Ivanova, I. Mitov

TABLE OF CONTENTS

<i>Preface</i>	3
<i>Table of Contents</i>	5
<i>Index of Authors</i>	7
<u>Decision Making</u>	
Application of Temporal Reasoning and Case-based Reasoning in Intelligent Decision Support Systems <i>Alexander Ereemeev, Ivan Kurilenko, Pavel Varshavskiy</i>	9
A General Purpose Software System for Linear Optimization and Decision Support "Optima-Plus" Vassil Vassilev , <i>Mariyana Vassileva, Krassimira Genova, Boris Staykov</i>	17
Interactive Methods for Group Decision Making <i>Filip Andonov</i>	25
Some Aspects of Comparative Analysis of Banks Functioning <i>Alexander Kuzemin, Vyacheslav Lyashenko</i>	31
Application of Technology Foresight Strategy in Innovation Activity <i>Nataliya Pankratova</i>	39
Functional Approach to Emergency Risk Minimization <i>Oleksandr Kuzomin, Illya Klymov</i>	47
Эволюционная парадигма как универсальный интегрирующий элемент методологии прогнозирования <i>Виталий Снитюк</i>	52
Построение интегральных показателей оценки результативности научных проектов <i>Алексей Петровский, Григорий Ройзензон, Игорь Тихонов</i>	59
Анализ малых возмущений линейных экономико-математических моделей <i>Алексей Волошин, Всеволод Богаенко, Владимир Кудин</i>	67
Исследование задачи оптимизации инвестиционного портфеля в нечетких условиях <i>Юрий Зайченко, Малихех Есфандиярфард</i>	74
Сравнение полиинтервальных оценок в методе ОИО <i>Михаил Стернин, Геннадий Шепелёв</i>	83
Идентификация оценки противопожарного состояния объекта для формирования базы знаний ИАС «Безопасность» <i>Александр Джулай, Артем Быченко, Андрей Левченко</i>	89
Мультиагентный <i>H</i> -метод в комбинаторной оптимизации <i>Леонид Гуляницкий, Денис Гобов</i>	97
Разработка и исследование алгоритмов решения задачи прогнозирования третичной структуры протеина <i>Леонид Гуляницкий, Виталина Рудык</i>	104

Business Intelligence

Systemological Business Modelling for Information Systems Development <i>Mikhail Bondarenko, Nikolay Slipchenko, Kateryna Solovyova, Olexiy Ukrayinets, Larysa Vinogradova.....</i>	113
Quality and Productivity Analysis in a Very Small Enterprise Using TSPi <i>Henry Caballero, Jose Calvo-Manzano, Gonzalo Cuevas, Tomás San Feliu.....</i>	121
Stochastic Models and Methods of Operative-dispatch Control over the Gas-transport System of Ukraine <i>Andriy Tevyashev</i>	129
Интеллектуальный анализ данных в системах управления процессом фрахтования судов <i>Сергей Шехунов, Антон Моторкин.....</i>	138
Использование технологии Semantic Web для интеллектуального управления в динамических распределенных системах <i>Анатолий Гладун, Юлия Рогушина</i>	143
Агрегирование прямых и двойственных балансовых моделей «затраты-выпуск» <i>Игорь Ляшенко, Андрей Онищенко, Игорь Онищенко</i>	154
Влияние информационной устойчивости на информационные инновации в фирмах <i>Александр Фролов</i>	161
Автоматизация управления бизнес-процессами на основе языка GBPR <i>Антон Рычков</i>	169
Механизмы распределения квот на выбросы по Киотскому протоколу <i>Алексей Волошин, Ирина Горицына</i>	175
<i>In memoriam: Prof. Dr Vassil Vassilev.....</i>	183

INDEX OF AUTHORS

Filip Andonov	25	Всеволод Богаенко	67
Mikhail Bondarenko	113	Артем Быченко	89
Henry Caballero	121	Алексей Волошин	67, 175
Jose Calvo-Manzano	121	Анатолий Гладун	143
Gonzalo Cuevas	121	Денис Гобов	97
Alexander Eremeev	9	Ирина Горицына	175
Krassimira Genova	17	Леонид Гуляницкий	97, 104
Illya Klymov	47	Александр Джулай	89
Ivan Kurilenko	9	Малихех Есфандиярфард	74
Alexander Kuzemin	31, 47	Юрий Зайченко	74
Vyacheslav Lyashenko	31	Владимир Кудин	67
Nataliya Pankratova	39	Андрей Левченко	89
Tomás San Feliu	121	Игорь Ляшенко	154
Nikolay Slipchenko	113	Антон Моторкин	138
Kateryna Solovyova	113	Андрей Онищенко	154
Boris Staykov	17	Игорь Онищенко	154
Andriy Tevyashev	129	Алексей Петровский	59
Olexiy Ukrayinets	113	Юлия Рогушина	143
Pavel Varshavskiy	9	Григорий Ройзензон	59
Vassil Vassilev	17	Виталина Рудык	104
Mariyana Vassileva	17	Антон Рычков	169
Larysa Vinogradova	113	Виталий Снитюк	52
		Михаил Стернин	83
		Игорь Тихонов	59
		Александр Фролов	161
		Геннадий Шепелёв	83
		Сергей Шехунов	138

Decision Making

APPLICATION OF TEMPORAL REASONING AND CASE-BASED REASONING IN INTELLIGENT DECISION SUPPORT SYSTEMS

Alexander Ereemeev, Ivan Kurilenko, Pavel Varshavskiy

Abstract: *In this paper the problem of the application of temporal reasoning and case-based reasoning (CBR) in intelligent decision support systems (IDSS) is considered. To reduce static algorithm complexity of temporal reasoning some methods are investigated. The dynamic algorithm for qualitative temporal reasoning on the base of point algebra is presented. CBR method for a solution of problems of real-time diagnostics and forecasting in IDSS is described. This paper demonstrates how the temporal reasoning system and CBR system can be used in IDSS of the car access control. This work was supported by RFBR and grant of President of Russia (MK-6009.2008.9).*

Keywords: *Intelligent decision support systems, temporal reasoning, analogous and case-based reasoning.*

ACM Classification Keywords: *H.4.2 [Information systems applications]: Types of systems – Decision support; I.2.4 [Artificial intelligence]: Knowledge Representation Formalisms and Methods – Temporal logic; I.2.5 [AI]: Programming Languages and Software – Expert system tools and techniques; I.2.6 [AI]: Learning – Analogies.*

Conference: *The paper is selected from XVth International Conference "Knowledge-Dialogue-Solution" KDS 2009, Varna, Bulgaria, June-July 2009*

Introduction

Temporal reasoning and "common sense" reasoning, in particular, CBR can be used in various applications of artificial intelligence (AI) and for solving various problems [1], e.g., for diagnostics and forecasting or for machine learning. The problem of presentation of time and temporal interconnections in AI systems (AIS) and especially in IDSS is very actual nowadays. A lot of basic notions, such as "alteration", "cause", "consequence/effect" and relations among them can be described by time notions [2]. However, the problem of creation of formal systems of presentation and operating by temporal information became really actual after the appearance and development of AIS, oriented towards open and dynamic problem domains. Typical representatives of these systems are real time IDSS (RT IDSS), designed for monitoring and control of complex objects and processes in rather strict time conditions and for different types of uncertainties of obtained information [1]. AI experts model CBR and temporal reasoning by computers in order to develop more flexible models of search for solutions and learning. The generalized structure of RT IDSS is given in Fig. 1.

Formally, a RT IDSS can be defined by the tuple $SS = \langle M, R(M), F(M), F(SS) \rangle$, where

- $M = \{M_1, \dots, M_n\}$ is the set of formal or logic-linguistic models, implementing defined intelligent functions;
- $R(M)$ is the function for selection of the necessary model in a current situation;
- $F(M) = \{F(M_1), \dots, F(M_n)\}$ is the set of modification functions of models M_1, \dots, M_n ;
- $F(SS)$ is the function for modification of SS system, i.e. its base components $M, R(M), F(M)$.

The main problems, solved by RT IDSS, are: *diagnostics and monitoring* – revealing of problem situations; *decision searching* – searching an optimal or admissible sequence of actions allowing to achieve the desired goal or to solve the problem situations; *forecasting* – assessing the recommended actions related to the goal achievement and the sanction to solve the problem situation.

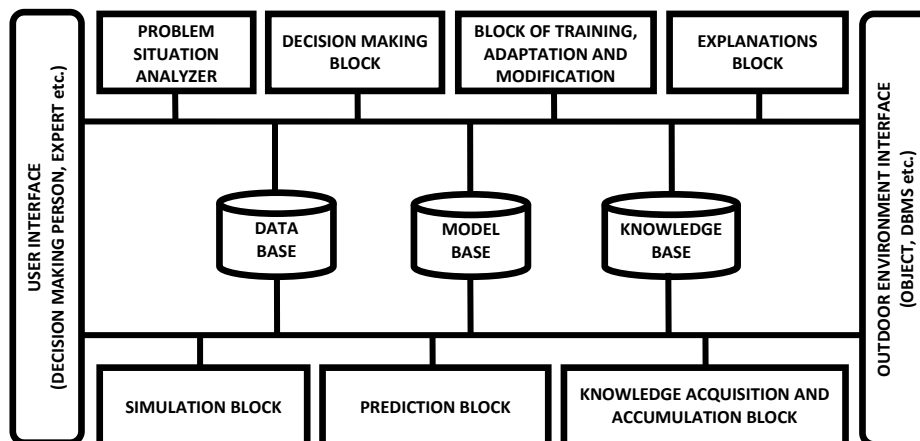


Fig. 1 Base RT IDSS structure

The methods of temporal reasoning and CBR may be applied in different blocks of RT IDSS. The necessity to present data and knowledge, changing in the course of time (indications of sensors, values of control parameters, information from decision making person (DMP), etc.) appears when solving many problems. RT IDSS must solve diagnostics, monitoring, decision searching and forecasting problems uninterruptedly in real time scale in order to help DMP to find efficient control effects in different operation modes of controlled objects, especially in abnormal modes. Using information about time while solving these problems permits to decrease search parameters greatly, which naturally positively effects reactivity of the whole system [3]. Thus, the use of the respective methods in RT IDSS broadens the possibilities of RT IDSS and increases the efficiency of making decisions in various problem (abnormal) situations.

Temporal reasoning

To implement the mechanism of temporal reasoning (MTR), it is necessary to formalize the notion of time and to provide the possibility to present and to discuss temporary aspects of knowledge. Modern approaches to presentation of time and temporal dependences in software systems can be divided into two large classes – based on *modeling of time changes* and on *explicit time modeling*. In the approaches which use changes modeling, the basic features are entities (actions), transforming one state of the system to the other. These states are regarded as momentary pictures of the world, which don't have any time duration. Time itself is regarded implicitly, via modeling of the system changes within time.

Approaches, based on changes modeling, have constraints when presenting difficult time dependences (events, which have duration, continuance of processes, competing or time overlapping events, causal relations etc). In literature different ways of constraints elimination can be found, however in the most cases they are reduced to introduction of an explicit time model. Explicit time modeling provides the possibility to make “flexible” formalized languages, which help to do reasoning on the basis of expressions, truth values of which are timed to the definite moment or time interval, and they can change in the course of time. Time is presented explicitly, taking into consideration its properties. Different *temporal logics* and *models* are included to this class. Time can be presented both syntactically (via explicit temporary structures) and semantically (modal logics are typical representatives of this approach). In the case of explicit time modeling the following specific *tasks of temporary reasoning* are raised: The support of information coordination about time – check of coordination of knowledge base when adding the new information. In the case of inconsistency of temporal constraints it is necessary to

localize plenty of expressions, responsible for this inconsistency; Answers to queries, dealing with time aspects of knowledge. These queries can be divided into finding a simple fact, true in the definite of moment and definition when a set of expressions is true at the same moment of time.

In this paper we consider the models, based on the presentation of information about time as constraints (dependences) between time primitives [4]. In temporal logics using the concept of constraint satisfaction, information about time is presented as dependences between temporal primitives (moments, intervals or their combinations). Dependences between primitives are interpreted as constraints to real time of their appearance. Main aim of the MTR is a generations of conclusions on the basis of sets of temporary constraints, i.e. new constraints for consistent input sets. Usually sets of temporal primitives and relations among them are presented as the Temporal Constraint Satisfaction Problem (TCSP), which is detailing of a more general Constraint Satisfaction Problem (CSP), what permits to use CSP methods to solve the TCSP.

The TCSP is specified by the following way $Z = (V, D, BTR, C)$ [5]:

- 1) $V = \{V_1, V_2, \dots, V_m\}$ – a finite set of temporal variables;
- 2) D – a value domain of temporal variables;
- 3) $BTR = \{R_1, R_2, \dots, R_n\}$ – a finite set of binary basic temporary constraints, and constraints entering there are mutually exclusive, but their total join is the universal constraint U ;
- 4) $C = \{C_{ij} \mid C_{ij} = \{r_1, \dots, r_k\}, k > 0, r_1, \dots, r_k \in BTR, i < m, j < m\}$, a finite set of temporary constraints (STC), where C_{ij} – is the constraint for temporary variables V_i and V_j . Each constraint C_{ij} from set C is interpreted as $(V_i R_1 V_j) \vee \dots \vee (V_i R_k V_j)$. In case C_{ij} consists only of one clause, it is called an *exact restriction*.

It is necessary to find such a STC $C^* = \{C_{ij}^* \mid C_{ij}^* = \{r_j\}, r_j \in C_{ij}\}$, so that exact constraints, entering it, do not conflict with each other.

Elements of V set can be interpreted as moments, time intervals or duration. The range of values of D variables, corresponding to moments of time and duration, represent a set of numbers, and for interval variables – a set of ordered value pairs.

In case in C set only exact constraints are entered, the TCSP is called *exact TCSP*, and the TCSP itself is regarded as a check of constraints consistency in C set [5].

The task of defining r constraint valid for variables V_i and V_j with the constraint $C_{ij} = \{r_1, \dots, r_k\}$, having more than one clause, is called the task of defining the inexact constraint C_{ij} . The constraint itself is called a *single mark of the constraint* C_{ij} , and the set $C^E = \{C_{ij}^E \mid C_{ij}^E = \{r_j\}, r_j \in C_{ij}\}$ – a *single mark of the TCSP*. In this case solution of the TCSP is its *consistent single mark*. The TCSP is *consistent* only when it has at least one solution.

The constraint C_{ij} is *feasible* for variables V_i and V_j if and only if at least there is one solution of the TCSP, where C_{ij} is a constraint between these variables. The minimal constraint C_{ij}^{min} is the set, consisting only of feasible constraints for V_i and V_j . The TCSP is called *minimal*, if all its constraints are minimal. It is known that for any TCSP it is always possible to find the equivalent minimal one or to show inconsistency of constraints [3].

Main operations for temporary constraints are the following:

- 1) complement (\neg): $\neg L_{ij} = U \setminus L_{ij}$;
- 2) inversion (\sim): $\sim(r_1, \dots, r_k) = (\sim r_1, \dots, \sim r_k)$;
- 3) intersection (\cap): $S \cap T = \{r : r \in S, r \in T\}$ – the set, consisting of equal constraints in S and T ;
- 4) composition (\bullet): $T \bullet S = (t_1, \dots, t_k) \bullet (s_1, \dots, s_q) = ((t_1 \bullet s_1), (t_1 \bullet s_2), \dots, (t_k \bullet s_q))$ – disjunction of individual compositions of all elementary constraints in T and S .

Main subtasks of the task of TCSP are the following: *checking consistency* - checking whether there is a way of building the C^* set from C set (which is the solution of the TCSP); *finding consistent scenarios* - definition of all possible C^* sets; *search of minimal presentation* - transition of initial TCSP to the minimal; *search for feasible constraints* for the given pair of temporary variables.

Allen proposed the interval algebra of temporal constraints wherein time intervals are taken as primitive. Reasoning within this algebra is NP-complete [6]. The point algebra is based on time points as primitives [4]. The major advantage of the point algebra is ability to construct the reasoning algorithms with polynomial complexity [7]. Further we will consider the point algebra as the base to construct TRS.

To solve the TCSP, a set of temporal variables and constraints are transformed into a graph, weighted by temporal information [7]. A *temporally labeled graph (TL-graph)* is a graph $G=(V,E)$ with at least one vertex ($V \neq \emptyset$) and a set of labeled edges, where each edge (v, l, w) connects a pair of distinct vertices v,w . The edges are either directed and labeled \leq or $<$, or undirected and labeled \neq . Every vertex of a TL-graph has at least one name attached to it. If the vertex has more than one name, than these names are alternative for the same time point. The name sets of any two vertices are required to be disjoint. A path on a TL-graph is called \leq -path if each edge on the patch has a label $<$ or \leq . A \leq -path is called $<$ -path if at least one of the edges has label $<$. Given a TL-graph G an interpretation of G is a triple $\langle T, I, R \rangle$ where T is a totally ordered set (with ordering $<$), I is a function $I: P \rightarrow T$ such that for all $p_i, p_j \in P$ if $\mu(p_i) = \mu(p_j)$ than $I(p_i) = I(p_j)$; R is a function mapping each label l on the edges of G into corresponding binary constraint $R(l)$ on T . Given a TL-graph G a model of G is an interpretation such that if (v_1, l, v_2) is an edge of G , than for all $p_i, p_j \in P$, satisfying $\mu(v) = v_1$ and $\mu(p_j) = v_2$ $\langle I(p_i), I(p_j) \rangle \in R(l)$. TL-graph is consistent if and only if it has at least one model. Two or more TL-graphs are logically equivalent if and only if they has same models [7].

TL-graph G contains an *implicit $<$ constraint* between two vertices v_1, v_2 when the strongest constraint entailed by the set of constraints from which G has been build, is $v_1 < v_2$ and there is no $<$ -path from v_1 to v_2 . A TL-graph without implicit $<$ constraints is an explicit TL-graph. An explicit TL-graph entails $v = w$ if and only if v and w are alternative names of same vertex; $v < w$ if and only if there is a $<$ -path from v to w ; $v \leq w$ if and only if there is a \leq -path from v to w , and there is no any $<$ -path from v to w ; $v \neq w$ if and only if there is a $<$ -path from v to w , or there is a $<$ -path from w to v , or there is an edge (v, \neq, w) . A *Time-graph* is an acyclic TL-graph portioned into a set of time chains (\leq -path), such that each vertex is on one and only one time chain. Search of solution of the TCSP is based on transformation of TL-graph to Time-graph, because if we pass from TL-graph to Time-graph, problems of checking consistency and definition of all feasible constraints will be solved automatically [7].

Algorithms for solving TCSP

During processing inexact information and after solving the task for set of exact constraints search algorithms with returns for processing set of inexact point constraints are used $D = \{D_i; D_i = (x\{R_1\}y) \vee (w\{R_2\}z) \vee \dots \vee (t\{R_k\}u)\}$.

Disjunctive Time-graph (D-Time-graph) is the pair $\langle T, D \rangle$, where T – Time-graph and D – a set of disjunctions in point algebra (PA). Set elements $D : D_i = (x\{R_1\}y) \vee (w\{R_2\}z) \vee \dots \vee (t\{R_k\}u)$ (x, y, z, w, \dots, t, u – temporal variables, R_1, R_2, \dots, R_k – constraints, $i=1..n$). *Realization* of D disjunctions set for Time-graph T is a STC in PA M , where one clause out of each D set disjunction enters, and Time-graph, received by adding T constraint from M , is consistent. D-Time-graph $\langle T, D \rangle$ is consistent only when there is the realization of D disjunctions set for Time-graph T . D-Time-graph $\langle T, D \rangle$ is *exact* in case it is consistent and doesn't contain implicit relations. In order to get explicit D-Time-graph it is necessary to define realization of sets of binary disjunctions D for graph T . In the general case, in order to solve this task for k disjunctions in D set it is necessary to check 2^k possible variants of solution in the worst case. In order to find solutions let's use modification of backtracking algorithm:

Modified backtracking algorithm

Input: D – set of inexact constraints; C – consistent set of exact point constraints.

Output: M – realization of inexact constraints set, solvability flag.

Operations: For constraint $D_i = (x_1\{R_1\}y_1) \vee \dots \vee (x_k\{R_k\}y_k)$ defined operations: $|D_i| = k, D_i[m] = \{(x_m\{R_m\}y_m)\}$

01: $M \leftarrow \emptyset$

02: Rollback \leftarrow **false**

03: **foreach** ($j \in [0, |D|)$) Active[j] \leftarrow 0

04: $j \leftarrow 0$

05: **while** ($j < |D|$) {

```

06: Decided ← false
07:  $i \leftarrow \text{Active}[j]$ 
08: if (Rollback)  $i \leftarrow i+1$ 
09: Rollback ← false
10: while (( $i < |D_j|$ ) && !Decided) {
11:    $M \leftarrow M \cup D_j[m]$ 
12:   if (TCSP with STC  $C \cup M$  is consistent) Decided ← true else  $M \leftarrow M \setminus D_j[m]$ 
13:    $i \leftarrow i+1$ 
14: }
15: if (Decided)  $j \leftarrow j+1$  else {
16:   Rollback ← true
17:    $j \leftarrow j-1$ 
18: }
19: if ( $j < 0$ ) return (false, M)
20:}
21: return (true, M)

```

During modification of initial TCSP the following situations are possible: set of exact constraints has changed; set of inexact constraints has changed; both set of exact and set of inexact constraints have changed. Backtracking algorithm takes significant part of time, necessary to solve the TCSP, that's why during step-by-step search of solutions it is desirable to minimize the number of its recurrent calls, what is reached by deleting corresponding constraints. In the situation when only inexact constraints change (i.e. there exists some set of inexact constraints D^* , which is necessary to add to D), it is possible to initiate the backtracking algorithm not from the very beginning, but from the moment of processing the new constraints (obviously, if we start alg. 1 for the set $D \cup D^*$, we'll spend time to calculate earlier received set M for set D , and only after it will be finalized up to M^* in the result of constraints analysis from D^*). It is possible to build algorithm, which significantly reduces the number of complete repeated analysis of set of disjunctive constraints, because the introduction of the constraint α to the STC of the TCSP C , not requiring the solution of the TCSP with STC $C \cup \alpha$.

Case-based reasoning

CBR, like analogous reasoning, is based on analogy, however, there are certain differences in their implementation [8]. A precedent is defined as a case that took place earlier and is an example or justification for subsequent events of this kind. As the practice shows, when a new problem situation arises, it is reasonable to use CBR method without drawing an analogy. This is caused by the fact that humans operate with these reasoning schemes at the first stages, when they encounter a new unknown problem. CBR solves new problems by adapting previously successful solutions to similar problems. The processes involved in CBR can be represented by a CBR-cycle [9]. Usually, CBR-cycle includes the four main stages:

- *RETRIEVE* the most similar case(s) from the case library (CL);
- *REUSE* the retrieved case(s) to attempt to solve the current problem;
- *REVISE* the proposed solution in accordance with the current problem if necessary;
- *RETAIN* the new solution as a part of a new case (precedent).

The main advantages of CBR include the possibility to use the experience gained by the system for solution new problem situation, without the intensive involvement of experts in a particular problem domain, and the exception of the repeated erroneous decision. In addition, CBR does not require an explicit problem domain model.

The disadvantages of CBR may include the following: the description of cases is usually limited to superficial knowledge of problem domain; a large number of cases may lead to a decrease in system performance; complexities in definition of criteria for indexation and case comparison.

The successful implementation of CBR is necessary to ensure the correct case retrieval from CL. The choice of case retrieval method directly linked to the way of case representation.

Methods of case representation and case retrieval

CL could be a part of the knowledge base of IDSS, but may act as an independent component of the system. There are different ways of representation and storage of cases [10] - from the simple (linear) to the complex hierarchical. It should be noted that the simple methods, based typically on the technology of databases, require much less costly to implement, as well as maintenance and support of CL than the complex. However, the time to search a solution with a simple case representation may require substantially more than in other complex ways of representation and storage of cases.

Usually a case comprises: the *problem* that describes the object state when the case occurred, the *solution* of the problem (diagnosis of problem situation and recommendations for DMP), and/or the *outcome* which describe the object state after the case occurred.

Cases which comprise problems and their solutions can be used to derive solutions to new problems. Other cases which comprise problems and outcomes can be used to evaluate new situations. If, in addition, such cases contain solutions they can be used to evaluate the outcome of proposed solutions and prevent potential problems. Cases can be represented in a variety of forms using the full range of AI representational formalisms including frames, objects, predicates, semantic nets and rules.

In most cases, the simple parametrical representation of precedents can be used, i.e. case description in the form of a set of parameters with specific values and the solution:

$$CASE = (x_1, x_2, \dots, x_n, R),$$

where $x_1 \dots x_n$ – the parameters of the problem situation, describing this case ($x_1 \in X_1, \dots, x_n \in X_n$), R - the solution, n – the number of case parameters, and X_1, \dots, X_n - the domains of permissible values for case parameters .

Well known methods for case retrieval (nearest neighbor, induction et al.) can be used alone or combined into hybrid retrieval strategies [8, 11].

1) *The Nearest Neighbor (NN) method.* This is the most common method of cases retrieval. The main advantages of this method are simplicity of implementation and universality in the sense of independence from the specifics of a particular problem domain. This approach involves the evaluation of similarity between stored cases and the new input case. A main constraint of this approach is the linear dependence of the retrieval time to the number of cases in the CL. Therefore this approach is more effective when the CL is relatively small. This method is also widely used to solve the problems of classification, clustering, regression and pattern recognition.

2) *Induction method of case retrieval.* Induction algorithms (e.g. ID3) generate a decision tree type structure to organize the cases in the CL. This method involves retrieval the required cases by resolving the decision tree tops. This approach is recommended for the grate CL in order to reduce the retrieval time.

3) *Method of case retrieval on the basis of knowledge.* In contrast to the methods described above, this method allows to take into account the knowledge of experts (DMP) in a specific problem domain (the importance of object parameters, identified dependencies and so on). The method can be successfully applied in combination with other methods of case retrieval, especially when the CL is great and problem domain is open and dynamic.

4) *Method of case retrieval, taking into account the applicability of cases.* This method of case retrieval is based not only on similarity of cases with a new problem situation, but also on their applicability in the circumstances. In some systems, this problem is solved by maintaining cases, together with comments of their application.

Use of the mechanism of cases for RT IDSS consists in output of the decision to the operator (DMP) for the current situation on the basis of cases which is contained in system. As a rule, the last stage in a CBR-cycle is excluded and performed by the expert (DMP) because the CL should contain only reliable information confirmed by the expert. Revising and adaptation of the taken decision is required seldom because the same object (subsystem) is considered. More detailed the modified CBR-cycle for RT IDSS is considered in [8, 11]. Thus,

CBR for RT IDSS consists in definition of similarity degree of the current situation with cases from CL. For definition of similarity degree in simple case (parametrical representation of precedents) the NN-algorithm and its modifications are used [11]. For more complex structure of cases like Time-graph the methods of case retrieval on the basis of structural analogy [12] (structural analogy using the context and structural analogy based on Structure Mapping Theory (SMT)) are used.

Application of temporal and CBR in RT IDSS of the car access control

One of the fields where the methods and algorithms, described above, are used is creation of distributed RT IDSS for access control. The considered apparatus is used in the payable car access control system (ACS) sPARK [3]. Modern parking solutions is the complicated complexes, which are equipped with an automatic barriers, the video cameras, fire and access alarm, etc. The major target of the car ACS is passage control of the cars, registration of the visitors and the car owners, stealing prevention and control of equipment in real time. The object of access in the car ACS is the car. The system should control that the car successfully entered to the parking territory. The necessity to control the passage process leads to take into account the temporal dependencies. Let's consider the simple rule, which system should check at the time of car passage [3]:

(1) *If operation of the entrance A_i was activated at the moment t_1 and to the time moment $t_2: t_2 - t_1 > 2.5$ min. operation A_i is still active, then display to operator the notice "delay at the entry".*

The quantity of such or more complicated rules for the basic car access point can reach several tens. The ability of analysis of the sequences of observed by the system actions permits to implement more reliable automatic parking complex control. For example, when the system recognizes the temporal situation, which represented at fig. 2, it can make decision, that visitor making attempt to leave parking in the car, which is not the car in which he had enter to the parking, and notice guards about it.

The developed CBR system (Case Libraries Constructor – CLC) was applied in RT IDSS of the car ACS for diagnostics and detection of different problem (abnormal, critical) situations on object.

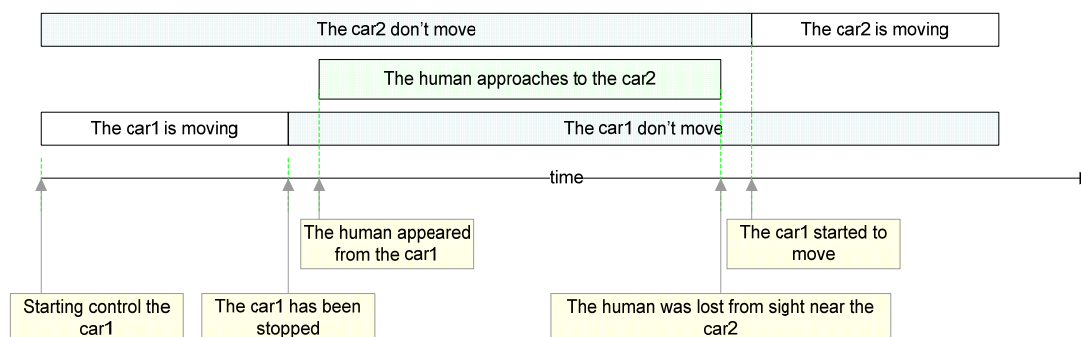


Fig. 2 The temporal diagram of the situation, which is suspicious to car stealing

Conclusion

Possibility to process information about time in the process of its coming is a very important task for plenty of field of AIS. In this work perspective algorithms of the TCSP solution are considered, which are oriented towards the use in distributive IDSS, oriented for dynamic value domains, in particular distributed RT IDSS and described algorithms are implemented in TRS PointTime [3], TRS Singularity and have been approbated in real RT IDSS – the system of access control and security provision – sPARK. The CBR method was considered from the aspect of its application in RT IDSS, in particular, for a solution of problems of real-time diagnostics, forecasting and detection of problem situations on object. The CBR-cycle is considered and different methods of case representation and retrieval are investigated. The proposed CBR method was implemented in CBR system CLC.

Bibliography

1. Vagin V.N., Yeremeyev A.P. Modeling Human Reasoning in Intelligent Decision Support Systems // Proc. of the Ninth International Conference on Enterprise Information Systems. Volume AIDSS. Funchal, Madeira, Portugal, June 12-16, INSTICC, 2007, pp.277-282.
2. Allen, J. F. 1984. Towards a General Theory of Action and Time // Artificial Intelligence Vol./23, pp.123-154.
3. Ereemeev A.P., Kurilenko I.E. Implementation of temporal reasoning in Intelligent Systems // Journal of Computer and Systems Sciences International, Vol. 2, 2007, pp. 120-136
4. Van Beek. P. Reasoning about qualitative temporal information // In Proceedings the AAAI National Conference on Artificial Intelligence, 1990, pp. 728-734.
5. Ereemeev, A.P. and Troitskii, V.V. Models of Representation of Temporal Relations in Intelligent Decision Support Systems // Journal of Computer and Systems Sciences International, Vol.42, No.5, 2003, pp.732-743.
6. Drakengran T. and Jonsson P. Maximal tractable subclasses of Allen's interval algebra. Preliminary report // In Proceedings of the AAAI National Conference on Artificial Intelligence, Portland, OR, 1996, pp. 389-394.
7. Gereviny A. and Schubert L. Efficient Algorithms for Qualitative Reasoning about Time. Technical report 496, Department of Computer Science, University of Rochester, Rochester, NY, 1993.
8. Ereemeev A., Varshavsky P. Analogous Reasoning and Case-Based Reasoning for Intelligent Decision Support Systems // International Journal INFORMATION Theories & Applications (ITHEA) 2006, Vol.13, No.4, pp. 316-324.
9. Aamodt A., Plaza E. Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches // AI Communications, IOS Press, Vol. 7: 1, 1994, pp. 39-59.
10. Leake D.B. CBR in Context: The Present and Future // Case-Based Reasoning: Experiences, Lessons and Future Directions, AAAI Press / The MIT Press, 1996, pp. 3–31.
11. Ereemeev A., Varshavsky P. Application of Case-based reasoning for Intelligent Decision Support Systems // Proceedings of the XIII-th International Conference "Knowledge-Dialogue-Solution" – Varna, vol. 1, 2007, pp. 163-169.
12. Ereemeev A., Varshavsky P. Methods and Tools for Reasoning by Analogy in Intelligent Decision Support Systems // Proc. of the International Conference DepCoS - RELCOMEX 2007. Szklarska Poreba, Poland, 2007, IEEE, pp. 161-168.

Authors' Information

Alexander P. Ereemeev – Applied Mathematics Department of the Moscow Power Engineering Institute (Technical University), Krasnokazarmennaya str., 14, Moscow, 111250, Russia; e-mail: eremeev@appmat.ru

Ivan E. Kurilenko – Applied Mathematics Department of the Moscow Power Engineering Institute (Technical University), Krasnokazarmennaya str., 14, Moscow, 111250, Russia; e-mail: ivan@appmat.ru

Pavel R. Varshavskiy – Applied Mathematics Department of the Moscow Power Engineering Institute (Technical University), Krasnokazarmennaya str., 14, Moscow, 111250, Russia; e-mail: varp@appmat.ru

A GENERAL PURPOSE SOFTWARE SYSTEM FOR LINEAR OPTIMIZATION AND DECISION SUPPORT "OPTIMA-PLUS"

Vassil Vassilev, Mariyana Vassileva, Krassimira Genova, Boris Staykov

Abstract: The paper presents a general purpose software system for optimization and decision support, called *Optima-Plus*. It consists of two independent parts – LIOP-1 system and MKO-2.1 system. *Optima-Plus* system is designed to support the decision maker (DM) in modeling and solving different problems of linear and linear integer single-criterion and multicriteria optimization. The system implements three methods for single-criterion optimization, as well as an innovative generalized interactive method for multicriteria optimization with variable scalarization and parameterization, which can apply most of the well-known scalarizing problems and is applicable for different ways of defining preferences by the DM. The class of the problems solved, the system structure, the implemented methods and the graphical user interface of the system are discussed in the paper. *Optima-Plus* system can be used for education, research and for solving a great variety of real-life problems.

Keywords: Decision Support Systems, Linear, Integer Optimization, Multicriteria Optimization, Interactive Methods.

ACM Classification Keywords: H.4.2. Information Systems - Information Systems Applications – Types of Systems

Conference: The paper is selected from Seventh International Conference on Information Research and Applications – i.Tech 2009, Varna, Bulgaria, June-July 2009

Introduction

Many application problems are related to one criterion, which has to be optimized or multiple conflicting criteria, which should be optimized simultaneously. Optimization is a problem of finding one or more solutions which are optimizing (minimizing or maximizing) one or more objective functions (criteria) in a feasible set of solutions, determined by preliminary given constraints. A single-objective optimization problem includes a single objective function (criterion) and if the problem has a solution, usually results in one final solution, which is called an optimal solution. The solution obtained is not subjective and does not depend on decision makers' preferences. There does not exist one optimal solution of multicriteria optimization problems, which optimizes simultaneously all the objective functions (criteria). There is a set of feasible alternatives with different trade-offs [Ehrgott and Wiecek, 2005], characterized by the following property: each improvement in the value of one criterion leads to deterioration in the value of at least one other criterion. This set of alternatives is called a set of the non-dominated or Pareto optimal alternatives (solutions). Each alternative in this set could be a solution of the multicriteria decision making problem being solved. Usually in practice only one of these multiple Pareto optimal solutions has to be chosen as the final solution of the problem being solved. In order to select one alternative, it is necessary to have additional information set by the decision maker (DM). In other words, compared to the single-objective optimization, in a multiobjective optimization problem there are two (instead of one) phases: an optimization one (for finding Pareto optimal solutions with the help of the software system), and a decision one (for choosing a single most preferred Pareto optimal solution as the final solution). The solutions obtained are to a great extent subjective and depend on DM's preferences.

Many real-life problems in planning, control, analysis and monitoring in economy, transport, industrial production, education, ecology and other spheres may be modelled as single-criterion or multicriteria optimization problems

[Pardalos and Resende, 2002, Korhonen, 1998, Joos, 1999, Rajesh et al., 2001, Thibault et al., 2002, Hämäläinen et al., 2003, Kaleta et al., 2003].

The developed software systems supporting the solution of optimization problems may be divided in two classes – problem-oriented software systems and software systems of general purpose. The problem-oriented software systems serve to support the solution of one or several types of specific optimization problems and very often they are included in other information-control systems. The interface of these systems has been developed in terms of the particular problem solved and depending on the users' preferences, either an exact method is chosen, or a heuristic method is developed especially for the problem solution. That is why, every change in the model or in some of the users' preferences leads to additional investments in new software development and education. The general purpose software systems aid the solution of different optimization problems realized by different types of users or decision makers. These types of systems are oriented to solving optimization problems, which are described in a general (standard) form of the particular class. In general, several algorithms (exact and heuristic) are included in these systems. Each of these algorithms has different advantages and disadvantages, depending on the requirements for high speed, accuracy and necessity to set different input information and parameters, etc. Users with different qualification can work easily with these systems and solve various real-life problems, modeled in terms of the corresponding class of problems, without additional software development.

Various types of software systems [Fourer and Goux, 2001, Pardalos and Resende, 2002, Mittelman and Spellucci, 2005] have been developed for solving linear and linear integer programming problems. The greater part of the professional (commercial) general purpose systems (CPLEX, LINDO, Xpress-MP, and others) are too expensive, but they are intended to solve problems of hundred of thousands continuous variables and constraints, and thousands integer variables and constraints. In some of them, e.g. LINDO system, some modeling languages are included, which enable the entering process of variables, constraints and objective functions to be done in a simple and intuitive way, close to the models developed by the users. Others, like CPLEX and Xpress-MP have their own algebraic input format. There are also libraries of solvers, realizing separate algorithms for solving linear and linear integer programming problems (e.g. LINDO Callable Library). The interface to such software is through subroutine calls, which are used to build up the model, to modify the input data in the corresponding format (usually MPS), to solve the problem and to present the results. Such solvers are used for building in bigger or problem-oriented systems.

Some well known general purpose software systems [Vassileva et al., 2008, Weistroffer et al., 2005], which solve problems of multicriteria optimization, are the following: MKO-2, PROMOIN, VIG, DIDAS, DINAS, MOLP-16, LBS, WWW-NIMBUS, NLPJOB and MOMILP. The Multicriteria Decision Support System for River Water-quality Planning [Lotov et al., 1997] and the ADELAI system for portfolio selection [Zopounidis et al., 1998] are two attractive problem-oriented multicriteria optimization systems. The development of web-based systems in recent years is quite intensive. They are oriented to remote solving of problems. Users submit data to the server and receive results through an e-mail or Web interface to Internet.

The paper describes the basic features of the general purpose software system for optimization and decision support, called *Optima-Plus*. This system is designed to support the DM in solving different single-criterion and multicriteria linear and linear integer optimization problems and it works under MS Windows operating system. *Optima-Plus* consists of two separate parts – *LIOP-1* system and *MKO-2.1* system. The basic idea for developing the system is the traditional users of single-criterion general purpose software systems to be able to make experiments during the solution process of the problem in terms of adding new additional criteria. This basic feature of such a system is innovative and important, because real-life problems for optimization and decision support can be modeled more realistically when using multiple criteria. The class of the solved problems, the system structure, the operation with the interface modules for entering information about DM's local preferences and for visualization of the current and final solutions, as well as the help information, given in a digital and graphical form, are discussed.

LIOP-1 system

LIOP-1 system, which is the first part of *Optima-Plus* system, is designed to solve a broad class of problems of linear, linear integer and mixed integer programming problems.

The formal statement of the Mixed Integer Programming problem (MIP) is the following:

$$F_1 = \sum_{j=1}^n c_j x_j \rightarrow \max \quad (\min)$$

$$\text{s.t.} \quad \sum_{j=1}^n a_{ij} x_j \begin{matrix} > \\ = \\ < \end{matrix} b_i, i = 1, \dots, m$$

$$l_j \leq x_j \leq d_j, j = 1, \dots, n,$$

x_j – integers, $j = 1, \dots, n_1$,

where: $n_1 \leq n$; F_1 is an objective function (criterion); $x_j, j = 1, \dots, n$ are the problem variables; $a_{ij}, b_i, i = 1, \dots, m, j = 1, \dots, n$, are respectively coefficients and the right –hand side of the constraint of index i ; $l_j, u_j, j = 1, \dots, n$, are upper and lower bounds of the variable x_j ; n_1 is the number of the integer variables. If n_1 is equal to n , the problem is said to be an integer programming problem (IP), while at n_1 equal to zero, the problem is a linear programming problem (LP)

LIOP-1 system is realized in Visual Basic and C programming languages. It consists of three main modules: a control module, optimization modules and interface modules. The control module is an integrated software environment for creating, processing and storing of files associated with the system (with a “*.lip” extension), as well as for linking and execution of different types of software modules. The basic functional possibilities of the control module may be separated into three groups. The first group includes the possibilities to use the standard for MS Windows applications menus and system functions – “File”, “Edit”, “View”, “Window”, “Help” and others in the system own environment. The second group of functional possibilities includes the control of the interaction between the modules realizing:

- creating, modification and saving of “*.lip” files, which contain input data and data for the intermediate and final results of the solution of linear and linear integer programming problems;
- solving of the linear and linear integer programming problems entered;
- localization and identification of the errors occurring in the process of operation with *LIOP-1* system.

The third group of functional possibilities of the control module consists in the possibilities to visualize significant information about the operation of *LIOP-1* system as a whole.

The control module is designed in Multi Document Interface style. There is one main form (window), containing the main menu and all the other forms (windows) are its “daughter” forms. Several “daughter” forms can be visualized at one and the same time. In order to accomplish communication with the libraries for dynamic linking of different single-criterion optimization modules, the independent “CSolver” class is used, which is a data processor. It includes modules for input data entry, for output data obtaining and for different setups of the type of the variables and the type of the solution required. It works after the black box principle – an optimization problem in certain format and syntax is set at its input as a parameter, syntactic analysis is done, the names and the type of the variables (continuous/integer) are defined, the type of the solution (continuous/integer) is also determined and the optimization problem is transformed in a format, required by the corresponding single criterion optimization module. The information about the solutions obtained is recorded at the output points of the class, from where it can be obtained. The optimization modules realize two simplex methods of linear programming and the branch and bounds method of linear integer programming.

The interface modules guarantee the dialogue with the system during the input and correction of the input data of the problems being solved, during the interactive process of their solution and also during the digital and graphical visualization of the main parameters of this process. With the help of the editing module, the descriptions of the criteria, of the constraints, as well as the type and bounds of alteration of the variables are

entered, changed or preserved. Another interface module realizes two types of graphical presentation of the information about the values of the criteria at different steps and the possibilities to compare them. Dynamic information is supplied about the purpose and the ways to use the fields and radio buttons.

LIOP-1 system presents to the DM different windows intended for entry and correction of the problem objective function and constraints and for visualization of the results obtained. Figs. 1 and 2 show two of these windows. The window, presented in Fig. 1, is the basic window of the editor for input data entry, called *LIOP-1-Editor*. Entering of the linear and linear integer models can be done in a standard algebraic form, which is more comfortable for the users, because it allows setting of the wanted names of the objective function, of the constraints and variables.

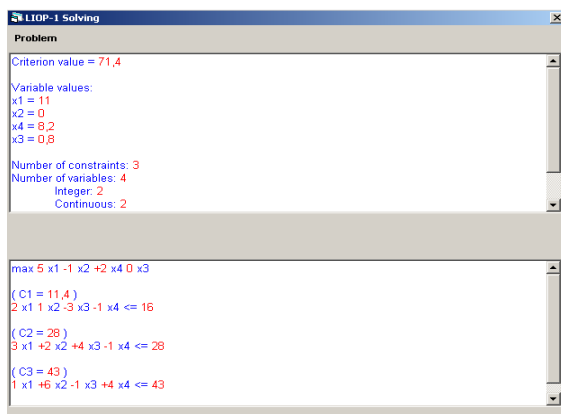


Fig. 1. *LIOP-1-Editor*

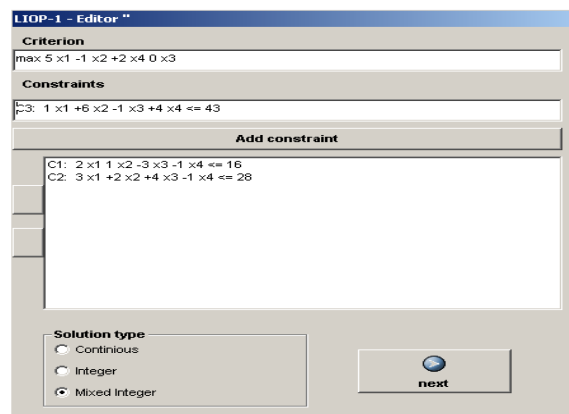


Fig. 2. *LIOP-1-Solving*

The results obtained are shown in “*LIOP-1 Solving*” window (Fig. 2). “*LIOP-1-Solving*” window contains two text fields. In the first field the values of the criterion and of the variables for the obtained solution are shown. In this field the iterations number and some parameters of the problem solution are also given. In the second field the problem being solved is viewed again, but the constraint values for the obtained solution are shown in addition. When solving linear integer problems of larger dimension, some messages for possible breaking of the solution process can be shown periodically (including information about the feasible solutions found). The time of the messages display is fixed in advance, but it can be changed.

The printing of short or detailed information about the problem currently solved is done with the help of “Print” command. A form is opened for selecting the information to be printed. Printing of the solution, the problem being solved, the input variables type and bounds, and detailed information about the process solution may be selected. Requirements about the quality of printing, the number of copies and the character style can also be set.

MKO-2.1 SYSTEM

MKO-2.1 system, which is the second part of *Optima-Plus* software system, is developed on the basis of *MKO-2* system [Vassileva et al., 2007]. It is realized in Visual Basic and C programming languages and designed to aid the DM in the solution of linear and linear integer problems of multicriteria optimization.

The linear and linear integer multicriteria optimization problem may be formulated as follows:

To maximize simultaneously the criteria: $\{f_k(x), k \in K\}$

$$\text{s.t.: } \sum_{j \in N} a_{ij} x_j \leq b_i, \quad i \in M$$

$$0 \leq x_j \leq d_j, \quad j \in N$$

$$x_j - \text{integers}, \quad j \in N'; \quad N' \subset N,$$

where: $f_k(x)$, $k \in K$ are linear criteria of the type: $f_k(x) = \sum_{j \in N} c_j^k x_j$; $x = (x_1, \dots, x_j, \dots, x_n)^T$ is the vector of the variables; $K = \{1, 2, \dots, p\}$, $M = \{1, 2, \dots, m\}$, $N = \{1, 2, \dots, n\}$ and $N' = \{1, 2, \dots, n' / n' \leq n\}$ are the sets of the indices of the linear criteria, linear constraints, variables and integer variables, respectively.

The main approach in solving multicriteria optimization problems is the scalarizing approach [Miettinen, 2003]. The interactive methods are the major representatives of the scalarizing approach. The multicriteria optimization problem is treated in these methods as a decision making problem and the emphasis is put on the real participation of the DM in the process of its solution. The interactive methods are the most developed and widespread due to their basic advantages – a small part of the Pareto optimal solutions must be generated and evaluated by the DM; in the process of solving the multicriteria problem, the DM is able to learn with respect to the problem; the DM can change his/her preferences in the process of problem solving; the DM feels more confident in his/her preferences concerning the final solution of the problem being solved.

MKO-2.1 system is developed on the basis of the generalized interactive method with variable scalarization and parameterization GENWS–IM [Vassileva, 2005, Vassileva, 2006]. This method is a generalization of a large part of the multicriteria optimization interactive methods developed up to the present moment. The generalization is with respect to the classes of the problems solved, the type of the defined preferences, the number and type of the applied scalarizing problems, and the strategies used in the search for new Pareto optimal solutions. Starting from the current (weak) Pareto optimal solution, the generalized scalarizing problem GENWS may be used. Altering some of its parameters the following twelve scalarizing problems can be obtained: the scalarizing problem of the Weighted Sum, the scalarizing problem of ε -constraints, the Chebyshev scalarizing problem, the scalarizing problem STEM, the scalarizing problem STOM, the scalarizing problem of the Reference Point, the scalarizing problem GUESS, the scalarizing problem of the Modified Reference Point, the External Reference Direction scalarizing problem, the Reference Direction scalarizing problem, the classification-based scalarizing problem NIMBUS, and the classification-based scalarizing problem DALDI. The generalized interactive method GENWS-IM has the following main characteristics: the DM may set his/her preferences in very different ways – with the help of the criteria weights, ε -constraints, desired and acceptable levels of change of the criteria values, desired and acceptable levels, directions and intervals of alteration in the criteria values; during the process of the multicriteria problems solving, the DM may change the way of presenting his/her preferences. Starting from one and the same current (weak) Pareto optimal solution and applying different scalarizing problems (with respective alteration of GENWS parameters), the DM may obtain different new Pareto optimal solutions at a given iteration, and this opportunity is especially useful in the education process and in the comparison of different scalarizing problems.

MKO-2.1 software system consists of three main groups of modules – a control program, optimization modules and interface modules. The control program is integrated software environment for creation, processing and saving of files associated with *MKO-2.1* system, as well as for linking and executing of different types of software modules. The basic functional possibilities of the control program may be separated in three groups. The first group includes the possibilities to use the applications, menus and system functions being standard for MS Windows – “File”, “Edit”, “View”, “Window”, and “Help” in the environment of *MKO-2.1* system. The second group of functional possibilities encloses the control of the interactions between the modules realizing the creation, modification and saving of files associated with *MKO-2.1* system, which contain input data and data connected with the process of solution of linear and linear integer problems of multicriteria optimization, the interactive solution of linear and linear integer multicriteria optimization problems entered, as well as the localization and identification of the errors occurring during the process of operation with *MKO-2.1* system. The third group of functional possibilities of the control program includes the possibilities for visualization of essential information about the DM and information of the system operation as a whole.

The optimization modules realize the generalized interactive algorithm GENWS–IM, two simplex algorithms for solving continuous single-criterion problems [Vanderbei, 1996], an algorithm of “branches and bounds” type for exact solution of linear integer single-criterion problems [Wolsey, 1998].

The interface modules provide the dialogue between the DM and the system during the entry and correction of the input data of the multicriteria problems being solved, during the interactive process of these problems solution, as well as the dynamic numerical and graphical visualization of the main parameters of the solving process. With the help of an editing module the formulations of the criteria and constraints are input, altered and stored, and also the type and limits of the variables alteration. Another interface module serves to supply two types of graphic presentation of the information about the values of the criteria at different steps of the solving process, as well as the possibilities for their comparison.

In addition to *MKO-2* system, when using the extended version *MKO-2.1*, the DM could use a variety of new ways for defining the preference information and for choosing the most appropriate method, concerning the problem being solved. In such a way, the DM could build and execute different solving scenarios for a certain multicriteria optimization problem, using, for example, different solving methods starting from a current Pareto optimal solution at a given iteration and compare the results in order to decide in which direction to search for the final solution. The DM could also use another new feature of *MKO-2.1* system – he/she could use different searching strategies at each iteration, as well as he/she could return to a Pareto optimal solution found at a previous iteration and start a new search with a different method, using the same or different preference information.

MKO-2.1 system presents to the DM different windows intended for entry and correction of the criteria and constraints of the multicriteria problem being solved, for setting his/her preferences, for choosing the solving method and for visualizing the current and the final solutions. Figs. 3-6 show four of these windows.

Fig. 3 shows the main window of *MKO-2.1*. It is intended for entering and editing of the initial data of the problem being solved. The same algebraic form for setting the objective functions, the constraints and variables, similar to the one in *LIOP-1* is used, and it is also possible to use the data already entered, to edit or complement them. The window, presented in Fig. 4, is designed to identify the type of DM's preferences. The DM may select among five types of preferences, it is assumed that he/she has selected to set the preferences in the form of desired and acceptable (aspiration) levels, directions and intervals of alteration in the criteria values.

The screen, presented in Fig. 5 shows setting of information for improvement in the value of the first criterion.

Choosing “Graphic” command enables the visualization of two types of graphical information about the process of problem solution. For this purpose, a window with two types of graphics is opened (Fig. 6). With the help of the upper bar-graphic, visual comparison can be made of the solutions found at two iterations, selected in the fields for step selection, given at the bottom of the window.

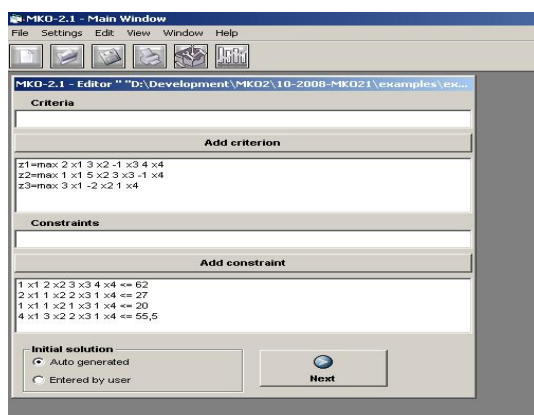


Fig. 3. MKO-2.1-Main Window

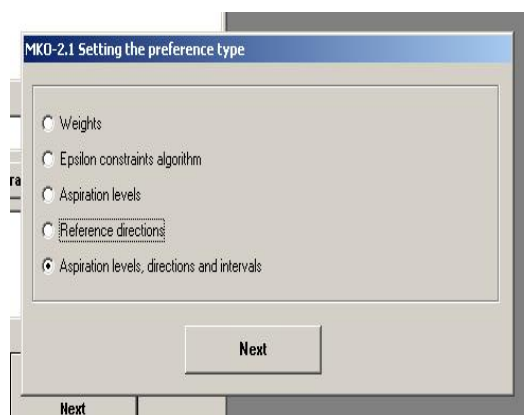


Fig. 4. Selecting the type of DM's preferences

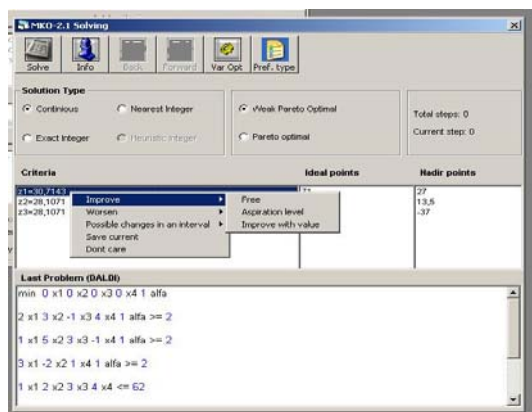


Fig. 5. Setting of an improvement



Fig. 6. Two types of graphics

Conclusion

The general purpose system *Optima-Plus* is designed to support DM in solving different single-criterion and multicriteria optimization problems. The first part of the system – *LIOP-1*, is designed to model and solve a broad class of problems of linear and linear integer programming. The second part of the system – *MKO-2.1*, is designed to model and solve linear and linear integer problems of multicriteria optimization. The DM can solve one and the same problem for one or several criteria. The user friendly interface of *MKO-2.1* and *LIOP-1* systems facilitates the operation of decision makers with different qualification level relating to the single-criterion and multicriteria optimization methods and the software tools. *MKO-2.1* and *LIOP-1* systems can be used for education and for real-life problems solving as well. A number of Bulgarian and European universities use the system for the purposes of education, as well as for experimental and research problems solving. A number of Bulgarian ministries, governmental and private organizations use the system to solve real-life single-criterion and multicriteria decision making problems. The future development of *Optima-Plus* system will be realized in two directions. The first one is connected with the implementing and adding of new methods (including heuristic methods) for single-criterion and multicriteria optimization. The second direction refers to the development of a web-based version of the system, enabling distant decision making.

Acknowledgments

This paper is supported by the European Social Fund and Bulgarian Ministry of Education and Science under the Operation Programme "Human Resources Development", Grant BG051PO001/07/3.3-02/7.

Bibliography

- [Ehrgott and Wiecek, 2005] Ehrgott M., Wiecek M.. Multiobjective Programming. In: Multiple Criteria Decision Analysis. State of the Art Surveys. (Figueire J., Greco S., Ehrgott M. Eds.). Springer, 2005.
- [Hamalainen et al., 2003] Hamalainen, J., Miettinen, K., Tarvainen, P., Toivanen, J.. Interactive Solution Approach to a Multiobjective Optimization Problem in Paper Machine Headbox Design. – In: *Journal of Optimization Theory and Applications*, 116, 2, 265-281, 2003.
- [Joos, 1999] Joos, H. D.. A Methodology for Multi-Objective Design Assessment and Flight Control Synthesis Tuning. – In: *Aero Science and Technology*, 3, 161-176, 1999.
- [Kaleta, et al., 2003] Kaleta, M., Ogryczak W., Toczytowski E., Zottowska I.. On Multiple Criteria Decision Support for Suppliers on the Competitive Electric Power Market. – In: *Annals of Operations Research*, 121, 1-4, 79-104, 2003.
- [Korhonen, P., 1998] Korhonen, P.. Multiple Objective Linear Programming in Supporting Forest Management. In *Multiple Use of Forests and Other Natural Resources*, (F. Hellens, H. Andersen and L. Wichmann, Eds.). Kluwer, Dordrecht, 85-95, 1998.

-
- [Lotov, et al., 1997] Lotov, A.V., Bushenkov, V.A., Chernykh, O.L.. Multicriteria DSS for River Water-Quality Planning. – In: *Computer-Aided Civil and Infrastructure Engineering*. Blackwell Publishing, 12, 1, 57-67, 1997.
- [Mittelman and Spellucci, 2005] Mittelman, H., P. Spellucci. Decision Tree for Optimization Software, 2005. <http://plato.la.asu.edu/guide.html>.
- [Miettinen, 2003] Miettinen, K.. Interactive Nonlinear Multiobjective Procedures. – In: *Multiple Criteria Optimization: State of The Art Annotated Bibliographic Surveys (M. Ehrgott et al., Eds.)*. Springer, New York, 227-276, 2003.
- [Pardalos, and Resende, 2002] Pardalos, P. and M. Resende. Handbook of Applied Optimization. Oxford University press, 2002.
- [Rajesh et al., 2001] Rajesh, J. K., Gupta, S. K., Rangaih, G. P., Ray, A. K.. Multi-Objective Optimization of Industrial Hydrogen Plants. – In: *Chemical Eng. Sci.*, 56, 999-1010, 2001.
- [Thibault et al., 2002] Thibault, J., Lanouette, R., Fonteix, C., Kiss, L. N.. Multicriteria Optimization of a High-Yield Pulping Process. – In: *The Canadian Journal of Chemical Engineering*, 80, 5, 897-902, 2002.
- [Vanderbei, 1996] Vanderbei, R.. Linear Programming: Foundations and Extensions, Kluwer Academic Publishers. Boston, 1996.
- [Vassileva et al., 2008] Vassileva, M., Poles, S. and D. Sasaki. Multiobjective Software. – In: *Multiobjective Optimization: Interactive and Evolutionary Approaches (Branke, J. et al., Eds.)*, *Lecture Notes in Computer Science*, Springer-Verlag, Berlin/Heidelberg, vol. 5252, 329-348, 2008.
- [Vassileva, 2006] Vassileva, M.. Generalized Interactive Algorithm For Multicriteria Optimization. – In: *Problems of Engineering Cybernetics and Robotics*, 56, 69-77, 2006.
- [Vassileva, 2005] Vassileva, M.. Generalized Scalarizing Problem Of Multicriteria Optimization. – In: *Comptes Rendus de l'Academie Bulgare des Sciences*, 58, 5, 537-544, 2005.
- [Vassileva et al., 2007] Vassileva, M., Vassilev, V., Staykov, B., Dochev, D.. Generalized Multicriteria Optimization Software System MKO-2. – In: *Proceedings of the 9th International Conference ICEIS 2007*, (G. Cordoso et al., Eds.). Madeira, Portugal, 283-289, 2007.
- [Weistroffer et al.,] Weistroffer, H., Smith, C., Narula, S.. Multiple Criteria Decision Support Software. – In: *Multiple Criteria Decision Analysis: State of the Art Surveys*, (J. Figueira, S. Greco and M. Ehrgott, Eds). Springer-Verlag, London, 990-1018, 2005.
- [Wolsey, 1998] Wolsey, L. A.. Integer Programming, *Wiley-InterScience*, 1998.
- [Zopounidis et al., 1998] Zopounidis, C., Despotis, D. K., Kamaratou, I.. Portfolio Selection Using The ADELAIIS Multiobjective Linear Programming System. □ In : *Computational Economics*, 11, 189–204, 1998.

Authors' Information

Vassil Vassilev – Prof., PhD, Institute of Information Technologies, BAS, Acad. G. Bonchev St., bl. 29A, Sofia 1113, Bulgaria

Mariyana Vassileva – Assoc.Prof., PhD, Institute of Information Technologies, BAS, Acad. G. Bonchev St., bl. 29A, Sofia 1113, Bulgaria; e-mail: mvassileva@iinf.bas.bg

Krasimira Genova - Assoc.Prof., PhD, Institute of Information Technologies, BAS, Acad. G. Bonchev St., bl. 29A, Sofia 1113, Bulgaria; e-mail: kgenova@iinf.bas.bg

Boris Staykov – Research Associate, Institute of Information Technologies, BAS, Acad. G. Bonchev St., bl. 29A, Sofia 1113, Bulgaria; e-mail: bstaykov@iinf.bas.bg

INTERACTIVE METHODS FOR GROUP DECISION MAKING

Filip Andonov

Abstract: *This paper aims to present new methods for group decision making that take into account different decision making styles and their implementation in decision making using the Group Multichoice system.*

Keywords: *decision support systems, group decision support, client-server, multicriteria analysis*

Conference: *The paper is selected from XVth International Conference "Knowledge-Dialogue-Solution" KDS 2009, Varna, Bulgaria, June-July 2009*

1. Introduction

Decision making problems are unformalized or weakly-formalized problems that in order to be solved require a decision maker (DM). The resulting solutions are subjective, depending on the DM's preferences. The decision making problems are divided into three major classes: multicriteria problems, decision making problems in risk conditions and decision making problems in uncertainty conditions.

Different tasks ranging from planning, management, analysis and control, transport, education to ecology, etc. can be formulated as multicriteria problems. Multicriteria problems are divided in two major classes depending on formal formulation. Multicriteria optimization problems are defined by a finite number of explicitly given functions, which describe an infinite number of alternatives. Multicriteria analysis gives a finite number of alternatives in table form.

In multicriteria analysis problems, a set of criteria are optimized simultaneously in an admissible set of alternatives. Generally, there is no alternative, optimal by all criteria (in this case the solution of the problem is trivial). For practical purposes, it is necessary to select only one alternative – with the help of additional information, extracted from the DM. According to the types of information given by the DM the methods are divided into: weighting methods, outranking methods and interactive methods.

With the development of information technologies, Internet and electronic communications more attention is paid to group decision making problems (GDMP). They are an extension to existing problems. However, in GDMPs the decision is not made by a single DM but a group of DMs, often referred to as experts. There are three different approaches, apriori, aposterior and interactive approach, for aggregating the information provided by the decision makers in GDMP. The methods implementing the apriori approach aggregate the DM's preferences. It is assumed that the group is working as one entity with one hierarchy. Different decision makers lose their identity and they are very consistent. The methods implementing the aposterior methods focuses the final result, final personal ranking of the alternatives of each decision maker where the rankings have to be aggregated in one final ranking. These methods are the crosspoint of multicriteria analysis and voting methods. In voting there are multiple decision makers and multiple alternatives/candidates and the task is to sort them in a priority list. Multicriteria analysis methods provide the personal rankings of the decision makers and voting methods provide the aggregation of the rankings in one final ranking.

So, the development of Internet-enabled decision support systems (DSS), supporting group decisions, is very important in the emerging information society.

In this context the main goal of this paper is to propose new interactive methods for group solving of multiple criteria analysis problems, suitable for different kind of organisational structures.

The rest if the paper is organized as follows. In sections 2 to 4 the author presents several group decision support methods. In section 5 classification of group decision support systems are made, in section 6 the group decision support system Group Multichoice is described and results are summarized in section 7.

2. GECBIM Method

This method is based on the concept that the final ordering vectors received from each DM can be aggregated by solving a multicriteria problem in which all the criteria are ranging criteria. Like the qualitative criteria, the alternatives of ranging criteria have no numerical values. Like qualitative criteria, they have no verbal values. They are simply ordered (ranged) according to value, weight, importance. The definition of the newly-created multi-criteria problem is Matrix A. The group leader solves the problem with the method, described below. Overtly or covertly, he/she may attribute different weight to the separate DMs. The decision of his problem shall be the ordering of the alternatives based on the orderings of the DMs taking part in the process and the leader's attitude to them. Thus the decisions are aggregated into a summarised solution.

	Expert 1	Expert 2	...	Expert n
Alternative 1	$V_{1,1}$	$V_{1,2}$...	$V_{1,n}$
Alternative 2	$V_{2,1}$	$V_{2,n}$
Alternative 3		
Alternative 4		
Alternative 5		$V_{m,n}$

Matrix A

Definitions:

$K_h^{\geq o} \cup K_h^{> o}$ - is a set of ordering vectors with indexes $j \in R$, for which the leader wants to increase their values compared to the values of the currently preferred alternative, where:

$K_h^{\geq o}$ - is a set of ordering vectors with indexes $j \in R$, for which the leader wants to improve their value by aspiration value - Δ_{hj} .

$K_h^{>}$ - is a set of ordering vectors with indexes $j \in R$, for which the leader wants to improve their value in the desired direction.

$K_h^{\leq o} \cup K_h^{< o}$ - is a set of ordering vectors with indexes $j \in R$, for which the leader agrees to deteriorate their values compared to the values of the currently preferred alternative, where:

K_h^{\leq} - is a set of ordering vectors with indexes $j \in R$, for which the leader agrees to deteriorate their values by no more than δ_{hj} ;

$K_h^{<}$ - is a set of ordering vectors with indexes $j \in R$, for which the leader agrees to deteriorate their values in the desired direction;

K^{\times} - is a set of ordering vectors with indexes $j \in R$, for which the leader wants their values in the interval $(a_{hj} - t_{hj}^- \leq a_{hj} \leq a_{hj} + t_{hj}^+)$ compared to their values in the currently preferred alternative;

K^- - is a set of ordering vectors with indexes $j \in R$, for which the leader does not want their values changed;

K^0 - is a set of ordering vectors with indexes $j \in R$, for which the leader is indifferent;

\bar{a}_{hj} - is desired aspiration level of ordering vectors with indexes $j \in K_h^{\geq}$:

$\bar{a}_{hj} = a_{hj} + \Delta_{hj}$, $j \in K_h$;

λ'_j – is the different between the minimal and maximal value of ordering vectors with index j :

$$\lambda'_{ij} = \max_{i \in I} a - \min_{i \in I} a.$$

With the notations defined above the scalarisation problem is:

$$(S) : \min_{i \in I'} S(i) = \min_{i \in I'} \left\{ \max \left[\min_{j \in K_h^{\geq}} (a_{hj} - a_{ij}) / \lambda_j, \max_{j \in K_h^{\leq} \cup j \in K_h^{\leq}} (a_{hj} - a_{ij}) / \lambda_j \right] + \max_{j \in K_h^{\geq}} (a_{hj} - a_{ij}) / \lambda_j \right\},$$

with constraints:

$$a_{hj} \geq a_{ij}, \quad i \in I', \quad j \in K_h^{\geq} \cup K_h^{\leq};$$

$$a_{hj} \geq a_{ij} - \delta_{hj}, \quad i \in I', \quad j \in K_h^{\leq};$$

$$a_{hj} \geq a_{ij} - t_{hj}^-, \quad i \in I', \quad j \in K_h^{\times};$$

$$a_{hj} \leq a_{ij} + t_{hj}^+, \quad i \in I', \quad j \in K_h^{\times};$$

where:

$$\lambda'_j = \begin{cases} \varepsilon & \text{ako } \lambda'_j \leq \varepsilon, \\ \lambda'_j & \text{ako } \lambda'_j > \varepsilon, \varepsilon > 0 \end{cases}$$

When solving the discrete optimisation problem S, the value S(i) is calculated for every alternative with index i, satisfying the constraints. Function S(i) represents the the distance between the alternative with index i and the aspiration alternative, using Tchebishev's metric.

3. Interactive Method for Group Decision Support GCBIM

The GCBIM method is designed for supporting group decision making in the individual consultative style (1). In this style the leader defines the problem and shares his vision with the members of the group. The leader solicits ideas regarding causes of problems and their potential solutions. The leader may also use these individuals' expertise in evaluating alternative solutions. Once this information is obtained, the leader makes a choice of the individual alternative solution which is to be implemented.

In the GCBIM method all decision makers are working in one and the same space of alternatives. In GCBIM it is the rankings, not the preferences that are aggregated and because of that the criteria by which the alternatives are evaluated can be different. The whole process of decision making is iterative. In the first iteration every DM starts with a given alternative and gives aspiration levels, directions and intervals of change and finally gets the ranking of the currently admissible alternatives.

All rankings are sent to the leader. He has solved the problem by himself and has a ranking of his own. When all decision makers send their rankings to the server, the aggregated ranking is calculated with the help of borda score. The leader determines "the chosen" ranking for this iteration by evaluating the rankings, which are closest to the aggregated ranking or to his/her ranking. If he is satisfied with the solution the process is terminated and the first alternative in the chosen ranking is the winner. If the current iteration does not give a solution which satisfies the leader, the process goes on with the next iteration where the best (first in the ranking) alternative in preferred ranking is taken as a reference alternative for the next iteration.

If a_{ij} is the position of the alternative i in the ranking of expert j ; a_{ic} is the position of the alternative i in the aggregated ranking and λ_i – is the difference between the maximal and the minimal position on which alternative i is placed, then we are seeking $\min(\max(a_{ij}-a_{ic})/\lambda_i)$ which gives us the ranking closer to the aggregated ranking.

4. GCBIM-bee method

This method is inspired by the decision making process bees use when searching, evaluating and choosing a new hive, as described in (2).

This method is different from the previous one because it assumes that group goals are the same for all members of the group and they differ only in their vision of achieving these goals. In this method no leader is required, so it is suitable for leaderless team style. The process is iterative and the group work in turns, but instead of sending every ranking for aggregating, the experts can move through several points in the alternatives space with the help of the CBIM method (3) before choosing an alternative to vote for in this turn. Every alternative already voted for is added to the list of “discovered” alternatives. If a DM votes for an alternative more than once, every time after first the weight of his vote is dropped by 15%. This stimulates the DM to search for other possible solutions (alternatives) and gives chance to later found alternatives. The DM can keep voting for one and the same alternative at every turn, but he/she can also choose other alternatives to vote for from the list of “discovered” alternatives.

A system implementing this method should offer a communication between decision makers and using this feature they should be able to “advertise” their currently preferred alternatives at every turn, so that other decision makers will eventually be able to choose an alternative to vote for.

Variant 1

At every iteration/turn alternatives with lowest borda score (the borda score is accumulated through iterations) are dropped. The number of dropped alternatives may depend on the number of iterations already made, but this may not be necessary as this would mean to compensate for the initial “wandering” of the decision makers. However, this issue is resolved due to the fact that decision makers can make more than one step before finding an alternative to vote for.

Variant 2

After the number of “discovered” alternatives reaches a certain threshold the procedure goes on with the final iteration and experts choosing only among alternatives in the list of “discovered” alternatives. The final solution of the problem is the alternative, chosen at this iteration.

The GCBIM, GCBIM-bee 1 and GCBIM-bee 2 methods are intended for implementation in the Group Multichoice system, which already implements four methods for solving multicriteria analysis problems – one weighting method, two outranking and one interactive. The group decision support features are provided by the interactive GCBIM, GCBIM-bee1 and GCBIM-bee2 methods and one aposterior method, based on borda score powered aggregated ranking.

The GCBIM-naname method is designed for facilitating decisions in groups of type Participative Style, Leaderless group and Group consulting style. These type of decisions are based on the idea of equal rights of all participants in the decision process. This allows situations, where there are disagreements in the group and achieving satisfying decision is not possible. In these cases it is important to locate the focus of the disagreement and to be eventually discussed and overcome, and to achieve final decision on this part of the problem for which there is consensus. These goals are achieved by calculating the coefficient of agreement λ_i , which represents a measure of agreement for alternative with index i on a given iteration. To find it, first the table P is calculated:

	P ₁	P ₂	...	P _k
A ₁	p ₁₁	p ₁₂		p _{1k}
A ₂	p ₂₁	p ₂₂		p _{2k}
...
A _k	p _{k1}	p _{k2}		p _{kk}

Table 1

p_{ij} – how many times alternative with index i is positioned on position j

q_i – number of not empty position on row with index i in table 1

m_i – maximal value of row with index i in table 1

Then the coefficient of agreement is $\lambda_i = m_i / q_i$.

On each iteration the alternatives with agreement coefficient λ above given threshold are placed on their positions in the final ranking. This process actually sifts the “easy” alternatives -those, for which there is agreement for their position. For the rest there is obviously no consensus. Therefore discovering of these alternatives is important because:

- in the decision process takes part supra DM, for him/her is important to see this critical alternatives
- finding these alternatives helps experts to focus on critical alternatives
- in usual cases those are the alternatives headed for the first positions, otherwise the problem would be trivial.
- Makes calculating the overall agreement for an iteration easy: $\lambda = \sum_{i=1}^n [m_i / q_i]$

Definition:

An alternative A_i is conflict, if in t_k iterations its agreement coefficient λ_i does not improve. After t_k iterations without increase in its agreement coefficient the alternative is positioned on its place based on borda score. If the position is not empty, the alternative occupies the next free position.

5. Classification of group decision support systems

Decision support systems are divided in two major classes – universal systems and problem-oriented systems.

Depending on the number of the decision makers, the systems can be local, intranet and Internet-based. Local systems are installed on one computer and are used either by one DM, or (in group decision making) by a number of decision makers using the computer in turns. Intranet based systems are deployed in LAN environment, where every decision maker uses his/her own computer. Depending on the software architecture there can be a central server. The time for decision making/using the system is limited due to the use of telecommunication technologies for providing communication between the group participants. The minimum is a text chat, but audio or video conference lines are also recommended.

Internet based systems allow the team to be spread over a large geographic area. This systems has a central server used not only for relaying information but also for storing it, and that is what provides the opportunity for a decision making process extended over a longer periods of time. Every GDSS should provide two functions: communication and decision support.

6. Description of the *Group Multichoice* system

The system interface is designed as a wizard – a sequence of steps where each step is a logical operation. Every stage of the work with the system corresponds to one or more windows in the wizard. The DM is able to go forward or to go back in order to make some changes or to try another option.

Group Multichoice system operates in client or server mode. Every decision making session has only one server and all clients are connected to this server. The leader/group facilitator works with the server.

Group Multichoice offers dynamic context help. The user can get information for every element of the interface just by moving the mouse over it. The system operates in two languages – Bulgarian and English. Whole interface is translated and language can be switched at every stage of working with the system. The translation module is designed in a way that makes adding new languages easy. The system performs automatic software updates if Internet connection is available when it is started.

7. Conclusions and Future Work

GCBIM, GCBIM-bee1 and GCBIM-bee2 are interactive method for solving group decision making problems of multicriteria analysis. Their application covers large part of the specter of different decision making styles.

Group Multichoice is an experimental software system, supporting group decision making and solving multicriteria analysis problems. The system implements four methods for this class of problems and four methods for supporting group decision making. The group decision can be taken in autocratic style with advisors or in democratic style, when all participants are equal. The systems allows network communication, provides user-friendly interface and rich help information. Every decision maker is able to reach a final result using HIS/HER preferred method.

Future work targets the development of commercial system, implementing the presented methods.

Acknowledgements

This work is partially financed by Bulgarian National Science Fund under the project D002-308/19.12.2008 "Automated Metadata Generating for e-Documents Specifications and Standards".

Bibliography

- [1] Fathi El Nadi, 2007 - <http://fathielnadi.blogspot.com/2007/10/tgm-as-tool-to-affect-organizational.html>
 - [2] Group Decision Making in Honey Bee Swarms, Thomas D. Seeley, P. Kirk Visscher and Kevin M. Passino, 2006
 - [3] Partition-based interactive method to solve discrete multicriteria problems, Subhash Narula, Vassil Vassilev, Krassimira Genova, Mariana Vassileva, 2003 –Cybernetics and Information Technologie, vol..4, No 2, 41-60.
-

Authors' Information

Filip Andonov –New Bulgarian University, Sofia, bul. Montevideo 20, tel. 02/8110610; e-mail: fandonov@nbu.bg

SOME ASPECTS OF COMPARATIVE ANALYSIS OF BANKS FUNCTIONING

Alexander Kuzemin, Vyacheslav Lyashenko

Abstract: General problems of carrying out the comparative analysis of the banks functioning and development are considered. The essence of a microsituation for carrying out the comparative analysis of the banks functioning and development is revealed. Various types of microsituations are generalized when carrying out the banks functioning comparative analysis. The approach to comparison of the banks functioning and development based on Wilcoxon criterion is offered.

Key words: bank, analysis, microsituation, statistical conclusion, Wilcoxon criterion.

Conference: The paper is selected from Seventh International Conference on Information Research and Applications – i.Tech 2009, Varna, Bulgaria, June-July 2009

Introduction

When considering directions of solving different aspects of analysis of the market economy objects spatial-temporal dynamics the necessity arises, in any event, to carry out the comparison between the characteristics of the subjects being studied. The performance of such analysis is related, first of all, to elucidation of the arisen situation in the estimation of the economic entity functioning being studied and comparison of such object development with other similar market objects development. As a whole, this contributes not only to revealing of the reasonable approaches to the arising problems solution, but to the possibility to justify look-ahead actions in decision-making relating to the stable functioning and development of the economic entity being studied.

In this work the banks are chosen as an example of the specific economic entities. This is motivated by the fact that the stable and systematic development of the banking sector has a profound impact on the reproduction structure of economy, as through the banking the flow and reallocation of monetary resources and capital funds are organized. At the same time, the analysis of the finance flows both of the banking system, as a whole, and individual banks, in particular, is one of the key components of building of the adequate economic security system of the economic entity operating in the market economy [1, 2]. Eventually, it is precisely this that defines the urgency of the given direction of the investigation, the importance and utility of its consideration as far as the banks is concerned.

Justification of the investigation purpose

The basis for performance of the banks functioning comparative analysis involves, as a rule:

- the models based on generalization and consideration of the normative indices, coefficients of economic agents activity standards with the subsequent estimation of their rating [3, 4];
- the methods of statistical conclusion based on econometric models and methods having in their origin the game theory approaches [5, 6];
- the methods of the fuzzy sets theory [1, 7].

In this case the prevalent approaches of the comparative analysis, as a rule, are:

- either relative generalization of dynamics of the corresponding activity indices of the economic entity being analyzed [8],
- or construction of cluster models which allow to rank the degree of development of the objects being compared [9, 10].

But in any case the mathematical basis of investigations consists, mainly, of the probabilistic methods for data analysis. In this case the solution key aspect consists in determination of the analyzed data distribution parameters with a view to obtain an adequate model, often this is connected with violation of the distribution normality law of the data sampling involving such objects. At the same time, the main problem, arising when constructing an adequate model of the banking comparative analysis, is connected with that the economic development laws assume the presence of such interaction between different subject of the market and account for the action on these interactions of various surroundings manifestations, not having a definite statistical nature in the classical meaning. Solution of the given problem is reached, in a way, through introduction of different aspects of information saturation, banking indices being considered, into the problem treatment. But in the given case another problem arises associated with the necessity to consider the procedure of various manifestations ranking of information saturation of one or other banking indices.

Thus, openness of the problem of construction or choice of the model for comparative analysis of the banks functioning and development stimulates the performance of investigations directed to its solution.

Some methodological aspects of comparative analysis of banks functioning

Classical approach to solution of the formulated problem is the use of the finance flows, which makes it possible to give the most complete description of the banking on the basis of multiple presentation of the initial data (separate indices of activity) x_t^γ of their sets of γ at a certain temporal interval t in terms of the finance flows – $\{x_t^\gamma\}$.

This is associated with that the basis of the flow approach comprises the possibility to realize the structuring of data for complex dynamic systems, it is precisely the structuring that opens different directions for carrying out the necessary analysis [11].

At the same time, the flow processes involve all spheres of the market economy, this is rather important as far as the banks is concerned as the centers of redistribution of monetary and reallocation of capital. This also allows taking into account the degree of various environment factors action, governing thereby the information saturation of the indices being considered.

It should be noted in this case that the flow approach can serve not only as the set of instruments for the banks functioning and development, but also act as the combining center of various approaches applications for carrying such analysis.

At the same time one of the shortages of the flow approach consists in performance of the banks generalizing comparative analysis as the financial flow concept assumes only consideration of some sets of such flows while their structuring is also significant. Therefore, the following part of the given investigation is just devoted to the processes of more precise bank finance flows structuring.

Microsituation concept as the foundation for performance of the banks functioning comparative analysis and development

Some problems of finance flows structuring for carrying out of the economical processes dynamics comparative analysis were considered in [1]. Nevertheless, the problematic aspects concerning comparison of special and general finance flows for their further structuring and analysis remain beyond the scope of investigations. First and foremost, such a generalization concerns, first of all, the problems of description of the situation of functioning of the banks as a complex system and the banks taken individually.

One of the specified problem solution directions can be the use of the microsituation concept which found the proper application when solving a number of problems arising in emergency situations [12].

In the given aspect, to perform the comparative analysis of the banks functioning and development based on the flow approach, by the microsituation, variety of the banking description with the help of the corresponding parameters and indices should be meant.

In this case the concrete microsituation S^L can be described in the form of a separate finance flow or some set of them being defined with a set of data $\gamma, (\gamma = \overline{1, m})$, characterizing the banking of some bank $L, (L = \overline{1, n})$:

$$S^L = (\{x_t^{k1}\}^L, k1 \in \gamma, \quad (1)$$

$$S^L = (\{x_t^{k1}\}, \{x_t^{k2}\}, \{x_t^{k3}\})^L, k1, k2, k3 \in \gamma. \quad (2)$$

Thus, comparing banks between themselves we, first of all, compare the micro situations which in the given case describe the state of the banks functioning and development in terms of some parameter or their totality:

$$S^1 = (\{x_t^{k1}\}, \{x_t^{k2}\}, \{x_t^{k3}\})^1 \approx S^2 = (\{x_t^{k1}\}, \{x_t^{k2}\}, \{x_t^{k3}\})^2, 1, 2 \in L, \quad (3)$$

where S^1 – is the microsituation describing the first of the banks being analyzed, S^2 – is the microsituation describing the second of the banks being analyzed.

At the same time it is possible to carry out comparison of the banks development and functioning as a whole fixing parameter t. Then, in the given case, variation of some of the banking parameters being analyzed x_{tp}^γ for a

fixed date tp in terms of the whole variety of banks– $\{x_{tp}^\gamma\}^L, L, (L = \overline{1, n})$ is considered as a finance flow.

In this case the concrete microsituation can be presented in the following form:

$$S_{tp}^L = (\{x_{tp}^\gamma\}^L, tp \in t, \gamma, (\gamma = \overline{1, m}), \quad (4)$$

or

$$S_{tp}^L = (\{x_{tp}^\gamma\}, \{x_{tp}^\gamma\}, \{x_{tp}^\gamma\})^L, tp \in t, \gamma, (\gamma = \overline{1, m}). \quad (5)$$

Then the comparison consists in performance of the analysis between the microsituations describing the state of the banking system functioning as a whole at some fixed dates of time:

$$S_{tp1}^L = (\{x_{tp1}^\gamma\}, \{x_{tp1}^\gamma\}, \{x_{tp1}^\gamma\})^L \approx S_{tp2}^L = (\{x_{tp2}^\gamma\}, \{x_{tp2}^\gamma\}, \{x_{tp2}^\gamma\})^L, tp1, tp2 \in t. \quad (6)$$

Ultimately, we receive some set of microsituations $\Omega = \{S^L, S_{tp}^S\}$, completely describing functioning and development of the banking system. Since, as mentioned above, not all microsituations can have the normal distribution, then we shall consider nonparametric tests to verify the hypothesis for coincidence of the microsituations being investigated. In the given case it is expedient to use the test Wilcoxon for bound samplings [13], which answers the question: whether some event essentially changing the microstructure hierarchy took place in the analyzed data which characterize different samplings.

In other words, when carrying out the comparative analysis of banks functioning and development the analyzed microsituations distinguishability is studied. Then the value of the Wilcoxon test can be used as the measure of distinction (agreement) of the microsituations being considered. The greater is the value of the test being considered, the more distinguishable as a whole are the microsituations being considered and vice versa, the less is the value of the test being considered the closer are the microsituations being considered.

The initial data and results of the comparative analysis of banks functioning in Ukraine

The foregoing approach is being considered as an example of the banking in Ukraine in terms of such index as a share of the granted credits in the overall totality of bank assets. The paramount importance of consideration of such banking values is associated with that just the credits

- on the one hand, constitute a considerable part of bank operations and, respectively, operating profits in total gains of a bank from such operations,
- on the other hand, the granted credits growth results in credit risks and, consequently, in the banks development destabilization.

Thus the problem associated with the succession of the development dynamics of relation between the granted credits and total volume of banks assets both for the banking system as a whole, and in terms of separate banks functioning is rather significant. The more so the generalized dynamics of the relation between the granted credits and total volume of banks assets as a whole is indicative of the rise in the banks preferred weight with the increased part of the granted credits in their assets volume (Fig.1, generalized using the site www.finance.ua).

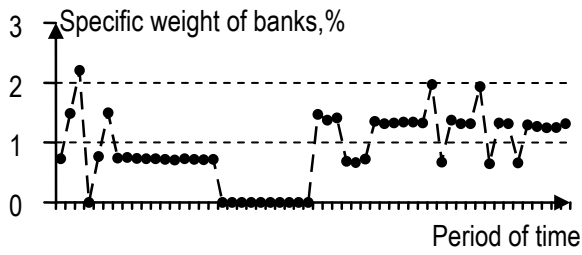
Hence the essence of the first question as to carrying out the comparative analysis of the banking activity consists in estimation of the succession in variation of the granted credits preferred weight in their assets volume during each year of the period being investigated. To analyze such a succession is possible on the basis of investigation of the microsituations each of them describes the state of the banking system functioning as a whole for the fixed date of time t_p in terms of the banking activity index x_{tp}^y – the credits preferred weight in the banks assets (see Eq.7). The results of such investigation obtained within the periods of 2004, 2005, 2006 and 2007 years in section of each month represent a separate microsituation shown in Figs.2-5 (generalized on the basis of the above approach and data of the site www.finance.ua). In this case the black circles mark the microsituations the most consistent between themselves, the microsituations less consistent are not shown at all.

The dimension of each circle represents the degree of correspondence (consistency) of microsituations being investigated in section of every month of the years covered. The smaller is the circle, the greater is the consistency between the microsituations.

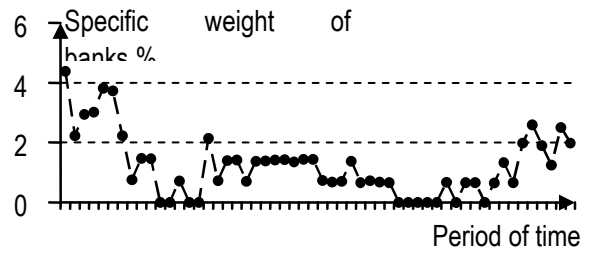
As can be seen from the data in Fig.2 the corresponding consistency between microsituations in the banking system development of Ukraine by the results of 2004 in terms of the credits specific weight in the banks assets is the least one.

At the same time the analysis of data from Fig.2-Fig.5 testifies that year after year the consistency between the microsituations becomes stronger. This is apparent both from the increase in the microsituations number and from the increase in such consistency, the decrease in the circles dimensions demonstrates this. Hence a dangerous situation forms in the banking system development in Ukraine as a whole, which is marked by the rise in the credits specific weight in the banks assets structure, this can cause the rise in credits risk level. Moreover, continuity in such development is observed.

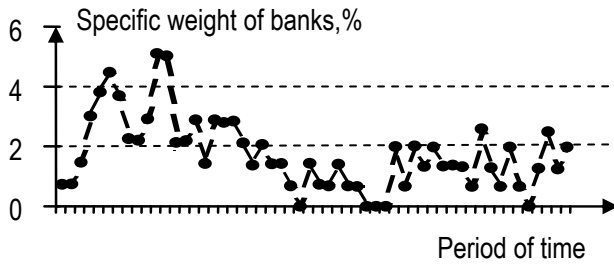
In other hand the analysis performed according to the above methods of consistency in development of separate banks is not less interesting in the considered aspect. To perform such an analysis let us consider a group of 12 banks representing those representing and operating in the same region that makes it possible to consider indirectly the action of various factors on their functioning and development. For the microsituations, their comparison will represent consistency of separate banks development, generalization of their finance flows appears, this represents the specific weight of credits in the structure of such banks assets. Further comparison is carried out on the basis of Wilcoxon criterion according to Eq.3. Fig.6 shows results of such consistency.



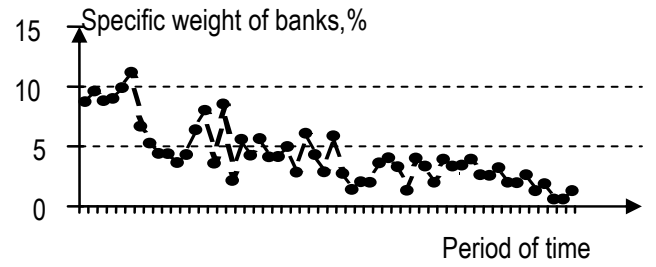
a) specific weight within 0%-10%



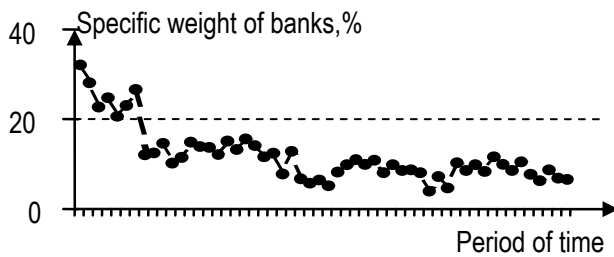
b) specific weight within 10%-20%



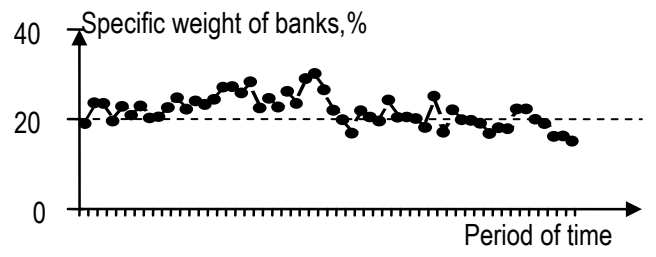
c) specific weight within 20%-30%



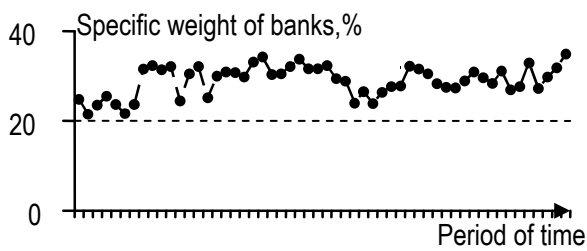
d) specific weight within 30%-40%



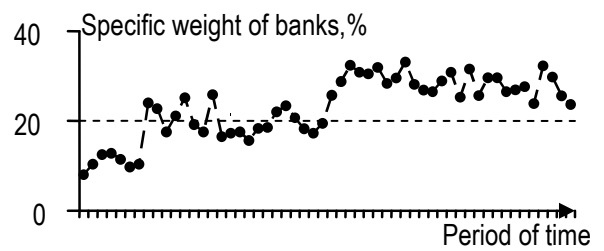
e) specific weight within 40%-50%



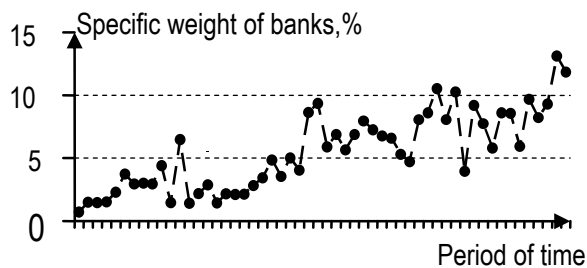
f) specific weight within 50%-60%



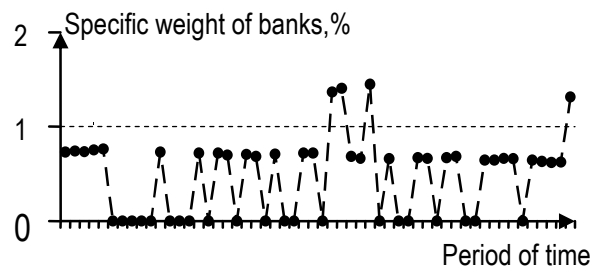
g) specific weight within 60%-70%



h) specific weight within 70%-80%



i) specific weight within 80%-90%



j) specific weight within 90%-100%

Fig.1 The specific weight dynamics of the granted credits to the total assets volume in the banking system as a whole during the period from 01.01.2004 till 01.05.2008

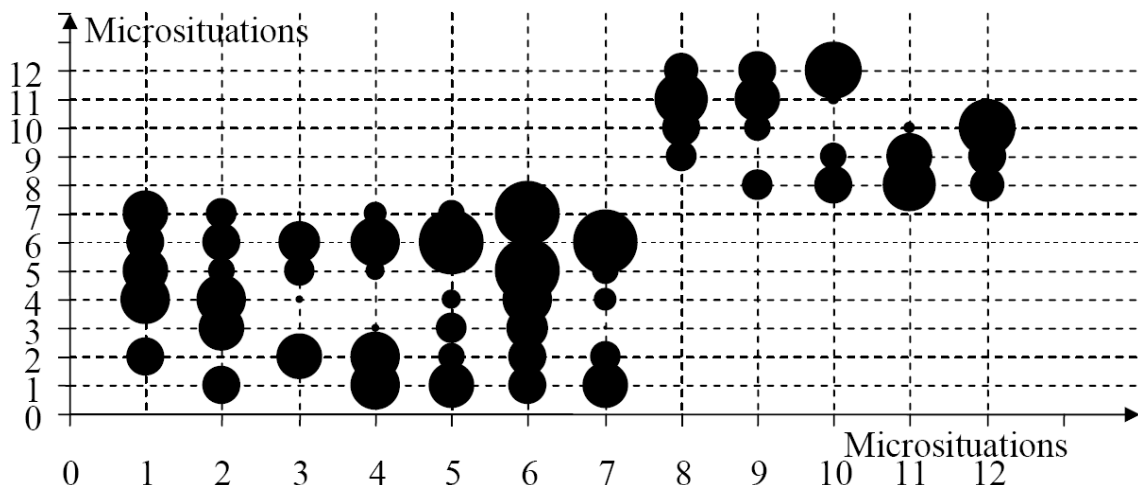


Fig.2 Consistency of microsituations representing variation of the credits specific weight in the banks assets volume according to the results of the banking system work in 2004.

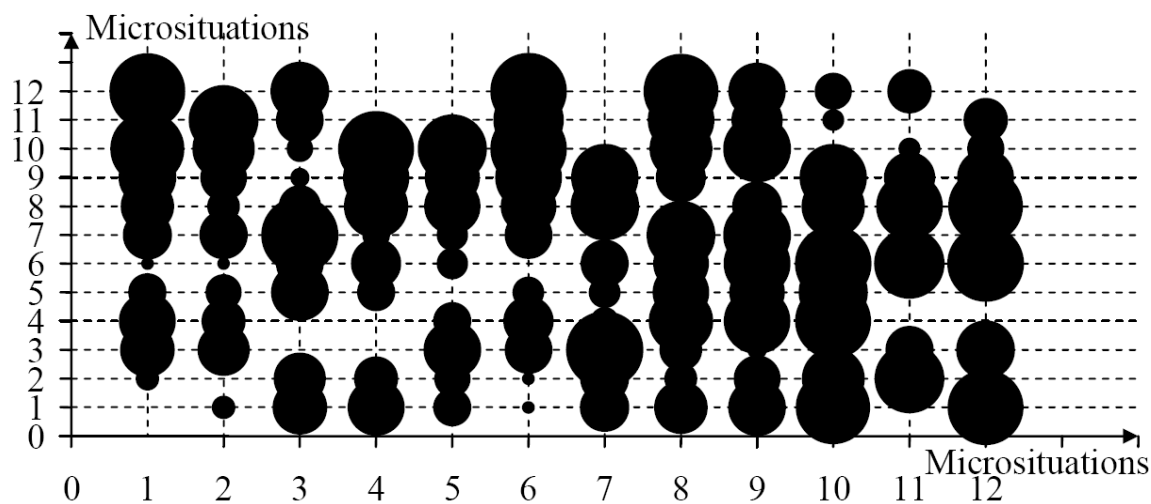


Fig.3 Consistency of microsituations representing variation of the credits specific weight in the banks assets volume according to the results of the banking system work in 2005

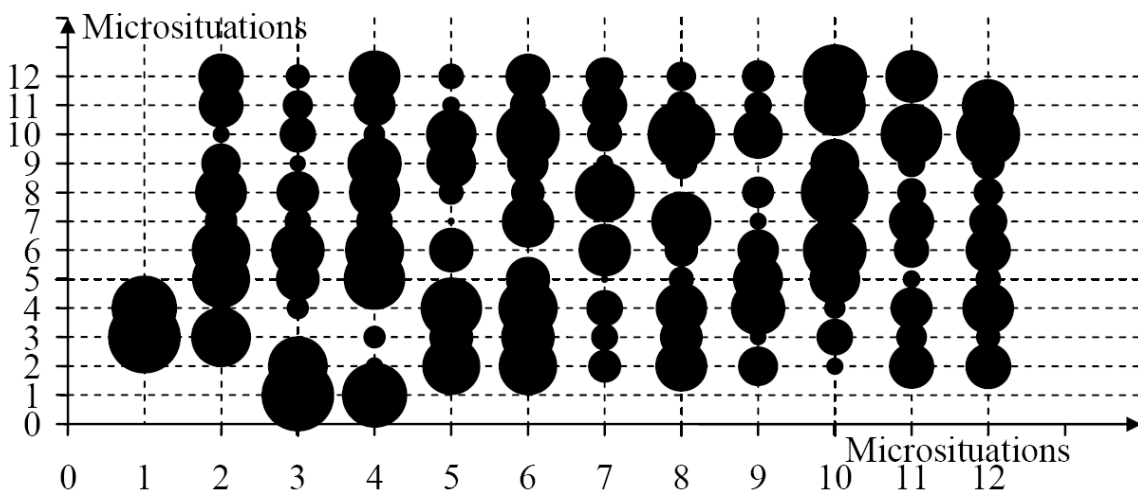


Fig.4 Consistency of microsituations representing variation of the credits specific weight in the banks assets volume according to the results of the banking system work in 2006.

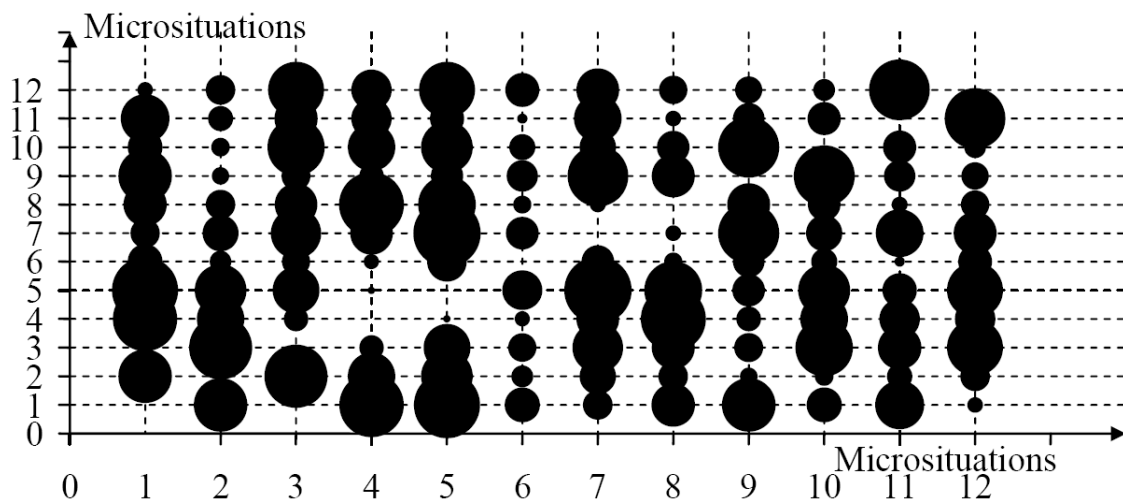


Fig.5 Consistency of microsituations representing variation of the credits specific weight in the banks assets volume according to the results of the banking system work in 2007

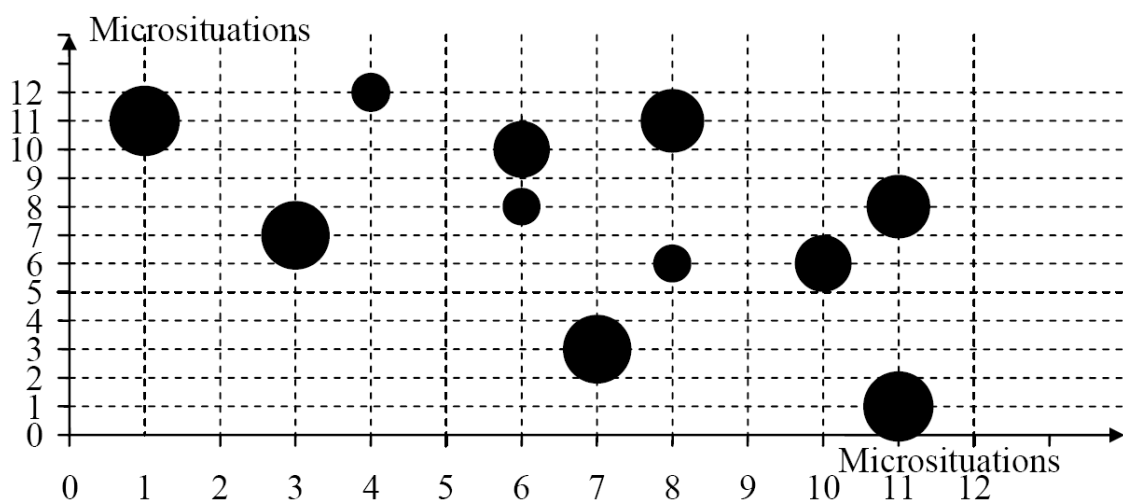


Fig.6 Consistency of microsituations representing variation of the credits specific weight in the banks assets volume according to separate banks of the group being studied by the results of their work within the period from 01.01.2004 till 01.05.2008

As evident from the data in Fig.6 the microsituations consistency in the considered aspect is not observed for the group of banks under study. Thus it may be concluded that each of the banks chooses its own strategy of increase of the credits being granted. Nevertheless, according to the data from Figs.2-5 such a strategy as a whole is aimed at increasing the credits specific weight in the banks assets structure. Consequently, the problem of the credit risk rise remains an urgent one.

Conclusions

Hence the methods for performing the comparative analysis of functioning and development of both the banking system as a whole, and separate banks in particular are offered in the given work. The essence of such methods consists in presentation of the banking activity in the form of a set of microsituations, each of them characterizing such activity based on definite finance flows which in turn reflect one or another index of separate banks activities. In this case non-parametric tests based on Wilcoxon criterion are used for the microsituations comparison. Adequacy and efficiency of the offered methods are approved using the real data concerning one of the banking activity directions. This makes it possible to use the given methods for carrying out the extended comparative analysis of various directions of both separate banks, and banking system as a whole.

Acknowledgements

The paper is partially financed by the project **ITHEA XXI** of the Institute of Information Theories and Applications FOI ITHEA and the Consortium FOI Bulgaria (www.ithea.org, www.foibg.com).

Bibliography

1. Kuzemin A., Lyashenko V. Fuzzy set theory approach as the basis of analysis of financial flows in the economical security system // International Journal "INFORMATION THEORIES & APPLICATIONS" – 2006. – Vol. 13, Num. 1. – P. 45–51.
2. Kuzemin A., Lyashenko V. Analysis of Spatial-temporal Dynamics in the System of Economic Security of Different Subjects of Economic Management // International Journal "Information Technologies and Knowledge" – 2008. – Vol. 2, Num. 3. – P. 234–238.
3. Kuzmenko E.S. Methods for banks safety rating estimation// Urgent problems of economics. – 2007. – № 1. – P. 164–175.
4. Indicators of financial stability. Instruction manual. – Washington, Columbia District, USA: International Monetary Fund, 2007. – 326 p.
5. Azarenkova G.M. Models and methods for financial flows analysis. – KharkovX: ВКФ "Grif", 2005. – 119 p.
6. Watshem T.G., Parramaw K. Quantitative methods in finances. – M.: Finances, ЮНИТИ, 1999. – 527 p.
7. Nedosekin A.O. Application of fuzzy models to banks finances management // <http://sedok.narod.ru>.
8. E.V. Comparative analysis of the banking systems of Russia and Czech Republic under condition of the transitional period in economics // Materials of the V International scientific-practical conference "Countries with transitional economy under conditions of globalization". – M.: РУДН, 2006.
9. Golovan' S.V., Karminsky A.M., Kopylov A.V., Peresetsky A.A. Models of the Russian banks default probability. Preliminary banks partitioning into clusters // Preprint # 2003 XXX. – M.: ПЭШ, 2003. – 49 p.
10. Snityuk V. Evolutionary clusterization of complex objects and processes // XI-th International Conference «Knowledge-Dialogue-Solution» – Varna, 2005. – Vol. 1. – P. 232–237.
11. Kuzemin A., Lyashenko V., Bulavina E., Torojev A. Analysis of movement of financial flows of economical agents as the basis for designing the system of economical security (general conception) // Third international conference «Information research, applications, and education». 27-30 June. Varna, Bulgaria. – Sofia: FOI-COMMERCE – 2005. – P. 204–209.
12. Kuzemin A., Lyashenko V. Conceptual Foundations of Construction of the Models and Procedures for Prediction of the Avalanche-dangerous Situations Initiation // International Journal INFORMATION THEORIES & APPLICATIONS. – 2008. – Volume 15. – №2. – P. 153–158.
13. General theory of statistics: Manual// T.V. Ryabushkin, M.R. Efimova, I.M. Ipatova, N.I. Yakovleva. — M.: Finance and statistics, 1981.

Authors' Information

Kuzemin A.Ya. – Dr. Prof. of Information Department, Kharkov National University of Radio Electronics, Head of IMD, (Ukraine), kuzy@kture.kharkov.ua

Lyashenko V.V. – senior scientific employee, Kharkov National University of Radio Electron (Ukraine)

APPLICATION OF TECHNOLOGY FORESIGHT STRATEGY IN INNOVATION ACTIVITY

Nataliya Pankratova

Abstract: *The methodological tool of the technology foresight strategy in innovation activity which is based on the new expert estimation principles and software for processing the results of the foresight under the real conditions are proposed. Obviously, this strategy should determine the main development directions of strategically important industries, the prospects of competitiveness products, the sensible use of natural, human, industrial, and scientific resources, and the technological potential of the country, and the strategy should consider the needs of the world market of high technologies and hi-tech products take into account the processes of sustainable development. The real projects were realised on the basis of proposed methodological instrument and created software.*

Keywords: *foresight system strategy, new principles, innovation activity, mathematical tool, software*

ACM Classification Keywords: *H.4.2. Decision Support*

Conference: *The paper is selected from XVth International Conference "Knowledge-Dialogue-Solution" KDS 2009, Varna, Bulgaria, June-July 2009*

Introduction

The modern stage of the world's development is characterized by high globalization rates of economical, social, ecological, and other processes. Global processes have created the new development effect, which M. Godet – the French economist – has completely and precisely characterized by a short phrase: «The future ceased to look like the past» [1]. In our opinion, the essence of this effect is based on the fact that the typical approaches and methods to prognosis targeted on investigation of the evolutionary gradual development cannot detect and foresee the changes in the process that are quick in time and discontinuous. At the same time, such processes characterize the modern global dynamics of the world system. Such dynamic characteristics given by the continuous influence of the semistructured, multilevel, hierarchical, practically unlimited set of continuously changing positive and negative correlations, interdependencies, and interactions between different processes, factors, and situations. The dynamic results started to reveal themselves as «the unforeseen and unpleasant consequences» [2]. Under the existing conditions of the world dynamics, the innovation activity becomes the defining tendency of the economical and social progress.

Practical necessity in new, system coordinated principles, approaches, methods of revelation of possibilities and measurement of prospects and tendency of civilization and international economy development and prospects of innovative development of specific country as well have appeared. In some countries and international organizations the new apparatus of qualitative and quantitative foresight of different dynamic processes, which in general calls science and technology or technology foresight, begins to develop intensively.

Goal of this paper is to present methodological tool for scenario analysis information platform creation in real conditions of innovation activity [2]. It is necessary to note, that when deciding practical foresight problems on the basis of the scenario analysis platform, expert estimations are received, as a rule, via Internet in on-line mode. One of the most important aspects of scenario analysis information platform is development of expert estimation procedure conception, methodology and software.

1. Peculiarities of innovative activity management

The complexity of dynamics and uncertainty of properties in intercommunications, interdependencies, interactions between different processes and factors require prior investigations on the basis of the models which define existing interconnections between the processes and practically required interconnections between the innovation activity subjects. The structural model is proposed on Fig. 1, which defines the structure and goals of systemic relations between the main organizational innovation activity subjects (industry, science, education) themselves and between them and the market of intellectual products [3].

We should note that the innovation activity strategy of the country is totally different from the development program of the country under the planned economy conditions. It is so not only due to the underlying differences between the market economy and the planned economy, but it is also due to the peculiarities of innovation activities under the modern conditions of high dynamism and globalization of international processes. The main difference is the conceptual uncertainty; unlike the informational uncertainty, such an uncertainty is conceptual in a sense of understanding it as the uniform system of uncertainty, ambiguity, and inconsistency of interrelated and interdependent elements of the set of uncertainties of various types. This set contains the uncertainty of goals of development and perspectives of innovation product competitiveness; the dynamics of the demand and distribution markets of competitive products, active competitors' opposition; the situation risk uncertainty in the dynamics of developing, producing, distributing, and using the innovation products.

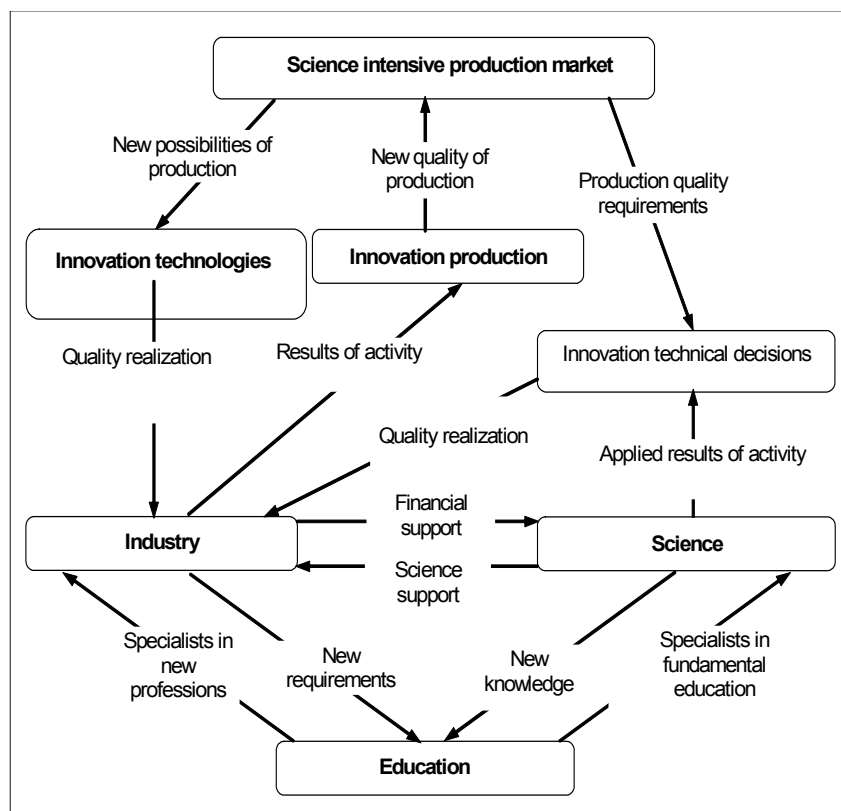


Fig. 1. System subject correlation structure of innovation activity

2. Scenario analysis as a basic foresight methodology

Formally by the term foresight we shall understand a complex system with a human factor that may be presented by a company, an enterprise, an industry, or a country as a whole, uniting certain parts of society (humans or any

social groups) with technological, ecological, economic, and other components that are characteristic of such systems. First of all, these systems may be subjected to various external influences and restrictions: legislative, political, economic, etc. Systems with a human factor consist of various subsystems with complex interconnections between them, of both quantitative and qualitative character. They function according to various purposes that, in most cases, clash among themselves. For complex systems with a human factor considerable uncertainties of data and information are characteristic, and risks of various sorts are inherent in their behavior. Expert judgments concerning their qualitative character always have a subjective character. Thus, taking into account all peculiarities of such systems, regarding their behavior in the future quite definite decisions should be taken in the form of scenarios and strategies of their development. Qualitative and quantitative methods are used to solve such problems in a complicated human-machine procedure. It is important to note that the influence of the human factor on foresight results determines the major level of subjectivity of the given procedure. It is connected with a combination of objective knowledge and subjective human attitudes to the object under investigation.

Every expert taking part in the foresight process expresses his or her personal opinion in the form of subjective estimation; however, at the same time, he or she should be maximally guided by objective knowledge. At the same time, such a glance into the future requires hypotheses and assumptions. The convergence of objective knowledge and creative assumptions in an interactive human-machine procedure allows one to increase the accuracy of the development scenarios for processes, phenomena, and events under investigation. Such scenarios can be constructed with the help of the universal set of means and approaches called scenario analysis methodology [2], which is a complex of mathematical, software, logical, and organizational means and tools for determining the sequence of method application, their interconnections, and the formation of the foresight process in general. In the first stage one should study the problem and the object of foresight using qualitative and quantitative analysis. After that qualitative and quantitative information is reduced to a unique platform. Then the consequences of the methods will be defined and the interconnection between them will be established. This makes it possible to form a complete foresight process and to develop a group of scenarios of the future behavior of the foresight object (complex system with a human factor).

Analyzing characteristics and peculiarities of each of the developed scenarios the group of strategic decision makers chooses scenarios that interest them, develops a plan of action according to the foresight object, and ensures implementation of this plan. (Fig.2). Comprehensive analysis of these scenarios is carried out in accordance with the following procedure: determine the reality and feasibility level for each scenario; estimate event probability based on scenarios; estimate risks connected with each of the scenarios; construct simulation models; select most acceptable scenarios based on above indicated criteria. The realness and realizability level of each scenario (confidence level) is determined on the basis of searching for each event of the scenario "contrary instance" or "antievnet," which excludes the possibility of realization of the examined event. If such antievnet are found, the confidence level for the investigated scenario decreases. The procedure is based on the application of methods of combinatorial mathematics of variants of enumeration of combinations.

3. Development of principles of technology foresight strategy

The formulated requirements raise a number of completely new problems related to the strategy of technology foresight; among these problems, first of all, we should underline those related to the development of new expert estimation principles in technology foresight. A technology foresight strategy should not only satisfy the new requirements, but also correspond to new conditions in innovation activities and correct itself in synchrony with the changing market for the corresponding product type. Adjustments of one's strategy should be adapted to dynamic market changes while preserving the strategic goals in the national manufacturing. In turn, strategic

goals should be coordinated with the long-term development trends of the world market and the development dynamics of the corresponding national industry.

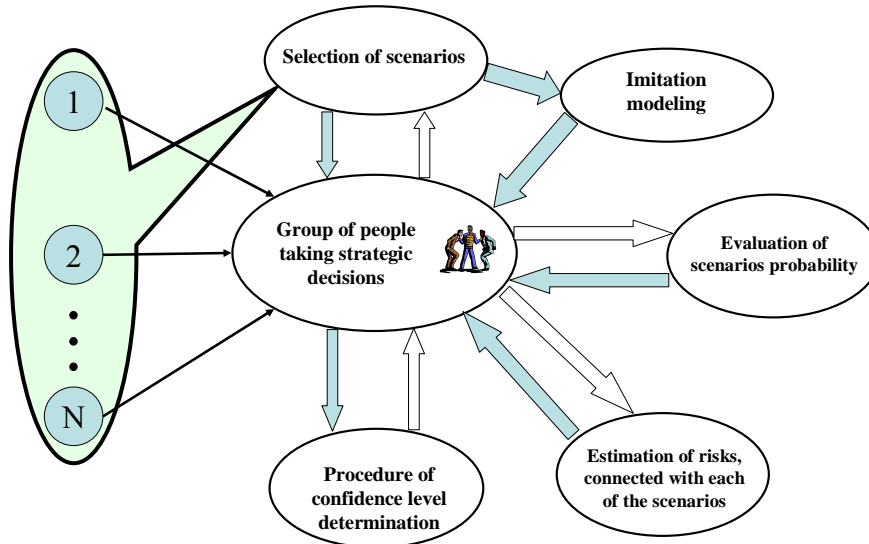


Fig.2. Submission of scenarios to the group of people taking strategic decisions, their analysis and selection

The necessity of developing new principles of expert procedures is dictated not only by the listed earlier factors and peculiarities of innovation activities, but also by the essential differences between technology foresight conditions and typical conditions of expert estimations. Typical expert procedures are oriented toward intuitive execution of the logical analysis of objects or processes of reality on the basis of comparing, ranking, systematizing, grouping, uncorrected proof measuring, classifying, and other procedures. In particular, the classification may be performed as the ranking of objects by increasing (or decreasing) values of a given parameter the quantitative (or qualitative) value of which is known for every object, and may be measured or calculated. Other expert procedures are implemented in the same way. Essentially, the ability to obtain initial information with the required completeness level means that an expert (each expert individually or the expert group as a whole) knows a priori that the given expert evaluation procedure can certainly be performed. Moreover, when information is provided with the use of expert evaluation results, it is possible to make a valid statement about the feasibility of implementing the given technologies, products, or other products, processes, and actions under investigation. Such conditions are satisfied in practice for a wide range of problems. From the theoretical viewpoint, this approach corresponds to the *potential realization principle*, which is accepted in intuitive logic [4]. The main point is that an expert knows a priori, proves, or postulates that the procedure is potentially realizable, and that is why he can ignore limitations in space and time resources. Furthermore, in the case of implementing the potential realization principle, the conditions are satisfied for the invariance principle of intuitive logic.

The main point of the invariance principle is the following: if the validity of any statement, opinion, or conclusion is determined or proved, it will hold in the future too [4]. Note that the conditions and principles mentioned earlier allow us not only to simplify the solutions to many important theoretical and practical problems, but it also allows us to create various automatic systems for classifying, comparing, measuring, and rejecting various products, making it possible to exclude human participation in expert procedures. However, fundamentally different

conditions are typical for technology foresight expert procedures. The high competition dynamism of innovation products in the world market place has created fundamentally different conditions of innovation activities, which are characterized not only by the conceptual uncertainty of market dynamics, but also by the multifactorial risk of overdue realization and quick obsolescence of an innovation product proposed in any project, and also due to the absence of technological possibilities for product realization. In particular, information incompleteness and uncertainty of many properties and peculiarities of attitudes toward the innovation product in the market are typical for an innovation project, for example, the information on how potential buyers and competitors feel about this product. Thus, the technical foresight expert procedures cannot be implemented for a typical expert estimation, which requires complete and valid information about the object under investigation. Therefore, technology foresight expert evaluation procedures should not only conform to the fundamentally new requirements, but it should also contain the fundamentally new characteristics and principles of organization and realization of expert procedures.

Expert procedures in technology foresight should be organized and realized on the basis of technologies that allow the insufficiency and uncertainty of the initial information about the investigated innovation object to be supplemented and compensated with knowledge, experience, intuition, and human foresight. In this case, the expert estimation results become dependent on many new risk factors that, in the group strategy of a typical expert estimation, are practically excluded. We need to emphasize that, among these factors, the expert estimation results depend on the peculiarities, knowledge, experience, and ability of a customer to determine the expert estimation goals and problems, a manager to choose the expert estimation strategy and form experts' groups, and, what is most important, the results depend on the abilities, knowledge, experience, intuition, and foresight of each expert. Note that, under conditions of uncertainty, the dependence arises on external factors and time, on the amount and level of research in scientific areas that have a direct or indirect effect on the innovation object under investigation. The essence of such dependence is the fundamental possibility to change estimations of an innovation object during a comparatively short period of time. If something is not known today, that does not mean that it will be impossible to realize an innovation object in the future. It may become known, possible to realize, and very useful and needed tomorrow. The most impressive example is a personal computer. But it is not excluded that tomorrow someone proves that a given innovative technological solution or a product is unquestionably impossible to realize.

Therefore, conditions and factors that exclude a priori postulation of potential realization of an innovation project or product take place in innovation activities. And, as a result, the possibility is excluded of realizing the invariance principle in time for the correctness of an expert statement, opinion, or conclusion. It follows that other principles that take into account its conceptual uncertainty must work for innovation activities; therefore, they differ completely in their properties from the principles of intuitive logic that are the potential realization principle and invariance principle [4].

Based on the expert procedure analysis of innovation activities, the following principles can be proposed. Instead of the potential realization principle, *the possible realization principle* is suggested [5]. This principle postulates that for certain innovation objects (scientific ideas or technical solutions, or projects of industrial products, or manufacturing technologies), the initial expert estimation results cannot guarantee their practical feasibility or prove the impossibility of being realized. This principle postulates that for the listed innovation objects, based on the expert estimation results of the presented information, a reliable estimation cannot be obtained a priori that would allow one, for the object under investigation, being grounded and valid, to exclude the possibility of being unrealizable. The estimation for an innovation project retains the uncertainty of the conclusion about the possibility of realization, until, theoretically or exuncorrected proof perimentally, the possibility of technical or technological realization is proved for the product.

Instead of the *truth invariance principle*, which postulates the invariability of the theoretical or technical statement, judgement, conclusion, or opinion about the object for a comparatively long time, a completely different principle is needed. Such a necessity is brought about by the previous principle and the innovation activity practice, as expert estimations under conditions of conceptual uncertainty cannot stay the same for a long time. In scientific research and experimental design processes, not only new knowledge about a product under development is accumulated, but the conception about the product's characteristics, use, and application areas may change; new inventions, technical solutions, and other know-how may emerge. Thus, the new principle must reflect the probabilistic characteristics of invariance in time of the initial estimation results of an innovation object, and that is why we shall call it the *probabilistic invariance principle*. This principle postulates that initial expert estimation results certain innovative ideas or technical solutions, industrial products, or manufacturing technologies are probabilistic and do not guarantee that they will be saved in time. Initial expert estimation results obtained under conditions of conceptual uncertainty as positive or negative findings, proposals, or recommendations are not invariant and may substantially change, be confirmed, or be disproved as time passes. Therefore, we do not exclude in a certain time frame both safekeeping of truth expert statements, opinions, or findings, and the possibility of disproving them.

4. Technology foresight expert estimation procedure requirements

Assertions given earlier prove that it is possible for innovative objects that do not have analogies and prototypes of objects that are actually produced to be producible or nonproducible. Therefore, there is a series of certain practically important requirements for technology foresight expert estimation procedures.

First, while performing these procedures, one must not only estimate the properties, advantages, and disadvantages of innovative objects, but also estimate the probabilities of them being realized in practice. Then, it is advisable, taking into consideration the subjectivity of expert estimations, to introduce an additional index that should determine the expert's confidence in the estimation. The introduction of such indices opens the possibility to substantially improve the reliability of the total object estimation, because the group of experts is formed by taking into account the known a priori uncorrected proof level of each expert's competence and the possibility of taking this level into account while processing the expert estimation results.

Second, while doing the expert estimation of innovation objects in the technology foresight process, it is advisable to introduce latent indices of project quality estimation and innovation objects proposed in this process. In particular, such indices characterize the practical necessity, technology possibility, and economy expediency of an investigated object [5]. In order to obtain the coordinated qualitative and quantitative estimations of these indices, it is necessary to provide each expert with a simple understanding of the content and the sense of estimated characteristics of expert estimation objects.

The main feature of the latent indices of innovation projects and products is their mutually exclusive dependence. It expresses itself in such a mutual systemic coordination of conditions for realizing characteristics that an absence of any of them excludes the necessity of realizing other characteristics. For example, an absence of market demand for innovation production excludes the necessity and expediency of its industrial production. The next main feature is the discrepancy of requirements for the internal and external characteristics of innovation products. This follows from the contrast of technical and economic interests of producers, which are expressed by internal characteristics (prime cost, technical and economic effectiveness, convenience, etc.), and the social and market interests of its potential customers, which are presented by external characteristics (price, aesthetics, quality, usage convenience etc.).

Other features are the impossibility of direct qualitative estimation of combined characteristics, as they depend not only on qualitative and quantitative indices of innovation products, but also on sets of different factors. We

cannot always control and forecast the structure and level of their influence. Among uncontrolled factors the most important are the various risk factors in the dynamics of the sequence of stages in the product's life cycle. We need to emphasize the importance of risks in the production cycle and the risks of market demand for new products. Their impact may substantially lessen the technical and economic effectiveness of production or even make the production unprofitable.

Modeling top-priority problems of Kiev city on basis of analytic hierarchy process (AHP) is shown on fig.3. AHP was developed by T. Saaty [6, 7] and is based on the description with multiple criteria of a problem and allows to turn the analysis even of very complex problems into a sequence of pair-wise comparisons of their individual components. During the problem stating stage, the group of experts performs decomposition of a complex problem – represents it as a hierarchical sequence of interconnected and interdependent systems and subsystems; determines its elements and relations between them. Then, the hierarchy is built which is the systematic model of reality. The top of the hierarchy is the general goal, sub-goals lie lower, then forces that influence these sub-goals, people, goals of people, policies, strategies, and, at the end, the outcomes that are the results of strategies. At the next stage of solving, the individual components of hierarchy are compared to each other. As a result, the relative degree of interaction intensity between elements in the hierarchy can be determined. Then, these judgments are expressed numerically. At end of the problem analysis, AHP includes procedures for the multiple judgments synthesis, determination of the priority of criteria and alternative solutions. Having completed the processes of decomposition and formalization of the considered problem, we apply the Analytic Hierarchy Process to the formalized problem with the consecutive account of its comprising procedures such as the usage of the fundamental integer scale, calculation of the matrix of paired comparisons, coincidence of matrices, local priorities, and carrying out the synthesis of priorities.

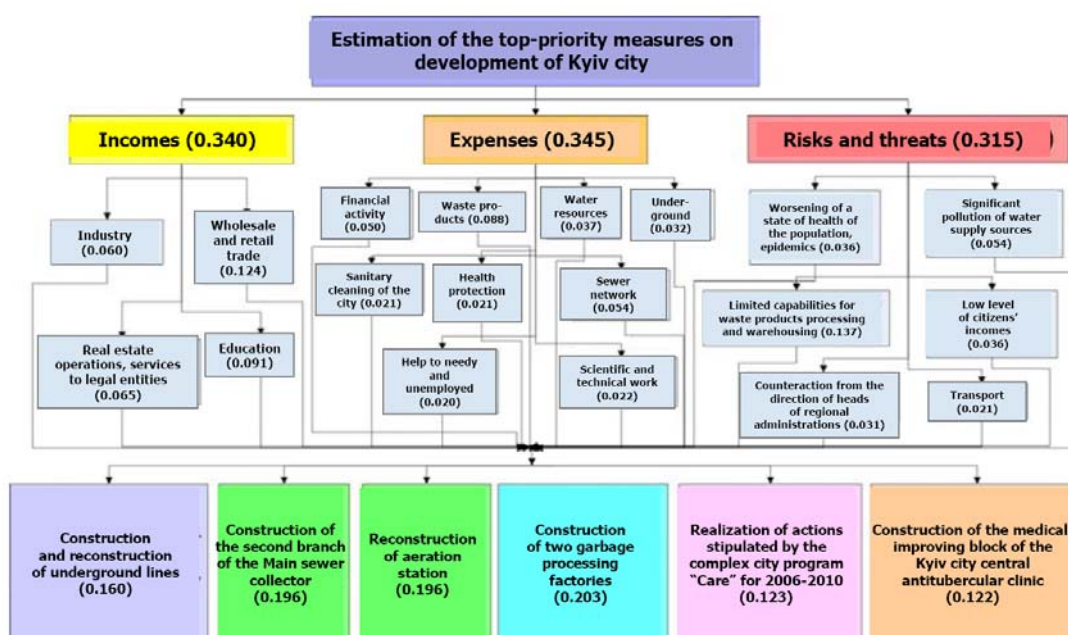


Fig.3. Estimations of top-priority problems of Kiev city

Having calculated the global priorities for every variant, we obtained (fig.3), that construction in Kiev of two garbage processing factories with priority 0.203; reconstruction of aeration station and construction of the second branch of the main sewer collector with priority 0.196 are more preferable than stipulated by the complex city program "Care" for 2006-2010 (0.123); construction of the medical-improving block of the Kyiv city central antitubercular clinic (0.122) and construction and reconstruction of underground lines (0.160)

Conclusion

The realization of the expert estimation procedure of an innovation activity is carried out according to stages of expertise realization with the help of the developed software focused on automation of the expert estimation procedure.

Because the expert estimation needs to involve experts who are on different continents, in different time zones, with different levels of computer knowledge, the system is implemented in the form of a web-site and works in an on-line mode. It is possible also to explain this choice by the chosen strategy of expertise organization – implementing the independent estimation in which the information about the structure and size of the expert group is anonymous. Besides, the program allows saving the results of the analysis of the quality of an expert estimation in files.

The software was created using the programming language Python with the elements of Z-object publishing environment. The core of the system is the database constructed on the basis of MySQL 3.23.21, which allows storing the large volume of information on expertise results and ensuring the independence between the developed program and the area under research.

The real projects were realised on the basis of proposed methodological instrument and created software.

Acknowledgement

The paper is partially financed by the project **ITHEA XXI** of the Institute of Information Theories and Applications FOI ITHEA and the Consortium FOI Bulgaria (www.ithea.org, www.foibg.com).

Bibliography

- [1] Godet M. Reducing the Blunders in Forecasting // Futures, 1983. — 15, № 3. — P. 181–192.
 - [2] Zgurovsky M.Z., Pankratova N.D. System analysis: Theory and Applications. Springer. —2007. — 475 p.
 - [3] Pankratova N.D. Problems of market formation of research production //Confirmation of innovation model of economic Ukraine development. Proceedings of research-technical conference. Kyiv. –2003. —P. 337-346. (in Russian)
 - [4] Dragalin A. Mathematical intuitionism. Introduction into proof theory. — M.: Nauka, 1979. — 256 p. (in Russian)
 - [5] Pankratova N.D., Oparina E.L. Problems and principle of system strategy in innovation activity// Materialy Naukowe. XI Miedzynarodowe Sympozjum «Geotechnika - Geotechniks 2004», Gliwice-Ustron, Polan. –2004. –P. 89-96.
 - [6] Saaty T., Kerns K. Analytic planning. System organisation. M.: Radio and communication, 1991. – 224 p.
 - [7] Saaty Thomas L. How to make and justify a decision: the Analytic Hierarchy Process (AHP). Part 1. Examples and applications. // System research and information technologies. – 2002. - №1. – p. 95 – 107.
-

Authors' Information

Nataliya Pankratova – Depute director of Institute for applied system analysis, National Technical University of Ukraine “KPI”, Av. Pobedy 37, Kiev 03056, Ukraine; e-mail: natalidmp@gmail.com

FUNCTIONAL APPROACH TO EMERGENCY RISK MINIMIZATION

Oleksandr Kuzomin, Illya Klymov

Abstract: *An alternate approach to emergency risk minimization, based on lazy calculations and high-level functional interpretation is presented in this article. This approach is compared with classical imperative approach, based on statistical analysis on real data in order to prove its effectiveness.*

Keywords: *risk minimization, emergency, functional programming, lazy calculations*

ACM Classification Keywords: *H.1.1 Systems and Information Theory*

Conference topic: *Information Systems.*

Introduction

Emergency risk minimization is a common task for geo-informational systems, which are implemented for everyday use in governments of many countries. A great number of different approaches were developed in order to provide an effective model for describing current situation, forecasting of possible emergencies and ways of reducing losses.

However, there are a number of challenges for such systems, which include, but not limited to:

- *Time factor* – usually an estimation of current state is needed urgently. That means that system is strictly limited in amount of possible calculation. The main challenge is to find optimal balance between complexity of developed system and precision of forecast.
- *Information factor* – in emergency it is a common situation, when obtained information could be incomplete or even contradictory. An intellectual approach is required for solving such conflicts.
- *Human factor* – experts situation estimation is an important information which is difficult to express in mathematical form in order to include into system estimation

A way to solve this problem greatly depends on type of emergency and can vary, depending on characteristics of geo-informational system, where such decision system is implemented. A challenging task is development of generalized self-adopting system, which can automatically configure its internal structure to fit current situation and requirements.

In other words the following ideas should be implemented in such system:

- Modular system of mathematical models – a framework, designed for describing peculiar properties of emergencies in region, their kind, type, and risk factors
- Map server – a common component of any geo-informational system to provide actual data of current relief, state of different objects on map (geo-information stations, rivers, dams, etc. – any object which can dramatically affect current state of situation)
- Decision adaptation system – tool, allowing quick adaptation of previous solutions to current situations, estimation of its effectiveness and quality.
- User interface subsystem – a way of visualization and obtaining information

Task formalization

Risk minimization system task can be expressed in a way, described in (1).

$$L(Sit, t) = \min F(X, R, A) \quad (1)$$

In this formula we introduce “losses function”, which depends on current situation Sit , and period of time t for taking decision. Amount of losses depends on our “actions function” F describing our actions depending on current state of nature environment X , available resources R , and database of previous situations A .

Current state of nature environment is presented as superposition of quantitative parameters (such as speed of the wind, angle of slope, etc.) and qualitative (weather state, type of surface, etc.). Available resources are expressed in form of triplets (*Type, Amount, Availability*) – which accordingly describe type of resource, its amount. Third parameter (availability) is intended to describe a risk of resource loss during emergency. For example after snow avalanche an automobile roads could be destroyed rendering safety cars useless, and making helicopters the only way to access avalanche site.

Database of previous situations in this formula should be interpreted not only as storage of data, but also a system for access and providing compare function to search for nearest to current situation in database. That could be implemented as ontology of situations or knowledge base, gathering required information through data mining process.

In order to decrease complexity of solving task, following assumptions are made to minimization function:

1. Each situation in database could be compared with current situation, and a numeric value – “distance” between situation could be calculated. There are no situations of different types in database (in other words databases of solutions for different kinds of situations are separated);
2. Qualitative parameters of situation could be expressed as a value from limited subset. No human language descriptions are supported;
3. Available resources are sufficient for handling any kind of emergency – there are always at least one solution of function which leads to decreased losses, comparable to other variants, and so called “zero variant” – amount of losses taken in case of no action;
4. Human losses coefficient is supposed to have a priority over material losses. No further comparison between two solutions is performed if one leads to increased human losses;
5. No pre-calculations are performed. System is not using stored data in order for synthetic decrease of function calculation time;
6. Amount of time after taking decision is sufficient for any of actions, proposed by system;
7. Recommended actions will be performed according to system recommendation. Effectiveness of such actions will be estimated by experts in order to obtain quality coefficient of current solution (learning procedure).

Such assumptions allows decreasing available space of possible solutions almost two times, leading to no significant reduce in risk estimation effectiveness. The function result is a numerical value describing losses factor and actions which should be taken to obtain this result.

The quality coefficient is added to a system when action was taken and it is an integral value, describing experts' estimation of system effectiveness.

Functional Approach to Risk Calculation

In computer science, functional programming is a programming paradigm that treats computation as the evaluation of mathematical functions and avoids state and mutable data. It emphasizes the application of functions, in contrast to the imperative programming style, which emphasizes changes in state. Functional programming has its roots in the lambda calculus, a formal system developed in the 1930s to investigate function definition, function application, and recursion. Many functional programming languages can be viewed as embellishments to the lambda calculus.

Functional approach in any situation handling interprets change of nature environment state (for example avalanche fall) as a result of function, depending on current parameters, and former value of this function [Hudak, 2006]. We interpret change of natural environment state as a result of low-level processes, which could be accumulated over time (2.)

$$R(fSit, t) = r(Sit) * R(fSit_{t-1}, t - 1) \quad (2)$$

It is important to realize that * in this formula should be interpreted not as common multiplication, but as an additional function taking two parameters, and producing a set of numerical value, which is expressed as risk of emergency, and internal data not visible to user describing previous states of function.

fSit - is a high-level function, describing current situation change. This function is a key element of risk estimation and it is composed under following rules:

1. Each parameter of nature environment is integrated in function one or more times;
2. Each qualitative parameter is integrated into function only combined with other qualitative parameters or as an argument in predicate form;
3. Each part of function could be evaluated independently and do not affect evaluation of other parts.

A source of creating such function could be different. In this paper for testing purposes a simple predicate system was created, based on automated analysis of statistical data. Any kind of low-level functions can be used, including data, obtained from regression and correlation analysis, experts' opinion, etc. This data are integrated into high-level function in such way they form a weighted graph for risk estimation. The functioning of this high-level function greatly depends on weight of this graph. In order to improve performance of system different parts of it are evaluated simultaneously, with no guarantee of computational order (this is safe because of rule 3). Irrelevant points of graph would not be evaluated; also in case of conflicting data a point with fewer sums of weights attached to it is eliminated.

Lazy calculations allow an effective usage of conflicting data. For example if we have a conflicting node in our graph, which describes a snowfall of grade 3, which is marked in conflict with node "snowfall – grade 5", and an evaluation functions takes a predicate "if it was snowing?" as argument – weights of this nodes will be summarized in final calculations since there is no conflict in appearance of snow and conflict evaluation would not be performed.

Additional data and functions could easily expand working model on-the-fly without any requirements of previous data recalculation.

Testing technique

In order to test effectiveness of proposed system a subset of mudflows in Carpathian region in period from June 2006 up to July 2008 was analyzed. Nature environment data are obtained both from geo-informational stations logs from that period, official meteorological data and records of commissions worked on that mudflows

elimination. Due to lack of information about available resources, the primary aim of a check was to compare existing methods of mudflows analysis (implemented in GIS system of Ukraine Ministry of emergency situations) and proposed functional method in terms of time, required for prognosis and its effectiveness.

A limited set of previous situations (about 1000) were passed to a system in order to obtain elementary predicate rules for possible mudflow. These rules were integrated into system with equal coefficients and a learning process was performed. A learning was performed on different sets of previous situations (from 100 up to 1000) in order to show the effectiveness of self-refining functional system.

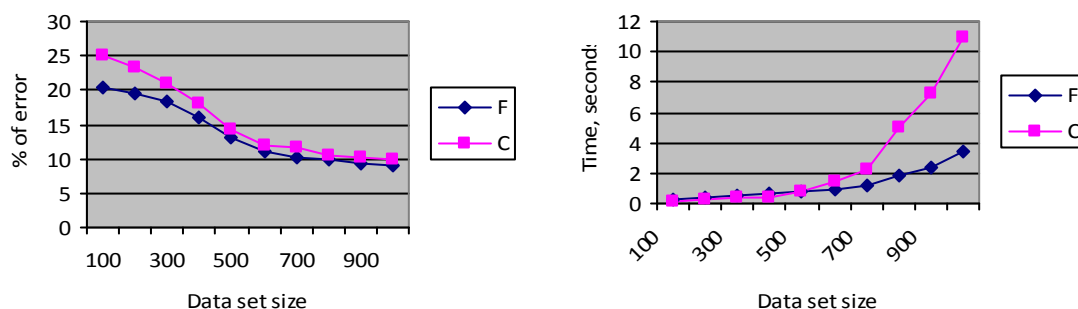


Fig. 1. Test results

On the first graph (see Fig. 1) we can see a percentage of error prognosis for implemented in governmental GIS system algorithm (marked “C” in graph) and our functional system (marked “F”). As it seen from the graph, the precision of prognosis on big sets (more than 500 situations) are almost equal. The situation with time of prognosis is significantly different. We can see almost a clear exponential grow for “C” algorithm, when grow of proposed functional algorithm after initial complexity change is almost linear, as expected.

According to statistical data [Ishchuk, 2002] an average amount of situations taken into account is about 5000 and process of taking decision usually takes up to 10-12 minutes, depending on amount of available data. Implementation of functional approach allows to significantly decreasing this value without loss of prognosis precision.

Modeling of decision taken was performed on snow avalanche data in Kirgizia during the period 2003-2009. Due to small amount of real avalanches in this period it is not possible to build a mathematical model to check system effectiveness. The main task of the system was proposal of resources and actions for avalanche controlling. A system solutions were effective (have an error less than 10%, compared with real situation) in 89% of situations.

However, testing revealed several problems, which should be solved before real system implementation:

1. A problem converting data from different sensors to one numerical format, suitable for processing with high-level functions;
2. Problem of building an effective high-level functions, describing complex emergency processes, for example earthquakes and water flows – system is quite effective analyzing big sets of similar data, but it’s speed decrease greatly on growing complexity of mathematical model
3. An effective algorithm for checking high-level functional models should be developed in order to test different function variations against real data.

Further improvements of an algorithm should allow benefiting from already clustered data about situations and providing a way of obtaining generalized rules for similar emergency in different regions all over the world.

Conclusion

The experimental check demonstrates that functional approach to emergency risk minimization is an effective solution for taking decision in limited time. The speed of proposed algorithm combined with lazy calculations, which allows greatly decrease time for making model less complex, allows performing just-in-time estimations of current nature state in order to perform a correct prognosis. However, the main disadvantages of this approach are problems of implementation and creation of basis functions in order to fully describe situation development. Creation of such mathematical tool should allow rapid increase in complexity of modern risk estimation systems.

Acknowledgements

The paper is partially financed by the project **ITHEA XXI** of the Institute of Information Theories and Applications FOI ITHEA and the Consortium FOI Bulgaria (www.ithea.org, www.foibg.com).

Bibliography

- [Hudak, 2006] Conception, evolution, and application of functional programming languages. ACM Computing Surveys 21 (3): 359–411, 2006
- [Ishchuk, 2002] Іщук О.О., Ободовський О.Г., Коноваленко О.С. Взаємодія ГІС та проблемно-орієнтованих моделюючих комплексів в системах прогнозування та оцінки наслідків надзвичайних ситуацій, пов'язаних з паводками // Науковий збірник КГУ «Гідрологія, гідрохімія, гідроекологія», 3т, 2002 р. –С.53- 59.

Authors' Information

Kuzomin Oleksandr – Professor, Ph.D., Chief of information marketing department. Kharkiv National University of Radio Electronics, Ukraine, Kharkiv, Lenina pr. 14, e-mail: kuzy@kture.kharkov.ua

Illya Klymov – Postgraduate (informatics chair), Kharkiv National University of Radio Electronics, Ukraine, Kharkiv, s. Mir 118, ap. 55, e-mail: illya.klymov@gmail.com

ЭВОЛЮЦИОННАЯ ПАРАДИГМА КАК УНИВЕРСАЛЬНЫЙ ИНТЕГРИРУЮЩИЙ ЭЛЕМЕНТ МЕТОДОЛОГИИ ПРОГНОЗИРОВАНИЯ

Виталий Снитюк

Аннотация: В статье рассмотрены составляющие элементы методологии прогнозирования, в основе которой лежит эволюционная парадигма. Рассмотрены объективные и субъективные аспекты уменьшения неопределенности, а также уровневая схема процессов принятия решений. Показано, что решение, принятое на каждом из низших уровней, позволяет уменьшить неопределенность, вызванную отсутствием данных, их неполнотой, а также неоднозначностью на следующих этапах. Определены аспекты применения методологии при создании сложных технических систем и программировании их жизненного цикла.

Ключевые слова: эволюционная парадигма, прогнозируемые, сложные системы

ACM Classification Keywords: [H.4 Information Systems Applications](#)

Conference: The paper is selected from XVth International Conference "Knowledge-Dialogue-Solution" KDS 2009, Varna, Bulgaria, June-July 2009

Введение

Процессы современного мира отличаются стремительной динамикой с присутствием качественных изменений, которые, в большинстве случаев, можно назвать катастрофами духовной или материальной субстанции. Существующие сложные взаимозависимости объектов естественной и искусственной сред определяют необходимость прогнозирования будущих процессов как основы рациональной деятельности, в т.ч. и в настоящем. Значительное количество научных работ посвящено задачам прогнозирования, моделям и методам их решения [Zgurovsky, 2007; Snytyuk, 2008]. Выбор той или иной модели и метода определяется задачами структурной и параметрической идентификации, а также структурой и мощностью множества исходных данных, желаемым типом и качеством результата.

Применение большинства технологий прогнозирования связано с необходимостью проверки большого количества требований и ограничений. Предложим технологию, одним из основных аспектов которой является минимальное количество таких требований и ограничений. Ее составляющими являются:

- уровневая схема принятия решений в условиях неопределенности;
- эволюционное моделирование как интегрирующий элемент.

1. Технология прогнозирования как уровневый процесс уменьшения неопределенности

Общая схема реализации такой технологии приведена на рис. 1. Известно, что у истоков создания сложных технических систем, которыми являются промышленные предприятия, сложные технические объекты, комплексные системы автоматизации объектов и процессов, информационно-аналитические системы и многие другие находится идея либо необходимость. Учитывая ресурсную ограниченность современного мира, необходимость имеет место в подавляющем большинстве случаев. Таким образом, релевантный процесс имеет следующую линию развития.

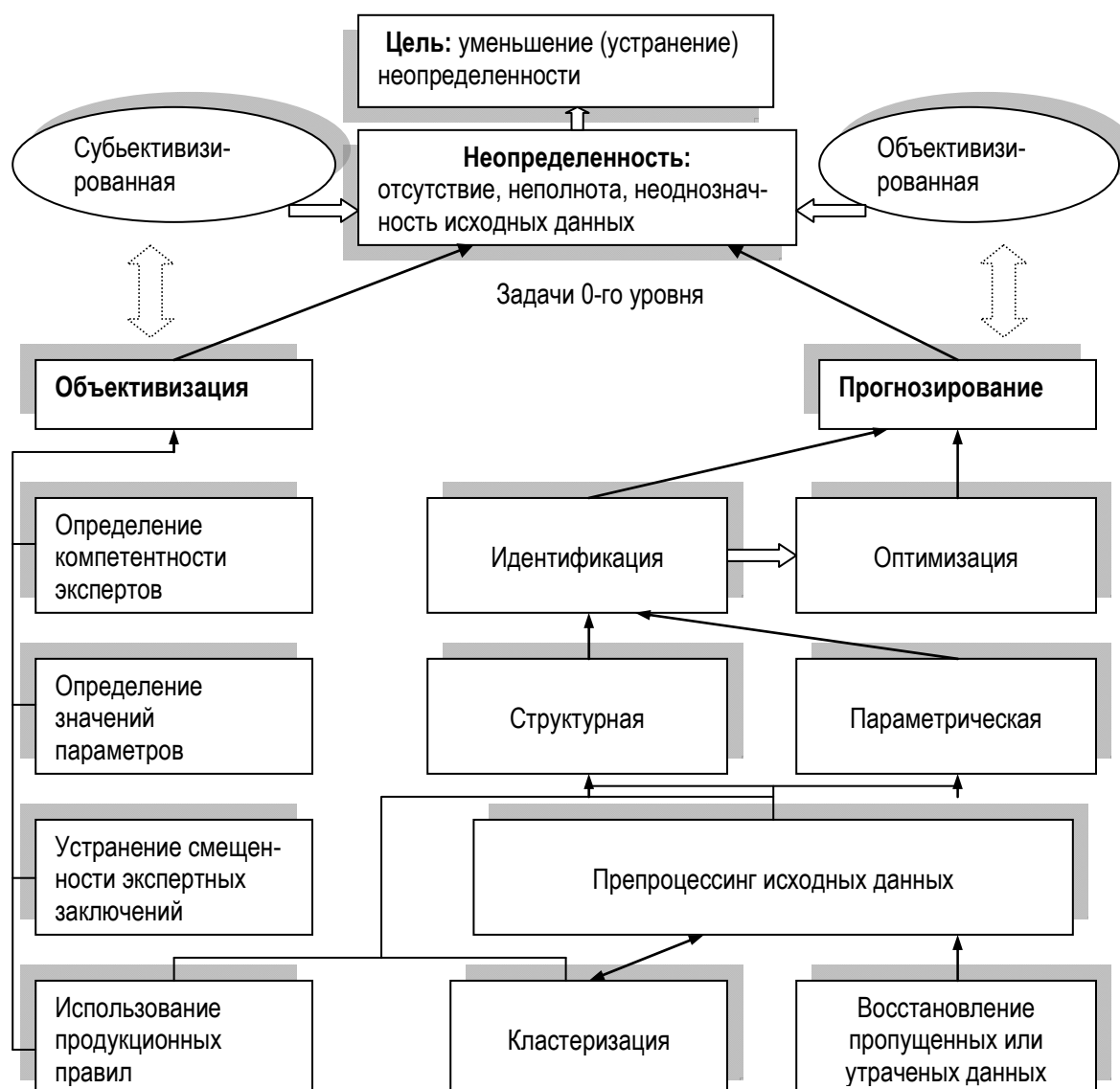


Рис. 1. Уровневая система уменьшения неопределенности

Возникновение необходимости создания системы двойственно определению главной цели функционирования такой системы. Достижение главной цели сопровождается реализацией множества второстепенных целей, причем среди них содержатся как цели, сопутствующие достижению главной цели, так и цели, препятствующие ее достижению. Реализация последних определяется несознательным антагонистическим влиянием природы и сознательным противодействием отдельных людей и общества в целом. Кроме того, важно учитывать динамику процессов в окружающей среде и их влияние на жизненный цикл создаваемой системы.

Очевидно, что процессы движения сложной системы по этапам жизненного цикла имеют прямой и обратный аспекты. В первом случае имеют место следующие преобразования:

$$\langle Z, M_0, M_e, A, R, T, K \rangle \rightarrow E, \quad (1)$$

$$\langle K_{i-1}, E_{i-1}, \theta, \xi \rangle \rightarrow E_i, \quad (2)$$

$$\langle K_{i-1}, E_{i-1} \rangle \rightarrow \langle A_i, R_i, M_{o_i}, M_{e_i} \rangle, \quad (3)$$

где Z – задача, M_o – соответствующая модель, M_e – метод решения задачи, A – исходные данные, R – ограничения, T – решающие процедуры, K – критерий оценки решения, E – решение задачи, E_{i-1} – решение задачи на $(i-1)$ -м этапе жизненного цикла, E_i – решение задачи на k -м этапе, K_{i-1} – критерий оценки решения задачи $(i-1)$ -го этапа, $A_i, R_i, M_{o_i}, M_{e_i}$ – исходные данные, ограничения, модели и методы решения задачи на i -м этапе, соответственно, θ – неучтенные факторы, ξ – случайные воздействия.

Исходя из (1)-(3), делаем вывод о том, что решения задач $(i-1)$ -го этапа и сопутствующие элементы являются исходными данными и ограничениями для задач i -го этапа.

Очевидно, что приобретаемый опыт изготовления, использования и модернизации сложных технических систем выражается в необходимости внесения изменений на предыдущих этапах жизненного цикла. Имеет место желаемое преобразование:

$$\langle E_{i+1}, \dots, E_{i+p-1}, E_{i+p} \rangle \rightarrow E_i, p \geq 1, p \in Z. \quad (4)$$

Получаем противоречие, поскольку осуществить преобразования (1)-(3) и (4) одновременно невозможно. Такое противоречие порождает неопределенность процессов принятия решений при создании сложных технических систем. Уменьшить ее можно с помощью доопределения исходной информации, что позволяет решать задачи (1)-(3), и прогнозирования, что позволит заменить «виртуальное» преобразование (4). Доопределение информации может быть осуществлено с помощью экспертных заключений и здесь сразу же возникает необходимость их объективизации или прогнозирования с использованием методов как интегро-дифференциальной, так и предлагаемой эволюционной парадигмы. Отметим, что под парадигмой мы понимаем идеи и принципы, позволяющие на единой методологической основе решать научно-технические проблемы. В данном случае (эволюционная парадигма) такой основой являются принципы эволюции, а также эволюционные модели и методы.

Вернемся к иерархической схеме уменьшения неопределенности (рис. 1), к ее нижнему уровню. Очевидно, что для решения задачи прогнозирования необходимо иметь статистическую базу, которую в данном случае представляет ретроспективная информация о прототипах, сосредоточенная в таблицах типа «объект-свойство». Оставляя проблемы, связанные с формированием таких таблиц и качественным составом информации, заметим, что в большинстве случаев, данные будут иметь пропуски, связанные с невозможностью измерений, потерями при хранении и т.п. Не учитывать такую информацию с пропусками означает потерять данные об объектах, значениях совокупности параметров и выходных характеристик.

Поэтому, на первом этапе доопределения данных необходимо решить задачу восстановления информации. Известные подходы к ее решению делятся на несколько классов, определяемых структурой пропусков [Snytyuk, 2007]. Отметим главные проблемы, сопровождающие решение этой задачи при реализации эволюционной парадигмы. К ним относится способ представления потенциального решения и определения критериальной функции. И, если в первом случае потенциальное решение может состоять из множества всех пропусков, имеющих некоторую форму представления, то вторая проблема является отнюдь не такой тривиальной. Необходимо получить решение задачи

$$\min_{\substack{x \in X_\Omega \\ y \in Y_\Omega}} |F(X, Y, x, y) - F(X, Y, x^*, y^*)|, \quad (5)$$

где X – множество входных факторов, Y – множество результирующих характеристик, x^* и y^* – точные значения пропущенных данных, X_Ω, Y_Ω – генеральные совокупности, состоящие из потенциальных решений x и y . Поскольку значения x^* и y^* неизвестны, то определение целевой функции F является приоритетной задачей.

Отсутствие четкой структуры исходных данных и отказ от априорного предположения о виде функции F указывает на рациональность использования нейронных сетей как моделей, реализующих преобразование $X \rightarrow Y$, в качестве одного из элементов решения задачи (5). Выбор нейросети из ансамбля по критерию минимизации ошибки реализации отображения $X \rightarrow Y$ на контрольной последовательности на следующем этапе позволяет найти решение (5). Отметим, что поиск такого решения – эволюционный процесс, поскольку потенциальные решения, участвующие в обучении нейросети, определяются путем хоть и случайного, но целенаправленного поиска с использованием эволюционных алгоритмов. Таким образом, на первом уровне уменьшение неопределенности достигается тем, что таблица данных приобретает комплектный вид.

Необходимость решения второй задачи определяется соображениями, которые частично уже были указаны выше. Изменчивость процессов современного мира, стремительная их динамика ведут к структурным изменениям. Значительное количество факторов, которые оказывают влияние на результаты прогнозирования не учитываются. Эти два обстоятельства приводят к выводу о неадекватности построения искомым зависимостей на всем множестве исходных данных и необходимости сегментации области исследования. Невыполнение такого требования приводит к тому, что вместо вывода «на $[A, B]$ цвет белый, а на $[B, C]$ – черный» получим заключение о том, что «на $[A, C]$ цвет серый», не отображающее действительного состояния.

Эволюционная кластеризация должна приводить к получению кластеров примерно одинаковых объектов или процессов, где, как известно, задача идентификации может быть решена достаточно эффективно. Две сопровождающие решение проблемы аналогичны первой задаче. В данном случае потенциальные решения имеют однозначную структуру – множество центров кластеров. Целевая функция будет такой:

$$\sum_{i=1}^K \sum_{j=1}^{m_i} d(X_j, \bar{X}_i) \rightarrow \min, \quad (6)$$

где число кластеров K задано и $K \ll m$, m – количество объектов, m_i – количество объектов в i -м кластере, $\bar{X}_i, i = \overline{1, K}$ – среднее значение в кластере, $d(X_j, \bar{X}_i)$ – расстояние между объектом и центром кластера. Решением задачи (6) являются центры кластеров \bar{X}_i , которые могут находиться среди данных объектов, что является достаточно строгим условием, и могут быть любыми точками области исследования.

Особенностью эволюционной кластеризации [Snytyuk, 2006] является определение центров кластеров и прямой расчет целевой функции без осуществления дополнительных операций. Результат решения задачи кластеризации также позволяет существенно уменьшить неопределенность и повысить точность решения задачи идентификации.

Зависимости, существующие между входными факторами, неравномерность распределения их значений приводит к значительному смещению результатов прогнозирования, а также к проблемам вычислительного характера. Алгоритм, позволяющий найти решение указанной проблемы, сводится к следующим шагам. Вначале из совокупности входных факторов удаляют линейно зависимые, затем из тех факторов, что остались, удаляют нелинейно зависимые факторы, применяя методику «box-counting» и вычисление энтропии.

Значения оставшихся независимых факторов чаще всего имеют неравномерное распределение и поэтому их использование для идентификации недостаточно информативно. Повысить информативность можно, используя метод «выбеливания» входов, результатом которого являются некоррелируемые факторы с нулевым средним значением и единичной дисперсией. Если количество факторов осталось сравнительно большим, используем аналог метода главных компонент, который реализуем с помощью автоассоциативной нейросети, в результате применения которого уменьшается количество факторов и

количество вычислений. Осуществление вышеуказанных операций также приводит к уменьшению неопределенности, поскольку уменьшаются вероятность неверного прогноза, «паралича» нейросети, плохой обусловленности матрицы исходных факторов и другие.

Решения описанных задач создают необходимый информационный базис для решения задач следующего уровня, а именно структурной, параметрической идентификации и оптимизации. Заметим, что и структурная, и параметрическая идентификация являются оптимизационными задачами, поэтому эволюционные технологии имеют место при решении каждой из них.

Как уже было указано выше, ретроспективная информация сосредоточена в таблицах данных. Никаких предположений о виде существующих зависимостей на начальном этапе не существует. В таких условиях рациональным является выбор в качестве соответствующих моделей нейронных сетей. Заметим, что выбор вида сети и ее архитектуры определяется задачей и исходными данными. Обучение сети приводит к получению искомой зависимости

$$F : X \rightarrow Y.$$

Она существует в «синтетической» форме, поскольку аналитический вид неизвестен (или достаточно сложный), и ее параметрами являются значения весовых коэффициентов. В то же время такая модель позволяет выполнять прогнозирование и решать другие задачи анализа данных.

2. Soft Computing метод решения задачи многофакторной оптимизации

Важной задачей, используемой при прогнозировании, является задача оптимизации. Рассмотрим новый подход и метод решения задачи нелинейной многофакторной оптимизации, которые базируются на использовании элементов теории вероятностей, теории нечетких множеств и эволюционного моделирования. Метод имеет такую пошаговую схему:

1. Положить $i = 1$. Определить генеральную \mathcal{E} и выборочную совокупности \mathcal{E}_i потенциальных решений.
2. Определить процентное соотношение p количества точек, которые переходят на следующий шаг поиска экстремума.
3. Вычислить значения функции, оптимум которой ищут, в точках выборочной совокупности.
4. Построить функцию принадлежности $\mu_{|y-y_{opt}|<\varepsilon}^i f(x)$, которая определяет меру уверенности в том, что решение-представитель x близко к оптимальному. Одним из параметров функции принадлежности, которая указывает на меру уверенности в этом, является h – высота. Определим также \tilde{h} – срез множества $D(f)$ как множество

$$M_h^i = \{x \in D(f) / \mu_{|y-y_{opt}|<\varepsilon}^i f(x) > h\}. \quad (7)$$

Построение функции принадлежности осуществляем с использованием элементов метода иерархий Т. Саати. Для этого нормируем значения функции f_i , $i = \overline{1, p}$. Пусть они находятся на отрезке $[0, 1]$. Формируем матрицу парных сравнений таким образом. Если нормированное значение $\tilde{f}_i \in [0, 1(k-1); 0, 1k]$, то $a_i = k$, $k \in \{1, 2, \dots, 10\}$. Тогда элемент матрицы парных сравнений $b_{ij} = \frac{a_i}{a_j}$. Достаточно определить эле-

менты одной строки этой матрицы. Остальные рассчитаем по формуле $b_{ij} = \frac{b_{qi}}{b_{qj}}$, где q – известная

строка. Тогда функция принадлежности рассчитывается так: $\mu(f_i) = 1 / (b_{1i} + b_{2i} + \dots + b_{pi})$.

5. С учетом шага 2, определяем множество точек $\{x_i\}$, которое принадлежит множеству M_h^i и для которых выполнено неравенство $|x_i - x_j| > \delta^*$, где δ^* – заданное положительное число.
 6. Для каждой точки из множества $\{x_i\}$ генерируем нормально распределенные последовательности $\{z_i^j\}$ с математическим ожиданием $Mz_i = x_i$ и среднеквадратичным отклонением σ_i . Заметим, что их общее количество равно количеству точек выборочной совокупности.
 7. Из точек множества $\{z_i^j\}$ формируем новую выборочную совокупность, допуская мутации каждого элемента с вероятностью P_m .
 8. Если выполняется критерий останова, то переход к шагу 9, иначе $i = i + 1$ и переход к шагу 3.
 9. Выполнение дополнительных процедур для уточнения оптимального решения.
- Такая композиция позволяет уменьшить временные затраты на случайный непроизводительный поиск.

3. Технология объективизации процессов прогнозирования

Одной из первых задач объективизации процессов прогнозирования является определение компетентности экспертов. Положим в их основу аксиому несмещенности: "суждение большинства компетентно" и ее следствие о том, что наиболее компетентен тот эксперт, суждение которого в большинстве случаев совпало с выводами большинства экспертов.

Формально задачу можно сформулировать так. Пусть n – количество экспертов, m – количество вопросов, причем $m = \sum_{i=1}^q m_i$, где m_i – количество вопросов i -го типа, $i = \overline{1, q}$, соответственно вышеизложенной классификации. Необходимо определить равные компетентности экспертов $\gamma_i, j = \overline{1, n}$.

Сущность метода состоит в определении элементов матриц, которые содержат значения близости суждений экспертов, их анализе и преобразованиях, в результате которых будут определены равные компетентности экспертов. Компетентность экспертов как результат выполнения шага 6 вычисляются по формуле

$$\gamma_p = \sum_{\substack{i, j=1 \\ i > j \\ (j=p) \vee (i=p)}}^n t_{ij} / \sum_{p=1}^n \sum_{\substack{i, j=1 \\ i > j \\ (j=p) \vee (i=p)}}^n t_{ij}, \quad p = \overline{1, n},$$

где t_{ij} – рассчитанное значение близости ответов i -го и j -го экспертов.

Непосредственно с задачей определения компетентности экспертов связана и задача установления значения определенного параметра системы в условиях неопределенности. Пусть α – неизвестное значение. Каждый эксперт $E_i, i = \overline{1, n}$ утверждает, что это значение равно α_i . Необходимо найти наиболее достоверную оценку значения α , если компетентности экспертов неизвестны.

Для каждого эксперта рассчитаем значения

$$\delta_j = \sum_{i=1}^n |\alpha_j - \alpha_i|, \quad j = \overline{1, n}.$$

Очевидно, что компетентность эксперта определяется так: $\gamma(E_j) = \frac{a}{\sum_{i=1}^n |\alpha_j - \alpha_i|}$, $\sum_{j=1}^n \gamma(E_j) = n$,

где a – некоторый параметр. Найдя значения a , получаем окончательное значение компетентности

$$\gamma(E_j) = \frac{n}{\sum_{j=1}^n \frac{1}{\sum_{i=1}^n |\alpha_j - \alpha_i|}} \cdot \sum_{i=1}^n |\alpha_j - \alpha_i|$$

Оценку неизвестного значения α рассчитываем как средневзвешенную по компетентностям величину оценок экспертов.

Очевидно, что значения компетентностей необходимо корректировать в зависимости от того, берет ли участие лицо, принимающее решение, в оценке компетентностей экспертов, какую роль сыграет его оценка и т.п. Очевидно, что субъективный фактор в проведении экспертиз приводит к смещенности экспертных суждений. Для того, чтобы избежать неправильных решений, лицо, принимающее решение, должно корректировать и адаптировать свои выводы к персональному составу экспертных комиссий и условий внешней среды. Модели, которые базируются на комбинации традиционных методов и методов теории нейронных сетей, указывают на один из путей решения этой проблемы. Важным также является применение процедуры корректирования экспертных заключений, применение которой разрешает нивелировать влияние субъективных факторов и находить близкие к оптимальным решения.

Заключение

В статье рассмотрены элементы технологии прогнозирования, которые базируются на эволюционной парадигме. Применение методов уменьшения неопределенности позволяет минимизировать субъективное влияние на процессы принятия решений и осуществлять доопределение отсутствующей информации, используя технологии прогнозирования.

Благодарности

Статья частично финансированна из проекта **ITHEA XXI** Института Информационных теории и Приложений FOI ITHEA и Консорциума FOI Bulgaria (www.ithea.org, www.foibg.com).

Список литературы

- [Zgurovsky, 2007] M.Z. Zgurovsky, N.D. Pankratova. System analysis: Theory and Applications. Springer. - 2007. - 475 p.
- [Snytyuk, 2008] V.E. Snytyuk. Forecasting. Models, methods, algorithms. – Kyiv, Maclaut, 2008. – 364 p.
- [Snytyuk, 2007] V.E. Snytyuk, S.O. Govorukhin. Evolutionary Technique for Filling Data Gaps in Outcome Variable Values Problems of market formation of research production. – Kharkiv, KHPI, 2007. – № 7. – Pp. 44-53.
- [Snytyuk, 2006] V.E. Snytyuk. Evolutionary clustering of complex systems and processes. – Information Theories and Applications. – 2006. – Vol. 13. – № 4. – Pp. 344-349.

Информация об авторе

Виталий Снитюк – заведующий кафедрой информационных технологий проектирования, Черкасский государственный технологический университет, бул. Шевченко 460, Черкассы 18006, Украина;
e-mail: snytyuk@gmail.com

ПОСТРОЕНИЕ ИНТЕГРАЛЬНЫХ ПОКАЗАТЕЛЕЙ ОЦЕНКИ РЕЗУЛЬТАТИВНОСТИ НАУЧНЫХ ПРОЕКТОВ

Алексей Петровский, Григорий Ройзензон, Игорь Тихонов

Аннотация: Изложен подход к агрегированию признаков и построению интегральных показателей оценки, основанный на интерактивном методе порядковой классификации многокритериальных альтернатив с последовательным снижением размерности признакового пространства. Представлен пример применения подхода для многокритериальной оценки результативности научных проектов.

Ключевые слова: вербальный анализ решений, многокритериальная порядковая классификация, снижение размерности признакового пространства, стратификация кортежей, агрегирование признаков, интегральные показатели оценки, результативность научных проектов

Conference: The paper is selected from XVth International Conference "Knowledge-Dialogue-Solution" KDS 2009, Varna, Bulgaria, June-July 2009

Введение

Непосредственная классификация или сравнение альтернатив, описываемых большим числом признаков, и в особенности качественных (вербальных) признаков, является трудоемкой процедурой, которая требует значительных временных затрат лица, принимающего решение (ЛПР), и нередко существенно затрудняет применение на практике известных методов принятия решений [Ларичев, 2006]. Кроме того, когда сравниваемых объектов мало (3-5), а их признаки различны по значениям и многочисленны (десятки и сотни), такие объекты, как правило, оказываются формально несравнимыми по своим свойствам. Эти обстоятельства диктуют необходимость разработки специальных методов обработки информации, обеспечивающих решение задач многокритериального выбора и классификации в пространствах большой размерности [Айвазян и др., 1989], [Глотов, Павельев, 1984], [Петровский, Ройзензон, 2008а,б].

В настоящей работе представлены интерактивный метод последовательного снижения размерности признакового пространства и результаты его использования для разработки интегральных показателей оценки результативности научных проектов, выполняемых в интересах Федеральных агентств и ведомств России по грантам Российского фонда фундаментальных исследований. Результативность научных проектов может рассматриваться как с позиций непосредственного применения полученных результатов на практике, так и для целевого формирования исследовательских программ. Построение интегрального показателя результативности научного проекта сведено к задаче многокритериальной порядковой классификации альтернатив по иерархической системе критериев. В качестве многопризнаковых объектов выступают комбинации исходных экспертных оценок проектов по принятым в Фонде критериям, градации интегрального показателя результативности играют роль классов решений. Применение процедуры иерархического агрегирования многокритериальных экспертных оценок проектов позволяет значительно снизить размерность исходного признакового пространства, что существенно сокращает время, затраченное ЛПР, для достижения поставленной цели. Предложенный подход к формированию интегральных показателей облегчает анализ и обоснование выбора окончательного варианта решения.

Снижение размерности признакового пространства

Формальная постановка задачи снижения размерности признакового пространства имеет следующий вид:

$$X_1 \times \dots \times X_m \rightarrow Y_1 \times \dots \times Y_n, n < m,$$

где X_1, \dots, X_m – исходный набор признаков, Y_1, \dots, Y_n – новый набор признаков, m – размерность исходного признакового пространства, n – размерность нового признакового пространства. Каждый из признаков имеет свою собственную шкалу $X_i = \{x_i^1, \dots, x_i^p\}$, $i=1, \dots, m$, $Y_j = \{y_j^1, \dots, y_j^h\}$, $j=1, \dots, n$ с упорядоченной градацией оценок.

Рассмотрим задачу снижения размерности признакового пространства как задачу многокритериальной классификации, в которой различные комбинации исходных признаков (кортежи оценок) последовательно агрегируются в меньшие наборы новых признаков, имеющих для ЛПР вполне определенный смысл. Итогом является иерархическая система критериев, верхний уровень которой определяется содержанием практической проблемы.

Назовем *составным критерием* интегральный показатель, который определяет выбранное ЛПР свойство вариантов, агрегирующее исходные характеристики. Каждая градация шкалы составного критерия является комбинацией оценок исходных показателей.

Процедура агрегирования показателей является многоуровневой иерархической структурой со «слабыми» связями, в которой элемент нижележащего уровня (оценки исходных показателей) подчинен двум и более вершинам вышестоящего уровня (оценкам составных критериев). Переходя шаг за шагом на более высокий уровень иерархии, ЛПР может сконструировать приемлемые для него составные критерии вплоть до одного единственного.

Представим процедуру построения шкал составных критериев в виде однотипных блоков. Блоки содержательно выделяются ЛПР в зависимости от специфики решаемой задачи. Каждый блок классификации i -го уровня иерархии состоит из некоторого набора признаков и одного составного критерия. В качестве объектов классификации выступают все градации оценок на шкалах признаков. Классами решений i -го уровня служат градации оценок на шкале составного критерия.

В блоке классификации $(i+1)$ -го уровня иерархии составные критерии i -го уровня считаются признаками, множество градаций оценок которых представляет собой новые объекты классификации в сокращенном признаковом пространстве, а классами решений будут теперь градации оценок на шкале составного критерия $(i+1)$ -го уровня. Процедура повторяется до тех пор, пока не останется единственный составной критерий верхнего уровня, являющийся искомым интегральным показателем, шкала оценок которого образует упорядоченные классы решений D_1, \dots, D_q . Тем самым устанавливается соответствие между классами решений D_1, \dots, D_q и совокупностью исходных показателей – множеством X_1, \dots, X_m всех возможных комбинаций градаций оценок на шкалах критериев $X_i = \{x_i^1, \dots, x_i^p\}$, $i=1, \dots, m$, и находятся границы классов, что позволяет легко построить классификацию реальных альтернатив (вариантов) V_1, \dots, V_p , оцененных по многим критериям K_1, \dots, K_m .

Для формирования шкал оценок по составным критериям ЛПР может воспользоваться несколькими способами из арсенала средств вербального анализа решений. Наиболее простым и легко воспринимаемым ЛПР способом конструирования порядковой шкалы составного критерия является метод стратификации кортежей, в котором используются однотипные (например, с одинаковым числом градаций) наборы порядковых вербальных шкал исходных показателей. Идея метода стратификации кортежей основана на сечении многомерного дискретного признакового пространства параллельными гиперплоскостями. Каждый слой (страта) состоит из однотипных комбинаций оценок на шкалах критериев X_i , а число таких сечений (оценок) определяется ЛПР из содержательных соображений. Максимально

возможное число слоев можно рассчитать по формуле $L=1-m+\sum_{i=1}^m g_i$. Каждый слой образуется как комбинация кортежей оценок, сумма которых фиксирована. Число классов $q \leq L$. Более сложные процедуры построения шкал критериев предполагают применение методов вербального анализа решений ЗАПРОС и ОРКЛАСС [Ларичев, 2006], в которых рассматривается множество всех возможных кортежей оценок в признаковом пространстве, образованном декартовым произведением значений оценок на шкалах критериев. В этих случаях число возможных комбинаций оценок (альтернатив) равно $p = \prod_{i=1}^m g_i$.

Рассмотрим небольшой пример. Пусть ЛПР необходимо построить шкалу составного критерия D , имеющую три вербальные порядковые градации $\{d^0, d^1, d^2\}$. Шкала составного критерия образуется из шкал оценок базовых показателей, каждая из которых также имеет три вербальные порядковые градации: $A=\{a_0, a_1, a_2\}$, $B=\{b_0, b_1, b_2\}$ и $C=\{c_0, c_1, c_2\}$. Например, x_0 – отлично, x_1 – хорошо, x_2 – плохо.

Применяя для построения порядковой шкалы составного критерия метод стратификации кортежей, ЛПР может объединить исходные оценки в обобщенную оценку, например, по принципу: все лучшие оценки по базовым показателям образуют одну лучшую оценку по составному критерию, все средние оценки – средние оценки, все худшие оценки – одну худшую оценку (рис. 1).



Рис. 1. Конструирование шкалы составного критерия с помощью метода стратификации кортежей.

Воспользовавшись методом ЗАПРОС, ЛПР может сконструировать единую порядковую шкалу составного критерия, формируя ее из оценок по отдельным исходным показателям (рис. 2).



Рис. 2. Конструирование шкалы составного критерия методом ЗАПРОС.

С помощью метода ОРКЛАСС строится полная непротиворечивая порядковая классификация многопризнаковых объектов, в качестве которых в рассматриваемом случае выступают наборы оценок по исходным показателям, образующим составной критерий (рис. 3).

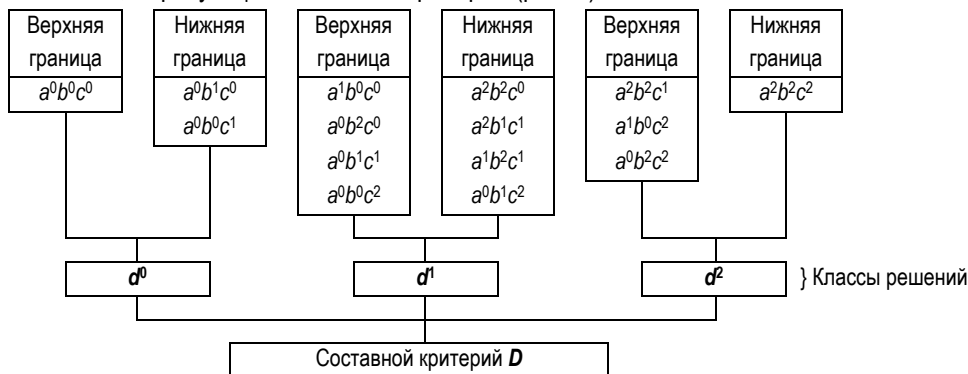


Рис. 3. Конструирование шкалы составного критерия методом ОРКЛАСС.

Предложенный новый методологический подход к снижению размерности пространства качественных признаков обладает определенной универсальностью, так как в общем случае можно оперировать как символьной (качественной), так числовой информацией.

Интерактивный метод снижения размерности признакового пространства

Схема решения задачи многокритериального выбора с использованием снижения размерности признакового пространства включает два этапа [Петровский, Ройзензон, 2008а]. На первом этапе проводится снижение размерности признакового пространства путем построения иерархической системы составных критериев. На втором этапе выполняется окончательное решение задачи выбора с использованием построенных составных (агрегированных) критериев.

Агрегирование признаков основано на предпочтениях ЛПР. Первоначально при участии ЛПР формируется базовый набор характеристик рассматриваемых объектов. В зависимости от специфики задачи эти характеристики могут быть либо заданы заранее, либо сформированы в процессе анализа проблемы. Для каждого базового показателя формируется шкала, которая может иметь числовые (точные, интервальные) или вербальные оценки. Шкалы оценок базовых показателей могут совпадать с обычно используемыми на практике, либо конструироваться специально.

Далее, исходя из опыта и интуиции ЛПР, базовые характеристики объединяются в критерии, обладающие вербальными порядковыми шкалами с небольшим числом градаций (3-5). ЛПР, по своему усмотрению, определяет число, состав и содержание критериев каждого уровня иерархии. В качестве критерия можно выбрать один из базовых показателей или несколько показателей, объединенных в составной критерий. ЛПР устанавливает, какие базовые показатели будут считаться самостоятельными критериями, а какие будут отнесены к тому или иному составному критерию. Смысловое содержание критериев и шкал оценок задается ЛПР. Критерии должны иметь такие шкалы оценок, которые, с одной стороны, будут отражать агрегированные качества объектов, а с другой стороны, будут понятны ЛПР при окончательном выборе наиболее предпочтительных объектов или их классификации.

Процедура агрегирования показателей имеет последовательный характер, т.е. полученные группы критериев могут быть, в свою очередь, объединены в новые группы (следующий уровень иерархии) и так далее. При конструировании шкал составных критериев на разных этапах могут использоваться различные способы. Например, один из составных критериев можно сформировать при помощи метода стратификации кортежей, а другой – методом многокритериальной порядковой классификации.

В зависимости от специфики задачи выбора иерархическая система критериев может быть известна заранее (например, перечень показателей организационной структуры предприятия), известна частично (например, характеристики сложной технической системы) и неизвестна вообще, т.е. иерархию требуется разработать «с нуля» (например, критерии оценки научных исследований, где присутствует высокая степень неопределенности и риска, связанная с получением нового знания). При построении системы критериев в первом случае основное внимание должно быть уделено определению шкал составных критериев. Особенностью разработки системы критериев во втором и в третьем случаях является возможность сформировать разные наборы составных критериев различными способами (например, последовательно объединяя критерии попарно или формируя группы критериев исходя из некоторой смысловой общности). Это позволяет сравнить полученные результаты для разных вариантов классификации и выбора с целью оценки качества решения исходной проблемы.

Интерактивный метод ИСКРА (Иерархическая Структуризация КРитериев и Атрибутов) предназначен для снижения размерности признакового пространства. Различные комбинации признаков разного уровня

иерархии рассматриваются в методе как многопризнаковые объекты, последовательно агрегируемые в составные критерии на основе предпочтений ЛППР. Метод включает в себя следующие шаги:

1. Выбор типа задачи T . Возможны следующие задачи: T_1 – выделить наилучшие альтернативы; T_2 – упорядочить альтернативы; T_3 – разделить альтернативы на упорядоченные группы (классы).
2. Формирование множества альтернатив V в зависимости от типа задачи T . Очевидно, что если необходимо проранжировать или выбрать наилучшую альтернативу, то для T_1 и T_2 $|V| \geq 2$.
3. Формирование множества базовых показателей (исходных признаков) $X: X_1, \dots, X_m$. $|X|=m \geq 2$. В зависимости от специфики задачи эти характеристики могут быть либо заданы заранее, либо сформированы при исследовании проблемы с помощью аналитика-консультанта или эксперта.
4. Формирование порядковых шкал $X_i = \{x_i^1, \dots, x_i^g\}$, $i=1, \dots, m$, базовых показателей в зависимости от типа задачи T . Для задач ранжирования или выбора наилучшей альтернативы необходимо рассматривать только те оценки базовых показателей, которые встречаются в описании альтернатив V , т.е. для T_1 и T_2 $g_i \leq |V|$. Тем самым можно предварительно сократить размерность исходного признакового пространства. Если изначально не задан набор реальных альтернатив, необходимо рассматривать множество всех возможных кортежей оценок в признаковом пространстве, образованном декартовым произведением градаций оценок на шкалах критериев K_1, \dots, K_m .
5. Формирование множества составных критериев, т.е. интегральных показателей, которые определяют выбранное ЛППР свойство вариантов, агрегирующее базовые характеристики: $Y: Y_1, \dots, Y_n$. ($n < m$).
6. Формирование порядковых шкал $Y_j = \{y_j^1, \dots, y_j^{h_j}\}$, $j=1, \dots, n$ составных критериев. Каждая градация шкалы составного критерия является комбинацией оценок базовых показателей.
7. Выбор способа W построения шкал составных критериев (агрегирования показателей) $X_1 \times \dots \times X_m \rightarrow Y_1 \times \dots \times Y_n$, $n < m$. Возможны следующие способы: W_1 – стратификация кортежей; W_2 – ранжирование кортежей; W_3 – многокритериальная порядковая классификация кортежей. Для разных составных критериев можно использовать различные способы построения шкал.
8. Если получен удовлетворительный результат решения задачи T (по составным критериям построена порядковая классификация альтернатив; получено упорядочение альтернатив; выделены наилучшие альтернативы), то алгоритм завершает работу, иначе переход к шагу 9.
9. Если на шаге 8 получен неудовлетворительный результат, то предлагается либо изменить способ построения шкалы составного критерия W (переход к шагу 7), либо изменить градации шкалы составного критерия (переход к шагу 6), либо сформировать новое множество составных критериев Y (переход к шагу 5).

Важной особенностью предложенного подхода к снижению размерности признакового пространства является возможность его использования в сочетании с другими методами принятия решений и обработки информации, позволяя представить каждую градацию шкалы составного критерия в виде комбинации оценок базовых показателей.

Многокритериальная оценка результативности научных проектов

Существует достаточно много практических задач, в которых необходима разработка интегральных показателей деятельности. Типичным примером является оценка результативности научных проектов. Значительный опыт в организации и проведении экспертизы фундаментальных исследований и полученных результатов, представляющих интерес с точки зрения их практического применения,

накоплен в Российском фонде фундаментальных исследований (РФФИ). Для оценки возможностей эффективного практического использования научных результатов было формализовано понятие «результативность научного проекта» применительно к фундаментальным исследованиям, выполняемым в интересах Федеральных агентств и ведомств России по грантам РФФИ.

Построение интегрального показателя результативности научного проекта рассмотрена как задача многокритериальной порядковой классификации по иерархической системе критериев с использованием снижения размерности признакового пространства [Петровский, Ройзензон, 2008а]. Классифицируемыми объектами являлись комбинации многокритериальных экспертных оценок отчетов по законченным проектам, классами решений служили градации интегрального показателя результативности. ЛПР имело возможность формировать понятие «результативность научного проекта» различным образом и сравнивать полученные интегральные показатели.

Каждый отчет оценивался несколькими экспертами по восьми критериями (базовым признаками) K_1, \dots, K_8 , имеющими шкалы X_i с двумя или тремя вербальными порядковыми оценками 0,1,2, где 0 обозначает лучшую оценку, 1 – среднюю (или худшую), 2 – худшую. Например, критерий K_1 характеризовал «Степень выполнения заявленных задач», которая может оцениваться как 0 – задачи выполнены полностью, 1 – задачи выполнены частично, 2 – задачи не выполнены. Критерий K_3 оценивал «Достижение поставленной цели в установленные сроки» как 0 – реальное, 1 – нереальное. Составным критерием верхнего уровня являлась «Результативность научного проекта», градации оценок по шкале которого (высокая, хорошая, средняя, низкая, неудовлетворительная) определяли 5 упорядоченных классов решений D_1, \dots, D_5 .

Модельная база данных включала экспертные оценки отчетов по проектам фундаментальных исследований РФФИ, завершённым в 2007 году. Для экспериментальной проверки работы алгоритмов и построения интегральных показателей результативности были взяты отчеты по следующим областям знаний: 01. Математика, механика и информатика (всего 48 отчетов); 03. Химия (всего 54 отчета); 07. Информационные, вычислительные, телекоммуникационные ресурсы (всего 21 отчет).

Интегральные показатели результативности проектов определялись по каждой области знания и для каждого эксперта с помощью различных сочетаний методов принятия решений, а именно: методом ОРКЛАСС (ОР); методом стратификации кортежей (СК); сначала (на нижнем уровне иерархии) методом стратификации кортежей, потом (на верхнем уровне иерархии) методом ОРКЛАСС (СК+ОР); сначала (на нижнем уровне иерархии) методом ОРКЛАСС, потом (на верхнем уровне иерархии) методом стратификации кортежей (ОР+СК). Пример оценки результативности проектов, полученных разными методами по оценкам двух экспертов, приведен на рис. 4.

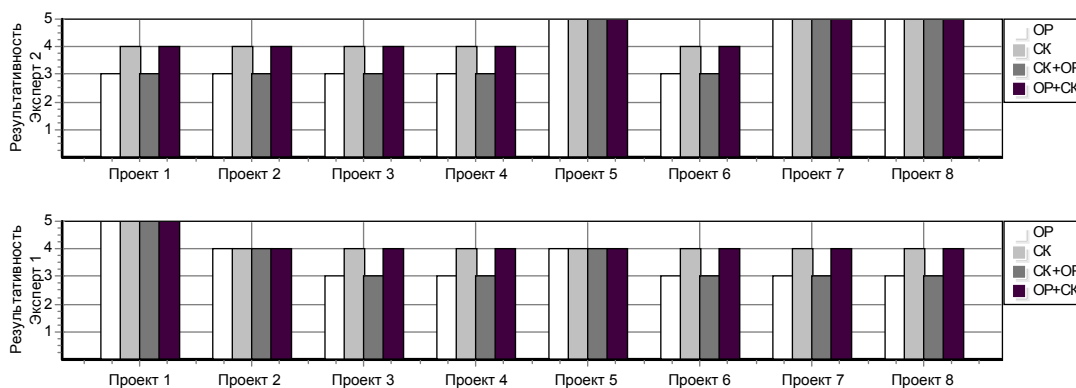


Рис. 4. Пример распределения альтернатив по классам решений.

Распределение проектов по классам результативности, полученное для четырех вышеперечисленных методов формирования шкал составных критериев, представлено в таблице 1. Первое число – число

проектов с оценками первого эксперта, второе число – число проектов с оценками второго эксперта. Значение интегрального показателя результативности совпадало у каждого из экспертов в следующем числе случаев: по области 01 (48 проектов) – для 34/23 проектов, по области 03 (54 проекта) – для 39/13 проектов, по области 07 (21 проект) – для 16/13 проектов, что также свидетельствует о достаточно высокой устойчивости результата относительно исходных данных и способа построения шкал составных критериев на всех уровнях иерархии.

Таблица 1. Распределение проектов по классам результативности

Область знания	Показатель результативности	Метод ОР	Метод СК	Метод СК+ОР	Метод ОР+СК
01. Математика, механика и информатика	Наивысший (5)	16/8	16/7	16/7	16/8
	Высокий (4)	15/14	28/34	14/15	28/35
	Средний (3)	17/26	4/4	18/25	4/4
	Низкий (2)	0/0	0/3	0/1	0/1
03. Химия	Наивысший (5)	4/3	4/3	4/2	4/2
	Высокий (4)	39/17	42/47	27/11	42/47
	Средний (3)	17/33	5/3	20/40	5/4
	Низкий (2)+(1)	3/1	3/2	3/1	3/1
07. Информационные, вычислительные, телекоммуникационные ресурсы	Наивысший (5)	6/4	6/4	6/4	6/4
	Высокий (4)	6/9	10/16	6/8	10/16
	Средний (3)	9/8	4/1	9/9	5/1
	Низкий (2)	0/0	1/0	0/0	0/0

Было также рассмотрено два варианта построения иерархической системы критериев: с агрегированием базовых оценок по критериям K_1, \dots, K_8 соответственно в два и три промежуточных критерия. Анализ распределения проектов для области 03 по показателю результативности исходя из их экспертных оценок показал следующее: наивысший показатель результативности получили по 1 варианту иерархической системы критериев 16 проектов, а по 2 варианту – 6; высокий показатель результативности – 75 и 40 проектов; средний показатель результативности – 13 и 59 проектов, низкий показатель результативности – 2 и 1 проекта, неудовлетворительный показатель результативности – 2 и 2 проекта.

Таким образом, 1 вариант агрегирования оценок по критериям (с выделением двух промежуточных критериев) дает более высокий интегральный показатель результативности, чем 2 вариант (с выделением трех промежуточных критериев). В целом по двум вариантам интегральные показатели результативности совпадают в 41 случае из 108 (54 проекта оценивались двумя экспертами). В остальных случаях интегральные показатели результативности отличались не более чем на одну градацию, что может рассматриваться как свидетельство достаточно высокой устойчивости результата относительно исходных данных и выбранных способов построения шкал составных критериев на всех уровнях иерархии.

Заключение

Использование метода ИСКРА при решении задач многокритериального выбора и классификации дает ЛПР возможность сравнить полученные решения для разных наборов составных критериев, сформированных с помощью различных способов. Оценкой эффективности выбранного способа решения задачи может служить число обращений к ЛПР, необходимых для построения полной непротиворечивой классификации для каждого набора составных критериев. Альтернативной оценкой эффективности

является сравнение распределений альтернатив по классам решений для одного и того же набора составных критериев, сформированных различным образом.

При решении конкретной прикладной задачи предложенный подход позволяет выбрать как наиболее предпочтительный набор составных критериев, так и метод или совокупность методов их построения. Последовательное разделение всех критериев на отдельные группы дает возможность «распараллелить» решение задачи, что обеспечивает ощутимую экономию трудозатрат. Применение процедуры иерархического агрегирования базовых показателей позволяет значительно снизить размерность исходного признакового пространства, что существенно сокращает время, затраченное ЛПП, для достижения поставленной цели.

Благодарности

Статья частично финансирована из проекта **ITHEA XXI** Института Информационных теории и Приложений FOI ITHEA и Консорциума FOI Bulgaria (www.ithea.org, www.foibg.com).

Список литературы

- [Айвазян и др., 1989] Айвазян С.А., Бухштабер В.М., Енюков И.С., Мешалкин Л.Д. Прикладная статистика. Классификация и снижение размерности / Под ред. С.А.Айвазяна. — М.: Финансы и статистика, 1989.
- [Глотов, Павельев, 1984] Глотов В.А., Павельев В.В. Векторная стратификация. — М.: Наука, 1984.
- [Ларичев, 2006] Ларичев О.И. Вербальный анализ решений / Под ред. А.Б.Петровского. — М.: Наука, 2006.
- [Петровский, Ройзензон, 2008а] Петровский А.Б., Ройзензон Г.В. Снижение размерности признакового пространства в задачах многокритериальной классификации // Decision Making and Business Intelligence Strategies and Techniques / Ed. by K. Markov, K. Ivanova, I. Mitov. — No. 3. — Sofia: FOI ITHEA, 2008. — Pp. 81–86.
- [Петровский, Ройзензон, 2008б] Петровский А.Б., Ройзензон Г.В. Интерактивная процедура снижения размерности признакового пространства в задачах многокритериальной классификации // Поддержка принятия решений: Труды Института системного анализа Российской академии наук / Под ред. А.Б.Петровского. — М.: Едиториал УРСС, 2008. — Т. 35. — С. 43–53.

Сведения об авторах

Петровский Алексей Борисович – д.т.н., заведующий лабораторией Института системного анализа РАН, Россия, Москва 117312, пр-т 60-летия Октября, 9, тел. (495)135-8503, e-mail: pab@isa.ru

Ройзензон Григорий Владимирович – к.т.н., научный сотрудник Института системного анализа РАН, Россия, Москва 117312, пр-т 60-летия Октября, 9, тел. (495)135-8503, e-mail: rgv@isa.ru

Тихонов Игорь Петрович – к.т.н., начальник отдела целевых фундаментальных исследований Российского фонда фундаментальных исследований, Россия, Москва 117334, Ленинский пр-кт, 32А, e-mail: tikhonov@rfbr.ru

АНАЛИЗ МАЛЫХ ВОЗМУЩЕНИЙ ЛИНЕЙНЫХ ЭКОНОМИКО-МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ

Алексей Волошин, Всеволод Богаенко, Владимир Кудин

Аннотация. Методы последовательного анализа вариантов и базисных матриц применяются для анализа влияний малых возмущений в модели Леонтьева с плохообусловленной матрицей ограничений. Приведены результаты вычислительных экспериментов.

Ключевые слова: модель Леонтьева, количественный и качественный анализ, базисная матрица.

ACM Classification Keywords: H.4.2 Information Systems Applications: Types of Systems: Decision Support.

Conference: The paper is selected from XVth International Conference "Knowledge-Dialogue-Solution" KDS 2009, Varna, Bulgaria, June-July 2009

Введение

подавляющее большинство эколого-экономических процессов описываются в классе линейных моделей. В частности, основная макроэкономическая модель, модель Леонтьева (МЛ) [Пономаренко, 2004], [Волошин, 2004] - в виде системы линейных алгебраических уравнений (СЛАУ) или неравенств (СЛАН) с квадратной матрицей с ограничениями на переменные в виде гиперпараллелепипеда P [Волошин, 2004]. Модель Леонтьева лежит также в основе ряда оптимизационных задач, в частности, линейного программирования (ЗЛП). [Волошин, 2008], [Кудин, 2002]. Известно, что:

1. Нечеткость значений параметров модели предопределяет наличие в контуре принятия решения экспертов (лиц, принимающих решение - ЛПР), которые направлены качественно очертить структуру и указать механизм (процедуру) устранения неопределенностей и разногласий при ее формировании (означить индивидуальные и коллективные функции принадлежности и т.п.) [Орловский, 1981]. Это является существенным усложнением МЛ.
2. Учет неточности представления модели (так называемая проблема адекватности математической и машинной модели) предопределяет разработку механизма согласования результатов проведения вычислений при различной точности представления модели (длина мантиссы при представлении чисел с плавающей запятой).
3. Некорректность (даже линейных моделей) при проведении вычислений с различной точностью может существенно повлиять на качественные характеристики, например, величину ранга (области принадлежности решений). В частности, свойство некорректности может проявиться как плохая обусловленность. В этом случае возникает проблема интерпретации результатов вычислений при различных входных параметрах (локализации области решений).
4. Представление нелинейного экономического процесса линейным приближением "добавляет неадекватности" при моделировании, а поэтому учет слабых нелинейных связей в элементах модели направлено это сгладить.
5. Исследование наряду с классической МЛ СЛАУ более обобщенной модели (СЛАН или ЗЛП) предопределяет разработку единого метода изучения свойств таких моделей.
6. Структурные связи в элементах моделей типа МЛ могут подаваться в классе линейных систем с прямоугольной матрицей ограничений, которая делает актуальным изучение геометрических образов приведенных моделей - многогранных множеств.

Постановка задачи

Предметом исследования являются следующие варианты МЛ:

$$Au = C, \quad (1)$$

$$u \in \Pi, \quad (2)$$

$$\text{и } Au \leq C, \quad (3)$$

при ограничениях (2), в частности, при наличии целевой функции

$$\max_{u \in R^m} Bu, \quad (4)$$

и ограничений (2)-(3), где $A = \{a_{ij}\}_{\substack{i=1,m \\ j=1,m}}$ - квадратная матрица размерности $(m \times m)$,

$a_j = (a_{j1}, a_{j2}, \dots, a_{jm})$, $j \in J = I = \{1, 2, \dots, m\}$, - строки матрицы A , $u = (u_1, u_2, \dots, u_m)^T$ - вектор переменных, $B = (b_1, b_2, \dots, b_m)$, $C = (c_1, c_2, \dots, c_n)^T$ - векторы градиента целевой функции и ограничений, $a_j u \leq c_j$, $j \in J$, - полупространство, определенное гиперплоскостью $a_j u = c_j$, $j \in J$.

Будут изучаться свойства МЛ с матрицами ограничений структурно близких к некорректным (плохо обусловленным) и более общие модели, в которых $n > m$. По своей природе (2)-(4) являются двойственными к канонической (прямой) задачи линейного программирования, а потому в рассмотрение вводится линейная (порождающая) модель вида (4):

$$Au \leq C, \quad (5)$$

при ограничениях (2). В частности, модель (5) может иметь вид

$$Au = C, \quad (6)$$

при ограничениях (2), где $B = (b_1, b_2, \dots, b_m)$, $C = (c_1, c_2, \dots, c_n)^T$, $u = (u_1, u_2, \dots, u_m)^T$, $a_j = (a_{j1}, a_{j2}, \dots, a_{jm})$, $j \in J$, - строка матрицы A . Задача вида (4), (5) имеет n ограничений и m переменных. Модель (4), (5) исследуется в пространстве E^m .

Целью исследований является экспериментальный анализ влияний малых изменений в модели с учетом приведенных выше факторов (при разной мере плохой обусловленности) на ее свойства (точность решения, величину невязок и прочее). Проверить эффективность процедур последовательного анализа вариантов (ПАВ) [Волошин, 1987] и метода базисных матриц (МБМ) [Кудин, 2002]. Исследовать свойства известных вычислительных схем (Гаусса, SVD-разложения, формул Гревилля, метода базисных матриц).

Концепция анализа МЛ является двухстадийной. Первая стадия содержит анализ "порождающей" модели (эталонной, математической), вторая - анализ "возмущенной" (неточной, нечеткой, машинной).

Рассматриваются эталонные модели (с известными или установленными свойствами) и слабо возмущенные (с различным уровнем точности и учетом нечеткости задания параметров или малых возмущений). Это ставит как начальную задачу анализ непротиворечивости структурных элементов: установление невырожденности и величины ранга матрицы (1), направленной коррекции величины ранга матрицы ограничений (изменением отдельных элементов), исследование совместных свойств МЛ (1) и ограничений (2) - неразрешимость, установление свойства единственности или неединственности решений СЛАУ, исследование свойств (1),(2) и оптимизационных задач. Следующий этап - анализ влияния малых возмущений на свойства МЛ.

Основные положения метода базисных матриц (МБМ)

В предложенном МБМ введены в рассмотрение строчные базисные матрицы. Базисные матрицы последовательно изменяются замещением строк вспомогательной СЛАУ строками (нормалей ограничений) основной СЛАУ. В общем случае в модели МЛ количество ограничений превышает количество переменных вида (2), в данном случае $m = n$ (для анализа вводится в рассмотрение вспомогательная СЛАУ с известными свойствами соответствующей размерности) [Кудин, 2002].

Определение 4. Квадратная матрица $A_{\bar{\sigma}}$, составленная из m линейно независимых нормалей ограничений (вспомогательной СЛАУ), будем называть базисной, а решение соответствующей ей системы уравнений $A_{\bar{\sigma}}u = c^0$ базисным. Две базисные матрицы, отличающиеся одной строкой, будем называть смежными.

Пусть β_{ij} , $i, j = 1, 2, \dots, m$, элементы базисной матрицы $A_{\bar{\sigma}}$, e_{ri} - элементы матрицы $A_{\bar{\sigma}}^{-1}$, обратной к $A_{\bar{\sigma}}$; $e_k = (A_{\bar{\sigma}}^{-1})_k$ - столбец обратной матрицы. Решение $u_0 = (u_{01}, u_{02}, \dots, u_{0m})^T$ системы уравнений $A_{\bar{\sigma}}u = c^0$, где c^0 - подвектор C , компоненты которого состоят из правых частей ограничений, нормали которых образуют базисную матрицу $A_{\bar{\sigma}}$; $\alpha_r = (\alpha_{r1}, \alpha_{r2}, \dots, \alpha_{rm})$ - вектор разложения нормали ограничения $a_r u_1 \leq c_r$ по строкам базисной матрицы $A_{\bar{\sigma}}$, $\alpha_0 = (\alpha_{01}, \alpha_{02}, \dots, \alpha_{0m})$ - вектор разложения нормали целевой функции (4) по строкам базисной матрицы $A_{\bar{\sigma}}$, $\Delta_r = a_r u_0^T - c_r$ - невязка r -го ограничения в вершине u_0 ; $J_{\bar{\sigma}}, J_H$, $J = J_{\bar{\sigma}} \cup J_H$, - множества индексов базисных и небазисных ограничений. Установлены формулы связи базисного решения, коэффициентов разложения нормалей ограничений и целевой функции, коэффициентов обратной матрицы, невязок ограничений и значений целевой функции при переходе к базисной матрице $\bar{A}_{\bar{\sigma}}$ (смежной), которая образовывается из матрицы $A_{\bar{\sigma}}$ заменой ее строки a_k на a_l , которая не входит в базисную матрицу $A_{\bar{\sigma}}$ [Волошин, 2008]. При нахождении формул и основных соотношений между элементами метода при переходе от одной базисной матрицы к следующей считаем $a_{i1}, a_{i2}, \dots, a_{im}$ нормальями ограничений, $a_j u^T \leq c_j, j \in J_{\bar{\sigma}}$ $J_{\bar{\sigma}} = \{i_1, i_2, \dots, i_m\}$ - индексы ограничений, нормали которых образуют строки базисной матрицы $A_{\bar{\sigma}}$ a_l - нормаль ограничения $a_l u \leq c_l$, $\alpha_l = (\alpha_{l1}, \alpha_{l2}, \dots, \alpha_{lm})$ - коэффициенты разложения вектора a_l по строкам базисной матрицы $A_{\bar{\sigma}}$.

Теорема 1. Между элементами МБМ в смежных базисных матрицах имеют место соотношения

$$\bar{\alpha}_{rk} = \frac{\alpha_{rk}}{\alpha_{lk}}, \quad \bar{\alpha}_{ri} = \alpha_{ri} - \frac{\alpha_{rk}}{\alpha_{lk}} \alpha_{li}, \quad r = \overline{0, n}; \quad i = \overline{1, m}; \quad i \neq k; \quad (7)$$

$$\bar{e}_{rk} = \frac{e_{rk}}{\alpha_{lk}}, \quad \bar{e}_{ri} = e_{ri} - \frac{e_{rk}}{\alpha_{lk}} \alpha_{li}, \quad r = \overline{1, m}; \quad i = \overline{1, m}; \quad i \neq k; \quad (8)$$

$$\bar{u}_{0j} = u_{0j} - \frac{e_{jk}}{\alpha_{lk}} \Delta_l, \quad j = \overline{1, m}, \quad (9)$$

$$\bar{\Delta}_k = -\frac{\Delta_l}{\alpha_{lk}}, \quad \bar{\Delta}_r = \Delta_r - \frac{\alpha_{rk}}{\alpha_{lk}} \Delta_l, \quad r = \overline{1, n}; \quad r \neq k; \quad (10)$$

$$B \bar{u}_0^{-T} = B u_0^T - \frac{\alpha_{0k}}{\alpha_{lk}} \Delta_l, \quad (11)$$

причем условием невырожденности является условие $\alpha_{lk} \neq 0$, допустимости опорного базисного решения - $\alpha_{lk} < 0$, роста значения целевой функции - $\alpha_{0k} < 0$.

На основе (7)-(11) построена схема типа симплекс-метода анализа МЛ - метод базисных матриц.

Теорема 2. Если \exists индекс k такой, что $\alpha_{0k} < 0$ и $\alpha_{rk} \geq 0$ для всех небазисных r , то целевая функция (4) задачи принимает неограниченные значения на множестве допустимых решений.

Теорема 3. Для существования единственного решения (1) необходимо и достаточно, чтобы $\alpha_{lk}^{(i)} \neq 0$, $i = \overline{1, m}$, где $\alpha_{lk}^{(i)}$ - ведущие элементы симплексной итерации МБМ по замещению строк базисной матрицы нормальями ограничений (1).

Следствие 1. Ранг матрицы ограничений (1) определяется количеством корректных замещений строк матрицы ограничений вспомогательной системы векторами нормальями (1), согласно формулам (7)-(11).

Теорема 4. Оптимизационная задача (3),(4) с квадратной невырожденной матрицей ограничений имеет единственное решение тогда и только тогда, когда $\alpha_{0i} > 0$, $i = \overline{1, m}$.

Пусть имеется возмущение в целевой функции, которая задается функциональной зависимостью вида

$$\bar{\alpha}_0(t) = B(t)(A_b^{-1}) = (\bar{\alpha}_{01}(t), \bar{\alpha}_{02}(t), \dots, \bar{\alpha}_{0m}(t)), \quad t \in [t_0, t_k].$$

Следствие 1. При $t \in T = [t_H, t_B]$ таких, что $\bar{\alpha}_{0k}(t) \geq 0$, $t \in [t_0, t_k]$, $\overline{k=1, m}$, сохраняется оптимальность; $\bar{\alpha}_{0k}(t) > 0$, $t \in [t_0, t_k]$, $\overline{k=1, m}$, - оптимальное решение единственное; $(\exists k_0, \bar{\alpha}_{0k_0}(t) = B(t)(A_b^{-1})_{k_0} = 0, t \in [t_0, t_k])$ - оптимальное решение не единственное; $\exists k_0, \bar{\alpha}_{0k_0}(t) < 0$, $t \in [t_0, t_k]$, решение становится не оптимальным (знаки компонентов вектора $\bar{\alpha}_0(t)$ указывают на свойства решений (2),(3).

Пусть для a_l ограничения $a_l u^T \leq c_l$ относительно A_b выполняется условие $\alpha_{lk} = 0$, а для возмущения: $(a_l + a'_l) u^T \leq c_l + c'_l$, $\bar{\Delta}_l = \Delta_l + \Delta'_l$, $\bar{\alpha}_l = (\alpha_l + \alpha'_l) = (a_l + a'_l) A_b^{-1}$.

Теорема 5. Необходимым условием невырожденности новой \bar{A}_b , образованной замещением нормали a_l (которая занимает k -ю строку в базисной матрице) возмущенной нормалью $(a_l + a'_l)$, являются выполнения условий $\exists i e_{ik} \neq 0$, где e_{ik} элемент A_b^{-1} , такой, что $\alpha'_{li} \neq 0$. Решение будет

определяться соотношением $\bar{u}_{oj} = u_{oj} - \frac{e_{jk}}{\bar{\alpha}'_{lk}} \bar{\Delta}_l$, причем, если $\bar{\Delta}_l = 0$, то $\bar{u}_{oj} = u_{oj}$, $j = \overline{1, m}$.

Следствие 3. Если ведущий элемент симплексной итерации $\alpha_{lk} \neq 0$, то при возмущении, сохраняющем невырожденность, должно выполняться $\alpha'_{lk} + \alpha_{lk} \neq 0$.

Теорема дает условия направленного восстановления свойства невырожденности базисной матрицы (1) и анализа влияния возмущений на решение.

Вычислительный эксперимент

Для вычислительного эксперимента были выбраны модели систем линейных алгебраических уравнений с 100% заполнением матрицы ограничений. Генерировались системы полного ранга (плохо обусловленные) с матрицей вида

$$A = \begin{cases} a_i = (\text{rnd}(1.0), \dots, \text{rnd}(1.0)), & i = 0 \\ a_i = a_{i-1} + \frac{\alpha}{n} (\text{rnd}(1.0), \dots, \text{rnd}(1.0)), & i > 0 \end{cases},$$

где $\|a_i - a_j\|^2 < \alpha$, α -параметр, который коррелирует число обусловленности. Малые возмущения в элементах таких моделей существенно влияли на характеристики решения. Исследовались свойства широко используемых алгоритмов и методов на плохо обусловленных моделях.

Проводилось два типа тестов:

А. Определялись невязки решений и точность разными алгоритмами для матриц размерностью 256x256 с разными значениями α и машинным нулем ε (использовался вариант МБМ без процедуры выбора главного элемента и уточнение).

В. Исследовалась зависимость невязки и быстродействия от размерности при $\alpha = 10^{-4}$ и $\varepsilon = 10^{-12}$.

Для эксперимента были выбраны алгоритмы Гаусса (GS), SVD, Гревилля (G), метода базисных матриц.

Из серии тестов А (Рис.1) можно сделать следующие выводы:

- при $\alpha < 10^{-4}$ метод базисных матриц дает машинно-точное решение за наименьшее время;
- при $\alpha = 10^{-4}$ метод базисных матриц имеет допустимые погрешности решения;
- при $\alpha > 10^{-4}$ только метод Гауса (GS) дает допустимые погрешности решения;

В среднем для всех алгоритмов уменьшение коэффициента α на 2 порядка приводит к снижению точности на 4 порядка. Под машинно-точным решением с плавающей запятой двойной точности понимается решение со значением максимальной невязки, меньшей 10^{-12} . Под допустимыми погрешностями – невязки в диапазоне от 10^{-12} до 10^{-5} . Решение с невязками большими 10^{-5} считаются неверными.

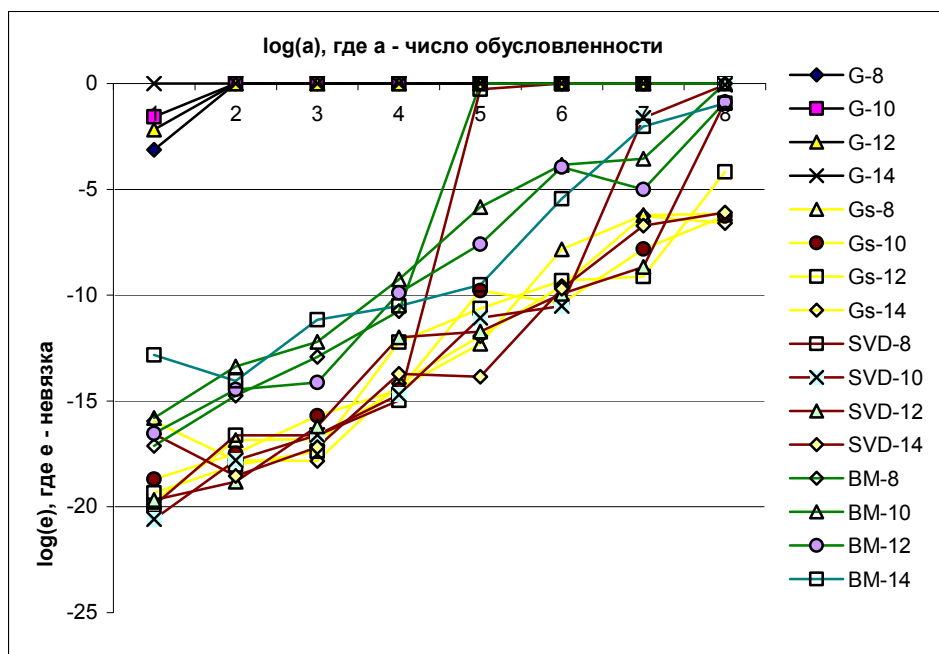


Рис 1. Точность решения в зависимости от значения машинного нуля

Серия тестов В (Табл.1) показывает зависимость точности решения от размерности матриц. Для матриц размерностью $n > 1000$ при увеличении размерности на 1000 элементов точность падает на порядок для всех алгоритмов. Однако, точность метода МБМ в среднем ниже, чем у алгоритмов метода Гаусса и SVD.

Таблица 1. Зависимость точности решения от размерности матрицы

Gauss				
N	$\alpha = 1$	1.00E-02	1.00E-04	1.00E-06
1000	1.41E-20	1.39E-17	4.51E-13	1.06E-10
2000	2.4E-20	1.58E-15	2.97E-12	8.51E-08
3000	1.8E-20	6.43E-15	9.5E-11	2.31E-06
4000	4.88E-17	2.01E-14		
5000	1.18E-16	9.31E-13		

SVD				
N	$\alpha = 1$	1.00E-02	1.00E-04	1.00E-06
1000	4.65E-21	2.5E-17	1.42E-13	3.99E-09
2000	1.13E-19	9.38E-16	1.81E-12	1.24E-07
3000	1.69E-19	4.33E-15	1.06E-10	1.63E-08

BM-I				
N	$\alpha = 1$	1.00E-02	1.00E-04	1.00E-06
1000	7.59E-16	1.02E-12	1.37E-08	0.004137
2000	1.25E-15	5.93E-12	6E-07	5897
3000	1.24E-13	2.1E-09	0.013045	1.007973
4000	5.09E-11	1.55E-07		
5000	1.27E-10	2.03E-07		

Исследованы и построены оценочные множества принадлежности решений при возмущениях (Рис. 4) для плохо обусловленных систем. Пример демонстрирует существенное влияние малых возмущений на множество принадлежности решений (значительная растянутость множества по оси абсцисс).

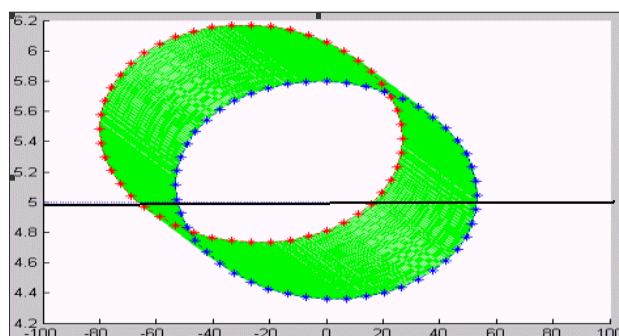


Рис. 4. Эллипсы принадлежности решений при малых возмущениях (оси координат — компоненты вектора решения)

Выводы

Применение симплексной идеологии на основе МБМ и ПАВ при анализе МЛ дает возможность:

- исследовать свойства решений МЛП (1)- (3) при изменениях в (3);
- использовать решение исходной МЛ при анализе возмущенной модели;
- контролировать или направленно изменять величину ранга системы;
- находить решение квадратной системы уравнений за фиксированное число шагов;
- проводить анализ свойств системы при изменении значений отдельных элементов и ее компонент;
- использовать решения начальной системы при анализе возмущенной задачи;
- разрабатывать механизм преобразование вырожденной матрицы ограничений в невырожденную направленной коррекцией элементов;
- анализировать корректность построения модели.

Благодарности

Статья частично финансированна из проекта **ITHEA XXI** Института Информационных теории и Приложений FOI ITHEA и Консорциума FOI Bulgaria (www.ithea.org, www.foibg.com).

Литература

- [Пономаренко, 2004] Пономаренко О.І. та ін. Макроекономіка.- К. :Вища школа,- 2004. - 207с.
- [Волошин, 1987] Волошин А.Ф. Метод локализации области оптимума в задачах математического программирования // Докл. АН СССР. - 1987. - 293, N 3.- С. 549-553.
- [Волошин, 2004] Волошин О.Ф. Методи аналізу статичних балансових еколого-економічних моделей великої розмірності // Наукові записки, Т.VII, КПДВ "Педагогіка", Київ, 2004. - С. 43-55.
- [Орловский, 1981] Орловский С.А. Принятие решений при нечёткой исходной информации.- М.: Наука, - 1981. - 206с.
- [Волошин, 2008] Кудин. В., Кудин Г., Волошин А. Анализ свойств модели Леонтьева при нечетко заданных параметрах с применением метода базисных матриц // Information Science & Computing, International Book Series, Number 7, Volume 7, ITHEA, SOFIA, 2008.- P. 86-90.
- [Кудин, 2002] Кудин В.И. Применение метода базисных матриц при исследовании свойств линейной системы // Вестник Киевского университета. Серия физ.-мат. науки. - 2002. - 2. - С. 56-61.

Информация об авторах

Алексей Ф. Волошин, Киев,. Киевский национальный университет имени Тараса Шевченко, факультет кибернетики, Украина, д. т. н., профессор, e-mail: ovoloshin@unicyb.kiev.ua

Всеволод А. Богаенко, Киев,. Институт кибернетики имени В.Г. Глушкова НАН Украины, Украина,, к. т. н., н. с., e-mail: sevab@ukrnet.ua

Владимир И. Кудин, Киев,. Киевский национальный университет имени Тараса Шевченко, факультет кибернетики, Украина, д. т. н., с. н. с., e-mail: V_I_Kudin@mail.ru

ИССЛЕДОВАНИЕ ЗАДАЧИ ОПТИМИЗАЦИИ ИНВЕСТИЦИОННОГО ПОРТФЕЛЯ В НЕЧЕТКИХ УСЛОВИЯХ

Юрий Зайченко, Малихех Есфандиярфард

Abstract: The problem of fuzzy portfolio optimization is considered and analyzed. The dependence “profitability-risk” for optimal fuzzy portfolio was obtained and investigated. The sufficient conditions for the curve “profitability-risk” to be monotonous decreasing were obtained. The experimental investigations of this dependence were carried out which fully confirmed the theoretical results.

Keywords: fuzzy portfolio, optimization, curve “optimal profitability-risk”, sufficient conditions.

ACM Classification Keywords: H.4.2. Information system Applications: Types of Systems Decision Support

Conference: The paper is selected from XVth International Conference “Knowledge-Dialogue-Solution” KDS 2009, Varna, Bulgaria, June-July 2009

Введение

Задача оптимизации инвестиционного портфеля в условиях неопределенности в последние годы вызывает значительный интерес. Для решения этой проблемы был предложен аппарат нечетких множеств в работах [Зайченко,2008, Zaychenko, 2008], согласно которому доходности акций и в целом доходность инвестиционного портфеля рассматриваются как нечеткие числа с заданной функцией принадлежности, а риск трактуется как возможность (субъективная вероятность) ситуации, когда реальная доходность портфеля оказывается ниже ожидаемой доходности нечеткого портфеля. С целью более обоснованной оценки доходности акций на фондовом рынке по предыстории в задаче портфельной оптимизации был предложен метод прогнозирования, применение которого позволило улучшить результаты [Зайченко,2007]. В ходе экспериментальных исследований разработанного метода были построены зависимости «оптимальная доходность- риск» для нечеткого портфеля, которые во многих случаях имели прямо противоположный вид по сравнению с классической моделью Марковица-Тоббина, а именно были монотонно убывающими, а не возрастающими [Зайченко,2008]. Поэтому представляет значительный теоретический и практический интерес выяснение достаточных аналитических условий, при которых эта зависимость имеет монотонно падающий характер. Целью настоящей работы является получение аналитических условий, при которых зависимость «оптимальная доходность- риск» для нечеткого портфеля будет монотонно убывающей и их экспериментальная проверка..

Анализ модели оптимизации нечеткого портфеля

Задача нечеткой портфельной оптимизации базируется на допущении, что доходность портфеля является нечетким числом (НЧ), описываемым тройкой параметров

$$r = (r_{\min} = \sum_{i=1}^N x_i r_{i1}; \tilde{r} = \sum_{i=1}^N x_i \tilde{r}_i; r_{\max} = \sum_{i=1}^N x_i r_{i2})$$

где $(r_{1i}, \tilde{r}_i, r_{2i})$ – доходность i -той ценной бумаги. x_i – доля i -той бумаги в портфеле

Для того, чтобы определить структуру портфеля, который обеспечит максимальную доходность при заданном уровне риска, требуется решить следующую задачу оптимизации:

$$\tilde{r} = \sum_{i=1}^N x_i \tilde{r}_i \rightarrow \max \quad (1)$$

$$\beta = \text{const} \quad (2)$$

$$\sum_{i=1}^N x_i = 1, \quad x_i \geq 0, \quad i = \overline{1, N} \quad (3)$$

$0 < \beta < 1$, где β - уровень риска.

В работе [Зайченко,2008] было показано, что этот случай возможен когда $\sum_{i=1}^N x_i r_{i1} \leq r^* < \sum_{i=1}^N x_i \tilde{r}_i$, либо

когда $\sum_{i=1}^N x_i \tilde{r}_i \leq r^* < \sum_{i=1}^N x_i r_{i2}$.

Пусть $\sum_{i=1}^N x_i r_{i1} \leq r^* < \sum_{i=1}^N x_i \tilde{r}_i$. Тогда используя (результаты работы [Зайченко,2008,], задача (1)-(3)

сводится к следующей задаче нелинейного программирования:

$$\tilde{r} = \sum_{i=1}^N x_i \tilde{r}_i \rightarrow \max \quad (4)$$

$$\frac{1}{\sum_{i=1}^N x_i r_{i2} - \sum_{i=1}^N x_i r_{i1}} \left(\left(r^* - \sum_{i=1}^N x_i r_{i1} \right) + \left(\sum_{i=1}^N x_i \tilde{r}_i - r^* \right) \cdot \ln \left(\frac{\sum_{i=1}^N x_i \tilde{r}_i - r^*}{\sum_{i=1}^N x_i \tilde{r}_i - \sum_{i=1}^N x_i r_{i1}} \right) \right) = \beta \quad (5)$$

$$\sum_{i=1}^N x_i r_{i1} \leq r^* \quad (6)$$

$$\sum_{i=1}^N x_i \tilde{r}_i > r^* \quad (7)$$

$$\sum_{i=1}^N x_i = 1, \quad x_i \geq 0, \quad i = \overline{1, N} \quad (8)$$

Учитывая вид для r_{\min} , \tilde{r} и r_{\max} , а также формулу (1), данную задачу нахождения интервала доходности для нечеткого портфеля с n ЦБ можно свести к задаче с двумя ЦБ (x_1, x_2) с минимальной и максимальной доходностями.

Пусть первая ЦБ имеет минимальную доходность в портфеле, описываемую НЧ $r_1 = [r_{11}, \tilde{r}_1, r_{12}]$, а вторая ЦБ – максимальную доходность в портфеле $r_2 = [r_{21}, \tilde{r}_2, r_{22}]$.

Пусть также критериальное значение описывается НЧ $r^* = [r_1^*, r^*, r_2^*]$.

При этом допустим $r_{11} < r_{21}$; $\tilde{r}_1 < \tilde{r}_2$.

Тогда $\tilde{r} = \tilde{r}_1 x_1 + \tilde{r}_2 x_2$; $x_1 + x_2 = 1$. Причем $x = x_2$, тогда $x_1 = 1 - x$.

Доходность портфеля описывается в этом случае НЧ вида $r = [r_{\min}, \tilde{r}, r_{\max}]$, где:

$$r_{\min} = r_{11}(1-x) + r_{12}x; \quad \tilde{r} = \tilde{r}_1(1-x) + \tilde{r}_2x; \quad r_{\max} = r_{21}(1-x) + r_{22}x.$$

Требуется найти такое x , что

$$\tilde{r} = \tilde{r}_1(1-x) + \tilde{r}_2x \rightarrow \max \quad (9)$$

при условиях:

$$\beta(x) \leq \beta_{\text{зад}}, \quad (10)$$

где выражение $\beta(x)$ для случая $0 < \beta < 1$ и $\tilde{r} < \tilde{r}_2$ при подстановке $x_1 = 1-x$; $x_2 = x$; $n = 2$ в выражение (5) дает:

$$\beta(x) = \frac{1}{(1-x)(r_{12} - r_{11}) + x(r_{22} - r_{21})} \{r^* - ((1-x)r_{11} + xr_{21}) + ((1-x)\tilde{r}_1 + x\tilde{r}_2 - r^*)\} \cdot \ln \frac{\tilde{r}_1(1-x) + \tilde{r}_2x - r^*}{x(\tilde{r}_1 - r_{11}) + (1-x)(\tilde{r}_2 - r_{21})} \leq \beta$$

Достаточные условия монотонно убывающего характера зависимости $\tilde{r}_{\text{opt}}(\beta)$ являются следующие:

$$\frac{\partial \tilde{r}_{\text{opt}}}{\partial x} \geq 0; \quad \frac{\partial \beta}{\partial x} \leq 0 \quad (11)$$

Рассмотрим сначала $\frac{\partial \tilde{r}}{\partial x}$.

Очевидно $\frac{\partial \tilde{r}}{\partial x} = \tilde{r}_2 - \tilde{r}_1$ (12) и т.к. по предположению $\tilde{r}_1 < \tilde{r}_2$, то $\frac{\partial \tilde{r}}{\partial x} > 0$.

Рассмотрим теперь зависимость $\frac{\partial \beta}{\partial x}$.

Для нахождения сложной производной $\frac{\partial \beta}{\partial x}$, введем следующие обозначения

$$A(x) = r^* - ((1-x)r_{11} + xr_{12}); \quad B(x) = (1-x)\tilde{r}_1 + x\tilde{r}_2 - r^* \quad (13)$$

$$C(x) = (1-x)(\tilde{r}_1 - r_{11}) + x(\tilde{r}_2 - r_{21}) \quad (14)$$

$$D(x) = (1-x)(r_{12} - r_{11}) + x(r_{22} - r_{21}) \quad (15)$$

Тогда

$$\beta(x) = \left[A(x) + B(x) \ln \frac{B(x)}{C(x)} \right] \frac{1}{D(x)} \quad (16)$$

Найдем производную $\frac{\partial \beta}{\partial x}$:

$$\frac{\partial \beta}{\partial x} = \left\{ \left[A'(x) + B'(x) \ln \frac{B(x)}{C(x)} + \frac{B'(x)C(x) - B(x)C'(x)}{C(x)} \right] D(x) - D'(x) \left[A(x) + B(x) \ln \frac{B(x)}{C(x)} \right] \right\} \frac{1}{D^2(x)} \quad (17)$$

Выясним условия, при которых $\frac{\partial \beta}{\partial x} < 0$.

Для этого необходимо, чтобы выражения в фигурных скобках $\{ \}$ в (17) было меньше 0.

$$\begin{aligned} \text{Очевидно} \quad A'(x) &= r_{11} - r_{21}; \quad B'(x) = -\tilde{r}_1 + \tilde{r}_2; \\ C'(x) &= -(\tilde{r}_1 - r_{11}) + (\tilde{r}_2 - r_{21}) = (\tilde{r}_2 - \tilde{r}_1) - (r_{21} - r_{11}); \\ D'(x) &= (r_{22} - r_{21}) - (r_{12} - r_{11}). \end{aligned} \quad (18)$$

Подставим (18) в (17) и из условия $\frac{\partial \beta}{\partial x} < 0$ после несложных преобразований в (17) получим следующее условие

$$C(x)D(x) \left[A'(x) + B'(x) + B'(x) \ln \frac{B(x)}{C(x)} \right] - DC'(x)B(x) - C(x)D'(x)(A(x) + B(x) \ln \frac{B(x)}{C(x)}) < 0 \quad (19)$$

Подставляя выражение для $A(x); A'(x); B(x); B'(x); C(x); C'(x); D(x); D'(x)$ в (19), получим

$$\begin{aligned} &[(1-x)(r_{12} - r_{11}) + x(r_{22} - r_{21})][(1-x)(\tilde{r}_1 - r_{11}) + x(\tilde{r}_2 - r_{21})]^* \\ &* \left[(r_{11} - r_{21}) + (\tilde{r}_2 - \tilde{r}_1) + (\tilde{r}_2 - \tilde{r}_1) \ln \frac{(1-x)\tilde{r}_1 + x\tilde{r}_2 - r^*}{(1-x)(\tilde{r}_1 - r_{11}) + x(\tilde{r}_2 - r_{21})} \right] - [(1-x)\tilde{r}_1 + x\tilde{r}_2 - r^*][(\tilde{r}_2 - r_{21}) - (\tilde{r}_1 - r_{11})] \\ &[(1-x)(r_{12} - r_{11}) + x(r_{22} - r_{21})] - [(1-x)(\tilde{r}_1 - r_{11}) + x(\tilde{r}_2 - r_{21})][(r_{22} - r_{21}) - (r_{12} - r_{11})]^* \\ &*(r^* - (1-x)r_{11} - xr_{21} + [(1-x)\tilde{r}_1 + x\tilde{r}_2 - r^*])^* \ln \frac{(1-x)\tilde{r}_1 + x\tilde{r}_2 - r^*}{(1-x)(\tilde{r}_1 - r_{11}) + x(\tilde{r}_2 - r_{21})} \end{aligned} \quad (20)$$

Пусть функции принадлежности являются треугольными симметричными.

Тогда

$$\tilde{r}_1 - r_{11} = r_{12} - \tilde{r}_1 = \Delta_1; \quad \tilde{r}_2 - r_{21} = r_{22} - \tilde{r}_2 = \Delta_2 \quad (21)$$

и, очевидно, $r_{12} - r_{11} = 2\Delta_1$; $r_{22} - r_{21} = 2\Delta_2$.

Подставляя эти выражения в (20) получим

$$\begin{aligned} &[(1-x)2\Delta_1 + x2\Delta_2][(1-x)2\Delta_1 + x\Delta_2]^* \\ &* \left[(\Delta_2 - \Delta_1) + (\tilde{r}_2 - \tilde{r}_1) \ln \frac{(1-x)\tilde{r}_1 + x\tilde{r}_2 - r^*}{(1-x)(\tilde{r}_1 - r_{11}) + x(\tilde{r}_2 - r_{21})} \right] - [(1-x)\tilde{r}_1 + x\tilde{r}_2 - r^*][\Delta_2 - \Delta_1] \\ &[(1-x)2\Delta_1 + x2\Delta_2] - [(1-x)2\Delta_1 + x\Delta_2][\Delta_2 - \Delta_1]^* \\ &*(r^* - (1-x)r_{11} - xr_{21}) + [(1-x)\tilde{r}_1 + x\tilde{r}_2 - r^*]^* \ln \frac{(1-x)\tilde{r}_1 + x\tilde{r}_2 - r^*}{(1-x)(\tilde{r}_1 - r_{11}) + x(\tilde{r}_2 - r_{21})} \end{aligned} \quad (21)$$

Преобразуем правую часть выражения (21) и получим

$$\begin{aligned}
& -[(1-x)\tilde{r}_1 + x\tilde{r}_2 - r^*](\Delta_2 - \Delta_1)[(1-x)2\Delta_1 + x2\Delta_2] - [(1-x)\Delta_1 + x\Delta_2](\Delta_2 - \Delta_1)\{r^* - (1-x)r_{11} - xr_{21}\} = \\
& = -2[(1-x)\Delta_1 + \Delta_2](\Delta_2 - \Delta_1)[(1-x)\tilde{r}_1 + x\tilde{r}_2 - r^* + r^* - (1-x)r_{11} - xr_{21}] = \\
& = -2[(1-x)\Delta_1 + \Delta_2](\Delta_2 - \Delta_1)[(1-x)\Delta_1 + x\Delta_2] = 2[(1-x)\Delta_1 + \Delta_2]^2(\Delta_2 - \Delta_1)
\end{aligned} \quad (22)$$

Анализируя выражение (22) нетрудно проверить, что

$C(x)D(x)[A'(x) + B'(x)] = D(x)C'(x)B(x) + C(x)D'(x)A(x)$, поэтому после сокращения подобных членов в (22) останутся члены

$$C(x)D(x)B'(x)\ln\frac{B(x)}{C(x)} - C(x)D'(x)B(x)\ln\frac{B(x)}{C(x)} \quad (23)$$

И подставляя выражение для $C(x); D(x); B(x); B'(x)$ из (23), получим

$$\begin{aligned}
& [(1-x)(r_{12} - r_{11}) + x(r_{22} - r_{21})] \\
& [(1-x)((r_{12} - r_{11}) + x(r_{22} - r_{21}))(\tilde{r}_2 - \tilde{r}_1) - (1-x)\tilde{r} + x\tilde{r}_2 - r^*)(\tilde{r}_2 - \tilde{r}_1)(r_{22} - r_{21})]^* \\
& * \ln\frac{(1-x)\tilde{r}_1 + x\tilde{r}_2 - r^*}{(1-x)(\tilde{r}_1 - r_{11}) + x(\tilde{r}_2 - r_{21})}
\end{aligned} \quad (24)$$

Откуда

$$2[(1-x)\Delta_1 + x\Delta_2]\ln\frac{(1-x)\tilde{r}_1 + x\tilde{r}_2 - r^*}{(1-x)(\tilde{r}_1 - r_{11}) + x(\tilde{r}_2 - r_{21})}[(1-x)2\Delta_1 + x2\Delta_2(\tilde{r}_2 - \tilde{r}_1) - (\Delta_2 - \Delta_1)[(1-x)\tilde{r}_1 + x\tilde{r}_2 - r^*]] \quad (25)$$

Заметим, что

$$(1-x)(\tilde{r}_1 - r_{11}) + x(\tilde{r}_2 - r_{21}) = (1-x)\tilde{r}_1 + x\tilde{r}_2 - (1-x)r_{11} - xr_{21} = (1-x)\tilde{r}_1 + x\tilde{r}_2 - r_{\min}$$

и так как $r^* > (1-x)r_{11} + xr_{22} = r_{\min}$, то

$$\frac{B(x)}{C(x)} = \frac{(1-x)\tilde{r}_1 + x\tilde{r}_2 - r^*}{(1-x)(\tilde{r}_1 - r_{11}) + x(\tilde{r}_2 - r_{21})} < 1 \quad \text{и} \quad \ln\frac{B(x)}{C(x)} < 0 \quad (26)$$

Следовательно, для того, чтобы выражение (25) было меньше нуля необходимо и достаточно, чтобы $(1-x)2\Delta_1 + x2\Delta_2(\tilde{r}_2 - \tilde{r}_1)(\Delta_2 - \Delta_1)[(1-x)\tilde{r}_1 + x\tilde{r}_2 - r^*]$ было больше 0.

Исследуем условия, при которых

$$\{2[(1-x)\Delta_1 + x\Delta_2](\tilde{r}_2 - \tilde{r}_1) - (\Delta_2 - \Delta_1)[(1-x)\tilde{r}_1 + x\tilde{r}_2 - r^*]\} > 0 \quad (27)$$

а) неравенства $\tilde{r}_2 - \tilde{r}_1 > \Delta_2 - \Delta_1 = (\tilde{r}_2 - \tilde{r}_1) - (r_{21} - r_{11})$, выполняются, если $r_{21} > r_{11}$

б) неравенства $2[(1-x)\Delta_1 + x\Delta_2] > (1-x)\tilde{r}_1 + x\tilde{r}_2 - r^*$ выполняется, если

$$\begin{aligned}
& 2(1-x)(\tilde{r}_1 - r_{11}) + 2x(\tilde{r}_2 - r_{21}) - (1-x)\tilde{r}_1 - x\tilde{r}_2 + r^* = (1-x)(\tilde{r}_1 - r_{11}) + x(\tilde{r}_2 - 2r_{21}) + r^* \\
& = (1-x)(\tilde{r}_1 - r_{11}) + x(\tilde{r}_2 - r_{21}) + r^* - [(1-x)r_{11} + xr_{21}] = (1-x)\Delta_1 + x\Delta_2 + r^* - r_{\min} > 0
\end{aligned} \quad (28)$$

и (28) выполняется, если $r^* - r_{\min} > 0$.

Таким образом мы получили следующие достаточные условия для $\frac{\partial \beta}{\partial x} < 0$ (см. рис 1)

а) $\tilde{r}_1 < r^* < \tilde{r}_2$; б) $r_{11} < r_{21}$;

$$в) r^* > r_{\min} = (1-x)r_{11} + x_1r_{21}; \quad г) r^* < \tilde{r} = (1-x)\tilde{r}_1 + x_1\tilde{r}_2.$$

При выполнении этих условий, зависимость «оптимальная доходность-риск» будет иметь монотонно-убывающий характер.

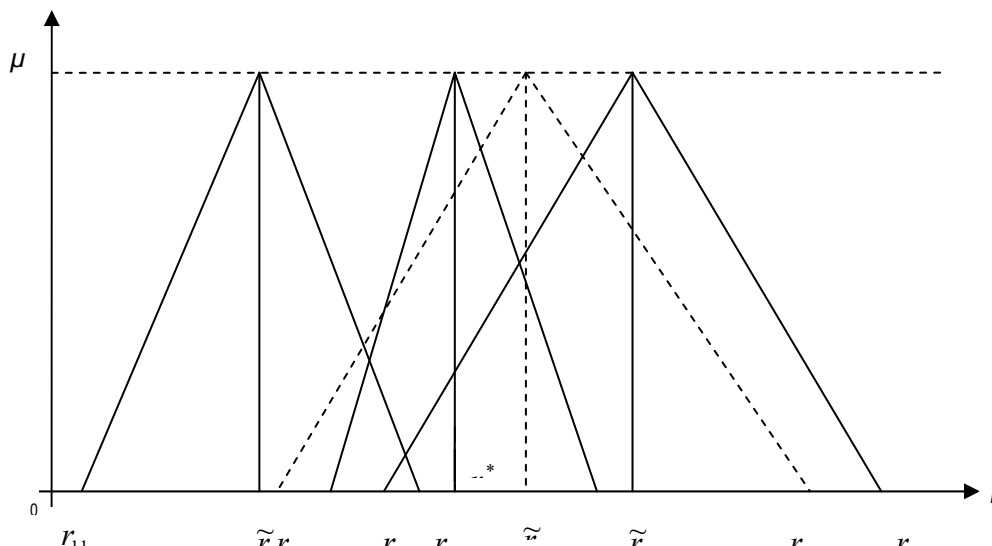


Рис. 1

Экспериментальные исследования зависимости «оптимальная доходность- риск»

Проведем экспериментальные исследования зависимости «оптимальная доходность- риск» для нечеткого инвестиционного портфеля. При этом рассмотрим случаи двух ЦБ с малой доходностью и высокой доходностью с различными волатильностями.

При этом будем варьировать величину критериального значения.

Исследуем случаи, когда эта зависимость будет иметь монотонно-убывающий характер.

Когда первой ЦБ была задана высокая волатильность и низкая доходность, а другой ЦБ- средняя волатильность и высокая доходность, то зависимость доходность-риск носила монотонно – убывающий характер. Это иллюстрируется графиками на рис.2.

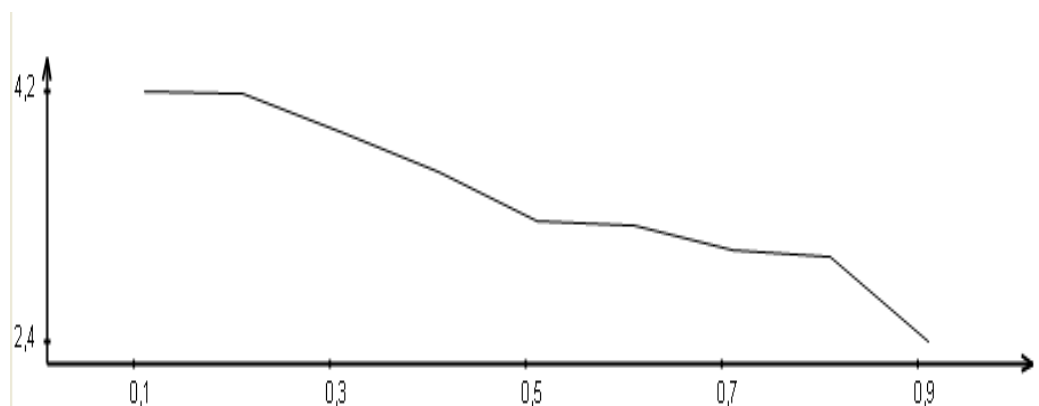


Рис .2. График зависимости доходность- риск нечеткого портфеля

Приведем примеры взаимного расположения доходностей ценных бумаг с различной волатильностью и нечеткого критериального значения

Пример 1

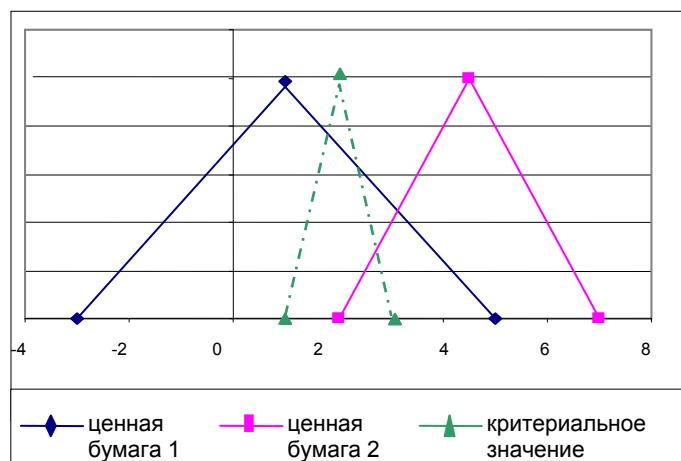


Рис.3. Взаимное расположение нечетких доходностей ЦБ и критериального значения для треугольных ФП

Пример 2

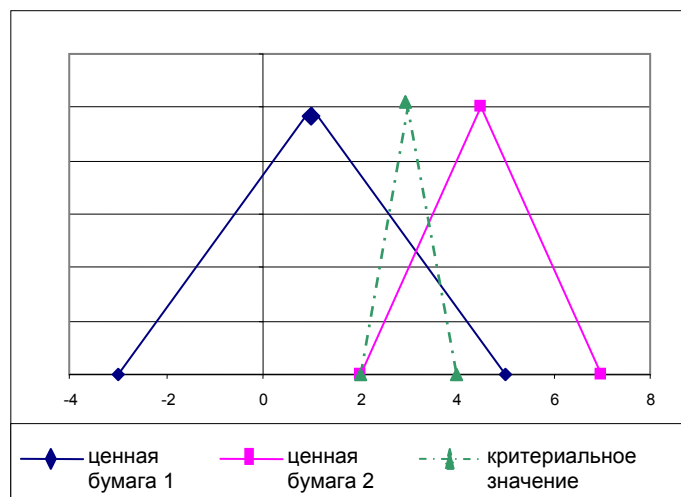


Рис. 4 Взаимное расположение нечетких доходностей ЦБ и критериального значения

Пример 3

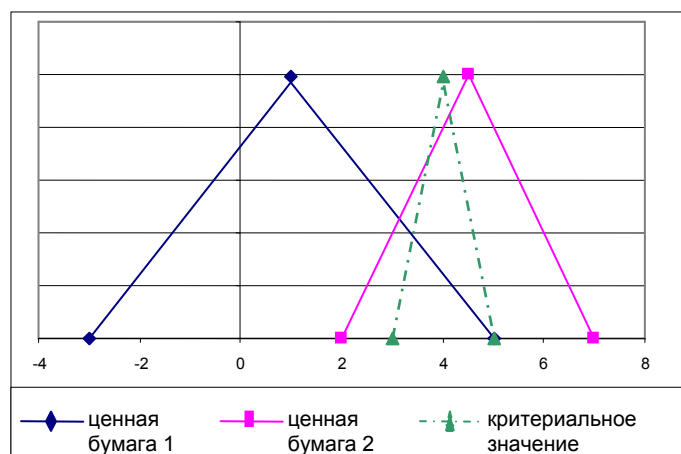


Рис.5. Взаимное расположение нечетких доходностей ЦБ и критериального значения

Пример 4.

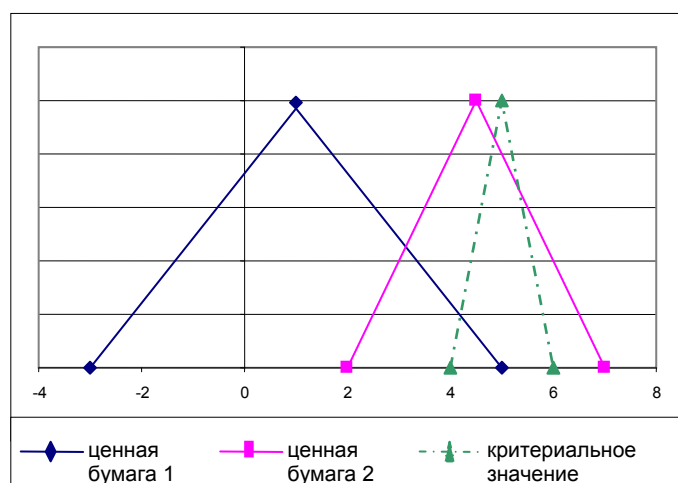


Рис.6. Взаимное расположение нечетких доходностей ЦБ и критериального значения

Пример 5

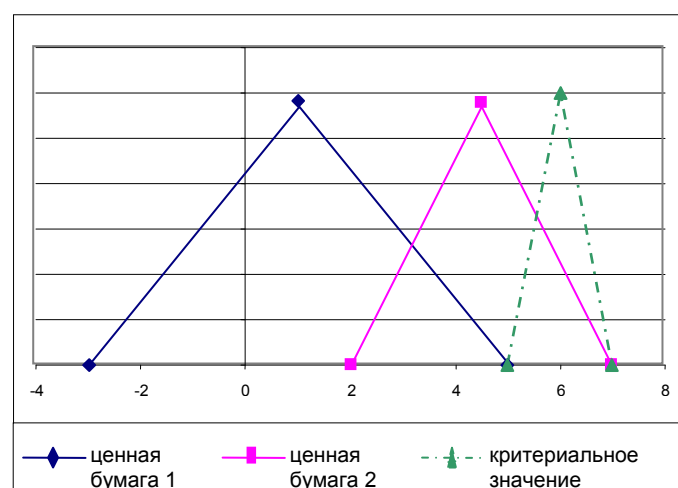


Рис.7. Взаимное расположение нечетких доходностей ЦБ и критериального значения

Анализ представленных результатов экспериментов позволяет сделать следующие **выводы**.

Зависимость «доходность- риск» нечеткого портфеля носит преимущественно убывающий характер в случаях когда менее доходная ЦБ имеет высокую или среднюю волатильность, а более ценная ЦБ имеет высокую доходность при малой и средней волатильности. При этом критериальное значение находится между доходностями обеих ЦБ.

Зависимость «доходность-риск» носит возрастающий характер, в случаях, когда менее доходная ЦБ имеет малую или среднюю волатильность, а более доходная ЦБ имеет высокую волатильность, а критериальное значение меньше доходности менее ценной ЦБ или расположено между ожидаемыми доходностями обеих ЦБ.

Выводы

В работе исследована модель задачи оптимизации нечеткого инвестиционного портфеля. Установлены достаточные условия, при которых зависимость «оптимальная доходность-риск» будет иметь убывающий характер, прямо противоположный зависимости для классической модели Марковица-Тоббина. Проведены экспериментальные исследования условий и оптимальных решений задачи портфельной оптимизации для убывающей и возрастающей зависимостей «доходность-риск», которые полностью подтвердили теоретические результаты.

Благодарности

Статья частично финансирована из проекта **ITHEA XXI** Института Информационных теории и Приложений FOI ITHEA и Консорциума FOI Bulgaria (www.ithea.org, www.foibg.com).

Bibliography

- [Зайченко,2008] Зайченко Ю.П., Малихех Есфандиярфард. Анализ инвестиционного портфеля для различных видов функций принадлежности //системні дослідження та інформаційні технології.-№2-2008.-с. 59-76.
- [Зайченко, Малихех 2008] Зайченко Ю.П., Малихех Есфандиярфард. Нечеткий метод индуктивного моделирования для прогнозирования курсов акций в задачах портфельной оптимизации //Вісник Черкаського державного технологічного університету.-№ 1.- 2008.- с. 9-14.
- [Зайченко,2007] Зайченко Ю.П., Малихех Есфандиярфард, Заика А.И. Анализ инвестиционного портфеля на основе прогнозирования курсов акций// Вісник національного технічного університету України «КПИ». «Інформатика, управління та обчислювальна техніка.» Київ ТОО «ВЕК+», № 47 – 2007, - с. 168-179.
- [Zaychenko,2008] Zaychenko Yurii, Esfandiyarfard Maliheh. Optimization of the Investment Portfolio in the Conditions of Uncertainty/ International Journal "Information Technologies and Knowledge" Volume 2 / 2008, Number 3. Sofia-Bulgaria.pp. 225-233

Authors' Information

Зайченко Юрий Петрович, профессор, д.т.н., кафедра «Институт прикладного системного анализа». Киев, НТУУ «КПИ», ул. Политехническая 14. тел: +8(044)241-86-93, e-mail: baskervil@voliacable.com

Малихех Есфандиярфард (Иран), аспирантка кафедры «Прикладная математика» НТУУ «КПИ», проспект Победы 37, тел: +38(096)6915857, e-mail: fard_sem@yahoo.com

СРАВНЕНИЕ ПОЛИИНТЕРВАЛЬНЫХ ОЦЕНОК В МЕТОДЕ ОИО

Михаил Стернин, Геннадий Шепелёв

Abstract: A problem concerning in comparison of poly-interval structures in generalized interval estimations approach is discussed. It's showed that to compare two poly-interval estimations (PIEs) we may compare their different mono-interval realizations, i.e. different pairs of mono-interval estimations from compared PIEs, using a Monte-Carlo family method. To compare mono-intervals people often use equivalent point estimations of intervals received on base of Hurwicz's criterion or similar criteria. We offer a method of direct comparison of mono-interval estimations that interprets the problem as a problem of decision-making. We suggest distinguishing a favorable zone (zone where one interval in comparable pair may be more preferable than another), a zone of uncertainty (zone where compared intervals are equivalent to each other) and an unfavorable zone. Preferences of the decision maker during analysis of relative widths of the zones allow analyze the problem. The results of direct comparison are compared with results received on the base of equivalent point estimations. The arguments are given that method of direct comparison of intervals is more adequate to decision-making tasks. Relationship between numbers of preferable, equivalent and no preferable comparisons of mono-intervals from PIEs gives an information that support process of decision making during comparison of poly-interval estimations by decision maker.

Keywords: generalized interval estimations, comparison of intervals, decision support

ACM Classification Keywords: H.4.2 Types of Systems – Decision support; G.1.0 Numerical Analysis - Interval Arithmetic, G.3 Probability and Statistics – Distribution functions

Conference: The paper is selected from XVth International Conference "Knowledge-Dialogue-Solution" KDS 2009, Varna, Bulgaria, June-July 2009

Введение

Для расширения возможностей выявления и представления экспертных знаний о количественных параметрах, известных с неопределенностью, нами предложен метод обобщенных интервальных оценок (ОИО) [Sherelyov, Sternin, 2003]. В отличие от традиционного моноинтервального оценивания в методе ОИО, для учета возможной неопределенности экспертной оценки по размаху и положению, она представляется совокупностью интервалов, каждый из которых служит сценарием возможной реализации анализируемого параметра. Шансы на реализацию описываются совместной функцией распределения вероятностей двух случайных переменных α , D с плотностью $\Psi(\alpha, D) = f_1(\alpha)f_2(D|\alpha)$. Переменная α «отмечает» место интервалов в ОИО. Соответствующее ей распределение $f_1(\alpha)$ характеризует шансы различных полос интервальных оценок в их совокупности. Переменная D описывает неопределенность значений параметра на каждой из интервальных оценок совокупности, а плотности $f_2(D|\alpha)$ - шансы на реализацию тех или иных значений D . На плоскости (D, α) ОИО удобно представлять криволинейной трапецией, названной полиинтервальной оценкой (ПИО) параметра.

Изучены некоторые возможные направления использования ОИО. Во-первых, разработаны процедуры усреднения, позволяющие свести задачи в ОИО постановке к задачам в моноинтервальной постановке [Стернин, Шепелев, 2003]. При этом на «базовом» интервале ОИО, образованном как теоретико-множественное объединение всех интервальных оценок совокупности, получаются усредненные результирующие функции распределения вероятностей («проекция» ОИО на базовый интервал). В русле этого направления показаны возможности исследования методом ОИО динамических задач [Стернин и др., 2005]. Удобным средством экспресс-анализа указанных задач явились распределения вероятностей, обобщающие традиционные. Так если на обеих осях трапеции (D, α) заданы равномерные

распределения, результатом усреднения служит обобщенное равномерное распределение [Стернин и др., 2007].

Во-вторых, предложены методы анализа структуры ОИО с автоматическим построением вероятностных трубок (p-tubes) и распределений вероятностей на различных сечениях таких трубок [Chugunov et al., 2008]. Наконец, найдены пути использования метода ОИО в сценарном анализе теории принятия решений [Стернин, Шепелев, 2008].

Для дальнейшего развития метода ОИО как инструмента поддержки процессов принятия решений необходимо исследовать возможности сравнения ПИО для результирующих индикаторов, описывающих эффективность альтернативных вариантов. Одной из таких возможностей является сравнение получаемых методом Монте-Карло различных реализаций моноинтервальных оценок в парах сопоставляемых ПИО с накоплением статистики результатов сравнения и дальнейшим использованием предпочтений лица, принимающего решение (ЛПР). Важнейшая составляющая этого подхода – разработка и выбор метода сопоставления моноинтервальных оценок. Эта, последняя, задача имеет и самостоятельное значение, не связанное с задачей сравнения ПИО.

Сравнение моноинтервальных оценок

Принятие решений часто сопровождается необходимостью прогнозирования значений исходных данных используемых расчетных схем (моделей) или результирующих индикаторов эффективности сопоставляемых в процессе решения альтернатив. Неизбежная неопределенность, присущая прогнозным оценкам, зачастую приводит к потребности в их представлении в виде числовых интервалов, выполнении операций над ними в ходе расчетов итоговых индикаторов, также получающих интервальное представление, и, наконец, в сравнении интервалов-индикаторов разных альтернатив при принятии решения. Методы интервального анализа [Калмыков и др., 1986] дают нам возможность проведения расчетов с интервальными оценками. Например, если мы имеем реализуемый в течение 5 лет инвестиционный проект с компонентами денежного потока $F_t = [F_t, F_t]$: $F_0 = [-105, -95]$, $F_1 = [70, 75]$, $F_2 = [83, 85]$, $F_3 = [100, 104]$, $F_4 = [191, 194]$, $F_5 = [94, 96]$ и ставкой дисконтирования d , принимающей значения в интервале $d = [10\%, 12\%]$, то результирующая оценка чистого дисконтированного дохода NPV, исчисленного в соответствии с правилами интервальной арифметики, будет лежать в интервале $NPV = [269.57, 313.68]$. Аналогичным образом может быть рассчитан чистый дисконтированный доход для какого-либо другого проекта. Чтобы выбрать один из них для финансирования, инвестор должен сравнить интервальные значения соответствующих NPV и решить, какая из этих оценок предпочтительнее («больше»). Таким образом, для принятия решения ЛПР должен обладать основанным на сопоставлении интервальных значений индикаторов эффективности методом выявления предпочтительной альтернативы или констатации того факта, что альтернативы несравнимы.

При сравнении моно-интервальных оценок часто используют полученные методом Гурвица [Hurwicz, 1951] или другими подобными методами точечные оценки, «эквивалентные» исходным интервальным. Подчеркнем, что такая замена не осуществляется автоматически, а базируется на предпочтениях ЛПР, ее/его склонности к риску. Вместе с тем переход от интервальных оценок к точечным не содержит процедур, явным образом отражающих связь получаемых точечных оценок с предпочтениями ЛПР. Недостаточная наглядность такого перехода не позволяет ЛПР в полной мере получить представление об обоснованности выбора той или иной эквивалентной точечной оценки.

Мы предлагаем метод непосредственного сопоставления интервальных величин, процесс сравнения в котором трактуется как задача принятия решений. В соответствии с методом в паре сравниваемых интервалов выделяются три зоны, зона, благоприятствующая проверяемой гипотезе о предпочтительности одной из оценок, зона, отвергающая гипотезу, и зона эквивалентности оценок. Вывод о предпочтительности или эквивалентности интервальных оценок делается на основе сопоставления размеров зон и предпочтений ЛПР. Метод позволяет на основе суждений ЛПР либо указать лучшую интервальную оценку из сравниваемых, либо рекомендовать ЛПР временно отказаться от принятия

решения из-за опасности совершить ошибку второго рода. Иначе говоря, из-за опасности выбора в качестве лучшего интервала, не являющегося таковым. В случае получения такого сигнала, вероятно, целесообразно воздержаться от немедленного принятия решения и попытаться сузить диапазон неопределенности. Мы покажем, что в ряде случаев полученные с помощью предлагаемого метода результаты отличаются от результатов, получаемых при сравнении точечных оценок, замещающих интервальные. Указанное сопоставление позволяет сделать вывод, что метод непосредственного сравнения интервалов более адекватен специфике задач принятия решений, для которых характерно наличие понятия несравнимости, отсутствующее при использовании точечных оценок.

Рассмотрим два числовых интервала $I_i = [L_i, R_i]$, заданных их левыми L_i и правыми R_i границами, $L_i < R_i$. С точностью до перестановки I_1 и I_2 возможны следующие четыре варианта их взаимного расположения:

- $R_1 < L_2$.
- $L_1 < L_2 < R_1 < R_2$.
- $L_1 < L_2 < R_2 < R_1$.
- $L_1 = L_2, R_1 = R_2$.

Они показаны на рисунке 1. Мы будем искать условия, при которых I_2 может быть предпочтительнее («больше»), чем I_1 .

Конфигурация а) не вызывает трудностей: при любых возможных будущих реализациях значений i_1 и i_2 в интервалах I_1 и I_2 соответственно, интервал I_1 не может превзойти I_2 , т.е. всегда $I_2 \succ I_1$.

В работе [Воцинин, Сотиров, 1989] указано, что а) – единственная конфигурация, для которой интервальные оценки сравнимы в строгом смысле слова. В этом случае фактически нет задачи принятия решений, поскольку нет неопределенности: интервал I_2 строго больше интервала I_1 .

Конфигурация д) характеризуется наибольшей неопределенностью, интервалы эквивалентны друг другу по предпочтительности. В этой ситуации ЛПР не в состоянии принять рациональное решение в пользу какого-либо из интервалов без дополнительной информации. Необходимо либо отказаться от выбора, либо, если сделать выбор все-таки нужно, оперативно получить дополнительную информацию невозможно, воспользоваться случайным выбором. Из этих примеров видно, что мерой неопределенности в рассматриваемой задаче сравнения может служить протяженность D пересечения $I_1 \cap I_2$.

Анализ конфигураций б) и с) позволяет выделить еще две меры, играющие важную роль при сравнении интервалов.

Именно, из рассмотрения конфигурации б) следует, что если для I_1 текущая реализация $i_1 \in [L_1, L_2]$ или для I_2 текущая реализация $i_2 \in [R_1, R_2]$, то эти ситуации благоприятствуют тому, чтобы I_2 оказался предпочтительнее I_1 . Назовем зону благоприятствования зеленой зоной. Протяженность интервала $[L_2, R_1]$ характеризует размеры зоны неопределенности при сравнении I_1 и I_2 . Назовем зону неопределенности желтой зоной. Таким образом, в конфигурации б) шансы на то, что I_2 окажется предпочтительнее I_1 , зависят от соотношения протяженностей зон. Можно видеть, что протяженность зеленой зоны $D_g = L_2 - L_1 + R_2 - R_1$, а желтой зоны $D_y = R_1 - L_2$.

В конфигурации с) появляется неблагоприятная («красная») зона протяженностью $D_r = R_1 - R_2$. Здесь протяженность зеленой зоны $D_g = L_2 - L_1$, а желтой зоны $D_y = R_2 - L_2$. Соотношения между

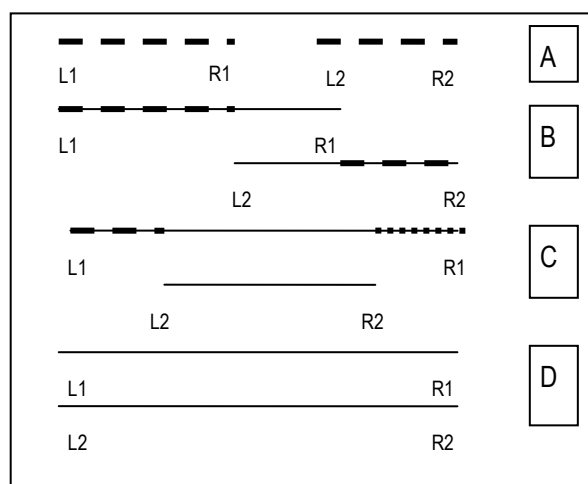


Рис. 1. Варианты расположения интервальных оценок

протяженностями зон в сопоставляемых парах оценок разных конфигураций определяют возможные результаты сравнения интервалов. Указанные соотношения удобно представить в виде числовых коэффициентов.

Среди нескольких возможных числовых коэффициентов, характеризующих ситуацию принятия решений, один показатель K , названный нами коэффициентом уверенности, имеет ясный экономический смысл. Для произвольной конфигурации этот коэффициент равен $K = [Dg - Dr]/D(I_1U_2)$, т.е. доле разности протяженностей зеленой и красной зон в общей протяженности сравниваемых интервалов с учетом их возможного пересечения.

Содержательно он характеризует относительный прирост возможного максимального выигрыша за счет верного принятия решений. Действительно, в случае конфигурации b), например, при выборе интервала I_2 в качестве предпочтительного максимум возможного выигрыша от такого выбора равен $R_2 - L_1$. При выборе интервала I_1 эта величина равна $R_1 - L_2$. Таким образом, относительный прирост возможного максимального выигрыша за счет верного принятия решений равен $[(R_2 - L_1) - (R_1 - L_2)]/(R_2 - L_1)$, т.е. K . Отметим, что если бы проверялась гипотеза о предпочтительности интервала I_1 , зеленая и красная зоны поменялись бы местами. Вычисленное в конкретной ситуации сравнения интервальных величин значение коэффициента уверенности K может быть использовано ЛПР при принятии решений. Именно, если при проверке гипотезы о предпочтительности интервала I_2 значение коэффициента K для сравниваемой пары интервальных оценок окажется «достаточно большим» и согласуется с представлениями ЛПР о приемлемой величине риска, который связан с принятием решения и который измеряется назначаемым ЛПР пороговым значением K_{th} (индивидуальным для каждого ЛПР и ситуации принятия решений), то принимается, что $I_2 \succ I_1$, если $K_{th} \geq K$. Из других соображений сходный с K «показатель интервального неравенства» предложен в работе [Ащепков, Давыдов, 2006].

Сопоставим результаты сравнения интервалов в конфигурации b) с соответствующими результатами для точечных оценок. Привлекая критерий Гурвица, для точечных оценок Es_i интервалов I_1 и I_2 получаем: $Es_i = (1 - \lambda)L_i + \lambda R_i$, $0 \leq \lambda \leq 1$. Для разности точечных оценок имеем: $Es_2 - Es_1 = (1 - \lambda)(L_2 - L_1) + \lambda(R_2 - R_1)$. Видно, что в конфигурации b), когда $L_2 > L_1$ и $R_2 > R_1$, $Es_2 > Es_1$ при всех λ , $0 \leq \lambda \leq 1$. Это означает, что в соответствии с точечными оценками для конфигурации b) всегда $I_2 \succ I_1$. В методе прямого сравнения интервалов это не всегда так. При большой зоне неопределенности (желтой зоне), значительно превосходящей размеры зеленой зоны, осторожный инвестор, опасаясь сделать ошибку второго рода, вряд ли согласится с выводом $I_2 \succ I_1$.

Сравнение интервалов конфигурации c) по точечным оценкам критерия Гурвица показывает, что $I_2 \succ I_1$, если $0 \leq \lambda \leq 1/[1 + (R_1 - R_2)/(L_2 - L_1)] = 1/[1 + Dr/Dg]$. Здесь вновь не учитываются (возможно, большие) размеры зоны неопределенности, что может не позволить ЛПР принять решение в соответствии с точечными оценками.

Например, если ЛПР принял $\lambda = 0.4$, то, в соответствии с точечными оценками, $I_2 \succ I_1$ при $Dr \leq 1.5 Dg$. Нетрудно подобрать пару интервалов, у которых это условие выполняется, но протяженность зоны неопределенности (D) столь велика, что не позволит с доверием относиться к выбору по точечным оценкам.

Следует отметить еще одно обстоятельство. При попарном сравнении трех интервалов A , B , C может случиться так, что для пары A , B интервал B предпочтительнее A с, например, $K = 65\% > K_{th} = 60\%$, а для пары A , C интервал C предпочтительнее A с $K = 80\% > K_{th}$. Это не означает, что C предпочтительнее B , потому что при непосредственном сравнении C и B может оказаться, что C «лучше» B с $K < K_{th}$, т.е. для ЛПР, принявшего для выбора величину K_{th} , эти интервалы несравнимы. Таким образом, всегда необходимо прибегать к непосредственному попарному сравнению всех интервалов-альтернатив с учетом предпочтений ЛПР.

Отметим также, что в некоторых задачах, таких как сопоставление чистых дисконтированных доходов инвестиционных проектов, вывод $I_2 \succ I_1$ не означает еще, что следует выбрать второй проект, если левая

граница I_2 лежит в отрицательной области. Дополнительно необходимо сравнить интервалы $[L_2, R_2]$ и $[L_2, 0]$. Если окажется, что первый интервал предпочтительнее, такой проект может быть рекомендован инвестору. Эта операция может быть проделана и до непосредственного сравнения исходных интервальных оценок.

Сопоставление с вероятностными мерами сравнения интервальных оценок

При прямом сравнении интервалов не предполагается наличия дополнительных знаний о шансах на реализацию значений рассматриваемых интервальных величин. Наличие такой дополнительной информации вносит определенные коррективы в результаты сравнения интервалов. Рассмотрим этот вопрос более подробно.

Пусть i_1, i_2 случайные величины, определенные на интервалах I_1, I_2 соответственно. Найдем вероятность того, что $I_2 \succ I_1$ для конфигураций b) и c). В случае конфигурации b) вероятность $P(I_2 \succ I_1)$ того, что интервал I_2 будет предпочтительнее интервала I_1 , равна вероятности P_g нахождения i_1 и i_2 в благоприятной зоне, $P_g = P(L_1 < i_1 < L_2) + P(L_2 < i_1 < R_1)P(R_1 < i_2 < R_2)$, а для конфигурации c) $P_g = P(L_1 < i_1 < L_2)$. Отметим, что для конфигурации c) P_g не зависит от типа распределения на I_2 . Будем считать, что $I_2 \succ I_1$, если $P_g \geq K_{pr}$, - порогового значения, выбранного ЛПР для сравнения альтернатив по вероятности.

Сравним «силу» решающих правил метода прямого сравнения и сравнения по вероятности для простого случая равномерных распределений на сопоставляемых интервалах. Тогда в случае b) $P_g = 1 - (R_1 - L_2)^2 / (D_1 D_2)$, где $D_i = R_i - L_i$ - протяженности сравниваемых интервалов. Сопоставляя предельные значения коэффициентов уверенности в конфигурации b) для метода прямого сравнения $K_{in} = 1 - (R_1 - L_2) / (R_2 - L_1)$ и сравнения по вероятности $K_{pr} = P_g$, получаем: $P_g = K_{in} + [(L_2 - L_1)(R_2 - R_1)(R_1 - L_2)] / [(R_2 - L_1)D_1 D_2]$. Поскольку второе слагаемое в этом соотношении положительно, то, если признать $I_2 \succ I_1$ с коэффициентом уверенности K_{in} при использовании метода прямого сравнения, этот вывод окажется справедливым с еще большим коэффициентом уверенности для метода сравнения по вероятности. К аналогичному выводу можно прийти для конфигурации c), где $P_g = K_{in} + (R_1 - R_2) / D_1$.

Этот вывод несправедлив, вообще говоря, для некоторых типов конфигураций при произвольных типах распределений вероятностей. Покажем это на примере треугольного распределения, которое является приемлемой моделью произвольного унимодального распределения, заданного на интервале конечной протяженности.

В случае конфигурации c) при $M_1 < L_2$, где M_1 - мода соответствующего треугольного распределения, P_g и K_{in} связаны соотношением $P_g = K_{in} + (R_1 - R_2) / D_1 + (R_1 - R_2)(L_2 - M_1) / [D_1(R_1 - M_1)]$, а при $M_1 > L_2$ $P_g = [K_{in} + (R_1 - R_2) / D_1](L_2 - L_1) / (M_1 - L_1)$. Видно, что $P_g \geq K_{in}$ и предыдущий вывод остается справедливым.

Для конфигурации b) отношения между коэффициентами P_g и K_{in} зависят от местоположения мод распределений. Мы не будем приводить громоздкого соотношения, связывающего P_g и K_{in} для этой конфигурации, а дадим формулы для вероятностей - компонентов P_g : $P(L_1 < i_1 < L_2) + P(L_2 < i_1 < R_1)P(R_1 < i_2 < R_2)$: $P(L_2 < i_1 < R_1) = 1 - P(L_1 < i_1 < L_2)$,

$$P(L_1 < i_1 < L_2) = \begin{cases} \frac{(L_2 - L_1)^2}{D_1(M_1 - L_1)}, & M_1 > L_2 \\ 1 - \frac{(R_1 - L_2)^2}{D_1(R_1 - M_1)}, & M_1 \leq L_2 \end{cases} \quad P(R_1 < i_2 < R_2) = \begin{cases} \frac{(R_2 - R_1)^2}{D_2(R_2 - M_2)}, & M_2 > R_1 \\ 1 - \frac{(R_1 - L_2)^2}{D_2(M_2 - L_2)}, & M_2 \leq R_1 \end{cases}$$

Можно показать, что при разных положениях мод возможны как ситуации с $P_g \geq K_{in}$, так и с $P_g < K_{in}$.

Заключение

Предложенный в работе метод непосредственного сравнения интервальных величин позволяет на основе предпочтений ЛПР либо указать лучшую интервальную оценку из сравниваемых, либо рекомендовать

ЛПР временно отказаться от принятия решения из-за несравнимости интервалов и опасности совершить ошибку второго рода. Мы полагаем, что этот метод более адекватен специфике задач принятия решений, чем распространенный метод замены интервальных оценок точечными, так как, ввиду своей наглядности, позволяет ЛПР более полно отразить свои предпочтения. При этом в ряде случаев полученные с помощью предлагаемого метода результаты отличаются от результатов, получаемых при сравнении точечных оценок, замещающих интервальные.

Предложенный метод сравнения моноинтервальных оценок позволяет реализовать один из возможных методов сопоставления полиинтервальных оценок, основанный на привлечении метода статистических испытаний. Другой возможный метод сравнения ПИО, связанный с анализом возможных конфигураций относительной локализации пар ПИО, также требует использования метода сопоставления моноинтервальных оценок. Эта возможность будет рассмотрена позже.

Благодарности

Статья частично финансирована из проекта **ITHEA XXI** Института Информационных теории и Приложений FOI ITHEA и Консорциума FOI Bulgaria (www.ithea.org, www.foibg.com).

Библиография

- [Ащепков, Давыдов, 2006] Л. Ащепков, Д. Давыдов. Показатель интервального неравенства: свойства и применение //Вычислительные технологии. Т. 11, сс. 13 – 22. 2006.
- [Вощинин, Сотиров, 1989] А. Вощинин, Г. Сотиров. Оптимизация в условиях неопределенности. М.: МЭИ, София: Техника. 1989.
- [Калмыков и др., 1986] С. Калмыков, Ю. Шокин, З. Юлдашев. Методы интервального анализа. Новосибирск: Наука. 1986.
- [Стернин, Шепелев, 2003] М. Стернин, Г. Шепелев. Метод представления знаний в интеллектуальных системах поддержки экспертных решений. // Новости искусственного интеллекта. Т. 4, сс. 11 – 15. 2003.
- [Стернин и др., 2005] М. Стернин, Н. Чугунов, Г. Шепелёв. Модели предметных областей в компьютерных системах, основанных на знаниях // Методы поддержки принятия решений. Труды Института системного анализа Российской академии наук. Т. 12, сс. 95 – 113. 2005.
- [Стернин и др., 2007] М. Стернин, Г. Шепелев, Н. Шепелев. Свойства обобщенного равномерного распределения вероятностей // Вторая международная конференция «Системный анализ и информационные технологии» (САИТ-2007). Труды конференции Т. 1, сс. 239 – 242. 2007.
- [Стернин, Шепелев, 2008] М. Стернин, Г. Шепелев. Обобщенные интервальные оценки в сценарном анализе. // Supplement to International Journal Information Technologies and Knowledge. V. 2, pp. 87 – 94. 2008.
- [Chugunov et al., 2008] N. Chugunov, G. Shepelyov, M. Sternin. The generalized interval estimations in decision making under uncertainty. // Int. J. Technology, Policy and Management. V. 8, pp. 298 – 321. 2008.
- [Hurwicz, 1951] L. Hurwicz. Optimality criteria for decision making under ignorance. 'Cowles Commission Discussion Paper', Statistics, 1951, # 370, New Haven.
- [Shepelyov, Sternin, 2003] G. Shepelyov, M. Sternin. Method of Generalized Interval Estimations for Intelligent DSS. // DSS in the Uncertainty of the Internet Age. The Karol Adamiecki University of Economics in Katowice, Katowice, pp. 367-377. 2003.

Информация об авторах

Михаил Стернин – старший научный сотрудник Института системного анализа РАН. Россия, 117312, Москва, просп. 60-летия Октября, ИСА РАН; e-mail: mister@isa.ru

Геннадий Шепелёв – заведующий лабораторией Института системного анализа РАН. Россия, 117312, Москва, просп. 60-летия Октября, ИСА РАН; e-mail: gis@isa.ru

ИДЕНТИФИКАЦИЯ ОЦЕНКИ ПРОТИВОПОЖАРНОГО СОСТОЯНИЯ ОБЪЕКТА ДЛЯ ФОРМИРОВАНИЯ БАЗЫ ЗНАНИЙ ИАС «БЕЗОПАСНОСТЬ»

Александр Джулай, Артем Быченко, Андрей Левченко

Аннотация: В статье рассмотрены задачи идентификации оценки противопожарного состояния объектов, используемой при разработке банка знаний системы поддержки принятия решений «Безопасность» и предназначенной для информационно-консультативного сопровождения процессов принятия решений руководителями пожарных подразделений.

Ключевые слова: идентификация, банк знаний, пожарная безопасность

ACM Classification Keywords: H.4 Information Systems Applications, J.6 Computer-aided Engineering

Conference: The paper is selected from XVth International Conference "Knowledge-Dialogue-Solution" KDS 2009, Varna, Bulgaria, June-July 2009

Аспекты проблемной области

Для обеспечения высокого уровня защиты населения пожарные подразделения должны использовать результаты прогнозирования. Особенно важно это для Украины в условиях ограниченности ресурсов. Использование методов прогнозирования, разработка и создание механизмов принятия решений, использующих результаты прогнозирования, и соответствующие способы реагирования приобретают в современных условиях особенную актуальность.

Каждый год в Украине регистрируют, в среднем, более 40000 пожаров, из которых значительная часть происходит в жилом секторе. Ведение и анализ такой большой информационной базы могли бы быть использованы для получения новых знаний, прогнозирования количества и структуры будущих пожаров и выполнения мероприятий по их эффективному предотвращению и ликвидации.

Разработку и реализацию такой базы знаний целесообразно разделить на две основных части: параметры распространения, исследование пожаров и оценка пожарной безопасности объекта.

Нечеткие базы знаний, как известно, отображают такую структуру: «если А, то В». При этом и А, и В в действительности являются функциями принадлежности, определяющими уверенность эксперта в том, что входной фактор и результирующая характеристика (время, скорость распространения огня) получают значения из определенных ограниченных множеств. База знаний может содержать значительное количество правил, которые определяются детализацией и точностью экспертных предположений. Функции принадлежности в такой базе представлены своими параметрами, количество которых определяется их формой. Традиционно используют треугольные, трапециевидные и гауссовские функции принадлежности. Расчеты в последнем случае упрощаются, но их получить и верифицировать достаточно сложно.

Постановка и решение задач

Скорость распространения и последствия пожаров на объектах определяют актуальность и необходимость разработки соответствующих моделей и использования моделирования для определения наиболее вероятных направлений путей и времени распространения огня. В работе [1] определены основные факторы, которые влияют на скорость развития пожара и изменение его периметра. Первой

задачей, которую необходимо решать при создании информационно-аналитической системы моделирования процессов развития пожаров, является формирование базы знаний. Она должна содержать статистические данные о прошедших пожарах, нормативных величинах, экспертных данных и данных, которые определяют размещение определенных объектов.

Значительная часть таких факторов имеет субъективный характер происхождения, что определяется:

- неполнотой данных, связанной с невозможностью осуществить измерение в критических условиях, а также их потерей;
- неопределенностью, которая возникает благодаря интервальному заданию множества значений параметров и экспертным предположением о возможных областях их изменения.

Известно, что прогнозирование базируется на анализе ретроспективной информации и идентификации искомой зависимости. В случае особо опасных объектов это невозможно, поскольку каждый объект является нетипичным, площадь и периметр пожара достаточно большие и его описание, в форме пригодной для расчетов, чаще всего, отсутствует. Поэтому традиционно используются справочные данные о скорости распространения пожара в помещениях в зависимости от их типа, скорости выгорания материалов, температуре пламени, влияющем на разрушение конструкций и веществ, экспертные оценки [2]. В критических условиях учесть все значимые факторы, которые влияют на скорость распространения огня и его направление, значения разнородных констант достаточно тяжело, поэтому экспертные оценки необходимы, в первую очередь, для того, чтобы учесть состояние материала, по которому распространяется огонь, а также наличие внутренних и внешних факторов, влияющих на его скорость. Разнотипность значений факторов, критические условия вызывают значительные расхождения экспертных заключений, что осложняет принятие решений.

Таким образом, имеем задачу идентификации зависимости:

$$V = F(V_l, V_m, T_f, T_p, X_{in}, X_{out}) \quad (1)$$

где V_l – линейная скорость распространения огня в помещениях определенного типа ($V_l \in [a, b]$),

V_m – средняя скорость выгорания материала ($V_m = const$),

T_f – температура пламени горения материала ($T_f \in [t_1, t_2]$),

T_p – температура плавления конструкции ($(T_p \in [q_1, q_2]) \vee (T_p = q)$),

X_{in}, X_{out} – внутренние и внешние факторы, соответственно.

Очевидно, что каждый из параметров V_l, V_m, T_f, T_p зависит от X_{in}, X_{out} . Неявным образом эти зависимости присутствуют в экспертных оценках, которые являются решающими в процессах принятия решений при пожаротушении. В то же время, такой субъективизм, как было указано выше, не способствует оптимизации выбора главного направления пожаротушения и распределения сил и средств, в связи с чем, предлагаем использовать нечеткие базы знаний для идентификации и прогнозирования скорости распространения огня.

Производной от задачи (1) является задача идентификации времени достижения огнем определенной точки

$$T_{x,y} = G(V, S) = G(V_l, V_m, T_f, T_p, X_{in}, X_{out}, S), \quad (2)$$

где S – путь распространения огня до точки с координатами (x, y) . Очевидно, что задача (2) является более общей и применимой для поддержки принятия решений при пожаротушении, поскольку, в отличие

от задачи (1), в ней учтено наличие препятствий на пути пожара, разнотипность помещений и другие факторы.

Следующим этапом после формирования такой базы данных является идентификация зависимости (2). При этом возникает проблема работы с нечеткими множествами, функциями принадлежности и процедурами их структурной и параметрической идентификации. Стоит также учитывать то, что использование (2) базируется на представлении входной информации в нечетком виде, а также то, что необходимо интерпретировать результат, который будет иметь вид функции принадлежности. Важную роль играет решение задачи параметрической оптимизации функций принадлежности.

Идентификация зависимости (2) может быть выполнена при разных условиях и предположениях, среди которых основными являются такие:

1. Рассматриваются суждения одного эксперта.
2. Учитываются суждения многих экспертов с одинаковыми или неизвестными уровнями компетентности.
3. Суждения экспертов имеют разный уровень значимости (компетентности экспертов разные).
4. Путь распространения огня является однородным, то есть структура помещений и препятствий является неизменной независимо от его маршрута.
5. Пожар может распространяться разными маршрутами с разным количеством помещений и препятствий на его пути с учетом замедляющих и ускоряющих факторов.

На начальном этапе решения задачи идентификации (2) будем считать, что принятие решений зависит от одного эксперта, а путь распространения огня является однородным, то есть выполняются первое и четвертое условия. Поскольку эксперт один, то рационально предполагать, что его суждения о значении параметров, определяющих время распространения огня от точки до точки, описываются симметричными треугольными функциями принадлежности.

Для определения оценки противопожарного состояния объекта необходимо решить задачу идентификации. Исходными данными является статистическая информация о пожарах. Традиционно, осуществляя ее последовательный анализ, определялись с видом зависимости, то есть выполняли структурную идентификацию. Этот процесс отмечался значительным присутствием субъективизма и ограниченностью мощности множества потенциальных зависимостей. Решение другой подзадачи – параметрической идентификации традиционно базировалось на методе наименьших квадратов, одной из особенностей которого является необходимость проверки значительного количества требований и предположений. Решив вышеуказанные проблемы, можно утверждать об эффективном оценивании уровня пожарной безопасности. В соответствии с этапами системного анализа [3], выполним формализацию внутренних и внешних факторов, определяющих пожарную безопасность объектов. Выделим основные факторы первой группы: X_1 – дата построения; X_2 – количество этажей; X_3 – их планирование; X_4 – наличие подвалов и чердаков; X_5 – структурные особенности дома, имеющие отношение к особенностям развития, тушения и последствий пожара; X_6 – соответствие нормативным требованиям материалов и конструкции; X_7 – наличие и состояние путей эвакуации. К факторам, которые определяют противопожарное состояние объекта, отнесем такие: Z_1 – расстояние до ближайшей пожарной части; Z_2 – количество пожарных частей в 20-километровой зоне; Z_3 – степень укомплектованности ближайшего пожарного подразделения кадрами; Z_4 – укомплектованность средствами пожаротушения; Z_5 – наличие и вид противопожарного водоснабжения; Z_6 – диаметр водопровода; Z_7 – состояние

ближайших источников водоснабжения; Z_8 – давление в водопроводе; Z_9 – наличие первичных средств тушения; Z_{10} – степень удобства подъезда; Z_{11} – наличие погодноклиматических факторов, осуществляющих влияние на эффективность пожаротушения.

К исходным характеристикам, которые необходимо идентифицировать, принадлежат: Y_1 – количество огнетушащих веществ, которые были использованы; Y_2 – время разворачивания сил и средств; Y_3 – время прибытия первого подразделения; Y_4 – время ликвидации пожара; Y_5 – убытки от пожара; Y_6 – ошибки пожарных подразделений.

Формализуем факторы как тройку элементов:

$$(Z, P, F), \quad (3)$$

где $Z = (Z_1, Z_2, \dots, Z_n)$ – способы и средства пожарной защиты; P – механизм прогнозирования. Результатом использования P является прогнозирование возникновения пожара, который может произойти в интервале времени $t_1 > t$ и оценка вероятности того, что она произойдет. Областью значений функции F является отрезок $[0;1]$, причем:

- - $F(z, p) = 1$, если за время t_2 система пожарной защиты способна применить все средства, являющиеся адекватными угрозе;
- - $F(z, p) = 0$, если время применения является большим t_2 или адекватных средств нет;
- - $F(z, p) \in (0;1)$ – в промежуточных случаях.

Таким образом, задача (3) заключается в идентификации векторной зависимости:

$$Y = F(X, Z), \quad (4)$$

где $Y = (Y_1, Y_2, \dots, Y_6)$, $X = (X_1, X_2, \dots, X_7)$, $Z = (Z_1, Z_2, \dots, Z_{11})$. Отметим, что статистические данные являются зашумленными, поскольку процессы, связанные с вычислительными процедурами и оценкой ситуации, происходят в условиях временного цейтнота. Не выполняя тестирования на мультиколлинеарность, гетероскедастичность и автокорреляцию, согласно метода наименьших квадратов для исходных характеристик получены такие результаты:

Для Y_1 :

$$Y_1 = 0,59X_1 + 0,31X_2 + 6,33X_3 - 0,92X_4 + 4,88X_5 + 10,4X_6 - 3,78X_7 - 1,3Z_1 - \\ - 0,8Z_2 + 7,5Z_3 - 15Z_4 - 5,3Z_5 + 1,9Z_6 - 5,8Z_7 + 4,3Z_8 + 6,72Z_9 - 6Z_{10} - 12Z_{11}.$$

Среднее квадратичное отклонение на тестовой выборке $\sigma_1 = 22,98$.

Для Y_2 :

$$Y_2 = 0,3X_1 + 0,49X_2 - 0,2X_3 - 3,66X_4 + 2,4X_5 + 7,19X_6 - 1,33X_7 - 0,55Z_1 - \\ - 0,5Z_2 + 3,7Z_3 - 7Z_4 - 3,9Z_5 + 0,4Z_6 - 2,85Z_7 + 3,6Z_8 + 3,72Z_9 - 2,6Z_{10} - 6Z_{11}.$$

Среднее квадратичное отклонение на тестовой выборке $\sigma_2 = 11,9$.

Для Y_3 :

$$Y_3 = 0,35X_1 - 0,84X_2 - 1,55X_3 - 8,87X_4 + 8X_5 + 10X_6 + 1,7X_7 + 3,76Z_1 + \\ + 0,77Z_2 + 6,97Z_3 + 21Z_4 - 24Z_5 - 1,7Z_6 - 2Z_7 + 0,8Z_8 + 3Z_9 - 5,7Z_{10} - 2Z_{11}.$$

Среднее квадратичное отклонение на тестовой выборке $\sigma_2 = 22,3$.

Для Y_4 :

$$Y_4 = 1,6X_1 - 0,47X_2 + 0,93X_3 - 23,3X_4 + 7,4X_5 + 25X_6 + 17X_7 + 7,5Z_1 + \\ + 1,8Z_2 + 15Z_3 + 7Z_4 - 46Z_5 - 2,5Z_6 - 2,3Z_7 - 1,3Z_8 - 14Z_9 - 11Z_{10} - 10,8Z_{11}.$$

Среднее квадратичное отклонение на тестовой выборке $\sigma_2 = 50,9$.

Для Y_5 :

$$Y_5 = -0,1X_1 + 0,28X_2 - 8X_3 - 2,1X_4 - 1,3X_5 + 5,2X_6 + 3,12X_7 + 0,4Z_1 + \\ + 0,57Z_2 + 3,05Z_3 - 4,24Z_4 - 6,2Z_5 - 0,7Z_6 - 4Z_7 + 2Z_8 - 3Z_9 - 0,2Z_{10} - 2,3Z_{11}.$$

Среднее квадратичное отклонение на тестовой выборке $\sigma_2 = 14,2$.

Для Y_6 :

$$Y_6 = 0,1X_1 + 0,018X_2 + 1,12X_3 - 0,2X_4 + 0,85X_5 + 1,12X_6 - 0,46X_7 + 0,2Z_1 - \\ - 0,1Z_2 + 0,4Z_3 - 0,4Z_4 + 0,07Z_5 + 0,6 - 0,1Z_7 + 0,4Z_8 + 0,3Z_9 - Z_{10} - 0,8Z_{11}.$$

Среднее квадратичное отклонение на тестовой выборке $\sigma_2 = 4,6$.

Анализируя коэффициенты при переменных, которые обозначают внутренние и внешние факторы, можно делать выводы об их влиянии на результирующую характеристику. Но такие оценки в результате перечисленных обстоятельств будут смещенными. Одним из путей преодоления такой проблемы является выполнение процедур препроцессинга данных. Более эффективным процесс идентификации будет при условии изъятия незначимых факторов и уменьшения шумовых эффектов.

При поступлении вызова в пожарной части по адресу объекта, на котором возник пожар, формируется информационно-аналитическая записка, которая содержит такие данные:

- значение внутренних и внешних параметров объекта;
- оценку уровня его пожарной безопасности с обоснованием;
- анализ ошибок, допущенных при тушении типичных объектов и в типичных ситуациях;
- рекомендованный маршрут проезда к месту пожара;
- наиболее вероятные пути и время распространения пожара.

Ядром информационно-аналитической системы «Безопасность» является банк математических моделей и методов (рис. 1).

Составляющими банка являются математические модели, реализованные в таких модулях:

- PLR – парная линейная регрессия;
- PNR – парная нелинейная регрессия;
- MLR – множественная линейная регрессия;
- NMR – множественная нелинейная регрессия;
- PKG – полином Колмогорова-Габора;
- NNFF – нейронные сети с прямым распространением сигнала;
- MKNN – нейронные сети с модульно-ядерной структурой;
- MATRIX – модели представления данных для задачи оптимизации проезда пожарного расчета.

Вычислительные методы инкапсулированы в таких модулях:

- PMLS – МНК для парной линейной регрессии;
- MMLS – МНК для множественной линейной регрессии;
- TEST1 – методы тестирования “негативных” явлений в исходных факторах;
- TESTN – другие методы тестирования;

- GMDH – многорядный алгоритм МГВА;
- CRA – методы корреляционно-регрессионного анализа;
- ALM – алгоритм Левенберга-Маркуарда;
- CGA – метод Глетчера-Ривса;
- BC – процедура “box-counting”;
- MMC – метод главных компонент;
- MWI – процедура “выбеливания” входов;
- GA – процедуры генетического алгоритма.

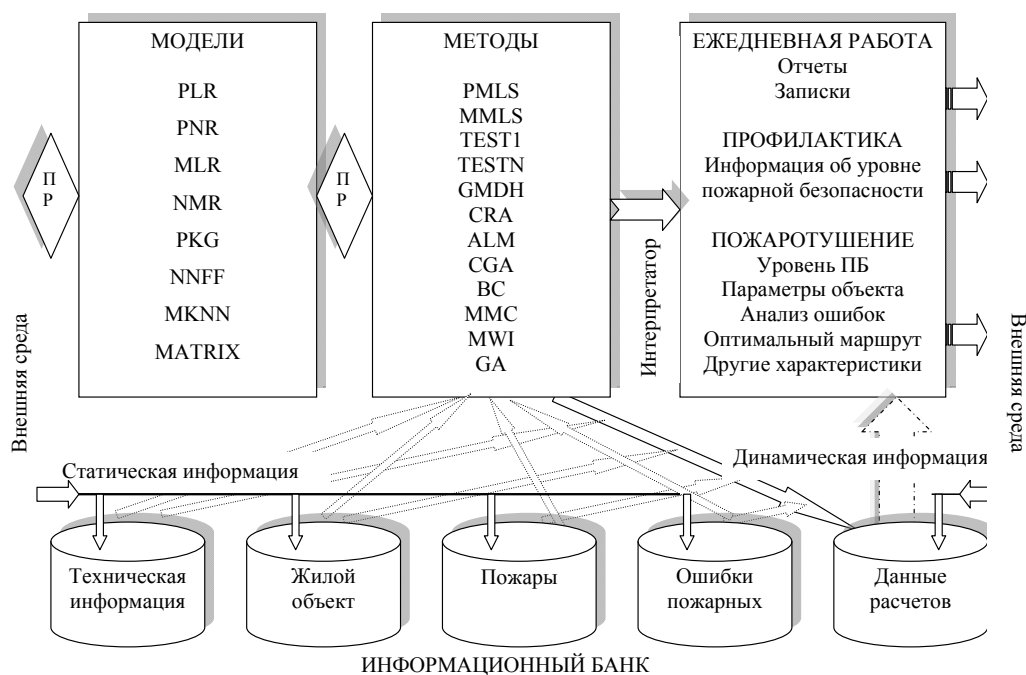


Рис. 1. Функциональная схема ИАС “БЕЗОПАСНОСТЬ”

Информационно-аналитическая система “БЕЗОПАСНОСТЬ” работает в двух режимах: текущем и оперативном. При текущем режиме: выполняется идентификация уровней пожарной безопасности объектов и формируются данные необходимые для отчетов;

- выполняются расчеты оптимальных путей проезда пожарного подразделения ко всем перекресткам зоны ответственности;
- при изменении параметров, от которых зависит время проезда, выполняется пересчет;
- выполняется анализ ошибок, устанавливаются связи между ними и определяются соотношения между пожаротушением определенных объектов и типами ошибок.

Для повседневного функционирования используется техническая информация, на базе которой формируются отчеты и строевые записки. Выполнение профилактических мероприятий базируется на данных расчетов, анализе приоритетности и значениях внутренних и внешних параметров жилых объектов.

Разработку базы знаний предлагается осуществлять на основе трех моделей: строения, функционирования и развития [3]. Модель строения является теоретико-множественной моделью и определяет состав базы знаний:

$$i_{\bar{a}} = \langle O, A_1(a_{11}, a_{12}, \dots, a_{1k}), A_2(a_{21}, a_{22}, \dots, a_{2k}), \dots, A_n(a_{n1}, a_{n2}, \dots, a_{nk}) \rangle, \quad (5)$$

где O - идентификатор объекта, $A_i(*, *, \dots, *)$ - идентификатор поля базы знаний, a_{ij} - 1-й формат i -го поля данных $i = \overline{1, n}, j = \overline{1, k}$.

Модель функционирования базы знаний позволяет определить, информация из каких полей будет использоваться при решении определенных задач. К таким задачам относятся:

- определение кратчайшего возможного маршрута распространения огня;
- определение наиболее возможного пути развития пожара;
- расчет усредненного времени и соответствующего маршрута распространения огня;
- формирование перечня возможных вариантов развития пожара в соответствии с заданными часовыми маркерами и тому подобное.

Модель функционирования является информационной моделью:

$$I_{\delta} = \langle P, A, \{A_{i_1}, A_{i_2}, \dots, A_{i_k}\} \rightarrow P_k, k = \overline{1, m} \rangle, \quad (6)$$

где P - непустое множество задач, A - множество атрибутов (полей) базы знаний, m - количество задач.

Модель развития формируется как определенное следствие процесса развития информационно-аналитической системы во времени, базирующееся на расширении знаний о пожарах, объектах пожаротушения, экспертной информации. Развитие базы знаний происходит в соответствии с параллельно-иерархической схемой, где ее уровни являются отображением процесса решения вышеуказанных задач в случаях привлечения дополнительных экспертов, изменения типа и формы функций принадлежности, которые являются отображением их суждений. На этапах такой схемы осуществляется последовательное уточнение решений, поскольку они ориентируют на решение одной задачи несколькими методами или на решение определенной проблемы путем решения последовательности уточняющих задач.

В модели развития интегрированы возможные варианты композиции задач и методов решения, а также учитывается необходимость расширения, уточнения и модификации записей в базе знаний. Модель развития формально представим так:

$$I_{\delta} = \langle t, P_{i_1} \rightarrow P_{i_2} \rightarrow (P_{i_1}, P_{i_2}) \rightarrow \dots \rightarrow (P_{i_1}, \dots, P_{i_q}), A_i^t \uparrow, A_i^t \downarrow, A_i^t \Downarrow \rangle, \quad (7)$$

где t - время, P_{i_j} - задачи, $A_i^t \uparrow, A_i^t \downarrow$ - увеличение или уменьшение атрибутов задачи, $A_i^t \Downarrow$ - их содержательная модификация.

Таким образом, разработка тройки моделей $\langle I_{\delta}, I_{\delta}, I_{\delta} \rangle$ является основанием для программируемого создания и сопровождения базы знаний для моделирования процессов распространения пожаров на различных объектах, что позволит реализовывать конструктивные построения, которые будут инвариантными к будущим ситуациям.

Перспективы исследований

Использование для поддержки принятия решений в процессах пожаротушения информационно-аналитических систем необходимо для объективизации субъективных решений. Предложенная процедура вычисления времени распространения пожара от одной точки к другой позволит руководителю процессом пожаротушения определять решающие направления боевого развертывания сил и средств, что особенно важно при ликвидации чрезвычайной ситуации. Моделирование с применением элементов теории нечетких множеств будет способствовать увеличению адекватности процессов принятия решений в реальной ситуации.

Такая информация позволит в режиме реального времени быстро оценивать ситуацию, принимать правильные решения, минимизировать вероятности ошибок, минимизировать время проезда пожарного подразделения к месту пожара, а также время его локализации и ликвидации, что будет способствовать предотвращению человеческих жертв и материальных убытков.

Благодарности

Статья частично финансированна из проекта **ITHEA XXI** Института Информационных теории и Приложений FOI ITHEA и Консорциума FOI Bulgaria (www.ithea.org, www.foibg.com).

Литература

1. Быченко А.А. Модели распространения пожара на особо опасных объектах в условиях неопределенности // Искусственный интеллект. – 2006. – № 3. – С. 359-365.
2. Миронов М.П., Маскаева Л.Н., Макурин Ю.Н. Применение комплексного вероятностного подхода к прогнозированию пожаров и оценке ущерба от них // Пожарная безопасность. – 2005.– № 5. – С. 110-114.
3. Тимченко А.А., Родионов А.А. Основы информатики системного проектирования объектов новой техники. – К.: Наук. думка, 1991. – 152 с.
4. Гнатиенко Г.Н., Снитюк В.Е. Экспертные технологии принятия решений. – К.: Маклаут, 2008. – 444 с.

Информация об авторах

Александр Джулай – нач. кафедры оперативно-тактической деятельности, Академия пожарной безопасности имени Героев Чернобыля, ул. Оноприенко, 8, Черкассы, Украина; e-mail: djulaj@ukr.net

Артем Быченко – доцент кафедры пожарной и промышленной автоматики, Академия пожарной безопасности имени Героев Чернобыля; ул. Оноприенко, 8, Черкассы, Украина; e-mail: bichenko@ukr.net

Андрей Левченко – нач. кафедры пожарной и промышленной автоматики, Академия пожарной безопасности имени Героев Чернобыля; ул. Оноприенко, 8, Черкассы, Украина; e-mail: lad_ck@rambler.ru

МУЛЬТИАГЕНТНЫЙ *H*-МЕТОД В КОМБИНАТОРНОЙ ОПТИМИЗАЦИИ

Леонид Гуляницкий, Денис Гобов

Аннотация: Предлагается метаэвристический алгоритм комбинаторной оптимизации, построенный на основе *H*-метода. В основе алгоритма лежит мультиагентный подход к исследованию пространства допустимых решений. Данный подход в отличие от стандартного *H*-метода, в котором между двумя точками в пространства решений строится только один отрезок специального вида, предполагает построение агентами нескольких отрезков. При работе агенты используют специальную модель решаемой задачи, что позволяет учитывать наряду с содержательной информацией о задаче и опыт, накопленный на предыдущих шагах алгоритма. Эффективность предложенного подхода проиллюстрирована на основе результатов вычислительного эксперимента по решению ряда задач коммивояжера и квадратичных задач о назначениях.

Ключевые слова: комбинаторная оптимизация, метаэвристики, *H*-метод, агенты, задача коммивояжера, квадратичная задача о назначениях.

ACM Classification Keywords: G.1.6 [Numerical Analysis] Optimization – Stochastic programming, G.2.1 [Discrete Mathematics] Combinatorics – Combinatorial algorithms, I.2.8 [Artificial Intelligence]: Problem Solving, Control Methods, and Search – Heuristic methods.

Conference: The paper is selected from XVth International Conference "Knowledge-Dialogue-Solution" KDS 2009, Varna, Bulgaria, June-July 2009

Введение

Трудности, возникающие при решении задач комбинаторной оптимизации, общеизвестны. Практически все они принадлежат к классу NP-сложных задач. Вследствие быстро возрастающей вычислительной сложности при увеличении размерности существует необходимость в разработки методов и алгоритмов, которые позволяли бы получать пригодные для использования решения задачи при условии использования ограниченных вычислительных ресурсов. Одним из перспективных направлений развития алгоритмов для решения задач комбинаторной оптимизации является разработка метаэвристических (гибридных) алгоритмов. Обобщая варианты определений [1,2], можно сказать, что метаэвристики являются высокоуровневыми стратегиями для исследования пространства решений с помощью использования различных базовых методов. Большинство известных метаэвристик базируются на объединении популяционных алгоритмов – в первую очередь генетических алгоритмов – и алгоритмов поиска в окрестностях (локальный поиск, табу-поиск, имитационный отжиг) [2]. Одним из метаэвристических алгоритмов, показавших высокую производительность при решении различных задач комбинаторной оптимизации, является *H*-метод [3], построенный на основе синтеза алгоритма ускоренного вероятностного моделирования (*G*-алгоритм) и дискретного метода, названного методом деформированных многогранников. Также нельзя не отметить активное развитие мультиагентных методов интеллектуальной оптимизации. В качестве примера можно привести метод оптимизации муравьиными колониями (ОМК) [4], метод пчелиной колонии [5] и другие. Данные методы основаны на моделировании поведения насекомых, птиц, животных, поведение которых носит коллективный характер, за счет чего достигается эффект называемый коллективным интеллектом. В основе поведения данных

систем лежит самоорганизация, которая представляет собой множество динамических механизмов. В соответствии с данными механизмами система регулируется на глобальном уровне за счет взаимодействия ее компонентов на нижнем уровне без прямого взаимодействия.

Ниже предлагается метаэвристический алгоритм, построенный на основе H -метода и использующий мультиагентный подход для исследования пространства решений.

Общая схема H -метода

Под задачей комбинаторной оптимизации далее будем понимать поиск такого $x \in \Omega$, при котором заданная целевая функция $f(x)$ достигает оптимума, где Ω – некоторое локально конечное пространство.

Основная идея H -метода состоит в следующем. Формируется начальное множество решений, которое играет роль, отдаленно напоминающую роль многогранника в методе Нелдера-Мида [6]. Из этого множества некоторым образом выбираются несколько пар точек, для каждой из которых строится полуинтервал в пространстве вариантов решения задачи, проходящий через выбранные точки. Далее на нем ищется наилучшая с точки зрения целевой функции точка. Найденный вариант решения передается в качестве начального приближения для некоторого метода локальной оптимизации **Local_search**. В результате таких действий находится совокупность улучшающих точек. Эти точки включаются в многогранник вместо "наихудших" с точки зрения некоторого выбранного критерия оптимальности, и описанная процедура повторяется до выполнения критерия останова.

В данном методе существенную роль играет понятие полуинтервала, которое вводится на основании d -отрезка [7]. Пусть на пространстве Ω определена метрика d .

Определение 1. d -отрезком $/x, y/$, который соединяет две произвольные точки $x, y \in \Omega$, называется упорядоченная совокупность точек $x_i \in \Omega, i = 1, \dots, k$, которые удовлетворяют условиям:

- 1) $d(x, x_i) + d(x_i, y) = d(x, y), i = 1, \dots, k$,
- 2) $x_1 = x, x_k = y$, а $d(x, x_i) < d(x, x_{i+1}), i = 1, \dots, k - 1$,
- 3) не существует такой точки $z \in \Omega$, что $d(x_i, z) + d(z, x_{i+1}) = d(x_i, x_{i+1}), z \neq x_i, z \neq x_{i+1}, i = 1, \dots, k - 1$.

В дальнейшем будем рассматривать такие пространства, для которых $/x, y/ \setminus \{x, y\} \neq \emptyset$, если $d(x, y) > h$, где $h = \inf(d(x, y))$ для $x \neq y$ (в работе [7] они названы регулярными в метрике d).

Определение 2. Полуинтервал $< x, y/$ определим как совокупность $< x, y/ = /x, y/ \setminus \{x\}$.

Приведем более формально вычислительную схему H -метода, используя терминологию эволюционных вычислений (рис.).

В предлагаемом алгоритме используются параметры:

m – число вершин многогранника (число особей в популяции – в терминах популяционных алгоритмов);

k – количество пар вершин, которые выбираются для построения полуинтервалов;

l – количество вершин, которые подвергаются мутации;

Отбор Популяции – процедура отбора, например, $(m + (k + l))$ – стратегия: выбор m точек из $(m + k + l)$ точек популяции;

$L(y)$ – заданная окрестность точки $y \in \Omega$.

```

procedure H-Algorithm (x);
begin
   $h := 0; P^0 = \emptyset;$ 
  for  $j:=1$  to  $m$  do
     $x :=$  ГенерацияДопустимогоРешения;
     $P^h := P^h \cup \text{Local\_search}(x);$ 
  endfor;                                     { формирование начальной популяции }
  repeat
     $P := P^h;$ 
    for  $i:=1$  to  $k$  do
      ОтборДляВариации ( $x, y \in P$ );
       $z := \arg \min \{f(u) : u \in \langle x, x^\infty \setminus L(y) \rangle \& y \in \langle x, x^\infty \rangle\};$ 
       $P^h := P^h \cup \text{Local\_search}(z);$ 
    endfor;                                     {сформирована популяция из  $(m+k)$  точек}
    for  $i:=1$  to  $l$  do
      ОтборДляМутации ( $x \in P$ );
       $x :=$  Мутация ( $x$ );
       $P^h := P^h \cup \text{Local\_search}(x);$ 
    endfor;                                     {сформирована популяция из  $m+k+l$  точек}
     $P^h :=$  ОтборПопуляции( $P$ );
     $h := h + 1;$ 
  until не выполняется условие останова;
   $x := \operatorname{argmin}\{f(y) : y \in P\};$ 
  return  $x;$ 
end.

```

Рисунок. Схема *H*-метода

Мультиагентный *H*-метод

Важную роль для *H*-метода играет способ построения полуинтервалов $\langle x, x^\infty \rangle$ в комбинаторных пространствах. Такие пространства имеют две следующие характерные особенности. Во-первых, для произвольного $x \in \Omega$ "максимально удаленной" является точка, удовлетворяющая условию: $x^\infty = \max\{d(x, u) : u \in \Omega\}$, где $d(x, u)$ – метрика на Ω . Отметим, что зачастую это значение равняется диаметру пространства Ω . Во-вторых, между двумя не соседними точками $x, y \in P$ можно построить не один, а несколько d -отрезков. Поэтому при реализации *H*-метода необходимо отличать случай, когда из множества допустимых d -отрезков $\langle x, y \rangle$, $d(x, y) > 2$, всегда по некоторому конкретному правилу/способу построения выбирается только один d -отрезок, или когда могут рассматриваться все возможные d -отрезки. Вариант с одним возможным d -отрезком приводит к быстрой сходимости алгоритма, тогда как вариант с рассмотрением всех возможных d -отрезков приводит к более качественному анализу пространства в определенной области. Однако, исходя из схемы алгоритмы, можно сказать, что второй вариант будет отличаться повышенной трудоемкостью, так как для каждого полуинтервала, построенного

на основании d -отрезка, будет осуществляться поиск наилучшей точки с последующим применением к ней алгоритма локального поиска. Для решения этой проблемы и использования преимуществ мультиагентных методов предлагается строить агентами несколько полуинтервалов одновременно, используя при этом специальную модель задачи, и выбирать лучшую точку с учетом всех полуинтервалов. В роли агента выступает алгоритм пошагового построения полуинтервала, который при выборе следующей точки из допустимого множества использует определенное вероятностное правило.

Рассмотрим пространство перестановок, как одно из наиболее часто используемого пространства в задачах комбинаторной оптимизации. Вероятность выбора для точки x^k очередной точки $y^{ij} \in < x, x^\infty /$ такой, что $y^{ij} = \omega^{ij} x^k$, где ω^{ij} - оператор транспозиции (или подобный ему в случае другой метрики), с помощью которого генерируются соседние для x^k точки, будем рассчитывать на основании:

1. эвристических проблемно-зависимых значений, например $\Delta_{ij} = \frac{f(x^k) - f(y^{ij})}{f(x^k)}$;
2. степени «желательности» оператора ω , которая рассчитывается на основании модели задачи и использования накопленного опыта.

Для расчета степени «желательности» оператора ω^{ij} транспозиции для перестановки $x = (x_1, \dots, x_n)$ были предложены два варианта.

Модель задачи представляется в виде матрицы (τ_{ij}) размерности $n \times n$ (n – размерность задачи), где элемент τ_{ij} определяет «желательность» следования j после i (в варианте решения задачи).

Вариант 1. Транспозиция $i \leftrightarrow j$ компонентов x_i и x_j перестановки x будет определять следующую величину степени «желательности»:

$$\nabla_{ij} = \tau_{x_{i-1}x_j} + \tau_{x_jx_{i+1}} + \tau_{x_{j-1}x_i} + \tau_{x_ix_{j+1}} - \tau_{x_{i-1}x_i} - \tau_{x_ix_{i+1}} - \tau_{x_{j-1}x_j} - \tau_{x_jx_{j+1}}.$$

Данный метод расчета «желательности» имеет определенные аналогии с расчетом вероятности включения очередной вершины в маршрут на основании количества феромона в ОМК.

Вариант 2. Транспозиция $i \leftrightarrow j$ компонентов x_i и x_j перестановки x будет определять следующую величину степени «желательности»: $\nabla_{ij} = \tau_{x_ij} + \tau_{x_ji} - \tau_{x_i} - \tau_{x_j}$. Данный метод расчета

«желательности» имеет определенные аналогии с известным алгоритмом оценки распределения (Estimation of Distribution Algorithm) [8].

Оба варианта предполагают одинаковую процедуру обновления значений матрицы (τ_{ij}) . Она схожа с процедурой обновления феромонных следов в циклическом ОМК и состоит из двух этапов:

1. увеличение пропорционально найденному решению элементов матрицы, которые определенным образом ассоциируются с найденным решением;
2. применение уменьшающего коэффициента ко всем элементам (τ_{ij}) (испарение феромона).

Пусть α и β – степени значимости величин Δ_{ij} и ∇_{ij} , а $J_{st} = \{(s, t) : y^{st} \in \Pi(x^k)\}$. Тогда выбор оператора транспозиции ω^{ij} , который осуществляет создание очередной точки из допустимого множества $\Pi(x^k)$ для продолжения построения полуинтервала,

будет определяться на основании вероятности p_{ij} , для вычисления которой предлагается следующее выражение:

$$p_{ij} = \min \left\{ 1, e^{\alpha\Delta_{ij} + \beta\nabla_{ij}} / \sum_{(s,t): y^{st} \in \Pi(x^k)} e^{\alpha\Delta_{st} + \beta\nabla_{st}} \right\}.$$

Несложно увидеть, что данное выражение может быть преобразовано к следующему виду:

$$p_{ij} = \min \left\{ 1, 1 / \sum_{(s,t): y^{st} \in \Pi(x^k)} e^{\alpha(\Delta_{st} - \Delta_{ij}) + \beta(\nabla_{st} - \nabla_{ij})} \right\}. \quad (1)$$

Опишем схему предложенного алгоритма, который назовем мультиагентным H -методом или MH -методом.

1. Инициализация (τ_{ij}) : $\tau_{ij} = \tau_{\min}$, $i, j = \overline{1, n}$. $v :=$ <количество агентов >
 Другие начальные действия для обычного H -метода.
 2. Очередная итерация:
 - 2.1 Построение и обработка полуинтервала каждый агентом $u \in \{1, \dots, v\}$:
 - а) строим полуинтервал: на каждом шаге с вероятностью p_{ij} , рассчитанной по (1), выбираем следующую точку из $\Pi(x^k)$;
 - б) находим наилучшую точку \tilde{x}^u на построенном полуинтервале.
 - 2.2 Выбор лучшей (или по правилу Метрополиса) точки \tilde{x} из множества \tilde{x}^u .
 - 2.3 Шаги обычного H -метода.
 - 2.4 Обновление (τ_{ij}) .
 3. Если критерий завершения не выполняется, переход на п.2.

Вычислительный эксперимент

Для анализа практической эффективности предложенного алгоритма был проведен вычислительный эксперимент по решению задач коммивояжера (ЗК) и квадратичных задач о назначениях (КЗН). Задачи были взяты из библиотек QAPLIB [9] и TSPLIB [10], в которых для каждой задачи приведено наилучшее известное на данный момент решение. Это позволило не только получить сравнительные оценки по сравнению с H -VNSG алгоритмом (одной из самых успешных реализаций H -метода), но и оценить точность работы MH -метода.

Для всех алгоритмов использовались следующие значения параметров: $m=25$, $k=5$, $l=0$. Радиус окрестности $L(y)$ равнялся 2. Для MH -метода коэффициент испарения $\rho=0.5$, а $\alpha=1$, $\beta=3$, $v=10$. В качестве алгоритма локального поиска использовался, как и H -VNSG алгоритме, модифицированный алгоритм локального поиска с переменными окрестностями [11]. С целью получения статистически достоверных результатов, для каждой задачи проводилась серия экспериментов с одинаковыми параметрами алгоритма и разными начальными приближениями. Объем серии для каждой задачи составлял 50. Результаты решения ряда задач приведены в табл.1 и табл.2. Здесь N – размерность задачи, f^* - лучшее известное решение задачи, ε – средняя относительная погрешность алгоритма (%), ε_{\min} – лучшее

значение ε в серии экспериментов, t – среднее время, за которое было найдено лучшее решение в этой серии.

Анализ полученных результатов свидетельствует, что для большинства задач алгоритмом *МН*-метода позволил повысить точность решения, для остальных – показал такой же результат, что и *H-VNSG* алгоритм. Также следует отметить, что предложенная схема алгоритма, включая и использованный встроенный алгоритм локального поиска, является универсальной для всех задач комбинаторной оптимизации, что значительно расширяет область ее применения.

Таблица 1. Результаты решения ЗК

Название задачи	N	f^*	<i>H-VNSG</i> алгоритм			<i>МН</i> -алгоритм		
			ε	ε_{min}	t	ε	ε_{min}	t
Els19	19	17212548	0,00	0,00	16,88	0,00	0,00	18,75
chr20a	20	2192	0,20	0,00	20,66	0,00	0,00	19,39
Chr20b	20	2298	1,38	0,00	25,20	0,2	0,00	12,86
Chr20c	20	14142	0,00	0,00	2,82	0,00	0,00	5,95
Nug21	21	2438	0,00	0,00	11,35	0,00	0,00	12,59
Lip40a	40	31538	0,50	0,00	23,79	0,00	0,00	20,20
Lip40b	40	476581	0,00	0,00	5,63	0,00	0,00	5,65
Lipa50a	50	62093	0,27	0,00	32,70	0,23	0,00	39,34
Lipa50b	50	1210244	0,00	0,00	16,06	0,00	0,00	16,06
Sco81	81	90998	0,32	0,19	61,94	0,24	0,17	84,87
Will100	100	273038	0,23	0,18	91,92	0,17	0,14	199,54
Tai100b	100	1185996137	0,34	0,21	153,84	0,20	0,08	221,39

Таблица 2. Результаты решения КЗН

Название задачи	N	f^*	<i>H-VNSG</i> алгоритм			<i>МН</i> -алгоритм		
			ε	ε_{min}	t	ε	ε_{min}	t
Ftv33	34	1286	0,00	0,00	7,94	0,00	0,00	10,57
Ftv44	45	1613	1,43	0,00	27,69	1,36	0,00	47,99
Ry48p	48	14422	0,61	0,17	50,83	0,55	0,00	51,41
Eil51	51	426	0,70	0,00	55,31	0,47	0,00	50,1
Berlin52	52	7542	0,58	0,58	38,52	0,58	0,58	40,70
Ftv55	56	1608	3,65	1,74	58,33	3,26	1,37	58,68
Ftv64	65	1839	3,81	1,30	100,87	3,24	0,38	115,42
Sto70	70	675	1,41	0,00	235,94	1,28	0,00	212,33
KroA100	100	21282	1,30	0,79	983,13	0,86	0,00	789,30

Заключение

Предложенный метод, который базируется на мультиагентном подходе и идеях *H*-метода, показал свою высокую эффективность при экспериментальном исследовании. Комбинирование механизма сканирования пространства и процедуры обучения позволило повысить точность решения ряда известных задач коммивояжера и квадратичных задач о назначениях в сравнении с *H*-методом.

Целью дальнейших исследований может стать исследование эффективности алгоритма в зависимости от изменения процедуры обновления значения матрицы «желательности», количества агентов и других параметров *MH*-метода. Представляет интерес получение оценок сходимости и трудоемкости.

Благодарности

Статья частично финансирована из проекта **ITHEA XXI** Института Информационных теории и Приложений FOI ITHEA и Консорциума FOI Bulgaria (www.ithea.org, www.foibg.com).

Список литературы

- [1] Gendreau M., Potvin J-Y. Metaheuristics in Combinatorial Optimization // Annals of Operations Research. – 2005. – 140. – P. 189–213.
- [2] Hoos H.H., Stützle T. Stochastic Local Search: Foundations and Applications. – San Francisco: Morgan Kaufmann Publ., 2005.
- [3] Гуляницкий Л.Ф., Гобов Д.А. Применение *H*-метода для решения задач комбинаторной оптимизации на перестановках // Системные исследования и информационные технологии. – 2007. – №2. – С. 74–86.
- [4] Dorigo M., Stützle T. Ant Colony Optimization. – Cambridge: MIT Press, MA, 2004.
- [5] Chong C.S., Low M.Y.H. A bee colony optimization algorithm to job shop scheduling // Winter Simulation Conference: Proceedings of the 38th conference on Winter Simulation. – Monterey: Monterey Press, 2006. – P. 1954–1961.
- [6] Химмельблау Д. Прикладное нелинейное программирование. – М.: Мир, 1974.
- [7] Сергиенко И.В., Гуляницкий Л.Ф. Фронтальные алгоритмы для многопроцессорных ЭВМ // Кибернетика. – 1981. – №6. – С.1–4.
- [8] Larrañaga P., Lozano J. A. (Eds.) Estimation of distribution algorithms: A new tool for evolutionary computation. – Kluwer Academic Publishers, 2001.
- [9] QAPLIB // <http://www.opt.math.tu-graz.ac.at/qaplib/>
- [10] <http://www.akira.ruc.dk/~keld/research/LKH/LKH-1.2/TSPLIB/>
- [11] Hansen P., Mladenovic N. Variable neighborhood search: Principles and applications // European Journal of Operational Research. – 2001. – №3. – P. 449–467.

Информация об авторах

Леонид Гуляницкий (*Hulianytskyi*) – д.т.н., ведущий научный сотрудник Института кибернетики им. В.М. Глушкова НАН Украины, пр. Глушкова, 40, Киев, 03680, Украина. e-mail: lh_dar@hotmail.com

Денис Гобов – аспирант, Институт кибернетики им. В.М.Глушкова НАН Украины, пр. Глушкова, 40, Киев, 03680, Украина. e-mail: gda1980@ukr.net

РАЗРАБОТКА И ИССЛЕДОВАНИЕ АЛГОРИТМОВ РЕШЕНИЯ ЗАДАЧИ ПРОГНОЗИРОВАНИЯ ТРЕТИЧНОЙ СТРУКТУРЫ ПРОТЕИНА

Леонид Гуляницкий, Виталина Рудык

Аннотация: Рассматривается оптимизационная задача вычислительной биологии – прогнозирование структуры протеина по последовательности аминокислотных остатков. Описывается HP-модель, которая, не смотря на свою простоту, достаточно точно отображает процессы сворачивания в реальных молекулах. Предложены локальные и глобальные алгоритмы решения. Сравнительные характеристики методов и их комбинаций проиллюстрированы результатами вычислительного эксперимента по решению задач разной размерности.

Ключевые слова: комбинаторная оптимизация, прогнозирование третичной структуры протеина, метод ветвей и границ, стохастический локальный поиск, алгоритмы муравьиных колоний.

Conference: The paper is selected from XVth International Conference "Knowledge-Dialogue-Solution" KDS 2009, Varna, Bulgaria, June-July 2009

Введение

Задача прогнозирования третичной структуры протеина по последовательности аминокислотных остатков – одна из самых важных и сложных задач вычислительной биологии. Поскольку процессы сворачивания протеинов изучены не в полной мере, исследователями предложен ряд упрощенных моделей, которые базируются на физических свойствах молекул и приводят к задачам комбинаторной оптимизации. Но и предложенные точные алгоритмы не решают задачи за удовлетворительное время даже для последовательностей длины 50, что дает основания для разработки приближенных алгоритмов. В настоящее время для решения возникающих задач предложен ряд таких алгоритмов популяционного типа [1, 2, 3], а также алгоритмов с оценкой точности [4, 5]. Необходимость в получении более точных решений побуждает к разработке новых детерминированных и метаэвристических алгоритмов, которые и являются предметом дальнейшего исследования.

Постановка задачи

Задача прогнозирования третичной структуры протеина состоит в определении его формы в пространстве, исходя из последовательности аминокислотных остатков в его цепи. В большинстве известных моделей пространство представляется дискретной решеткой. Соседние в последовательности остатки должны располагаться в соседних узлах решетки. Тогда структура протеина задается самонепересекающимся путем в решетке – сверткой. Задача прогнозирования третичной структуры состоит в поиске свертки с минимальной энергией, которая подсчитывается согласно определенному правилу. Таким образом, для описания задачи необходимо выбрать тип решетки и функцию энергии.

В данной работе рассматривается гидрофобно-полярная модель (HP-модель) Дилла [6]. В HP-модели 20 аминокислотных остатков делятся на две группы – гидрофобные (неполярные) и гидрофильные (полярные). Тогда входную последовательность можно рассматривать как слово из алфавита $\{H, P\}$, где P обозначает полярные, а H – гидрофобные остатки. В заданной структуре контактирующие остатки определяются как остатки, которые не являются соседними в последовательности, но расположены в соседних узлах решетки. Между контактирующими гидрофобными остатками возникают связи. Энергия

каждой связи принимается за -1 , а энергия структуры – это сумма энергий всех связей в ней. С биологической точки зрения это объясняется тем, что при сворачивании протеина гидрофобные остатки для предотвращения контакта с водой перемещаются внутрь молекулы, в то время как гидрофильные остаются извне. Согласно гипотезе, протеин принимает ту структуру, которая минимизирует площадь воздействия гидрофобных остатков с водой или полярными остатками. Несмотря на простоту NP-модели, процессы свертывания в ней имеют сходство с таковыми в реальных протеинах.

Будем рассматривать двумерную модель протеинов. Для нее изначально Диллом была выбрана квадратная решетка. Но одним из ее свойств является то, что два остатка могут быть соседними в решетке только тогда, когда количество символов между ними будет четным. Так, строка $(HP)^n$ в квадратной решетке не будет иметь ни одной связи, что не соответствует логическим представлениям. Кроме того, такое строгое ограничение усложняет анализ приближенных алгоритмов, определяя искусственно низкую нижнюю границу энергии, вследствие чего оценка точности алгоритма не будет иметь существенного значения для реальных задач. Поэтому в работе была выбрана треугольная решетка, в которой таких проблем не возникает, так как для любых двух позиций в строке существует свертка, в которой соответствующие остатки будут соседями по решетке.

Доказано, что задача прогнозирования третичной структуры протеина в NP-модели является NP-трудной [7, 8].

Детерминированный локальный поиск

Прежде всего, рассмотрим алгоритм детерминированного локального поиска для минимизации энергии протеина в NP-модели. В общем, он не дает удовлетворительных результатов, поскольку функция энергии имеет много локальных минимумов, но может быть использован в более сложных алгоритмах.

Пусть n – длина аминокислотной последовательности протеина. Каждую свертку в треугольной решетке можно закодировать в виде строки длины $n-2$ из алфавита $T = \{F, Fl, Fr, Bl, Br\}$, элементы которого обозначают поворот на каждом шагу – Front, Front-left, Front-right, Back-left, Back-right.

Назовем свертку допустимой, если она задает путь без самопересечений на решетке. Обозначим множество допустимых сверток D . Определим расстояние между свертками $\rho(s, v)$ как количество символов, на которое отличаются их кодировки s и v . Обозначим $O_\delta(s)$ окрестность точки s размера δ , т.е. $O_\delta(s) = \{v \mid \rho(s, v) \leq \delta\}$, а $E(s)$ – энергию свертки s . В приведенных обозначениях схему детерминированного локального поиска показано на рис. 1 [9].

```

procedure LocalSearch ( $s_0$ )
    while (окрестность  $O_1(s_0)$  не просмотрена полностью) do
         $v :=$  некоторый вариант из  $O_1(s_0)$ ;
        if  $v \in D$  and  $E(v) < E(s_0)$  then
             $s_0 := v$ ;
    end while;
    return  $s_0$ ;
end procedure.

```

Рисунок 1 Схема алгоритма детерминированного локального поиска

Таким образом, алгоритм гарантирует улучшение текущей свертки в окрестности, если оно существует. Сложности возникают в том, что условие допустимости является достаточно строгим, поэтому в окрестности $O_1(s)$ может оказаться малое количество допустимых вариантов решения (сверток) или не оказаться вообще. Поэтому предлагается в процедуре LocalSearch заменить окрестность $O_1(s)$ окрестностью большего размера $O_2(s)$, что позволяет улучшать полученные результаты, но, с другой стороны, увеличивает время выполнения. Поэтому для повышения эффективности предлагается использовать стохастические методы локального поиска.

Алгоритм имитационного отжига

Алгоритм имитационного отжига – один из широко используемых алгоритмов локальной оптимизации [10]. Его схема похожа на детерминированный локальный поиск, но за счет стохастической составляющей алгоритм может выходить из локальных минимумов, возвращая более точное значение. Основная идея состоит в том, что переход к лучшим значениям осуществляется всегда, а к "худшим" (в смысле значений целевой функции) – с определенной вероятностью, зависящей от значения целевой функции и номера итерации. Его схема показана на рис.2.

```

procedure SA ( $S_0$ )
     $S_{rec} := S_0$ ;  $T :=$  начальное значение;
    while окрестность  $O_1(s_0)$  не просмотрена полностью do
        while не достигнуто равновесие do
             $v :=$  некоторый вариант из  $O_1(s_0)$ ;
            if ( $v \in D$ ) then
                 $\Delta := E(v) - E(s_0)$ ;
                 $p := \min \{1, \frac{\Delta}{T}\}$ ;
                 $\xi := \text{random}[0, 1]$ ;
                if ( $p \geq \xi$ ) then
                     $s_0 = v$ ;
                    if ( $E(v) < E(s_{rec})$ ) then
                         $S_{rec} := v$ ;
                    end if;
                end if;
            end while;
             $T :=$  следующее значение;
        end while;
    return  $S_{rec}$ ;
end procedure.

```

Рисунок 2 Схема алгоритма имитационного отжига

Для задачи прогнозирования структуры протеина алгоритм имитационного отжига был предложен и проанализирован в работе [11], где можно найти рекомендации по выбору параметров алгоритма и особенностям реализации.

G-алгоритм

Эффективным алгоритмом комбинаторной оптимизации также является G-алгоритм [12]. В отличие от алгоритма имитационного отжига, вероятность перехода к худшему значению на каждом шаге не зависит от номера итерации, но изменяется пороговое значение, которое определяет условия отсева ухудшающих вариантов (рис. 3).

```

procedure G-algorithm ( $S_0$ )
   $S_{rec} := S_0$ ;  $\mu_0 := 0$ ;  $h := 0$ ;  $t := 0$ ;
  while (окрестность  $O_1(S_0)$  не просмотрена полностью) do
    while (не достигнуто равновесие) do
       $v :=$  некоторый вариант из  $O_1(S_0)$ ;
      if ( $v \in D$ ) then
         $\tilde{\varphi} := \varphi(S_0, v)$ ;
         $p := \min\{1, \max\{0, \tilde{\varphi}\}\}$ ;
         $\xi := \mu_t + \text{random}[0,1] \cdot (1 - \mu_t)$ ;
        if ( $p \geq \xi$ ) then
           $s_0 := v$ ;  $h := h + 1$ ;
          if ( $E(v) < E(S_{rec})$ ) then
             $S_{rec} := v$ ;
          end if;
        end if;
      end while;
       $\mu_{t+1} := G(\mu_t)$ ;  $t := t + 1$ ;
    end while;
  return  $S_{rec}$ ;
end procedure.

```

Рисунок 3 Схема G-алгоритма

В качестве функции $\varphi(x, y)$ выберем функцию $\varphi(x, y) = 1 - \frac{E(y) - E(x)}{\gamma \cdot (E(x) - E_{\min})}$,

где E_{\min} – нижняя граница энергии протеина. Для ее подсчета используется свойство, что каждый гидрофобный аминокислотный остаток, кроме первого и последнего, в последовательности в треугольной решетке может образовывать максимум четыре связи, а первый и последний остатки, если они гидрофобные -- по пять связей. Тогда, если обозначить n_H количество гидрофобных остатков в последовательности, E_{\min} можно посчитать по формуле:

$$E_{\min} = -\frac{4n_H + 2}{2} = -(2n_H + 1).$$

В качестве функции G выбрано $G(x) = (\eta_k(x) + H)^k$,

где $0 < H < 1$, $k \in \{1, 2, 3\}$ – параметры алгоритма,

$$\eta_k(x) = \begin{cases} 0, & x = 0, \\ x^{1/k}, & x > 0. \end{cases}$$

Алгоритм имитационного отжига и G-алгоритм в среднем дают более заметное улучшение допустимого решения задачи, чем детерминированный локальный поиск, при этом время расчетов увеличивается несущественно, что дает основания для их дальнейшего использования.

Алгоритм свертывания стрелкой

В [13] предложен алгоритм для решения поставленной задачи с априорной оценкой точности. Гарантированная точность – $\frac{1}{2}$ от оптимального решения. Алгоритм является детерминированным, конструктивным. Сворачивание производится таким образом, чтобы все гидрофобные остатки находились на одном из трех уровней – так называемом костяке свертки (см. рис. 4)

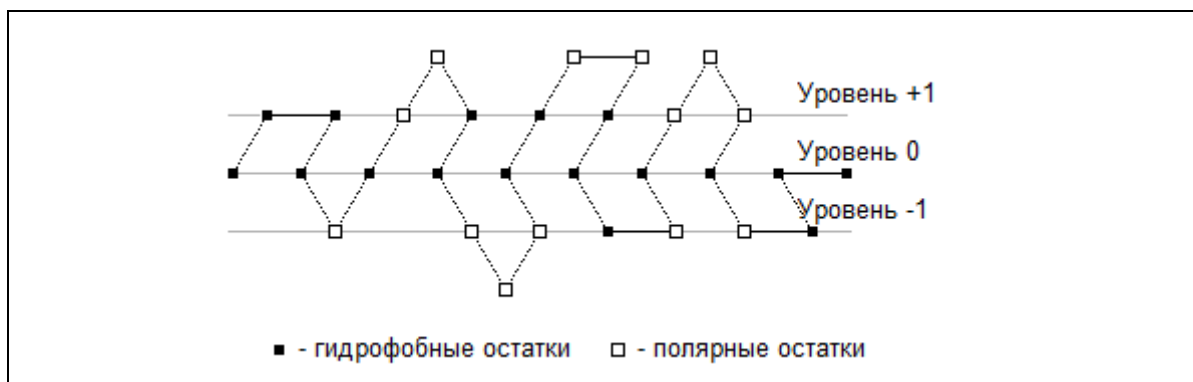


Рисунок 4 Сворачивание стрелкой для последовательности RHPRHPRRHPRRHNNRHHNPPRHNNRH

Алгоритм более эффективен для последовательностей с высоким соотношением числа гидрофобных к числу полярных остатков. Важное преимущество – линейные затраты по времени, что, во-первых, важно для практических задач больших размерностей, и, во-вторых, позволяет использовать его в других более сложных алгоритмах. Недостатком алгоритма является невозможность достигнуть желаемой точности.

Метод ветвей и границ

Для поиска точного решения поставленной задачи предлагается метод ветвей и границ, идея которого состоит в использовании двух процедур – ветвления и поиска границ. Ветвление – деление множества допустимых решений на подмножества меньших размеров, а поиск границ – поиск верхней и нижней оценки значения целевой функции на определенном множестве. При этом для задачи на минимум при условии, что нижняя оценка для подмножества допустимых решений A больше верхней оценки для подмножества B , то A не содержит в себе оптимального решения и дальше может не рассматриваться. Ветвление предлагается делать последовательно по каждой координате на пять областей – по числу возможных направлений.

Для нижней границы предлагается следующая формула:

$$m = -(b_{ee} + b_{en} + b_{nn} + 1),$$

$$b_{en} = 4 * \min \{p_f, p_n\}, \quad b_{nn} = 2 * (p_n - \min \{p_f, p_n\}),$$

где b_{ee} – количество связей между фиксированными остатками, b_{en} и b_{nn} – максимальное количество связей, которая может быть образована свободными остатками, которые свяжутся с фиксированными и свободными соответственно, p_n – количество свободных гидрофобных остатков, p_f – количество мест в решетке, в которых фиксированные остатки могут образовывать связи.

Для получения верхней оценки используется описанный выше алгоритм свертывания стрелкой. Свободные остатки располагаются соответственно алгоритму, а в том случае, когда такое положение не является допустимым – равновероятно в любом другом положении. Предусмотрено генерацию допустимого решения в том случае, если оно существует в данном подмножестве. В качестве верхней оценкой взято значение целевой функции на построенном решении.

Как и любой точный метод, предложенный алгоритм требует больших затрат по времени, что делает его непрактичным для реальных задач.

Алгоритм оптимизации муравьиными колониями

Одним из современных стохастических методов комбинаторной оптимизации является алгоритм оптимизации муравьиными колониями [14]. Его общая схема показана на рис. 5, где k – параметр алгоритма.

```

procedure ACO()
   $s_{rec} := null$ ;
  ИнициализироватьФеромоннуюМатрицу(M);
  while (не выполняется условия завершения) do
    for  $i = 1, \dots, k$  do
       $v_i := \text{ДопустимоеРешение}(M)$ ;
       $v_i := \text{ЛокальныйПоиск}(v_i)$ ;
      if ( $E(v_i) < E(s_{rec})$ ) then
         $s_{rec} := v_i$ ;
    end for;
    for  $i = 1, \dots, k$  do
      ОбновитьФеромоннуюМатрицу(M,  $v_i$ );
    end while;
  return  $s_{rec}$ ;
end procedure.

```

Рисунок 5 Схема алгоритма оптимизации муравьиными колониями

Феромонная матрица в рассматриваемой задаче имеет размерность $5 \times (n - 2)$. На каждой итерации вероятность совершить i -й поворот после $(j + 1)$ -го аминокислотного остатка в структуре пропорционален элементу матрицы с номером $[i, j]$ и оценке позиции $e(pos)$, которая считается по следующему правилу: $e(pos) = \sum_{p \in O(pos)} e_{f(pos)f(p)}$, где pos – узел решетки, на котором может находиться следующий остаток, $O(pos)$ – соседние узлы, $f(\cdot) \in \{H, P, 0\}$ – функция, которая возвращает H , если в узле находится гидрофобный остаток, P , если полярный и 0 в случае, если узел не занят. Числа $e_{HH}, e_{HP}, e_{H0}, e_{PH}, e_{PP}, e_{P0} \geq 0$ – параметры алгоритма, должны удовлетворять условиям:

$$e_{HP} \leq e_{H0} \leq e_{HH}, \quad e_{PH} \leq e_{P0} \leq e_{PP}.$$

Обновление феромонной матрицы происходит отдельно по частям последовательности, элементы которой образуют связи только внутри этой части. В качестве процедуры локального поиска может быть выбран один из трех вышеописанных алгоритмов.

Вычислительный эксперимент

Для сравнения показателей эффективности алгоритмов был проведен вычислительный эксперимент, результаты которого приведены в табл. 1 и 2.

В табл. 1 отражены результаты применения алгоритмов локального поиска – детерминированного (ДЛП), детерминированного с окрестностью радиуса 2 (ДЛП-2), имитационного отжига (АИО) и G-алгоритма. Для тестирования была случайным образом сгенерирована последовательность задач размерности от 15 до 300 с соотношением гидрофобных остатков к полярным 1:4, 1:2, 1:1, 2:1, 4:1 для каждой размерности – всего 40 задач. Все алгоритмы в качестве начального приближения получали одну и ту же свертку в каждой задаче. В этой табл. приведены усредненные результаты, сгруппированные по длине входящей последовательности – n . Здесь $\overline{f_0}$ – усредненная энергия свертки, которая поступает на вход алгоритма, \overline{t} – усредненное время выполнения в миллисекундах, $\overline{f^*}$ – усредненное значение оптимальной энергии, выданное алгоритмом, \overline{q} – усредненное значение улучшения $q = \frac{f_0 - f^*}{|f_0|} \cdot 100\%$. В контексте таких подсчетов показатель \overline{q} является более корректным, чем $\overline{f^*}$, так как чем больше в молекуле соотношение количества гидрофобных остатков к количеству полярных, тем меньше нижняя граница энергии, следственно величина $\overline{f^*}$, в отличие от \overline{q} , неравномерно взвешена по задачам. Стохастические алгоритмы для каждой задачи запускались по три раза, поэтому для них выводится также показатель $\overline{f_{\min}^*}$ – усредненное значение лучшего результата для каждой задачи.

Табл. 1

n	$\overline{f_0}$	ДЛП			ДЛП-2			АИО				G			
		$\overline{t}, \text{мс}$	$\overline{f^*}$	$\overline{q}, \%$	$\overline{t}, \text{мс}$	$\overline{f^*}$	$\overline{q}, \%$	$\overline{t}, \text{мс}$	$\overline{f^*}$	$\overline{f_{\min}^*}$	$\overline{q}, \%$	$\overline{t}, \text{мс}$	$\overline{f^*}$	$\overline{f_{\min}^*}$	$\overline{q}, \%$
15	-1,2	3	-4,2	260	16	-6,0	410	5	-5,4	-6,4	360	10	-5,4	-6,4	360
50	-10,6	34	-23,0	177	220	-29,2	328	46	-23,8	-26,2	180	61	-21,9	-28,4	175
75	-14,2	70	-28,4	156	518	-35,2	289	89	-28,8	-29,8	197	127	-26,3	-30,6	181
100	-24,8	141	-41,0	137	750	-49,6	236	191	-44,1	-46,6	177	270	-45,3	-48,8	201
150	-34,6	383	-66,2	82	2208	-71,4	125	457	-63,9	-67,6	83	546	-67,0	-69,0	89
200	-39,0	838	-79,0	140	4343	-97,2	210	790	-76,0	-81,6	134	937	-74,3	-83,0	115
250	-51,2	941	-91,8	106	5481	-115,6	156	1240	-92,6	-96,6	110	1373	-93,9	-95,0	113
300	-38,0	589	-81,6	38	3397	-84,9	41	574	-72,5	-128,6	30	872	-71,8	-133,0	29

В табл. 2 приведены результаты, полученные алгоритмами ветвей и границ (АВГ), свертывания стрелкой (АСС), свертывания стрелкой + G-алгоритм (С+G), оптимизации муравьиными колониями без локального поиска (ОМК) и с детерминированным локальным поиском (ОМК+ДЛП). Поскольку в отличие от алгоритмов локального поиска, в приведенных алгоритмах нет начального приближения, для подсчета относительного улучшения, задачи не группировались по размерности. В табл. 2 n обозначает размер последовательности, $h : p$ – соотношение количества гидрофобных остатков к количеству полярных, t – время работы алгоритма в секундах, f^* – полученное оптимальное значение. Как и в случае стохастических локальных алгоритмов, алгоритмы ОМК и ОМК+ ДЛП выполнялись по три раза для каждой

задачи, f_{\min}^* – наименьшее из полученных значений. Критерий остановки: оптимальное значение не изменялось на протяжении количества итераций, которое равняется длине входящей последовательности. Алгоритм ветвей и границ выполнялся с ограничениями по времени, в качестве которого выбиралась величина в пять раз большая, чем время, использованное алгоритмами ОМК и ОМК+ ДЛП.

Табл.2

n	$h:p$	АВГ		АСС		АСС+G		ОМК			ОМК+ ДЛП		
		t, c	f^*	t, c	f^*	t, c	f^*	t, c	\bar{f}^*	f_{\min}^*	t, c	\bar{f}^*	f_{\min}^*
15	4:1	10	-12	<1	-11	<1	-12	<1	-11,7	-12	<1	-12,0	-12
15	2:1	10	-12	<1	-9	<1	-9	<1	-11,0	-12	<1	-11,7	-12
15	1:1	10	-6	<1	-5	<1	-5	<1	-5,0	-6	<1	-6,0	-6
15	1:2	10	-6	<1	-4	<1	-4	<1	-5,3	-6	<1	-4,7	-5
15	1:4	10	-5	<1	-3	<1	-4	<1	-4,7	-5	<1	-4,3	-5
50	4:1	25	-52	<1	-44	<1	-49	2,0	-44,0	-46	5,0	-51,0	-52
50	2:1	12	-51	<1	-44	<1	-48	1,3	-48,7	-52	2,3	-52,7	-54
50	1:1	22	-34	<1	-29	<1	-31	1,3	-37,7	-41	4,3	-37,0	-38
50	1:2	15	-12	<1	-12	<1	-12	1,0	-8,3	-9	3,0	-8,3	-9
50	1:4	15	-17	<1	-15	<1	-16	1,0	-17,7	-19	3,0	-16,7	-18
100	4:1	660	-113	<1	-99	<1	-99	11,3	-98,7	-104	131,7	-112,0	-114
100	2:1	200	-97	<1	-78	1	-87	10,3	-82,7	-87	40,0	-90,0	-92
100	1:1	257	-57	<1	-52	<1	-53	8,3	-52,7	-55	51,3	-55,0	-56
100	1:2	127	-34	<1	-33	1	-34	10,3	-34,0	-35	25,3	-34,3	-35
100	1:4	65	-24	<1	-22	1	-24	2,7	-20,0	-23	13,0	-22,0	-24
300	4:1	2797	-350	<1	-294	6	-324	189,3	-300,7	-326	559,3	-342,3	-359
300	2:1	5259	-280	<1	-249	6	-261	185,3	-244,0	-260	1051,7	-272,7	-279
300	1:1	1413	-167	<1	-157	4	-164	146,0	-137,0	-151	282,7	-150,3	-154
300	1:2	1677	-98	<1	-95	7	-96	177,3	-72,3	-80	335,3	-79,3	-89
300	1:4	2035	-66	<1	-66	3	-66	160,3	-51,3	-55	407,0	-59,7	-64

Вычислительный эксперимент показал существенную зависимость показателей эффективности алгоритмов от их параметров, чем обуславливается необходимость разработки методологии их настройки в зависимости от типа решаемых задач.

Заключение

Практически каждый из "базовых" алгоритмов имеет свои преимущества при решении поставленных задач в зависимости от их характеристик. Возможность создания на их основе метаэвристических алгоритмов позволяет достичь лучших показателей эффективности.

Целью дальнейших исследований может быть: теоретическое исследование сходимости некоторых из вышеописанных алгоритмов; формализованный выбор значений параметров алгоритмов в зависимости от длины входящей последовательности и, возможно, соотношения гидрофобных и полярных остатков в ней для повышения эффективности; применение более агрегированных алгоритмов для решения поставленной задачи; обобщение полученных результатов на трехмерную решетку.

Благодарности

Статья частично финансированна из проекта **ITHEA XXI** Института Информационных теории и Приложений FOI ITHEA и Консорциума FOI Bulgaria (www.ithea.org, www.foibg.com).

Список литературы

- [1] R.Unger, J.Moult. Genetic Algorithms for protein folding simulations // J. of Molecular Biology.–1993. – 231(1). –P.75–81.
- [2] A.Piccolboni, G.Mauri. Application of Evolutionary Algorithms to Protein Folding Prediction // Artificial Evolution. – 1997. – P. 123–136.
- [3] M.Khimasia, P.Coveney. Protein Structure Prediction as a Hard Optimization Problem: the Genetic Algorithm Approach // Molecular Simulation. – 1997. – 19. – P. 205–226.
- [4] W.Hart, S.Istrail. Fast protein folding in the hydrophobic-hydrophilic model within three-eighths of optimal // J. of Computational Biology. – 1996 –3(1). –P. 53–96.
- [5] R.Agarwala, S.Batzoglou, V.Dancik, SE.Decatur, M.Farach, S.Hannenhalli, S.Muthukrishnan, S.Skienna. Local rules for protein folding on a triangular lattice and generalized hydrophobicity in the HP model // J. of Computational Biology. – 1997. – 4, №3 . – P. 75–96.
- [6] K.Dill, S.Bromberg, K.Yue, K.M.Fiebig, D.Yee, P.Thomas, H.Chan. Principles of protein folding - a perspective from simple exact models // Protein Science. – 1995. – 4. – P. 561– 602.
- [7] B.Berger, T.Leighton. Protein folding in the hydrophobic-hydrophilic (HP) model is NP-complete // J. of Computational Biology. – 1998. –5(1). – P. 27–40.
- [8] P.Crescenzi, D.Goldman, C.Papadimitriou, A.Piccolboni, M.Yannakakis. On the complexity of protein folding // J. of Computational Biology. –1998. – 5(3) . – P. 423–465.
- [9] И.В. Сергиенко. Математические модели и методы решения задач дискретной оптимизации. – К.: Наукова думка, 1985. – 384 с.
- [10] H.Hoos, T.Stützle. Stochastic Local Search: Foundations and Applications. – San Francisco: Morgan Kaufmann Publ., 2005. – 658 p.
- [11] A. Albrecht, K. Steinhöfel. Run-time Estimates for Protein Folding Simulation in the H-P Model. – 9th Int. Symp. on Artificial Intelligence and Mathematics, Fort Lauderdale, Florida, 2006.
- [12] Гуляницкий Л.Ф. Решение задач комбинаторной оптимизации алгоритмами ускоренного вероятностного моделирования // Компьютерная математика. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 2004. – №1. – С. 64–72.
- [13] R.Agarwala, S.Batzoglou, V.Dancik, S.E.Decatur, M.Farach, S.Hannenhalli, S.Muthukrishnan, S. Skienna. Local rules for protein folding on a triangular lattice and generalized hydrophobicity in the HP model // J. of Computational Biology. – 1997 – 4.– P. 275–296.
- [14] M.Dorigo M.,T. Stützle. Ant Colony Optimization. – Cambridge: MIT Press, MA, 2004. – 348 p.

Об авторах

Леонид Гуляницкий (*Hulianytskyi*) – д.т.н., ведущий научный сотрудник Института кибернетики им. В.М.Глушкова НАН Украины, пр-т Глушкова, 40, Киев, 03680, Украина. e-mail: lh_dar@hotmail.com

Виталина Рудык (*Rudyk*) – студентка, Киевский Национальный Университет им. Тараса Шевченка, ф-т кибернетики, пр-т Глушкова, 2, корп. 6, Киев, 03680, Украина. e-mail: vitalina_@rambler.ru

Business Intelligence

SYSTEMOLOGICAL BUSINESS MODELLING FOR INFORMATION SYSTEMS DEVELOPMENT

**Mikhail Bondarenko, Nikolay Slipchenko, Kateryna Solovyova,
Olexiy Ukrayinets, Larysa Vinogradova**

Abstract: *One of the effective ways of the representation and analysis of the high level-purposes and business requirements to the information program system is modelling of a context of the system functioning – the organization which is automated by software creation and implementation. With this aim the domain specific language in the form of a UML profile for business modelling, where the basis is the systemological approach “Unit – Function – Object”, is defined. The profile allows to create syntactically correct UML models of business and to provide simplicity of the description of the dependences between them and UML models of the program system. Profile implementation, instruments of verification of syntax a business model and its transformation into a software model is fulfilled in a CASE tool for UML modelling.*

Keywords: *Business Modelling, Unified Modelling Language - UML, Systemology, Systemological Approach “Unit – Function – Object” (UFO-approach), UML profile, Requirements, Model Driven Architecture - MDA.*

ACM Classification Keywords: *D.2 Software Engineering, D.2.1 Requirements/Specifications – Languages.*

Conference: *The paper is selected from Seventh International Conference on Information Research and Applications – i.Tech 2009, Varna, Bulgaria, June-July 2009*

Introduction

The research done in software development companies shows that errors in defining requirements to program systems (PS) are a problem which is most often faced during software development, and the cost and complexity of their elimination is the highest and can considerably affect the budget and execution time of all the project [Leffingwell et al, 2000].

The majority of the existing requirements modelling methods do not allow considering high-level business requirements or software customers, who are not its users, but are interested in development and deployment of such a PS to solve business problems. To avoid this problem, it is necessary to start the engineering process with the description of the functional environment of the system, i.e. with business model development [Украинец, 2008 А].

Some modern methods allow to model business in the software engineering process (for example, SOMA [Graham, 2000], Rational UML profile for Business Modelling [Johnston, 2004]), however not always provide strict formal link with the system model or do not guarantee a syntactic correctness of the constructed model. Besides, frequently these methods and technologies frequently have no methodological basis that rises doubts in adequacy of the developed models to the existing business. There is no guarantee that business models reflect

the reality and give correct mechanisms of its analysis, and after all it is the main criterion of any model estimation.

«UML + UFO» – the Key to the Problem Solving

For business processes modelling, it is undoubtedly possible to apply great number of methods. As far as methods of object-oriented visual modelling which are united in unified modelling language UML, are applied virtually in all new projects program systems development, it is natural to try to use similar methods in the field of business processes modelling.

It will provide availability of the method to a wide range of non-experts, and in this case, it is very important as far as a business model should be clear for many people interested in the system and its users. One of the advantages of the given approach is the simplicity of the description of the dependences between business and program system models. It increases productivity of engineering process, reduces the number of system troubles, and also helps to make sure that the developed system solves real business requirements [Украинец, 2007 А].

The usual practice is the application of a method of Use Cases for business modelling, but this approach evokes a number of problems [Graham, 2000]. Use Cases, undoubtedly, are useful to define systems, but do not allow to describe the requirement completely. Use Cases link goals that represent roles of the users in the system with the system itself. On the other hand, requirements can be set up by people and organizations that will never appear close to the system. Ignoring of interaction with non-users can lead to loss of some important technical possibilities. That is, a special object-oriented business modelling method compatible with Use Cases models is necessary.

Thus, undoubtedly, for business modelling it is reasonable to use UML, but for this purpose standard language device is not enough. The creators of UML foresaw such a situation and have provided mechanisms to extend the language which allow defining new domain specific languages (DSL) on the basis of UML, including for business modelling. To provide strictness, entirety, the coordination of the device of modelling of business in a newly created DSL it is suggested to use systemology and systemological approach “Unit – Function – Object” (UFO-approach) [Бондаренко и др., 2004]. This theoretical device has shown its strength when solving different kinds of tasks, including business modelling. It possesses a very important advantage from the point of view of a considered problem – it is coordinated with object-oriented methodology [Украинец, 2007 В].

Functional systemology (systemological approach is the noospheric stage of the science development for complicated systems analysis) unlike other system approaches can be used to solve low-formalized ill-structured problems in the information and other areas. It allows forming informational resources of an organization; consolidating, systematizing, classifying information and knowledge; guarantees the “phase transfer of knowledge into power” [Бондаренко и др., 1998]. Obtaining of new knowledge on the basis of a new systemological method and criteria of natural classification allows to model most objectively deep conceptual knowledge in view of objects essential properties and to develop powerful ontology of ill-structured subject domains [Solovyova, 2000 etc.]. The perspective systemological technology of the organizational system modelling and analysis, material and information processes which proceed there (UFO-approach) allows consulting with challenges of putting in order in the working organization, designing and modelling of the new organization, and also forecasting of its development [Bondarenko et al, 2006, Бондаренко и др., 2004].

Systemological UML metamodel of Business

Definition a new DSL on the basis of UML that is UML profile development always assumes definition of structure of a language metamodel, metamodel unit structure – UML stereotypes and their metaclasses, and also relations between metamodel units. Three interconnected views are suggested to represent business model in language

UML (fig. 1). The business-instances model is the main starting point for business processes modelling (the functional structure of a business system). Business units, which provide performance of system functions, are represented on the subject domain model of business (organizational structure of a business system). The details of business processes in the form of separate business operations and performers responsible for their carrying out are represented on a business activity model [Украинец, 2008 В].

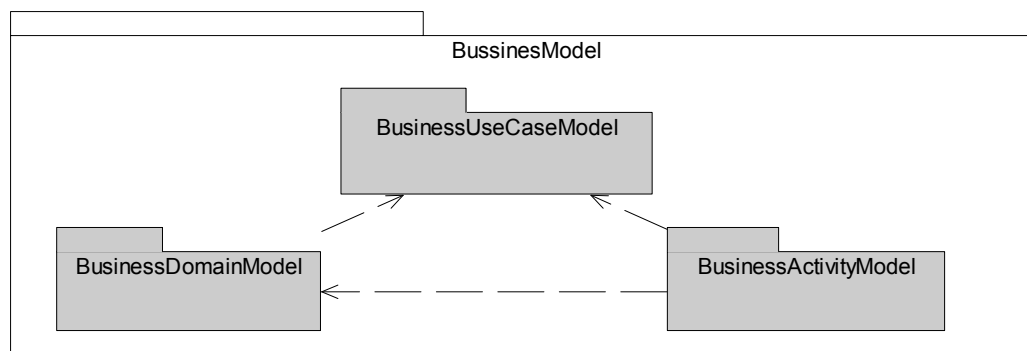


Fig. 1. Business model structure

The structure of each model's units (UML stereotypes) was defined according to the theoretical device of systemology and the UFO-approach to business modelling. To identify UML metaclasses corresponding to the new stereotypes the ontology of the common system modelling of Triune Continuum Paradigm was used [Naumenko, 2002]. Partly the results of this research [Украинец, 2008 В] are presented in tables 1, 2.

Table 1. Elements of Business Use Case Model

UFO Concept		UML Metaclass	UML Stereotype
UFO-Model		Model from Models	BusinessUseCaseModel
UFO-Element	Unit	1) Class from UseCases	1) BusinessSystem
		2) Class from UseCases	2) BusinessActor
	Function	1) UseCase from UseCases	1) BusinessUseCase
		2) Operation from Kernel, Interfaces	2) BusinessRequirement
	Object	1) Class from Kernel	1) BusinessWorker
2) Class from Kernel		2) BusinessCustomer	
Link		AssociationClass from AssociationClasses	BusinessEntity

Table 2. Elements of Business Activity Model

UFO Concept		UML Metaclass	UML Stereotype
UFO-Model		Model from Models	BusinessActivityModel
UFO-Element	Unit	Activity from BasicActivities, CompleteActivities, FundamentalActivities, StructuredActivities	BusinessActivity
	Function	Action from CompleteActivities, FundamentalActivities, StructuredActivities	BusinessAction
	Object	ActivityPartition from IntermediateActivities	BusinessActionWorker
Link		ObjectFlow from BasicActivities, CompleteActivities	BusinessEntityFlow

The relations between the business metamodel elements have been defined according to the theoretical device of systemology and the UFO-approach too. To provide strict correspondence of the UML metamodel and UFO-metamodel the mathematical apparatus of the theory of groups and the theory of graph was used [Украинец, 2008 A]. As a result a conceptual model of a UML profile has been obtained. The fragment of this model is presented on fig. 2.

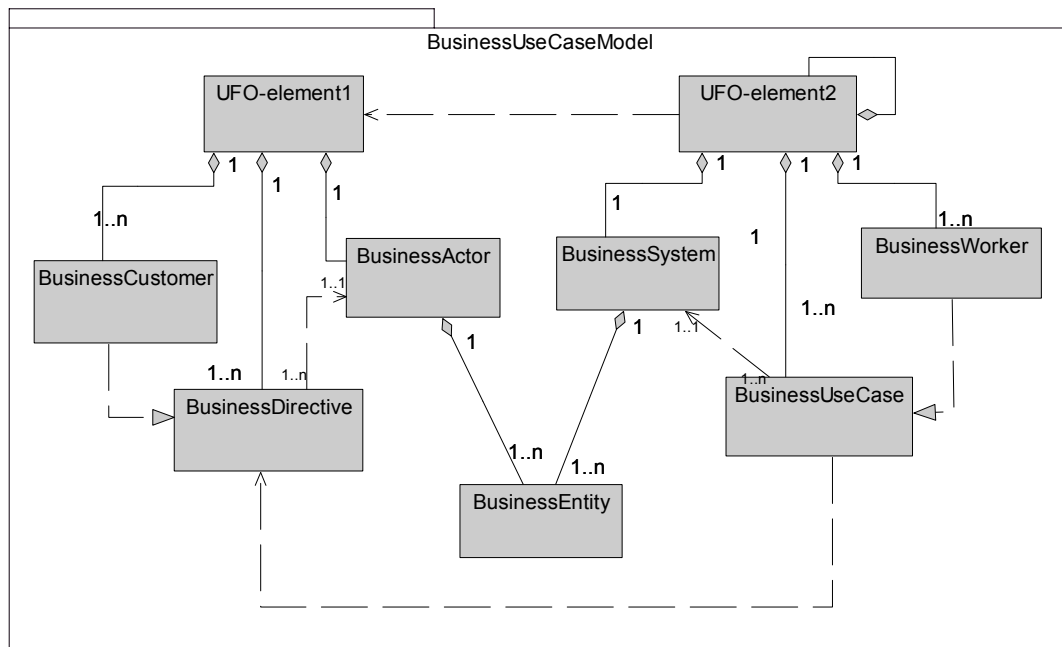


Fig. 2. Conceptual graph-analytic model of Business Use Cases

Going from Business Model to System Model

During business modelling not only it is important to obtain a qualitative model describing the purposes and the requirements of business to the program system, but also to provide usage of the modelling results in the analysis and system designing. For this reason it is necessary to describe formal link of a business model with a program system model. This solution will allow coordinating models and will give the chance to automate transition from a business model to a system model. It raises productivity of engineering process, reduces the system defects number, and also helps to make sure that the developed system solves real business requirements.

This idea is well stated in the MDA (Model Driven Architecture) international standard for which OMG (Object Management Group) is responsible [MDA, 2003]. There are three kinds of models described in MDA, which used during software development. The first one is called CIM (Computational Independent Model) – a business model. The second one is PIM (Platform Independent Model) that is a program system model. Meanwhile in MDA it is supposed that PIM can be partially obtained (generated) on the basis of CIM. The process of conversion of a business model into a system model from the considered issue's point of view according to MDA is shown on fig. 3. It is necessary to mention that nowadays the ideas of transformation of automated models and consideration of the business model as the key model in the software engineering process are widely spread.

For example, IBM Company widely advertises and actively extends such technologies as MDD (Model Driven Development) and BDD (Business Driven Development) [Swithinbank et al, 2004].

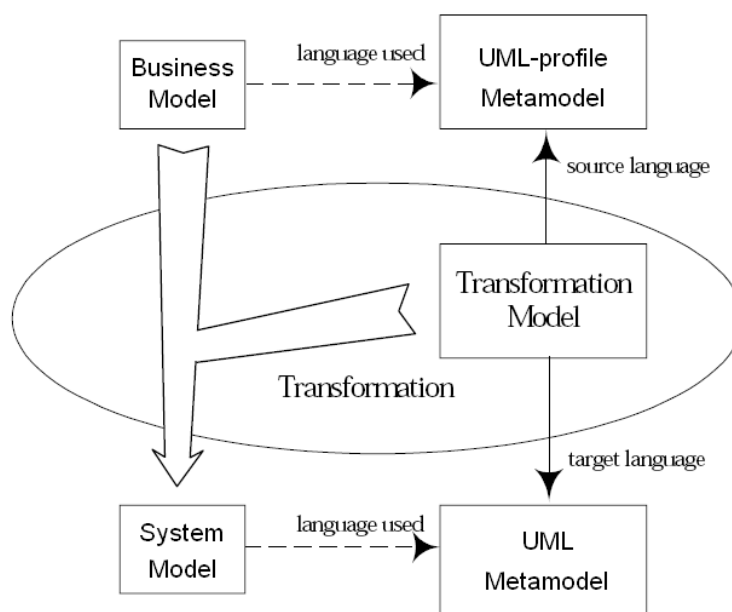


Fig. 3. Transformation of Business Model to System according to MDA

Before the conversion of the systemological business UML model in the system model some stages are supposed:

- creation of the business model
- syntax test of the business model on correctness, entirety and internal coordination that is expressed in the analysis of the admissible and necessary relations between the model elements;
- definition on the business model of those business processes and objects which will be automated at the expense of software development and deployment. They are designated by a stereotypic label "Automated".

The formal semantic link between the systemological business UML model and program system UML model has been defined using the ontology of the common system modelling of Triune Continuum Paradigm [Naumenko, 2002]. Partly the results of this research [Українець, 2008 В] are presented in table 3.

Table 3. Semantic link between a program system model and a business model

System model	Business model
Actor from UseCases	BusinessActor from BusinessUseCase Model
	BusinessWorker from BusinessUseCase Model, BusinessDomain Model
Classifier from UseCases	BusinessSystem from BusinessUseCase Model
UseCase from UseCases	BusinessUseCase from BusinessUseCase Model
	BusinessAction from BusinessDomain Model
Activity from BasicActivities, CompleteActivities, FundamentalActivities, StructuredActivities	BusinessActivity from BusinessActivity Model
Action from CompleteActivities, FundamentalActivities, StructuredActivities	BusinessAction from BusinessActivity Model
ActivityPartition from IntermediateActivities	BusinessActionWorker from BusinessActivity Model
ObjectFlow from BasicActivities, CompleteActivities	BusinessEntityFlow from BusinessActivity Model

Package from Kernel	BusinessSystem from BusinessUseCaseModel
Class from Kernel	BusinessSystem from BusinessUseCase Model
	BusinessWorker from BusinessUseCase Model, BusinessDomain Model
	BusinessEntity from BusinessUseCase Model, BusinessActivity Model, BusinessDomain Model
Operation from Kernel, Interfaces	BusinessAction from BusinessActivity Model
	BusinessActionWorker from BusinessDomain Model
Relationships between classes	Relationships between business model elements

Design and Implementation of the Business Modelling System

The results obtained during the research have been brought to the level of practical implementation in the form of the program unit of the analysis and business modelling in Case tool of object-oriented modelling of IBM Rational Software Architect (RSA) [Swithinbank et al, 2004]. On fig. 4 the main functions of this program system and the users interested in it are presented. The description of the developed system model is given further.

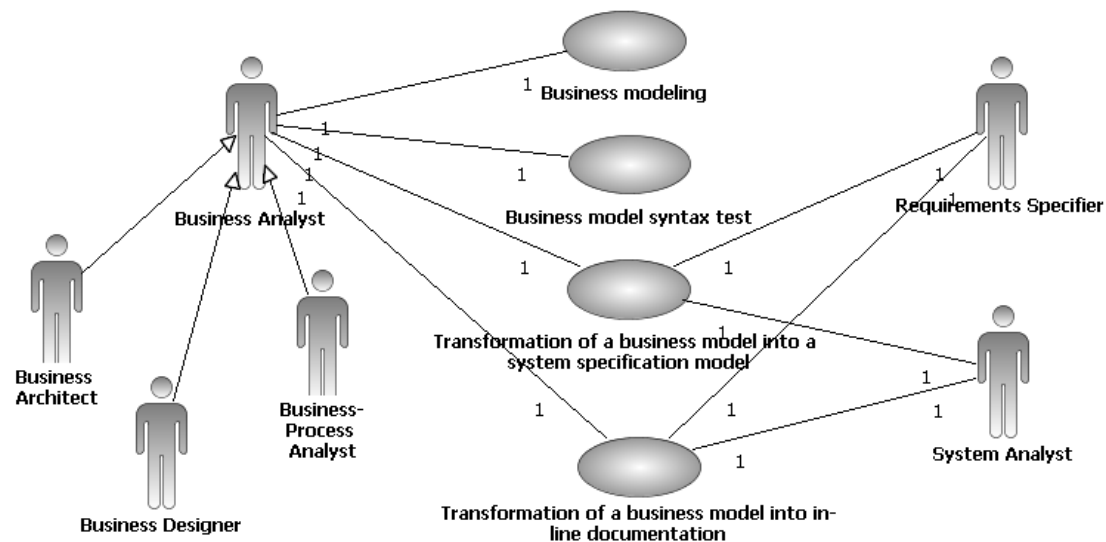


Fig. 4. Use case model of system of systemological UML modelling of business

Business analyst is the generalized role that includes common responsibilities for the business architect, business designer and business processes analyst. *Requirements Specifier* is the role specifies and supports the detailed requirements to the program system. *Systems analyst* is the role supervises and coordinates the process of requirements detection by definition of the fundamentals of the system functionality and boundaries.

Business modelling. This system use case allows a business analyst using a systemological UML profile to build business models which include a business Use Cases model, a business activity model and a business subject domain model. The business analyst can add, update or delete units of business model and the relations between them.

Business model syntax test. This system use case allows a business analyst to carry out verification on the syntax correctness of the constructed business UML model. During model verification the correctness of existing relations between units in a separate model, between units of different interconnected models, and correspondence of the specification of separate units of the model considered in different contexts in a business model is checked. Verification allows revealing the internal inconsistency of the business model, defining its

entirety, revealing duplicating and absent units of the model and the relations between them. As a result of the use case performance the business analyst obtains from system the report on containing errors in the model.

Transformation of a business model into a system specification model. This use case allows a business analyst, requirements qualifier and systems analyst to partially generate UML model of the program system specification from the systemological business UML model. As a result of the use case performance the use case model, activity model and class model of program system is partially generated.

Transformation of a business model into in-line documentation. This use case allows a business analyst, requirements qualifier and systems analyst to partially generate the in-line documentation from a systemological business UML model. As a result of the use case performance the use case model, a document which is the specification of a business system in the form of a structured text is partially generated.

The implementation of the business modelling system functions has been carried out by the various tools RSA presented in table 4.

Table 4. Implementation tools for business modelling system use cases in RSA

Use Case Name	Techniques in RSA
Business modelling	UML Profile Project
Business model syntax test	Constraints in OCL (Object Constraint Language) and built in RSA tools for analysis and constraints check
Transformation of a business model into a system specification model	Java Plug-in
Transformation of a business model into in-line documentation	Java Plug-in

Conclusion

Modern economic conditions, including the world financial crisis, dictate new rules of business behaviour in the area of information systems and technologies (IS / IT). They cause the necessity of strict control over return of investments and optimization of costs for software development and deployment in organizations along with other IS / IT. The developed language of business modelling, based on systemology and UML, is aimed at this problem solving. It allows connecting the purposes and business processes of an organisation with the software functions. Thus, it is possible to guarantee that the investments into software development and implementation are strictly aimed at the solution of business problems, perfection of its business processes and achievement of its competitive advantages.

While developing the language, its metamodel (a set of the model elements and the possible relations between them) and a strict formal connection between the business and system models were determined, the tools of modelling, business model syntax checking, as well as transformation of a business model into a system model and covering documentation were realized. The developed means of the language eliminate the defects of the analogues and allow avoiding errors in software requirements specifying and automating of its development. All this leads to decrease in effort, development and budget time especially in big software development projects.

Acknowledgements

The paper is partially financed by the project **ITHEA XXI** of the Institute of Information Theories and Applications FOI ITHEA and the Consortium FOI Bulgaria (www.ithea.org, www.foibg.com).

Bibliography

[Leffingwell et al, 2000] D. Leffingwell, D. Widrig. Managing Software Requirements. A Unified Approach. AW, 2000.

- [Украинец, 2008 А] Украинец, А.Г. Концептуальная модель UML профиля для моделирования бизнеса [Текст] / А.Г. Украинец // Системы управління, навігації та зв'язку. – К. : ЦНДІ НіУ, 2008. – Вип. 4(8). – С. 120 – 125.
- [Graham, 2000] I. Graham. Object-oriented methods. Principles and Practice. Addison-Wesley Professional; 2000.
- [Johnston, 2004] S. Johnston. Rational UML Profile for business modeling. IBM Rational Technical Library. [Online]. Available: <http://www.ibm.com/developerworks/rational/library/5167.html>
- [Украинец, 2007 А] Украинец, А.Г. Системологический подход к моделированию бизнеса как основа бизнес-ориентированного подхода к информационным технологиям. [Текст] / А.Г. Украинец // Радиоэлектроника и информатика. - 2007. – №4. - С. 142 - 149. - ISSN 1563-0064.
- [Бондаренко и др., 2004] Бондаренко, М.Ф. Моделирование и проектирование бизнес-систем: методы, стандарты, технологии. [Текст]: уч. пособ. / М.Ф. Бондаренко, С.И. Маторин, Е.А. Соловьева; – Х.: «СМИТ» - 2004. - 272с.
- [Украинец, 2007 В] Украинец, А.Г. Разработка модели согласования объектно-ориентированной методологии и системологического подхода к моделированию бизнеса. [Текст] / А.Г. Украинец // Вісник Міжнародного Слов'янського університету. Серія «Технічні науки». - 2007. - Т.Х, №2. - С.29-37 - 500 экз. - ISSN 1563-9961.
- [Бондаренко и др., 1998] Бондаренко М.Ф., Соловьева Е.А., Маторин С.И. Основы системологии. Харьков: 1998. 118с.
- [Solovyova, 2000] E Solovyova. Mathematical and Systemological Foundations of Natural Classification. Automatic Document and Mathematical Linguistics. New York: Allerton Press, Inc., 2000. V. 33, No. 4.
- [Bondarenko et al, 2006] M. Bondarenko, K. Solovyova, S. Matorin and V. Matorin. "Embedding Conceptual Models of Knowledge in Technology of Systemological Analysis." Papers from the IFIP WG8.3 International Conference on Creativity and Innovation in Decision Making and Decision Support, vol. 2, p. 904-928, 2006, London, UK.
- [Украинец, 2008 В] Украинец, А.Г. Структура и стереотипы UML-профиля для моделирования бизнеса [Текст] / А.Г. Украинец // Системи обробки інформації. – 2008. – вип. №6(73). - С. 142 - 149.
- [Naumenko, 2002] A. Naumenko. Triune Continuum Paradigm: a paradigm for General System Modeling and its applications for UML and RM-ODP. Ph.D. thesis 2581, Swiss Federal Institute of Technology - Lausanne, June 2002. [Online]. Available: <http://stella.epfl.ch/tcp/Naumenko-PhD-Thesis.pdf>.
- [MDA, 2003] MDA Guide Version 1.0.1. Object Management Group, Inc. [Online]. Available: <http://www.omg.org/docs/omg/03-06-01.pdf>
- [Swithinbank et al, 2004] P. Swithinbank, M. Chessell, T. Gardner, C. Griffin, J. Man, H. Wylie, L. Yusuf. Patterns: Model-Driven Development Using IBM Rational Software Architect. IBM Corp, 2004. [Online]. Available: <http://www.redbooks.ibm.com/redbooks/SG247105>

Authors' Information

Mikhail Bondarenko - Rector, Kharkiv National University of Radio Electronics (KhNURE), Lenin Ave., 14, Kharkiv, 61166, Ukraine; e-mail: rector@kture.kharkov.ua

Nikolay Slipchenko - Chief of Scientific Department, Kharkiv National University of Radio Electronics, Lenin Ave., 14, Kharkiv, 61166, Ukraine; e-mail: slipchenko@kture.kharkov.ua

Kateryna Solovyova - Chief of Joint Scientific Educational Knowledge Management Center and Social Informatics Department, KhNURE, Lenin Ave., 14, Kharkiv, 61166, Ukraine; e-mail: si@kture.kharkov.ua

Olexiy Ukrainets – Senior lecture of Social Informatics Department, Kharkiv National University of Radio Electronics, Lenin Ave., 14, Kharkiv, 61166, Ukraine; e-mail: aguqr@inbox.ru

Larysa Vinogradova – Student of Social Informatics Department, Kharkiv National University of Radio Electronics, Lenin Ave., 14, Kharkiv, 61166, Ukraine; e-mail: si@kture.kharkov.ua

QUALITY AND PRODUCTIVITY ANALYSIS IN A VERY SMALL ENTERPRISE USING TSPI

Henry Caballero, Jose Calvo-Manzano, Gonzalo Cuevas, Tomás San Feliu

Abstract: *This article shows the benefits of developing a software project using TSPI in a "Very Small Enterprise" based in quality and productivity measures. An adapted process from the current process based on the TSPI was defined and the team was trained in it. The pilot project had schedule and budget constraints. The workaround began by gathering historical data from previous projects in order to get a measurement repository, and then the project metrics were collected. Finally, the process, product and quality improvements were verified.*

Keywords: *TSPI, Very Small Enterprise, Software Quality, Process improvement.*

ACM Classification Keywords: *D.2.9 Software Engineering – Management, Software Process Models*

Conference: *The paper is selected from Seventh International Conference on Information Research and Applications – i.Tech 2009, Varna, Bulgaria, June-July 2009*

Introduction

Large, medium and small enterprises have common problems related to the management and quality [IPRC, 2006] of software projects. This generates cost overruns, low quality and cancelled projects [Standish Group, 2007].

Some process models like the CMMI are not affordable for the small organizations [IPRC, 2006].

Organizations have recognized that the control of their software processes affects the success of their projects, "they know what to do but not how to apply it" [Noopur, 2003]. A new research line based on process improvement in small enterprises is arising in order to facilitate competitive capabilities for this environment in a global market [Glazer, 2006]. These are named Small Settings and include small and medium organizations, and small software projects [Garcia 2006]. (*This paper focuses on small organizations only*).

Team Software Process (TSP) is a framework that provides a customizable process based in an excellent experience in planning and managing software projects [Humphrey, 2006]. It guides teams in managing cost, schedule and quality [Noopur, 2003].

This article shows the quality and productivity results of using an adapted process based on the introduction of the Team Software Process (TSPI) in a very small enterprise. The following goals have been established (see Table 1):

Table 1 – Project goals

Number	Goal
1	To reduce the estimation deviation.
2	To verify the productivity improvement.
3	To verify the process and quality improvement.

The organization decided to use TSPI in order to accomplish the previous goals assuming the risk of modifying its current process. In order to verify the project goals, measures were evaluated comparing the pilot project and Historical average data.

Therefore, the organization decided to apply the basic TSPi principles, getting a customized process as a result of combining TSPi with the organizational process. A basic training was provided for the new process, and historical data were collected in order to facilitate the estimation of the pilot project.

In the following section, the article shows the organization, the development context and the pilot project attributes. Later, the historical data collection will be described, the new process will be analyzed and the advantages pointed out. The project goals will be verified using the project measures and the measurement repository, and finally, the conclusions will be described.

The schema showed in Figure 1 resumes the factors considered in the project.

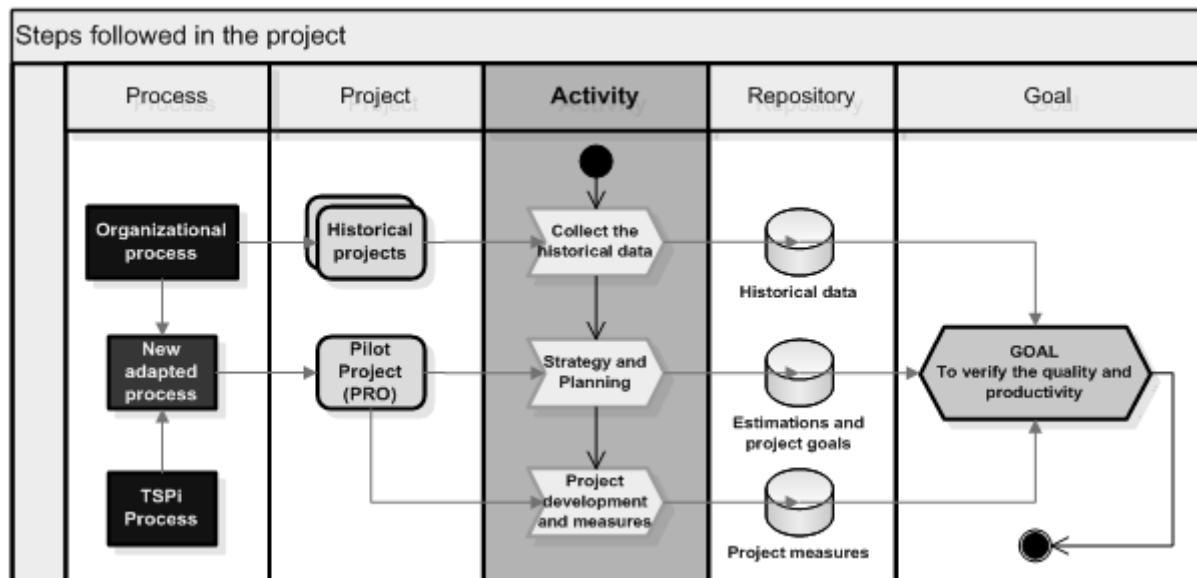


Figure 1 – Factors considered in the project

Context: The organization and the pilot project

UPTA is a Spanish intersectorial organization who takes care of all the scopes of economic activity which they are exerted by self – employment. UPTA leads a lot of projects which generally need a specific software development and in the last year, the number of software projects has increased.

The working scenario has changed to a new environment where many projects were developed simultaneously, and with a greater number of involved people. As a result of an internal assessment, senior management detected that projects were delayed, dedicating additional efforts to accomplish the objectives. Moreover, products quality had decreased.

UPTA was interested in introducing a process model such as CMMI-DEV, but it could not afford for it.

Besides this handicap, UPTA had a project (called PRO) with schedule and budget constraints, and was delayed. The organization selected this project as the pilot.

Collecting the historical data

Data on previous UPTA projects were not enough. There were only schedule and budget data, but in order to verify the project goals, defects and phases efforts data were needed.

Measures related to schedule, size, effort and defects were collected. In addition, some derived metrics were calculated in order to analyze the project results.

In order to support the analysis, historical projects were divided into three phases (see Table 2):

Table 2 – Project phases

Phase	Description
Development	From launch project until unit tests.
Test	It includes integration and system tests.
Operation	From the deployment until three months up.

Effort and defect data of these phases are approximated values because there was no previous data repository.

The new process

The process is a customized process as a result of blending the basic TSPi principles and the previous organizational process. Once the new process was defined, the project started with a training on the new process and the launching meeting.

The TSPi phases were used in the new process in order to get benefits from its procedures and metrics, but the intermediate products, such as requirements or design specifications, were based on the previous organizational process in order to reduce the change impact.

The focus on quality is the main difference with the previous organizational process. Examples of this approach are the quality plan relative to the phases and process performance, inspections and reviews.

The team was empowered to estimate and plan the project balancing the workload, and so, they were more committed. Also, a role definition was adopted.

Table 3 shows the basic TSPi principles applied in the new process and the difference with the previous process.

Table 3 –TSPi principles applied in the new process

New process	Previous process
Process well defined. It makes easier the estimation and monitoring project	Process with inconsistencies. The phases are not well defined
Team motivated, participative and collaborative	Only a project leader elaborates the project plan and the task assignement
Quality focus based in an early defect detection and reduction	Since the schedules are restricted, the quality was not considered
Introduction of inspection activities in the process	Only personal reviews without a quality control
Detailed plan in order to avoid schedule and effort deviation	Projects begin with cost and schedule pre-established and restricted
Monitoring and project visibility with the earned value method	There is no mechanism to monitor the project status
Weekly meeting to analyse the project and to resolve process issues	There are no formal meetings and they are preformed only when there are problems

With respect to project management, weekly meetings and the earned value method were introduced. These gave to the project an actual visibility and an effective monitoring. The schedule, goals, risks, and change requests were evaluated in the weekly meetings.

The weekly meetings and the earned value method allowed the improvement of the project management [Humphrey, 1995] (see Figure 2).

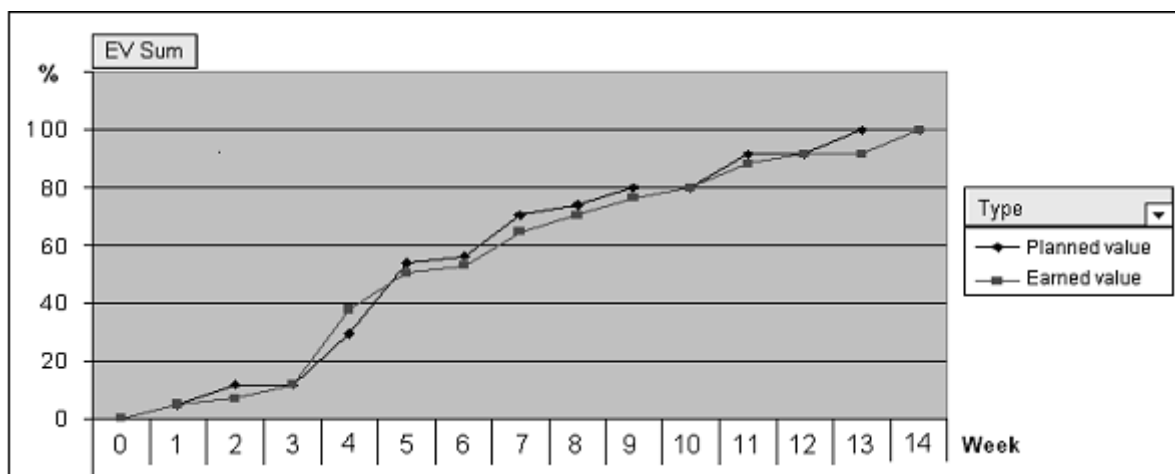


Figure 2 – Earned value monitoring

Verifying the project goals

In order to verify the project goals, measures were evaluated comparing the PRO and the historical average data. The formula applied for the reduction is: $\% \text{ Reduction} = (\text{Historical} - \text{PRO}) / \text{Historical} * 100\%$

Goal 1: To reduce the estimation deviation.

The estimation deviation related to schedule, size and effort were reduced (see Table 4).

Table 4 – Goal 1 results

Goal	Historical	PRO	Reduction
% Schedule deviation	21.4%	7.7%	64.0%
% Effort deviation	55.9%	18.00%	67.8%
% Size deviation	33.7%	22.6%	33.0%

Figure 3 shows graphically the improvement and demonstrates the utility of the new process in the organization.

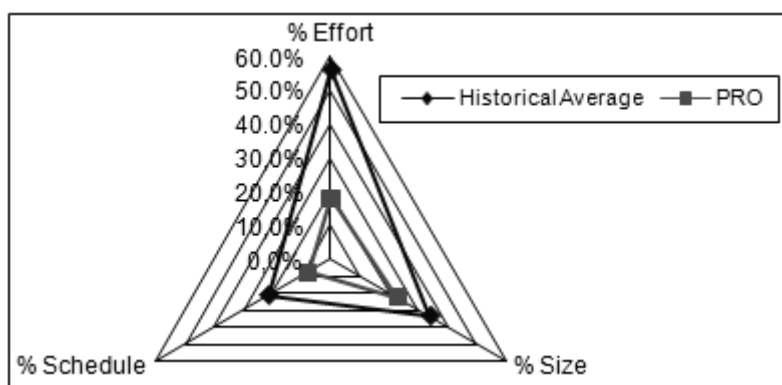


Figure 3 – Estimation deviation

Goal 2: To verify the productivity improvement

The released defect density (ρ), the test productivity and the process productivity were analyzed in order to verify the productivity improvement.

Table 5 – Goal 2 results

Goal	Historical	PRO	Reduction
ρ Released defects [def./KLOC]	2.0	1.2	40.9%
Test Productivity [hour/KLOC]	33.4	13.2	60.5%
Process Productivity [LOC/Hour]	7.3	7.6	- 3.9%

According to Table 5, the test productivity has been increased due to the defect reduction applied before the test phase through formal reviews and inspections. However, the process productivity has not been altered because the test effort was substituted by the quality activities balancing the total effort (process productivity was calculated using the LOC number coded by hour).

The released defects density has improved, but it is far from the TSPi indicators. Figure 4 shows the relationship between ρ Released defects and the product size.

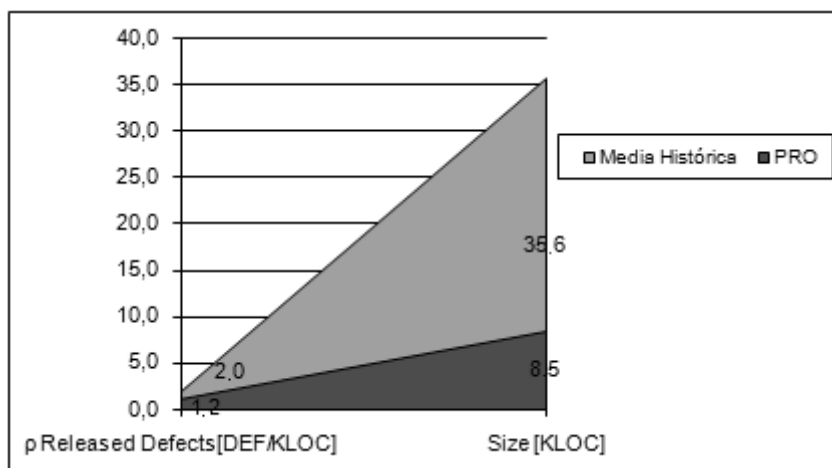


Figure 4 – Released defect density vs. Size

Figure 5 shows graphically the productivity improvement and confirms the benefits of the new adapted process.

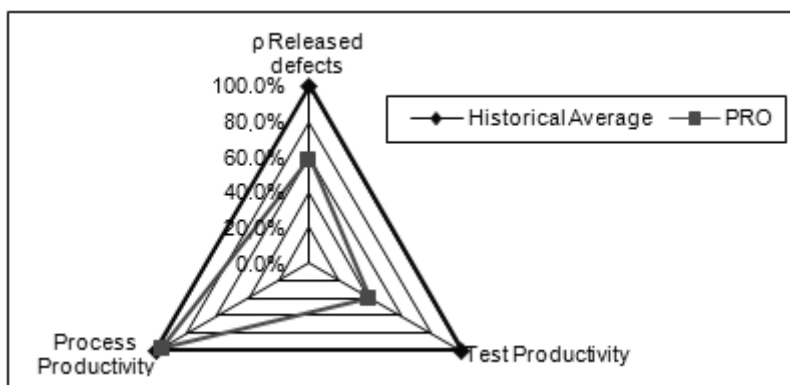


Figure 5 – Productivity improvement

Goal 3: To verify the process and quality improvement

TSPi uses the defect density and process yield indicators in order to verify this goal.

The defect density allows analyzing the process. The process is considered effective when every phase has less or equal defect density than the last one [Humphrey, 1999]. Figure 6 shows that the new adapted process based in TSPi is more effective than the last one because the defect density was reduced in every project phase.

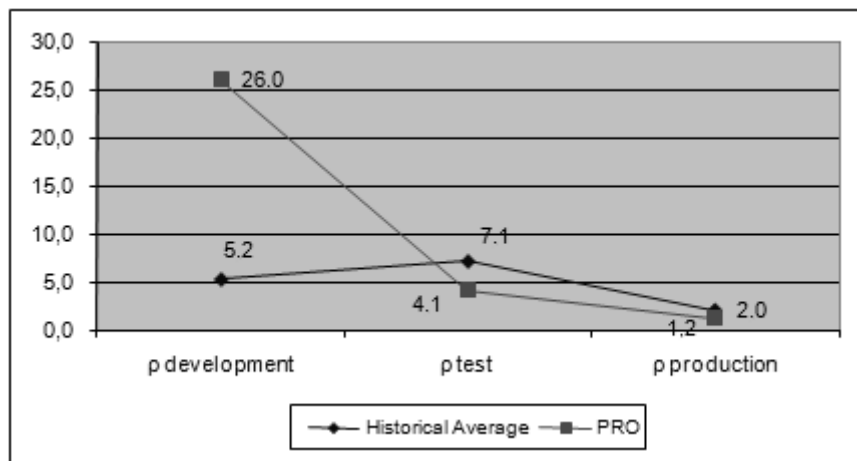


Figure 6 – Process improvement based in defect density

The process yield is a measure that allows analyzing the quality, which measures the percentage of removed defects before a given phase. Process yield goal with TSPi is to remove 99% of all defects before the test. With such yield, process will generally produce very high quality products [Humphrey, 1999].

In the old process, the most of the defects were found during the test phase, however in the newest this was reduced drastically because the early defect reduction was applied. Nevertheless, the value is far from the expected by the TSPi (see Figure 7).

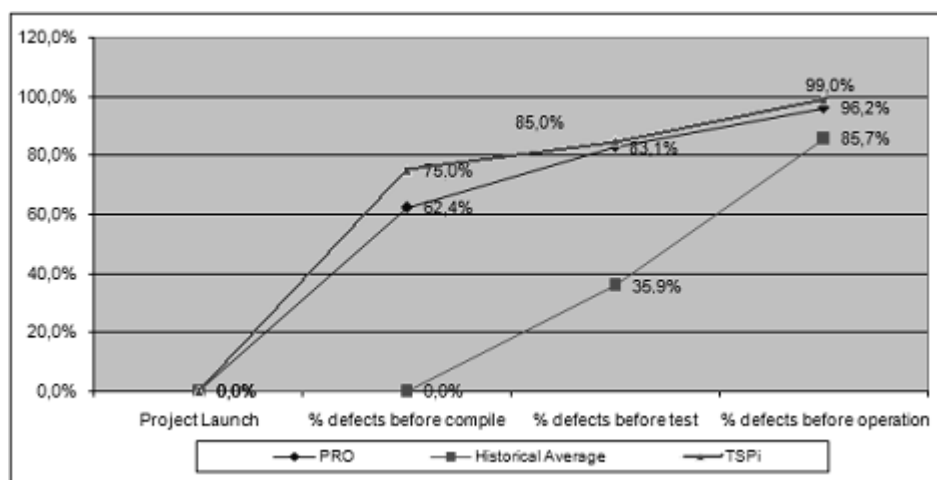


Figure 7 – Quality improvement based in process yield

The Cost of Quality (COQ) is a measure that allows quantifying the size of the quality [Humphrey, 1995].

It has three components, but TSPi only works with two (COQ = Appraisal Costs + Failure Costs):

Appraisal Costs: the cost of evaluating the product to determine its quality level (reviews and inspections).

Failure Costs: the cost of diagnosing a failure, making necessary fixes, and getting back into operation (compilation and test).

Figure 8 shows that the COQ has not changed between the old and the new process (34.7% vs. 35.5%).

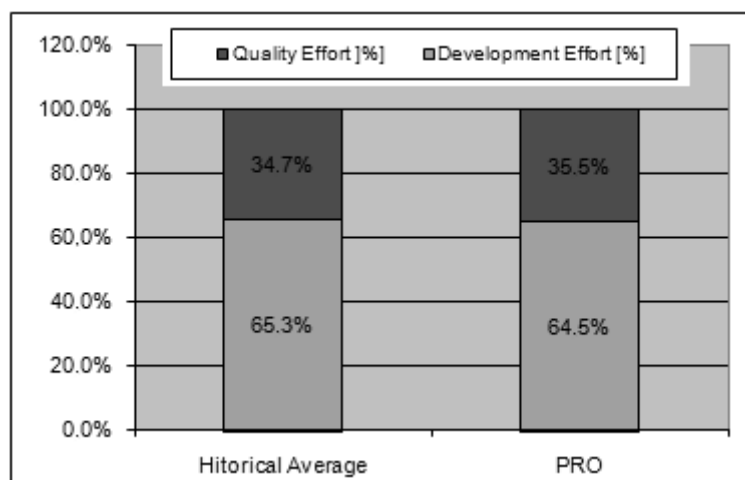


Figure 8 – Cost of Quality

However, if the appraisal costs and failure costs are analyzed in a separated way, the improvement is evident. The new process adopted inspections and formal reviews because it was based on early defect reduction. These new activities allowed increasing the appraisal costs and reducing the failure costs (see Figure 9).

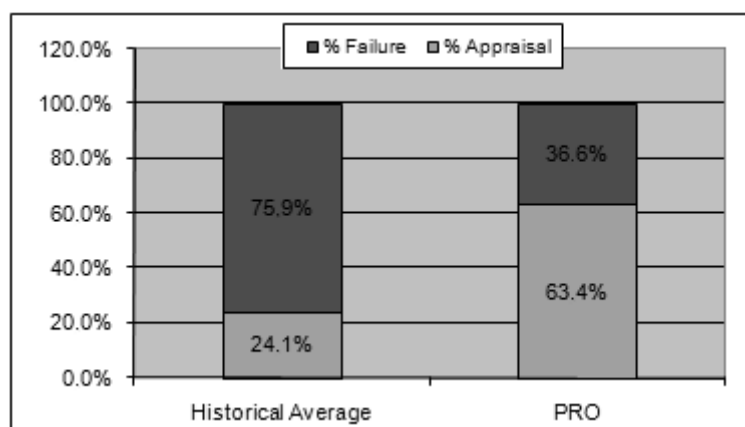


Figure 9 – Appraisal costs vs. Failure costs

Finally, Table 9 summarizes some quality and productivity measures used to verify the improvement.

Table 9 – Goal 3 results

Goal	Historical	PRO	Reduction
Defects before test [%]	35.9%	83.0%	-131.5%
Defects before operation [%]	85.7%	96.2%	-12.3%
Development Effort [%]	65.3%	64.5%	1.2%
Quality Effort [%]	34.7%	35.5%	-2.3%
Appraisal Cost [%]	24.1%	63.4%	-162.8%
Failure Cost [%]	75.9%	36.6%	51.7%
Defects removed [%]	85.7%	96.2%	-12.3%
Defects released [%]	14.3%	3.8%	73.7%

Conclusion

The use of TSPi principles in the new process allowed the accomplishment of project goals based on the following considerations.

1. The team integration, the detailed plan, the change management, the weekly meetings and the earned value method allowed the accomplishment of these goals.
2. Applying the early defect reduction principle, the test productivity has been increased and the re-work decreased.
3. The reviews, inspections and quality plans allowed a quality improvement. The team members understood the test phase as a quality evaluation and not as a defect detection activity.

With an affordable investment in process definition, it has been demonstrated that TSPi is an effective alternative solution for process, quality and productivity improvement in a Very Small enterprise.

Bibliography

- [Standish Group, 2007] Standish group. CHAOS Report, 2007.
- [Serrano, 2006] Serrano, M., Montes, C., Cedillo, K. An Experience on Implementing the CMMI in a Small Organization Using the TSP, 81-92. <http://www.sei.cmu.edu/pub/documents/06.reports/pdf/06sr001.pdf>, 2006.
- [IPRC, 2006] International Process Research Consortium. IPSS White Paper. Improving Process in Small Settings, 2006 <http://www.sei.cmu.edu/iprc/ipss-white-paper-v1-1.pdf>
- [Noopur, 2003] Noopur, D. The Team Software Process in Practice: A Summary of Recent Results. SEI Technical Report CMU/SEI-2003-TR-014, 2003.
- [Humphrey, 2006] Humphrey, W. TSP: Coaching Development Teams. Ed. Addison-Wesley Publishing Company, 2006
- [Humphrey, 1999] Humphrey, W. Introduction to the Team Software Process. Ed. Addison-Wesley Publishing Company, 1999.
- [Humphrey, 1995] Humphrey, W. A Discipline for Software Engineering. Ed. Addison-Wesley Publishing Company, 1995
- [Garcia, 2006] Garcia, S. Graettinger, C., Kost K. Proceedings of the First International Research Workshop for Process Improvement in Small Settings. SEI Special Report CMU/SEI-2006-SR-001, 2005.
- [Garcia, 2005] Garcia, S. Thoughts on Applying CMMI in Small Settings. <http://www.sei.cmu.edu/cmmi/adoption/pdf/garcia-thoughts.pdf>, 2005.
- [Glazer, 2006] Glazer, H. Time to Market vs. Process Discipline. <http://www.sei.cmu.edu/iprc/sepg2006/glazer.pdf>, 2006.
- [Calvo-Manzano, 2008] Calvo-Manzano, J., Gonzalo, C., San Feliu, T., Caballero, E. International Journal "Information Technologies and Knowledge" Vol.2 / 2008, 245-250. <http://www.foibg.com/ijitk/ijitk-vol02/ijitk02-3-p07.pdf>
-

Authors' Information

Caballero E. Henry – Universidad Politécnica de Madrid, Facultad de Informática, Campus Montegancedo, Boadilla del Monte-28660, Madrid-España; e-mail: ecaballero@zipi.fi.upm.es

Calvo-Manzano Jose A. – Universidad Politécnica de Madrid, Facultad de Informática, Campus Montegancedo, Boadilla del Monte-28660, Madrid-España; e-mail: jacalvo@fi.upm.es

Cuevas Gonzalo – Universidad Politécnica de Madrid, Facultad de Informática, Campus Montegancedo, Boadilla del Monte-28660, Madrid-España; e-mail: gcuevas@fi.upm.es

San Feliu Tomás – Universidad Politécnica de Madrid, Facultad de Informática, Campus Montegancedo, Boadilla del Monte-28660, Madrid-España; e-mail: tsanfe@fi.upm.es

STOCHASTIC MODELS AND METHODS OF OPERATIVE-DISPATCH CONTROL OVER THE GAS-TRANSPORT SYSTEM OF UKRAINE

Andriy Tevyashev

Abstract: Stochastic models and methods of solving the problem of operative-dispatch control of the gas-transport system of Ukraine are considered. It is shown that the problem of operative-dispatch control of the gas-transport system is reduced to solving of two linked problem: operative planning and stabilization of gas-transport system conditions.

Keywords: Gas-transport system: operative planning, operative-dispatch control, stochastic model, stochastic method.

ACM Classification Keywords: H.4.2 Types of Systems, G.1.6 Optimization: Stochastic programming.

Conference topic: Information Systems.

Conference: The paper is selected from Seventh International Conference on Information Research and Applications – i.Tech 2009, Varna, Bulgaria, June-July 2009

Introduction

Gas industry is one of the most important constituents of the Ukrainian fuel and energy complex, and it is hard to overestimate its influence on other branches of the national economy and national safety of Ukraine. The gas-transport system (GTS) is a basis of the Ukrainian gas industry.

The Ukrainian GTS is a unified technological complex and its two main functions are: transport and distribution of natural gas to Ukrainian consumers and transit of natural gas over Ukrainian territory to the countries of Central and Western Europe.

The main technological elements of the Ukrainian GTS include: six multiple-strand gas-main pipelines (GMP) 37,2 thousand kilometers length, 72 multiple-shop compressor stations (CST), 112 compressor shops (CSH), 786 gascompressor units (GCU), over 1420 gas-distributing stations (GDS) and 225 nonassociated gas fields.

Ukrainian gas-main pipelines (GMP) are connected with each other by system-defined connective gas pipelines, that ensure increasing of reliability indexes of the GTS and provide with the ability to manipulate with gas stream directions.

13 underground gas-holders (UGH) are used to guarantee technological stability of the Ukrainian GTS while volume of supplies and volume of natural gas usage fluctuate essentially. The main of those are located beside western borders of Ukraine, and their technical characteristics allow to smooth over considerable unevenness of gas extraction by importing countries and Ukrainian consumers. Moreover, Ukrainian UGH are used to store natural gas that has been preliminarily purchased by importing countries in Russia.

Spreading over great space and technical complexity of the Ukrainian GTS results in the fact that the processes of supply, transport and distribution of natural gas are performed by the three-level system of operative-dispatch control: amalgamated dispatching administration (ADA) DC "Ukrtransgas", control points (CP) of gas-main pipelines administrations (GMA), control points of line production administrations of gas-main pipelines (LPAGMP).

At present time “The Ukrainian energy strategy until 2030” has been elaborated and approved by the Ukrainian Government. It determines the objects and the main lines of development of the gas industry in a long-term and near-term outlook.

It is being planned to supply a demand for natural gas on the Ukrainian home market at the cost of both increasing of volume of output from home fields (particularly, at the cost of forthcoming development of the Ukrainian part of shelves in Black Sea and Sea of Azov) and organization of hydrocarbon extraction and import of them. The import of natural gas policy should be guided by principles of diversification of supplies and routes of natural gas delivery with the object of ensuring the energy security of Ukraine. However, at present time the Russian Federation is the main source of export natural gas in Ukraine. This determines a strict technological and organizational interconnection between unified gas-supply system (UGSS) of Russian Federation and the Ukrainian GTS.

Signed contracts and technical agreement are the main documents to regulate operating modes of the Ukrainian GTS. This contracts include yearly, quarterly and monthly volumes of receiving and delivery of natural gas, and operation factors of gas flows (pressure, temperature, composition and humidity of natural gas), that should be guaranteed at inlets and outlets of the Ukrainian GTS. Monthly balances of supply, transport and distribution are created on basis of these documents, and they are detailed to 24 hours periods.

While realization of operational control over operating modes of the Ukrainian GTS the following factors are taken into consideration: actual operating mode of the technological equipment of the GTS, predictable volume of consumption of natural gas by home consumers (which is linked with the corresponding GDS's), and daily requests of importing countries for volumes of delivered natural gas, that come to the ADA DC “Ukrtransgas” from the control centre “Gasexport”, LLC and the CPDD of JSC “Gasprom”.

Signing of long-term contracts allows to forecast operating mode of the Ukrainian GTS for the nearest future and perform necessary technical maintenance, repair, reconstruction and modernization measures.

From a formal point of view, the Ukrainian GTS belongs to the class of object-orientated, multidimensional, manifold non-linear stochastic systems with distributed parameters, that are characterized by networked multilevel structure, availability of a DM in control circuit, availability of discrete and continuous command variables, high level of indeterminacy of objects, structure, parameters and conditions and influence of environment.

At present time there is no integrated solution of the problem of computer simulation and optimization of operating modes of the Ukrainian GTS both due to complexity of an adequate describing of reaction of the system to great number of deterministic and stochastic internal and external perturbations, and due to complexity of an adequate describing these perturbations themselves.

The results of system analysis of the Ukrainian GTS as a control object led to necessity of the Ukrainian GTS being considered as a stochastic object which is functioning in a stochastic environment [1].

1. Determination of the problem class of the problems of operative-dispatch control over the Ukrainian GTS.

In order to formalize the problem class of the operative-dispatch control of the Ukrainian GTS we consider it as an interconnected system of multiple-strand gas-main pipelines with multiple-shop compressor stations. From a technological point of you it is possible to represent the structure of the GTS as an ordered sequence of multiple-strand linear sections (LS) $LU = \{1, 2, 3, \dots, lu\}$ and multiple-shop compressor stations (CS)

$CS = \{1, 2, 3, \dots, cs\}$, From a formal point of view the structure of the GTS may be represented as a directed graph $G(V, E)$, where

$V = W \cup N = \{1, 2, \dots, w, w+1, w+2, \dots, n\}$ – indexing set of points of graph;

$W = \{1, 2, \dots, w\}$ – indexing set of points of graph (inputs), that natural gas comes into the GTS through;

$N = \{w+1, w+2, \dots, w+n\}$ – indexing set of points of graph (oupputs), that natural gas is taken out from the GTS from.

$E = L \cup M = \{1, 2, \dots, l, l+1, l+2, \dots, l+m\}$ – indexing set of arcs of graph of the GTS;

$L = \{1, 2, \dots, l\}$ – indexing set of arcs of graph that correspond to active elements of the GTS, i.e. CS.

$M = \{l+1, l+2, \dots, l+m\}$ – indexing set of arcs of graph that correspond to passive elements of the GTS, i.e. gas cleaning plants, gas cooling plants, pipeline sections, intercepting and controlling faucets etc.;

(Ω, B, P) – probability space with the defined flow of σ -algebras $\{B_t, t \in [0, T]\}$ on it, i.e. family of σ -algebras $B_t \subset B$, that from $t < t_1$ follows $B_t \subset B$ for;

Ω – space of elementary events;

B – σ -algebra of events from Ω ;

P – probability measure on B ;

$[0, T]$ – control interval.

Condition of the GTS in every instant of time $t \in [0, T]$ is characterized by a random vector

$$S(t, \omega) = \langle X(t, \omega), Y(t, \omega), U(t) \rangle, \quad (1)$$

where

$$X(t, \omega) = \langle X_1(t, \omega), X_2(t, \omega), \dots, X_i(t, \omega), \dots, X_w(t, \omega) \rangle \quad (2)$$

random column-vector of parameters of natural gas on outputs of the GTS, and its every i -th, $i \in W$ component is of the form of

$$X_i(t, \omega) = \langle P_i(t, \omega), q_i(t, \omega), T_i^0(t, \omega), N_i(t, \omega), S_i(t, \omega) \rangle, i \in W \quad (3)$$

where $P_i(t, \omega), q_i(t, \omega), T_i^0(t, \omega)$ – random values that characterize, correspondingly, pressure, consumption and temperature of natural gas on the i -th input of the GTS at the instant of time t ;

$$N_i(t, \omega) = \langle N_{1i}(t, \omega), N_{2i}(t, \omega), \dots, N_{ji}(t, \omega), \dots, N_{ni}(t, \omega) \rangle \quad (4)$$

random vector, which characterizes hydrocarbon composition of natural gas, and $S_i(t, \omega)$ – humidity of natural gas on the i -th input of the GTS at the instant of time t ;

$$Y(t, \omega) = \langle Y_1(t, \omega), Y_2(t, \omega), \dots, Y_j(t, \omega), \dots, Y_n(t, \omega) \rangle \quad (5)$$

random column-vector of parameters of natural gas on outputs of the GTS (inputs of the GDS), and its every j -th component is of the form analogous to vector $X_i(t, \omega)$:

$$Y_j(t, \omega) = \langle P_j(t, \omega), q_j(t, \omega), T_j^0(t, \omega), N_j(t, \omega), S_j(t, \omega) \rangle, j \in N, \quad (6)$$

where $P_j(t, \omega), q_j(t, \omega), T_j^0(t, \omega), N_j(t, \omega), S_j(t, \omega)$ – random values that characterize, correspondingly, pressure, consumption, temperature, hydrocarbon composition and humidity of natural gas on the j -th input of the GTS at the instant of time t ;

$$U(t) = \langle d(t), b(t), \beta(t) \rangle \quad (7)$$

deterministic control vector, which determines work structure of LU – $d(t)$, work structure of KC – $b(t)$ and actual parameters of KC – $\beta(t)$ at the instant of time t . By-turn, each component of the control vector (7) is of the form of

$$d(t) = \langle d_1(t), d_2(t), \dots, d_i(t), \dots, d_{lu}(t) \rangle \quad (8)$$

vector of discrete, Boolean variables which determines work structure of all the lu linear sections of the gas-main pipelines of the GTS; each i -th component is of the form of

$$d_i(t) = \langle d_{1i}(t), d_{2i}(t), \dots, d_{ki}(t) \rangle, \quad i = 1, 2, \dots, lu \quad (9)$$

vector of discrete, Boolean variables that characterize the state (open/closed) of each of k intercepting faucets of the i -th linear section of the gas-main pipelines of the GTS.

$$b(t) = \langle b_1(t), b_2(t), \dots, b_i(t), \dots, b_{kc}(t) \rangle \quad (10)$$

vector of discrete, Boolean variables that characterize state work structure of all the kc linear sections of compressor stations; each j -th component is of the form of

$$b_j(t) = \langle b_{1j}(t), b_{2j}(t), \dots, b_{rj}(t) \rangle, \quad j = 1, 2, \dots, kc \quad (11)$$

vector of discrete, Boolean variables that characterize the state (open/closed) of each of r intercepting faucets of the j -th CS, which determines the work structure of the j -th CS (layouts of engaging GCUs, LSs and GCPs in each of CS)

$$\beta(t) = \langle \beta_1(t), \beta_2(t), \dots, \beta_j(t), \dots, \beta_{kc}(t) \rangle \quad (12)$$

vector of continuous variables which determines actual controllable parameters for all the kc compressor stations, each its j -th component is of the form of

$$\beta_j(t) = \langle \beta_{1j}(t), \beta_{2j}(t), \dots, \beta_{sj}(t) \rangle \quad (13)$$

vector s of continuous controllable variables of the j -th CS (operating revolutions of CAP GCU drive, state of regulating(anti-surge) faucets, steering angles of distributor motors etc.)

According to designation of the GTS on a specified interval of control $[0, T]$ a deterministic column-vector is defined:

$$Z(t) = \langle Z_1(t), Z_2(t), \dots, Z_j(t), \dots, Z_n(t) \rangle \quad (14)$$

It is the vector of needed (predictable or stated in a contract) values of parameters of gas flows on all the outputs of the GTS (inputs of the GDS); each its j -th component is of the form of

$$Z_j(t) = \langle P_j(t), q_j(t), T_j^0(t), N_j(t), S_j(t) \rangle, j \in N, \quad (15)$$

where $P_j(t), q_j(t), T_j^0(t), N_j(t), S_j(t)$ are deterministic values, that characterize needed (predictable or stated in a contract) values of parameters of natural gas (pressure $P_j(t)$, consumption $q_j(t)$, temperature $T_j^0(t)$, hydrocarbon composition $N_j(t)$ and humidity $S_j(t)$) on the j -th output of the GTS (input of the j -th GDS)

Besides:

$$J_{1j} [X(t, \omega), Y(t, \omega), U(t)], j \in N \quad (16)$$

$$J_{2i} [X(t, \omega), Y(t, \omega), U(t)], i \in KC \quad (17)$$

are given random functions that are measured on the strength of all the variables and, if it is fixed, are measured in regard to σ -algebra B_t ; they characterize, correspondingly, (16), i.e. quality of functioning of the GTS in regard to the j -th consumer (extent of execution of contract), and (17), i.e. efficiency of production of the i -th CS (unit costs in energy and value terms)

Using expressions (16)-(17), we can determine quality and efficiency indexes of functioning of the entire GTS.

$$J_1 [X(t, \omega), Y(t, \omega), U(t)] = \varphi_1 \{ J_{11} [X(t, \omega), Y(t, \omega), U(t)], \dots, J_{1n} [X(t, \omega), Y(t, \omega), U(t)] \} \quad (18)$$

where $J_1 [X(t, \omega), Y(t, \omega), U(t)]$ is a measured random function, which characterize quality of functioning of the GTS at the instant of time t in regard to all the consumers; $\varphi_1 \{ \}$ is a deterministic contraction function.

Index of efficiency of functioning of all the CSs at the instant of time t is also determined via indexes of efficiency of each CS and can be given as

$$J_2 [X(t, \omega), Y(t, \omega), U(t)] = \varphi_2 \{ J_{21} [X(t, \omega), Y(t, \omega), U(t)], \dots, J_{2kc} [X(t, \omega), Y(t, \omega), U(t)] \}, \quad (19)$$

where $\varphi_2 \{ \}$ is a deterministic contraction function of local indexes of efficiency of each CS.

In this case overall quality and efficiency index of the functioning of the GTS can be given as

$$J [X(t, \omega), Y(t, \omega), U(t)] = \varphi \{ J_1 [X(t, \omega), Y(t, \omega), U(t)], J_2 [X(t, \omega), Y(t, \omega), U(t)] \}, \quad (20)$$

where $\varphi \{ J_1 [X(t, \omega), Y(t, \omega), U(t)], J_2 [X(t, \omega), Y(t, \omega), U(t)] \}$ is a deterministic contraction function of quality and efficiency indexes of the functioning of the GTS at the instant of time t .

We select contraction functions $\varphi \{ \}, \varphi_1 \{ \}, \varphi_2 \{ \}$ the way so maximum of quality and efficiency of the functioning of the GTS is reached at the instant of time t , given the expression (20) reaches its minimum.

Since $J [X(t, \omega), Y(t, \omega), U(t)]$ is a random function, it is necessary to bring a functional of this function in order to state the exact problem. In lieu of this function mathematical expectation, probability or variance of a random function on a fixed interval can be used. It is reasonable to use mathematical expectation in order to minimize average losses. In this case the problem of operational control over transport modes and distribution of

natural gas in the GTS on a time interval $[0, T]$ reduces to the problem of minimization of the functional of the form of

$$J(T) = M_{\omega} \int_0^T J [X(t, \omega), Y(t, \omega), U(t)] dt \rightarrow \min_{U(t) \in \Omega}, \quad (21)$$

where M_{ω} is a mathematical expectation symbol.

The region of feasibility is described by a system of equations which determines interrelation between $X(t, \omega)$, $Y(t, \omega)$ and $U(t)$, and also systems of inequalities that determine area of technologically acceptable transport modes and distribution of natural gas in the GTS.

Technological limits should be fulfilled for all the instant of times $t \in [0, T]$ and for all the realizations of random components of vectors $X(t, \omega)$, $Y(t, \omega)$.

The problem (21) belongs to the problem class of optimal stochastic control of non-linear discrete-continuous type and of a large scale. At present time there are no general methods of solving this problem, so we consider a particular method of solution of this problem which allows for specifics of its informational supplying, also conditions of physical realizability and the structure of the control vector.

2. Structurization of the Problem Class of Operative-dispatch Control of the Transport Modes and Natural Gas Distribution in the GTS

Conditions of physical realizability of the selection and processing information systems define discrete nature of the information about values of realizations of random components of vectors $X(t, \tilde{\omega})$ and $Y(t, \tilde{\omega})$ on inputs and outputs of the GTS, which comes at discrete instants of time $k = 1, 2, \dots, K$ with a step $\Delta t = T / K$. In this case discrete analogue of the problem (21) is of the form of

$$J(K) = M_{\omega} \left\{ \sum_{k=1}^K J [X(k, \omega), Y(k, \omega), U(k)] \right\} \rightarrow \min_{U(k) \in \Omega} \quad (22)$$

Persistence of the GTS control system, which is related to considerable time and energy losses on changing of the structure of the GTS (structures of LS and CS), results in necessity of making the decision $U_0^*(k)$, $k = 1, 2, \dots, K$ at a null instant of time, i.e. until observation of realization of random vectors $X(t, \omega)$ and $Y(t, \omega)$ $k = 1, 2, \dots, K$. After observation of realization of $X(t, \omega)$, $Y(t, \omega)$ at the instants of times $k = 1, 2, \dots, K$ the ability to calculate the value $J [X(k, \tilde{\omega}), Y(k, \tilde{\omega}), U(k)]$ appear, and, on its basis calculate deflection of the actual mode from the planned one and correct the primary solution $U_0^*(k)$. Thus, it is needed to find an optimal solution of the problem (22) $U^*(k)$, $k = 1, 2, \dots, K$ of the form of sum of two vectors

$$U^*(k) = U_0^*(k) + \delta U^*(k), \quad k = 1, 2, \dots, K \quad (23)$$

where $U_0^*(k)$ is a control vector which determines an optimal plan of transport mode and natural gas distribution in the GTS; this plan is calculated at the null instant of time until observation of realization of random conditions of the problem.

$\delta U^*(k)$ is a control vector which determines an optimal compensation of the plan and it is calculated in each instant of time $k = 1, 2, \dots, K$ after observation of actual realizations of random conditions of the problem.

Conditions of physical realizability of the process of control over transport modes of the GTS determine the structure of vectors $U_0^*(k)$ and $\delta U^*(k)$:

$$U_0^*(k) = \langle d_0^*(k), b_0^*(k), \beta_0^*(k) \rangle, \quad (24)$$

$$\delta U^*(k) = \langle \beta^*(k) \rangle, \quad k = 1, 2, \dots, K, \quad (25)$$

This is concerned to the fact that it is not reasonable to change the structure of linear sections of gas-main pipelines and the structure of compressor stations shops in a real-time mode; moreover, it is impossible in most occasions. In other words, it is not advisable to use these variables for operative compensations of the plan mistakes.

At present time it is only possible to operatively change the parameters of operating of GCU CS relative to their planned values (settings). That's why only these variables are components of the vector of optimal plan compensation $\delta U^*(k)$.

The problem (22) under conditions (24), (25) belongs to the class of multiple-steps problems of non-linear stochastic programming at a tough statement, and a solution should be a pure strategy one, of the form of decision rule of a order zero which links unknown components of the control vector with the parameters of the GTS and parameters of stochastic processes $X(k, \omega), Y(k, \omega)$.

We consider, that at the null instant of time $k = 0$ we know all the previous history (informational condition of the system) $X(k, \tilde{\omega}), Y(k, \tilde{\omega}), U(k)$, $k = 0, -1, -2, \dots$. Then for $k = 1, 2, \dots, K$ we can get estimations of conditional expectations $\hat{X}_0(k), \hat{Y}_0(k)$ of vectorial stochastic processes $X(k, \omega), Y(k, \omega)$ and their correlation matrixes $K_X(k), K_Y(k)$, i.e. predictions of future values $X(k)$ and $Y(k)$ that are calculated at the null instant $k = 0$ with prediction $k = 1, 2, \dots, K$. It is also considered, that estimates $\hat{Y}_0(k)$ and $\hat{K}_Y(k)$ are calculated with some fixed values $U_0(k)$, $k = 1, 2, \dots, K$.

While expanding $J[X(k, \omega), Y(k, \omega), U(k)]$ in a Taylor series relative to a nominal trajectory $\langle \hat{X}_0(k), \hat{Y}_0(k), U_0(k) \rangle$ with saving all the terms not above the second one and applying the mathematical expectation operation to the both parts of the derived equality, we get:

$$\begin{aligned} M_{\omega} \{ J[X(k, \omega), Y(k, \omega), U(k)] \} &= J_0[\hat{X}_0(k), \hat{Y}_0(k), U_0(k)] + \partial J_0[\sigma_{X(k)}^2, \sigma_{Y(k)}^2, \sigma_{U(k)}^2], \quad (26) \\ \partial J_0[\sigma_{X(k)}^2, \sigma_{Y(k)}^2, \sigma_{U(k)}^2] &= \frac{1}{2} \left[\sum_{i=1}^w \left(\frac{\partial^2 J}{\partial X_i^2(k)} \right)_0 \sigma_{X_i(k)}^2 + \sum_{i < j} \left(\frac{\partial^2 J}{\partial X_i(k) \partial X_j(k)} \right)_0 K_X(k) + \right. \\ &\quad \left. + \sum_{j=1}^n \left(\frac{\partial^2 J}{\partial X_i^2(k)} \right)_0 \sigma_{Y_i(k)}^2 + \sum_{j < p} \left(\frac{\partial^2 J}{\partial Y_j(k) \partial Y_p(k)} \right)_0 K_Y(k) + \sum_{i=1}^l \left(\frac{\partial^2 J}{\partial U_i(k)} \right)_0 \delta_{U(k)}^2 \right] \end{aligned}$$

$$\delta J_0 \left[\sigma_{x(k)}^2, \sigma_{y(k)}^2, \sigma_{u(k)}^2 \right] = \frac{1}{2} \left[\sum_{i=1}^w \left(\frac{\partial^2 J}{\partial X_i^2(k)} \right)_0 \sigma_{x_j(k)}^2 + \sum_{i < j} \left(\frac{\partial^2 J}{\partial X_i(k) \partial X_j(k)} \right)_0 K_X(k) + \right. \\ \left. + \sum_{j=1}^n \left(\frac{\partial^2 J}{\partial X_j^2(k)} \right)_0 \sigma_{y_j(k)}^2 + \sum_{j < p} \left(\frac{\partial^2 J}{\partial Y_j(k) \partial Y_p(k)} \right)_0 K_Y(k) + \sum_{i=1}^l \left(\frac{\partial^2 J}{\partial U_i(k)} \right)_0 \delta_{u(k)}^2 \right]; \quad (27)$$

where

$$\sigma_{X_i(k)}^2 = M_{\omega} \{ X_i(k, \omega) - X_{i0}(k) \}^2, \quad i \in W \quad (28)$$

is a variance of random value relative to its conditional mathematical expectation $Y_{i0}(k)$;

$$\delta_{iU(k)}^2 = \{ U_i(k, \omega) - U_{i0}(k) \}^2, \quad i \in KC, \quad (30)$$

is a square of deviation of the i -th component of the control vector $U_i(k)$ from its nominal value $U_{i0}(k)$.

If we substitute (2) into (22), we get

$$J(K) = \sum_{k=1}^K \left\{ J_0 \left[\hat{X}_0(k), \hat{Y}_0(k), U_0(k) \right] + \delta J_0 \left[\sigma_{\hat{X}(k)}^2, \sigma_{\hat{Y}(k)}^2, \sigma_{U(k)}^2 \right] \right\} \rightarrow \min_{\langle U_0(k), \delta U(k) \rangle \in \Omega} \quad (31)$$

Thus, solution of the stochastic problem (22) can be get in two steps as a result of resolving two deterministic problems.

On the first step, until observation of realization of random conditions, the problem of operational planning of transport modes and natural gas distribution in the GTS is being solved:

$$J_0(K) = \sum_{k=1}^K J_0 \left[\hat{X}_0(k), \hat{Y}_0(k), U_0(k) \right] \rightarrow \min_{U_0(k) \in \Omega} \quad (32)$$

The solution of the problem (22) allows to get at the null instant of time the optimal control vector $U_0^*(k)$, which determines the structures of all the LSs and CSs and parameters of CSs that guarantee realization of the optimal transport mode and natural gas distribution in the GTS on a time interval $[0, T]$.

On the second step, after each observation of realizations of random conditions at null instants of time $k = 1, 2, \dots, K$, the following problem is being solved:

$$\delta J_0(K) = \sum_{k=1}^K \delta J_0 \left[\sigma_{\hat{X}(k)}^2, \sigma_{\hat{Y}(k)}^2, \sigma_{U(k)}^2 \right] \rightarrow \min_{\delta U(k) \in \Omega}; \quad (33)$$

it is a problem of stabilization of transport mode and natural gas distribution in the GTS.

The problems (32), (33), as this follows from the (31), are roughly linked between themselves and actually are the solution of one problem (22). A feature of this solution is the fact that the optimal plan $U_0^*(k)$ should be the one so the problem of the second step (33) is resolvable for all $k = 1, 2, \dots, K$ and all possible realizations of random values. Thus, resolvability conditions of the second step problem induce additional limits on a constraint region of the problem of planning regime Ω_0 .

The problem of operative planning is being solved in a conventional real-time mode by the use of facilities of IAS ADA in ADA DC "Ukrtransgas" with specification on all the six GMAs.

The problem of stabilization of transport mode is being solved in a real-time mode by the use of facilities local automatics (SAC and ACS TC of technological objects and machines of the GTS).

Conclusion

1. Stochastic models and methods of solving the problem of operative-dispatch control of the gas-transport system are considered.
 2. It is shown that the problem of operative-dispatch control of the gas-transport system is reduced to solving of two linked problems:
 - operative planning of the conditions according to the forecasting of the basic disturbing factor of the gas-transport system. This factor is gas consuming;
 - stabilization of gas-transport system conditions with local automatic machinery after supervision over actual NG consuming in each GDS.
 3. Conversion of operative planning problem is carried out and it is shown that its decision can be received as a result of solving the two interconnected problems:
 - structural and parametric optimization of the LS GTS (local subsystem of GTS);
 - coordination of boundary conditions determined the LS GTS work regime.
-

Bibliography

1. Евдокимов А.Г., Тевяшев А.Д., Дубровский В.В. Моделирование и оптимизация потокораспределения в инженерных сетях. – М.: Стройиздат, 1990. – 365 с.
 2. Евдокимов А.Г., Дубровский В.В., Тевяшев А.Д., Потокораспределение в инженерных сетях. – М.: Стройиздат, 1979.- 199 с.
 3. Тевяшев А.Д., Тевяшева О.А., Фролов В.А. Стохастичний підхід до оцінювання ступеня технологічної стійкості режимів роботи газотранспортних систем.// Науково-виробничий журнал України "Нафтова и газова промисловість" - 2006 р., №4. - С.49-52.
 4. Тевяшев А.Д., Тевяшева О.А., Фролов В.А. Багатокритеріальна оптимізація планових режимів роботи газотранспортних систем // Науково-виробничий журнал України "Нафтова и газова промисловість" - 2006 р. №2 - С. 46-47.
 5. Артюх А.Ю., Тевяшев А.Д., Фролов В.А. Стохастические модели и методы оптимизации режима работы компрессорных станций. // Всеукраинский межведомственный научно-технический сборник. «Автоматизированные системы управления и приборы автоматики» - Харьков, ХНУРЭ, -2005, Выпуск 133 - С. 26-36.
-

Authors' Information

Andriy Tevyashev – Professor, Dr, Head of the Department of applied mathematics, Kharkiv National University of Radioelectronics, Lenina pr. 14, Kharkiv, Ukraine, e-mail: tad45@mail.ru

ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ В СИСТЕМАХ УПРАВЛЕНИЯ ПРОЦЕССОМ ФРАХТОВАНИЯ СУДОВ

Сергей Шехунов, Антон Моторкин

Аннотация: В данной статье описан процесс фрахтования судов и его основные проблемы. Описаны проблемы анализа фрахтового рынка. Рассмотрены системы автоматизации деятельности менеджера по отфрахтованию: веб-портал с хранилищем данных и клиентское приложение для анализа данных и поддержки принятия профессиональных решений. Предложены методы решения основных проблем автоматизации и анализа данных. Предложен метод поддержки принятия решений о выборе оптимальных пар заявок для заключения контракта.

Ключевые слова: автоматизация, поддержка принятия решений, теория полезности, OLAP, фрахтование.

Conference: The paper is selected from XVth International Conference "Knowledge-Dialogue-Solution" KDS 2009, Varna, Bulgaria, June-July 2009

Введение

Фрахтование судов далеко не древняя профессия. Она возникла в последней половине XIX века. Ее появление было, с одной стороны, обусловлено нуждами торговли и промышленности, а с другой, — заинтересованностью судовладельцев в создании некоего органа, с помощью которого можно было бы эффективнее использовать быстрорастущий флот в обстановке роста и усложнения международного товарооборота.

Фрахтование судна, в простейшем рассмотрении, является договором аренды, по которому фрахтователь нанимает судно у его владельца. Обычно этот фрахтователь владеет грузом, который необходимо перевезти из одной части мира в другую, и, пока в его собственном распоряжении нет судов, он будет зависеть от других компаний. Деньги за наем (фрахт) выплачиваются судовладельцу в качестве вознаграждения за использование его судна.

Прежде чем выступить на фрахтовом рынке и заключить договор фрахтования от имени фрахтователя (продавца) морской брокер изучает общую обстановку (коммерческие условия судоходства) сложившуюся на фрахтовом рынке. Изучаются наличие свободного тоннажа во фрахтовой секции, район освобождения судов, предложения на них, состав фирм-судовладельцев, возможные каналы заключения фрахтовых сделок, уровень фрахтовых ставок и тенденции к их изменению и т.п.

После обзора конъюнктурной ситуации фрахтового рынка или отдельных секций начинается предметное изучение котировок, исследуемых с точки зрения возможности заключения сделки с тем или иным судовладельцем.

Этап изучения рынка является наиболее трудоемким, так как брокеру приходится иметь дело с большим и постоянно растущим количеством заявок. Поэтому предоставление средств к быстрому и удобному обзору фрахтового рынка является важнейшей задачей систем автоматизации деятельности морских брокеров.

После анализа рынка происходит выбор наиболее экономически выгодного и целесообразного предложения. Для этого оценивается возможность быстроты заключения договора, производится расчет экономических результатов в случае реализации каждого из предложенных вариантов.

Конечной стадией всей процедуры заключения договора фрахтования судна является согласование условий чартера. После окончания согласования условий отфрахтования судна брокер уведомляет стороны. Затем брокер составляет и подписывает от имени фрахтователя чартер, копии которого рассылаются заинтересованным в реализации сделки сторонам и участникам перевозки груза или аренды судна.

Проблема анализа фрахтового рынка

Современный мировой открытый фрахтовый рынок - это сфера обращения специфического товара - морских транспортных услуг. Для фрахтового рынка характерны неустойчивость, стихийность развития, постоянная подверженность кризисам, оживлениям и спекулятивным факторам. Одной из особенностей мирового открытого фрахтового рынка является значительные колебания уровня фрахтовых ставок, происходящие вследствие влияния многочисленных факторов экономического, политического и природного характера.

Целью анализа текущего уровня ставок фрахта рынка является оперативное планирование деятельности брокерских компаний. Изучение активности рынка и факторов, влияющих на состояние рынка, позволяет выявить закономерности изменения ставок, делать прогнозы для стратегического планирования.

Фрахтовый рынок активно исследуется аналитиками, которые с определенной периодичностью публикуют в различных СМИ свои отчеты. Например, в отчет о состоянии и прогнозе развития фрахтового рынка сухогрузного тоннажа входят следующие сводные таблицы: уровень ставок при отфрахтовании сухогрузного тоннажа в рейсовый тайм-чартер, цены на новострой, уровень тайм-чартерных ставок в зависимости от продолжительности отфрахтовок для балкерного тоннажа и т.п.

Для решения задач автоматизации деятельности морских брокеров на рынке программного обеспечения уже разработано несколько систем. Невозможность их использования для анализа фрахтового рынка обусловлена общими для такого рода систем недостатками:

- Частично, благодаря возможностям обновления заявок с сервера, обработки писем от других менеджеров, пользователь имеет представление только об отдельных предложениях, являющихся частью ограниченного рынка.
- Не реализована функция анализа данных, позволяющая брокеру принять решение о начале работы в наиболее перспективных направлениях. Отсутствует интеллектуальная поддержка принятия решений пользователя.
- Все данные в основном представлены в виде одномерных таблиц. Такое представление очень неудобно для анализа.

Автоматизация деятельности морского брокера с помощью централизованного хранилища данных и web-портала CargoGuru

Глобальный масштаб и географическая широта оказания услуг морских брокеров, в силу своей природы, предполагают хранение, сравнение и анализ многих тысяч предложений и заявок сторонних компаний. Часто подобная задача сводится к поиску иголки в стоге сена. Подобная проблема в своем решении предполагает наличие единого централизованного хранилища данных, в которое будут "стекаться" предложения и заявки со всего мира, что позволит проинформировать о требуемой или оказываемой услуге сразу всех кому она интересна.

Подобным хранилищем является база данных системы CargoGuru, взаимодействие с которой осуществляется с помощью web-портала CargoGuru или клиентского приложения CargoGuruClient.

Основные цели и задачи хранилища CargoGuru:

- Объединение тысяч морских брокеров в одну глобальную систему и создания единого хранилища данных.
- Быстрый анализ и поиск встречных предложений сторонних морских брокеров.
- Поиск оптимально подходящих встречных предложений, основанный на базе знаний.
- Доступность из любой точки мира (при наличии подключения к Интернет и web-браузера).
- Возможность взаимодействия с приложением CargoGuruClient для расширения аналитических функций над данными, автоматизации взаимодействия с хранилищем, уменьшения однообразной работы и повышения эффективности трудозатрат.
- Моментальное оповещение брокера о появлении новых интересных для него предложений и заявок.
- Уменьшение звеньев в цепочках посредников таких как «Грузовладелец – Брокеры – Судовладелец», «Фрахтователь – Брокеры – Судовладелец» и др., посредством прозрачности и расширения представления о большей части рынка услуг морских брокеров.
- Динамические структуры предложений брокеров.
- Создание гибких и удобных интерфейсов доступных для неопытных пользователей персональных компьютеров и мобильных устройств.

Анализ информации о заявках в системе CargoGuruClient

Система автоматизации деятельности брокерских компаний CargoGuruClient вобрала в себя всю необходимую функциональность для решения основных задач профессиональной деятельности морского брокера:

- Систематизация информации о заявках и рабочих контактах, поступающей брокеру от инородных компаний.
- Организация поиска необходимой информации по любым параметрам в систематизированных данных.
- Организация доступа к широкому рынку заявок.
- Поддержка брокера в выборе наиболее подходящих пар заявок для дальнейшего заключения контрактов.
- Анализ состояния рынка фрахтовых заявок.

Задачи систематизации и простого поиска решены путем обработки базы данных, содержащей информацию о заявках, поступающих от брокеров или инородных компаний. Все манипуляции (добавление, обновление, удаление, поиск) производятся при помощи SQL запросов к базе данных.

Доступ к рынку заявок обеспечен через взаимодействие с Интернет порталом, предоставляющим специально спроектированные и согласованные по передаваемым структурам данных веб-службы. Эти службы дают возможность принимать заявки от клиентской программы и отсылать ей требуемые данные, обновлять базу справочных данных по портам, вести поиск заявок, размещенных другими пользователями в сети. Для этого были созданы библиотеки функций на основе WSDL-спецификаций (Web Service Description Language) и классы-посредники между клиентской и серверной частью системы.

Передача данных к portalу производится по протоколу SOAP, который включает в себя спецификацию сервисов на языке WSDL и технологии передачи запросов и ответов от сервера HTTP и SMTP.

Возможность локализации системы для нескольких языков заметно расширяет аудиторию пользователей системы, что позволяет пользователю работать с заявками всего мирового рынка.

Для того чтобы система могла давать брокеру поддержку в пользу принятия решения по выбору наиболее подходящих пар заявок, в CargoGuruClient включена интегрированная экспертная подсистема. В основе этой подсистемы лежит метод многокритериальной полезности, применение которого позволяет упорядочить все пары предложений по возрастанию степени оптимальности их выбора для заключения контракта.

В основу многокритериальной теории полезности (МТП, MAUT - Multi-Attribute Utility Theory) положено предположение, что варианты решений имеют оценки по многим критериям. Функция полезности имеет вид суммы произведений значения каждой альтернативы по критерию на весовой коэффициент этого критерия.

Подход МТП подразумевает следующие этапы:

- Разработать перечень критериев.
- Построить функции полезности по каждому из критериев.
- Оценить все имеющиеся альтернативы и выбрать наилучшую.

Перечень критериев вместе с функциями полезности в системе CargoGuruClient формируется в специальном редакторе знаний, где в виде пошагового диалога пользователю предлагается ввести условия для критериев и оценить их вес. Поиск лучшей альтернативы происходит автоматически на основе единожды введенных знаний.

Преимуществами предложенного метода являются:

- Возможность проверки большого количества альтернатив в короткие сроки без участия пользователя, кроме как для изначального назначения весов критериев.
- Возможность изменения цели поиска оптимальной пары путем изменения весов критериев.

Анализ фрахтового рынка реализован за счет набора функций многомерной отчетности, основанной на особой технологии репортинга OLAP-систем. Технология OLAP позволяет пользователям сформировать свое собственное видение данных, используя быстрый, единообразный, оперативный доступ к разнообразным формам представления информации.

OLAP-отчет может иметь экранную и бумажную формы. OLAP-отчет в экранном виде позволяет манипулировать данными и формой отображения. Любой полученный экранный отчет можно распечатать на бумагу в том виде, как он выглядел на экране.

OLAP-отчет отличается рядом принципиальных особенностей:

- отчет, предоставляющий пользователю высоко-интерактивный способ работы с данными;
- агрегированный отчет, позволяющий углубиться в детали;
- отчет, позволяющий легко изменять аналитические разрезы путем изменения порядка следования полей;
- отчет, позволяющий мгновенно фильтровать данные по всем возможным сочетаниям
- экранный отчет с возможностью вывода на принтер;
- отчет, настраиваемый пользователем без программирования;
- отчет, имеющий простую, регулярную структуру.

Заключение

В данной статье были предложены к рассмотрению решения для автоматизации деятельности морского брокера: системы CargoGuru и CargoGuruClient.

Основным преимуществом системы CargoGuru является широта и полнота функций обработки данных, предоставляемых клиентским приложением CargoGuruClient. Но в тоже время все необходимые задачи брокера могут быть решены посредством использования web-сайта CargoGuru, который воплотил в себе все необходимое для этого. Система CargoGuru - это связующее звено между тысячами пользователей приложения CargoGuruClient расположенных по всему миру, решающих свои профессиональные задачи и помогающие решить их другим.

Система CargoGuruClient синтезировала в себе возможности рядовых систем обработки заявок и инновационный подход к обеспечению интерактивного взаимодействия менеджера по отфрахтованию с мировым рынком и процессу поддержки принятия профессиональных решений и может быть использована как для автоматизации рутинных операции профессиональной деятельности морского брокера, так и для сложного интеллектуального анализа данных о заявках.

Благодарности

Статья частично финансированна из проекта **ITHEA XXI** Института Информационных теории и Приложений FOI ITHEA и Консорциума FOI Bulgaria (www.ithea.org, www.foibg.com).

Источники литературы

Правовое регулирование внешнеторговых перевозок : [курс. раб.] / Волжская Гос. академия водного транспорта; рук. д.т.н. Асеев С.В; исполн. Максимов С.В. – Нижний Новгород, 1998. – 31 с.

«Грузоперевозки России» — портал о грузоперевозках / 2007. - <http://tranzgruz.info>

Внедрение OLAP [Электронный ресурс] / Interface Ltd. - <http://www.interface.ru/>

Акимова О.В. К вопросу изучения конъюнктуры фрахтового рынка / Одесский национальный морской университет, Украина

Информация об авторах

Сергей Шехунов, Антон Моторкин– *Дальневосточный государственный университет, Институт математики и компьютерных наук, Адрес: 690048, Россия, Владивосток, пр-кт 100 лет Владивостоку 46 - 62, e-mail: musersheik@yandex.ru*

ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИИ SEMANTIC WEB ДЛЯ ИНТЕЛЛЕКТУАЛЬНОГО УПРАВЛЕНИЯ В ДИНАМИЧЕСКИХ РАСПРЕДЕЛЕННЫХ СИСТЕМАХ

Анатолий Гладун, Юлия Рогушина

Аннотация: Предложен новый подход к использованию знаний о предметной области для интеллектуального управления сложными гетерогенными динамическими системами. Распределенная система рассматривается как набор гетерогенных информационных ресурсов, распределенных в телекоммуникационной сети и доступных при помощи общесетевой инфраструктуры. Для интеллектуализации этой инфраструктуры предлагается использовать знание-ориентированные технологии Semantic Web, позволяющие оптимизировать распределение сетевых ресурсов.

Ключевые слова: Service oriented architectures, network management, Semantic Web, inductive reasoning, ontologies, agent, multimedia service.

ACM Classification Keywords: I.2.4 Knowledge Representation Formalisms and Methods

Conference: The paper is selected from XVth International Conference "Knowledge-Dialogue-Solution" KDS 2009, Varna, Bulgaria, June-July 2009

Введение

Повсеместное распространение Internet, мобильная телефония и переносные компьютеры позволяют сегодня пользователю использовать целый спектр новых приложений, поддерживаемых разработчиками сервисов. Одновременно возрастает сложность как пользовательских устройств самой сети (в частности, на базе Интернет формируется новая концепция развития сетей NGN – сети следующего поколения), так и сервисов, которые предоставляются пользователям, и сервисных платформ, которые увеличивают доступность и количество дополнительных сервисов в сети. При этом задачи управления такими сетями усложняются [1], а это, в свою очередь, требует поиска новых путей, технологий, методов и средств проектирования, разработки и интеллектуализации сложных систем управления.

Новые технологии в телекоммуникациях вызывают потребность поддержки качества сервисов, защиты и мобильности для новых сервисов, которые могут быть доступны сегодня «в любом месте, в любое время» в гетерогенной сети. Существующие проблемы значительно замедляют процесс разработки прикладных сервисов и не обеспечивают выполнение пользовательского запроса вовремя.

Один из путей решения этой проблемы – сервис-ориентированный подход к реализации задач управления сетью и обслуживания клиентов в мобильных сетях следующего поколения, вследствие чего сеть переопределяется как коллекция гетерогенных информационных ресурсов (ГИР), распределенных в сети и доступных при помощи общесетевой инфраструктуры. При этом необходимо оптимизировать распределение этих ресурсов таким образом, чтобы увеличить эффективность сети. Для этого нужно увеличить "уровень интеллекта" сети, т.е. использовать знания предметной области (ПрО) и механизмы принятия решения, которые позволяют трансформировать в знания собранную в сети информацию.

Развитие Интернет и распространение мобильной телефонии обострили проблемы доступности ГИР, качества передачи информации и предоставления услуг. Последние достижения в микроэлектронике и, стимулируемые ею, новые технологии связи, вызвали концептуальный сдвиг в предоставлении сервисов – от простого наличия сервиса к уже приемлемому качеству сервиса, от автономных вычислений – к

вычислениям, ориентированным на группу, и от голосовой связи – к мультимедиа. Эти проблемы являются открытыми для будущих исследований и для дальнейшего развития основных протоколов и сетевых технологий. Кроме того, расширения в существующих технологиях, типа протокола Internet (IPv6), подсистема мультимедиа (IMS) увеличивают сложность сети еще больше.

Одной из главных задач, требующих сегодня немедленного решения, является качество обслуживания (QoS). QoS определяет гарантию, данную сетью пользователю при выполнении сервисов с учетом набора сервисных ограничений, касающихся показателей качества: непрерывная задержка, флуктуация (джиттер), доступная полоса пропускания и вероятность потери пакета [2]. Для мультимедиа QoS связано в первую очередь с распределением ресурсов сети, цель которого – резервировать ресурсы таким образом, чтобы QoS смогло удовлетворить требования всех приложений [3].

Кроме того, в задачи QoS входит обеспечение надежной, высокоскоростной, непрерывной связи в гетерогенной сети следующего поколения NGN (Next Generation Network) [4,5].

Большую роль в поддержке QoS и обеспечении надежной связи играет механизм хендовера (handover). Необходимо обеспечить динамичную адаптацию к мобильности абонентов (процедура хендовера), а также сегодня возникает потребность в защите от разнообразных злонамеренных нападений. Таким образом, сеть должна иметь базовый механизм, который обеспечивает функциональные возможности: QoS, защиту и мобильность.

Пользователь стремится получить сервис, который удовлетворяет его потребности за приемлемую цену. Этот подход может быть осуществлен на основе предлагаемой интеллектуальной сетевой инфраструктуры, обеспечивающей компоненты (QoS, защита, мобильность) средствами для достижения поставленных целей. Таким образом, требуется более интеллектуальная сеть и более интеллектуальные сервисы.

Постановка задачи

Развитие новых технологий в телекоммуникациях вызывает необходимость поддержки качества сервисов, доступных «в любом месте, в любое время» в гетерогенной сети, которую можно представить, как коллекцию гетерогенных информационных ресурсов (ГИР), доступных при помощи общесетевой инфраструктуры. Необходимо оптимизировать распределение этих ресурсов между пользователями и сервисами таким образом, чтобы увеличить эффективность всей сети. Средством этого является увеличение интеллектуальности системы управления сетью на основе знаний, применения агентных технологий, метаданных, онтологий и других средств Semantic Web. Для решения этих задач необходимо разработать соответствующие модели и методы использования технологии Semantic Web для управления гетерогенными сетями.

Сети NGN

Сети NGN базируется на Интернет – постоянно растущей глобальной IP-сети. Фундаментальная цель Интернет – обеспечение связи, то есть разработка технологий, позволяющих связать существующие сети. Эти сети соединены с Интернет пакетной коммутацией посредством шлюзов. Вспомогательные цели Интернет – обеспечение живучести, поддержка разных типов сервисов и приспособленность к разнообразию сетей. Сегодня мы наблюдаем *конвергенцию* сетей, т.е. реализацию общих архитектурных принципов построения телекоммуникационных сетей, поддержку единых совместимых протоколов и интерфейсов, а также возможностей разных сетевых платформ предоставлять аналогичные виды сервисов независимо от типа технологий доступа [6].

NGN – мультисервисная сеть на базе систем коммутации пакетов и широкополосных транспортных технологий, которая предоставляет пользователям выбор телекоммуникационных услуг разных провайдеров, поддерживает конвергенцию сетей и обеспечивает при этом требуемое качество услуг QoS и мобильность абонентов [7,8]. Поскольку в NGN увеличиваются как мобильность пользователей, так и гетерогенность сети и количество интерактивных и мультимедийных приложений, то надо более строго выполнять требования производительности (к задержке и потере пакетов), чем в традиционной сети Интернет.

Функция хендовера (handover) – возможность перехода абонента из зоны действия одной базовой станции в зону действия другой без потери разговора или передачи мультимедийного трафика. Процесс хендовера может ввести дополнительную задержку из-за поиска сети, конфигурации, аутентификации и авторизации [9]. Чтобы уменьшать эту задержку, группа IEEE 802.21 предложила стандарт, который определяет функцию хендовера независимую от среды (MIHF, Media Independent Handover Function), что позволяет обеспечить абстрактные сервисы [10], которые уведомляют о возникших событиях, управляют поведением канала, связанного с хендовером и мобильностью, а также передают другим сервисам информацию относительно соседних сетей и их возможностей.

Стандарт IEEE 802.21 поддерживает алгоритмы, допускающие "безшовный" хендовер (seamless handover), обеспечивающий непрерывную, непрерывающуюся связь абонентам, как в сетях одного и того же самого типа, так и в гетерогенных сетях, содержащих фрагменты сетей различного типа. Стандарт обеспечивает информацию, позволяющую осуществлять передачу данных в сотовых сетях, GSM, UMTS, CDMA2000, GPRS, WiFi, Bluetooth, WPAN, IEEE 802.11 и IEEE 802.16 сетях, используя различные механизмы хендовера.

На рис.1 представлена обобщенная архитектура знание-ориентированной мобильной сети, представляющая собой некоторую интеллектуальную надстройку над существующей традиционной архитектурой. Главной целью новой архитектуры NGN является 1) обеспечение качества и разнообразия сервисов пользователям, 2) эффективное управление службами QoS, 3) обеспечения мобильности (хендовер и местоположение абонента) в сети и 4) служба защиты. Интеллектуальный узел управления сетью работает на основе управляющей информации, поступающей от узлов, абонентов и других компонент сети (используя традиционные протоколы SNMP, CMIP) в базу данных MIB, где она накапливается и хранится для дальнейшей обработки.

Блок индуктивного вывода анализирует эту статистическую информацию (некоторые отклонения в функционировании сети требующие управления) и извлекает из данных определенные закономерности – новые знания о правилах функционирования сети для того, чтобы затем принять решение о выполнении определенных управленческих действий, которые обеспечить бесперебойное и качественное функционирование сети. Для обобщения данных используется алгоритм, являющийся обобщением ID3 Куинлана для произвольного числа целевых значений.

Полученные закономерности преобразовываются, касающихся конкретной возникшей ситуации в сети и хранящиеся в онтологии (базе знаний). Эти правила позволяют использовать накопленный в процессе функционирования сети опыт (так, если ранее некоторое решение привело к удачному результату, то в дальнейшем это же решение будет применяться при сходных условиях без дополнительного анализа). Это позволяет также ускорить реакцию системы на изменения внешней среды. Знания, полученные от семантической базы данных, поступают на блок принятия решений узла, который формирует сигналы управления для компонент мобильной сети для исправления (улучшения, оптимизации) ее работы.

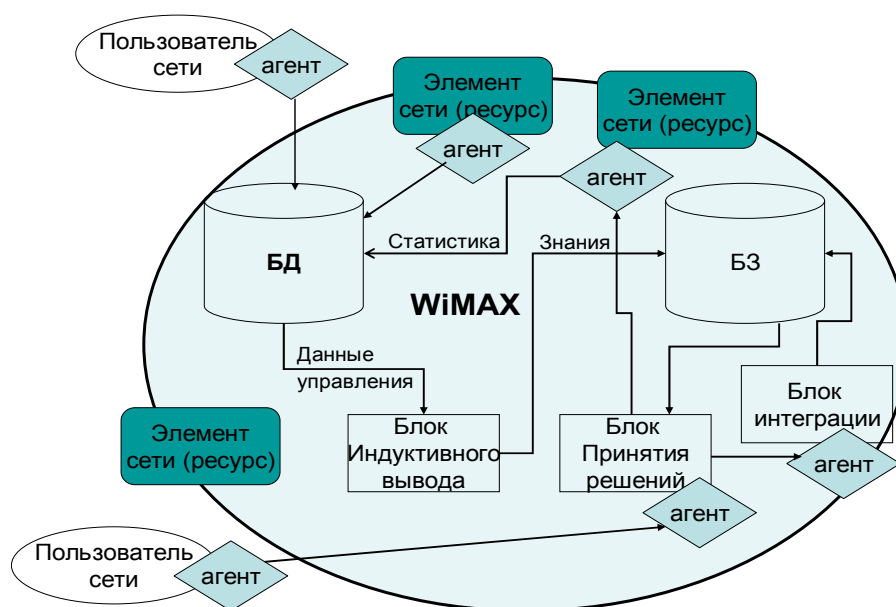


Рис.1 Архитектура знание-ориентированной мобильной сети

Архитектура интеллектуального узла управления сетью

Архитектура интеллектуального узла управления сетью (ИУУС), представленная на рис.2, позволяет ему извлекать информацию для поддержки функций QoS, мобильности и защиты в сети (для основных сетевых уровней согласно эталонной модели) и преобразует их в требуемые функциональные возможности. Все аспекты каждого сервиса представлены таким абстрактным способом, что техническая сложность их реализации скрыта от пользователя, поставщика сервиса и сетевого оператора.



Рис. 2 Архитектура узла, обеспечивающая сервисную функциональность сети

Многоуровневую архитектуру узла необходимо спроектировать таким образом, чтобы он был организован согласно семантической и контекстной управляющей информации, собранной из сети. Термин «семантика» касается изучения и изменения значения. Данные управления, собранные из компонентов сети, являются необработанными и не пригодны для того, чтобы быть представленными в качестве

семантической и контекстной информации. Однако, существует потребность доступа к этим необработанным данным для того, чтобы компилировать их, получать и интерпретировать на их основе результаты, которые будут затем интегрированы, чтобы сгенерировать семантическую и контекстную информацию.

Таким образом, новая архитектура узла должна разделять информацию по уровням протокола в общей семантической базе данных, и интерпретироваться интеллектуальным механизмом принятия решения (ИМПР), который будет обеспечивать требуемые сервисы в универсальном формате.

На рис. 3 показаны главные логические компоненты платформы предоставления сервисов и их взаимоотношения с семантической информационной базой данных (СИБД), упомянутой выше.

Платформа предоставления сервисов состоит из нескольких функциональных блоков, названных логическими объектами – интерпретаторами информации (QoS, безопасности, мобильности, предоставления сервисов), каждый из которых отвечает за один определенный аспект сетевой поддержки, которая должна обеспечить сервисы на прикладном уровне. Такой логический объект способен собирать информацию в сети на нескольких уровнях с помощью интеллектуальных агентов и использовать эту информацию совместно с другими локальными логическими объектами посредством промежуточного логического объекта – посредника, координирующего связь между логическими объектами.

Логические объекты также сохраняют информацию в их общей СИБД, которая будет использоваться логическими объектами других узлов в сети. Для удаленного логического объекта доступ к этой информации и формирование запроса к семантической базе данных другого узла должно осуществляться при помощи посредника – промежуточного логического объекта.

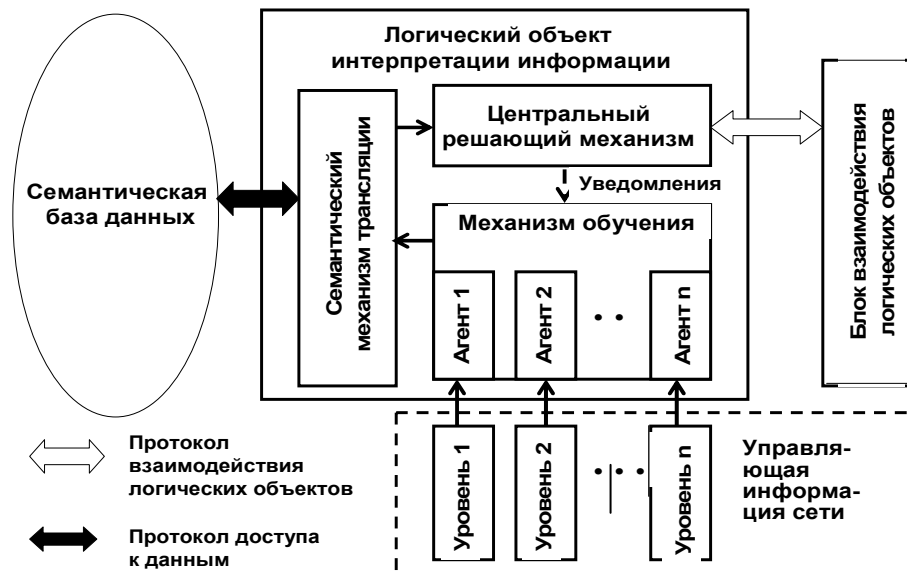


Рис.3 Главные компоненты логического объекта предоставления сервисов

Таким образом, два узла могут совместно использовать семантическую информацию только через посредников. Каждый логический объект – интерпретатор информации выполняет определенную функцию платформы предоставления сервиса в узле, такие, как QoS, мобильности (хендовер), защиты или других сетевых задач, используя релевантную внутреннюю и внешнюю информацию узла. Например, для *информационного элемента безопасности* – это состояние защиты некоторого туннеля IPSec, тогда

как текущая скорость передачи данных, местоположение и IPv6 – примеры *информационных элементов мобильности*.

Информационные элементы определены и сгруппированы согласно их релевантности для логических объектов интерпретации информации. Каждый объект ответственен за извлечение соответствующих ему *информационных элементов* в их форматах, за их интерпретацию и преобразование в общий семантический формат базы данных, а также сохранение их в базу данных. Позже *логические объекты интерпретации* информации используют интерпретируемые семантические версии локальных и удаленных информационных элементов в своих решениях.

Каждый *механизм обучения* содержит *агентов* обучения для тех уровней, на которых логический объект определил типы информации, которая будет извлекаться. Эти агенты обучения используют различные методы для извлечения управляющей информации. Они сканируют сетевые пакеты, читают переменные операционной системы, взаимодействуют с драйверами устройств и используют интерфейсы API, выполняющиеся на узле.

Семантический механизм трансляции конвертирует интерпретируемую информацию, полученную от механизма обучения, в семантическое представление. Он реализует протокол доступа к данным и сохраняет эту информацию в форме отчетов в семантическую базу данных. Он также читает ответы базы данных на запрос от центрального механизма принятия решений, который использует семантическую информацию и протокол между логическими объектами для того, чтобы обратиться к релевантным информационным элементам и выполнить это решение.

Центральный механизм принятия решения отвечает за эффективность сети в соответствии с целями уровня сервиса и уровня прикладных программ узла. Он также сохраняет целостность интерпретатора информации. Именно этот модуль делает выводы на основании результатов запросов к базе данных и состояния сети. Он также управляет запросами, основанными на этих заключениях, и уведомляет механизм обучения об изменениях. Этот модуль – ядро *интерпретатора информации*, который формирует замкнутую систему управления через обратную связь, поддерживая другие модули и гарантируя, что каждый из них работает адаптивно и самообучается, и своими действиями они не противоречат друг другу.

Архитектура узла расчленяет большую задачу на меньшие компоненты. Части этой задачи – семантическое представление запроса, его классификация и распознавание. Эти части разделены и реализованы интеллектуальными механизмами, составляющими внутреннюю структуру логического объекта интерпретации информации (рис. 3).

Семантическое коммуникационное пространство

Главная задача – создание семантического коммуникационного пространства Semantic Web в пределах нашей проектируемой сети. Целью концепции Semantic Web, в которой используются программные агенты, онтологии и Web-сервисы, является преобразование всей совокупности информационных ресурсов в единую базу знаний, пользоваться которой могут как люди, так и программы. Для этого необходимо снабдить каждый ресурс описанием его семантики и предоставить средства для автоматизированной обработки этих описаний и представления знаний о нем.

Использование технологий Semantic Web при разработке программных систем позволяет существенно упростить проблему совместимости систем из смежных областей и является первым шагом к построению высокоинтеллектуальных компонентов и агентов. Центральным компонентом концепции является применение онтологий. Онтологии разрабатываются и могут быть использованы при решении различных

задач, в том числе для совместного применения людьми или программными агентами, для возможности накопления и повторного использования знаний в предметной области, для создания моделей и программ, оперирующих онтологиями, а не жестко заданными структурами данных, для анализа знаний в предметной области.

Три технологии могут быть использованы для сетей NGN: агенты для того, чтобы представлять реальные объекты и автоматизированную разрешающую способность задачи от имени их владельца; онтологии для семантического расширения информации, которой обмениваются и обрабатывают Web-приложения; Web-сервисы как вычислительные средства, доступные через Интернет.

Целью Semantic Web является преобразование всей совокупности информационных ресурсов Web (или же сети, в нашем случае) в единую базу знаний, пользоваться которой могут как люди, так и программы. Для этого необходимо снабдить каждый ресурс описанием его семантики и предоставить средства для автоматизированной обработки этих описаний и представления знаний о них. Автором этой концепции является Т.Бернес-Ли, который ранее задумал и разработал Web. Несмотря на нечеткость концепции Semantic Web, несколько основных направлений повторяются во всех ее определениях:

- описание семантики ресурсов при помощи метаданных;
- использование метаданных для поиска информационных ресурсов на семантическом уровне;
- представление Web как большой интероперабельной базы данных;
- обнаружение, компоновка и использование Web-сервисов на основе описания их семантики;
- предоставление информации интеллектуальным программным агентам в форме, пригодной для машинной обработки их значения.

Нередко онтология используется как посредник между пользователем и информационной системой. Онтология может формализовать договоренности о терминологии между членами сообщества, например, между пользователями некоторого хранилища данных сети.

Онтология – это набор определений (на формальном языке) фрагмента декларативных знаний, ориентированный на совместное многократное использование различными пользователями в своих приложениях. В онтологии вводятся термины, типы и соотношения (аксиомы), описывающие фрагмент знания.

Формальная модель онтологии O представляет собой упорядоченную тройку $O = \langle X, R, F \rangle$, где X – конечное множество концептов (понятий, терминов) предметной области, которую представляет онтология O ; R – конечное множество отношений между концептами заданной предметной области; F – конечное множество функций интерпретации, заданных на концептах и отношениях онтологии O [11-13].

Онтология позволяет формализовать семантику произвольной предметной области (PrO). Можно сказать, что онтология – это точная спецификация некоторой области, которая включает в себя словарь терминов этой области и множество связей (типа «элемент-класс», «часть –целое», «оказывает воздействие», «похоже на» и т.п.), которые описывают, как эти термины соотносятся между собой. Онтологии позволяют представить понятия так, что они становятся пригодными для машинной обработки.

Для описания метаданных в семантической базе данных использован язык RDF (Resource Description Framework), который использует XML-синтаксис. RDF описывает ресурсы в виде ориентированного размеченного графа – каждый ресурс может иметь свойства, которые в свою очередь также могут быть ресурсами или их коллекциями.

Чтобы упростить и унифицировать создание метаописаний ресурсов, пользователям нужно предоставить определенные шаблоны и стандарты описания типичных ресурсов. Из таких средств наиболее

основательно проработан набор элементов для создания метаданных "Dublin Core Metadata Elements", который состоит из 15 базовых элементов [14].

Обработка контекста

Контекст (от лат. contextus – «соединение», «связь») – относительно законченный фрагмент записи, текста, общий смысл которого позволяет уточнить значение отдельных входящих в него слов, предложений, и т. п. В более широком значении контекст – среда, в которой существует объект. Архитектура интеллектуального узла мобильной сети снабжена средствами обработки контекста для того, чтобы повысить качество принятия решений, сделанных интерпретаторами информации. Контекст зависит от пользовательских предпочтений, текущей ситуации и возможностей мобильных устройств [15].

Обработки контекста – это способность интеллектуального устройства обеспечить своего пользователя наиболее релевантными представлением приобретенного сервиса. Она требует дополнительной информации, которая собрана либо заранее перед тем, как выбран сервис, либо реактивно после развития ситуации. Обработка контекста начинается с определения устройством собственных свойств, возможностей и ограничений.

Интеллектуальное устройство способно адаптироваться таким образом, чтобы представить пользователю сервис в правильном формате и при приемлемых условиях. На более абстрактном уровне, контекст может быть определен как набор обстоятельств и фактов, которые соответствуют конкретной ситуации. В NGN понятие контекста расширяется на компьютерную связь, в которую вовлекаются и люди, и компьютеры, и где каждый узел в сети – производитель и потребитель контекста (подобно концепции Web 2.0). Каждый узел производит свое подмножество контекста, доступное другим узлам через семантическую базу данных и протокол посредника.

Каждый узел также использует эту информацию, так как он формирует локальную базу знаний о других узлах, находящихся в прямой или косвенной связи с ним. Такой подход необходим, чтобы более эффективно использовать каналы связи посредством понимания контекста и выполнять локальные решения, используя глобальные вводы.

Предложенная архитектура узла позволяет собирать и использовать локальную и удаленную контекстную информацию. Данные о местоположении – типичный пример для информации, которая может использоваться для обработки контекста. Доступность данных местоположения дает мобильному устройству дополнительную информацию, которая может использоваться сервисом.

Другая важная часть информации – окружение устройства, которое определяется персональной локальной сетью пользователя (PAN). Пример PAN – ноутбук с подключенным к нему мобильным телефоном, КПК и Web-камерой. Персональное интеллектуальное оборудование пользователя может быть обнаружено соответствующей технологией связи типа Bluetooth. Окружение может также быть расширено за счет других публично доступных устройств, находящихся в поле зрения пользователя.

Новые знания могут быть получены разными способами. Можно вывести их как логическое следствие из уже существующих знаний, вывести общее правило из имеющихся фактов или перенести факты и знания, истинные для одних объектов, на другие объекты на основании их сходства. Интеллектуальность поведения системы связана с ее способностью учиться на собственном опыте, то есть обобщать известные системе факты в некоторые общие правила.

Для извлечения знаний из семантической базы данных используется *метод индуктивного вывода* (от частного к общему). Индуктивный вывод – вывод из имеющихся данных (наблюдений, фактов) общих закономерностей (правил), что их объясняет. Основное назначение индуктивного вывода – генерация

гипотез, которые могут потом опровергаться или подтверждаться пользователем. Полученная гипотеза используется для объяснения для объяснения имеющихся данных, классификации и прогнозирования новых данных.

Индуктивная логика – формальная система, которая описывает правила формирования общих утверждений на основе конечного множества отдельных утверждений. Интеллектуальность поведения любой системы в значительной мере определяется ее способностью учиться на собственном опыте, то есть индуктивно обобщать известные системе факты в некоторые общие правила. Индуктивное обобщение состоит в том, чтобы за набором примеров построить для функции f функцию-гипотезу h , что аппроксимирует f .

В статье предлагается использовать методы индуктивного обобщения [16] для автоматизированного извлечения онтологических знаний о ПрО из набора информационных ресурсов (хранящихся в БД системы), релевантных этой ПрО.

Практическая реализация сервиса

Рассмотрим, как предложенная архитектура может использоваться пользователем, поставщиком сервиса и сетевым оператором. Этот подход увеличивает доступность сервисов в сети. Предположим, пользователь хочет получить качественный доступ к мультимедийному сервису через мобильный терминал (МТ) во время своего передвижения. Он начинает сеанс, используя соответствующий протокол сеансового уровня. Провайдер сервиса (service provider node, SPN) способен предоставить мультимедийный сервис в различных форматах, каждый с различными требованиями полосы пропускания. Полоса пропускания – ценный ресурс для провайдера сервиса, так как он получен в аренду от сетевого оператора, и цель состоит в том, чтобы перепродать эту полосу пропускания эффективным способом для многих пользователей так, насколько это возможно. Чтобы выбрать самый эффективный формат, логический объект предоставления сервисов (рис.2) запрашивает у логического объекта QoS качество доступной полосы пропускания.

Логический объект QoS входит в контакт со своим одноранговым узлом сетевого оператора (network operator node, NON) посредством протокола посредника, чтобы запросить возможности МТ (мобильного терминала) и AP (точек доступа), которые расположены в пределах зоны передачи пользователя. Когда одноранговый логический объект QoS получает этот запрос, он обращается к своей семантической базе данных, где хранится профиль мобильности пользователя, который логически вычислен из местоположения и скорости МТ, так же как и местоположение точки доступа AP через которую вошел пользователь.

Эта информация и полоса пропускания, которая является доступной посредством этих AP (точек доступа), дают результирующий диапазон полосы пропускания, с которым пользователь может поддерживаться во время мультимедиа сеанса, и который возвращается к провайдеру сервиса SPN как профиль мобильности QoS пользователя. После того как логический объект QoS получил требуемый профиль пользователя, провайдер сервиса принимает интеллектуальное решение относительно формирования полосы пропускания, которая должна быть выделена пользователю, и которая не должна превышать возможности, описанные в профиле мобильности QoS и сообщает затем логическому объекту предоставления сервиса, что тот может выбрать соответствующий формат.

Логический объект мобильности провайдера сервиса также регистрирует себя в своем одноранговом узле сетевого оператора с целью быть информированным относительно изменений в семантической базе данных касающихся состояния мобильности МТ (передвижения абонента) таким образом, чтобы провайдер сервиса мог адаптировать мультимедийный сервис к изменяющимся условиям связи.

Логический объект QoS провайдера сервиса сообщает своему одноранговому узлу сетевого оператора о распределенной полосе пропускания и также регистрируется в нем для возможности учета изменений в состояниях QoS так, чтобы мультимедийный сервис мог модифицироваться в соответствии с изменениям в сетевой нагрузке, управляемой логическим объектом QoS. На основе предложенной структуры может быть реализован межуровневый распределенный сервисный доступ и схема адаптации как ответ на мобильность пользователя и динамику сетевых условий без потребности в определении отдельных протоколов для каждого интерфейса. Такой сценарий гарантирует, что пользователь получает мультимедийный сервис при перемещении, в то время как провайдер сервиса и владелец сети имеют самое эффективное использование своих соответствующих ресурсов.

Заключение

В работе представлена концепция, которая определяет расширенную архитектуру, позволяющую интегрировано реализовать основные требования к NGN на основе использования знаний и средств управления знаниями.

Для этого необходим интеллектуальный подход, позволяющий конвертировать необработанные данные управления в семантическое представление, принимать решения основанные на семантике и учиться на опыте. Важно применять для этого технологии Semantic Web и методы, используемые в задачах искусственного интеллекта.

Таким образом, новые приложения и сервисы могут также быть реализованы намного проще чем сегодня, и ожидается, что они будут поддерживаться в будущем.

Благодарности

Статья частично финансированна из проекта **ITHEA XXI** Института Информационных теории и Приложений FOI ITHEA и Консорциума FOI Bulgaria (www.ithea.org, www.foibg.com).

Литература

1. Chakrabarti, S. and A. Mishra, "QoS Issues in Ad Hoc Wireless Networks", IEEE Communications Magazine, Vol. 39, No. 2, pp. 142-148, February 2001.
2. Zhang, Q., W. Zhu, G.J. Wang and Y.Q. Zhang, "Resource Allocation with Adaptive QoS for Multimedia Transmission over W-CDMA Channels", Proceedings of WCNC 2000, pp. 179-184, Chicago, USA, 23-28 September 2000.
3. Cavalcanti, D., D. Agrawal, C. Cordeiro, B. Xie and A. Kumar, "Issues in Integrating Cellular Networks, WLANs, and MANETs: A Futuristic Heterogeneous Wireless Network", IEEE Wireless Communications, Vol. 12, No. 3, pp. 30-41, June 2005
4. Капитонова Ю.В. Общие принципы построения знание-компьютерных систем // Кибернетика и систем. анализ. — 2006. — 42, N 4. — С. 81-101.
5. Валькман Ю.П. Распределенный искусственный интеллект и многоагентные системы в исследовательском проектировании сложных объектов// Труды Международного семинара DAIMAS'97 «Распределенный искусственный интеллект и многоагентные системы», Санкт-Петербург, 1997. — С.269-279.
6. Анисимов А.В., Романик А.Н, Тарануха В.Ю. Эвристические алгоритмы для определения канонических форм и грамматических характеристик слов // Кибернетика и систем. анализ. — 2004. — 40, N 2. — С. 3-14.
7. International Standard ITU-T Y.2001(12/2004). General overview of NGN.
8. International Standard ITU-T G.902 (11/95). Framework Recommendation on functional access networks (AN) – Architecture and functions, access types, management and service node aspects.

9. Гладун А.Я., Свиридов Е.А. Сети нового поколения и технология Softswitch // Корпоративные системы, 2008, №3.- С.41-49.
10. Dutta, A., S. Das, D. Famolari, Y. Ohba, K. Taniuchi, T. Kodama and H. Schulzrinne, "Seamless Handover across Heterogeneous Networks – An IEEE 802.21 Centric Approach", Proceedings of IEEE WPMC'05, Aalborg, Denmark, 18-22 September 2005.
11. Anatoly Gladun, Julia Rogushina, Francisco Garcí'a-Sanchez, Rodrigo Martínez-Béjar and Jesualdo Tomás Fernández-Breis An Application of Intelligent Techniques and Semantic Web Technologies in e-Learning Environments// Journal of Expert Systems with Applications, ELSEVIER, New-York, USA, v.23, 2008.-P.72-83.
12. Гладун А.Я., Проценко О.І. Аналіз моделей процедур хендвера в бездротових мережах, базованих на технологіях ATM \\\ Наукові записки, сер. Комп'ютерні науки, К.:НаУКМА, том 21, 2003.-- С.85-95.
13. Рогушина Ю.В., Гладун А.Я. Онтологический подход к поиску Web-сервисов в распределенной среде Интернет// Информатика, №5, 2006, Минск, Беларусь.-- С.134-146.
14. Рогушина Ю.В., Гладун А.Я. Мереологические аспекты онтологического анализа интеллектуальных Web-сервисов // Збірник праць VII-ої Міжнародної конференції „Інтелектуальний аналіз інформації” IAI-2007 Київ.— 2007.-- С.312-321.
15. Рогушина Ю.В., Гладун А.Я. Формирование и применение онтологий предметных областей для поиска Web-сервисов на семантическом уровне. // Труды Межд. конф. „Знания-Онтологии-Теория” ЗОНТ-2007, т.2, РАН И-тут математики им. С.П.Соболева, Новосибирск, Россия, 2007.--С.177-186.
16. Anatoly Gladun, Julia Rogushina, and Victor Shtonda Ontological Approach to Domain Knowledge Representation for Information Retrieval in Multiagent Systems // International Journal "Information Theories and Applications", V.12, 2006.-P.72-84.

Информация об авторах

Гладун Анатолий Ясонович – *Международный научно-учебный центр информационных технологий и систем НАНУ, Киев-187 ГСП, 03680, просп. акад. Глушкова, 40, email: glanat@yahoo.com*

Рогушина Юлия Витальевна – *Институт программных систем НАНУ, Киев-187 ГСП, 03680, просп. акад. Глушкова, 40, Киевский Славистический Университет, email: jjj_@ukr.net*

АГРЕГИРОВАНИЕ ПРЯМЫХ И ДВОЙСТВЕННЫХ БАЛАНСОВЫХ МОДЕЛЕЙ «ЗАТРАТЫ-ВЫПУСК»

Игорь Ляшенко, Андрий Онищенко, Игорь Онищенко

Аннотация: Исследование экономических процессов на макро уровне используют агрегированные показатели. Детализированные показатели определяются на микро уровне. Связь между детализированными и агрегированными показателями осуществляется с помощью методов агрегирования. Авторы основное внимание уделяют методу точного агрегирования и возможностям его использования для статичных и динамичных межотраслевых моделей, исследование которых позволяют сбалансировать производственные мощности, выпуск продукции и расходы на производство.

Ключевые слова. Экономико-математическое моделирование, точное агрегирование, модель Леонтьева.

Conference: The paper is selected from XVth International Conference "Knowledge-Dialogue-Solution" KDS 2009, Varna, Bulgaria, June-July 2009

Введение.

Изучение экономических показателей требует от исследователя оперировать большим количеством данных, которые в силу объективных причин не могут быть учтены в полном объеме. Поэтому исследователь не может эффективно их обработать без использования надежных методов агрегирования входящей информации.

II. Постановка задачи.

Одной из первых линейных балансовых моделей построенных на основании детализированных показателей является межотраслевая балансовая модель Леонтьева, известная под названием «затраты-выпуск» [1,2]:

$$x = Ax + y, \quad x \geq 0, \quad y > 0 \quad (1.1)$$

где $x = (x_1, x_2, \dots, x_N)^T$ – вектор-столбец объемов выпуска продукции, $y = (y_1, y_2, \dots, y_N)^T$ – вектор-столбец конечного потребления продукции, $A \geq 0$ – неотрицательная матрица коэффициентов прямых затрат производства, размерности $N \times N$, элемент a_{ij} которой показывает, какое количество продукции i необходимо затратить для производства единицы продукции j .

Модель «затраты-выпуск» строится как в натуральном, так и в стоимостном выражении. Во втором случае натуральные элементы модели заменяют их денежным эквивалентом. В данной работе будем рассматривать материальный баланс.

III. Результаты.

1. Рассмотрим задачу агрегирования межотраслевого баланса (1.1) в единый продукт. Пусть $T = (p_1, p_2, \dots, p_N)$ – агрегирующий вектор, где p_i – цена i -го продукта. Умножим обе части уравнения (1.1) на вектор T слева, в результате получим

$$Tx = TAx + Ty.$$

Учитывая, что $X = Tx$ – агрегированный показатель валового выпуска продукции в денежном выражении, $Y = Ty$ – агрегированный показатель конечного потребления продукции в денежном выражении, можем записать уравнение агрегированного баланса в виде

$$X = TAx + Y. \quad (1.2)$$

Или в скалярном виде

$$X = \tilde{a}X + Y,$$

где \tilde{a} – числовой коэффициент прямых затрат производства.

Последнее уравнение можно записать также в виде

$$X = \tilde{a}Tx + Y. \quad (1.3)$$

Из уравнений (1.2) и (1.3) следует равенство

$$TA = \tilde{a}T. \quad (1.4)$$

Равенство (1.4) – условие точного агрегирования, известное под названием условия Хатанаки [3]. Раскроем экономический смысл этого условия. Уравнение (1.4) запишем в виде

$$T(A - \tilde{a}E) = 0, \quad (1.5)$$

где E – диагональная единичная матрица N -го порядка.

На элементы уравнения (1.5) накладываем следующие дополнительные условия:

А) $T > 0$ – следует из того, что T – вектор цен.

Б) $A \neq \tilde{a}E$ – иначе технологическая матрица будет иметь вид диагональной, что сильно сужает круг исследуемых задач.

При таких условиях и нераскладной продуктивной матрице A решениями уравнения (1.5) будут [4]:

$\tilde{a} = \lambda_A$ – корень Фробениуса матрицы A ,

$T = p_A$ – левый вектор Фробениуса матрицы A .

Подводя итоги сказанного можно утверждать

Утверждение 1: Для продуктивной нераскладной модели «затраты-выпуск» $x = Ax + y$, $y > 0$ всегда существует единственное точное агрегирование к одномерному уравнению $X = \tilde{a}X + Y$, $Y > 0$, в котором $\tilde{a} = \lambda_A < 1$ – корень Фробениуса, а $T = p_A$ – левый вектор Фробениуса матрицы A .

2. Рассмотрим задачу агрегирования межотраслевого баланса из N отраслей в n , причем $1 < n < N$. Как и раньше, будем изучать модель Леонтьева в виде (1.1).

Так как мы используем показатели матрицы A в натуральной форме, то для агрегирования модели (1.1) будем считать матрицу T матрицей цен:

$$T = \begin{pmatrix} p_1^1 \dots p_{N_1}^1 & 0 \dots 0 & \dots & 0 \dots 0 \\ 0 \dots 0 & p_1^2 \dots p_{N_2}^2 & \dots & 0 \dots 0 \\ \dots & \dots & \dots & \dots \\ 0 \dots 0 & 0 \dots 0 & \dots & p_1^n \dots p_{N_n}^n \end{pmatrix}, \quad (2.1)$$

$$N_1 + N_2 + \dots + N_n = N,$$

Размерность агрегирующей матрицы – $n \times N$, элемент p_j^s матрицы соответствует цене j -го продукта из множества $\{N_s\}$, где $\bigcup_{s=1}^n \{N_s\} = \{N\}$, $\{N_k\} \cap \{N_r\} = \emptyset$ при $k \neq r$.

Матрицы A и T соответственно с (2.1) запишем в клеточном виде

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \dots & \dots & \dots & \dots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{pmatrix}, T = \begin{pmatrix} p^1 & 0 & \dots & 0 \\ 0 & p^2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & p^n \end{pmatrix}.$$

Умножим уравнение (1.1) на T слева, получим

$$Tx = TAx + Ty.$$

Тут $Tx = X$, где $X = (X_1, X_2, \dots, X_n)^T$ – вектор-столбец агрегированных валовых выпусков продукции; $Ty = Y$, где $Y = (Y_1, Y_2, \dots, Y_n)^T$ – вектор-столбец агрегированных выпусков конечной продукции. Можем переписать последнее равенство в виде:

$$X = TAx + Y. \quad (2.2)$$

С другой стороны уравнение агрегированного баланса имеет вид:

$$X = \tilde{A}Tx + Y. \quad (2.3)$$

Сравнивая равенства (2.2) и (2.3), приходим к условию Хатанаки:

$$TA = \tilde{A}T,$$

где T – агрегирующая матрица; A, \tilde{A} – соответственно матрицы коэффициентов прямых материальных затрат детализированного и агрегированного балансов. Причем матрица \tilde{A} имеет вид

$$\tilde{A} = \begin{pmatrix} \tilde{a}_{11} & \tilde{a}_{12} & \dots & \tilde{a}_{1n} \\ \tilde{a}_{21} & \tilde{a}_{22} & \dots & \tilde{a}_{2n} \\ \dots & \dots & \dots & \dots \\ \tilde{a}_{n1} & \tilde{a}_{n2} & \dots & \tilde{a}_{nn} \end{pmatrix}.$$

Условие точного агрегирования можем также записать в виде

$$(p^s A_{sk})_1 : (p^s A_{sk})_2 : \dots : (p^s A_{sk})_{N_k} = p_1^k : p_2^k : \dots : p_{N_k}^k, \quad (2.4)$$

$$s = 1, 2, \dots, n; k = 1, 2, \dots, n.$$

Из полученных выше результатов можно сделать следующий вывод:

Утверждение 2: *Необходимым и достаточным условием существования точного агрегирования модели межотраслевого баланса из N - размерного к n - размерному уравнению ($1 < n < N$) есть требование, чтобы во всех n частичных агрегированиях удельные материальные расходы отраслей были пропорциональными ценам продукции.*

3. Если в задаче возникает необходимость исследовать как производственные затраты, так и затраты на расширение производства, то в таком случае используют динамическую модель Леонтьева.

$$x = Ax + B\dot{x} + y, \quad (3.1)$$

где A – матрица затрат на производство продукции, B – матрица затрат на расширение производства, \dot{x} – вектор приростов производства за счет создания новых мощностей.

Поставим задачу агрегирования динамической модели Леонтьева (3.1) в единый продукт. Для решения поставленной задачи выберем агрегирующий вектор $T = (p_1, p_2, \dots, p_N)$, элемент p_i которого равен цене i -го продукта, $i = \overline{1, N}$. Умножим обе части уравнения (3.1) на вектор T слева. Получим

$$Tx = TAx + TB\dot{x} + Ty. \quad (3.2)$$

С другой стороны, агрегированная динамическая модель Леонтьева имеет вид

$$Tx = \tilde{a}Tx + \tilde{b}T\dot{x} + Ty. \quad (3.3)$$

Из уравнений (3.2) и (3.3) следуют условия

$$\begin{aligned} TA &= \tilde{a}T, \\ TB &= \tilde{b}T, \end{aligned} \quad (3.4)$$

которые есть обобщенными условиями Хатанаки. Условия (3.4) могут быть записаны в виде

$$\begin{aligned} T(A - \tilde{a}E) &= 0, \\ T(B - \tilde{b}E) &= 0. \end{aligned}$$

Если рассматривать общий случай, когда матрицы A и B не диагональные или нулевые, то из последних уравнений получим следующие результаты. Число \tilde{a} – корень Фробениуса матрицы A , число \tilde{b} – корень Фробениуса матрицы B . Агрегирующий вектор $T = p_A = p_B$, где p_A и p_B – левые вектора Фробениуса матриц A и B соответственно. Подводя итоги, можем утверждать:

Утверждение 3: *Необходимым и достаточным условием существования точного агрегирования для продуктивной нераскладной динамической модели Леонтьева $x = Ax + B\dot{x} + y$, $y > 0$ к одномерному уравнению $X = \tilde{a}X + \tilde{b}\dot{X} + Y$, $Y > 0$, где $\tilde{a} = \lambda_A < 1$, $\tilde{b} = \lambda_B$ – корни Фробениуса матриц A и B , есть условие равенности их левых векторов Фробениуса. Тогда агрегирующий вектор $T = p_A = p_B$.*

4. Перейдем к задаче агрегирования динамической модели Леонтьева (3.1) из N -размерной к n -размерной. Причем выполняется условие $1 < n < N$.

В таком случае агрегирующая матрица T имеет размерность $n \times N$

$$T = \begin{pmatrix} p_1^1 & \dots & p_{N_1}^1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & \dots & 0 & p_1^2 & \dots & p_{N_2}^2 & \dots & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 & \dots & p_1^n & \dots & p_{N_n}^n \end{pmatrix}, \quad (4.1)$$

$$N_1 + N_2 + \dots + N_n = N,$$

элемент p_j^s матрицы соответствует цене j -го продукта из множества $\{N_s\}$, где $\bigcup_{s=1}^n \{N_s\} = \{N\}$,

$$\{N_k\} \cap \{N_r\} = \emptyset \text{ при } k \neq r.$$

Матрицы затрат A и B , агрегирующую матрицу T в соответствии с (4.1) запишем в клеточном виде

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \dots & \dots & \dots & \dots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{pmatrix}, B = \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1n} \\ B_{21} & B_{22} & \dots & B_{2n} \\ \dots & \dots & \dots & \dots \\ B_{n1} & B_{n2} & \dots & B_{nn} \end{pmatrix}, T = \begin{pmatrix} p^1 & 0 & \dots & 0 \\ 0 & p^2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & p^n \end{pmatrix}.$$

Для агрегирования динамической модели Леонтьева умножим уравнение (3.1) на матрицу T слева. Получим

$$Tx = TAx + TB\dot{x} + Ty. \quad (4.2)$$

Также уравнение агрегированного баланса можно записать в виде

$$Tx = \tilde{A}Tx + \tilde{B}T\dot{x} + Ty. \quad (4.3)$$

Из последних двух равенств приходим к условиям точного агрегирования

$$\begin{aligned} TA &= \tilde{A}T, \\ TB &= \tilde{B}T. \end{aligned} \quad (4.4)$$

Причем агрегированные матрицы затрат \tilde{A} и \tilde{B} имеют вид

$$\tilde{A} = \begin{pmatrix} \tilde{a}_{11} & \tilde{a}_{12} & \dots & \tilde{a}_{1n} \\ \tilde{a}_{21} & \tilde{a}_{22} & \dots & \tilde{a}_{2n} \\ \dots & \dots & \dots & \dots \\ \tilde{a}_{n1} & \tilde{a}_{n2} & \dots & \tilde{a}_{nn} \end{pmatrix}, \tilde{B} = \begin{pmatrix} \tilde{b}_{11} & \tilde{b}_{12} & \dots & \tilde{b}_{1n} \\ \tilde{b}_{21} & \tilde{b}_{22} & \dots & \tilde{b}_{2n} \\ \dots & \dots & \dots & \dots \\ \tilde{b}_{n1} & \tilde{b}_{n2} & \dots & \tilde{b}_{nn} \end{pmatrix}.$$

Учитывая условие положительности всех векторов цен, $p^s > 0$, $s = \overline{1, n}$, приходим к выводу, что

$$\tilde{a}_{ss} = \lambda_{A_{ss}}, p^s = p_{A_{ss}}, s = \overline{1, n}.$$

Аналогично для второго уравнения системы (4.4) получаем

$$\tilde{b}_{ss} = \lambda_{B_{ss}}, p^s = p_{B_{ss}}, s = \overline{1, n}.$$

Условие точного агрегирования можем переписать в виде

$$(p^s A_{sk})_1 : (p^s A_{sk})_2 : \dots : (p^s A_{sk})_{N_k} = p_1^k : p_2^k : \dots : p_{N_k}^k. \quad (4.5)$$

Аналогичное условие для второго уравнения системы (4.4), запишется в виде

$$\begin{aligned} (p^s B_{sk})_1 : (p^s B_{sk})_2 : \dots : (p^s B_{sk})_{N_k} &= p_1^k : p_2^k : \dots : p_{N_k}^k, \\ s &= 1, 2, \dots, n; k = 1, 2, \dots, n. \end{aligned} \quad (4.6)$$

Подводя итоги результатов полученных в уравнениях (4.5), (4.6), можно сделать вывод

Утверждение 4: *Необходимым и достаточным условием существования точного агрегирования динамической модели Леонтьева из N к n - мерному уравнению ($1 < n < N$) есть требование, чтобы во всех n частичных агрегированиях удельные материальные затраты отраслей на производство были пропорциональными удельным затратам на расширение производства и цене соответственного продукта.*

5. Рассмотрим модель двойственную к модели межотраслевого баланса. Которую запишем в виде

$$p = pA + r, \quad (5.1)$$

где p – вектор цен на продукцию, A – технологическая матрица модели Леонтьева, r – вектор добавленной стоимости.

Пусть для модели (5.1) нужно решить задачу агрегирования в единый продукт. Для этого введем агрегирующий вектор-столбик $T = (x_1, x_2, \dots, x_N)^T$, где x_i – объемы выпусков i -й продукции. Умножим уравнение (5.1) на агрегирующий вектор T справа. Получим

$$pT = pAT + rT. \quad (5.2)$$

Учитывая, что $pT = P$, $rT = R$ – агрегированные показатели, уравнение (5.2) можно записать в виде

$$P = pT\tilde{a} + R. \quad (5.3)$$

Сравнивая равенства (5.2) и (5.3) получим условие точного агрегирования для двойственной модели

$$AT = T\tilde{a}. \quad (5.4)$$

Можем записать последнее условие так

$$(A - \tilde{a}E)T = 0. \quad (5.5)$$

Для нетривиального случая матрица A – не диагональная или нулевая. Поэтому выполнение равенства (5.5) требует чтобы $\tilde{a} = \lambda_A$, то есть агрегированный показатель \tilde{a} равен корню Фробениуса матрицы A , а $T = x_A$, где x_A – правый вектор Фробениуса технологической матрицы A . Полученные результаты сформулируем в форме утверждения.

Утверждение 5: Для продуктивной нераскладной модели (5.1) всегда существует единственное точное агрегирование к одномерному уравнению $P = P\tilde{a} + R$, $R > 0$. Причем $\tilde{a} = \lambda_A$, $T = x_A$, где λ_A и x_A – соответственно корень Фробениуса и правый вектор Фробениуса матрицы A .

6. Рассмотрим двойственную модель к динамической модели Леонтьева

$$p = pA + \dot{p}B + r, \quad (6.1)$$

где A , B – технологические матрицы, r – вектор добавленной стоимости, $\dot{p}B$ – стоимость инфляционных процессов.

Введем агрегирующий вектор-столбик $T = x = (x_1, x_2, \dots, x_N)^T$, где x_i – объем валового выпуска i -го продукту.

Умножим левую и правую часть уравнения (6.1) на вектор-столбик $T = x$ справа. Получим следующее равенство

$$px = pAx + \dot{p}Bx + rx. \quad (6.2)$$

С другой стороны, соответствующий агрегированный динамический баланс имеет вид

$$px = px\tilde{a} + \dot{p}x\tilde{b} + rx. \quad (6.3)$$

Уравнения (6.2) та (6.3) дают нам условия точного агрегирования модели (6.1)

$$\begin{aligned} Ax &= x\tilde{a}, \\ Bx &= x\tilde{b}. \end{aligned} \quad (6.4)$$

или

$$\begin{aligned} (A - \tilde{a}E)x &= 0, \\ (B - \tilde{b}E)x &= 0. \end{aligned}$$

Если рассматривать общий случай, когда матрицы A и B не диагональные или нулевые, то из последних равенств приходим к таким результатам. Число \tilde{a} – корень Фробениуса матрицы A , число \tilde{b} – корень Фробениуса матрицы B . При этом агрегирующий вектор-столбик $T = x_A = x_B$, где x_A и x_B – правые векторы Фробениуса для матриц A и B соответственно.

Утверждение 6: *Необходимым и достаточным условием существования точного агрегирования для продуктивной нераскладной модели динамического баланса цен к одномерному уравнению $P = \tilde{a}P + \tilde{b}\dot{P} + R$, $R > 0$, где $\tilde{a} = \lambda_A < 1$, $\tilde{b} = \lambda_B < 1$ – корни Фробениуса, есть условие $x_A = x_B$ – равенства правых векторов Фробениуса матриц A и B . Тогда агрегирующий вектор $T = x_A = x_B$.*

IV. Выводы.

Сформулированы условия точного агрегирования модели Леонтьева «затраты-выпуск» к одномерному уравнению. Полученные результаты обобщены для случая агрегирования модели «затраты-выпуск» из N в n отраслей, а также для агрегирования динамической модели Леонтьева в один продукт и из N в n продуктов, где $1 < n < N$. Получены условия агрегирования моделей двойственных к модели «затраты-выпуск».

Благодарности

Статья частично финансированна из проекта **ITHEA XXI** Института Информационных теории и Приложений FOI ITHEA и Консорциума FOI Bulgaria (www.ithea.org, www.foibg.com).

Литература

- [1] Леонтьев В.В. Экономические эссе. Теории, исследования, факты и политика: Пер. с англ./Под ред. С.С. Шаталина, Д.В. Волового. – М.: Политическая литература, 1990. – с. 415
- [2] Леонтьев В.В. Межотраслевая экономика. – М.: ОАО «Издательство «Экономика», 1997. – с. 479.
- [3] Итеративное агрегирование и его применение в планировании / Под ред. Л.М. Дудкина. – М.: Экономика, 1979. – с. 328
- [4] Пономаренко О.І., Перестюк М.О., Бурим В.М. Основи математичної економіки. – К.: Інформтехніка. 1995. – с. 320.

Информация об авторах

Игорь Ляшенко – д-р физ.-мат. наук, заслуженный профессор Киевского национального университета имени Тараса Шевченко

Андрій Онищенко – канд. екон. наук, доцент, докторант Киевского национального университета имени Тараса Шевченко; e-mail: onyshchenko@yandex.ru

Игорь Онищенко – аспирант Киевского национального университета имени Тараса Шевченко; e-mail: iony@yandex.ru

ВЛИЯНИЕ ИНФОРМАЦИОННОЙ УСТОЙЧИВОСТИ НА ИНФОРМАЦИОННЫЕ ИННОВАЦИИ В ФИРМАХ

Александр Фролов

Abstract: В статье вводится понятие информационной устойчивости (инфоустойчивости) фирмы. Предлагается использовать показатель инфоустойчивости в управлении фирмой. Проанализированы понятия комплексной информационной системы (КИС) фирмы и информационных инноваций (инфоинноваций). На основе теории перестроек устанавливается связь между инфоустойчивостью КИС и инфоинновациями, делаются выводы о характере этой связи.

Keywords: фирма, комплексная информационная система, информационная устойчивость, информационные инновации.

ACM Classification Keywords: H.1.1 Information Systems: Models and Principles – Systems and Information Theory.

Conference: The paper is selected from Seventh International Conference on Information Research and Applications – i.Tech 2009, Varna, Bulgaria, June-July 2009

Введение

Собственный опыт по внедрению инноваций в информационные системы предприятий, опыт коллег, публикации в литературе говорят о том, что это – крайне непростая задача с мало предсказуемым результатом. В этой статье мы попытаемся сделать выводы о причинах указанной сложности. Сначала выведем причинную связь между асимметриями информации и появлением и существованием фирм. Рассматривая фирму с точки зрения информации, мы, фактически отождествляя фирму с её комплексной информационной системой (КИС), выявляем зависимость конкурентоспособности фирмы от эффективности КИС, информационной устойчивости (инфоустойчивости) и информационных инноваций (инфоинноваций). Модель связи указанных понятий показана на рис. 1. Интерес в этой работе представляет влияние инфоустойчивости на инфоинновации. Для его исследования используются выводы теории перестроек.

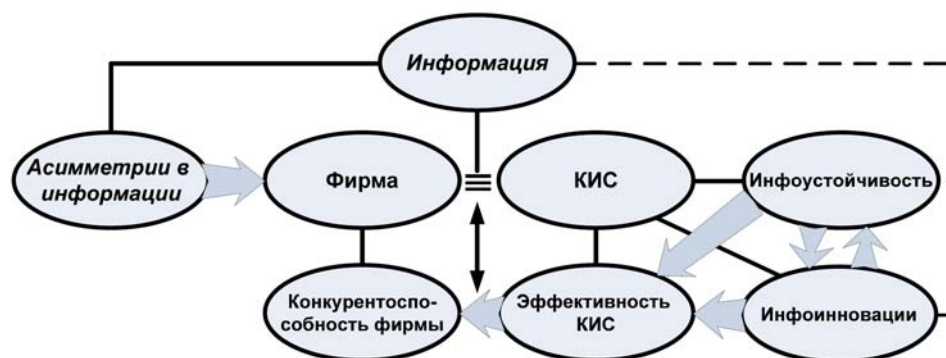


Рис. 1. Модель взаимосвязи асимметрий информации, инфоустойчивости, инфоинноваций

Природа фирмы и несовершенство информации

В настоящее время не существует единой устоявшейся теории фирмы. В рамках данной работы исследование проводится в терминах информационной теории фирмы. В 1937 году Роналд Коуз (лауреат

Нобелевской премии 1991 г.) опубликовал свою статью «О природе фирмы» [4], в которой и нескольких последующих работах ввёл и развил важнейшее понятие *транзакционные издержки*, т.е. издержки рыночной деятельности. Фирмы, как показал Коуз, возникают как ответ на дороговизну рыночной координации. Он обратил внимание на фактор неопределённости в рынке, а также рассмотрел причины увеличения и уменьшения размеров фирм, которые видит в эффективности и результативности управления. Важно отношение Коуза к новшествам (инновациям) – он перечисляет инновации, которые влияют на размер фирмы как те, что позволяют пространственно сблизить факторы производства (в т.ч. средства коммуникации), так и те, которые улучшают технику управления.

Дальнейшее развитие теории фирм (или в более широком смысле институтов) связано с трудами экономистов неонинституциональной и неокейнсианской школ. Из них наибольший интерес в рамках статьи представляет теория Экономии информации Д. Стиглица (лауреат Нобелевской премии 2001 г.). Развивая идею транзакционных издержек, Стиглиц сделал вывод, что рыночные (экономические) издержки *имеют информационные истоки и связаны с неполнотой и несовершенством (асимметриями) информации* в реальном мире (рис. 2).

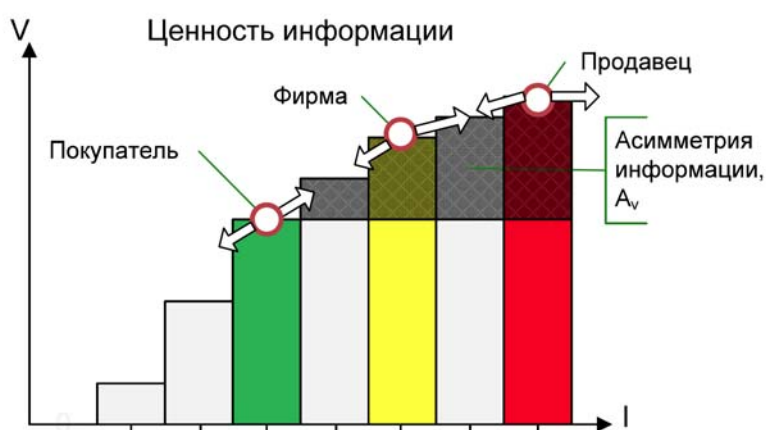


Рис. 2. Асимметрии информации между тремя экономическими агентами на рынке

В своей нобелевской речи [9] Стиглиц сказал: «...Проблемы информации являются центральными для понимания не только рыночной экономики, но также и политической экономики...». Несовершенства информации, способность эффективно устранять эти несовершенства, уменьшать риск и, соответственно, снижать издержки являются глубинной причиной появления и существования фирм. В мире с совершенной информацией никакой нужды в появлении фирм нет. Сознательное управление ресурсами и факторами производства в условиях неопределённости с помощью механизма управления является отличительной чертой фирмы, и это сознательное управление оказывается более эффективным в локальном смысле, чем механизм рыночной саморегуляции (рис. 2).

Комплексная информационная система фирмы

Появление фирм (организаций), как было показано выше, связано с асимметриями в информации и с тем, что «ручное» управление некоторыми рыночными транзакциями при помощи механизма управления фирмы является более эффективным, чем при помощи механизма рыночного саморегулирования. Рассмотрим более подробно внутренние системы организаций, отвечающие за обработку информации. Вообще, в широком смысле, вся фирма в целом, особенно её система управления, занята в различных информационных процессах. Недаром, некоторые авторы сравнивают организации с большими *информационными процессорами* [3].

В настоящее время общепринятой точкой зрения является то, что основной работой служащих в организациях является работа с информацией. Для выделения данной группы работников, то есть тех, кто главным образом создает, обрабатывает или распространяет информацию, экономисты используют специальный термин *информационные служащие*.

Информация (данные, знания), информационные технологии и информационные служащие в организациях системно выделяются в виде *информационной системы (ИС)*, т.е. «системы, основным объектом в которой является информация (в отличие от, скажем, физических тел), и в которой присутствуют информационные процессы (ПИ) хранения, передачи и преобразования информации» [6].

Выделим четыре класса ИС по степени их автоматизации [7]:

1) *вербальные ИС* – традиционные системы коммуникаций, информация которых хранится в памяти отдельных членов сообщества, передаётся между ними вербально или визуально, возможно с помощью технических средств (например, телефонная связь), при этом, как правило, не фиксируясь на носителях;

2) *документарные ИС* – традиционные бумажные технологии ведения информации в журналах, книгах, документах, записях и т.д.;

3) *автоматизированные ИС (АИС)* – настолько популярный класс ИС, что обычно термином ИС в настоящее время называют именно АИС, к ним относятся ИС, оперирующие информацией, представленной в электронном виде;

4) *автоматические ИС* – различные роботы.

Каждый следующий по возрастанию номера класс описанных систем позволяет обрабатывать информацию более эффективно (по соотношению «объём и сложность информации / число информационных служащих»), принимать более качественные решения. Исходя из определения и классификации ИС, сформулируем интегральное определение комплексной информационной системы предприятия. *Комплексная ИС* – это взаимосвязанная совокупность всех ИС, функционирующих в организации, основным объектом в которой является информация, необходимая для принятия взаимосвязанных решений на всех уровнях управления, в которой протекают взаимосвязанные информационные процессы хранения, передачи и преобразования информации. Комплексная ИС организации (КИС) является неотъемлемой частью её системы управления.

Информационная устойчивость фирмы

Современный мир стремительно развивается, увеличиваются, усложняются и гуманизируются потребности человечества, средства и технологии для удовлетворения этих потребностей, естественным образом это сопровождается необычным ростом информации. Проведённые в 2000-2003 гг. исследования с целью определения ежегодных объёмов производимой человечеством информации показали, что объём информации в мире растёт в геометрической прогрессии, удваиваясь примерно каждые три года [8]. Увеличение объёмов информации влечёт увеличение неопределённости и информационных асимметрий, и является причиной появления различного рода институтов, в том числе фирм. Чтобы повысить свою конкурентоспособность фирмы улучшают техники управления, используют информационно-коммуникационные технологии (ИКТ), эксплуатируют больше более глубоких асимметрий. Обратная сторона этого процесса – рост размера и сложности управления фирмой. Поддержание баланса между размером фирмы, эффективностью и результативностью её системы управления – одна из важнейших задач её менеджмента. Среди новых *информационных* характеристик фирмы, которые следует учитывать при управлении, выделим как одну из важнейших – *инфоустойчивость* [7].

Устойчивость системы связывают с её способностью противостоять воздействиям на неё. Устойчивость является интегральным свойством и может объединять различные свойства: прочность, стойкость к

воздействию внешних факторов, защищённость, стабильность, надёжность, живучесть и т.д. Степень устойчивости системы зависит от многих факторов: числа и разнообразия элементов, характера и силы связей между ними, направленности и значительности воздействий и т.д. Устойчивость важна для сохранения системы, но имеет и отрицательную сторону – сопротивление к позитивным переменам, в том числе и инновациям. Связанное понятие *равновесия* системы ассоциируется с процессом её нормального стабильного функционирования в динамично меняющихся условиях, когда возмущения уравновешиваются адаптационными свойствами системы. Если система устойчива, то она способна вернуться в исходное или близкое к нему состояние после внесения в неё возмущений. Потеря устойчивости системы может произойти из-за изменения её параметров, повлёкшего за собой бифуркацию (катастрофу), либо из-за наличия непредусмотренных при создании системы воздействий, либо, наконец, из-за нарушения связей в системе, вызвавших изменение её структуры (структурная неустойчивость). Потеря устойчивости – наиболее распространённая причина гибели систем. Следовательно, наиболее эффективным способом разрушения системы является её *дестабилизация*, т.е. целенаправленное отрицательное воздействие на устойчивость системы [2]. В неустойчивом состоянии нельзя прогнозировать динамику процессов. Система становится неуправляемой, т.к. её состояние не поддаётся измерениям, и, следовательно, отсутствует требуемое позиционирование системы во времени и пространстве. В конечном итоге система разрушается путём потери целостности. Информационная устойчивость социально-экономических систем затрагивалась в нескольких теориях (табл. 1).

Таблица 1.

<i>Теория</i>	<i>Понимание инфоустойчивости</i>	<i>Примечание</i>
Теория информации Шеннона	Устойчивость информации к помехам в каналах, приёмниках и передатчиках, других частях системы (помехоустойчивость).	Механистический подход. Нет семантики информации.
Теория информационной безопасности фирмы	Обеспечение надёжного и безопасного функционирования компьютерной сети фирмы, учитывая риски нападений из внешних компьютерных сетей и утечки информации через инсайдеров.	Используются методы прикладной информатики. Механистический подход.
Социология и политология	Стабильное развитие государств и корпораций в состоянии информационной войны при агрессивных выбросах клеветы, компромата, слухов и т.д. и т.п.	Социальные методы исследования.

В данной статье *инфоустойчивость* понимается как *интегральное свойство фирмы, отражающее стабильность её параметров в динамично меняющейся информационной среде, где информация имеет несовершенную природу.*

Для исследования устойчивости систем применяется математический аппарат, объединяющий *теории устойчивости* непрерывных и дискретных систем (исследуют устойчивость системы при наличии начальных возмущений, критерием служит возвращение системы в установившийся режим) и *теорию катастроф*, которая предоставляет универсальный метод для исследования скачкообразных переходов, разрывов, внезапных качественных изменений в процессах. Инфоустойчивость фирм может исследоваться и классическими теориями, и теорией катастроф. Например, на рынках естественных монополий асимметрии постоянны. В то же время, есть отрасли с катастрофичным характером изменения равновесия в асимметриях информации, связанным, например, с появлением новой ИКТ, выходом на рынок новой фирмы, принятием законов. Кстати, теория катастроф подтверждает, что очень трудно бороться с катастрофой, если уже видны её признаки, вследствие бесконечно быстрого нарастания

самоподдерживающихся колебаний или опять же бесконечно высокой скорости конвергенции при столкновении с неустойчивым режимом [5].

Эффективность КИС и инфоустойчивость (в порядке постановки проблемы)

КИС фирмы представим в виде совокупности элементов: $\Phi = (S, X, I)$, где S – структура, X – поведение (множество функций передачи, хранения, преобразования информации), I – множество информационных структур данной КИС.

Проиндексируем множество информационных структур через параметр $\theta \in \Theta$ и введём функцию ценности информации $V(\theta)$, которую интерпретируем как ценность информационной структуры θ . Под эффективностью КИС далее будем понимать ценность информации, продуцируемую КИС в единицу времени $E = \frac{dV}{dt}$.

Определим динамическую модель КИС (E-модель): $\frac{dv}{dt} = H(V, A, B, X, t)$, где A – множество информационных асимметрий, эксплуатируемых фирмой при рыночных транзакциях, B – множество внутренней информации, используемой в процессах управления фирмы (тоже подверженной асимметриям), t – время.

Представляет интерес исследование устойчивости полученной E-модели, для которого возможна следующая последовательность шагов:

- 1) на основе результатов теории экономики информации выдвинуть гипотезу о виде модели;
- 2) оценить параметры модели;
- 3) проверить адекватность и качество построенной модели;
- 4) исследование информационной устойчивости модели;
- 5) применение в различных приложениях, например, внедрения инфоинноваций.

В данный момент времени исследование находится на начальной стадии. Одно из предположений о возможном поведении модели представлено на рис. 3, который наглядно представляет различные варианты потери устойчивости: скачки, разрывы, разбалансировку, что может объясняться внедрением инноваций, сбоями в АИС, деградацией системы управления и т.д.

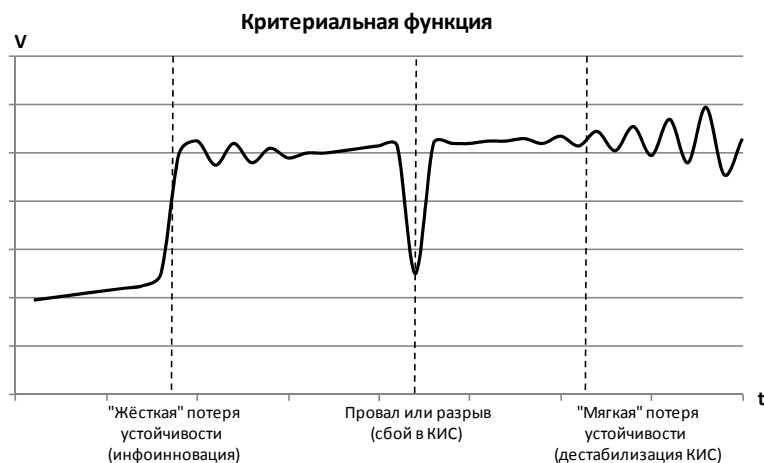


Рис. 3. Поведение критерияльной функции

Информационные инновации в фирме и их оценка

В настоящее время внедрение инноваций в фирмах в основном оценивается финансово-стоимостными методами (ФСМ), например, методами оценки окупаемости инвестиций (NPV, IRR, Payback и т.д.) и стоимости владения (ТСО), что не совсем приемлемо для инфоинноваций. Влияние инфоинноваций на эффективность организации зачастую трудно оценить в деньгах, поскольку информационные процессы являются обеспечивающими. Финансовый эффект от инфоинноваций может быть достоверно вычислен только скрупулёзным анализом, который часто экономически нецелесообразен. Стоимостные методы показывают затратную составляющую инфоинноваций. Поэтому ФСМ дают приблизительные оценки. Более достоверны качественные и вероятностные методы оценки, например Метод сбалансированных показателей (BSC), но он требует внедрения и поддержки на всех уровнях предприятия соответствующей системы, что требует больших усилий, вплоть до реорганизации компании.

Теоретический аппарат инфоустойчивости практически можно использовать при оценке КИС существующих предприятий и распространении лучшего опыта, анализе эффекта альтернативных мероприятий (в т.ч. инноваций) на инфоустойчивость предприятий и обоснованном выборе альтернатив. Использование анализа инфоустойчивости, не отменяя существующие методики оценки, дополняет их предложением оценки инфоинноваций с точки зрения КИС, анализа их информационного эффекта.

В силу быстрых изменений в информационной среде инфоустойчивость предприятия должна постоянно поддерживаться комплексом мероприятий, направленных на развитие КИС, т.е. инфоинновациями.

Из определения КИС, ориентированной на информационных служащих, информационные процессы и ценность информации как источника принятия решений, приведём основные группы информационных инноваций в организациях [7]:

- 1) *регистрирующие инфоинновации* – новшества, которые направлены на перевод информации из вербального и документарного вида в более пригодный для обработки электронный (цифровой) вид;
- 2) *структурирующие инфоинновации* – новшества, упорядочивающие информацию из неструктурированных источников в массивы с чётко определённой структурой, позволяющей их понимание и обработку людьми и машинами;
- 3) *преобразующие инфоинновации* – новшества, преобразующие имеющуюся информацию в качественно новые её виды с точки зрения использования в процессе управления при принятии решений, т.е. различные показатели, индикаторы, прогнозы и т.п.

Взаимосвязь инфоустойчивости и инфоинноваций

Рассмотрим одно из приложений предложенной модели – влияние инфоустойчивости на инфоинновации в фирме. Оно проиллюстрировано на графике производительности КИС (см. рис.4), где отображены два положения устойчивого равновесия для КИС, каждое из которых притягивает ближайшие состояния.

Можно утверждать, что инфоустойчивость в данном случае будет играть и положительную и отрицательную роль для внедрения инфоинноваций в зависимости от их природы. Если инфоинновация имеет эволюционную природу и направляет КИС к ближайшему положению равновесия, то её успех обеспечен, однако всё наоборот, если инфоинновация требует перестройки КИС, перевода её в новое устойчивое состояние.

Математическая модель, связывающая инфоустойчивость фирмы с инфоинновациями, может быть построена на основе теории перестроек В.И. Арнольда [1]. Трудности проблемы перестройки состоят в её нелинейности. Общепринятые методы управления, при которых результаты пропорциональны усилиям,

здесь не действуют и необходимо выработать управляющие воздействия на основе порой парадоксальных выводов нелинейной теории. В.И. Арнольд делает несколько качественных выводов для нелинейной системы, находящейся в установившемся устойчивом состоянии, признанном плохим, поскольку в пределах видимости имеется лучшее, предпочтительное устойчивое состояние системы. Перефразируя их в терминах КИС, инфоустойчивости и инфоинноваций получим: при постепенном внедрении инфоинноваций (движение по графику на рис. 4 из текущего положения вправо) сопротивление КИС, связанное с её инфоустойчивостью, растёт, при этом состояние КИС ухудшается; пройдя пик сопротивления (точка перегиба) и затем точку худшего состояния КИС (локального минимума), система попадает в зону притяжения нового состояния, в котором сопротивление текущего состояния уже полностью отсутствует и система начинает притягиваться к новому состоянию; для слаборазвитых систем есть возможность перейти в новое состояние почти без ухудшения, в то время как развитая КИС, в силу своей инфоустойчивости, на такое постепенное непрерывное улучшение неспособна; если систему удаётся сразу (скачком), а не непрерывно, перевести из плохого устойчивого состояния достаточно близко к хорошему, то дальше она будет сама собой эволюционировать в сторону хорошего состояния. Эти качественные выводы очень хорошо согласуются с известными огромными цифрами неудач при первом внедрении больших АИС (например, ERP) в устоявшихся организациях, что, как теперь становится понятно, связано с сильной инфоустойчивостью их КИС. Вопросам создания научно-обоснованных комплексов управляющих воздействий для преодоления инфоустойчивости путём оптимальной дестабилизации систем при внедрении инноваций в фирмах следует посвятить отдельное исследование.

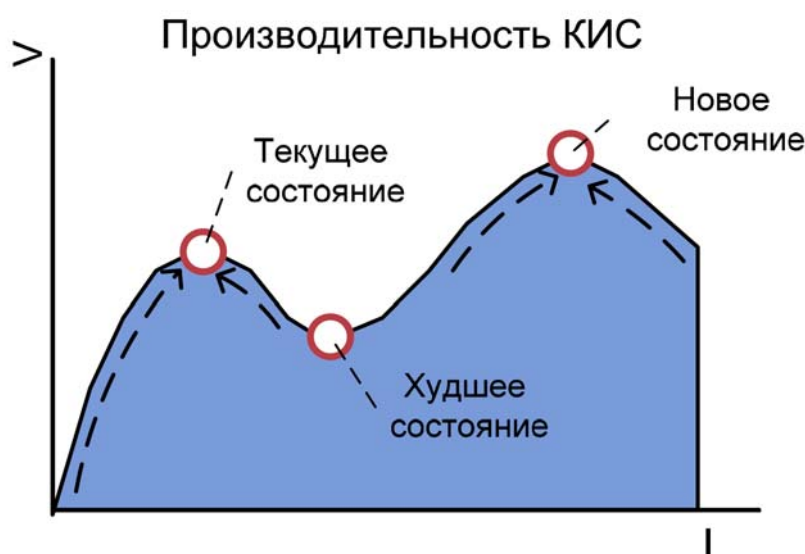


Рис 4. Состояния КИС в процессе внедрения инноваций

Инфоинновации внедряются в инструментальной, методологической и организационной сферах. Отличительной чертой инфоинноваций является их высокая наукоёмкость и технологичность, использование теоретических и прикладных научных результатов. Отметим, что инфоинновации требуют соответствующего профессионального роста информационных служащих. Ведь, несмотря на любые инновации, пренебрежение законами природы и общества, падение компетентности и отсутствие личной ответственности специалистов за принимаемые решения рано или поздно приводит к катастрофе.

Заключение

В представленной статье (в порядке постановки проблемы) вводится модель инфоустойчивости фирмы. Предлагается использовать показатель инфоустойчивости в управлении фирмой. На основании методов теории перестроек сделаны выводы о влиянии инфоустойчивости на внедрение инфоинноваций в фирме. Сделаны выводы о необходимости продолжения исследования инфоустойчивости, построения её модели, выявления факторов и условий дестабилизации, построения различных прикладных моделей.

Благодарности

Статья частично финансированна из проекта **ITHEA XXI** Института Информационных теории и Приложений FOI ITHEA и Консорциума FOI Bulgaria (www.ithea.org, www.foibg.com).

Библиографический список

1. Арнольд В.И. Теория катастроф. – М.: Наука. Гл. ред. физ.-мат. лит., 1990. – 128 с.
2. Дрогобыцкий И.Н. Системный анализ в экономике: учеб. пособие. – М.: Финансы и статистика, 2007. – 512 с.: ил.
3. Информационные технологии в бизнесе / Под ред. М. Желены. – СПб: Питер, 2002. – 1120 с.: ил.
4. Коуз Р. Фирма, рынок и право / Пер. с англ. М.: Новое издательство, 2007. – 224 с.
5. Острейковский В.А. Анализ устойчивости и управляемости динамических систем методами теории катастроф: Учебное пособие для вузов. – М.: Высш.шк., 2005. – 326 с.
6. Романов В.П. Интеллектуальные информационные системы в экономике. – М.: Экзамен, 2003.
7. Фролов А.П. Информационная устойчивость и инновации на предприятиях / Сборник Формирование стратегии инновационного развития экономических систем / Под ред. д-ра экон. наук, проф. В.В. Глухова, д-ра экон. наук, проф. А.В. Бабкина: труды конф. – СПб.: Изд-во Политехн. ун-та, 2008. С. 382-390.
8. Lyman P., Varian H. How Much Information? [Электронный ресурс] [Режим доступа: <http://www.sims.berkeley.edu/research/projects/how-much-info-2003>].
9. Stiglitz J. Information and the Change in the Paradigm in Economics. Part 1, 2. NY: The American Economist, Vol.47, No.1 (Fall 2003), Vol.48, No.1 (Spring 2004).

Сведения об авторах

Александр Фролов – Пермский государственный университет, аспирант кафедры информационных систем и математических методов экономики, Россия, г. Пермь, 614990, ул. Букирева, д. 15;
e-mail: frolov@pmz.ru

АВТОМАТИЗАЦИЯ УПРАВЛЕНИЯ БИЗНЕС-ПРОЦЕССАМИ НА ОСНОВЕ ЯЗЫКА GBPR

Антон Рычков

Аннотация: Рассматриваются возможности разработанного автором языка GBPR (General Business Process Representation) для решения проблем интеграции бизнес-процессов структурных подразделений учреждений и предприятий-партнеров, а также оптимизации бизнес-процессов и их динамической настройки. Представленный язык основывается на *pi*-исчислении и является его расширением, ориентированным на автоматизацию управления бизнес-процессами. Основой языка GBPR, как и *pi*-исчисления, являются информационные каналы, по которым передаются сообщения (сервисные каналы – каналы, отвечающие за обращение к сервису или процессу; маршрутные каналы, отвечающие за маршрутизацию потоков управления и данных внутри процесса). Базовой структурой языка GBPR является действие, выполняющее обращение к сервису. Процесс состоит из множества действий, выполняемых параллельно, последовательность действий реализуется через последовательную передачу сообщений между действиями. Процесс является частным случаем сервиса и может быть вызван как сервис посредством обращения по соответствующему информационному каналу. В статье приведено описание языка, его возможностей, в частности, показана возможность процессной интеграции, основанной на ретрансляции сообщений и отношении наследования процессов.

Ключевые слова: бизнес-процесс, автоматизация управления, *pi*-исчисление, наследование, интеграция, оптимизация.

ACM Classification Keywords: H.4 Information Systems Applications: H.4.1 Office Automation – Workflow management.

Conference: The paper is selected from Seventh International Conference on Information Research and Applications – i.Tech 2009, Varna, Bulgaria, June-July 2009

Введение

В настоящее время в области автоматизации управления существуют нерешенные проблемы *интеграции* на основе общих бизнес-процессов различных предприятий с учетом специфики каждого из участников интеграции. Существуют также проблемы *частной оптимизации бизнес-процессов* (оптимизации частных случаев реализации) и *динамического изменения* версий процесса.

Процессно-ролевой подход к управлению является основой современной эффективной организации. Автоматизируются процессы (Workflow Management, WfM) и функции по их управлению (Business Process Management, BPM). Современные информационные технологии позволяют создавать приложения по модели бизнес-процессов, собирая их функциональность из отдельных «кубиков» (Service Oriented Architecture, SOA). Интеграция партнеров в системах B2B (Business-to-Business) также выполняется на основе процессов. Все это приводит к повышению степени важности модели бизнес-процессов организации и повышенным требованиям к скорости ее изменения.

Управление бизнес-процессами

Технология BPM является комплексной и включает в себя WfM, SOA и B2B. Действия, которые составляют управление бизнес-процессами, можно разделить на несколько групп: разработка, моделирование, выполнение, мониторинг и оптимизация. Управление фокусируется на повышении эффективности бизнес-процессов, приносящих «прибыль» организации, и их непрерывной оптимизации.

Таким образом, важными аспектами являются возможности моделирования бизнес-процессов, скорость их внедрения (изменения) и мониторинг показателей.

Современные языки описания бизнес-процессов ориентируются на возможность последующей автоматизации управления этими бизнес-процессами. Для целей автоматизации управления были созданы языки BPEL и BPML, а языки, созданные ранее, были дополнены (ARIS).

Бизнес-процессы организации включают в себя и межорганизационное взаимодействие, которое также может быть описано как процесс, в котором участвуют две и более организаций. Автоматизация таких процессов требует технической реализации обмена данными. В настоящее время стандартом такого рода взаимодействия являются Web-сервисы. Более того, Web-сервисы стали стандартом и для взаимодействия различных информационных систем внутри организации, что позволяет создавать программные приложения на основе существующих бизнес-процессов организации (SOA). Возможности использования Web-сервисов в бизнес-процессах описаны в работах [1, 2].

Web-сервис реализует определенную функциональность, скрывая реализацию этой функциональности, однако в ряде случаев требуется более тесная интеграция, позволяющая организациям или подразделениям разделять один экземпляр бизнес-процесса и контролировать его выполнение от начала до конца. В этом случае Web-сервисы могут предоставлять протоколы нижнего уровня для разделения экземпляра бизнес-процесса между различными информационными системами одной или нескольких организаций, вовлеченных в выполнение процесса. Однако остается проблема реализации методов взаимодействия на верхнем уровне.

Интеграция организаций привела к разделению описания бизнес-процессов на две группы:

- внешние процессы описывают процессы взаимодействия с внешними источниками (другие организации или клиенты);
- внутренние процессы описывают реальные процессы организации, являются более детальными.

Внешние процессы организации являются «видимой» стороной выполнения *внутренних процессов*. Иными словами, внешние процессы описывают внешний интерфейс внутренних процессов организации. Описание внешних процессов открыто для партнеров, в то время как описание внутренних процессов зачастую является коммерческой тайной, так как определяет конкурентное преимущество.

Такое разделение описания процессов требует создания *технологии верификации* соответствия внутреннего процесса внешнему. Одним из вариантов решения этой проблемы является *механизм наследования процессов*. Отношение наследования в данном случае отличается от аналогичного отношения в UML. Так в UML наследуются свойства и методы класса – статические характеристики, в то время как для процесса большее значение имеет поведение – динамическая характеристика. Производный процесс должен поддерживать поведение базового процесса с возможностью его расширения. Таким образом, производный процесс должен выполнять все действия базового процесса в порядке, определенном для базового процесса.

Наследование бизнес-процессов также позволяет решать такие проблемы как:

- проблема динамического изменения – возникает при внедрении новой версии существующего процесса и требует миграции существующих экземпляров процесса на новую версию;
- проблема агрегации управленческой информации выполняющихся процессов в случае наличия различных версий этих процессов [3];
- проблема частной оптимизации – для процесса существуют различные варианты, оптимизированные под частные случаи общего процесса (например, «обслуживание клиента» и такие варианты как «обслуживание частного лица» и «обслуживание организации») [4].

Описание языка GBPR

Для решения приведенных выше задач был разработан язык описания бизнес-процессов GBPR (General Business Process Representation) [7]. Данный язык основывается на *пи-исчислении* [5, 6] и является его расширением, ориентированным на автоматизацию управления бизнес-процессами.

Основой языка GBPR, как и *пи-исчисления*, являются *информационные каналы*, по которым передаются сообщения, но разделяются на группы:

- *сервисные каналы* – каналы, отвечающие за обращение к сервису или процессу;
- *маршрутные каналы* – каналы, отвечающие за маршрутизацию потоков управления и данных внутри процесса.

Базовой структурой языка GBPR является *действие*, выполняющее обращение к сервису:

$$ACTIVITY ::= MESSAGEMAP.[CONDITION] SERVICE(PARAMS)$$

Описание действия включает в себя карту необходимых сообщений (маршрутизация), условие (проверка данных) и обращение к сервису (с указанием сервиса и передаваемых параметров). Так как взаимодействие с сервисами является асинхронным (сервис не может передать ответное сообщение по каналу обращения), то в параметрах указывается канал для ответа в случае, если он требуется.

Карта сообщений состоит из информационных каналов, по которым требуется получить сообщение, связанных различными операциями:

- операция *суммирования* «+» определяет возможные альтернативы (иными словами, для вызова сервиса необходимо получить сообщение *по одному* (не важно по которому) из каналов);
- операция *распараллеливания* «|» определяет список каналов, по которым ожидаются сообщения, и для обращения к сервису необходимо получить сообщения *по всем* каналам из списка.

Описание сервисов в языке GBPR определяет информационный канал для обращения, список необходимых параметров и правила обращения (протокол, адрес и другая техническая информация).

Процесс является частным случаем сервиса, и может быть вызван как сервис посредством обращения по соответствующему информационному каналу. Процесс состоит из множества действий, выполняемых параллельно, последовательное выполнение действий реализуется через последовательную передачу сообщений между действиями. Процесс может содержать блоки обработки исключительных ситуаций и блок завершения.

$$PROCESS ::= ACTIVITY | \{ACTIVITY\}$$

GBPR удовлетворяет большинству шаблонов, описанных в работах [9, 10]. Данные шаблоны описывают встречающиеся ситуации в описании маршрутизации бизнес-процессов. Описания, созданные на других наиболее распространенных языках моделирования бизнес-процессов, могут быть преобразованы в GBPR-описания, как это показано в [8].

Интеграция на основе процессов в GBPR

Как уже говорилось выше, интеграция организаций (или подразделений одной организации) на основе бизнес-процессов является актуальной задачей.

Поскольку маршрутизация потоков управления и данных в языке GBPR выполняется на основе передачи сообщений по информационным каналам, то процесс интеграции сводится к *ретрансляции сообщений* во все системы-участники процесса.

Рассмотрим пример процесса ремонта автомобиля (рис. 1). В процесс вовлечены автосервис «А» и страховая компания «В»:

start. «Приезд клиента» (e1)

e1() + e3(«Состояние»). [«Состояние» != 'Согласовано'] «Составление калькуляции» (e2)

e2(). «Согласование калькуляции» (e3)

e3(«Состояние»). [«Состояние» = 'Согласовано'] «Ремонт автомобиля» (e4)

e4(). «Оплата ремонта» (finish)

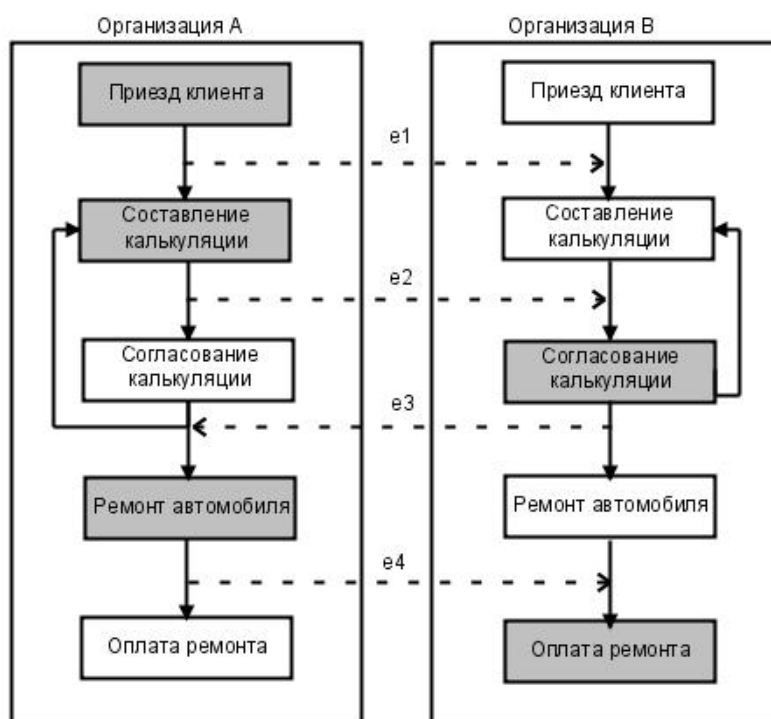


Рис. 1. Процессная интеграция в GBPR

На рис. 1 показано, как ретрансляция сообщений позволяет поддерживать оба экземпляра процесса в идентичном состоянии. При этом могут быть выбраны *различные схемы ретрансляции сообщений*, как *централизованные* (выделенный центральный узел, отвечающий за ретрансляцию), так и *децентрализованные* (система, в которой сгенерировано сообщение, передает его в каждую из систем-участниц процесса). В большинстве случаев подойдет централизованная схема ретрансляции, где в качестве центрального узла выступает система, породившая процесс. Такой подход позволяет снизить издержки, связанные с содержанием каналов для ретрансляции.

Наследование процессов в GBPR

Введение отношения наследования позволяет решить ряд проблем, в том числе и уточнения общего процесса в случае процессной интеграции, за счет чего учитывается специфика каждой организационной участницы процесса и сохраняется в секрете описание внутреннего процесса.

Процесс можно определить как набор работ, набор ресурсов и набор правил. Для упрощения мы не будем рассматривать ресурсную составляющую в описании бизнес-процессов, тогда процесс можно разделить на набор сервисов, выполняемых в бизнес-процессе, и набор правил, определяющих порядок выполнения этих сервисов.

Представим описание процесса как *класс*, содержащий описание необходимых параметров, правил маршрутизации и связей с сервисами, и определим отношение наследования.

Параметры процесса могут быть дополнены и уточнены, иными словами, могут появиться дополнительные параметры, а тип существующих параметров может быть изменен на производный от указанного. Уточнение возможно за счет реализации отношения наследования для классов, определенного в UML.

Отношение наследования определяется в рамках пи-исчисления. При этом, так как сервисы являются внешними по отношению к процессу, то они моделируются посредством исполнителя [5]:

$$Exec(x) ::= x(y).y^{\wedge}.$$

Если процесс Q находится в бинарном отношении наследования от процесса P ($P \leq Q$), то:

- $fn(P)$ принадлежит $fn(Q)$ – все свободные имена процесса P также являются свободными именами процесса Q ;
- для любой транзакции $a(x)$, такой что $P \rightarrow P'$ под действием $a(x)$, найдется такой процесс Q' и последовательность транзакций a_1, a_2, \dots, a_n , таких, что для любого y из области определения x справедливо $Q \rightarrow Q'$ под действием $a_1, a_2, \dots, a_n, a(x)$ и $P'\{y/x\} \leq Q'\{y/x\}$, транзакции a_1, a_2, \dots, a_n могут быть входящими $a(x)$ или исходящими $a^{\wedge}(x)$, транзакция $a(x)$ – входящая, и a является свободным именем для процессов P и Q или является сервисным каналом этих процессов;
- для любой транзакции $a(x)$, такой что $P \rightarrow P'$ под действием $a(x)$, найдется такой процесс Q' и последовательность транзакций a_1, a_2, \dots, a_n , таких, что для любого y из области определения x , справедливо $Q \rightarrow Q'$ под действием $a_1, a_2, \dots, a_n, a(x)$ и $P' \leq Q'$, транзакции a_1, a_2, \dots, a_n могут быть входящими $a(x)$ или исходящими $a^{\wedge}(x)$, транзакция $a(x)$ – любая транзакция, кроме случая из предыдущего пункта;
- для любой транзакции $a(x)$, такой что $Q \rightarrow Q'$ под действием $a(x)$ и не найдется процесс P' , такой, что справедливо $P \rightarrow P'$ под действием $a(x)$, тогда для любого y из области определения x выполняется $P \leq Q'\{y/x\}$, транзакция $a(x)$ – входящая, и a является свободным именем для процесса Q или сервисным каналом этого процесса;
- для любой транзакции $a(x)$, такой что $Q \rightarrow Q'$ под действием $a(x)$ и не найдется процесс P' , такой, что справедливо $P \rightarrow P'$ под действием $a(x)$, тогда для любого y из области определения x выполняется $P \leq Q'$, транзакция $a(x)$ – любая транзакция, кроме случая из предыдущего пункта.

Отношение наследования *не является симметричным*, иными словами, из $P \leq Q$ не следует $Q \leq P$.

С учетом отношения мы можем построить *иерархию процессов*, на самом верхнем уровне которой будет *абстрактный процесс*. Такая иерархия помогает управлять вариантами процесса, определяя частные случаи как производные от общего.

Рассмотрим пример процесса «обращение» и производного от него процесса «обращение от организации»:

```
«обращение» ::= start. «прием обращения» (e1) |
              e1(). «оформление заказ-наряда» (e2) |
              e2(). «оказание услуг» (finish)

«Обращение от организации» ::= start. «прием обращения» (e1) |
                               e1(). «оформление договора» (e2) |
                               e2(). «оформление заказ-наряда» (e3) |
                               e3(). «оказание услуг» (finish)
```

Процесс «обращение» находится на первом уровне иерархии процессов, в то время как процесс «обращение от организации» является производным от него и находится на втором уровне. Аналогично,

если есть производный процесс «обращение от физического лица», то оно также будет находится на втором уровне. Следует заметить, что изменение базового процесса (*эволюционная оптимизация*) не всегда приводит к перепроектированию производных процессов, а создание варианта процесса не приводит к изменению базового процесса. Это упрощает частную оптимизацию (оптимизацию частных случаев процесса).

Аналогично может быть использовано отношение наследования в случае процессной интеграции для уточнения разделяемого процесса с учетом специфики каждой из организаций-участниц интеграции.

Заключение

Язык GBPR позволяет проводить основанную на процессах интеграцию организаций с возможностью контроля всего процесса. Асинхронная передача сообщений, лежащая в основе GBPR, позволяет использовать различные протоколы передачи данных, что придает гибкость интеграции. Возможность прозрачного уточнения общего процесса благодаря отношению наследования для процессов позволяет учесть специфику участников интеграции без ущерба для общего процесса. Наследование помогает также в проведении частной оптимизации и решении проблем динамического изменения версий процессов.

Благодарности

Статья частично финансированна из проекта **ITHEA XXI** Института Информационных теории и Приложений FOI ITHEA и Консорциума FOI Bulgaria (www.ithea.org, www.foibg.com).

Библиографический список

- [1] Feng G., Wang C., Li H. Web services based cross-organizational business process management // Web Technologies Research and Development – APWeb 2005. Springer Berlin, Heidelberg, 2005.
- [2] Papazoglou M., Yang J. Design methodology for Web services and business processes // Technologies for E-Services. Springer Berlin, Heidelberg, 2002.
- [3] van der Aalst W.M.P. Inheritance of business processes: a journey visiting four notorious problems // Petri Net Technology for Communication-Based Systems. Springer Berlin, Heidelberg, 2004.
- [4] Hallerbach A., Bauer T., Reichert M. Managing process variants in the process life cycle // Technical report. University of Twente, Enschede, 2008.
- [5] Milner R., Parrow J., Walker D. A calculus of mobile processes, part 1 // Information and Computation. Academic Press, Duluth, 1992.
- [6] Milner R., Parrow J., Walker D. A calculus of mobile processes, part 2 // Information and Computation. Academic Press, Duluth, 1992.
- [7] Рычков А.Ю. Язык описания бизнес-процессов на основе пи-исчисления // Математика программных систем: Межвузовский сборник научных трудов / Перм. ун-т. Пермь, 2008.
- [8] Рычков А.Ю. Преобразование описаний бизнес-процессов в представление на языке GBPR // Математика программных систем: Межвузовский сборник научных трудов / Перм. ун-т. Пермь, 2008.
- [9] van der Aalst W.M.P., ter Hofstede A.H.M., Kiepuszewski B., Barros A.P. Workflow patterns // Distributed and Parallel Databases, vol.14, num.1. Springer Netherlands, 2004.
- [10] van der Aalst W.M.P., ter Hofstede A.H.M., Kiepuszewski B., Barros A.P. Advanced workflow patterns // Cooperative information systems: 7th International Conference, CoopIS 2000. Springer-Verlag, New York, 2000.

Сведения об авторе

Антон Рычков – аспирант кафедры математического обеспечения вычислительных систем, Пермский государственный университет; ул. Букирева, д. 15, г. Пермь, 614990, Россия;
e-mail: anton@rychkov.biz

МЕХАНИЗМЫ РАСПРЕДЕЛЕНИЯ КВОТ НА ВЫБРОСЫ ПО КИОТСКОМУ ПРОТОКОЛУ

Алексей Волошин, Ирина Горицына

Аннотация: Проблема изменения климата не сводится к обычным экологическим вопросам. Затрагиваются базовые экономические и политические интересы. Киотский протокол – первый международный документ, использующий рыночный механизм для решения глобальных экологических проблем. Предлагаются математические модели для реализации механизмов Киотского протокола.

Ключевые слова: Киотский протокол, углеродный рынок, задача распределения квот.

ACM Classification Keywords: ACM Classification Keywords: I. Computing Methodologies – I.6. Simulation and modeling (Time series analysis) – I.6.8 Types of Simulation – Gaming.

Conference: The paper is selected from XVth International Conference "Knowledge-Dialogue-Solution" KDS 2009, Varna, Bulgaria, June-July 2009

Введение

Среди глобальных экологических вызовов, угрожающих цивилизации в самом начале XXI века, на первое место, безусловно, должна быть поставлена проблема изменения климата. Мировое сообщество, понимая все потенциальные угрозы глобального изменения климата, предпринимает ряд согласованных мер по адекватному реагированию на этот вызов. Наблюдаемые изменения климатической системы Земли ученые связывают с аномальным ростом концентрации в атмосфере так называемых парниковых газов (ПГ), в первую очередь, CO₂. За последние десятилетия концентрация в атмосфере CO₂ выросла на треть, метана – в 2.5 раза, выброшены сотни миллионов тонн хлорфторуглеродов (до середины прошлого века их на Земле вовсе не было).

Абсолютное доказательство подобной гипотезы невозможно в принципе: у нас нет второй Земли, чтобы поставить контрольный опыт, поместив ее в точно такие же условия, но исключив влияние человека. Хотим мы этого или нет, все суждения о будущем Земли будут не более, чем предположением. Человек, вынужденный продвигаться в полной тьме, будет стараться это делать с максимальной осторожностью. Точно также человеческой цивилизации в условиях неопределенности предпочтительнее исходить из той гипотезы, которая предписывает ей более осторожное собственное поведение. Именно такой гипотезой и является антропогенная теория изменения климата. У этой теории есть одно важное преимущество: на сегодня она лучше согласована со всеми известными фактами, чем любая другая.

В 2007 г. Межправительственная группа экспертов по изменению климата (МГЭИК) - представила свой Четвертый оценочный доклад, в котором содержатся обобщенные выводы о прошлых и нынешних климатических изменениях, их воздействии на природу и человека, а также о возможных мерах по противодействию таким изменениям. В этом Докладе говорится, что с вероятностью 0.9 наблюдаемые изменения климата связаны с деятельностью человека. Несмотря на расхождения во взглядах ученых относительно причин изменения климата, сам факт глобального потепления уже никто не отрицает. При этом недостаточная научная определенность не должна использоваться в качестве причины для отсрочки принятия мер по снижению антропогенного влияния на атмосферу Земли.

Развитые страны ответственны за 75% парниковых газов, содержащихся на сегодняшний момент в атмосфере, если не принимать в расчет выбросы от вырубки лесов. Эти страны имеют наибольшие финансовые ресурсы и мощнейший технологический потенциал для сокращения своих выбросов. Таким образом, они должны взять на себя наибольшую тяжесть борьбы с изменением климата в следующем десятилетии. Для того, чтобы удержать глобальное потепление в пределах 2°C выше температуры доиндустриального периода, ЕС и другим развитым странам необходимо сделать следующий важный шаг в рамках нового международного соглашения – снизить общий уровень выбросов к 2020 г. на 30% по сравнению с уровнем 1990 г. Существуют заслуживающие самого серьезного внимания научные доказательства того, что дальнейшее повышение температуры сверх указанного порога может вызвать необратимые и потенциально катастрофические изменения. Хотя в ближайшее время основное бремя борьбы с изменением климата ляжет на развитые страны, одних их усилий будет недостаточно для сокращения выбросов парниковых газов в мировом масштабе.

С ростом потенциала развивающихся стран увеличиваются и их выбросы в атмосферу, которые, по прогнозам, к 2020 г. превысят объемы выбросов развитых стран. Таким образом, необходимо, чтобы развивающиеся страны, в частности, с наиболее быстро развивающейся экономикой, начали как можно скорее ограничивать рост выбросов с перспективой сокращения в абсолютном выражении с 2020 г. и далее. Как только развивающиеся страны достигнут уровня развития, сопоставимого с уровнем развитых стран, они должны также принять на себя обязательства по сокращению выбросов. Эти обязательства должны соответствовать их уровням выбросов, а также техническим и финансовым возможностям стран.

Прогнозируется, что в долгосрочной перспективе изменение климата может привести к сокращению валового внутреннего продукта (ВВП) от 5% до 20% в год, если его не сдерживать путем уменьшения выбросов парниковых газов. Таким образом, проведение международной кампании по борьбе с изменением климата является стратегией с перспективой роста в долгосрочном плане. Чем раньше мы начнем действовать, тем меньше будет цена наших действий. Анализ Европейской Комиссии показывает, что объем инвестиций для перехода на народное хозяйство с низким уровнем выбросов углеродосодержащих газов составит лишь 0,5% мирового ВВП в период между 2013 и 2030 гг. Согласно этим прогнозам, проведение международной кампании по борьбе с изменением климата сократило бы рост мирового ВВП всего лишь на 0,14% в год до 2020 г. Рост мирового ВВП в период с 2005 по 2020 гг. составит 53%, что лишь немного ниже прогнозируемого уровня в 55%, если никаких мер предприниматься не будет. И в этих цифрах не учитываются преимущества сокращения выбросов, такие как снижение ущерба от изменения климата, которого удалось избежать, большая энергетическая безопасность и экономия расходов здравоохранения в связи со снижением уровня загрязненности воздуха.

В качестве первого шага была подготовлена и принята 9 мая 1992г. рамочная Конвенция ООН об изменении климата (РКИК). А в июне 1992 г. на Всемирной конференции ООН по окружающей среде и развитию в Рио-де-Жанейро ее подписали 154 страны. Затем встал вопрос о конкретных обязательствах и в декабре 1997 г. на Третьей конференции Сторон РКИК был принят **Киотский протокол** (КП) к РКИК, главной особенностью которого является юридически обязывающие количественные обязательства развитых стран и стран с переходной экономикой по ограничению и снижению поступления ПГ в атмосферу. В самом широком смысле Киотский протокол – первый международный документ, использующий рыночный механизм для решения глобальных экологических проблем. Развитые страны и страны с переходной экономикой могут совместно осуществлять проекты по снижению выбросов ПГ в атмосферу. Предусмотрен также механизм «торговли квотами» на выбросы. Обязательство той или иной страны не должны превысить в среднем за 2008-2012 гг. определенный уровень как наличие у страны общенациональной квоты на выбросы – разрешения на выброс, равный обязательствам. Даже страны, не

ратифицировавшие Киотский протокол, начинают все больше осознавать опасность изменения климата, что приводит к региональным инициативам по сокращению выбросов. ЕС и другим развитым странам необходимо оставаться во главе движения, снизив уровень своих выбросов к 2020 г. до 30% ниже показателя 1990 г. с перспективой дальнейшего сокращения выбросов на 60-80% к 2050 г.

Зависимость загрязнения окружающей среды от экономического роста

Ведет ли экономический рост к загрязнению окружающей среды? Эмпирические исследования показывают [Самуэльсон, 2008], что с повышением дохода загрязнение, как правило, следует инверсной U-образной кривой. Чем выше доходы страны, тем больше, как правило, средств инвестируют в строительство очистных сооружений, а структура экономики эволюционирует в сторону сферы услуг, постепенно избавляясь от тяжелой и химической промышленности, что также способствует снижению загрязнения окружающей среды. Поскольку в работе исследуются вопросы, связанные с выбросами ПГ в атмосферу, авторы попытались исследовать эту зависимость на примере выбросов основного парникового газа. International Energy Annual 2006 дает информацию «World Per Capita Carbon Dioxide Emissions from the Consumption and Flaring of Fossil Fuels, 1980-2006» по всем странам мира. Исследовав зависимость этих данных в зависимости от валового национального продукта на душу населения (доклад ООН "State of World Population 2008" р.р. 90-93 , показатель GNI Per Capita PPP\$ 2006), авторы получили зависимость, представленную на рис.1.

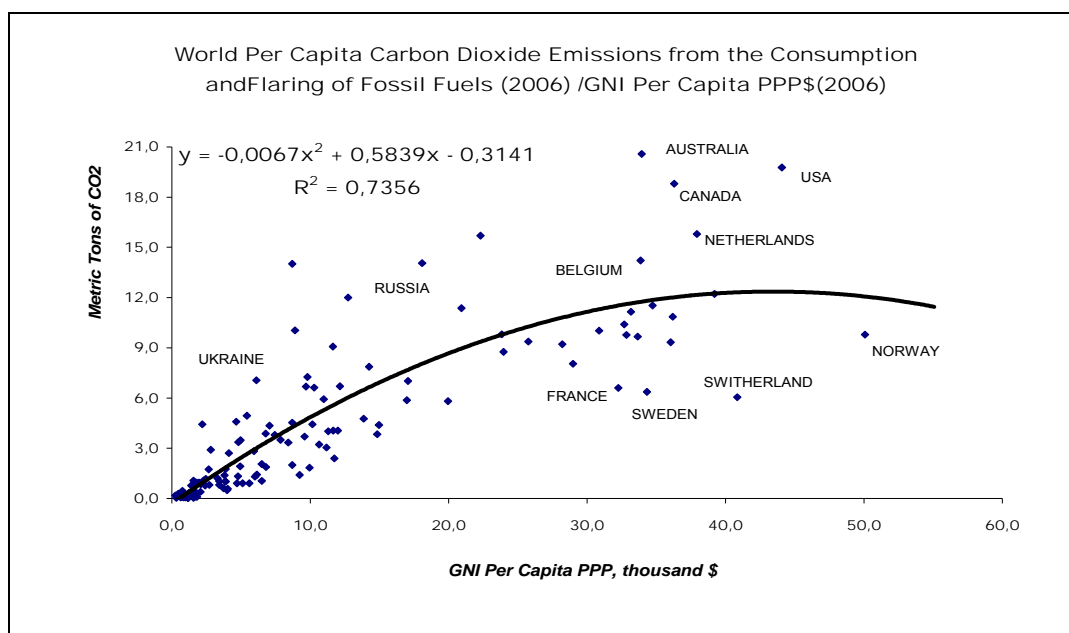


Рис. 1. Зависимость объемов выбросов от ВВП на душу населения

Этим графиком также подтверждается вид так называемой кривой Кузнеца. Безусловно, нельзя рассчитывать на полную адекватность этого графика, поскольку понятие загрязнения намного шире, чем выбросы одного парникового газа.

Реализация механизмов Киотского протокола

По различным оценкам, ежегодные объемы превышения выбросов ПГ над обязательствами (т.е. спроса на квоты) в 2008-2012 гг. оцениваются суммарно примерно в 150 млн. тонн CO₂ для стран ЕС и 300-400 млн. тонн CO₂ для Японии Канадцы, Новой Зеландии, Норвегии. Большинство развитых стран, скорее

всего, столкнется с невозможностью выполнения своих количественных обязательств по снижению выбросов парниковых газов исключительно за счет реализации внутренней политики и мер. Повышение энергоэффективности и снижение углеродоемкости производства на предприятиях промышленно развитых стран в настоящее время представляется либо дорогостоящим, либо вовсе невозможным из-за отсутствия в современном мире более совершенных технологий. По количеству выбросов в 2006 г. страны ЕС на 8% превышают квоты Киотского протокола. Не все страны ЕС смогут самостоятельно снизить выбросы ПГ до квот КП к 2012г. Европейское агентство по окружающей среде [EEA Report, 2008] констатирует, что Евросоюзу в составе 12 стран, так называемая группы EU-15 для выполнения обязательств по Киотскому протоколу необходимо снизить выбросы ПГ на 8%. И это снижение будет достигнуто не только за счет внутренней политики, но и посредством привлечения механизмов КП. Ситуация по отдельным странам разнится в значительной степени разниться. Все страны ЕС можно условно разбить на три группы: 1-ая группа – страны, которые, ни в , ни в будущем не смогут выполнить обязательства по КП; 2-ая группы – страны, которые в 2006 г. не достигли обязательств, но к 2012 г. выполняют взятые обязательства; 3-я группа – страны, которые уже в 2006 г. имеют выбросы, ниже обязательств по КП (табл.1).

Таблица 1.

	Страны, которые в 2006 г не достигли обязательств, и к 2012 г. тоже не смогут выполнить	Страны, которые в 2006 г не достигли обязательств, но к 2012 г. выполняют взятые обязательства	Страны, которые уже в 2006 г. имеют выбросы, ниже обязательств по КП
EU-15	Denmark, Italy, Spain	Austria, Belgium, Finland, Germany, Ireland, Luxembourg, Netherlands, Portugal	France, Greece, UK, Sweden
EU-12		Slovenia	Bulgaria, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovak Republic
EU candidate countries			Croatia
Other EEA member countries		Iceland, Lichtenstein, Norway	Switzerland

Совсем недавно, в ноябре 2008 г. Украина реализовала квоты на выбросы парниковых газов. 1.46 миллиона тонн условных выбросов было передано Японии. Это первое в мире соглашение такого масштаба. Договор реализуется двумя равными траншами в 2009 и 2010 гг.

Распределение квот должно (и по-видимому, в какой-то мере – неизвестной авторам – реализуется) осуществляться с учетом иерархической структуры интенсивности загрязнений регионов, их технологического потенциала, плотности населения и т.п. Предположим, что выделяются 5 уровней иерархии, представленных на рис.2.

Пусть Киотским протоколом определен объем выбросом, например, для стран Евросоюза (ЕС). С учетом того, что некоторые страны превышают свои квоты на выбросы, а некоторые недобирает допустимые объемы (см. табл.1), возникает вопрос - каким образом должны распределяться квоты среди государств, входящим в ЕС? Каким образом квоты на выбросы должны распределяться между промышленными районами Украины – Донбассом и Карпатским регионом? Эти проблемы (на любом уровне иерархии) предлагаем формализовать моделью «распределения затрат» в следующей постановке [Волошин, 2006].

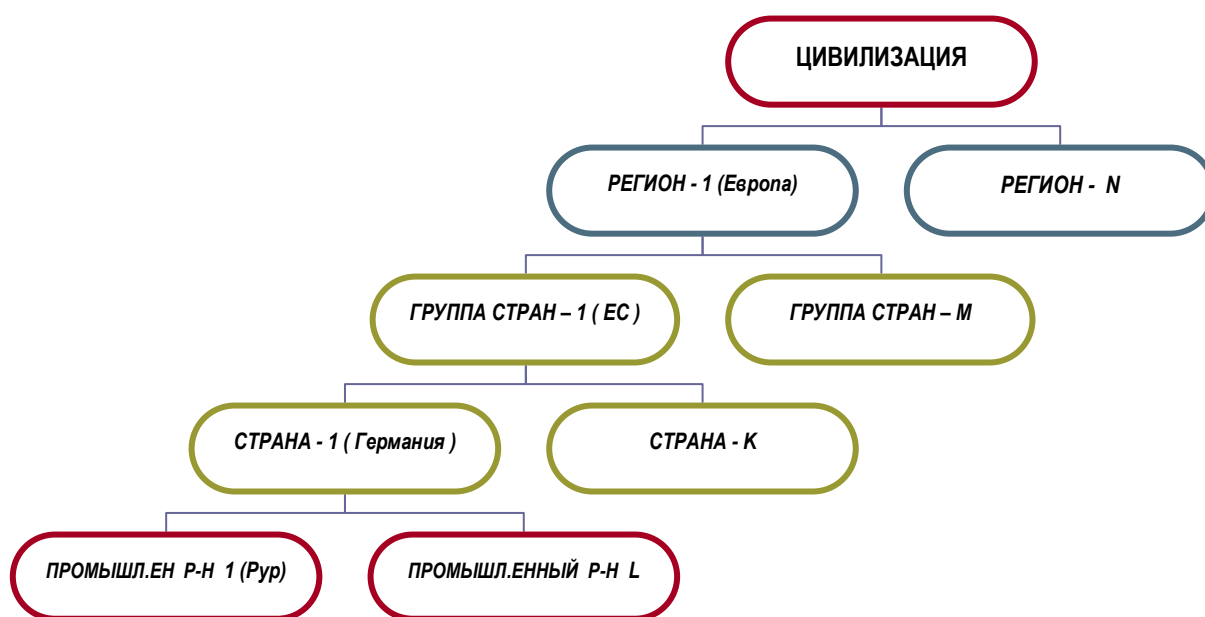


Рис. 2. Иерархия распределения квот

Имеется $n, n \geq 2$, агентов; $c, c > 0$, - их общие затраты; $b_i, b_i \geq 0$, - «потенциальный» доход i -го агента от эксплуатации коллективного объекта, созданного в результате вложения c единиц затрат.

Предполагается, что кооперация эффективна: $\sum_{i=1}^n b_i > c$.

Возникает две проблемы: 1. Как формировать «потенциальные» доходы каждого агента? 2. Какие механизмы распределения затрат использовать?

Рассмотрим вначале вторую проблему. Три основных способа распределения сводятся к следующему:

1. Выравнивание затрат: затраты i -го агента $x_i = c/n, i = \overline{1, n}$;

2. Выравнивание прибыли: $b_i - x_i = b_j - x_j = \frac{\sum_i b_i - c}{n}$ для $\forall i \neq j$;

3. Пропорциональное распределение: $x_i = \frac{b_i}{\sum_i b_i} c, \forall i$.

Каждый из приведенных способов наряду с очевидными преимуществами обладает недостатками, приводящими к распаду коалиции [Волошин, 2006]. Так, первый способ может привести к существованию агента, затраты которого могут превысить его доход ($\exists k : x_k > \frac{c}{n}$) и он выйдет из коалиции. Второй способ может привести к существованию дотируемого агента ($\exists k : x_k < 0$) и его не воспримет коалиция. Третий способ может привести к ничтожной прибыли некоторого агента и он выйдет из коалиции (без которого она не сможет покрыть расходы величиной c).

Модификация обсуждаемых способов распределения осуществляется двумя способами. На распределение затрат вводится ограничение вида $0 \leq x_i \leq b_i, i = \overline{1, n}$, что приводит к так называемым подушному и уровневому налогам, которые также не лишены недостатков [Волошин, 2006].

Второй способ состоит в сведении задачи распределения к кооперативной игре (N, S) , где для любой

коалиции $S, S \subseteq N$, вводится прибыль $v(S) = \max\left(\sum_{i \in S} b_i - c, 0\right)$ и предполагается выбирать

распределение из «ядра» игры, которая задается системой неравенств:

$\sum_{i \in S} x_i \geq v(S), \forall S \subseteq N, \sum_{i \in N} x_i = v(N)$. Это условие означает, что любой коалиции не выгодно

«выделяться». Однако, в этом случае остается проблема выбора единственного решения из ядра, в [Волошин, 2006] предлагается эгалитарное распределение, так называемое « N -ядро». Применяя утилитарный принцип распределения, учитывающий вклад агента в каждую коалицию, приходим к так называемому распределению (вектору) Шепли, которое может не принадлежать ядру игры.

Применительно к задачам распределения квот авторы утверждают, что, во-первых, основные значения имеют принципы формирования «потенциальных» доходов агентов $b_i, i = \overline{1, n}$; во-вторых, главное здесь не свойства получаемого решения (например, N -ядра или вектора Шепли), а именно принципы распределения (например, принцип равных расходов или равной прибыли), поскольку все человечество равно ответственно за экологию Земли и в равной мере (или «почти в равной мере») страдает от загрязнения окружающей среды.

При формировании «потенциального дохода» предлагается учитывать два эффекта влияния на загрязнение окружающей среды: 1. ВВП на душу населения; 2. Выбросы ПГ на 1 кв.км. и плотность населения. Очевидно, что плата за выбросы уменьшает величину ВВП, а с другой стороны, увеличение выбросов происходит за счет увеличения объемов производства. Таким образом, может быть выгодным одновременно платить за выбросы и наращивать объемы производства. Если плата за дополнительные выбросы будет меньше, чем объемы прироста ВВП, то это будет означать, что стране выгоднее заплатить за наращивание выбросов ПГ, чем сокращать производство. Так, если США присоединятся к КП, они обязаны выплачивать сумму, которая в 4 раза превышает их «пропорциональную долю» в выбросах [Самуэльсон, 2008]. Страна может также устанавливать дополнительное ограничение на эту плату (% от ВВП = G), $\alpha G, 0 < \alpha < 1$. Поскольку действие КП продолжается до 2012 года, страна может устанавливать верхнюю границу оплаты за превышение выбросов ПГ. Пусть C_i – «пропорциональная выплата» i -ой страны. Тогда $C_i^* = \beta C_i, \beta \leq 1$, – верхняя граница оплаты и при формировании дохода необходимо учитывать величины $b_i^* = C_i - \gamma C_i = (1 - \gamma)C_i$, где $\gamma, 0 \leq \gamma \leq 1$, – коэффициент «положительного эффекта» от платы за выбросы. Тогда доход i -ой страны, учитывающий эффект влияния ВВП per capita на загрязнение окружающей среды равен $b_i' = \min\{\alpha G, C_i^*\}$.

Аналогично, при формировании доли дохода b_i'' , учитывающего эффект влияния удельных выбросов и плотности населения в стране, предлагается учитывать положительный эффект от уменьшения удельных выбросов и негативный эффект от уменьшения ВВП (уменьшение объема производства и увеличение расходов на экологию и здравоохранение).

Тогда «потенциальный» доход i -го агента $b_i = \delta b_i' + (1 - \delta)b_i''$, где $\delta, 0 < \delta < 1$, - весовой коэффициент, учитывающий особенности конкретной страны или региона.

Выводы

Одним из основных недостатков большинства международных договоров об охране окружающей среды является отсутствие конкретных механизмов их реализации, в первую очередь, формализованных правил распределения ответственности, в частности, финансовой, которую можно более или менее точно просчитать. На основе этого прогноза агенты соглашения могут принимать соответствующие решения о кооперации. Об отсутствии подобных механизмов свидетельствует и отказ США ратифицировать Киотский протокол.

В развитие предлагаемых в докладе механизмов распределения квот целесообразно рассмотреть нечеткие постановки моделей распределения коллективных затрат, поскольку параметры, введенные при формировании индивидуальных «потенциальных доходов» агентов являются эмпирическими, а значит неточными, нечеткими, неопределенными.

Благодарности

Статья частично финансированна из проекта **ITHEA XXI** Института Информационных теории и Приложений FOI ITHEA и Консорциума FOI Bulgaria (www.ithea.org, www.foibg.com).

Литература

[Самуэльсон, 2008] Самуэльсон П., Нордхаус В. Экономика.–М.: Вильямс, 2008.–1360с.

[Волошин, 2006] Волошин О.Ф., Мащенко С.О. Теорія прийняття рішень.– К.: ВПЦ «Київський ун-т», 2006.– 304 с.

[EEA Report, 2008] Greenhouse Gas Emission Trends and Projections in Europe 2008.Tracking Progress towards Kyoto Targets. Executive Summary // www.eea.com

Информация об авторах

Волошин Алексей – доктор технических наук, профессор, Киевский национальный университет имени Тараса Шевченко, Украина, 01017 Киев, ул. Владимирская, 64; e-mail: ovoloshin@unicyb.kiev.ua

Горицына Ирина– кандидат экономических наук, старший научный сотрудник; Киевский национальный университет имени Тараса Шевченко, Украина, 01017 Киев, ул. Владимирская, 64; e-mail: goritsyna@mail.ru

In memoriam



Prof. Dr Vassil Vassilev, 1946 - 2009

It is our sad duty in the present issue to announce the sudden and unexpected death of Prof. Dr. Vassil Vassilev. Prof. Vassilev, a member of the Editorial board, passed away on 4th of February 2009.

This heavy loss is not only of a good colleague and a friend, but also of a prominent scholar in the Bulgarian and international field of Cybernetics and Information Science.

Prof. Dr. Vassil Vassilev was born on 11th of November 1946 in the village of Lopjan, Sofia district. He graduated in 1970 from the Wroclaw Polytechnic University, Poland, with a Master degree in Automation. In 1973 Prof. Dr. Vassilev earned a second Master degree in Applied Mathematics at Wroclaw University, Poland, and in 1976 he defended Doctor's Thesis in Operations Research at Wroclaw University, Poland.

During the period 1970-1972, Prof. Dr. Vassil Vassilev worked as a research associate at the Institute of Engineering Cybernetics of Bulgarian Academy of Sciences; then until 1984 at the Institute of Engineering Cybernetics and Robotics of Bulgarian Academy of Sciences (IECR – BAS). In 1985, he raised to the post of an Associate Professor at IECR - BAS. Later on, until 1993, he was an Associate Professor at the Institute of Informatics of Bulgarian Academy of Sciences.

Since 1994, Prof. Dr. Vassil Vassilev was the head of the Decision Support Systems department as well as the Director of the Institute of Information Technologies (IIT - BAS). In 2006, he earned a full Professorship position at IIT-BAS.

Prof. Dr Vassil Vassilev is one of the founders of Engineering Cybernetics and Operations Research in Bulgaria. He conducted rigorous research in the areas of discrete optimization, linear and non-linear continuous optimization, as well as in the development of decision support systems. The research work resulted in the development of efficient approximate methods algorithms and programs for solving certain classes of discrete problems. This software delivered great practical results and economic value. Prof. Dr Vassil Vassilev was a prominent participant in the development of the software for the first Bulgarian automatic control systems.

For more than 25 years of research activity, Prof. Dr Vassil Vassilev was devoted to his work in the area of Multiple Criteria Decision Making (MCDM). His research and tutorial work at IIT-BAS and

some of the biggest universities in Bulgaria has created a scientific school, that has conducted intensive scientific research in the area of MCDM. His team had achieved notable results in linear, nonlinear and discrete Multiple Criteria Decision Making. His unremitting work, the legacy of his teaching, as well as the application of new methods of classification and of the generalized scalarizing problem, earned him international acknowledgement and respect.

Prof. Dr Vassil Vassilev actively participated in international and local conferences and editorial boards. He also taught a number of lecture courses at the Sofia University “St. Kliment Ohridski”, the New Bulgarian University, the Technical University – Sofia, and the Bourgas Free University. He successfully tutored 9 Ph D students as well as 35 post-graduate students. The professional success of his students is also a testament what an outstanding teacher Prof. Dr. Vassil Vassilev was.

Prof. Dr. Vassil Vassilev authored more than 180 articles in international and Bulgarian journals and in the proceedings of international and local conferences, 4 patents and more than 200 quotations in international publications.

Prof. Dr Vassil Vassilev was the founding force behind the creation of the journal “Cybernetics and Information Technologies” and proved to be the most active member of its Editorial Board.

Our departed colleague Prof. Dr. Vassil Vassilev, with his devoted work and profound scientific knowledge, developed IIT - BAS to a nationally and internationally recognized institution in the field of Information Technologies. He will be remembered as a kind, gentle and generous person, an inspirational leader, a respected colleague and a reliable friend. He will be greatly missed by all of us, who were fortunate enough to have worked with and known him.

His untimely death leaves us with a large profound space.

The Editorial Board

