

Vitalii Velychko, Olexii Voloshyn, Krassimir Markov  
(editors)

**XX-th International Conference**  
**“Knowledge-Dialogue-Solution”**



**Proceedings**

**I T H E A**

**2014**

**ITHEA International Scientific Society**

**XX-th International Conference**

**Knowledge-Dialogue-Solution**

**September 8-10, 2014, Kyiv (Ukraine)**

**P R O C E E D I N G S**

**ITHEA<sup>®</sup>**

**Kyiv - Sofia, 2014**

**Olexii Voloshyn, Vitalii Velychko, Krassimir Markov (eds.)**

**Proceedings of the XX-th International Conference “Knowledge-Dialogue-Solution”**

ITHEA®

2014, Kyiv, Ukraine, Sofia, Bulgaria,

ISSN 1313-0087 (printed)

ISSN 1313-1206 (online)

ITHEA IBS ISC No.: 31

First edition

Printed in Ukraine

The XX-th International Conference “Knowledge-Dialogue-Solution” (KDS 2014) continues the series of annual international KDS events organized by Association of Developers and Users of Intelligent Systems (ADUIS).

The conference is traditionally devoted to discussion of current research and applications regarding three basic directions of intelligent systems development: knowledge processing, natural language interface, and decision making.

Edited by:

Association of Developers and Users of Intelligent Systems, Ukraine

Institute of Information Theories and Applications FOI ITHEA, Bulgaria

Publisher: ITHEA®

Sofia-1090, P.O.Box 775, Bulgaria

e-mail: [office@ithea.org](mailto:office@ithea.org)

[www.ithea.org](http://www.ithea.org)

All Rights Reserved

© 2014 ITHEA®, Bulgaria - Publisher

© 2014 Association of Developers and Users of Intelligent Systems, Ukraine - Co-edition

© 2014 Institute of Information Theories and Applications FOI ITHEA, Bulgaria - Co-edition

© 2014 Vitalii Velychko, Olexii Voloshyn, Krassimir Markov - Editors

© 2014 Krassimira B. Ivanova - Technical editor

© 2014 For all authors in the issue

**ISSN 1313-0087 (printed)**

**ISSN 1313-1206 (online)**

## PREFACE

The scientific XX<sup>th</sup> International Conference "Knowledge-Dialogue-Solution" (KDS) took place in September, 08-10, 2014 in Kyiv, Ukraine. It continues the series of international scientific meetings, which were initiated more than twenty years ago. The KDS conference is traditionally devoted to discussion of current research and applications regarding three basic directions of intelligent systems development: knowledge processing, natural language interface, and decision making.

Now we can affirm that the international conferences "Knowledge-Dialogue-Solution" in a great degree contributed to preservation and development of the scientific potential in the East Europe.

We would like to express our sincere thanks to everybody who helped to make conference success.

*Vitalii Velychko, Olexii Voloshyn, Krassimir Markov*

### CONFERENCE ORGANIZERS

- ITHEA International Scientific Society
- V.M.Glushkov Institute of Cybernetics of National Academy of Sciences of Ukraine
- Association of Developers and Users of Intelligent Systems (Ukraine)
- National Academy of Sciences of Ukraine
- Institute of Information Theories and Applications FOI ITHEA (Bulgaria)
- Taras Shevchenko National University of Kiev (Ukraine)
- Kharkiv National University of Radio Electronics (Ukraine)
- School of Information Technologies and Cybersecurity of V.S. Martinovsky Institute of Refrigeration Cryotechnologies and Ecoenergetics at Odessa National Academy of Food Technologies (Ukraine)
- Educational-scientific complex «Institute for applied system analysis» (Ukraine)
- Institute of Mathematics and Informatics, BAS (Bulgaria)
- Institute of Mathematics of SD RAN (Russia)

### STEERING COMMITTEE

Ivan Sergienko(Ukraine), Alexander Palagin (Ukraine), Iurii Kryvonos (Ukraine), Natalia Pankratova (Ukraine), Krassimir Markov (Bulgaria), Vitalii Velychko (Ukraine), Olexii Voloshyn (Ukraine), Volodymyr Romanov (Ukraine)

### PROGRAM COMMITTEE

Alexander Reznik (Ukraine)	Nadezhda Kiseleyova (Russia)	Vitaliy Snituk (Ukraine)
Albert Voronin (Ukraine)	Natalia Filatova (Russia)	Vladimir Donchenko (Ukraine)
Anatoly Krissilov (Ukraine)	Nikolay Zagoruiko (Russia)	Vladimir Donskoy (Ukraine)
Anatoliy Gupal (Ukraine)	Olga Nevzorova (Russia)	Vladimir Jotsov (Bulgaria)
Constantine Gaidric (Moldova)	Rumyana Kirkova (Bulgaria)	Vladimir Ryazanov (Russia)
Galina Gaivoronskaya (Ukraine)	Tatyana Gavrilova (Russia)	Vladislav Shelepov (Ukraine)
Ekaterina Solovyova (Ukraine)	Vadim Vagin (Russia)	Xenia Naidenova (Russia)
Iliya Mitov (Bulgaria)	Valery Tarasov (Russia)	Yevgeniy Bodyanskiy (Ukraine)
Igor Gorban (Ukraine)	Vassil Sgurev (Bulgaria)	Yury Valkman (Ukraine)
Leonid Hulianytskyi (Ukraine)	Vitaliy Lozovskiy (Ukraine)	Yury Zaychenko (Ukraine)

### ORGANIZING COMMITTEE

Oleksandra Skripnik, Kirill Malahov, Anatoliy Zayonchkovski, Natalia Kosach, Vitalii Velychko.

## TABLE OF CONTENTS

<i>Preface</i> .....	3
<i>Table of Contents</i> .....	4
<i>Index of Authors</i> .....	7
Аксиоматика Армстронга: повнота, критерій повноти та незалежність	
<i>Дмитро Буй, Анна Пузікова</i> .....	9
Гештальты в процессах образного мышления	
<i>Юрий Валькман</i> .....	13
Концепція аналізу та прийняття рішень при моделюванні сталого розвитку національної економіки	
<i>Олексій Волошин, Володимир Кудін, Володимир Кулик</i> .....	17
Анализ свойств нечётких обобщений методов распределения	
<i>Алексей Волошин, Василий Лавер</i> .....	21
Дистрибуция и развёртывание программной системы "ИКОН" с применением современных технологий виртуализации и облачных сервисов	
<i>Виталий Величко, Кирилл Малахов</i> .....	25
Спосіб автоматизованого виділення відношень між термінами з природномовних текстів технічної тематики	
<i>Віталій Величко, Віталій Приходнюк</i> .....	27
Медицинская грид-система для накопления и обработки электрокардиограмм	
<i>Виталий Вишнеевский</i> .....	29
Багатопараметричний вибір найбільш доцільних маршрутів ліній доступу з використанням моделі балансних мереж	
<i>Галина Гайворонська, Світлана Сахарова</i> .....	31
Феномен статистической устойчивости	
<i>Игорь Горбань</i> .....	35
Прогнозирование социально-экономического развития с помощью ассоциативных сетей	
<i>Валентин Григорьевский</i> .....	38
Прогнозирование на основе генетического алгоритма обучения	
<i>Леонид Гуляницкий, Анна Павленко</i> .....	41
Повышение эффективности алгоритма омп для прогнозирования третичной структуры протеинов методами распараллеливания	
<i>Леонид Гуляницкий, Виталина Рудьк</i> .....	43
К вопросу об особенностях построения знаниеориентированных социальных сетей в интернете в сфере обучения	
<i>Андрей Данилов</i> .....	48
Матричная множественная регрессия	
<i>Владимир Донченко, Инна Назарага, Ольга Тарасова</i> .....	50

---

Разработка хранилища данных для информационно-ресурсного центра <i>Анна Жолнарская</i> .....	54
Прогнозирование риска банкротства корпораций в условиях неопределенности с использованием нечетких нейронных сетей <i>Юрий Зайченко, Ови Нафас Агаи аг Гамиш</i> .....	57
Анализ и автоматическое выделение мимических проявлений жестового языка в видеопотоке <i>Юрий Крак, Антон Тернов, Владислав Кузнецов</i> .....	61
Метод линейной полосной классификации сложноразделимых дактилем украинского жестового языка <i>Юрий Крак, Григорий Кудин, Инга Соломянюк</i> .....	63
Соответствие системы управления потребностям объекта управления <i>Виталий Косс</i> .....	66
Автоматическое построение терминологических онтологий <i>Дмитрий Ландэ, Андрей Снарский</i> .....	68
Деякі задачі оцінки платоспроможності підприємств у сфері кредитування <i>Микола Маляр, Володимир Поліщук</i> .....	72
От физических отражений материального мира к основаниям информатики и математики <i>Анатолий Мерзвинский</i> .....	74
Алгебри квазіарних відношень та їх властивості <i>Нікітченко М.С., Шкільняк С.С., Матвіюк Д.А.</i> .....	78
Моделирование рыночных механизмов в агентно-ориентированной модели функционирования экономики <i>Диана Омелянчик</i> .....	79
Анализ существующих классификаций компетенций <i>Юлия Панасовская, Максим Вороной</i> .....	81
Разработка методов семантического поиска web-сервисов на основе онтологического аннотирования <i>Юлия Рогушина</i> .....	83
Кластерный анализ: проблема адекватности <i>Ирина Рясная</i> .....	85
Автоматизированная система агрегации, поиска, визуализации данных <i>Михаил Соколов, Сергей Николаев</i> .....	87
О следующем этапе компьютерной семантики <i>Владимир Сторож</i> .....	90
О применении системологии в трансдисциплинарных исследованиях <i>Екатерина Соловьева</i> .....	92
Классификация биоэлектрических сигналов на основе нейроподобных моделей <i>Наталья Н. Филатова, Дмитрий М. Ханеев</i> .....	94

Применение шаблонов при проектировании онтологий для извлечения знаний из правовых документов	
<i>Екатерина Хала</i> .....	96
Проект создания электронного мозга интеллектуальных роботов (ЭМИР)	
<i>Виталий Яценко</i> .....	99
A Formal Semantics of Additional Operations of Multiset Table Algebra	
<i>Dmitriy Buy, Iryna Glushko</i> .....	103
General Approach for the Classification Problem Solution	
<i>Mykola Budnyk</i> .....	105
Ontology Driven Context Based Decision-Making	
<i>Yuriy Chaplinsky, Olena Subbotina</i> .....	109
Framework to Manage Scrum Meeting Artifacts	
<i>Elena Chebanyuk</i> .....	111
Example of Multi-Layer Knowledge Representation by Means of Natural Language Addressing	
<i>Krassimira Ivanova</i> .....	115
Ontology Using in Geometrical Models Data Processing of Complex Technical Object	
<i>Dmytro Konotop</i> .....	118
Global System for Monitoring, Diagnostic and Managing Objects, as an Integral Part of the Information Society of the Future	
<i>Alexander Kosianchuk</i> .....	120
Energy Versus Information	
<i>Krassimir Markov</i> .....	122
A Computer Dialog with Wind Power Gives the Non-Emergency Wind Energy	
<i>Valery Pisarenko, Julia Pisarenko, Vladimir Malachinskij</i> .....	126
Wireless Sensor Network for Precision Farming and Environmental Protection	
<i>Volodymyr Romanov, Igor Galelyuka, Oleksandr Voronenko</i> .....	128
Research of the Fluorometer «Floratest» Sensitivity Impact of Stress Factors on the Plants	
<i>Volodymyr Romanov, Volodymyr Hrusha, Oleksandra Kovyrova</i> .....	130
Architecture of Smart Sensor Network for Agricultural Purposes	
<i>Yevgenia Sarakhan</i> .....	132
Joint Study of Visual Perception Mechanism and Computer Vision Systems that Use Coarse-To-Fine Approach	
<i>Anton Sharypanov, Alexandra Antoniouk, Vladimir Kalmykov</i> .....	134
Matrix-Evolutionary Method for Timetable Forming	
<i>Vitaliy Snytyuk, Olena Sipko</i> .....	136
A Hierarchical Approach to Multicriteria Problems	
<i>Albert Voronin, Yuriy Ziatdinov, Igor Varlamov</i> .....	138
A Hibrid Technologies to Optimize Constrained Eq Model	
<i>Olha Yegorova</i> .....	142
Построение процесса удаленного интеллектуального обучения	
<i>Александр Бабенко</i> .....	144

## INDEX OF AUTHORS

### Names in Russian and Ukrainian

Дмитро	Буй	9	Сергей	Николаев	87
Александр	Бабенко	144	Микола	Нікітченко	78
Юрий	Валькман	13	Віталій	Приходнюк	27
Олексій	Волошин	17, 21	Анна	Павленко	41
Максим	Вороной	81	Инна	Назарага	50
Виталий	Величко	25, 27	Ови	Нафас Агаи аг Гамиш	57
Виталий	Вишневский	29	Виталина	Рудык	43
Галина	Гайворонська	31	Світлана	Сахарова	31
Валентин	Григорьевский	38	Инга	Соломянюк	63
Леонид	Гуляницкий	41, 43	Ольга	Тарасова	50
Игорь	Горбань	35	Антон	Тернов	61
Андрей	Данилов	48	Диана	Омельянчик	79
Владимир	Донченко	50	Анна	Пузікова	9
Анна	Жолнарская	54	Юлия	Панасовская	81
Юрий	Зайченко	57	Володимир	Поліщук	72
Юрий	Крак	61, 63	Юлия	Рогущина	83
Владислав	Кузнецов	61	Ирина	Рясная	85
Володимир	Кулик	17	Михаил	Соколов	87
Володимир	Кудін	17	Андрей	Снарский	68
Григорий	Кудин	63	Владимир	Сторож	90
Віталій	Косс	66	Екатерина	Соловьева	92
Василий	Лавер	21	Наталья	Филатова	94
Дмитрий	Ландэ	68	Екатерина	Хала	96
Кирилл	Малахов	25	Дмитрий	Ханеев	94
Микола	Маляр	72	Степан	Шкільняк	78
Дмитро	Матвіюк	78	Виталий	Яценко	99
Анатолий	Мержвинский	74			



---

**INDEX OF AUTHORS****Names in English**

Alexandra	Antoniouk	134
Dmitriy	Buy	103
Mykola	Budnyk	105
Yuriy	Chaplinskyy	109
Elena	Chebanyuk	111
Igor	Galelyuka	128
Iryna	Glushko	103
Volodymyr	Hrusha	130
Krassimira	Ivanova	115
Vladimir	Kalmykov	134
Dmytro	Konotop	118
Alexander	Kosianchuk	120
Oleksandra	Kovyrova	130
Krassimir	Markov	122
Vladimir	Malachinskij	126
Valery	Pisarenko	126
Julia	Pisarenko	126
Volodymyr	Romanov	128, 130
Oleksandr	Voronenko	128
Yevgenia	Sarakhan	132
Anton	Sharypanov	134
Olena	Sipko	136
Vitaliy	Snytyuk	136
Olena	Subbotina	109
Albert	Voronin	138
Igor	Varlamov	138
Olha	Yegorova	142
Yuriy	Ziatdinov	138

## АКСІОМАТИКА АРМСТРОНГА: ПОВНОТА, КРИТЕРІЙ ПОВНОТИ ТА НЕЗАЛЕЖНІСТЬ

Дмитро Буй, Анна Пузікова

**Анотація:** Дана робота є узагальненням отриманих авторами результатів з побудови строгого та повного доведення повноти аксіоматики Армстронга щодо функціональних залежностей в реляційних базах даних через встановлення збіжності відношень синтаксичного та семантичного слідувань, встановлення критерію повноти аксіоматики Армстронга в термінах потужностей множини атрибутів і універсального домену та побудови доведення незалежності її складових.

**Ключові слова:** аксіоматика Армстронга, повнота, незалежність, критерій повноти.

### Вступ

Огляд наукової та навчальної літератури показав, що в ній відсутні обґрунтування відомих результатів в теорії реляційних баз про повноту аксіоматики Армстронга та незалежність її компонентів, які б задовольняли стандартним вимогам до строгого математичного доведення. Відсутній також критерій повноти аксіоматики Армстронга. Результати щодо обґрунтування повноти аксіоматики Армстронга та критерію повноти наведені нижче без доведень (але з посиланнями на роботи авторів, що містять відповідні доведення), встановлення незалежності компонентів аксіоматики Армстронга наводиться з доведенням.

### Повнота аксіоматики Армстронга

Всі неозначувані тут поняття та позначення використовуються в розумінні монографії [Редько, 2001].

Нехай  $t$  – таблиця,  $R$  – схема таблиці  $t$  (скінченна множина атрибутів),  $X, Y, Z$  – підмножини схеми  $R$ ,  $s, s_1, s_2$  – рядки таблиці  $t$ .

На таблиці  $t$  виконується функціональна залежність (ФЗ) (див., наприклад, [Редько, 2001])  $X \rightarrow Y$ , якщо для двох довільних рядків  $s_1, s_2$  таблиці  $t$ , які збігаються на множині атрибутів  $X$ , має місце їх рівність і на множині атрибутів  $Y$ , тобто:

$$(X \rightarrow Y)(t) = \text{true} \stackrel{\text{def}}{\Leftrightarrow} \forall s_1, s_2 \in t (s_1 | X = s_2 | X \Rightarrow s_1 | Y = s_2 | Y).$$

В цьому означенні  $s | X$  – обмеження рядка  $s$  як функції на множину атрибутів  $X$ .

Таблиця  $t$  схеми  $R$  є моделлю множини ФЗ  $F$ , якщо кожна ФЗ  $X \rightarrow Y \in F$  виконується на таблиці  $t$ :

$$t \text{ модель } F \stackrel{\text{def}}{\Leftrightarrow} \forall (X \rightarrow Y) (X \rightarrow Y \in F \Rightarrow (X \rightarrow Y)(t) = \text{true}).$$

ФЗ  $X \rightarrow Y$  семантично слідує ( $|=$ ) з множини ФЗ  $F$ , якщо на кожній таблиці  $t(R)$ , яка є моделлю множини ФЗ  $F$ , виконується також ФЗ  $X \rightarrow Y$ :

$$F \models X \rightarrow Y \stackrel{\text{def}}{\Leftrightarrow} \forall t(R) (t \text{ модель } F \Rightarrow (X \rightarrow Y)(t) = \text{true}).$$

Вище запис  $t(R)$  означає, що таблиця  $t$  має схему  $R$ .

Доведення нижченаведених лем 1-3 та їхніх наслідків 1-3 можна знайти в [Буй, 2011].

**Лема 1** (аксіома рефлексивності Армстронга).  $\forall t (X \rightarrow Y)(t) = \text{true}$ , де  $Y \subseteq X$ .  $\square$

**Наслідок 1.**  $\emptyset \models X \rightarrow Y$  для  $Y \subseteq X$ .  $\square$

**Лема 2** (правило поповнення Армстронга).  $\forall t ((X \rightarrow Y)(t) = \text{true} \Rightarrow (X \cup Z \rightarrow Y \cup Z)(t) = \text{true})$  для  $Z \subseteq R$ .  $\square$

**Наслідок 2.**  $F \models X \rightarrow Y \Rightarrow F \models X \cup Z \rightarrow Y \cup Z$  для  $Z \subseteq R$ .  $\square$

**Лема 3** (правило транзитивності Армстронга).  $\forall t ((X \rightarrow Y)(t) = \text{true} \& (Y \rightarrow Z)(t) = \text{true} \Rightarrow (X \rightarrow Z)(t) = \text{true})$ .  $\square$

**Наслідок 3.**  $F \models X \rightarrow Y \& F \models Y \rightarrow Z \Rightarrow F \models X \rightarrow Z$ .  $\square$

Функціональна залежність  $X \rightarrow Y$  синтаксично слідує ( $| -$ ) з множини ФЗ  $F$  ( $F | - X \rightarrow Y$ ), якщо існує скінченна послідовність ФЗ  $\varphi_1, \varphi_2, \dots, \varphi_{m-1}, \varphi_m$ , така, що  $\varphi_m = X \rightarrow Y$  і  $\forall i = \overline{1, m-1}$  кожна ФЗ  $\varphi_i$  є або аксіомою рефлексивності, або належить множині ФЗ  $F$ , або отримана за яким-небудь правилом виведення (поповнення, транзитивності) з попередніх у цій послідовності ФЗ  $\varphi_j, \varphi_k$ ,  $j, k < i$  [Ліндон, 1968].

Нехай задана деяка множина ФЗ  $F$ . *Замикання*  $[F]$  – це множина усіх ФЗ, які синтаксично слідують з  $F$ :

$$[F] \stackrel{def}{=} \{X \rightarrow Y \mid F | - X \rightarrow Y\}$$

**Лема 4.** Виконуються властивості (див., наприклад, [Мейер, 1987]):

- 1)  $F \subseteq [F]$  (зростання);
- 2)  $[[F]] = [F]$  (ідемпотентність);
- 3)  $F \subseteq G \Rightarrow [F] \subseteq [G]$  (монотонність).  $\square$

Таким чином, за термінологією [Скорняков, 1982] оператор  $F \mapsto [F]$  є оператором замикання.

Нехай задана деяка множина ФЗ  $F$ , і нехай  $\mathfrak{F}_F$  – сімейство усіх множин ФЗ  $G$ , які містять  $F$ , таких, що при застосуванні до них аксіом та правил виведення Армстронга не можна отримати жодної функціональної залежності, яка б не належала  $G$ :

$$\mathfrak{F}_F \stackrel{def}{=} \{G \mid F \subseteq G \wedge [G] \subseteq G\}.$$

Сім'я  $\mathfrak{F}_F$  непорожня, вона містить, наприклад, множину всіх ФЗ  $\{X \rightarrow Y \mid X \subseteq R \wedge Y \subseteq R\}$ .

**Лема 5.** Сім'я  $\mathfrak{F}_F$  замкнена відносно довільних перетинів.  $\square$

Нехай  $F^* \stackrel{def}{=} \bigcap_{G \in \mathfrak{F}_F} G$  – перетин всіх множин ФЗ з сім'ї  $\mathfrak{F}_F$ .

**Наслідок 4.**  $F^* \in \mathfrak{F}_F$ .  $\square$

Сім'я ФЗ  $\mathfrak{F}_F$  є муровською сім'єю в розумінні абстрактної теорії решіток (див., наприклад, [Біркгоф, 1984]).

**Лема 6.** Виконується рівність:  $[F] = F^*$ .  $\square$

**Наслідок 5.** Замикання  $[F]$  – це найменша (в розумінні включення  $\subseteq$ ) множина, що містить  $F$ , така, що при застосуванні до неї аксіом та правил виведення Армстронга не можна отримати жодної функціональної залежності, яка б не належала  $[F]$ .

Доведення лем 5-6 та їхніх наслідків 4-5 можна знайти в [Буй, 2011].

*Замиканням*  $[X]_F$  множини  $X$  (відносно множини ФЗ  $F$ ) називається об'єднання правих частин всіх ФЗ вигляду  $X \rightarrow Y$ , які слідує (синтаксично) з множини  $F$ :

$$[X]_F \stackrel{def}{=} \bigcup_{X \rightarrow Y \in [F]} Y.$$

**Лема 9.** Виконуються властивості.

- 1)  $X \subseteq [X]$ ;
- 2)  $F | - X \rightarrow [X]$ ;
- 3)  $X \rightarrow Z \notin [F] \Rightarrow Z \notin [X] \wedge [X] \subset R$ .  $\square$

**Твердження 1** (коректність аксіомати Армстронга). Якщо ФЗ  $X \rightarrow Y$  синтаксично виводиться з множини ФЗ  $F$ , то ФЗ  $X \rightarrow Y$  виводиться з  $F$  семантично:

$$F | - X \rightarrow Y \Rightarrow F \models X \rightarrow Y. \square$$

*Доведення* проводиться індукцією за довжиною доведення (див. [Буй, 2011]).

**Твердження 2** (повнота аксіоматики Армстронга). Якщо ФЗ  $X \rightarrow Y$  семантично виводиться з множини ФЗ  $F$ , то ФЗ  $X \rightarrow Y$  синтаксично виводиться з  $F$ :

$$F \models X \rightarrow Y \Rightarrow F \vdash X \rightarrow Y. \square$$

Твердження про повноту аксіоматики Армстронга наводиться в [Мейер, 1987].

**Теорема 1.** Відношення семантичного та синтаксичного слідування збігаються:

$$F \models X \rightarrow Y \Leftrightarrow F \vdash X \rightarrow Y. \square$$

### Критерій повноти аксіоматики Армстронга в термінах потужностей універсального домена та множини атрибутів

Детальний аналіз наведеного доведення основного результату про збіжність відношень  $\models$  та  $\vdash$  показує, що воно проведено в припущенні:  $|D| \geq 2$  та  $|R| \geq 2$ , тобто універсальний домен містить не менше 2 елементів та схема  $R$  має щонайменше 2 атрибути. Розгляд випадків, коли  $|D| \leq 1$  або  $|R| \leq 1$  здійснено у роботі [Буй, TAAPSD'2011], основним результатом якої є наступна теорема, що містить уточнені умови, в межах яких аксіоматика Армстронга є повною.

**Теорема.** Відношення  $\models$  та  $\vdash$  співпадають тоді і тільки тоді, коли  $|D| \geq 2$  або  $|R| = 0$ .  $\square$

Доведення наведено у роботі [Буй, TAAPSD'2011].  $\square$

### Незалежність аксіоматики Армстронга

**Лема 10.** Аксіома рефлексивності є незалежною від правил виведення аксіоматики Армстронга.  $\square$

Доведення є тривіальним, оскільки інших аксіом в аксіоматиці Армстронга немає.  $\square$

**Лема 11.** Правила транзитивності є незалежними від аксіом рефлексивності і правил поповнення.  $\square$

Доведення. Для доведення достатньо вказати множину ФЗ  $F$ , для якої  $[F]_{\vdash-tr} \subset [F]_{\vdash}$ , де  $[F]_{\vdash-tr}$  – замикання множини ФЗ  $F$ , побудоване за допомогою аксіом рефлексивності і правил поповнення, а  $[F]_{\vdash}$  – замикання множини ФЗ  $F$  відносно всіх аксіом і правил виведення.

Зафіксуємо множини атрибутів  $X, Y, Z$ , причому  $X \cap Z = \emptyset$ , і розглянемо множину ФЗ  $F$ , яка задовольняє умовам  $\forall Y \subseteq R (X \rightarrow Y \in F \wedge Y \rightarrow Z \in F \Rightarrow X \rightarrow Z \notin F)$ .

Покажемо, що ФЗ  $X \rightarrow Z$  не може бути виведена лише із застосуванням аксіом рефлексивності та правил поповнення.

Оскільки  $Z \cap X = \emptyset$ , то  $Z \not\subseteq X$ , отже, ФЗ  $X \rightarrow Z$  не є тривіальною.  $\square$

Припустимо, що ФЗ  $X \rightarrow Z$  отримана з деякої ФЗ  $U \rightarrow V$  за правилом поповнення, тобто існує така множина атрибутів  $W \neq \emptyset$ , що  $U \cup W = X$  і  $V \cup W = Z$ . Тоді  $Z \cap X = (V \cup W) \cap (U \cup W) \subseteq W \neq \emptyset$ , що суперечить умові  $Z \cap X = \emptyset$ .

Таким чином, для вказаної множини ФЗ  $F$  ФЗ  $X \rightarrow Z$  не може бути отримана за аксіомами рефлексивності і правилами поповнення, отже,  $[F]_{\vdash-tr} \subset [F]_{\vdash}$  і правило транзитивності є незалежним.  $\square$

**Лема 12.** Правило поповнення є незалежним від аксіом рефлексивності і правил транзитивності.  $\square$

Доведення. Розглянемо множину ФЗ  $F = \{X \rightarrow \{A\} \mid A \notin X \wedge X \neq \emptyset\}$ . Покажемо, що ФЗ вигляду  $X \rightarrow X \cup \{A\}$ , де  $A \notin X \wedge X \neq \emptyset$  не можна вивести лише із застосуванням аксіом рефлексивності і правил транзитивності.

Припустимо, що ФЗ  $X \rightarrow X \cup \{A\} \in F$ , тоді її права частина є одноелементною множиною. Останнє виконується лише в наступних випадках:

- $A \in X$  і  $X = \{A\}$ ;
- $A \notin X$  і  $X = \emptyset$ .

Кожен з зазначених випадків суперечить означенню множини ФЗ  $F$ . Отже,  $X \cup \{A\}$  не є одноелементною множиною і ФЗ  $X \rightarrow X \cup \{A\}$  не належить множині  $F$ .

Оскільки  $A \notin X$ , то ФЗ  $X \rightarrow X \cup \{A\}$  не є тривіальною.

Припустимо, що ФЗ  $X \rightarrow X \cup \{A\}$  отримана за правилами транзитивності. Враховуючи, що

- множина тривіальних ФЗ замкнена відносно правила транзитивності ([Буй, ТАAPSD'2011], лема 3);
- результатом застосування правила транзитивності до ФЗ з одноелементною правою частиною є ФЗ, права частина якої є знову одноелементною множиною;
- результатом застосування правила транзитивності до тривіальної ФЗ  $X \rightarrow X'$  і ФЗ з одноелементною правою частиною є ФЗ з одноелементною правою частиною;
- результатом застосування правила транзитивності до ФЗ виду  $X \rightarrow \{A\}$  і тривіальної ФЗ  $\{A\} \rightarrow \{A\}$  є ФЗ виду  $X \rightarrow \{A\}$ ,

приходимо до висновку, що ФЗ  $X \rightarrow X \cup \{A\}$  повинна бути або тривіальною або її права частина є одноелементною множиною, що суперечить означенню цієї ФЗ.

Таким чином, для вказаної множини ФЗ  $F$  ФЗ  $X \rightarrow X \cup \{A\}$  не може бути отримана за аксіомами рефлексивності і правилами транзитивності, отже, правило транзитивності є незалежним.  $\square$

**Теорема 2.** Аксіоматика Армстронга є незалежною.  $\square$

Доведення впливає безпосередньо з лем 10, 11 і 12.  $\square$

---

## Висновки

В роботі отримані наступні результати щодо аксіоматики Армстронга в теорії (однозначних) функціональних залежностей реляційних баз даних: повнота та коректність аксіоматики Армстронга в термінах збіжності відношень синтаксичного та семантичного слідувань (Теорема 1), критерій повноти аксіоматики Армстронга в термінах потужностей універсального домена  $D$  та множини атрибутів  $R$  (Теорема 2; зауважимо, що множини  $R, D$  виступають параметрами аксіоматичної системи), взаємна незалежність складових аксіоматики Армстронга (Теорема 3; складові – це аксіоми рефлексивності, а також правила поповнення та транзитивності). Все це говорить про створення фрагменту теорії нормалізації (з приводу сучасного стану теорії нормалізації в реляційних базах даних див. огляд авторів [Буй, 2014]).

---

## Література

- [Біркгоф, 1984] Г. Биркгоф. Теория решеток. – Москва: Наука, 1984.
- [Буй, 2011] Д. Б. Буй. Повнота аксіоматики Армстронга. Д. Б. Буй, А. В. Пузікова // Вісник Київського національного університету імені Тараса Шевченка, Серія: Фізико-математичні науки, випуск № 3, 2011.
- [Буй, 2014] Д. Б. Буй. Огляд теорії нормалізації в реляційних базах даних. Д. Б. Буй, А. В. Пузікова // Матеріали міжнародної конференції «КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ В НАУКОЕМКИХ ТЕХНОЛОГИЯХ – КМНТ-2014» (Харьков, 28-31 мая 2014 года).
- [Буй, ТАAPSD'2011] Д. Б. Буй. Критерій повноти аксіоматики Армстронга. Д. Б. Буй, А. В. Пузікова // Матеріали міжнародної конференції «Теоретичні та прикладні аспекти побудови програмних систем» – ТАAPSD'2011 (Україна, Ялта, 19-23 вересня 2011 року).
- [Ліндон, 1968] Р. Линдон. Заметки по логике. – Москва: Мир, 1968.
- [Мейер, 1987] Д. Мейер. Теория реляционных баз данных. – Москва: Мир, 1987.
- [Редько, 2001] В.Н.Редько. Реляційні бази даних: табличні алгебри та SQL-подібні мови. В.Н. Редько, Ю.Й. Брона, Д.Б. Буй, С.А. Поляков. Київ: Академперіодика, 2001.
- [Скорняков, 1982] Л.А. Скорняков. Элементы теории структур. – Москва: Наука, 1982.

---

## Інформація про авторів

**Дмитро Буй** – д-р фіз.-мат. наук, професор, Київський національний університет ім. Т. Шевченка, Київ, Україна, e-mail: [buy@unicyb.kiev.ua](mailto:buy@unicyb.kiev.ua)

**Анна Пузікова** – аспірантка, Київський національний університет ім. Т. Шевченка, Київ, Україна, e-mail: [anna\\_inf@mail.ru](mailto:anna_inf@mail.ru)

---

## ГЕШТАЛТЫ В ПРОЦЕССАХ ОБРАЗНОГО МЫШЛЕНИЯ

Юрий Валькман

**Аннотация:** Исследовано определение гештальта, его свойства. Рассмотрены некоторые критерии их классификации. Рассмотрены понятия априорных и апостериорных гештальтов. Показано, что они являются необходимым атрибутом семиозиса и основой моделирования мышления в вычислительной среде. Гештальты «участвуют» не только в процессах восприятия, но и во многих других процессах нашего мышления и жизнедеятельности. Классифицировать их целесообразно по различным критериям. Они обладают многими интересными структурными свойствами, что позволит нам разрабатывать их модели для представления операций мышления в вычислительной среде.

**Ключевые слова:** Образ, мышление, гештальт, семиозис, семиотика, когнитивные процессы

---

### 1. Введение

Данная работа является продолжением исследований процессов образного мышления. Объектом этих исследований является моделирование процессов образного мышления. Здесь, предмет исследования – гештальты в процессах образного мышления. Цель исследования – разработка моделей представления структур образной информации и методов их использования в практической деятельности. С нашей точки зрения, проблемы представления образной информации и моделирования образного мышления тесно связаны с когнитологией и семиотикой. Основными ресурсами в когнитивных семиотических операциях являются гештальты, а базовыми процессами – семиозисы. Здесь акцентируем внимание на гештальтах. Нас интересуют проблемы синтеза и анализа гештальтов, структур их хранения и использования.

---

### 2. Об определении гештальта

Гештальт основное понятие гештальтпсихологии. Он обозначает *цельные, несводимые к сумме своих частей, образования сознания*. Понятие о гештальте зародилось при изучении процессов восприятия и структурирования ощущений человека. Примеры гештальта: инсайт; восприятие мелодии, не сводимое к сумме ощущений звуков этой мелодии, неразложимые на части запахи, вкусовые ощущения и т. д. Возникнув в психологии, сегодня термин «*гештальт*» используется также для обозначения некоторых целостных и неделимых физических, физиологических и социальных феноменов.

В буквальном переводе с немецкого Gestalt означает «*форма, вид, фигура*». Другие более или менее подходящие по смыслу русские аналоги — «*целостность*», «*структура*», «*система*», «*образ*», «*архетип*» и «*модель*». Еще И. Кант говорил о том, что мы не можем воспринимать физический мир непосредственно. Человек всегда взаимодействует с информацией, полученной от органов чувств — дорабатывая ее в своем сознании. Таким образом, любое целое для нас больше суммы его частей, потому что мы вкладываем в него и свое восприятие. Образно говоря, когда мы смотрим на картину Ренуара или другого импрессиониста, мы не замечаем и не оцениваем каждый мазок, а видим единое целое и именно это общее сочетание цветов и форм нас впечатляет.

Чтобы ухватить смысл гештальта, *невозможно его однозначно сопоставить с чем-то уже понятным, нужно синтезировать в своей голове новое содержательное понятие*. Поэтому, четкого определения этого понятия не существует. Мы попытаемся приблизиться к смыслу слова "гештальт". В данной работе лишь определены некоторые концепции будущих исследований. Здесь мы хотим показать, что, во-первых, гештальты «участвуют» не только в процессах восприятия, но и во многих других процессах нашего мышления и жизнедеятельности. Поэтому, можно классифицировать их по различным критериям. И, во-вторых, они обладают многими интересными структурными (системными) свойствами, что позволит нам разрабатывать их модели для представления операций мышления в вычислительной среде.

---

### 3. Образ и гештальт

---

Для нас приложение методов и моделей гештальтов и семиозиса к *образам* и *образному мышлению* представляет интерес по следующим причинам.

- *Во-первых*, образ, как и гештальт всегда является отражением некоторого объекта (материального, ментального, знакового). Вопрос о тождественности образов и гештальтов мы рассмотрим потом. Здесь рассматриваются только те образы, которые имеют денотат, концепт и знак. Или мы хотим на основе некоторых из этих компонент синтезировать остальные для данного образа.

- *Во-вторых*, уже давно известно разделение образов на две категории: *объективные* (существующие в реальности внешнего, по отношению к человеку, мира) и *субъективные* (ментальные, существующие в мыслях человека).

Ментальные образы психологи делят на три класса: образы *восприятия*, *представления* и *воображения*. Соответственно, можно говорить о четырех формах представления знаний. В этой работе мы коснемся отношений и взаимодействия этих четырех классов образов, точнее закодированных в них знаний. Заметим, поскольку в образах отражаются некоторые знания, информация, то образ можно считать *моделью представления знаний*. Представляется, что отношения между всеми четырьмя разновидностями образов (моделей знаний) мы можем рассматривать только на уровне их знаковых структур. Тем более это касается образных операций (интерпретации образов, их синтеза, анализа, объединения и т.д.)

- *В-третьих*, фиксация, систематизация и передача знаний невозможна без их представления в некоторой форме. А эта форма всегда имеет знаковую структуру. Фактически «кодирование-декодирование» знаний осуществляется в системе «*знак - значение*».

- *В-четвертых*, построение интеллектуальных компьютерных технологий предполагает погружение в вычислительную среду знаний с целью их практического использования в решении сложных проблем. А это невозможно без представления этих структур знаний в символической (знаковой) форме. Более того, компьютер должен имитировать работу со смыслами; особо это касается моделирования образного мышления. Известно высказывание Ч. С. Пирса "*мы думаем только в знаках*". Мы бы его расширили так: «*И мыслим только на основе гештальтов*». Таким образом, необходимо смыслы (гештальты) сопоставить с синтаксическими структурами (знаками).

---

### 4. О классификации гештальтов

---

Рассмотрим некоторые признаки классификации гештальтов.

1. Тот внутренний врожденный (или встроенный в автомат) интенциональный образ, который позволяет опознавать определенные внешние объекты без предварительного научения, только благодаря правильной работе рецепторов, назовем *априорным гештальтом* (термин Г. П. Мельникова). Наличие априорного гештальта позволяет не только обеспечивать отнесение внешних отражаемых объектов к определенному универсальному множеству, но и ускорять процесс выработки новых интенциональных образов для различения подмножеств этого множества и даже индивидуальных его представителей. Такие образы можно бы назвать *апостериорными гештальтами*. Без априорных гештальтов невозможно замыкание цепочек семиозиса. Естественно предположить, что, всегда, любой акт опознания (в семиозисе) должен завершаться возбуждением априорного гештальта.

2. Очевидна соотнесенность априорного гештальта с конкретными рецепторами например, обоняние, зона вкусовой рецепции и т.д. В чувственном образе может быть воплощено любое содержание. Поэтому мы можем говорить о: *зрительных гештальтах*, *слуховых*, *тактильных*, *мышечных*, *вестибулярных*, *вкусовых*, *обонятельных* и т. д.

3. По аналогии с образами мы можем рассматривать: *гештальт восприятия* — отражение в идеальном плане внешнего объекта (сцены), воздействующего на органы чувств, *гештальт представления* — отражение (вспоминание) объекта без его наличия (сенсорного контакта с ним) и *гештальт воображения* — вымышленный образ, данный в представлении, но не имеющий аналогов в реальной действительности и, поэтому, никогда ранее не воспринимавшийся.

4. Кроме этого, могут быть гештальты различных комбинаций чувственных образов. Дж. Лакофф предполагает что: *мысли, восприятия, эмоции, процессы познания, моторная деятельность и язык* организованы с помощью одних и тех же структур, которые он называет гештальтами.

5. Рассмотрим классификатор Пирса. Он выделяет следующие множества знаков: *иконы, индексы, иконические индексы и символы*. Заметим, здесь классифицируются не знаки, а их отношения к гештальтам.

6. Известна классификация знаков Ч. Морриса. На основе разных способов обозначения он выделяет пять типов знаков: *знаки-идентификаторы* (т. е. те, которые отвечают на вопрос "где?"); *знаки-десигнаторы* (знаки, отвечающие на вопрос "что такое?"); *оценочные* (связанные с предпочтением, отвечающие на вопрос "почему?"); *прескриптивные* (отвечающие на вопрос "как?"); *формирующие, или знаки систематизации* (направляющие поведение интерпретатора в отношении других знаков). А так как знак становится знаком, только тогда, когда осуществлена его интерпретация (Пирс), т.е. соотнесение данного знака с «его гештальтом» то и здесь исследуются отношения между знаками и гештальтами.

7. Р. Уфимцев предлагает рассматривать три класса гештальтов.

- *Врожденные гештальты* - каждый человек рождается с некоторым набором унаследованных гештальтов. Иными словами, его восприятие и внимание изначально оказывается способным распознавать в мире некоторые целостные конфигурации.
- *Базовые приобретённые гештальты* - результат воспитания, обучения и социальной адаптации ребенка. Естественно, что каждая культура, даже каждая географическая местность обогащает человека собственным уникальным набором гештальтов.
- *Индивидуально приобретённые гештальты* – результат личного жизненного опыта, поэтому эта часть когнитивного словаря может существенно различаться у разных людей. Обычно индивидуальное расширение когнитивного словаря происходит в направлении социальных гештальтов и профессиональных.

Далее рассмотрим и другие классификаторы гештальтов, в частности, анализ гештальта как метафоры..

## **5. Некоторые свойства гештальтов**

Как известно, *цельность* – это антипод *структуры*, она представима только как гештальт и, будучи антиподом структуры, *гештальт не имеет составных частей*. Дж. Лакофф, отказываясь от определения гештальта, рассматривает некоторые его свойства. С нашей точки зрения, по-сути, это свойства системы, т. е. гештальт рассматривается как система, с соответствующей структурой

1. Гештальты являются одновременно целостными и анализируемыми. Они состоят из частей, но не сводимы к совокупности этих частей. Гештальты обладают определенными дополнительными свойствами, и их части могут располагать дополнительными свойствами, если рассматривать их в рамках гештальта в целом. Содержание восприятия вообще гештальтно; оно таково изначально, а не в результате акта соединения отдельных ощущений. Однако бесспорно и то, что в восприятие любого предмета входит многое такое, что само может быть предметом самостоятельного восприятия.

2. Гештальты разложимы более чем одним способом, в зависимости от принятой точки зрения. Это означает, что возможны различные способы «правильного» разложения гештальта на части.

3. Части гештальта связаны внутренними отношениями, которые бывают различных классов (типов – у Лакоффа). Способ отношений частей в рамках гештальта входят в содержание самого гештальта.



4. Гештальт может быть связан внешними отношениями с другими гештальтами. Он может составлять часть другого гештальта или проецироваться на другой иным. При проецировании части одного гештальта отображаются на части другого гештальта. Результатом проецирования будет «наследование» частями второго гештальта свойств и отношений первого.
  5. Внешние соотношения гештальта с другими гештальтами являются свойством гештальта в целом. Эти внешние отношения также могут быть различными в зависимости от принятой точки зрения.
  6. Проецирование гештальта на другой гештальт может быть частичным. Гештальты могут также иметь частичное сходство друг с другом, степень сходства зависит от близости связи между ними.
  7. Гештальт, включенный в качестве части в другой гештальт, может приобрести новые свойства.
  8. Одним из свойств гештальта может быть то, что он находится в отношении оппозиции к другому гештальту.
  9. Некоторые свойства гештальта могут выделяться как базисные свойства. Гештальты, находящиеся в отношении оппозиции, обычно имеют одинаковые базисные свойства.
  10. Гештальты - это структуры, используемые в процессах — языковых, мыслительных, перцептуальных, моторных или других.
  11. Процессы сами по себе можно рассматривать как гештальты.
  12. При анализе гештальтов не выдвигаются требования необходимости подразделения анализируемой сущности на элементарные части. Анализ, осуществляемый человеком, ограничен возможностями, намерениями и точкой зрения человека. При разных возможностях, намерениях и точках зрения результаты анализа будут различаться не только степенью детализированности, но и самими выделенными частями. Так, например, в семантике нет необходимости доводить анализ до элементарных, атомарных предикатов.
  13. Гештальты могут пересекаться.
  14. В гештальтах должны быть разграничены прототипические и непрототипические свойства.
  15. Гештальты часто относятся к смешанному типу. Это означает, что свойства гештальтов неоднородны. Так, сенсорно-моторные гештальты включают как перцептуальные, так и моторные свойства.
- В заключение Лакофф подчеркивает, что сказанное выше не является определением гештальта, а лишь дает представление об этом понятии (точнее, отражает точку зрения Лакоффа на Гештальт). Заметим, предлагаемое понятие гештальта весьма недоопределенно. И эта недоопределенность соответствует нашей концепции моделирования когнитивных процессов посредством компьютерных технологий. Далее мы рассмотрим и другие свойства гештальтов, точнее сущностей, которые они отражают.

---

## 6. Заключение

Исследовано определение гештальта, его свойства. Рассмотрены некоторые критерии их классификации. Гештальты «участвуют» не только в процессах восприятия, но и во многих других процессах нашего мышления и жизнедеятельности. Классифицировать их целесообразно по различным критериям. Они обладают многими интересными структурными свойствами, что позволит нам разрабатывать их модели для представления операций мышления в вычислительной среде. Полагаем, что это направление исследований должно обеспечить разработку формального аппарата для моделирования образного мышления в компьютерных технологиях.

---

## Информация об авторе

*Юрий Валькман – Международный научно-учебный центр информационных технологий и систем НАН и МОН Украины, 03680 ГСП, г. Киев, проспект акад. Глушкова, 40; e-mail: yur@valkman.kiev.ua*

---

## КОНЦЕПЦІЯ АНАЛІЗУ ТА ПРИЙНЯТТЯ РІШЕНЬ ПРИ МОДЕЛЮВАННІ СТАЛОГО РОЗВИТКУ НАЦІОНАЛЬНОЇ ЕКОНОМІКИ

**Олексій Волошин, Володимир Кудін, Володимир Кулик**

**Анотація:** Запропоновано концепцію аналізу та прийняття рішень при моделюванні процесів відтворення національної економіки за допомогою матриці соціальних рахунків, на основі якої будується математична модель у вигляді системи лінійних алгебраїчних рівнянь великої розмірності, для аналізу котрої використовується методологію послідовного аналізу варіантів.

**Ключові слова :** сталий розвиток, національна економіка, моделювання, матриця соціальних рахунків, метод базисних матриць, підтримка процесів відтворення, бюджетування.

---

### Вступ

Сучасне становище національної економіки потребує використання спеціального інструментарію для моніторингу, аналізу, розробки та актуалізації поточних та стратегічних завдань соціального і економічного відтворення. В міжнародній практиці загальнодержавного управління для цих цілей використовується система національних рахунків (СНР) [1], яка є як статистичним відображенням процесів відтворення національної економіки, так і інформаційно-аналітичною моделлю для аналізу та формування економічної політики, актуалізації окремих її аспектів (відтворювальних, галузевих, інституційних, регіональних тощо) [2]. Це в свою чергу, передбачає проведення аналізу моделі на основі математичних та інформаційних методів та алгоритмів, дослідження впливу зміни її параметрів з метою прийняття рішень щодо вибору їх значень для досягнення сталого розвитку економіки. Цілеспрямована підтримка макроекономічних пропорцій відтворення та подолання існуючих критичних обмежень (диспропорцій) сприятиме сталості системи доходів та реальному зростанню валового внутрішнього продукту (ВВП).

---

### Постановка проблеми

Враховуючи значну інерційність великих економічних систем (в першу чергу, світової та національної економіки), процес моделювання слід здійснювати «від досягнутого», де вихідною точкою для корегування економічної політики є саме соціально-економічна ситуація, що склалася на даний момент. Для сталих в соціально-економічному сенсі країн макроекономічні пропорції залишаються протягом тривалого часу практично незмінними і економічна політика, перш за все, спрямовується на підтримку сприятливої кон'юнктури, що підтримує дану пропорційність й надалі. Проте сучасні події вносять значні зміни в існуючі макроекономічні пропорції, що пов'язані із кардинальними змінами у зовнішньоекономічних відносинах (наприклад, пов'язаних із введенням санкцій), розгортанням торговельних війн із слабо прогнозованими наслідками та ін. Особливо негативно це впливає на невеликі за розміром недиверсифіковані економіки, що критично залежать від високотехнологічного імпорту і експорт яких складають сировинні товари.

В цих умовах необхідно зосередитися на розвитку внутрішнього ринку та потенціалу економічної системи.

Таким чином, підготовка та обґрунтування заходів економічної політики має ґрунтуватися як на детальному аналізі процесів соціально-економічного відтворення [1, 2], так і на передбаченні змін до національного бюджету в частині всього комплексу процесів відтворення, зумовлених поточним станом. Це потребує застосування такого інструментарію збалансування національного бюджету як матриці соціальних рахунків (МСР) (social accounting matrix), яка залежно від цілей може мати різну структуру.

### Інформаційна модель дослідження сталого розвитку

Для дослідження процесів сталого розвитку пропонується використовувати спрощену систему рахунків (модель) Петера фон дер Ліппе [Липпе, 1995].

Така модель дозволяє пов'язати різні суб'єкти економіки із притаманними їм економічними процесами та загальноекономічним процесом відтворення (див. таблицю нижче) :

- підприємства (інституційний сектор «нефінансових корпорацій» S.11 згідно [1]) виробляють на конкурентних засадах товари і послуги та здійснюють розподіл доходу таким чином, що це сприяє їх економічному зростанню;
- діяльність домогосподарств (інституційного сектору «домашніх господарств» S.14) спрямована на підвищення рівня життя громадян та задоволення їх потреб, всебічне стимулювання отримання доходу, раціональний його розподіл на поточні та стратегічні потреби, і відповідне використання;
- діяльність фінансової системи (інституційного сектору «фінансових корпорацій» S.12) забезпечує акумуляцію тимчасово вільних фінансових ресурсів (заощаджень) задля підтримки необхідного для економічного росту інвестиційного попиту;
- завдання ж державного сектору (інституційного сектору «сектор загального державного управління» S.13) та державного регулювання економіки у взаємоузгодженні різнопланових інтересів суб'єктів економіки, що й забезпечує в широкому розумінні соціальну справедливість та економічне зростання. Цей сектор в моделі не представлено.

Спрощена система рахунків процесів відтворення.

Таблиця.

	Використання		
	Нефінансові корпорації	Домашні господарства	Фінансові корпорації
Нефінансові корпорації	Виробництво і утворення доходу	<b>Споживання</b>	<b>Інвестиції</b>
Домашні господарства	<b>Кінцевий продукт</b>	Розподіл і використання наявного доходу	
Фінансові корпорації		<b>Заощадження</b>	Утворення капіталу

Діяльність кожного із суб'єктів економіки описується відповідним бюджетом, який відображає його основні функції і, на перший погляд, є автономним, непов'язаним з іншими бюджетами. Проте приведена система бюджетів тісно пов'язана системою інформаційних потоків, що вимагає як професійного і відповідального підходу кожного до виконання своїх інституційних функцій та формування власних доходів, так і зовнішнього регулювання в інтересах всіх учасників соціально-економічного відтворення.

Приведена модель відповідає принципам положенням економічної реформи в Україні, запропонованої урядом, і передбачає : фінансову децентралізацію; уніфікацію рахунків економічного (бухгалтерського) та податкового обліку; бюджетну автономію регіонів й інших суб'єктів економіки, стимулювання і вирівнювання доходів, формування довірчих взаємовідносин між учасниками соціально-економічного відтворення, взаємовигідну кооперацію через посилення горизонтальних зв'язків, притаманних ринковій економіці тощо.

### Математична модель дослідження сталого розвитку

Наведена вище модель (таблиця) може бути проінтерпретована як спрощена МСР. В математичному поданні вона може бути представлена як квадратна матриця  $A(1)$  відповідної розмірності, яка включає блоки рядків-стовпців (підматриць), що описують цілісну систему пов'язаних бюджетів:

$$\dot{A} = \begin{pmatrix} P & C & I \\ Y & D & 0 \\ 0 & S & K \end{pmatrix}, \quad (1)$$

де  $P, D, K$  – блоки (бюджети), що описують відповідно процеси виробництва і утворення доходу, розподілу і використання наявного доходу, утворення капіталу;  $Y, S, C, I$  – блоки, що описують формування відповідно кінцевого продукту, заощаджень, кінцевого споживання та інвестицій.

Вектор ресурсних та організаційних обмежень (фінансових, людських і т.д.) позначимо через  $B$ . Тоді зв'язок матриці  $A$  та вектора обмежень  $B$  можна подати як лінійну систему з шуканим 3-х компонентним вектором інтенсивностей відповідно виробництва, споживання та інвестування за допомогою системи лінійних алгебраїчних рівнянь (СПАР):

$$Au = B. \quad (2)$$

Якщо для початкового моменту розв'язок системи (2) при заданих  $A$  та  $B \in U_0$ , який відтворює поточний стан інтенсивностей при відповідних рівнях  $Y, S, C, I$ , то новий вектор інтенсивностей  $\bar{u}_0$  буде вказувати на еволюцію інтенсивностей в результаті сценарних змін в МСР (матриці  $A$ ). Такі зміни можуть бути зумовлені направленою еволюцією елементів матриці  $A$ , тобто цілеспрямованими змінами структури приведених бюджетів і відповідною системою утворюваних доходів  $Y, S$ .

Слід зазначити, що зміни в матриці  $A$  мають свої обмеження, які, з одного боку, визначаються соціальними умовами та економічними можливостями, а з іншого боку, обмеженнями інструментарію проведення аналізу впливу кількісних змін в моделі на якісні властивості параметрів. Математично врахування обмежень, вказаних вище, проводиться включенням додаткових обмежень в модель (2).

### Аналіз математичної моделі сталого розвитку

При аналізі властивостей лінійних систем типу (2) пропонується використовувати методологію послідовного аналізу варіантів (ПАВ) [Волошин, 2013] в його конкретній реалізації для аналізу СПАР у вигляді методу базисних матриць (МБМ) [Кудин, 2007].

Обґрунтування застосування ПАВ, зокрема, підтверджується результатами, отриманими в [Тинбэрхэн, 1967], щодо дослідження та управління великими економічними системами.

Моделювання складних економічних процесів лінійними моделями загалом не дають глобального розв'язку, а лише локальний. Це вказує на доцільність проведення короткострокового прогнозування змін (структури досліджуваних бюджетів та відповідних потоків), де дискретизація неперервного процесу проводиться на малі часові інтервали. Тоді процес побудови математичної моделі є багатокроковим (поетапним). Забезпечення виконання умов коректності побудови моделі обумовлює проведення аналізу властивостей та аналізу впливу кількісних змін вхідних параметрів моделі на розв'язок. Параметри коректної лінеаризації нелінійних зв'язків лінійними можуть бути досягнуті в класі слабонелінійних систем [Кудин, 2007] у вигляді (3).

$$\dot{A}(\cdot) = \begin{pmatrix} P(\cdot) & C(\cdot) & I(\cdot) \\ Y(\cdot) & D + \Delta D(\cdot) & 0 \\ 0 & S(\cdot) & K(\cdot) \end{pmatrix}, \quad (3)$$

де  $(\cdot) = (t, \lambda_p, \lambda_D, \lambda_K, \lambda_C, \lambda_I, \lambda_Y, \lambda_S)$ ,  $t$  – параметр часу,  $\lambda_p, \lambda_D, \lambda_K, \lambda_C, \lambda_I, \lambda_Y, \lambda_S$  – відповідно вектори параметрів, що впливають за нелінійні функціональні зміни в блоках матриць  $P, D, K, C, I, Y, S, \Delta D$  – описує зміни у розподілі та використанні кінцевого продукту.

Тоді формування дослідження основ сталого розвитку національної економіки буде ґрунтуватися на варіантах змін у матриці  $A$ , наприклад, щодо політики у сферах :

- *виробництва і утворення доходу* (шляхом розвитку конкуренції, зниження собівартості вітчизняної продукції і формування умов для інвестування за моделлю «зниження витрат → збільшення доходів → створення умов для інвестування», зміни системи регуляторної політики тощо);
- *розподілу та використання наявного доходу* (шляхом планування доходів, змін у формуванні і розподілі ВВП, у формуванні доходів та можливостях інвестування суб'єктів економіки та ін.);
- *утворення капіталу* (шляхом підтримки стійкості фінансового сектору та його ліквідності та ін.);
- *зовнішньоекономічної діяльності* (шляхом сприяння і розвитку експортного потенціалу в умовах міжнародного поділу праці та відкритості ринків, підтримки сприятливих внутрішніх і зовнішній макроекономічних пропорцій та сталості відповідної системи доходів та ін.).

---

## Висновки

Концепція аналізу та прийняття рішень при моделюванні сталого розвитку економічної системи полягає у: 1) математичному представленні інформаційної моделі Петера фон дер Ліппе у вигляді системи лінійних алгебраїчних рівнянь великої розмірності; 2) дослідженню зміни впливу параметрів моделі на її розв'язок; 3) застосуванню до аналізу чутливості моделі методології послідовного аналізу варіантів у його інтерпретації у вигляді методу базисних матриць.

---

## Бібліографія

- [1] System of National Accounts 2008. United Nations. – New York: UN, 2009, 722p. (ST/ESA/STAT/SER.F/2/REV.5).
- [2] Use of Macro Accounts in Policy Analysis. United Nations. – New York: UN, 2002, 333p. (ST/ESA/STAT/SER.F/81).
- [Ліппе, 1995] *Петер фон дер Ліппе*. Экономическая статистика: Статистические очерки – том 1 Европейского центра повышения квалификации. – Штуттгарт, Йена: ФСУ Германии, 1995. – 629 с.
- [Волошин, 2013] Волошин О.Ф., Гнатієнко Г.М., Кудін В.І. Послідовний аналіз варіантів: Технології та застосування: Монографія. – К: Стилос, 2013. – 304 с.
- [Кудін, 2007] Кудін В.И., Ляшко С.И., Хритonenко Н.В., Яценко Ю.П. Анализ свойств линейной системы методом искусственных базисных матриц // Кибернетика и системный анализ. 2007, № 4. – С. 119–127.
- [Тинбэрхэн, 1967] Тинбэрхэн Я, Бос Х. Математические модели экономического роста. – М.: Прогресс, 1967. – 176 с.

---

## Інформація про авторів

**Олексій Ф. Волошин** - Київський національний університет Тараса Шевченка, факультет кібернетики, д.т.н., професор, Україна; e-mail: olvoloshyn@ukr.net

**Володимир І. Кудін** - Київський національний університет Тараса Шевченка, факультет кібернетики, д.т.н., с.н.с., Україна; e-mail: V\_I\_Kudin@mail.ru

**Володимир В. Кулик** - Міжнародний науково-навчальний центр інформаційних технологій і систем НАНУ і МОНУ, Україна, к.е.н., с.н.с., Україна; e-mail: Volodymyr\_Kulyk@ukr.net

## АНАЛИЗ СВОЙСТВ НЕЧЁТКИХ ОБОБЩЕНИЙ МЕТОДОВ РАСПРЕДЕЛЕНИЯ

Алексей Волошин, Василий Лавер

**Аннотация:** Рассматривается применение теории нечётких множеств к классической задаче распределения. Предлагаются нечёткие обобщения классических методов распределения и дается анализ некоторых их свойств.

**Ключевые слова:** задачи распределения, нечеткие модели и методы, методы распределения.

### Введение

Проблемы распределения (ресурсов, затрат, прибыли) в современном глобализированном мире выходят на одно из ведущих мест в системе приоритетов цивилизации. Современный этап изучения проблем распределения начинается в 70-х годах прошлого столетия в связи с известными социально-экономическими потрясениями во многих странах мира. Одной из первых работ, посвященных данной тематике, была работа Б. О'Нила [O'Neill, 1982], в которой исследуются проблемы распределения наследства на базе материалов из Талмуда. Анализу проблем банкротства через призму теории кооперативных игр была посвящена и ключевая статья Р. Ауманна (лауреата премии им. Нобеля Академии наук Швеции по экономике за 2005 год) и М. Машлера «Теоретико-игровой анализ задачи о банкротстве из Талмуда» [Aumann, Maschler, 1985]. Среди других работ следует упомянуть также статью П. Янга [Young, 1988], в которой проблема распределения изучается через призму налогообложения. Методы распределения и их аксиоматическая характеристика исследуется и разрабатывается такими учеными как Э. Мулен [Moulin, 1991, 2001], К. Герреро [Herrero, Villar, 2001], В. Томпсон [Thomson, 2013] и другими, в том числе и [Voloshyn, Laver, 2013]. Современные результаты в моделировании распределения затрат (прибыли) изложены в обзоре В. Томпсона [Thomson, 2013].

### Постановка задачи

Задача распределения определяется тройкой  $(N, c, b)$  [Voloshyn, 2010], где  $N$  – конечное множество агентов, неотрицательное вещественное число  $c$  определяет количество ресурсов (затрат или прибыли), которые необходимо распределить, вектор  $b = (b_i)_{i \in N}$  определяет для каждого агента  $i$  его заявку  $b_i$

(запас денег), причем  $\sum_{i=1}^n b_i > 0$  и

$$0 \leq b_i, \forall i \in N : 0 \leq c \leq \sum_{i=1}^n b_i. \quad (1)$$

Решением задачи распределения является вектор  $x = (x_i)_{i \in N}$ , который ставит в соответствие каждому агенту  $i$  его долю  $x_i$ , причём

$$0 \leq x_i \leq b_i, \quad (2)$$

$$\forall i \in N : \sum_{i \in N} x_i = c. \quad (3)$$

Метод распределения определяется отображением  $r$ , которое каждой тройке  $(N, c, b)$  ставит в соответствие вектор распределения  $x = (x_i)_{i \in N}$ ,  $x = r(N, c, b)$ .

Основными методами распределения в «налоговой» интерпретации П. Янга [Voloshyn, 2010], [Thomson, 2013] являются следующие: 1) пропорциональный налог; 2) подушный налог (как обобщение выравнивания затрат); 3) уровневый налог (как обобщение выравнивания прибыли); 4) Талмудический метод (как обобщение подушного и уровневого налога, которое базируется на некоей задаче из Талмуда [Aumann, Maschler, 1985]); 5) «обратный» Талмудический метод (метод, двойственный к Талмудическому); 6) метод Пинила (распределение по этому методу совпадает с распределением по Талмудическому методу, если сумма запасов агентов меньше половины фонда распределения, и с распределением по обратному Талмудическому методу в иных случаях; этот метод менее известный, хотя предложен в талмудической литературе еще в 1861 году [Thomson, 2013]). Приведенные методы будем называть «чёткими», в противовес «нечётким», которые рассматриваются далее и которые базируются на теории нечётких множеств.

### Нечёткие обобщения методов распределения

При нечётком обобщении метода распределения затрат  $r(N, c, b)$  приходим к решению следующей параметрической (относительно параметра  $\lambda$ ,  $0 \leq \lambda \leq 1$ ) задачи [Voloshyn, Laver, 2013]:

$$\lambda \rightarrow \max, \quad (4)$$

$$\mu_k(x_k) \geq \lambda, \forall k \in N; \quad (5)$$

$$\sum_{i=1}^n x_i = c; \quad (6)$$

$$x_k \geq 0, k \in N. \quad (7)$$

Тут  $N_1, N_2$  ( $N_1 \cup N_2 = N$ ) – разбиение множества агентов, соответствующие доли затрат которых являются правосторонними нечёткими числами треугольного вида,  $x_i = (\underline{x}_i, \hat{x}_i), \forall i \in N_1$ , и  $x_j = (\hat{x}_j, \bar{x}_j), \forall j \in N_2$ ,  $x_i = (\underline{x}_i, \hat{x}_i)$ , с функциями принадлежности  $\mu_k, \forall k \in N$ , где  $\hat{x}_k = r(N, c, b)$  для всех  $k \in N$ ,  $\sum_{i=1}^n \hat{x}_i = c$ . Далее будем предполагать, что агенты упорядочены по увеличению их денежных запасов.

Пусть вектор  $(x_1, x_2, \dots, x_n; \lambda)$  является решением задачи (4)-(7). Тогда вектор  $(x_1, x_2, \dots, x_n)$  называется распределением затрат при нечётком обобщении метода  $r(N, c, b)$ , а параметр  $\lambda$  характеризует согласованность участников максимальной коалиции  $N$ .

В общем случае решение об отнесении конкретного агента к группе  $N_1$  или  $N_2$  остается на усмотрение лица, принимающего решения (ЛПР), хотя данный выбор может зависеть и от конкретного метода распределения.

Предлагается следующий алгоритм для поиска долей затрат агентов для нечёткого обобщения метода распределения  $r$ .

1. Вычисляются доли затрат агентов  $\hat{x}_i = r(N, c, b)$  по выбранному методу распределения  $r$ ;
2. Множество агентов по некоторому правилу делится на две подгруппы:  $N_1$  и  $N_2$ ;
3. Устанавливаются величины уступок для агентов из первой  $\alpha$  ( $i \in N_1$ ) и второй  $\beta$  ( $j \in N_2 = N \setminus N_1$ ) групп – на сколько процентов от величины  $\hat{x}_i$  агенты первой группы желают заплатить

меньше, а агенты второй –согласны заплатить больше величины, которую приписывает им метод распределения  $r$ . Находим  $\underline{x}_i = (1-\alpha)\hat{x}_i, i \in N_1$  и  $\bar{x}_j = (1+\beta)\hat{x}_j, j \in N_2$ .

4. Величины затрат агентов рассматриваются как правосторонние нечёткие числа треугольного вида, соответственно  $x_i = (\underline{x}_i, \hat{x}_i)$  с функцией принадлежности  $\mu_i(x_i)$  ( $i \in N_1$ ) и  $x_j = (\hat{x}_j, \bar{x}_j)$  ( $j \in N_2$ ) с функцией принадлежности  $\mu_j(x_j)$ ;

5. Формируется задача линейного программирования вида (4)-(7), решением которой является оптимальное распределение  $(x_1, x_2, \dots, x_n)$  и степень согласованности мнений агентов о коррекции долей затрат  $\lambda$  ( $0 \leq \lambda \leq 1$ ). Если полученные результаты не удовлетворяют ЛПР, переходим к п.3. Иначе работа алгоритма закончена.

Следует отметить, что решение  $(x_1, x_2, \dots, x_n; \lambda)$  задачи (4)-(7) всегда существует, причём  $\lambda \in (0, 1)$ . Более того, его можно найти аналитически [Voloshyn, Laver, 2014].

Поскольку функции принадлежности агентов являются правосторонними нечёткими числами, обозначим  $x_k^L$  левый конец интервала нечеткости соответствующего нечёткого числа, а через  $x_k^R$  - правый. Желаемая доля затрат агента  $k$  представляется нечётким числом  $(x_k^L, x_k^R), \forall k \in N$ .

Тогда

$$\lambda_0 = \mu_{\Sigma}(c) = \frac{\sum_{k=1}^n x_k^R - c}{\sum_{k=1}^n x_k^R - \sum_{k=1}^n x_k^L},$$

причем  $x_k = x_k^R - \lambda_0(x_k^R - x_k^L), \forall k \in N$ .

Для существования допустимого распределения затрат, необходимо выполнение неравенства

$$\alpha \sum_{i \in N_1} \hat{x}_i \leq \beta \sum_{j \in N_2} \hat{x}_j.$$

Предположим, что задан некоторый уровень  $\lambda_0$ . Тогда справедливым является неравенство

$$\lambda_0 = \frac{\beta \sum_{j \in N_2} \hat{x}_j}{\alpha \sum_{i \in N_1} \hat{x}_i + \beta \sum_{j \in N_2} \hat{x}_j} \geq \frac{\beta \sum_{j \in N_2} \hat{x}_j}{\beta \sum_{j \in N_2} \hat{x}_j + \beta \sum_{j \in N_2} \hat{x}_j} = \frac{1}{2}.$$

То есть,  $\frac{1}{2} \leq \lambda_0 < 1$  для любого допустимого распределения затрат, причём  $\lambda_0 = \frac{1}{2}$  только если

$$\alpha \sum_{i \in N_1} \hat{x}_i = \beta \sum_{j \in N_2} \hat{x}_j.$$

Для чётких методов распределения существует хорошо разработанная система аксиом, с помощью которых можно характеризовать эти методы. Для нечетких методов, в общем случае, данные характеристики не имеют места, хотя некоторые свойства («Равное отношение к равным», «Согласованность с нулём», «Независимость от шкалы», «Упорядочивание» и др..) характерны как чётким, так и нечётким методам [Voloshyn, Laver, 2014].



---

### Случай нечётких затрат

---

Пусть затраты  $c = (\underline{c}, \hat{c}, \bar{c})$  являются нечётким числом треугольного вида. Одним из вариантов решения данной задачи есть разбиение её на две подзадачи – задачу «оптимиста» (затраты «небольшие», т.е.  $c \in [\underline{c}, \hat{c}]$ ) и задачу «пессимиста» (затраты «большие», т.е.  $c \in [\hat{c}, \bar{c}]$ ). Для задачи оптимиста

$$c_0 = \frac{\hat{c} \sum_{i=1}^n x_i^R - \underline{c} \sum_{i=1}^n x_i^L}{\sum_{i=1}^n x_i^R - \sum_{i=1}^n x_i^L + \hat{c} - \underline{c}}. \quad (8)$$

Значения  $\lambda_0$  и  $x_k$ ,  $\forall k \in N$ , вычисляются по вышеприведенным формулам из расчета, что  $c = c_0$ .

Задача пессимиста, в свою очередь, сводится к поиску оптимального распределения для  $c_0 = \hat{c}$  при функциях принадлежности затрат агентов, рассчитанных для  $\bar{c}$ .

В качестве решения исходной задачи принимается вектор, для которого величина  $\lambda_0$  будет наибольшей.

---

### Выводы

Основной идеей нечётких обобщений методов распределения, которые предлагаются выше, является деление агентов на две группы таким образом, что агенты из первой группы платят меньше своей доли затрат, а агенты из второй группы покрывают этот дефицит. Данная идея фактически используется при прогрессивном налогообложении. Поскольку решение нечеткой задачи распределения существует и может быть найдено аналитически, то данные обобщения применимы в реальных условиях. Более того, они позволяют учитывать не только нечеткость данных, но и социальное неравенство агентов.

---

### Библиография

- [O'Neill, 1982] O'Neill B. A problem of rights arbitration from the Talmud. - Mathematical Sciences, 2, 1982. - pp. 345-371.
- [Aumann, Maschler, 1985] Aumann R. and Mashler M. Game theoretic analysis of a bankruptcy problem from the Talmud. - Journal of Economic Theory, 36, 1985, - pp. 195-213.
- [Young, 1988] Young, H.P. Distributive justice in taxation. - Journ. of Econ.Theory, 1988, - pp. 321-335.
- [Thomson, 2013] William Thomson. Axiomatic and game-theoretic analysis of bankruptcy and taxation problems: an update, Working Papers 578, University of Rochester, 2013. – pp.157-182.
- [Voloshyn, 2010] Voloshyn O., Mashchenko S. Decision making models and methods.. - K.: «Taras Shevchenko National University of Kyiv», 2010. – 336 с. (in Ukr.).
- [Herrero, Villar, 2001] Herrero, C. and Villar A. The Three Musketeers: four classical solutions to bankruptcy problems. - Mathematical Social Sciences, vol.42(3), 2001. - pp. 307-328.
- [Moulin, 2001] Moulin, H. Axiomatic Cost and Surplus-Sharing // Working Papers 2001-06, Rice University, Department of Economics, 2001. – 152p.
- [Moulin, 1991] Moulin, H. Axioms of cooperative decision making.- M: Mir, 1991. – 464p. (in Rus.).
- [Voloshyn, Laver, 2013] Voloshyn O., Laver V. Generalization of Distributing Methods for Fuzzy Problems // "Information Theories & Applications", Vol. 20, Number 4, 2013. – P. 303-310.
- [Voloshyn, Laver, 2014] Voloshyn O., Laver V. Axiomatic Characterization of the Fuzzy Rationing Methods // Bulletin of Taras Shevchenko National University of Kyiv. Series: Physics and Mathematics, № 2, 2014 (in Ukr.).

---

### Информация об авторах

**Алексей Волошин** – д.т.н., профессор, Киевский национальный университет имени Тараса Шевченко; e-mail: olvoloshyn@ukr.net

**Василий Лавер** – ассистент, Ужгородский национальный университет; e-mail: v.laver@gmail.com

---

## ДИСТРИБУЦИЯ И РАЗВЕРТЫВАНИЕ ПРОГРАММНОЙ СИСТЕМЫ “ИКОН” С ПРИМЕНЕНИЕМ СОВРЕМЕННЫХ ТЕХНОЛОГИЙ ВИРТУАЛИЗАЦИИ И ОБЛАЧНЫХ СЕРВИСОВ

Виталий Величко, Кирилл Малахов

**Аннотация:** В работе рассмотрены современные технологии и средства виртуализации, в частности, контейнеризация (LXC, Docker) как один из вариантов реализации виртуализации уровня операционной системы. Разработана технология дистрибуции и развёртывания программной системы «Инструментальный комплекс онтологического назначения» (ИКОН) с помощью виртуальных приложений и виртуальных операционных систем в среде облачных сервисов, на базе кластера Института кибернетики им. В. М. Глушкова НАН Украины. Программная система ИКОН интегрирована в модель сервис-ориентированных информационных систем – “Контейнер как Сервис”.

**Ключевые слова:** виртуализация, дистрибуция, контейнеризация, Docker, облачный сервис, программная система.

---

### Введение

Технологии виртуализации в настоящее время становятся одним из ключевых компонентов процесса разработки программных систем (ПС) и современной ИТ-инфраструктуры крупных предприятий (организаций) для повышения доступности серверов, бесперебойного предоставления сервисов клиентам. Такие этапы разработки программных систем, как внедрение (deployment - процесс настройки ПС под определенные условия эксплуатации, а также обучение пользователей работе с программным продуктом) и сопровождение (maintenance - процесс оптимизации и устранения, обнаруженных в процессе опытной эксплуатации дефектов функционирования ПС), являются одними из важнейших в жизненном цикле программного обеспечения, которые обеспечивают дистрибуцию, развёртывание и эксплуатацию конечным пользователям программной системы. С учётом важности изложенных выше этапов, необходимо разработать технологию дистрибуции и развёртывания программной системы ИКОН [Палагин, 2012] с использованием современных технологий и средств виртуализации. Существует несколько методов осуществления виртуализации через разные уровни абстракции, основными из них на сегодняшний день являются эмуляция оборудования, полная виртуализация и виртуализация уровня операционной системы. Ключевым методом виртуализации для разработки технологии дистрибуции и развёртывания программной системы ИКОН является метод виртуализации уровня операционной системы и ключевой его реализации – LXC (Linux Containers). Этот метод позволяет запускать изолированные и безопасные виртуальные машины на одном физическом узле.

---

### Программная система ИКОН в модели “Контейнер как Сервис”

При виртуализации уровня операционной системы не существует отдельного слоя гипервизора. Вместо этого сама хостовая операционная система отвечает за разделение аппаратных ресурсов между несколькими виртуальными машинами (выделенными серверами) и поддержку их независимости друг от друга. Виртуализация уровня операционной системы требует внесения изменений в ядро операционной системы, но при этом преимуществом является «родная» производительность, без потери ресурсов на виртуализацию устройств.

Для виртуализации ИКОН на уровне операционной системы была использована среда LXC, а также высокоуровневый интерфейс для автоматизации развёртывания и управления приложениями, и доступа к контейнерам LXC – Docker. Docker позволяет “упаковать” приложение со всем его окружением и зависимостями в так называемый контейнер, который может быть перенесён на любую Linux-систему с поддержкой cgroups в ядре, а также предоставляет среду по управлению контейнерами. С применением описанных средств и методов разработана технология развёртывания и дистрибуции программной

системы ИКОН в модели "Контейнер как Сервис". Обобщённая схема реализации программной системы ИКОН в модели "Контейнер как Сервис" представлена на рис. 1.

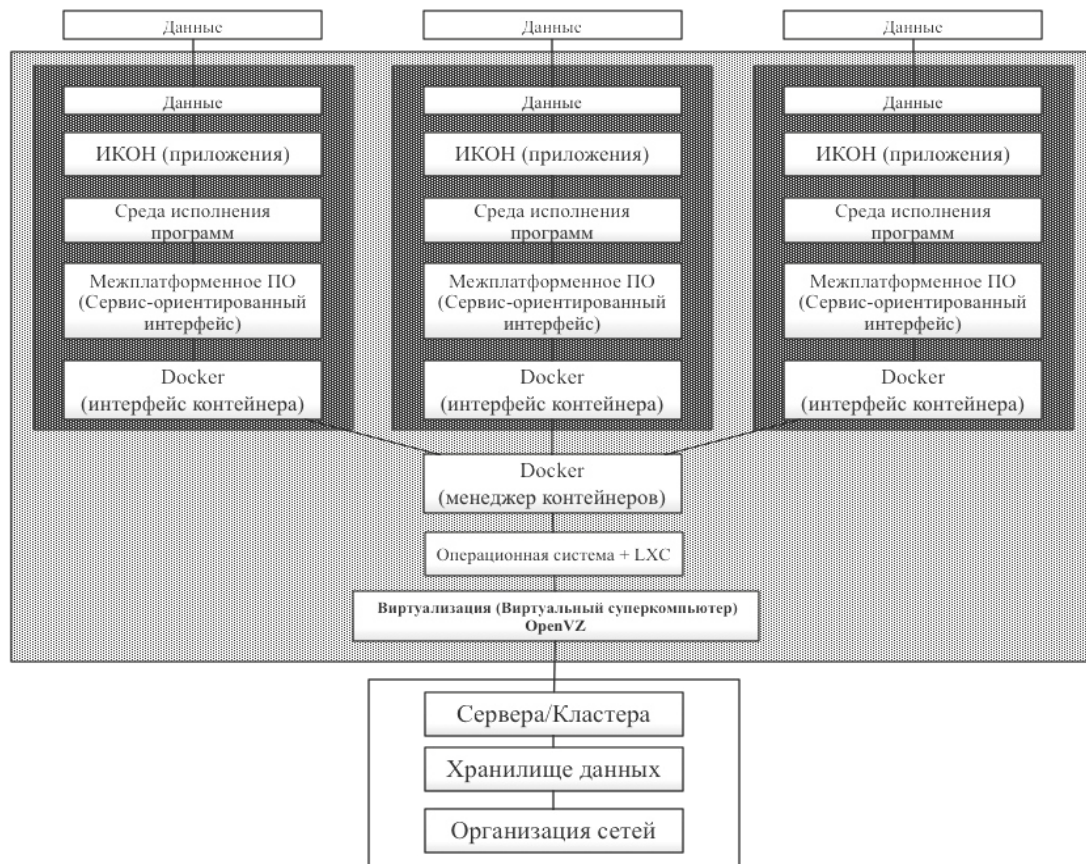


Рис. 1. Обобщённая схема реализации программной системы ИКОН в модели "Контейнер как Сервис".

## Выводы

В работе предложена модель развёртывания и дистрибуции программной системы ИКОН с применением технологий виртуализации, в частности, контейнеризация. Описана реализация разработанной модели в среде облачных сервисов на базе кластера Института кибернетики имени В. М. Глушкова НАН Украины. Разработанная модель позволяет: повысить уровень обслуживания и свести нарушения работы сервисов к минимуму, упростить процесс управления жизненным циклом программных систем, сократить время на их развёртывание и дистрибуцию, устранить проблемы совместимости приложений, повысить эффективность использования ресурсов оборудования.

## Литература

[Палагин, 2012] А.В. Палагин, С.Л. Крывый, Н.Г. Петренко. Онтологические методы и средства обработки предметных знаний. – [монография] – Луганск: изд-во ВНУ им. В. Даля, 2012. – 323 с.

## Информация об авторах

**Виталий Величко** – Институт кибернетики имени В.М. Глушкова Национальной академии наук Украины, с.н.с., проспект Академика Глушкова, 40, Киев, 03680 МСП, Украина; e-mail: [aduisukr@gmail.com](mailto:aduisukr@gmail.com)

**Кирилл Малахов** – Институт кибернетики имени В.М. Глушкова Национальной академии наук Украины, м.н.с., проспект Академика Глушкова, 40, Киев, 03680 МСП, Украина; e-mail: [malahov@live.com](mailto:malahov@live.com)

## СПОСІБ АВТОМАТИЗОВАНОГО ВИДІЛЕННЯ ВІДНОШЕНЬ МІЖ ТЕРМІНАМИ З ПРИРОДНОМОВНИХ ТЕКСТІВ ТЕХНІЧНОЇ ТЕМАТИКИ

Віталій Величко, Віталій Приходнюк

**Анотація:** В роботі розглядається спосіб виділення семантичних відношень з природномовних текстів. Для формалізації шаблонів семантичних відношень використовується числення предикатів.

**Ключові слова:** автоматизований аналіз природномовних текстів, семантичний аналіз.

### Вступ

В умовах постійного росту об'ємів інформації значно зростає роль систем автоматичного і автоматизованого аналізу неструктурованої інформації. Оскільки значна частина доступної на теперішній час інформації зберігається у вигляді природномовних текстів, актуальним є завдання аналізу таких текстів з метою перетворення інформації у формальний, машинно-зрозумілий, вигляд. В даній статті описується модель такого аналізатора, що дозволяє автоматизовано виділяти семантичну інформацію з природномовних текстів.

### Опис концепції аналізатора

Семантичний аналіз, який буде розглядатись нижче, базується на морфологічному і синтаксичному аналізах, що виконується програмою Texttermin [Андрущенко та ін., 2013]. В результаті морфологічного і синтаксичного аналізів формується послідовність лексем (нормалізованих слів та розділових знаків) і визначаються синтаксичні зв'язки між ними. Texttermin видає результати своєї роботи у вигляді двох структур даних: дерева термінів і власне лексичної згортки. Дерево термінів являє собою деревовидний граф, який вказує на підпорядкованість термінів у тексті. Лексична згортка має тривірневу ієрархію: вся згортка представляє текст, розбитий на речення, а речення, в свою чергу, розділяються на синтаксичні одиниці (групи). Групи складаються з лексем, кожна з яких може мати синтаксичний зв'язок з іншою лексемою цієї ж групи.

Запропонована модель оперує базою знань про природну мову, що аналізується, на основі формальних граматик. Для обробки тексту використовується модифікований GLR-парсер (Generalized Left-to-right Rightmost derivation parser) [Tomita, 1986]. Основна операція при аналізі тексту – підстановка лексем  $L = \{L_j | j \in 1..m\}$  з речень у предикатну формулу (шаблон), що являє собою послідовність предикатів  $P^n = \{P_i(x) | i \in 1..n\}$ . Предикатні формули зберігається в базі знань. Предикати  $P_i(x)$  бувають двох видів: для ключових та не ключових слів і мають однакову структуру. Предикати відрізняються інтерпретацією своїх елементів: для предикатів, що виділяють неключові слова, кожен елемент предиката являє собою певний семантичний клас, для предикатів, що виділяють ключові слова, – нормалізовану основу деякого слова.

Додатково вводиться процедура валідації отриманої підпослідовності лексем, яка перевіряє, чи зв'язані елементи вихідної підпослідовності потрібним чином. Для неї вводиться особливий тип предикатів – предикат виділення зв'язків. Такий предикат застосовується не до окремої лексеми, а до всього їхнього ланцюжка. Такий предикат містить два індекси – індекси елементів, між якими повинен існувати зв'язок, і власне тип зв'язку. Кожен шаблон містить набір таких предикатів, що застосовуються послідовно після отримання ідентифікації підланцюжка, який відповідає предикатам виділення лексем.

Додатково кожен шаблон містить маску результату – ще один шаблон, у який підставляється виділена підпослідовність і який задає перетворення вхідного ланцюжка лексем  $L'$  або графу результуючих відношень.

Пошук  $L'$  здійснюється за допомогою модифікованого GLR-парсера. Основною операцією при роботі з лексемами є застосування предикату, що замінює стандартну для парсерів на основі формальних

граматик операцію порівняння термінальних символів. Тоді кожен предикат вхідного шаблону можна розглядати як термінальний елемент правила (хоча реально він являє собою нетермінальний, а термінальними символами є елементи предикату).

Між важливими лексемами можуть зустрічатись другорядні (вставні слова, дієприслівникові звороти та ін.). Тому програма підстановки лексем повинна перевіряти різні способи підстановки, пропускаючи деякі вхідні лексеми. Додатково вводиться евристична процедура повернення. Повернення здійснюється, якщо в процесі підстановки лексем було досягнуто кінець послідовності лексем. Якщо ж предикатна формула відповідає послідовності, то кінець послідовності не буде досягнуто: підстановка лексем завершиться, як тільки буде знайдена відповідність до останнього предиката у формулі.

Оскільки через неоднозначність підстановки необхідна процедура повернення на попередні кроки, для автомата вводиться стек позицій ( $LS$ ). Для стека визначені операції: занесення поточної позиції в стек ( $LS+$ ) і її відновлення з стека ( $LS-$ ). В обох випадках операція виконується до зчитування наступної лексеми.

Власне виділення відношень відбувається в три етапи. Перший етап – виділення конструктів. Конструкти виділяються виключно в рамках одної групи. Шаблон виділення конструктів представляє собою звичайний шаблон з предикатів. Маска результату для шаблону виділення конструктів являє собою модель перетворення, а саме – кластеризації. В масці результату вказані позиції елементів, що сформують новий конструкт. Також в шаблоні вказано головний елемент конструкту. Перетворення полягає у видаленні з вхідного ланцюжка всіх лексем, які входять до вихідної підпослідовності (а також несуттєвих лексем, що розташовані між ними), і запису на їх місце одної лексеми, основа якої являє собою з'єднані у вказаному в масці порядку основи елементів підпослідовності, а семантичний клас дорівнює семантичному класу головного слова. Після заміни лексеми над групою виконується додаткова фільтрація: видаляються лексеми, які зв'язані з підпорядкованими членами конструкту і несуттєвими символами. Після обробки всіх груп в реченні, елементи цих груп, які пройшли фільтрацію, об'єднуються в один ланцюжок.

Другий етап виділення відношень – виділення зв'язків. Зв'язки виділяються на рівні речення. Алгоритм працює з ланцюжками, отриманими в результаті застосування до речення шаблонів виділення конструктів. Кожен зв'язок виділення зв'язку має центральний елемент – ключове слово. Справа і зліва від ключового слова знаходяться фільтри (предикати для неключових слів). Для більш складних речень можливі послідовності ключових слів і фільтрів.

Третій етап – уточнення ієрархічних зв'язків. Шаблон уточнення ієрархічних зв'язків є аналогічним до шаблону виділення конструктів і також застосовується на рівні груп. Відмінність полягає у масці результату: вона може виділяти кілька підпослідовностей, причому в загальному випадку підпослідовності вкладені одна в одну.

---

## Висновки

В даній статті описано модель виділення семантичної інформації з природномовних текстів. Для роботи модель вимагає формалізації відношень, що мають бути виділені при семантичному аналізі, у вигляді шаблонів.

---

## Література

[Андрущенко та ін, 2013] Методики написання наукових робіт на основі онтологічного аналізу текстів: методичний посібник / упоряд. Т. І. Андрущенко, В. Ю. Величко, С. А. Гальченко, Л. С. Глоба, К. Д. Гуляєв, Е. Я. Клімова, О.Б. Комова, О. В. Лісовий, М. А. Попова, В. В. Приходнюк, О. Є. Стрижак, Д. М. Стус. – К.: ТОВ «СІТІПРІНТ», 2013. – 124 с.

[Tomita, 1986] Tomita M. Efficient parsing for natural language. Kluwer Academic Publishers, Boston, 1986.

---

## Інформація про авторів

**Віталій Величко** – Інститут кібернетики імені В.М.Глушкова Національної академії наук України, с.н.с., проспект Академіка Глушкова, 40, Київ, 03680 МСП, Україна; e-mail: [aduisukr@gmail.com](mailto:aduisukr@gmail.com)

**Віталій Приходнюк** – Інститут телекомунікацій і глобального інформаційного простору Національної академії наук України, e-mail: [vitalik1700@yandex.ru](mailto:vitalik1700@yandex.ru)

## МЕДИЦИНСКАЯ ГРИД-СИСТЕМА ДЛЯ НАКОПЛЕНИЯ И ОБРАБОТКИ ЭЛЕКТРОКАРДИОГРАММ

Виталий Вишневский

**Аннотация:** В докладе обсуждается опыт построения в Украине специализированной медицинской грид-системы для накопления и обработки больших объемов цифровых электрокардиограмм при популяционных исследованиях. Показаны возможности современных грид-инфраструктур для проведения «отложенных телеконсультаций».

**Ключевые слова:** грид-система, медгрид, цифровые электрокардиограммы.

Современное состояние организаций здравоохранения в Украине характеризуется явным противоречием между достаточно распространенным цифровым диагностическим оборудованием и архаическим по форме и содержанию бумажным документооборотом медицинских записей.

Мало кто из пациентов обращает внимание на то простое обстоятельство, что цифровая по форме диагностическая информация меняет свои свойства в момент документирования на твердый носитель (бумагу, рентгеновскую пленку и т.д.) и уже не может обрабатываться современными компьютерами.

Это противоречие порождает, по крайней мере, два вопроса:

1. Нужно ли вообще хранить диагностическую информацию в цифровом виде (в том числе и цифровую ЭКГ) на протяжении всей жизни пациента?
2. Если все же диагностическая информация будет накапливаться, то можно ли организовать технологию доступа к ней в условиях бумажного документооборота?

На наш взгляд, преимущества представления диагностической информации в цифровом виде давно доказаны и потеря этого преимущества при документировании на бумажный носитель просто непростительная роскошь. Да и хранение всей диагностической информации на протяжении жизни пациента – скорее давно назревшее требование, нежели предмет для дискуссии.

Поэтому мы сконцентрировались на ответе на второй вопрос и дали его в ходе реализации проекта национальной программы развития и внедрения грид-технологий «Медицинская грид-система для популяционных исследований в кардиологии на базе кардиограмм» или просто «Медгрид», о котором пойдет речь ниже.

Первичной информацией для проекта «Медгрид» являются электрокардиограммы, получаемые с помощью сертифицированных цифровых кардиографов в специализированных медицинских учреждениях.

Технологический замысел проекта «Медгрид» в совокупности с компьютерными ресурсами виртуальной грид-организации (ВО) с одноименным названием приведены на рис.1.

Грид-хранилища проекта «Медгрид» позволяют накапливать деперсонализированные цифровые диагностические данные в масштабе популяции всей страны и обеспечивать обработку этой информации для целей практического здравоохранения и для научных исследований. Более подробно цели и перспективы проекта «Медгрид» проанализированы в [1].

Пока же важно понимать, что к распределенным грид-хранилищам цифровые кардиограммы доставляются через Веб-портал <http://medgrid.immsp.kiev.ua> при помощи специально разработанного транспортного модуля. Этот модуль может использоваться либо в виде самостоятельной программы, либо быть встроенным в программный комплекс рабочего места кардиолога.

Для работы с файлами цифровых кардиограмм используются специально разработанные интерактивные веб-сервисы. Один из этих Веб-сервисов генерирует для каждой доставленной в грид-хранилище проекта кардиограммы уникальную Интернет-ссылку и генерирует специальный документ в pdf-формате, который врач может распечатать или записать для пациента. Как видно из рис.2, этот документ имеет специальный QR-код, в котором закодирована ссылка доступа к цифровой кардиограмме и который легко распознается стандартными программными средствами ВЕБ-камер, в том числе современных планшетов и смартфонов.

На наш взгляд, бумажный документ с встроенным в него QR-кодом может рассматриваться как решение, которое обеспечивает стыковку современных технологий хранения цифровой диагностической информации с архаическим методом бумажного ведения медицинской документации. Кроме этого,

уникальную ссылку, которая используется при генерации QR-кода, можно использовать и при телемедицинских консультациях.

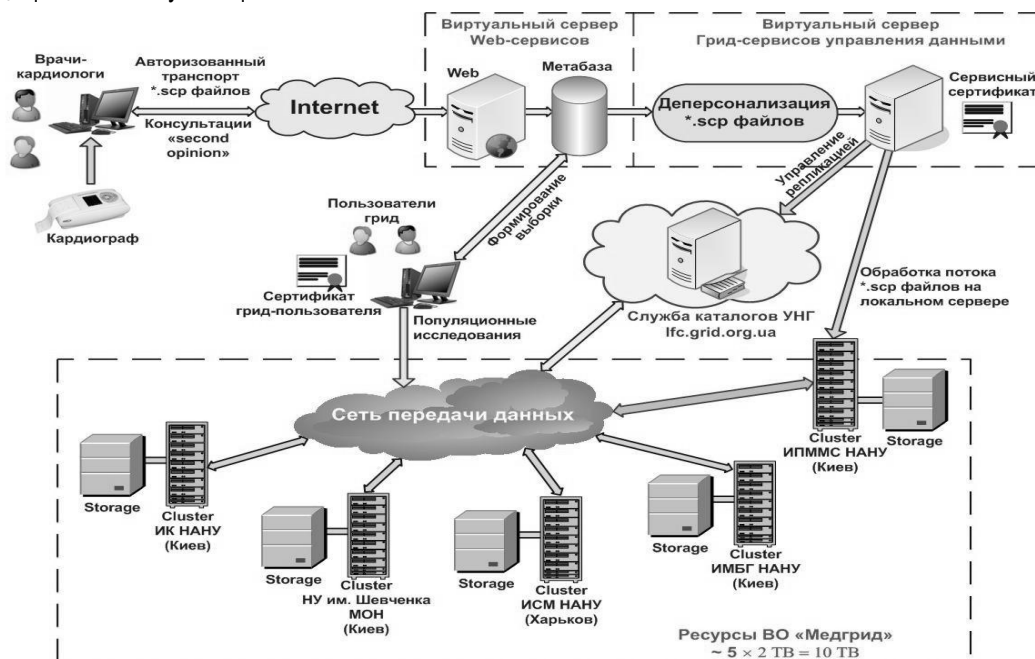



Рис.1 Технологический замысел проекта «Медгрид»


**Український Національний Грид**  
 Проект «МедГрид»  
<http://medgrid.immsp.kiev.ua>


Пациент: \_\_\_\_\_  
 ПІБ: И \_\_\_\_\_  
 Рік народження: 1959  
 Стать: Мужской  
 Поштовий код: 03187  
 Медичний заклад: Інститут проблем математических машин та систем НАН України, 03187, 03187, м.Київ-187, пр. Глушкова, 42  
 Основний симптом: \_\_\_\_\_  
 Лікар: \_\_\_\_\_ В. Вишневецький  
 Файл електрокардіограми: 80\_117\_20091214\_123100.scp  
 Дата зняття ЕКГ: 14.12.2009      Дата запису в МедГрид: 06.02.2012

URL електрокардіограми:  
[http://medgrid.immsp.kiev.ua/qz/eqa.asp?qrcode=5c0c61c85b3cf0d43d638745e72130641e43f1a42e16ca6bc946e71e341880\\_117\\_20091214\\_123100.scp](http://medgrid.immsp.kiev.ua/qz/eqa.asp?qrcode=5c0c61c85b3cf0d43d638745e72130641e43f1a42e16ca6bc946e71e341880_117_20091214_123100.scp)



Рис.2. Уникальная ссылка на кардиограмму пациента в виде QR-кода

В настоящее время в нашем проекте обрабатываются технологии динамического запуска виртуальных машин на грид-кластерах, которые позволяют работать пользователю с интерактивными приложениями. Таким образом, в Украине накоплен опыт использования современных грид-технологий для накопления, организации доступа и обработки деперсонализированной медицинской диагностической информации. Создана и развивается виртуальная организация "Медгрид", которая уже сейчас способна обрабатывать электрокардиограммы в масштабе популяции всей страны.

## Литература

1. Вишневецкий В.В., Чайковский И.А., Киржнер Г.Д., Ена Л.М., Дордиенко Н.А., Фролов Ю.А., Васильев В.Е. Медицинская ГРИД-система на базе электрокардиограмм: новый инструмент для клинической кардиологии и популяционных исследований // Международный научно-практический журнал "Кардиологи: от науки к практике". - 2012. - №2. - С. 108-116.

## Информация об авторе

**Виталий Вишневецкий** - Киев, Институт проблем математических машин и систем НАНУ, пр-т Академика Глушкова 42, г.Киев, 03187, Украина, к.т.н., с.н.с., e-mail: [vit@immsp.kiev.ua](mailto:vit@immsp.kiev.ua)

## БАГАТОПАРАМЕТРИЧНИЙ ВИБІР НАЙБІЛЬШ ДОЦІЛЬНИХ МАРШРУТІВ ЛІНІЙ ДОСТУПУ З ВИКОРИСТАННЯМ МОДЕЛІ БАЛАНСНИХ МЕРЕЖ

Галина Гайворонська, Світлана Сахарова

**Анотація:** В роботі на основі підходу, запропонованого проф. Гладуном В.П., сформульовано понятійну та категоріальну моделі синтезу мереж доступу. Запропоновано метод багатопараметричного вибору найбільш доцільного шляху прокладки ліній доступу на основі моделі балансних мереж, який враховує всю множину вхідних параметрів, необхідних при синтезі мережі доступу.

**Ключові слова:** мережа доступу, модель балансних мереж, управляючий вплив.

### Вступ

Найважливішим завданням розвитку телекомунікацій є створення мереж доступу (МД), оскільки цей фрагмент телекомунікаційної мережі є одним з її найбільш вузьких місць. Це обумовлено рядом питань, пов'язаних з реалізацією каналу зв'язку між терміналом користувача і мережею. До них відносяться: дороговизна оновлення обладнання та ліній; відсутність чітких показників попиту послуг тощо. Крім того, з кожним роком зростають вимоги користувачів до якості послуг, в результаті чого збільшуються необхідні швидкості обміну інформацією.

На сьогоднішній день тема створення МД є дуже молодою. Це пояснює відносно невеликий перелік робіт, пов'язаних із створенням цих мереж. Різні аспекти синтезу МД, застосування на МД проводових і безпроводових технологій освітлені в роботах Соколова Н.А., Гольдштейна Б.С., Бакланова І.Г., Михайлова В.Ф., Гайворонської Г.С., Денисьевой О.М., Балашова В. А., Бірюкова Н. Л. та ін. Деяким важливим аспектам синтезу МД присвячені дисертаційні роботи Павлова С.В., Котової О.І., Сахарової С.В. та Крижанівської А.А.

### Основна частина

В роботі на основі підходу, запропонованого проф. Гладуном В.П. в [Гладун, 2000], сформульовані понятійна та категоріальна моделі синтезу МД. Понятійна модель моделі синтезу МД має вигляд  $M_d = \langle W_d, L_d \rangle$  (рис. 1), де  $W_d$  – множина вершин, що відповідають поняттям синтезу МД;  $L_d$  – множина дуг, які представляють зв'язки між цими поняттями.



Рис.1. Понятійна модель синтезу МД

Сформована загальна схема прийняття рішення при синтезі МД, за якою створено алгоритм, що відповідає цільовим критеріям  $E = \{e\}$ , а значення  $f(x) = g(x) + h(x)$  є оцінкою оптимального шляху розв'язку задачі. На основі аналізу задач, розв'язок яких необхідний для синтезу МД [Соколов, 1991,



Соколов, 2002, Гайворонская, 2012, Гайворонская 2013], особливостей МД [Бирюков, 2003, Соколов, 2005, Гайворонская, 2013] та сформульованої понятійної моделі синтезу МД розроблено граф підцілей синтезу МД та запропоновано модель синтезу МД в цілому [Гайворонская, 2010, Гайворонская, 2012, Gayvoronska, 2012, Sakharova, 2012].

Задача, яку було розв'язано в роботі, пов'язана з визначенням найбільш доцільного шляху прокладки ЛД. Для розв'язання задачі використана модель балансних мереж (МБМ) [Гладун, 1995]. Ця модель призначена для управління процесом пошуку рішень і являє собою граф  $\langle U, V \rangle$ , де  $U$  – множина вузлових вершин, а  $V$  – множина зв'язків між ними. Принцип роботи моделі полягає у формуванні послідовності управляючих впливів  $hr$ , тобто плану досягнення кінцевого результату, де кожна вершина має своє значення, причому нижня вершина є аргументом для функції вищої вершини.

Враховуючи, що на прокладку ЛД впливає множина параметрів  $D = \{d^1, \dots, d^n\}$  виділено множини параметрів  $A = \{a^1, \dots, a^p\}$ , що характеризують вузлові вершини, та  $B = \{b^1, \dots, b^r\}$ , які відображають зв'язки між цими вершинами, при цьому  $A \cup B = D$ . Для кожної вузлової вершини визначено функцію

$$f_{ij} = f((i, j), a^1, \dots, a^p), \quad (1)$$

де  $i, j$  – ідентифікатори вузлової вершини. Для кожного зв'язку між вершинами визначено функцію

$$g_{ks} = g((k, s), b^1, \dots, b^r), \quad (2)$$

де  $k, s$  – ідентифікатори зв'язку.

На основі принципів МБМ та введених позначень отримана модель визначення найбільш доцільного шляху прокладки ЛД [Гайворонская и др., 2013 (1)], приклад реалізації якої зображено на рис. 2. Такий підхід дозволив виділити взаємозамінні фрагменти шляху, відкинути завідомо найгірші фрагменти та виявити варіанти зі схожими фрагментами шляху, що дозволило зменшити кількість розрахунків і скоротити час, необхідний для вирішення задачі.

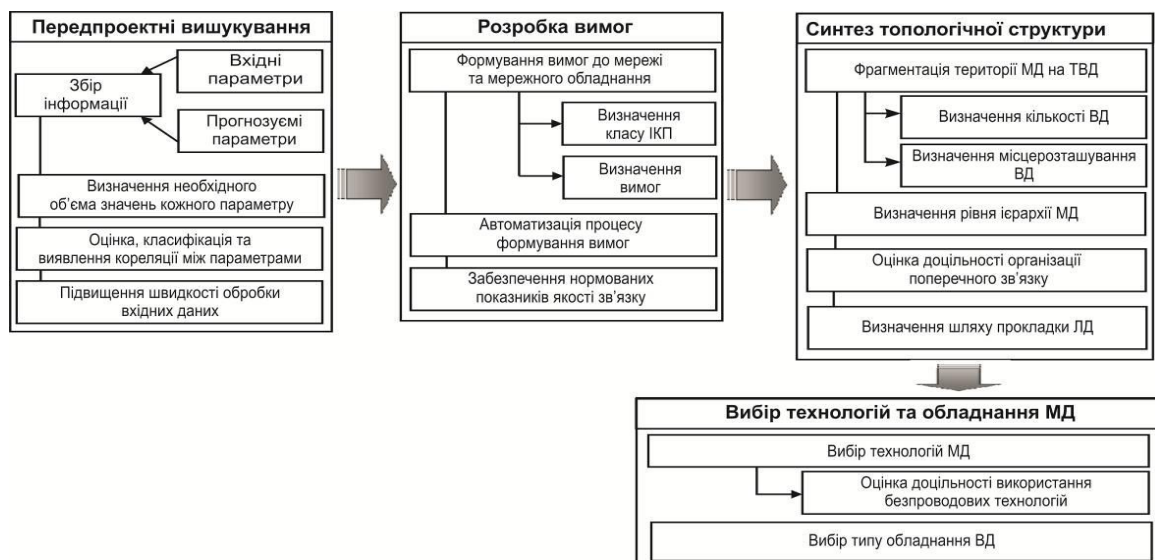


Рис.2. Модель синтезу МД в цілому

Враховуючи багатопараметричність системи, напрям пошуку описано як

$$hr = \sum_{ij} f_{ij} + \sum_{ks} g_{ks}, \quad (3)$$

де  $i, j$  та  $k, s$  – ідентифікатори відповідної вершини.

Для знаходження найбільш доцільного шляху прокладки ЛД сформовано множину альтернатив – всіх можливих шляхів прокладки ЛД між обраними пунктами  $L = \{l_1, \dots, l_y\}$ , де  $l_y$  – шлях, що описує послідовність вузлових вершин, через які він проходить.

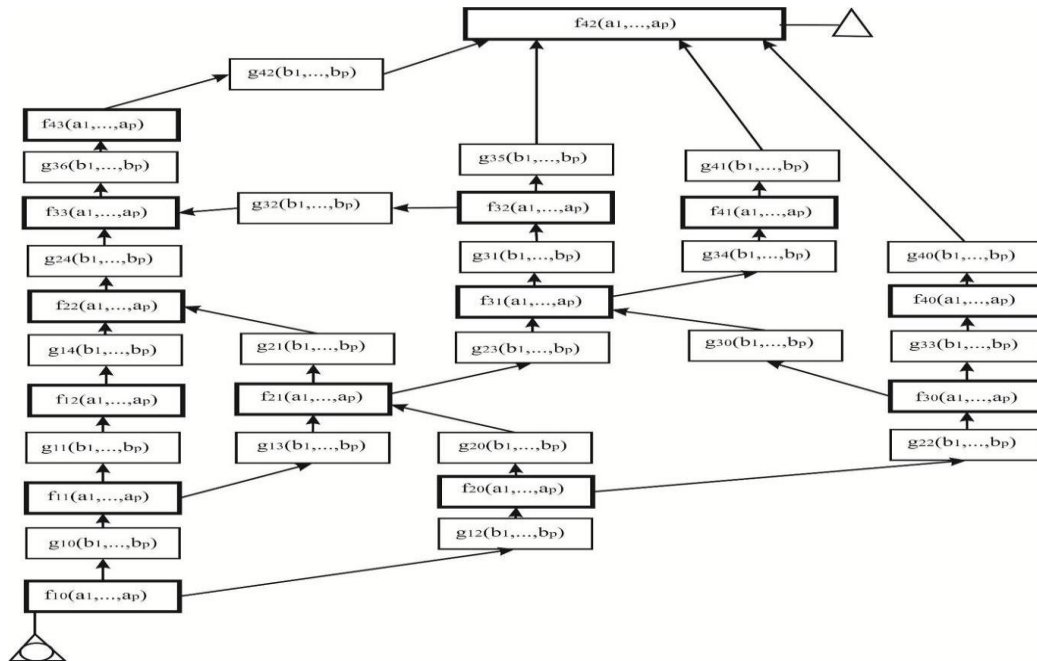


Рис.3. МБМ для визначення найбільш доцільного шляху прокладки ЛД

Нехай  $d^w$  – параметр, по якому оптимізується шлях, тоді  $D_w(l_y) = \sum_{k \in l} d_{ks}^w \rightarrow \min$ , де  $D_w(l_y)$  – сумарне значення параметру  $d^w$  для обраного  $y$ -го шляху та  $w = \overline{1..n}$ , а  $f_{ij}$  вузлові та проміжні вершини обраного шляху з (1-3), тоді загальний управляючий вплив, що характеризує обраний шлях має вигляд  $Hr(l_y) = \sum_{i,j \in l_y} f_{ij} + \sum_{k,s \in l_y} g_{ks} \rightarrow \min$ . Найбільш доцільний варіант прокладки ЛД отримано за критерієм

Байеса-Лапласа [Волошин, 2006, Гайворонская, 2013 (2)]

$$I_{BL} = \min_y | l_y | = \min_y | \sum_w D_w(l_y) \cdot q_w + Hr(l_y) \cdot q_{w+1} |, \quad (6)$$

де  $q_1, q_{w+1}$  – вага для  $D_w(l_y)$  та  $Hr(l_y)$ , яка встановлюється особою, що приймає рішення.

## Висновки

Таким чином, на основі моделі балансних мереж запропоновано метод багатопараметричного вибору найбільш доцільного шляху прокладки лінії доступу, якій враховує всю множину вхідних параметрів необхідних при синтезі мережі доступу. З використанням запропонованого методу розроблено алгоритм автоматизованої процедури синтезу топологічної структури МД.

## Література

[Гладун, 2000] В.П. Гладун. Партнерство с компьютером. Человеко-машинные целеустремленные системы, «Port Royal». Киев, 2000. – 128 с.

[Соколов, 1991] Н.А. Соколов. Сети абонентского доступа. Принципы построения. Пермь: ИПК Звезда, 1999. – 254 с.

[Соколов, 2002] Н.А. Соколов. Развитие сетей доступа и законы Мерфи. Матер. IX Всероссийского научно-технического семинара «Связь в деловой сфере. Новые информационные технологии». Москва, 2002.

- [Бирюков, 2003] Н.Л. Бирюков, С.Ф. Коваленко. Транспортные функции сети доступа в документах МСЭ. Вестник связи №4. 2003.
- [Соколов, 2005] Н.А. Соколов. Семь аспектов развития сетей доступа. Технологии и средства связи №3. Часть 2. – Специальный выпуск «Системы абонентского доступа». 2005.
- [Гладун, 1995] В.П. Гладун, Н.Д. Ващенко, Г.В. Реутов, С.И. Принада. Балансные методы планирования действий в системах управления. УСИМ, Киев, 1995. №1-2. – с.93-96.
- [Волошин и Мащенко, 2010] О.Ф. Волошин, С.О. Мащенко. Моделі та методи прийняття рішень. Навч. Посіб.: Видавничо-поліграфічний центр «Київський університет». Київ, 2010. – 336 с.

---

### Результати попередніх досліджень

---

- [Гайворонская, 2003] Г.С. Гайворонская. Особенности реализации транспортного сегмента при модернизации сети абонентского доступа. Вісник УБЕНТЗ №1. Київ, 2003. – с. 173-181.
- [Гайворонская, 2010] Г.С. Гайворонская, А.И. Котова. Основные задачи модернизации сетей пользовательского доступа. Зв'язок. №2 (90). Харків, 2010. – с. 32-36.
- [Гайворонская, 2012] Г.С. Гайворонская, С.В. Сахарова. Основные принципы проектирования сетей доступа. Зв'язок №4 (100). Харків, 2012. – с. 22-28.
- [Гайворонская, 2012] Г.С. Гайворонская. Процессы выбора альтернативных решений при проектировании информационных сетей. In: Problems of Computer Intellectualization. National Academy of Sciences of Ukraine V.M.Glushkov Institute of Cybernetics, ITHEA, 2012. №28. Kyiv-Sofia, 2012. – pp. 307-317.
- [Gayvoronska, 2012] G. Gayvoronska. Method for Determination of Interrelation between Access Network Characteristics. In: Problems of Computer Intellectualization. National Academy of Sciences of Ukraine V.M.Glushkov Institute of Cybernetics, ITHEA, 2012. №28. Ed. Galyna Gayvoronska, Svitlana Sakharova. Kyiv-Sofia, 2012. – pp. 41-46.
- [Sakharova, 2012] Svetlana Sakharova Analysis and Justification for Selection Parameters of Wired Access Systems / Svetlana Sakharova // International Journal "Information models and analyses". – Sofia: ITHEA, 2012. – № 3 (Volume 1) – P. 283-289.
- [Гайворонская, 2013] Г.С. Гайворонская. Особенности сетей доступа к NGN, требующие нового подхода к их синтезу. Материалы первой МНПК «Проблемы инфокоммуникаций. Наука и технологии» PICS&T. Харків, 2013. – с. 14-18.
- [Гайворонская и др., 2013 (1)] Г.С. Гайворонская, А.А. Бондаренко. Задача выбора топологической структуры сети доступа. International Journal «Information Models and Analyses». Vol. 2 № 2. Bulgaria, 2013. – pp. 252-261.
- [Гайворонская и др., 2013 (2)] Г.С. Гайворонская, А.А. Бондаренко. Методы теории принятия решений при выборе топологической структуры сетей доступа. 23-я Международная Крымская конференция «СВЧ-техника и телекоммуникационные технологии (КрыМиКо'2013)». Севастополь, Крым, Украина, 2013. – с.416-417.

---

### Інформація про авторів

---

**Галина Гайворонська** – кафедра інформаційно-комунікаційних технологій Одеської національної академії харчових технологій, д.т.н., професор, завідує кафедрою інформаційно-комунікаційних технологій; вул. Дворянська, 1/3, Одеса-26, 65026, Україна; тел. (048) -720-91-48, e-mail: gayvoronska@osar.odessa.ua

Головні галузі наукових досліджень: оптимізація перехідних періодів при еволюції телекомунікаційних мереж. Потіки викликів, навантаження і міжвузловими тяжіннями в мережах. Проблеми створення перспективних мереж доступу. Проблема побудови повністю оптичних систем комутації.

**Світлана Сахарова** – кафедра інформаційно-комунікаційних технологій Одеської національної академії харчових технологій, доцент кафедри; вул. Дворянська, 1/3, Одеса-26, 65026, Україна; моб. (38067)-483-39-47; e-mail switchonline@rambler.ru

Головні області наукового дослідження: Проблеми створення перспективних мереж доступу.

## ФЕНОМЕН СТАТИСТИЧЕСКОЙ УСТОЙЧИВОСТИ

Игорь Горбань

**Аннотация:** Представлена новая монография, посвященная исследованию физического феномена статистической устойчивости и изложению основ физико-математической теории гиперслучайных явлений, описывающей физические события, величины и процессы с учетом нарушений статистической устойчивости. В отличие от предыдущих двух монографий автора, в которых основное внимание уделено математическим аспектам теории гиперслучайных явлений, в новой книге акцент сделан на физической стороне вопроса. Книга рассчитана на научных работников, инженеров и аспирантов, исследующих статистические закономерности реальных физических явлений, разрабатывающих и использующих статистические методы высокоточных измерений, прогнозирования и обработки сигналов на больших интервалах наблюдения, а также для студентов старших курсов университетов физических, технических и математических специальностей.

**Ключевые слова:** феномен статистической устойчивости, теория гиперслучайных явлений, экспериментальные исследования, нарушение сходимости.

---

### Введение

---

Одним из удивительных физических феноменов является статистическая устойчивость массовых явлений, проявляющаяся в стабильности статистик. На этот феномен впервые обратил внимание торговец сукном Дж. Граунт в 1669 г. Его изучение привело к построению теории вероятностей, которая была аксиоматизирована А.Н. Колмогоровым в конце двадцатых годов XX века.

Хотя предложенный А.Н. Колмогоровым аксиоматический подход базируется исключительно на абстрактных математических понятиях и фактически игнорирует феномен статистической устойчивости, однако, физические корни теории вероятностей играют первостепенную роль во всех физических, технических, социологических и других дисциплинах, где эта теория нашла широкое применение.

Современная теория вероятностей (включая в широком понимании и математическую статистику) описывает массовые явления с помощью случайных (вероятностно-случайных или, иначе, стохастических) математических моделей, характеризуемых вероятностной мерой, опираясь (зачастую неявно) на физическую гипотезу идеальной статистической устойчивости, предполагающей наличие сходимости частоты реальных событий и средних значений физических величин.

Многие годы гипотеза идеальной статистической устойчивости не вызывала сомнений. Однако, последние экспериментальные исследования различных физических величин и процессов на больших интервалах наблюдения показали, что она не находит экспериментального подтверждения.

На относительно небольших временных, пространственных или пространственно-временных интервалах наблюдения увеличение объема данных приводит к уменьшению уровня флуктуаций статистик. Однако при больших объемах эта тенденция не прослеживается: достигнув определенного значения, уровень флуктуаций практически не меняется или возрастает. Это указывает на отсутствие сходимости реальных статистик (их несостоятельность).

Исследования нарушений статистической устойчивости физических явлений и разработка эффективных средств адекватного описания мира с учетом таких нарушений привели к построению новой физико-математической теории — теории гиперслучайных явлений [Gorban, 2007, 2011].

В теории вероятностей базовыми математическими объектами (моделями) являются случайные событие, величина и функция; в теории гиперслучайных явлений в таком качестве выступают гиперслучайные событие, величина и функция, представляющие собой множества несвязанных между собой случайных событий, величин и функций, рассматриваемые как единое целое.

Математическая составляющая теории гиперслучайных явлений базируется на классических аксиомах теории вероятностей А.Н. Колмогорова, физическая — на гиперслучайных физических гипотезах: гипотезе ограниченной статистической устойчивости реальных событий, величин, процессов и полей и на гипотезе адекватного описания этих физических явлений гиперслучайными моделями.

С точки зрения математики теория гиперслучайных явлений — ветвь теории вероятностей; с точки зрения физики — новая теория, основанная на новых представлениях об окружающем мире.

Целью настоящего доклада является презентация новой монографии [Gorban, 2014], посвященной исследованию нарушений статистической устойчивости реальных физических явлений и описанию реальных событий, величин, процессов и полей с помощью гиперслучайных моделей.

В новой монографии, в отличие от предыдущих книг [Gorban, 2007, 2011], основное внимание уделено физической стороне вопроса.

---

### **Краткое описание монографии**

---

Книга состоит из пяти частей, объединяющих 33 главы. Первая, вводная, часть монографии, включающая главы 1–8, посвящена описанию физического феномена статистической устойчивости, анализу его свойств и их учету в рамках известных теорий.

Акцентировано внимание на том, что феномен статистической устойчивости реальных физических явлений носит ограниченный характер, проявляющийся в отсутствии тенденции сходимости средних величин. Поэтому понятие вероятности, определяемое в статистике как предел частоты событий, не имеет физического смысла.

Проанализированы разные концептуальные взгляды на устройство мира с позиций детерминизма и неопределенности. Проведено сопоставление варианта описания физических явлений без учета нарушений статистической устойчивости, широко используемого в теории вероятностей, с альтернативным вариантом описания реальных явлений с учетом нарушений статистической устойчивости, предлагаемого теорией гиперслучайных явлений.

Формализовано понятие статистической устойчивости процессов, введены параметры статистической неустойчивости и единицы их измерения, описана методика оценки степени нарушения устойчивости на интервалах конечной длительности, установлены зависимости нарушений статистической устойчивости процессов от особенностей их временных и спектральных характеристик.

Во второй части книги (главы 9–13) представлены результаты экспериментальных исследований на больших интервалах наблюдения статистической устойчивости разнообразных процессов разной физической природы.

Исследованы, в частности, колебания напряжения городской сети, высоты морских волн и периода их следования, температуры и скорости звука в океане, интенсивности излучения астрофизических объектов, температуры воздуха и количества осадков, магнитного поля Земли, курса валют и др.

Итогом этих исследований явилось установление факта ограниченной статистической устойчивости рассмотренных процессов. Несмотря на разную физическую природу, все они имеют ограниченный интервал статистической устойчивости. Эти результаты позволяют предположить, что нарушение статистической устойчивости носит всеобщий характер или, иначе, феномен статистической устойчивости носит ограниченный характер.

Третья часть книги (главы 14–21) посвящена краткому изложению основных положений математической части теории гиперслучайных явлений. В отличие от стохастических моделей, характеризующихся исключительно конкретными параметрами и характеристиками, в гиперслучайных моделях присутствуют элементы неопределенности, ассоциируемые с непредсказуемо меняющимися условиями.

---

Описание гиперслучайных явлений (гиперслучайных событий, величин и функций) обеспечивается разными средствами. Наиболее полное описание гиперслучайного явления дает многозначная функция распределения, представляемая множеством функций распределения, соответствующих разным условиям, менее полное – границы многозначной функции распределения, моменты этих границ, границы моментов, а также другие параметры и характеристики.

Пятая часть (главы 26–33) содержит теоретические и экспериментальные исследования статистических закономерностей при нарушениях статистической устойчивости.

Установлено, что при нарушении статистической устойчивости, как ни странно может показаться, действует закон больших чисел и справедлива центральная предельная теорема, хотя смысл их оказывается не таким, как в случае идеальной статистической устойчивости.

Особое внимание уделено выяснению причин ограниченной точности реальных измерений и потенциальной точности реальных измерений, а также причин возникновения нарушений статистической устойчивости.

Предложен подход, позволяющий распространить понятие шенноновской энтропии и другие понятия, использующие вероятностные характеристики, на неопределенные величины, не имеющие вероятностной меры.

---

## Выводы

1. Представлена новая монография, посвященная исследованию нарушений статистической устойчивости физических явлений и описанию реальных событий, величин, процессов и полей с помощью гиперслучайных моделей.
2. Выяснено, что вероятность – математическое понятие, не имеющее физической интерпретации.
3. Установлено, что при нарушении статистической устойчивости действует закон больших чисел и справедлива центральная предельная теорема, хотя смысл их оказывается не таким, как в случае идеальной статистической устойчивости.
4. На базе математического аппарата теории гиперслучайных явлений разработаны основы математического анализа расходящихся и многозначных функций.

---

## Литература

- [Gorban, 2007] Gorban, I. I. (2007). Theory of hyper-random phenomena. Kiev: IMMSP, NAS of Ukraine, 181. ISBN 978-966-02-4367-5. From [http://www.immsp.kiev.ua/perspapes/gorban\\_i\\_i/index.html](http://www.immsp.kiev.ua/perspapes/gorban_i_i/index.html). (In Russian)
- [Gorban, 2011] Gorban, I. I. (2011). Theory of hyper-random phenomena: physical and mathematical basis. Kiev: Naukova dumka, 318. ISBN 978-966-00-1093-2. From [http://www.immsp.kiev.ua/perspapes/gorban\\_i\\_i/index.html](http://www.immsp.kiev.ua/perspapes/gorban_i_i/index.html). (In Russian)
- [Gorban, 2014] Gorban, I. I. (2007). Phenomenon of statistical stability. Kiev: Naukova dumka, 318. (In Russian)

---

## Информация об авторе

**Игорь Горбань (Igor Gorban)** – Principal scientist of the Institute of Mathematical Machines and Systems Problem, National Academy of Sciences of Ukraine, Glushkov ave., 42, Kiev, 03187, Ukraine; e-mail: [igor.gorban@yahoo.com](mailto:igor.gorban@yahoo.com)

Major Fields of Scientific Research: Phenomenon of Statistical Stability, Mathematical statistics, Probability theory, Physical-mathematical theory of hyper-random phenomena, Space-time signal processing in complicated dynamic conditions, Fast multichannel space-time signal processing

## ПРОГНОЗИРОВАНИЕ СОЦИАЛЬНО-ЭКОНОМИЧЕСКОГО РАЗВИТИЯ С ПОМОЩЬЮ АССОЦИАТИВНЫХ СЕТЕЙ

Валентин Григорьевский

**Аннотация:** *Анализируются понятия социально-экономической системы и ассоциативных сетей. Рассматривается задача прогнозирования показателей социально-экономического развития, а также задачи, решаемые с помощью инструмента ассоциативных сетей. Исследование проводится с целью последующей разработки алгоритма прогнозирования показателей социально-экономического развития региона, а также разработки программного обеспечения с целью реализации данного алгоритма.*

**Ключевые слова:** *Ассоциативные сети, прогнозирование, социально-экономическая система.*

### 1. Введение

Данная работа является первым шагом к разработке альтернативного полнофункционального инструмента решения задач прогнозирования показателей социально-экономического развития произвольного региона.

Объект исследования – социально-экономическая система (СЭС).

Предмет исследования – ассоциативные связи в показателях социально-экономического развития.

Цель исследования – разработка концепции модели для нахождения ассоциативных связей между показателями социально-экономического развития основных регионов и использование полученных результатов для прогнозирования значений аналогичных показателей по произвольным регионам.

### 2. О социально-экономических системах

Территориальная СЭС (ТСЭС) – совокупность всех экономико- и социально-географических объектов определенной территории с установившимися между ними связями. Это сложившееся и функционирующее на определенной территории сочетание всех элементов народного хозяйства (не только производственных), будь то поселения промышленные или сельскохозяйственные предприятия, звенья производственной и социальной структуры и др. Каждая ТСЭС — очень сложное образование, структура которого формировалась под воздействием как природных, так и общественных факторов. ТСЭС различаются не только функциональной структурой, но и размерами территории. ТСЭС по отношению к территориально-производственному комплексу являлись новой категорией с более высоким уровнем обобщения.

При этом под СЭС могут подразумеваться объекты совершенно различных уровней: от домохозяйств до глобальной экономической системы. Естественно, сравнивать настолько разные системы не корректно, но если посмотреть на задачу системно, то можно обнаружить общие элементы и принципы, которые имеют системы разного уровня и масштаба. Выявление таких элементов, определение закономерностей их взаимодействия даёт возможность сравнивать эти системы, сколь сильно они не отличались бы по масштабу. Подобное сравнение, в свою очередь, даёт возможность утверждать, что, например, некая система А функционирует лучше системы Б, а системы Б и В функционируют «приблизительно одинаково». Экстраполируя данное утверждение на временной ряд, получаем следующий пример: система А в периоды «1» и «3» функционировала одним образом, а в периоды «2», «4», и «5» – иначе.

Представим, что нам в полном объёме доступна необходимая информация о функционировании всех интересующих нас СЭС (различного масштаба) за длительный период времени. Всё, что остаётся сделать, найти «похожие» поведения и рассчитать необходимые метрики. Данная информация позволит решать различные задачи:

- Проводить анализ системы на высоком уровне для выполнения экспресс-оценок;
- Оценивать состояние «новой» системы по минимальному набору данных о ней – найти поведение, «похожее» на поведение других систем;
- Оценивать будущее состояние системы, имея данные по её текущему состоянию.

В следующей части рассмотрим принципы, используя которые возможно решение указанных выше задач.

### 3. Об ассоциативных сетях

Ассоциативная сеть – средство представления реляционных знаний с помощью маркированного ориентированного графа. Каждая вершина графа представляет собой концепт и каждая метка представляет собой соотношение между понятиями. Данное определение используется с точки зрения

концептуального описания ассоциативных сетей. Когда же речь идёт об их прикладном использовании (реализация программы), то под термином «ассоциативная сеть» возникает другой термин, детализирующий семантику понятия – «ассоциативная память на нейронных сетях». Указанный термин, в свою очередь, состоит из двух понятий: «ассоциативная память» и «нейронная сеть». Дадим определение этим понятиям.

Ассоциативная память (информатика) – устройство хранения данных, в котором расположение данных определяется их информационным содержанием, а не именем, адресом или относительной позицией, и из которого данные могут быть получены.

Ассоциативная память (психология) – восстановление в памяти информации об объекте, с которым мы взаимодействовали ранее, по причине того, что мы думаем о чём-то, что связано с этим объектом, и вызываем, таким образом, ассоциацию.

Нейронная сеть (искусственная) – математическая модель, а также её программная или аппаратная реализация, построенная по принципу организации и функционирования биологических нейронных сетей — сетей нервных клеток живого организма.

Анализируя все вышеуказанные определения, принимаем следующие утверждения, которые будут использованы в рамках данного исследования:

1. Ассоциативная сеть рассматривается как процесс анализа больших массивов данных с целью выделения общих закономерностей и сохранения результатов проведённых вычислений.
2. Ассоциативная память как свойство системы позволяет по заданному условию находить информацию о выявленных свойствах и закономерностях (то, что выше указано как «память по содержанию»).
3. Модель разрабатывается с использованием технологии нейронных сетей.

Одним из примеров, в котором уже довольно давно и успешно применяются алгоритмы выявления ассоциативных правил (и который послужил идеи использования подобных принципов для прогнозирования показателя социально-экономических процессов), является задача по анализу покупательской корзины. Рассмотрим её подробнее и на её примере приведём описание предпосылок к формированию концепции алгоритма прогнозирования показателей социально-экономического развития с помощью ассоциативных сетей.

Главные задачи модели анализа покупательской корзины:

1. Рассчитать все возможные наборы товаров и частоту их возникновения в покупках;
2. Отсеять те наборы, которые имеют низкий уровень поддержки (вероятность наступления некоего события ниже определенного уровня);
3. По заданному условию (по запросу пользователя) рассчитать условную вероятность определенного события (например, какова вероятность того, что, купив Товар А, клиент приобретёт Товар Б).

Существует множество различных методов для вычисления ассоциативных правил приобретения товаров, которые, впрочем, не могут быть применены к задаче поиска ассоциативных правил при прогнозировании значений показателей социально-экономического развития. Причина этого заключается в том, что первая из указанных задач является дискретной, а вторая – непрерывной. Следовательно, необходимо взять идею и применить её для решения новой задачи в новых условиях. Так, основная идея заключается в решении главных задач, которые описаны выше. В следующей части рассмотрим каждую из них по отдельности и на их примере опишем концепцию предлагаемого алгоритма.

---

#### 4. Концепция разрабатываемой модели

---

##### **Задача 1. Перечислить все возможные наборы и частоту их возникновения.**

Аналогом «товара» выступает состояние СЭС, выраженное набором показателей (матрицей значений). Такие наборы показателей не могут быть просто сгруппированы, как наборы товаров, ведь перечень товаров конечен, чего нельзя сказать о перечне состояний системы. Поэтому для группировки наборов показателей необходимо применить один из алгоритмов кластеризации, который сгруппирует подобные состояния системы. Для каждого полученного кластера следует рассчитать и сохранить информацию о характеристиках наборов, входящих в этот кластер, и частоте его возникновения.

Алгоритм решения данной задачи включает следующие шаги.

1. Для анализа выбираются данные по основным показателям уровня социально-экономического развития по 50 (пятидесяти) странам за последние 50 (пятьдесят) лет с детализацией по годам. Для того, чтобы модель была универсальной и её возможности можно было применить для СЭС любого уровня, к



показателям следует чётко сформулировать требования (например, данные должны отражать уровень социально-экономического развития системы; данные должны быть динамичными, т.е. изменяться каждый год; расчет показателей должен быть таким, что не требует значительных затрат и т.д.).

2. Данные приводятся к такому виду, в котором с ними может работать модель.
3. Из обработанных данных формируются наборы (матрицы) по  $N$  последовательных записей, где:
  - столбец соответствует значению конкретного показателя,
  - строка отвечает за период времени.

Например, при  $N = 2$  формируется набор показателей (матрица с 2 (двумя) строками) за 2 (два) последовательных года.

4. Для каждого значения  $N$  выполняется кластеризация наборов – все наборы подаются на вход нейронной сети Кохонена (т.н. «самоорганизующаяся карта Кохонена»). Данный инструмент зарекомендовал себя достаточно хорошо при решении задач кластеризации без учителя.

5. Для каждого кластера для каждого  $N$  рассчитываются и сохраняются характеристики кластера, необходимые для последующего определения принадлежности определённого набора к кластеру, а также поддержка (вероятность появления события) как количество наборов, вошедших в кластер, деленная на общее количество наборов.

**Задача 2. Отсеять те наборы, которые имеют низкий уровень поддержки (вероятность наступления некоего события ниже определенного уровня).**

Из памяти удаляются записи о тех кластерах, частота возникновения которых ниже заданного уровня.

**Задача 3. По заданным данным (по запросу пользователя) рассчитать условную вероятность определенного события.**

Пользователь задаёт набор показателей за известный (или желаемый) период (своеобразный аналог «Товара А») и шаг прогноза (своеобразный аналог «Товара Б»). Система среди сформированных кластеров выполняет поиск тех, параметры которых более всего соответствуют заданным условиям. На основе частоты возникновения кластеров рассчитывается условная вероятность наступления прогноза.

Алгоритм решения данной задачи включает следующие шаги.

1. На вход модели подаётся матрица показателей на период  $M$  ( $M$  определяет количество периодов, на основании которых необходимо построить прогноз – количество строк матрицы).
2. На вход модели подается значение  $P$  – количество периодов, на которые необходимо построить прогноз.
3. Из кластеров, полученных на этапе настройки, выбираются те, в которых  $N = M + P$  – среди отобранных кластеров производится поиск таких, в которых первые  $X$  элементов соответствуют значениям из матрицы  $M$  с установленным уровнем ошибки, а значение поддержки такого кластера записывается в переменную  $A$ .
4. Из кластеров, полученных на этапе настройки, выбираются те, в которых  $N = M$  с установленным уровнем ошибки, а значение поддержки такого кластера записывается в переменную  $B$ .
5. На выход модель выдаёт последние  $P$  строк найденного кластера, у которого  $N = M + P$  и значение вероятности выполнения прогноза как частное от деления значения переменной  $A$  на значение переменной  $B$  (согласно формуле расчета условной вероятности).

---

## 5. Заключение

Ассоциативные сети как инструмент моделирования и решения задач прогнозирования имеет довольно высокий потенциал, и подобные инструменты будут пользоваться всё больше и больше с ростом объёма анализируемых данных, с переходом к исследованию более сложных систем. Одним из примеров таких систем является СЭС произвольного региона. Кроме того, что данная система является довольно сложной для моделирования, решение связанных с данным объектом задач крайне важно для эффективного развития общества. Считаем, что данное направление исследований должно обеспечить разработку соответствующего алгоритма, его оптимизацию и внедрение в программный продукт с целью решения практических задач.

---

## Информация об авторе

**Валентин Григорьевский** – студент НТУУ «КПИ», УНК «ИПСА», группа СА-21.

## ПРОГНОЗИРОВАНИЕ НА ОСНОВЕ ГЕНЕТИЧЕСКОГО АЛГОРИТМА ОБУЧЕНИЯ

Леонид Гуляницкий, Анна Павленко

**Аннотация:** Исследуются методы обучения множеств правил на основе генетических алгоритмов для дальнейшего прогнозирования временных рядов. Предложен способ представления популяции генетического алгоритма, способы рекомбинации, формирования ниш, расчета приспособленности.

**Ключевые слова:** прогнозирование, временные ряды, генетический алгоритм, машинное обучение на основе правил, ниши, эволюционные алгоритмы

### Введение

Разработка и применение генетических алгоритмов и баз правил для решения задач машинного обучения в последнее время интенсивно развивается благодаря природной масштабируемости, параллелизму, устойчивости к шуму, универсальности вычислительной схемы и возможности использования эвристик для представления данных [Ханк et al, 2007].

### Задача прогнозирования временных рядов

Предлагается свести задачу прогнозирования временных рядов к задаче классификации: в заданном временном ряде ищутся закономерности – правила, которые используются для построения прогноза.

**Представление популяции хромосом.** Множество правил формируется генетическим алгоритмом и представляет собой конечную популяцию хромосом [Fernandez et al., 2010]. Хромосома состоит из произвольного числа условий  $\omega$  (атомарных подправил) и итогового значения  $\omega_c$ ,

$$\omega = (F_s - F_k) \in [y_b, y_{b+1}),$$

где  $F_t$  – значение временного ряда в момент времени  $t$ ;  $s, k$  – произвольные моменты времени,  $y_b, y_{b+1}$  – действительные числа, которые используются для ранжирования приращений,  $b = \overline{-m, m-1}$ .

Таким образом, полное правило имеет вид:

$$\Omega_j = \text{if} \left( \bigwedge_{i=1}^z \omega_i \right) \text{ then } \omega_c,$$

где  $\omega_i$  – атомарное подправило условной части,  $i$  – номер атомарного подправила,  $i = \overline{1, z}$ ,  $z$  – длина правила (количество атомарных подправил в правиле),  $\omega_c$  – атомарное подправило, которое используется в случае выполнения условия *if*.

**Приспособленность хромосомы** определяется с учетом количества положительных примеров  $\phi$ , которые покрывает правило, и количеством атомарных подправил  $z$ :

$$\varphi = \begin{cases} \alpha z + (1 - \alpha)\phi, & \phi > \delta \\ 0, & \phi \leq \delta \end{cases}$$

где  $\delta$  – минимальное допустимое количество положительных примеров,  $\alpha$  – параметр алгоритма.

Расчет степени приспособленности для каждого правила выполняется по принципу «непрямого попадания»: если определенный процент подправил покрывает пример положительно, то считается, что правило полностью покрывает рассмотренный пример.

**Рекомбинации.** Представление правила в виде неупорядоченного множества атомарных подправил позволяет определять точку сечения произвольно. Скрещивание происходит объединением частей двух

хромосом. С некоторой вероятностью наследники мутируют: могут изменяться моменты времени произвольных подправил  $s, k$  и границы приростов  $y_b, y_{b+1}$  с учетом допустимого пространства значений.

**Степень подобия хромосом. Диверсификация поиска.** Реализовано и исследовано несколько способов отбора для рекомбинации: панмиксия, инбридинг, аутбридинг, турнирный отбор. Для их реализации было введено понятие подобия хромосом – степени сходства их генотипов. Понятие расстояния между хромосомами определяется процентом подобных подправил и схемой их расположения относительно друг друга.

Отметим, что процедура нахождения подобия используется для исключения избыточности правил на этапе рекомбинации, при добавлении правил в популяцию на этапе отбора к выживанию с помощью процедуры вытеснения, на этапе отбора к рекомбинации для способов инбридинг и аутбридинг, но основным преимуществом использования введенной метрики является возможность формирования ниш.

Когда генетические алгоритмы используются для оптимизации, то целью, как правило, является возвращение единственного значения – наилучшего найденного решения. Фактически, если традиционный генетический алгоритм выполняется достаточное количество итераций, вся популяция обычно сходится к окрестности одного решения. Для прогнозирования такой подход непригоден, поскольку необходимо получить достаточно широкую базу правил для построения финального прогноза. Генетические алгоритмы, использующие механизмы формирования ниш, поддерживают наличие различных семейств правил в популяции и позволяют одновременную оптимизацию в нескольких областях пространства поиска за счет снижения конкуренции между достаточно разнородными особями.

**Построение прогноза.** Критерий остановки цикла при построении базы правил – количество проведенных итераций. Для построения прогноза разработан алгоритм разрешения конфликтов с учетом полного множества конфликтов и значений приспособленности хромосом для момента прогноза.

---

### Выводы и направления дальнейших исследований

Предложены подходы к представлению данных для задач прогнозирования, и функции приспособленности, а также разработаны специальный оператор рекомбинации и механизм формирования ниш при решении задач прогнозирования с использованием ГА. Дополнительного исследования требуют вопросы выбора оптимальных значений параметров алгоритма, построение прогноза с учетом влияния нескольких факторов, представленных временными рядами.

---

### Список литературы

- [Ханк et al, 2007] Ханк Д.Э., Уичерн Д.У., Райтс А.Дж. Бизнес-прогнозирование.– 7-е изд. / Пер. с англ. – М.: Издательский дом «Вильямс», 2003. – 656 с.
- [Fernandez et al, 2010] A Fernández, S García, J Luengo, E Bernadó-Mansilla, F Herrera. Evolutionary Computation, IEEE Transactions, 14 (6), 2010, pp. 913-941.

---

### Информация об авторах

**Гуляницкий Леонид Фёдорович** – д.т.н., ведущий научный сотрудник Института кибернетики им. В.М.Глушкова НАН Украины, пр-т Глушкова, 40, Киев, 03680, Украина; e-mail: leonhul.icyb@gmail.com

**Павленко Анна Игоревна** – аспирант Института кибернетики им. В.М.Глушкова НАН Украины, пр-т Глушкова, 40, Киев, 03680, Украина; e-mail: dmitrieva.anya@gmail.com

---

## ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ АЛГОРИТМА ОМК ДЛЯ ПРОГНОЗИРОВАНИЯ ТРЕТИЧНОЙ СТРУКТУРЫ ПРОТЕИНОВ МЕТОДАМИ РАСПАРАЛЛЕЛИВАНИЯ

Леонид Гуляницкий, Виталина Рудык

**Аннотация:** *Предлагается параллельный алгоритм оптимизации муравьиными колониями для решения задачи прогнозирования третичной структуры протеина исходя из его последовательности аминокислотных остатков. Проводится исследование эффективности предложенного алгоритма, обсуждаются результаты вычислительного эксперимента на суперкомпьютерах семейства SKIT.*

**Ключевые слова:** *комбинаторная оптимизация, прогнозирование третичной структуры протеина, методы оптимизации муравьиными колониями, параллельные алгоритмы, суперкомпьютеры SKIT.*

---

### Введение

Исследования в прикладных сферах часто приводят к разработке моделей проблем, которые удобно анализировать математическими методами. Так, на стыке математики и биологии возникла область вычислительной биологии, которая включает в себя совокупность математических моделей и методов для решения проблем, возникающих в биологии, генетике, фармакологии, медицине. При этом широко применяются статистические методы, анализ данных, методы комбинаторной оптимизации. Особая роль методов комбинаторной оптимизации выделяется в генетике, где молекулы ДНК и РНК кодируются последовательностями генов и возникают задачи оптимизации на строках, которые и приводят естественным образом к задачам комбинаторной оптимизации.

Рассматривается одна из задач вычислительной биологии – задача прогнозирования третичной структуры молекулы протеина в модели Дилла [Dill et al., 1995].

---

### Задача прогнозирования третичной структуры протеина

Цель задачи прогнозирования третичной структуры протеина – построить трехмерную форму молекулы (третичную структуру) исходя из ее состава – цепи последовательно соединенных аминокислот (первичной структуры белка). В модели Дилла для представления третичной структуры используется некоторая дискретная решетка, в узлах которой располагаются аминокислоты. Чтобы молекула оставалась связанной, стоящие рядом в последовательности аминокислоты должны располагаться в соседних узлах решетки. Все аминокислоты в соответствии с их физическими свойствами разделяют на гидрофобные и полярные. Между близко расположенными (в соседних узлах) гидрофобными аминокислотами возникают гидрофобные связи. Каждая структура в решетке обладает неположительной энергией, равной количеству гидрофобных связей в ней. Считается, что молекула принимает такую форму, в которой ее энергия достигает минимума. Формализация этой проблемы приводит к задаче комбинаторной оптимизации, которая является NP-сложной [Berger & Leighton, 1998].

---

### Параллельный алгоритм оптимизации муравьиными колониями для решения задачи

Метод оптимизации муравьиными колониями (ОМК) – разработанная для решения задач комбинаторной оптимизации схема приближенного алгоритма, которая имитирует процесс поиска колонией муравьев кратчайших путей к источнику пищи [Dorigo & Stützle, 2004]. При этом общение между особями осуществляется через феромонные следы, которые способствуют выбору оптимальных направлений движения. Сложность возникающей задачи комбинаторной оптимизации обуславливает актуальность вопросов разработки и исследования параллельных алгоритмов комбинаторной оптимизации.

Алгоритмы на основе метода ОМК разработаны в работах [Shmygelska & Hoos, 2005; Chu et al., 2005; Fidanova & Lirkov, 2008]. Разработанный нами параллельный алгоритм ОМК отличается от предложенных ранее, в частности используются другие схемы для обновления феромонной матрицы, подсчета

эвристических оценок, способы кодирования структур (использована q-кодировка, предложенная в [Гуляницкий & Рудык, 2013]). Его схема представлена на рис.1.

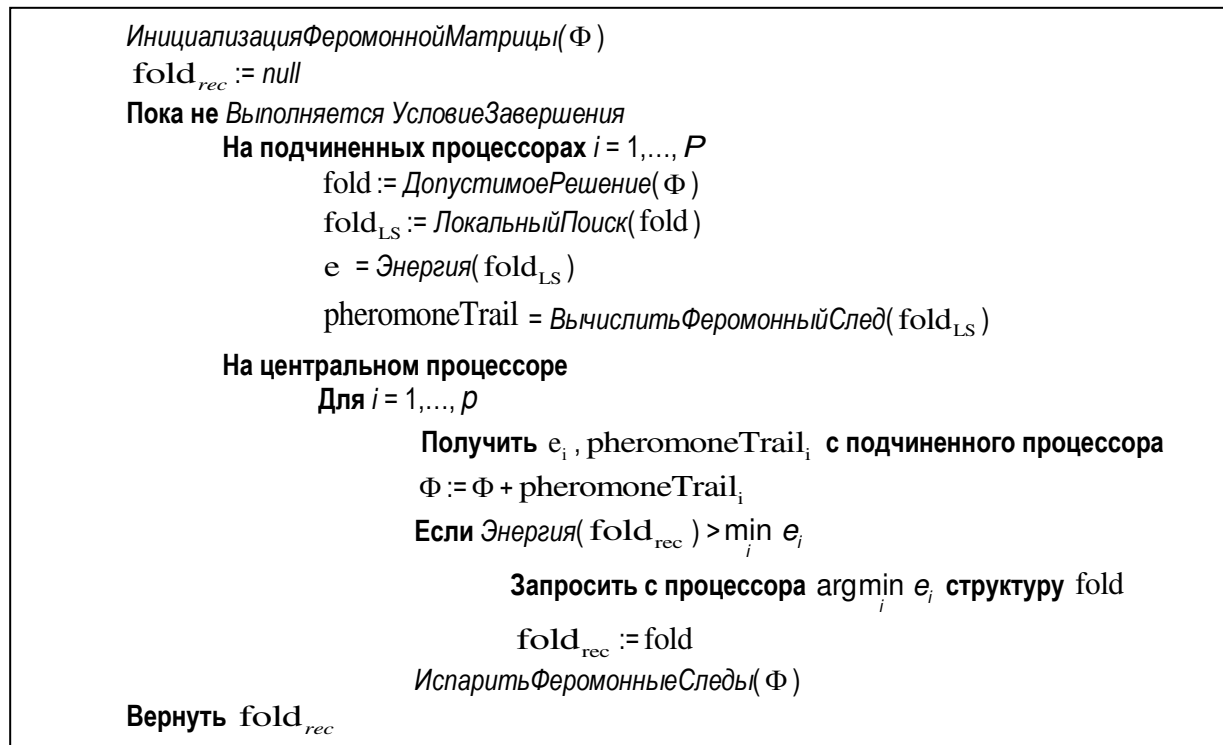


Рис. 1. Схема параллельного алгоритма

Выделяется один центральный процессор, остальные – подчиненные, попарные взаимодействия проходят между центральным и каждым из подчиненных процессоров. Разделение процессоров на центральный и подчиненные является условным, от архитектуры сети требуется только возможность быстрой передачи сообщений с одного из процессоров на остальные и обратно. Количество процессоров, задействованных в схеме, выбрано на один больше, чем количество элементов в популяции.

Первый шаг алгоритма – инициализация матрицы феромонных следов  $\Phi$  на центральном процессоре. Количество строк матрицы  $\Phi$  соответствует количеству элементов кода в выбранной решетке, количество столбцов – длине кода задачи. Изначально каждый элемент матрицы поставим равным некоторому значению  $\mu_0 > 0$ .

В процессе работы процедуры оптимальную (в контексте минимальной энергии) из найденных структур будем сохранять, обозначим ее  $fold_{rec}$ .

После инициализации запускается итерационный процесс. На каждой его итерации используя информацию, которую содержит феромонная матрица, строится новая популяция структур (количество элементов популяции является параметром алгоритма) после чего каждая из структур модифицирует феромонную матрицу, добавляя в нее феромонные следы, которые усилят пути, ведущие к низким за энергией структурам. Построение структуры начинается с первого элемента и проходит последовательно. Позиция аминокислоты выбирается случайным образом среди свободных с вероятностью

$$P_{i,d} = \frac{[\tau_{i,d}]^\alpha [\eta_{i,d}]^\beta}{\sum_{l \in D} [\tau_{i,l}]^\alpha [\eta_{i,l}]^\beta},$$

где  $D$  – множество элементов кода q-кодировки, которое отражает направления к свободным соседям текущего узла в заданной решетке,  $\tau_{i,d}$  – элемент феромонной матрицы, соответствующий направлению  $d$  для позиции  $i$ ,  $\eta_{i,d}$  – эвристическая оценка, которая зависит от уже построенной части структуры,  $\alpha$

и  $\beta$  – параметры алгоритма, задающие степень влияния феромонов и эвристической информации на построение решения.

Если на каком-то шаге процесс построения зашел в тупик – т.е. возникла ситуация, когда все соседние с крайней аминокислотой позиции уже заняты – делается откат назад на один шаг, а “тупиковое” направление добавляется в список табу. Генерация решений производится на подчиненных процессорах, которые получают на вход феромонную матрицу. Для повышения эффективности итерации каждый из подчиненных процессоров также выполняет процедуру локального поиска.

Далее, на каждом подчиненном процессоре вычисляется феромонный след  $pheromoneTrail$  в формате последовательности пар (индекс элемента матрицы – количество феромона)  $(d, \Delta_{i,d,c})$ , где  $d$  – направление, по которому расположен остаток в структуре  $c$ , а  $\Delta_{i,d,c}$  – относительное качество структуры  $c$  с учетом направления  $d$ . Для вычисления  $\Delta_{i,d,c}$  рассматриваются две схемы. В первой  $\Delta_{i,d,c}$  – относительное качество структуры, которое зависит от ее энергии. Вторая разработана с учетом того, что необходимо усиливать только те феромонные пути, которые влияют на энергию заданной молекулы. Сила  $\Delta_{i,d,c}$  определенного направления  $d$  в структуре  $c$  определяется как количество связей, в которых это направление задействовано. Такая схема позволяет более точно учитывать приемлемость той или иной части структуры молекулы.

Сгенерированная последовательность пар вместе со значением энергии наилучшего из найденных решения передается на центральный процессор, который анализирует и агрегирует полученную информацию: добавляет феромон к соответствующим элементам матрицы ( $\tau_{i,d} = \tau_{i,d} + \Delta_{i,d,c}^\gamma$ ,  $\gamma$  – параметр алгоритма), сравнивает значения энергий с сохраненным оптимальным значением энергии. Если нашлось меньшее значение энергии, производится дополнительный запрос на соответствующий процессор для получения структуры.

Для поддержания актуальности феромонной матрицы, следуя естественной аналогии, проводится процедура испарения феромона. Каждый элемент матрицы умножается на положительный параметр  $\rho$ , меньший единицы:  $\tau_{i,d} = (1 - \rho)\tau_{i,d}$ ,  $\rho$  – коэффициент испарения – характеризует, какая доля информации, собранной на предыдущих этапах, сохраняется. Испарение феромона также выполняется на центральном процессоре, после чего, если не выполняется условие завершения, производится следующая итерация алгоритма. В нашей реализации алгоритм завершается, когда оптимальное решение  $fold_{rec}$  не менялось на протяжении определенного количества итераций.

### Исследование эффективности параллельного алгоритма

Распараллеливание вычислений не всегда приводит к уменьшению времени вычислений. Поэтому важно провести анализ времени работы предложенной параллельной процедуры. Для этого введем следующие обозначения:

- $p$  – число подчиненных процессоров (равное количеству агентов в последовательном алгоритме);
- $t_{ini}$  – среднее время инициализации;
- $t_{slavelter}$  – среднее время, затрачиваемое подчиненным процессором на одну итерацию (включает в себя генерацию решения, локальный поиск и вычисления феромонных следов);
- $t_{phUpdate}$  – среднее время, затрачиваемое центральным процессором на обновление феромонной матрицы для одной структуры;

- $t_{vapor}$  - среднее время, затрачиваемое на модификацию (испарение) элементов феромонной матрицы;
- $t_{exch}$  - среднее время, затрачиваемое на обмен сообщениями центрального процессора с одним подчиненным процессором (включает пересылку феромонной матрицы с центрального процессора на подчиненный и пересылку феромонного следа с подчиненного процессора на главный);
- $I$  – количество итераций алгоритма.

Тогда среднее общее время выполнения параллельного алгоритма  $t_{parallel} = t_{ini} + (p(t_{exch} + t_{phUpdate}) + t_{slavelter} + t_{vapour}) \times I$ , при этом среднее общее время выполнения последовательного алгоритма  $t_{seq} = t_{ini} + (p(t_{phUpdate} + t_{slavelter}) + t_{vapour}) \times I$ .

Ускорение, которое будет получено при использовании параллельного алгоритма по отношению к последовательному равно

$$S_p = \frac{t_{ini} + (p(t_{phUpdate} + t_{slavelter}) + t_{vapour}) \times I}{t_{ini} + (p(t_{exch} + t_{phUpdate}) + t_{slavelter} + t_{vapour}) \times I} = 1 + \frac{((p-1)t_{slavelter} - pt_{exch}) \times I}{t_{ini} + (p(t_{exch} + t_{phUpdate}) + t_{slavelter} + t_{vapour}) \times I}.$$

При достаточно большом количестве итераций  $I$  временем инициализации  $t_{ini}$  можно пренебречь, тогда

$$S_p \approx 1 + \frac{(p-1)t_{slavelter} - pt_{exch}}{p(t_{exch} + t_{phUpdate}) + t_{slavelter} + t_{vapour}}.$$

Для дальнейшего анализа обозначим  $n$  число соседних узлов в рассматриваемой решетке, а  $m$  – число аминокислот в заданном протеине. Вычислим порядки времени выполнения операций в зависимости от значений  $n$  и  $m$ .

Начнем с времени обмена сообщениями. Центральный процессор отправляет на подчиненный процессор феромонную матрицу, которая состоит из  $m \times n$  элементов. В свою очередь, подчиненный процессор пересылает результат своих вычислений в форме последовательности пар (индекс элемента матрицы – количество феромона) – всего  $m \times 2$  элементов. Таким образом, объем пересылаемых сообщений –  $O(m \times n + m \times 2) = O(m \times n)$ . Предположив, что время пересылки сообщений линейно зависит от его размера, получим, что  $t_{exch} = O(m \times n)$ . При испарении феромона обновляется каждый из  $m \times n$  элементов феромонной матрицы, следовательно  $t_{vapour} = O(m \times n)$ .

Время, затрачиваемое на обновление феромонной матрицы одним агентом,  $t_{phUpdate} = o(m)$ , поскольку изменяется один элемент в каждом столбце феромонной матрицы (а их всего  $m$ ).

Осталось проанализировать время  $t_{slavelter}$ . В него входят две составляющих – время, которое требуется для генерации допустимого решения на основе заданной феромонной матрицы, и время вычисления феромонных путей, исходя из построенного решения. Начнем со второго. Процедура, вычисляющая количество феромонов, перебирает все гидрофобные контакты в структуре (а их количество линейно зависит от ее длины  $m$ ), и для каждого такого контакта прибавляет количество феромона на участке между элементами, образующими контакт (максимум  $m$  элементов). Таким образом, сложность описанной процедуры –  $O(m^2)$ .

Оценим время, затрачиваемое на генерацию допустимого решения. Напомним, что структура строится последовательным добавлением элементов, при невозможности такого добавления производится откат. В худшем случае процедура отката может свестись к полному перебору внутри некоторой области, а значит, сложность наихудшего случая –  $O(e^m)$ . На практике такие ситуации случаются редко, но следует

заметить, что проблема более актуальна в двумерных решетках, чем в трехмерных – в них более вероятно образование тупиков. Теоретический анализ средней сложности вычислений сводится к подсчету путей без самопересечений в определенной решетке, предполагается, что эта задача сама по себе является *NP*-сложной [Liśkiewicz et al., 2003]. Анализ вычислительных экспериментов показывает, что процедура генерации допустимого решения является наиболее времязатратной, особенно на последних итерациях алгоритма, когда феромонная матрица становится в значительной степени неоднородной.

Подытожим приведенные оценки:

$$S_p \approx 1 + \frac{(p-1)O(e^m) - pO(m \times n)}{p(O(m \times n) + O(m)) + O(e^m) + O(m \times n)}.$$

Из этого можно сделать два вывода. Во-первых, при достаточно больших  $m$  числитель больше нуля, следовательно ускорение при использовании параллельного алгоритма больше единицы. Во-вторых, при возрастании  $m$   $S_p$  тоже возрастает и приближается к  $p$ , что доказывает целесообразность использования МВК для задач больших размерностей.

---

### Выводы и направления дальнейших исследований

---

Для решения задачи прогнозирования третичной структуры сворачивания протеина на МВК разработан и реализован параллельный алгоритм на основе схемы ОМК. Получены оценки ускорения параллельного алгоритма по сравнению с аналогичным последовательным алгоритмом. Направлением дальнейших исследований будет сопоставление теоретических результатов с реальными, полученными путем вычислительного эксперимента.

---

### Литература

---

- [Berger & Leighton, 1998] B. Berger, T. Leighton. Protein folding in the hydrophobic-hydrophilic (HP) model is NP-complete // Journal of Computational Biology, 1998, 5(1), pp. 27-40.
- [Chu et al, 2005] D. Chu, M. Till, A. Zomaya. Parallel Ant Colony Optimization for 3D Protein Structure Prediction using the HP Lattice Model, 19th IEEE International Parallel and Distributed Processing Symposium (IPDPS'05), 2005, 7, pp.193-200.
- [Dill et al, 1995] 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, pp. 561– 602.
- [Dorigo & Stützle, 2004] M. Dorigo, T. Stützle. Ant Colony Optimization, Cambridge: MIT Press, MA, 2004, 348 p.
- [Fidanova & Lirkov, 2008] S. Fidanova, I. Lirkov. Ant Colony System Approach for Protein Folding, Int. Conf. Multiconference on Computer Science and Information Technology, 2008, pp. 887–891.
- [Liśkiewicz et al, 2003] M. Liśkiewicz, M. Ogihara, S. Todac. The complexity of counting self-avoiding walks in subgraphs of two-dimensional grids and hyper cubes, Theoretical Computer Science, 2003, 304, pp. 129-156.
- [Shmygelska & Hoos, 2005] A. Shmygelska, H. Hoos. An ant colony optimization algorithm for the 2D and 3D hydrophobic polar protein folding problem, BMC Bioinformatics, 2005, 6(30), pp. 30–52
- [Гуляницький & Рудык, 2013] Л. Гуляницький, В. Рудык. Проблема предсказания структуры протеина: формализация с использованием кватернионов // КиСА. – 2013. – 4. – С.130-137.

---

### Информация об авторах

---

**Леонид Гуляницький (Hulianytskyi)** – д.т.н., заведующий отделом Института кибернетики им. В.М.Глушкова НАН Украины, пр-т Глушкова, 40, Киев, 03680, Украина; e-mail: leonhul.icyb@gmail.com

**Виталина Рудык (Rudyk)** – младший научный сотрудник Института кибернетики им. В.М.Глушкова НАН Украины, пр-т Глушкова, 40, Киев, 03680, Украина; e-mail: vitalina.rudyk@gmail.com



---

## К ВОПРОСУ ОБ ОСОБЕННОСТЯХ ПОСТРОЕНИЯ ЗНАНИЕОРИЕНТИРОВАННЫХ СОЦИАЛЬНЫХ СЕТЕЙ В ИНТЕРНЕТЕ В СФЕРЕ ОБУЧЕНИЯ

**Андрей Данилов**

**Аннотация:** Рассмотрены аспекты разработки знаниеориентированных нишевых социальных сетей в области образования для повышения скорости обучения и качества образования за счет использования современного ноосферного метода – системологического классификационного анализа.

**Ключевые слова:** управление знаниями, социальные сети в Интернете, системология, системологический классификационный анализ.

---

### Введение

Развитие информационных знаниеориентированных технологий привело к широкому внедрению инструментов управления знаниями в компаниях различных отраслей деятельности. Особенно актуально использование методов и технологий управления знаниями в сферах связанных с обучением и управлением персоналом.

Использование нишевых социальных сетей в Интернете, как одного из наиболее эффективных инструментов управления знаниями, позволит повысить конкурентоспособность сотрудников за счет постоянного обмена и приобретения знаний из внешних источников. В данной работе рассмотрим некоторые особенности разработки и применения социальных сетей в Интернете в сфере обучения с использованием системологического классификационного анализа.

---

### Актуальность применения социальных сетей в Интернете в сфере обучения

Для разработки эффективных образовательных социальных сетей в Интернете целесообразно применять современные научные методы, которые соответствуют ноосферному этапу развития науки и ориентированы на работу с содержательной информацией. Наиболее эффективным современным методом решения сложных неформализованных задач является метод системологического классификационного анализа. Применение метода системологического классификационного анализа при формировании классификаций функций и системы ссылок внутри ресурса позволяет повысить скорость доступа к необходимой информации, а также уменьшает нагрузку на пользователя за счет размещения информационных блоков и пунктов меню согласно их функциональному назначению.

Применение предложенных социальных сетей в Интернете для обучения подразумевает не только создание учебных материалов в среде Интернет и самостоятельную работу сотрудников, а также возможность непосредственного контакта между сотрудниками, тренерами и консультантами в любое время суток. Такой подход в обучении позволит организовать обучение в виде мастер класса, что значительно повысит эффективность обучения за счет прямого контакта между участниками сети и приглашенными внешними «наставниками-консультантами» [1].

Применение знаниеориентированных социальных сетей в процессе обучения позволит организовать индивидуальное обучение без личного присутствия сотрудника на рабочем месте, что будет способствовать повышению качества профессиональной подготовки, позволит людям с ограниченными возможностями получить качественное образование в комфортных условиях[1].

## Разработка модели передачи знаний для функционирования социальных сетей в Интернете в сфере обучения

Для понимания эффективности использования социальных сетей, как инструмента обучения, в том числе в рамках самообучающейся организации, была разработана модель распространения и передачи знаний в социальных сетях.

Рассмотрим требования, которым должна удовлетворять социальная сеть в организации, точнее те функции обучения, которые мы должны рассматривать. На рисунке 1 рассмотрены основные возможности использования социальных сетей в Интернете по обучению сотрудников.

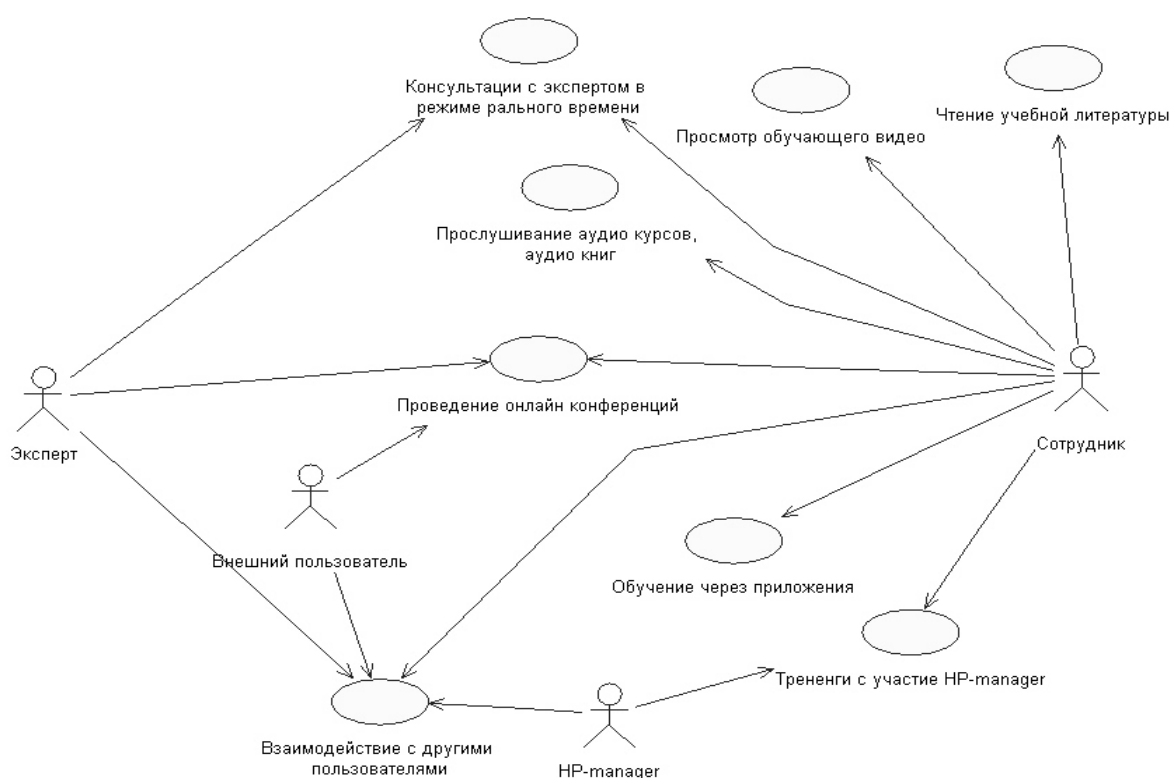


Рис. 1. Возможности использования социальных сетей для обучения сотрудников компании

### Выводы

Использование социальных сетей в Интернете для обучения специалистов позволит повысить скорость и качество обучения, профессиональный уровень общества и создать мощный фундамент для дальнейшего развития сферы образования в Украине.

### Список использованной литературы

1. Данилов А.Д. К вопросу об актуальности разработки знаниеориентированных социальных сетей в сфере образования. / 18-й Международный молодежный форум «Радиоэлектроника и молодежь в XXI веке». Сб. материалов форума. Т.9. Международная конференция «Управление знаниями и конкурентная разведка» - Харьков: ХНУРЭ. 2014. – с.50 - 51.

### Сведения об авторе

**Данилов Андрей Дмитриевич** – аспирант, ст. преподаватель кафедры СИ ХНУРЭ; e-mail: Skil06@ukr.net

## МАТРИЧНАЯ МНОЖЕСТВЕННАЯ РЕГРЕССИЯ

Владимир Донченко, Инна Назарага, Ольга Тарасова

**Аннотация:** Предложен к рассмотрению класс матричных функций набора матричных аргументов: матричную множественную регрессию. В рамках развития концепции кортежных операторов разработан математический аппарат сингулярного (SVD) разложения и техники псевдообращения (ПдО) за Муром-Пенроузом, который позволяет поставить и конструктивно решить задачу оценки методом наименьших квадратов (МНК) для множественной матричной регрессии.

**Ключевые слова:** матричная множественная регрессия, псевдообращение по Муру – Пенроузу.

### Введение

Задача группирования информации [Donchenko et al, 2012]: задача восстановления зависимостей и задача классификации, кластеризации и распознавания образов, - это задача принципиальной важности в прикладных исследованиях. Важным классом математических структур [Донченко, 2011], которые используются в решении такой задачи, являются евклидовы пространства  $R^n$  числовых векторов. Такое использование определяется, с одной стороны - богатством структурных связей евклидовых пространств в целом, с другой – развитой техникой оперирования в этих пространствах, включая сингулярное (SVD) разложение и псевдообращение (ПдО) по Муру – Пенроузу [Moore, 1920], [Penrose, 1955] (см. также [Albert, 1972]) в этих пространствах. Отметим выдающуюся роль Н.Ф. Кириченко и, в частности, его работу [Kirichenko, 1997], которая развивает ПдО-технику в  $R^n$ . Тем не менее, использование евклидовых пространств числовых векторов становится ограничивающим в решении прикладных задач, которыми являются, в частности, задачи распознавания речи и обработки изображений. В этих задачах естественным представителем исследуемых объектов являются матрицы: спектрограммы и матрицы изображений соответственно. В работе [Donchenko, Zinko, Skotarenko, 2012] предложена концепция кортежных операторов, которая позволяет перенести SVD – и ПдО – техники на евклидовы пространства матриц фиксированной размерности. Там же упомянутые техники использованы для построения расстояний соответствия в задачах кластеризации. В работах [Donchenko, Nazaraga, Tarasova, 2013-1], [Donchenko, Nazaraga, Tarasova, 2013-2] развит конструктивный аппарат метода наименьших квадратов для векторных и матричных наблюдений с примерами прикладного использования для макроэкономических показателей и микроэкономических показателей. В предлагаемой работе предлагается расширение концепции кортежных операторов, которое позволяет поставить и конструктивно решить задачу МНК-оценивания для матричных функций нескольких матричных аргументов.

### Теоретические основы матричной множественной регрессии

1. Теорема В1. (Базовая теорема про SVD-разложение [Donchenko, Zinko, Skotarenko, 2012].) Пусть  $A$  - линейный оператор между двумя абстрактными евклидовыми пространствами  $A: E_1 \rightarrow E_2$ ,  $A^*: E_2 \rightarrow E_1$  - сопряженный к нему,  $r = \text{rank} A = \dim \mathfrak{R}(A)$ . Тогда:

- общее количество ненулевых сингулярностей (пары собственный вектор – собственное число) двух операторов  $AA^*$ ,  $A^*A$  равняется  $r$ ;
- набор собственных чисел  $\lambda_1^2 \geq \dots \geq \lambda_r^2 > 0$  упомянутых операторов является общим;
- для  $(u_i, \lambda_i^2), (v_i, \lambda_i^2), i = \overline{1, r}$  справедливы соотношения:

$$u_i = \lambda_i^{-1} A v_i = \lambda_i^{-1} A^*, i = \overline{1, r}; \quad (1)$$

• оператор может быть представлен в таком виде:  $Ax = \sum_{i=1}^r \lambda_i u_i(v_i, x), x \in E_1$  (2)

2. Определение В2. Псевдо-обратным (ПДО) к оператору  $A: E_1 \rightarrow E_2$  - будем обозначать  $A^+$  - называется оператор  $A^+: E_2 \rightarrow E_1$ , который определяется соотношением:  $A^+y = \sum_{i=1}^r \lambda_i^{-1} v_i(u_i, y), y \in E_2$ .

3. Теорема В3. Значение  $A^+y$  является минимальным по норме решением оптимизационной задачи минимизации для  $x \in E_1$  функционала  $\|Ax - y\|^2$ :  $A^+y = \underset{z \in \text{Argmin}_{x \in E_1} \|Ax - y\|^2}{\text{argmin}} \|z\|^2$ .

4. Теорема В4. (Исследование СЛАУ.) В принятых выше обозначениях, необходимым и достаточным условием разрешимости относительно системы линейных алгебраических уравнений (СЛАУ)  $Ax = y, y \in E_2$ , является условие  $(y, (I_{E_2} - AA^+)y) = 0$ , а множество всех решений имеет такой вид:

$$A^+y + (I_{E_1} - A^+A)E_1. \quad (3)$$

В случае неразрешимости СЛАУ  $(y, (I_{E_2} - AA^+)y) > 0$ , множество, определяемое соотношением (3), описывает множество  $\underset{x \in E_1}{\text{argmin}} \|Ax - y\|^2$ , то есть множество всех решений оптимизационной задачи наилучшего квадратичного приближения правой части СЛАУ значениями левой части. Все элементы множества (3) в этом случае называются псевдо-решениями.

5. Множественную матричную регрессию (ММР) с матричными значениями определим в детерминированной постановке: как матричную функцию  $Y$ , определяемую матричным кортежем-строкой  $\alpha_{1,K} = (X_1 : \dots : X_K)$ ,  $X_k \in R^{m \times n}$ ,  $k = \overline{1, K}$ , вида  $Y(\alpha_{1,K}) = \beta_1 X_1 + \dots + \beta_K X_K, \beta_k \in R^1, k = \overline{0, K}$ , подлежащей определению на основе наблюдений  $(\alpha_{1,K}^{(i)}, Y_i), i = \overline{1, N}$ , где  $Y_i = Y(\alpha_{1,K}^{(i)}), \alpha_{1,K}^{(i)} = (X_1^{(i)} : \dots : X_K^{(i)}), i = \overline{1, N}$ .

6. Отображение  $\wp_{\alpha_{1,K}}: R^K \rightarrow R^{m \times n}$ , которое определяется соотношением:

$$Y(\alpha_{1,K}) = \beta_1 X_1 + \dots + \beta_K X_K \equiv \wp_{\alpha_{1,K}} \beta, \beta^T = (\beta_1, \dots, \beta_K), \beta \in R^K, \quad (4)$$

будем называть  $\wp_{\alpha_{1,K}}$ , следуя [Donchenko, Zinko, Skotarenko, 2012], кортежным оператором.

7. Будем также использовать обозначения:

- $\chi_Y$  - матричный кортеж - столбец, полученный из значений наблюдений  $Y_i, i = \overline{1, N}$ :

$$\chi_Y = \begin{pmatrix} Y_1 \\ \dots \\ Y_N \end{pmatrix}, i = \overline{1, N};$$

- $\chi_{N,j}^{(j)}, j = \overline{1, K}$ , - кортежи - столбцы длины  $N$ , полученные из  $j$ -тых компонент кортежей-строк  $\alpha_{1,K}^{(i)}, i = \overline{1, K}$ ;

- $\alpha_{N,K}$  - кортеж - строка длины  $K$ , полученный из кортежей - столбцов  $\chi_{N,1}^{(j)}, j = \overline{1, K}$ :  
 $\alpha_{N,K} = (\chi_{N,1}^{(1)} \dots \chi_{N,1}^{(K)})$ ;
- $\wp_{\alpha_{N,K}} : R^K \rightarrow R^{N,m \times n}$  - матричный кортежный оператор из евклидова пространства  $R^K$  в евклидово пространство кортежей - столбцов длины  $N$  из  $m \times n$  матриц с покомпонентным скалярным произведением.

8. Оценка  $\hat{\beta} \in R^K$  метода наименьших квадратов неизвестного параметра матричной множественной регрессии с матричными значениями естественным образом определяется как решение оптимизационной задачи минимизации функционала

$\sum_{i=1}^N \|Y_i - \wp_{\alpha_{1,K}^{(i)}} \beta\|_{m \times n}^2$  метода наименьших квадратов:

$$\hat{\beta} = \operatorname{argmin} \|\beta\|^2 \quad \beta \in \operatorname{Argmin}_{\beta \in R^K} \sum_{i=1}^N \|Y_i - \wp_{\alpha_{1,K}^{(i)}} \beta\|_{m \times n}^2 \quad (5)$$

9. Стандартным образом МНК-оценку  $\hat{Y}(\kappa_{1,K}), \kappa_{1,K} = (X_1 \dots X_K), X_k \in R^{m \times n}, k = \overline{1, K}$ , матричной множественной регрессии будем определять соотношением

$$\hat{Y}(\kappa_{1,K}) = \wp_{\kappa_{1,K}} \hat{\beta} = \sum_{k=1}^K \hat{\beta}_k X_k, X_k \in R^{m \times n}, k = \overline{1, K}.$$

10. Лема 1. Функционал метода наименьших квадратов матричной регрессии  $A(\beta)$  может быть представлен в виде:

$$A(\beta) = \sum_{i=1}^N \|Y_i - \wp_{\alpha_{1,K}^{(i)}} \beta\|_{m \times n}^2 = \|\chi_Y - \sum_{j=1}^K \beta_j \chi_{N,1}^{(j)}\|_{N, m \times n}^2 = \|\chi_Y - \wp_{\alpha_{N,K}} \beta\|^2. \quad (6)$$

11. Теорема 1. Решением оптимизационной задачи (5) является  $\hat{\beta}$ , которое определяется соотношением:

$$\hat{\beta} = \wp_{\alpha_{N,K}}^+ \chi_Y. \quad (7)$$

Доказательство Теоремы 1 вытекает из общего варианта SVD-разложения, определения ПдО и исследования СЛАУ для линейных операторов между абстрактными евклидовыми пространствами (П.3).

12. Обозначим  $(v_i, \lambda_i^2), i = \overline{1, r}$ , набор ненулевых сингулярностей матрицы Грамма  $\Gamma = ((\chi_{N,1}^{(i)}, \chi_{N,1}^{(j)})_{N, m \times n})_{i,j=\overline{1,K}}$ , набора  $\chi_{N,1}^{(i)}, i = \overline{1, K}$ ,  $r = \operatorname{rank} \Gamma, \lambda_1^2 \geq \dots \geq \lambda_r^2 > 0$ , а  $(\chi_{U,i}, \lambda_i^2), i = \overline{1, r}$ , - сингулярности оператора  $\wp_{\alpha_{N,K}} \wp_{\alpha_{N,K}}^*$ . В соответствии с общей теоремой о SVD-разложении

$$\chi_{U,i} = \frac{\wp_{\alpha_{N,K}} v_i}{\lambda_i}, i = \overline{1, r}.$$

13. Теорема 3. Оценка МНК вектора неизвестных параметров определяется соотношением

$$\hat{\beta} = \wp_{\alpha_{N,K}}^+ \chi_Y = \sum_{k=1}^r \lambda_k^{-1} v_k (\chi_{U,i}, \chi_Y)_{N, m \times n}.$$

Доказательство Теоремы 3 получается прямым применением Теоремы В1.

## Заключение

Предложен к рассмотрению класс матричных функций набора матричных аргументов: матричная множественная регрессия. В рамках развития концепции кортежных операторов разработан

---

математический аппарат сингулярного (SVD) разложения и техники псевдообращения по Муру-Пенроузу для кортежных операторов. Предложенное развитие концепции позволяет поставить и конструктивно решить задачу МНК-оценивания для множественной матричной регрессии.

---

### Литература

---

- [Moore, 1920] Moore E.H. On the reciprocal of the general algebraic matrix. Bulletin of the American Mathematical Society. – 26, 1920. – P.394 -395.
- [Penrose, 1955] Penrose R. A generalized inverse for matrices // Proceedings of the Cambridge Philosophical Society 51, 1955. – P.406-413.
- [Albert, 1972] Albert A.E. Regression and the Moore-Penrose pseudoinverse.-Academic Press., 1972, 180 p.
- [Kirichenko, 1997] N.F. Kirichenko. Analytical Representation of Perturbation of Pseudoinverse Matrices. Cybernetics and Systems Analysis. - March-April 1997. - Vol.33, No 2. - P.230-239.
- [Донченко, 2011] Владимир Донченко. Евклидовы пространства числовых векторов и матриц: конструктивные методы описания базовых структур и их использование. International Journal "Information technologies & Knowledge".- 2011.- Vol. 5.- Number 3.-P.203-216.
- [Donchenko, Krak, Krivonos, 2012] Donchenko V. Krak Ju., Krivonos Yu. Recurrent procedure in solving the Grouping Information Problem in Applied Mathematics//ITHEA: International Journal Information Theories and Applications. – Vol.1, No 1. – 20012. – P.62 -77.
- [Donchenko, Zinko, Skotarenko, 2012] Donchenko V, Zinko T, Skotarenko F. "Feature Vectors" in Grouping Information Problem in Applied Mathematics: Vectors and Matrixes. - Vitalii Velichko, Alexey Voloshin, Krassimir Markov(ed.) - Problems of Computer Intellectualization.- Institute of Cybernetics NASU, ITHEA.-Kyiv, Ukraine -Sofia, Bulgaria.- 2012.- P.111-124.
- [Donchenko, Nazaraga, Tarasova, 2013-1] Donchenko V, Nazaraga I, Tarasova O. Vectors and matrixes least square method: foundation and application examples. International Journal Information Theories and Applications. – 2013. – Vol.20, No 4.– P.311-322.
- [Donchenko, Nazaraga, Tarasova, 2013-2] Donchenko V, Nazaraga I, Tarasova O. Matrixes least squares method and examples of its application. International Journal Information Technologies & Knowledge. – 2013. – Vol.7, No 4. – P. 325–336.

---

### Сведения об авторах

---

**Владимир Донченко** – доктор физ.-мат. наук, профессор, Киевский национальный университет имени Тараса Шевченко, факультет кибернетики, Украина; e-mail: voldon@bigmir.net.

**Инна Назарага** – кандидат технических наук, младший научный сотрудник, Киевский национальный университет имени Тараса Шевченко, факультет кибернетики, Украина; e-mail: pna\_na@ukr.net.

**Ольга Тарасова** – аспирант, Киевский национальный университет имени Тараса Шевченко, факультет кибернетики, Украина; e-mail: olga\_ta@bigmir.net.

## РАЗРАБОТКА ХРАНИЛИЩА ДАННЫХ ДЛЯ ИНФОРМАЦИОННО-РЕСУРСНОГО ЦЕНТРА

**Анна Жолнарская**

**Аннотация:** В докладе определяется понятие "Информационно-ресурсный центр", который является техническим администратором единых и государственных реестров. Рассматриваются основные функции данного центра такие, как обеспечение функционирования реестров, сохранность и защита данных. Определяется специфика информации, которая передается в хранилище данных. Исследование проводится с целью последующего определения концепций организации хранилища данных для информационного центра.

**Ключевые слова:** Информационный центр, хранилище данных, специфика информации, обработка информации, реестр.

### Введение

Данная работа содержит результаты исследований, посвященных разработке теоретико-методологических основ создания информационных хранилищ для информационно-аналитических систем. *Объектом* исследования является процесс организации хранилища данных для информационно-ресурсного центра. *Предмет* исследования – средства организации хранилищ данных в информационно-аналитических системах. *Цель* исследования – определение концепций организации хранилищ данных в информационно-аналитических системах. В докладе также рассмотрены технологии и принципы организации хранилища данных в информационно аналитических системах.

### Понятие «Информационно-ресурсный центр» и его основные задачи

В начале 80-х годов, в период бурного развития регистрирующих информационных систем, появилось осознание ограниченности их применения для анализа данных и построения систем поддержки принятия решений. Регистрирующие системы создавались для автоматизации рутинных операций: выписки счетов, оформления договоров, проверки состояния предприятия и т.д., и предназначались для линейного персонала. Основными требованиями к таким системам были обеспечение транзакционности вносимых изменений и максимизация скорости, что и определило тогда выбор реляционных СУБД и модели представления данных «сущность-связь» в качестве основных технических решений при построении регистрирующих систем.

Для менеджеров и аналитиков, в свою очередь, требовались системы, которые бы позволяли: анализировать информацию во временном аспекте, формировать произвольные запросы к системе, обрабатывать большие объемы данных, интегрировать данные из различных регистрирующих систем.

«Информационно-ресурсный центр» является предприятием государственной формы собственности, осуществляет свою хозяйственную деятельность с 2005 года.

Основными рабочими целями данной организации является предоставление информационной, технической поддержки, а также консультаций разнообразным государственным или негосударственным организациям.

Одной из основных задач информационного центра является обеспечение функционирования реестров:

- Единого государственного реестра юридических лиц и физических лиц-предпринимателей (ЕГР)
- Реестра документов разрешительного характера (РДРХ)
- Единого лицензионного реестра (ЕЛР).

Каждый из вышеперечисленных реестров представляет собой автоматизированную систему сбора, накопления, защиты, учета и предоставления информации.

Обобщая вышеизложенное, можно сделать вывод, что информационно-ресурсный центр - это комплекс, включающий информационные, программно-аппаратные и научно-методические средства, а также современное оборудование и иные виды ресурсов для организации процесса функционирования реестров, сохранности и защиты данных.

Организация таких ресурсных центров дала потребность в появлении технологии хранилищ данных.

---

## Технологии и концепции организации хранилищ данных в информационно-аналитических системах

---

В основе концепции хранилища данных лежат две основные идеи:

- Интеграция разьединенных детализированных данных (детализированных в том смысле, что они описывают некоторые конкретные факты, свойства, события и т.д.) в едином хранилище. В процессе интеграции должно выполняться согласование рассогласованных детализированных данных и, возможно, их агрегация. Данные могут поступать из исторических архивов, оперативных баз данных, внешних источников.
- Разделение наборов данных и приложений, используемых для оперативной обработки и применяемых для решения задач анализа.

Интеллектуальный анализ информации в информационно-аналитических системах часто рассматривают как естественное развитие концепции хранилищ данных, в котором накапливаются все необходимые данные для осуществления определенных задач стратегического управления разных периодов.

Ключевыми компонентами такой организации являются технологии OLAP и Data Mining.

Технология OLAP и Data Mining "обогащают" друг друга. Эти два вида анализа должны быть тесно объединены, чтобы интегрированная технология могла обеспечивать одновременно многомерный доступ и поиск закономерностей. По словам N.Raden, "многие компании создали ... прекрасные хранилища данных, идеально разложив по полочкам горы неиспользуемой информации, которая сама по себе не обеспечивает ни быстрой, ни достаточно грамотной реакции на рыночные события".

Средство организации хранилищ данных в информационно-аналитических системах, основанное на методах искусственного интеллекта и инструментах поддержки принятия решений – называется Data Mining.

Функционирование Data Mining состоит в выявлении скрытых правил и закономерностей в больших массивах данных.

Методы Data Mining позволяют обнаружить стандартные закономерности. Выделяют пять стандартных типов закономерностей: ассоциация, последовательность, классификация, кластеризация и прогнозирование.

В OLAP, также как и в Data Mining, реализуется сложный интеллектуальный анализ данных с возможностью использования хорошо зарекомендовавших себя методов математической статистики, нейронных сетей, машинного обучения, визуализации данных, индукции правил, нечеткой логики, генетических алгоритмов.

В основе OLAP системы лежит многомерное концептуальное представление данных. Такая система, основана на технологии аналитической обработки данных в режиме реального времени (On-Line Analysis Processing).

В OLAP-системах основное внимание уделено гибкости доступа и манипулирования информацией.

В Data Mining, в отличие от OLAP, основную задачу формулирования гипотез и выявление необычных шаблонов выполняет компьютер. Data Mining позволяет выявить новые знания, которые невозможно получить методами статистического, регрессионного анализа или эконометрики.

На сегодняшний день немногие производители реализуют Data Mining для многомерных данных. Кроме того, некоторые методы Data Mining, например, метод ближайших соседей или байесовская классификация, в силу их неспособности работать с агрегированными данными неприменимы к многомерным данным.

Преимуществом этих технологий в том, что они постоянно развиваются, привлекают к себе все больший интерес со стороны научного мира, так и со стороны применения достижений технологий в различных отраслях.

---

## Проблемы интеграции данных

---

Остановимся на некоторых проблемах реализации хранилища данных:

- Неоднородность программной среды
- Распределенный характер организации



- Повышенные требования к безопасности данных
- Необходимость наличия многоуровневых справочников метаданных
- Потребность в эффективном хранении и обработке очень больших объемов информации .

При выборе средств реализации хранилища данных приходится учитывать множество факторов, включающих уровень совместимости различных программных компонентов, легкость их освоения и использования, эффективность функционирования. Приходится также решать вопросы аутентификации пользователей, защиты данных при их перемещении в хранилище данных из оперативных баз данных и внешних источников, защиты данных при их передаче по сети.

Безопасность при хранении и пересылке данных является одним из основных вопросов при работе с конфиденциальными и приватными данными. Возникшую необходимость можно организовать с помощью шифрованного хранилища для удаленного размещения файлов – так называемое облачное решение.

Надёжность, своевременность получения и доступность данных в облаке очень сильно зависит от многих промежуточных параметров, таких как: каналы передачи данных на пути от клиента к облаку, надёжность последней мили, качество работы Интернет-провайдера, доступность самого облака в данный момент времени.

Облачное хранилище данных — модель онлайн-хранилища, в котором данные хранятся на многочисленных распределённых в сети серверах, предоставляемых в пользование клиентам, в основном, третьей стороной. В противовес модели хранения данных на собственных выделенных серверах, приобретаемых или арендуемых специально для подобных целей, количество или какая-либо внутренняя структура серверов клиенту, в общем случае, не видна. Данные хранятся, а равно и обрабатываются, в так называемом облаке, которое представляет собой, с точки зрения клиента, один большой виртуальный сервер. Физически же такие серверы могут располагаться удалённо друг от друга географически, вплоть до расположения на разных континентах.

---

## Заключение

Определено понятие «Информационно-ресурсного центра». Проанализированы его основные функции и рабочие цели. Продемонстрированы технологии OLAP и Data Mining, которые являются ключевыми компонентами организации хранилищ данных. В Data Mining, в отличие от OLAP, основную задачу формулирования гипотез и выявление необычных шаблонов выполняет компьютер. Data Mining позволяет выявить новые знания, которые невозможно получить методами статистического, регрессионного анализа или эконометрики. Преимущество этих технологий в том, что они постоянно развиваются, привлекают к себе все больший интерес, как со стороны научного мира, так и со стороны применения достижений технологий в различных отраслях.

Данные исследования должны определить концепции для организации хранилища данных для информационного центра.

---

## Литература

- [Inmon, 1992] Inmon W. Building the Data Warehouse, New York: John Wiley & Sons, 1992.
- [Спирли, 2001] Спирли, Эрик. Корпоративные хранилища данных. Планирование, разработка, реализация. Том. 1: Пер. с англ. – М.: Издательский дом "Вильямс", 2001. – 179 с.
- [Дюк и др., 2001] Дюк В, Самойленко А, Data mining: учебный курс. – СПб: Питер, 2001.- 368 с.
- [Kimbal, 1996] Ralph Kimball. The Data Warehouse Toolkit: Practical Techniques for Building Dimensional Data Warehouses. John Wiley 1996.
- [Codd et al., 1993] E.F. Codd, S.B. Codd, C.T. Salley. Providing OLAP (On-line Analytical Processing) to User-Analysts: An IT Mandate.. Codd & Date, Inc, 1993. Retrieved on 2008. [http://www.olap.ru/basic/olap\\_arch.asp](http://www.olap.ru/basic/olap_arch.asp)
- [de Ville, 2001] B. de Ville. Microsoft Data Mining. Digital Press, 2001.
- [Codd et al., 1993] E. F. Codd, S.B.Codd. Providing OLAP. On-line Analytical Processing to User-Analysts: An IT Mandate. C. T. Salley, E. F. Codd & Associates, 1993.

---

## Информация об авторе

**Анна Жолнарская** - Главное управление региональной статистики Украины, 01601, г. Киев, ул. Эспланадная, 4-6 (Ст. М "Дворец Спорта"); e-mail: [azh2002@ukr.net](mailto:azh2002@ukr.net)

## ПРОГНОЗИРОВАНИЕ РИСКА БАНКРОТСТВА КОРПОРАЦИЙ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ С ИСПОЛЬЗОВАНИЕМ НЕЧЕТКИХ НЕЙРОННЫХ СЕТЕЙ

Юрий Зайченко, Ови Нафас Агаи аг Гашиш

**Аннотация:** В статье рассматривается проблема анализа риска банкротства корпораций в условиях неопределенности. Для ее решения предлагается применение каскадных нео-фаззи нейронных сетей. Приводится структура каскадной нео-фаззи нейросети, описаны ее свойства и алгоритмы обучения в пакетном и оперативном режимах. Приводятся результаты экспериментального исследования их применения в задаче прогнозирования риска банкротства предприятий Украины и сравнительный анализ с результатами использования классических методов Альтмана и Давыдовой-Беликова, также нечеткими нейронными сетями Мамдани и Цукамото.

**Ключевые слова:** прогнозирование риска банкротства корпораций, каскадные нео-фаззи нейронные сети, нечеткие нейронные сети Мамдани и Цукамото.

### Введение

Одной из актуальных проблем, связанных со стратегическим менеджментом, является анализ финансового состояния и оценка риска банкротства предприятий (корпораций).

Своевременное выявление признаков возможного банкротства позволяет руководству принимать срочные меры по исправлению финансового состояния и снижению риска банкротства.

В течение многих лет **классические статистические методы** широко использовались для прогнозирования рисков банкротства. Они включают процедуру классификации, которая относит ту или другую компанию к группе потенциальных банкротов или к группе компаний с благоприятным финансовым положением с определенной мерой точности. В настоящее время существует несколько общепризнанных статистических методов и методик оценки риска банкротства. Наиболее известной и широко применяемой является *методика профессора Альтмана* [Altman, 1983; Altman, 1968]. Вместе с тем, модель Альтмана имеет ряд недостатков и ее применение для экономики Украины сопряжено с определенными трудностями. Поэтому в последние годы разрабатываются альтернативные подходы и методы, учитывающие специфику анализа и принятия решений в условиях неопределенности. К их числу относятся аппарат нечетких множеств и нечеткие нейронные сети. В работах [Згуровский & Зайченко, 2013] были исследованы нечеткие нейронные сети с выводом Мамдани и Цукамото, а также нечетко - множественный „матричный” метод [Недосекин & Максимов, 2003; Недосекин, 2003] в задаче прогнозирования риска банкротства корпораций.

Целью настоящей работы является рассмотрение и сравнительный анализ эффективности применения классических методов и нового класса нечетких нейронных сетей – каскадных нео-фаззи нейросетей методов для прогнозирования риска банкротства предприятий применительно к экономике Украины.

### Модели оценки риска банкротства на основе многомерного дискриминантного анализа

К числу наиболее известных и распространенных моделей оценки риска банкротства относится модель профессора Е. Альтмана [Altman, 1983]. Модель Альтмана построена с использованием аппарата мультипликативного дискриминантного анализа (МДА), который позволяет подобрать такие показатели, дисперсия которых между группами была бы максимальной, а внутри группы минимальной. В данном случае классификация проводилась по двум группам компаний, одни из которых позднее обанкротились, а другие, наоборот, смогли выстоять и упрочить свое финансовое положение.

В результате применения МДА была построена модель Альтмана, имеющая следующий вид [Altman, 1968]:

$$Z=1.2 K_1 +1.4 K_2 +3.3 K_3+0.6 K_4+1.0 K_5 \quad (1)$$

где  $K_1$  = собственный оборотный капитал/ сумма активов;  $K_2$ = нераспределенная прибыль/ сумма активов;  $K_3$ = прибыль до уплаты процентов/ сумма активов;  $K_4$ = рыночная стоимость собственного капитала/ стоимость заемного капитала;  $K_5$ = объем продаж/ сумма активов.

В результате подсчета  $Z$  – показателя для конкретного предприятия делается заключение:

если  $Z < 1,81$  – очень высокая вероятность банкротства;

если  $1,81 \leq Z \leq 2,7$  – высокая вероятность банкротства;

если  $2,7 \leq Z \leq 2,99$  –возможно банкротство;

если  $Z \geq 3,0$  – вероятность банкротства крайне мала.

Модель Альтмана дает достаточно точный прогноз вероятности банкротства с временным интервалом 1 – 2 года.

В последние годы были выполнены исследования по применению модели Альтмана для стран СНГ с переходной экономикой. При этом коэффициенты модели должны быть скорректированы с учетом специфики данного типа экономики. К числу наиболее успешных моделей относится модель Давыдовой-Беликова, разработанная для экономики России [Давыдова & Беликов, 1999].

**Модель Давыдовой – Беликова** имеет вид:

$$R = 8.38K_1 + K_2 + 0.054K_3 + 0.63K_4, \quad (2)$$

где  $K_1$  - отношение оборотного капитала к сумме всех активов;  $K_2$  - отношение чистой прибыли к сумме собственного капитала;  $K_3$  - отношение объема продаж (выручки от реализации) к сумме активов (коэффициент оборачиваемости);  $K_4$  - отношение чистой прибыли к себестоимости.

При  $R < 0$ - вероятность банкротства максимальна (90 - 100%);  $0 < R < 0,18$  - вероятность банкротства высока (60 - 80%);  $0,18 < R < 0,32$  вероятность банкротства средняя (35 - 50%);  $0,32 < R < 0,42$  - вероятность банкротства низкая (15 - 20%);  $R \geq 0,42$  - вероятность банкротства минимальна.

В экономике Украины модель Альтмана пока не получила широкого применения по следующим причинам:

- 1) Требуется вычисление соответствующих коэффициентов при показателях  $K_i$ ,  $i = 1..5$ , которые, естественно, отличаются от их значений для зарубежных стран;
- 2) Информация о финансовом состоянии анализируемых предприятий, как правило, недостоверна, руководство ряда предприятий „сознательно” подправляет свои показатели в финансовых отчетах.

Поэтому задача оценки вероятности риска банкротства должна решаться в условиях неопределенности, неполноты исходной информации, и для ее решения предлагается использовать адекватный аппарат принятия решений – нечеткие нео-фаззи нейронные сети [Бодянский, 2011; Згуровский & Зайченко, 2013]. В докладе рассматриваются каскадные нео-фаззи нейронные сети, приводится их архитектура и рассматриваются алгоритмы их обучения для режимов пакетной обработки и режима „on-line”.

---

### **Экспериментальные исследования алгоритмов прогнозирования риска банкротства**

---

Для анализа разных методов оценки риска банкротства был разработан программный комплекс, в котором реализованы классический метод дискриминантного анализа Альтмана, метод Давыдовой-Беликова и нечеткие каскадные нео-фаззи нейронные сети. Используя разработанный программный комплекс, были проведены прогнозирования банкротства для пятидесяти восьми предприятий Украины, 29 из которых в 2011 году арбитражным судом были признанные банкротами.

Входными данными для расчетов были финансовые коэффициенты, которые исчислялись на основе данных из бухгалтерских отчетов за 2009 и 2010 годы. Прогнозирование проводилось с помощью моделей

Альтмана, Давыдовой-Беликова, нео-фаззи нейронных сетей. Анализ проводился на основе только количественных показателей. В Таблицах 1-2 приводятся результаты прогнозирования - процент ошибочной классификации банкротства предприятий за год до банкротства для статистических методов Альтмана и Давыдовой - Беликова соответственно. В Таблице 3 приведены результаты классификации с помощью каскадных нео-фаззи нейронных сетей.

**Таблица 1.** Результаты прогнозирования методом Альтмана за год до банкротства

	Альтман
Доля ошибок первого типа	0.3
Доля ошибок второго типа	0.344
Относительное количество ошибок	0.327

**Таблица 2.** Результаты прогнозирования моделью Давыдовой-Беликова за год до банкротства

	Модель Давыдовой - Беликова
Ошибка первого типа	0.206
Ошибка второго типа	0.31
Относительное количество ошибок	0.258

Таким образом, выборка состояла из 58 предприятий. По финансовым данным за 1 год до банкротства методом Альтмана 22 предприятия было признано банкротами, 8 предприятий - с большим риском банкротства и 26 предприятий - с удовлетворительным финансовым состоянием. По модели Давыдовой - Беликова 26 предприятий признано банкротами, 5 предприятий - с большим риском банкротства и 29 предприятий - с удовлетворительным финансовым состоянием

Таким образом, методом, который спрогнозировал банкротство с наибольшей точностью за год до банкротства, оказался метод с использованием каскадных нео-фаззи нейронных сетей. Точность прогнозирования составила 90% за год до банкротства и 85% за 2 года до банкротства. Такой результат целиком закономерный, поскольку нечеткие нейронные сети, который базируется на использовании теории систем с нечеткой логикой, лучше прогнозирует при условиях неопределенности, неоднородности данных, а также учитывает субъективные оценки экспертов.

**Таблица 3.** Результаты прогнозирования степени риска банкротства за год до банкротства с использованием каскадной нео-фаззи нейронной сети

	Учебная выборка	Тестовая выборка
Ошибка первого типа	0.128	0.143
Ошибка второго типа	0.0632	0.143
Количество ошибок первого типа	2	2
Количество ошибок второго типа	1	2
Относительное количество ошибок	0.096	0.143

Как, видим, статистическая модель Альтмана показала довольно большую ошибку прогнозирования, которая равна 32.7% (за 1 год) и 38% (за два года). Это ставит под вопрос целесообразность использования этих моделей для анализа финансового состояния украинских предприятий.

В работах [Згуровский & Зайченко, 2013] были исследованы матричный метод Недосекина [Недосекин & Максимов, 2003], а также нечеткие нейронные сети с выводом Мамдани и Цукамото в задаче прогнозирования риска банкротства. Были получены следующие результаты: точность прогнозирования за год до банкротства составила: для матричного метода - 86%, для ННС Мамдани - 90,5%, для ННС

Цукамото - 85%. За два года до банкротства точность прогнозирования составила: для матричного метода - 81%, для ННС Мамдани - 86%, для ННС Цукамото - 82,2%. Как видим, точность прогнозирования риска банкротства для каскадной нео-фаззи нейронной сети приближается к ННС Мамдани, но при этом нео-фаззи каскадные сети имеют преимущества:

- 1) не нужно разрабатывать и настраивать базу правил;
- 2) значительно более высокая скорость обучения.

---

## Заключение

В статье рассмотрены методы прогнозирования риска банкротства предприятий: классические методы дискриминантного анализа Альтмана и Давыдовой–Беликова, а также предлагаемый метод на основе каскадных нео-фаззи нейронных сетей. Проведены экспериментальные исследования указанных методов для прогнозирования риска банкротства предприятий Украины.

Проведенные исследования показали, что наиболее высокую точность прогнозирования риска банкротства применительно к экономике Украины дают нечеткие нейронные сети.

---

## Литература

- [Altman, 1968] Altman E.I., "Discriminant Analysis and the prediction of Corporate Bankruptcy", Financial Ratios, Journal of Finance, September 1968, pp. 589-609.
- [Altman, 1983] Altman E. I., "Corporate Financial Distress", New York, John Wiley, 1983, 192 p.
- [Бодянский, 2011] Е. В. Бодянский, "Каскадная эволюционная нейронная сеть с нео-фаззи нейронами в качестве", [Электронный ресурс], Режим доступа: [http://www.nbuu.gov.ua/portal/natural/vejpt/2011\\_4\\_3/2011\\_4\\_3/55-58.pdf](http://www.nbuu.gov.ua/portal/natural/vejpt/2011_4_3/2011_4_3/55-58.pdf)
- [Давыдова & Беликов, 1999] Давыдова Г.В., Беликов А.Ю., "Методика количественной оценки риска банкротства предприятий", Управление риском, № 3, 1999, с. 13 - 20.
- [Згуровский & Зайченко, 2013] Згуровский М. З., Зайченко Ю.П., "Основы вычислительного интеллекта", К.: Наукова Думка, 2013, 406 с.
- [Недосекин & Максимов, 2003] Недосекин А.О., Максимов О.Б., "Анализ риска банкротства предприятия. Метод. Указания по курсу „Антикризисное управление“", 2003 [Электронный ресурс], Недосекин А.О. Максимов О.Б., Павлов Г.С., Режим доступа до журн: [http://sedok.narod. Ru/sc\\_group.htm](http://sedok.narod. Ru/sc_group.htm).
- [Недосекин, 2003] Недосекин А.О. Система оптимизации фондового портфеля от Сименс Бизнес Сервисез, Банковские технологии, 2003, № 5, Также на сайте: <http://www. finansy. ru/publ/fin/004. htm>

---

## Информация об авторах

**Зайченко Юрий** - д.т.н., профессор ННК „Институт прикладного системного анализа“, 03056, Киев-56, проспект Победы, 37, Украина phone: 38044 -4068393; e-mail: [baskervil@voliacable.com](mailto:baskervil@voliacable.com),

**Области научных исследований:** теория принятия решений в условиях неопределенности, модели и методы вычислительного интеллекта в задачах прогнозирования и анализа в экономике и финансовой сфере, моделирование и оптимизация компьютерных сетей

**Ови Нафас Агаи Аг Гамиш** (Иран) - аспирант НТУУ „КПИ“; 03056, Киев-56, Украина e-mail: [ovinafas@yahoo.com](mailto:ovinafas@yahoo.com)

---

## АНАЛИЗ И АВТОМАТИЧЕСКОЕ ВЫДЕЛЕНИЕ МИМИЧЕСКИХ ПРОЯВЛЕНИЙ ЖЕСТОВОГО ЯЗЫКА В ВИДЕОПОТОКЕ

**Юрий Крак, Антон Тернов, Владислав Кузнецов**

**Аннотация:** Предложены средства автоматического выделения мимических компонент жестового языка в видеопотоке, которые показали результаты близкие к ручному транскрибированию. Для получения количественных характеристик мимических компонент видеопоток был обработан с использованием алгоритма Optical Flow из библиотеки алгоритмов компьютерного зрения. На основе координат отдельных датчиков были вычислены характеристические признаки для определения временных характеристик мимических компонент в видеопотоке.

**Ключевые слова:** транскрипция, видеоданные, эмоции, жестовый язык, компьютерное зрение.

---

### Введение

Предлагаемые исследования посвящены созданию средств анализа мимических компонент жестового языка в потоке видео. Жестовый язык – сложная взаимосвязанная структура, характеризующаяся многопоточностью, своей грамматикой и своими средствами передачи информации – жестами, мимикой и другими элементами (немануальные компоненты) [Krak, 2012]. Как правило, жестовый язык анализируется с привлечением экспертов (сурдопереводчиков) жестовой речи. Компоненты языка выделяются вручную с помощью систем нотации. Такой подход требует значительных трудозатрат и при большом количестве похожих друг на друга элементов может давать большие погрешности в идентификации их составляющих. В данной работе предложен подход к созданию системы автоматического выделения мимических компонент в видеопотоке, который включает:

- Видеофрагменты мимики и метаданные, которые описывают мимические проявления;
- Базу данных, которая реализует модель данных (элементы системы и взаимосвязи между ними);
- Средства выделения количественных характеристик мимических проявлений;
- Инструменты визуализации видеофрагментов и интервалов активности мимических компонент.

---

### Подготовка данных

На основе данных, полученных в исследовании [Krak, 2012] была предложена инфологическая модель, отображающая сведения о мимике и мимических проявлениях в виде структуры и связей между компонентами. Данная инфологическая модель также определяла области на лице, предназначенные для анализа после записи видеофрагментов. На основе инфологической модели был составлен сценарий записи мимических компонент (объектом исследования взят украинский жестовый язык). Его суть заключалась в том, что перед записью на лицо актера в особых областях наносились специальные датчики, которые позволяли получить численные характеристики мимических проявлений. Полученные видеозаписи, каждая из которых описывает одно конкретное мимическое выражение, сегментировались на отдельные фрагменты-тейки(takes). С помощью программного обеспечения для лингвистического анализа Elan [ELAN, 2014], каждому видеофрагменту ставились в соответствие метаданные, описывающие такие элементы: начало, конец, степень проявлений. Затем, с помощью алгоритма Optical Flow из библиотеки OpenCV [ITSEEZ, 2014]) анализировались кадры последовательности и в текстовый файл записывались изменения координат датчиков на лице актера. С целью обеспечения корректной работы алгоритмов на следующем этапе траектории движения датчиков сглаживались вейвлет-

фильтрами. Это позволило устранить влияние возможных потерь позиции датчиков на некоторых кадрах или погрешности вычисления центра маркера, не исказив при этом общий вид траектории.

---

### Получение обучающей выборки

---

На основе координат отдельных датчиков вычислялись характеристические признаки для выделения односторонних и двухсторонних мимических проявлений. Также была введена корректировка, которая позволяла анализировать мимические проявления с аффинными искажениями.

Полученные параметры верифицировались с помощью статистических величин – среднее, стандартное отклонение и др., что позволило установить информативность каждого признака. С помощью разработанного алгоритма на основе усредненной оценки положения начала и конца изменений для каждого из признаков, вычислялись координаты начала и конца изменения отдельного мимического компонента. Эти же параметры использовались для вычисления относительной величины изменения мимических компонент и представления их в виде кода и степени изменения (в относительных единицах) описания мимических проявлений FACS [Miller, 2011].

На основе предложенных алгоритмов было создано приложение - библиотека видеофрагментов. В библиотеку была включена база данных, являющаяся реализацией инфологической модели. Библиотека позволяет вывести на экран описанные выше параметры, графики их изменения, интервалы активности мимических компонент, траектории движения датчиков и сами видеофрагменты.

---

### Выводы

---

Разработаны компьютерные средства для автоматического анализа мимических проявлений в потоке видео, которые показали результаты близкие с результатами ручного транскрибирования. С помощью обработки данных от датчиков и предложенных характеристических признаков удалось осуществить разделение близких друг к другу мимических выражений, как по критерию наличия проявления, так и по степени данного проявления. Это позволяет использовать результат автоматического транскрибирования для идентификации основных мимических выражений. В дальнейшем, для более качественного изучения информации, полученной от датчиков, будут проведены эксперименты с большой группой экспертов, как носителей жестового языка, так и обычных пользователей.

---

### Литература

---

- [Krak, 2012] Krak Ju.V. Notation systems for describing facial expressions in sign languages. / Ju.V. Krak, A.S. Ternov, V.A. Kuznyetsov // Artificial intelligence. – 2012. – № 4. – P.172-182.
- [ELAN, 2014] EUDICO Linguistic Annotator. Hyperlynk: <https://tla.mpi.nl/tools/tla-tools/elan/>
- [ITSEEZ, 2014] Online resource on computer vision algorithms library OpenCV. Hyperlynk: <http://www.opencv.org/>
- [Miller, 2011] Miller F. P. Facial Action Coding System. / Miller F. P., Vandome A. F., McBrewster J. – Beau Bassin, International Book Marketing Service Ltd, 2011. – 80 p.

---

### Информация об авторах

---

**Юрий Крак** (Iurii Krak) – The Institute of Cybernetics of National Academy of Science of the Ukraine, the senior scientist, address: 40 Glushkov ave., Kiev, Ukraine, 03680; e-mail: [krak@unicyb.kiev.ua](mailto:krak@unicyb.kiev.ua)

**Антон Тернов** (Anton Ternov) – The Institute of Cybernetics of National Academy of Science of the Ukraine, the scientist, address: 40 Glushkov ave., Kiev, Ukraine, 03680

**Владислав Кузнецов** (Vladyslav Kuznietsov) – The Institute of Cybernetics of National Academy of Science of the Ukraine, postgraduate student, address: 40 Glushkov ave., Kiev, Ukraine, 03680

## МЕТОД ЛИНЕЙНОЙ ПОЛОСНОЙ КЛАССИФИКАЦИИ СЛОЖНОРАЗДЕЛИМЫХ ДАКТИЛЕМ УКРАИНСКОГО ЖЕСТОВОГО ЯЗЫКА

Юрий Крак, Григорий Кудин, Инга Соломянюк

**Аннотация:** Предложен метод классификации дактилем, которые являются подобными при формировании и демонстрации. Для формирования характеристических признаков используются геометро-топологические параметры кисти руки и пальцев человека, демонстрирующего дактилемы. Данные параметры объединены в шесть блоков, выполнена попарная классификация дактилем в рамках каждого блока. Показано, что наиболее эффективными для классификации являются дескрипторы контурного анализа и геометрические характеристики дактилем.

**Ключевые слова:** жестовый язык, классификация дактилем.

### Введение

Глухие люди используют дактильный язык для передачи информации в тех случаях, когда нет жестового аналога этой информации, в частности для показа цифр, имен, сокращений, географических названий и т.п. Информация передается при помощи движений кисти и пальцев человека, которые строго находятся в области видимости собеседника, поэтому при распознавании дактилем предпочитают анализировать видеоизображение рук человека и проследить изменение их состояния или формы с течением времени [Sing, 2010]. В силу этого, для построения систем классификации дактильной информации, в качестве признаков, важно использовать геометрические характеристики кисти и пальцев руки человека [Kрак, 2011]. В докладе предлагаются шесть различных способов выделения и формирования векторов характеристических признаков. Для дактилем, которые сложно разделить в силу их подобия формирования или демонстрации, предложено исследовать условия классификации по каждому из шести блоков признаков. К сложноразделимым дактилемам украинского жестового языка относятся дактилемы А, В, Е, Є, І, І, І, Л, М, Н, Р, С, Ф, Т, У, Ч, Х, Ю, Я - всего  $n = 19$ . В качестве алгоритма классификации предложено использовать простую в реализации линейную полосную классификацию [Kirichenko, 2009].

### Выделение характеристических признаков дактилем

Для формирования характеристических признаков будем использовать геометро-топологические параметры кисти руки и пальцев человека при демонстрации дактилем. Данные параметры объединены в шесть следующих блоков:

- 1) углы между векторами, проведенными из центра масс ладони к крайним точкам – слева, справа, сверху и снизу области – вектор признаков размерности 4;
- 2) нормированная длина векторов, проведенными из центра масс к крайним точкам – вектор признаков размерности 4;
- 3) компактность, направленность, вытянутость – вектор признаков размерности 3;
- 4) горизонтальное сканирование (разбит на 11 полос) – вектор признаков размерности 11;
- 5) отношение количества черных пикселей к общему количеству (область показа разделена на 25 клеток) – вектор признаков размерности 25;
- 6) дескрипторы контурного анализа – вектор признаков размерности 4.

Для каждой дактилемы в рамках каждого блока было получено 5 наборов. Предварительно для каждого блока осуществлено нормирование векторов признаков на промежуток  $[0, 1]$ .



### Метод полосной классификации

Предполагается, что известно к какому из классов (дактилем) относятся векторы обучающей выборки  $x(j)$ ,  $j = \overline{1, n}$ , в пространстве признаков. Определим условия линейной отделимости конечного множества точек в многомерном пространстве из условий существования решения системы линейных алгебраических уравнений  $x^T(j)a = y_j$ ,  $y_j \geq \Delta$ ,  $j = \overline{1, n}$ , относительно вектора  $a \in R^m$  для некоторых  $\Delta > 0$  и значений  $y_j$ . При фиксированном  $\Delta$  необходимое и достаточное условие существования решения этой системы с использованием операции псевдообращения матриц будет иметь вид:  $\min_{y \in D(\Delta)} y^T Z(X)y = y_*^T(\Delta)Z(X)y_*(\Delta) = 0$ , где  $X = (x(1) : \dots : x(n)) = (x_{(1)} \dots x_{(m)})^T$ ,  $Z(X) = I_n - X^+X$ ,  $D(\Delta) = \{y : y = (y_1, \dots, y_n)^T, y_j \geq \Delta, j = \overline{1, n}\}$ . При этом искомым вектор  $a$  принимает следующее значение:  $a(\Delta) = (X^T)^+ y_*(\Delta)$ , а толщина полосы  $\delta$ , отделяющей множество точек  $x(j)$ ,  $j = \overline{1, n}$ , от начала координат, будет равна величине  $y_* = \delta = \Delta / (y_*^T(\Delta)R(X)y_*(\Delta))^{\frac{1}{2}}$ , где  $R(X) = X^+(X^T)^+$ . Если положить, что  $\Delta = 1$ ,  $y_*(1) = y_*$ , то максимальная толщина полосы определяется из условия  $y_{opt} = \operatorname{argmin}_{y \in D} y^T R(X)y$ ,  $a_{opt} = (X^T)^+ y_{opt}$ , где  $D = \{y : y^T Z(X)y = y^T Z(X)y = 0\}$  для всех  $e_j^T y \geq 1$ ,  $j = \overline{1, n}$ ,  $e_j^T = (0, \dots, 0, 1, 0, \dots, 0)$ .

Найдем условия существования решения и алгоритм получения самого решения для задачи линейной полосной отделимости двух классов точек в пространстве признаков, используя введенные условия линейной отделимости точек в  $R^m$ . Пусть известно, что для последовательности точек  $x(j)$  в пространстве признаков  $R^m$ :  $x(j) \in R^m$ ,  $x(j) = (x_1^T(j) \dots x_m^T(j))^T$ ,  $x_m(j) = 1$ ,  $j = \overline{1, n}$ , точки  $x(i_k)$ ,  $k = \overline{1, n_1}$ , принадлежат первому классу, а точки  $x(j_s)$ ,  $s = \overline{1, n_2}$  - второму классу. Тогда под линейной полосной отделимостью этих классов понимается существование такого вектора  $a \in R^m$ , для которого  $a^T x(i_k) \geq 1$ ,  $k = \overline{1, n_1}$ ,  $a^T x(j_s) \leq -1$ ,  $s = \overline{1, n_2}$ .

Отсюда, условие линейной полосной отделимости принимает следующий вид:  $\min_{y \in D} y^T Z(X)y = 0$ ,  $D = \{y : e_{i_k}^T y \geq 1, e_{j_s}^T y \leq -1, k = \overline{1, n_1}, s = \overline{1, n_2}\}$ , а значение вектора  $a$  оптимально определяется в процессе максимизации толщины разделяющей полосы  $y_{opt} = \operatorname{argmin}_{y \in D_1} y^T R(X)y$ , где

$$D_1 = \{y : y^T Z(X)y = 0\} \cap D, \text{ из условия } a_{opt} = (X^T)^+ y_{opt}.$$

Полученные условия линейной полосной отделимости использованы при построении классификатора сложноразделимых дактилем жестового языка.

### Результаты классификации сложноразделимых дактилем

Для анализа эффективности полученных векторов признаков осуществлена попарная классификация всех 19 дактилем в рамках каждого блока. Для количественного сравнения предложено для каждой пары дактилем определять относительную величину: отношение между шириной полосы и суммой расстояний до соответствующих сторон полосы. Показано, что для хорошей классификации это отношение должно

быть больше 0,3. В таблице 1 приведены результаты попарного разделения некоторых дактилем, используя признаки разных блоков 1-6.

Таблица 1. Результаты попарного разделения дактилем

№ блока признаков	Пары дактилем	Отношение	Радиусы кластеров дактилем	
2,3,6.	Все 19	>0.30		
1.	Б,Я	0.29	0.1	1.0
	Б,Е	0.29	0.1	1.0
	Ю,К	0.26	1.0	1.5
5.	А,В	0.11	1.7	2.7
	А,Г	0.01	1.7	1.8
	В,Ю	0.11	2.7	2.3
4.	А,Б	0.24	0.66	0.9
	Я,Ц	0.06	1.1	1.6
	Я,М	0.03	1.1	1.2

Анализируя полученные результаты, можно отметить, что в рамках разработанных методов классификации и выделения геометро-топологических признаков дактилем, самыми эффективными были признаки, которые объединены в 1,2,3 и 6 блоки. Признаки 5 и 6 блоков оказались менее эффективными.

## Выводы

Предложенные характеристические признаки каждой дактилемы, их группирование в блоки и использование метода полосной классификации, показали эффективность и перспективность такого подхода для распознавания сложноразделимых дактилем. В дальнейших исследованиях предлагается: увеличить обучающую выборку признаков; для блоков с плохой классификацией и с вектором признаков большой размерности провести анализ влияния каждого признака; для слаборазделимых дактилем использовать линейную оптимизационную процедуру с целью улучшения процесса классификации.

## Литература

- [Sing, 2010] Sing Languages. Ed. by D. Brentari. Cambridge: Cambridge University Press, 2010, – 714 p.
- [Kрак, 2010] Крак І.В., Шкільчук Д.В. Analysis of dactyl language elements // Artificial Intelligence. – 2010. – № 3. – P. 322-328. (in Ukrainian).
- [Kirichenko, 2009] Kirichenko N.F., Kudin G.I. Analysis and synthesis of signal classification systems by means of perturbation of pseudoinverse and projection operations // Cybernetics and Systems Analysis. – 2009. – №3. – P. 47-57.

## Информация об авторах

**Юрий Крак** (Iurii Krak) – Taras Shevchenko National University, head of department, V.M.Glushkov Institute of Cybernetics NAS of Ukraine, senior scientist; address: 64/13 Volodymirska str., Kyiv, Ukraine, 01601; e-mail: [krak@unicyb.kiev.ua](mailto:krak@unicyb.kiev.ua)

**Григорий Кудин** (Grigorii Kudin) – Taras Shevchenko National University, dean deputy, address: 64/13 Volodymirska str., Kyiv, Ukraine, 01601

**Инга Соломянюк** (Solomyanyuk Inga) – Taras Shevchenko National University, magistr, address: 64/13 Volodymirska str., Kyiv, Ukraine, 01601

## СООТВЕТСТВИЕ СИСТЕМЫ УПРАВЛЕНИЯ ПОТРЕБНОСТЯМ ОБЪЕКТА УПРАВЛЕНИЯ

Виталий Косс

**Аннотация:** Подобие антропогенных и естественных систем позволяет формулировать их функцию целеполагания в терминах предназначения, устойчивости функционирования и взаимодействия со средой. Функция целеполагания становится критерием для анализа баланса базовых функциональных процессов системы. Результаты анализа служат основанием для обоснованных решений в системе управления. Соответствие системы управления потребностям объекта управления можно измерять показателем соотношения объективной и субъективной информации, циркулирующей в системе.

**Ключевые слова:** Система, управление, объект, информация, истина, ложь.

### 1. Введение

Важнейшим условием системной гармонии любой организационной (общественной) системы является соответствие системы управления потребностям объекта управления. Материал доклада направлен, в первую очередь, на освещение проблем реализации объектом своего предназначения. И, во вторую очередь, на соответствие системы управления решению задач устойчивого функционирования объекта.

### 2. Системные потребности объекта управления

Рассмотрим в качестве объекта управления государство, как систему организации общества, и сравним ее с человеком, как живой системой. Современные социальные науки пропагандируют демократическую организацию сообщества граждан в государстве как наиболее совершенную. Посмотрим на «демократию» в теле человека. Организм плода человека растет за счет ресурсов тела матери. Но неизбежно наступает период отделения плода от пуповины. Для организма ребенка это общесистемный кризис. Если органы тела ребенка способны реализовать свое предназначение во благо жизни единого тела человека, ребенок выживает. Если органы продолжают соперничать за ресурсы - плод умирает и разлагается. Период детства - есть период адаптации новорожденного ребенка к этапу реализации его предназначения. Чем старше становится человек, тем больше ограничений приходится вводить ему на функционирование своего организма (ограничение физических нагрузок, питания, режима труда и отдыха и т.п.). Аналогично в государстве конституция все больше и больше ограничивается кодексами законов. Период активного старения организма характеризуется тем, что процесс восстановления функций начинает отставать от процесса их деградации. Нехватка ресурсов провоцирует внутреннюю борьбу органов. Наступает период «демократизации» (дисгармонизации) единого организма. В государстве этого периода активно формируются олигархические кланы и политические партии, которые борются за доступ к ресурсам через власть, через монополизацию экономики. Государственная машина работает на перераспределение ресурсов народа в пользу олигархов, через механизмы приватизации. Стадия «демократии» неизбежно приводит к тирании. Тело человека активно теряет свою функциональность (подвижность, память и т.д.). Тело народа теряет ресурсы своей страны, деградирует культура, наука, экономика. Далее, тело человека как единый организм умирает и разлагается, а тело государства распадается на более мелкие образования или поглощается соседними государствами. Единый народ дробится по национальному признаку (см. рис.1).

Как вернуть распадающийся организм в стадию его гармоничного функционирования?

Для решения этой задачи системология [1] описывает базовые системные процессы:

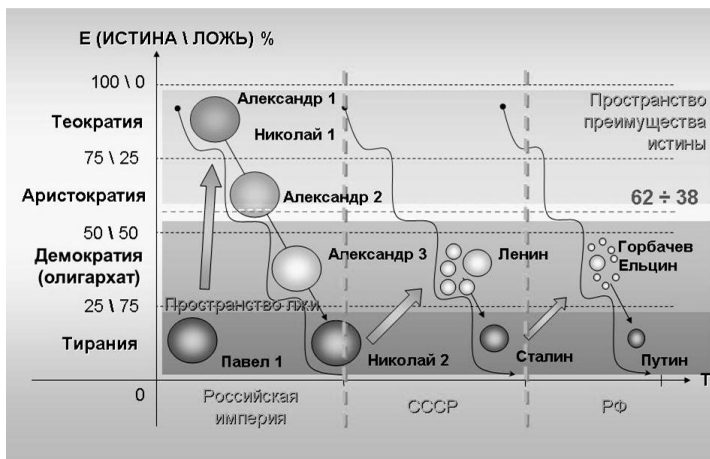


Рис. 1. Пример деградации государственной власти

1. Воплощение (рост) - разрушение (выделение).

2. Реализация предназначения (трансформация) - предотвращение болезни (бунта, преступлений).

3. Восстановление (исцеление) - реорганизация (возрождение).

Базовый процесс *возрождения* по своей сути есть процесс перепроектирования деградировавшей системы в новую систему, способную реализовать свое предназначение. Для человека это перерождение в новом

теле. Для народа это образование новой общественной формы единения. Подняться из состояния тирании или демократии к теократии возможно только через осознание *замысла Творца*.

### 3. Суть управления потребностями объекта управления

Задача системы управления – балансировать базовые системные процессы. Критерием эффективной работы организма любой системы является степень реализации ее предназначения путем регулирования баланса в базовых процессах. Основное влияние на качество управления оказывает показатель соотношения достоверной информации к информации искаженной, утраченной, неактуальной (см. шкалу E на рис.1). Ежесуточный, еженедельный, ежемесячный и ежегодный циклы процедур управления есть процесс восприятия информации, ее трансформации в документы и далее в управляющие стимулы. В настоящем докладе требуется сделать акцент на том, что *главным заданием системы управления есть:*

- *снабжение персонала актуальной и достоверной информацией;*

- *интеллектуальная поддержка процесса трансформации персоналом информации.* Поддержка необходимыми критериями, способствующими гармонизации базовых функциональных процессов.

### 4. Выводы

Процесс проектирования ошибочно воспринимают как процесс «придумывания» создаваемой системы. Современное научное творчество только тогда станет эффективным, когда проектировщики поймут, что процесс функционального проектирования любой системы есть попытка познать *замысел Творца*, а процесс инженерной реализации функций создаваемой системы есть процесс сотворчества с Творцом, где люди пытаются реализовать *замысел Творца* наилучшим образом.

### Библиография

1. Богданов А.А. Очерки организационной науки. <http://www.uic.unn.ru/pustyn/lib/bogdanov.ru.html>

### Информация об авторе

**Виталий Косс** – ИГММС НАН Украины, старший научный сотрудник, пер. Моторный 5/7, кв.25, Киев, Украина, 03083; e-mail: vkoss@ukr.net

---

## АВТОМАТИЧЕСКОЕ ПОСТРОЕНИЕ ТЕРМИНОЛОГИЧЕСКИХ ОНТОЛОГИЙ

Дмитрий Ландэ, Андрей Снарский

**Аннотация:** Представлен подход к автоматическому созданию терминологических онтологий на основе анализа массивов текстов по выбранной проблематике. Подход базируется на применении компактифицированных графов горизонтальной видимости для терминов, а также автоматическом установлении связей между ними

**Ключевые слова:** граф горизонтальной видимости, сеть иерархии терминов, терминологическая онтология, текстовый корпус

---

### Введение

Для решения актуальных задач построения онтологий (детальной формализации выбранных областей знаний) требуется проведение комплексных исследований, определенным этапом которых является построение словарных номенклатур, тезаурусов. Эффективный автоматический отбор отдельных терминов для таких конструкций – не решенная окончательно задача, а проблема автоматического построения сетей из таких терминов до сих пор остается открытой. Как терминологическую основу для формирования соответствующей терминологической онтологии предлагается использовать сеть естественной иерархии терминов (СЕИТ), которая базируется на информационно-значимых элементах текста, опорных словах и словосочетаниях [Ландэ, 2014].

---

### Постановка задачи

Опорные слова и словосочетания в теории информационного поиска выбираются с учетом такого их свойства, как дискриминантная сила. Вместе с тем, одного этого свойства часто оказывается недостаточно для отражения содержания предметной области. Иногда слова с низкой дискриминантной силой, в частности, наиболее частотные слова из выбранной предметной области (например, слова «Web», «Search», «Text» в корпусе текстов по тематике информационного поиска) оказываются важнейшими для рассматриваемой задачи. В данной работе для автоматического выявления терминологической сетевой основы при построении онтологий предметной области предлагается использовать сети естественных иерархий терминов, базирующейся на контенте аннотаций научных статей выбранной направленности. Связи в такой сети определяются естественным взаимным положением слов и словосочетаний, которые экстрагируются из текстов. Такая сеть, создаваемая полностью автоматически, может рассматриваться как основа для дальнейшего автоматизированного формирования терминологической онтологии с участием экспертов.

---

### Методика формирования СЕИТ

Методика формирования сети естественных иерархий терминов предусматривает реализацию последовательности шагов, охватывающей предварительную обработку исходного текста, определение и сортировку терминов, выбор из них необходимого количества наиболее весомых, непосредственное построение СЕИТ и ее отображение. Рассмотрим эти шаги более подробно. В начале формируется исходный текстовый корпус. Как пример такого корпуса рассматриваются аннотации электронных препринтов Arxiv (<http://arxiv.org>) по тематике информационного поиска (рубрика csir, свыше 500 документов за 5 лет). Предварительная обработка такого текстового корпуса предусматривала выделение фрагментов текстов (отдельных аннотаций, абзацев, предложений, слов), исключение нетекстовых символов, отсечение флективных окончаний (стемминг). На втором этапе каждому отдельному термину из

текста (слову, биграмме или триграмме) ставится в соответствие оценка их дискриминантная сила, а именно TFIDF, которая в каноническом виде равна произведению частоты соответствующего термина (Term Frequency) в фрагменте текста на двоичный логарифм от величины, обратной к количеству фрагментов текста, в которых этот термин встретился (Inverse Document Frequency) [Salton, 1983]. Следует отметить, что TFIDF – не единственный метод для вычисления весовых значений терминов. Могут использоваться также различные дисперсионные и весовые оценки, широко применяемые в практике информационного поиска. Затем для последовательностей терминов и их весовых значений по TFIDF строятся компактифицированные графы горизонтальной видимости (CHVG) [Luque, 2009], [Ландэ, 2014] и выполняется переопределение весовых значений слов уже по этому алгоритму, что позволяет учитывать в дальнейшем кроме терминов с большой дискриминантной силой также высокочастотные термины, имеющие большое значение для общей тематики. В качестве весовых оценок отдельных слов в дальнейшем используются степени соответствующих им узлов. После этого все термины текста сортируются по убыванию рассчитанных весовых значений CHVG. Дальнейшему анализу не подлежат термины из так называемого стоп-словаря, являющиеся важными для связности текста, но не несущие информационной нагрузки. Используемый авторами стоп-словарь был построен на основе различных стоп-словарей, представленных в доступном виде на различных веб-ресурсах. Экспертным методом определяется необходимый размер СЕИТ, после чего выбирается соответствующее количество униграмм (единичных слов), биграмм и триграмм с наибольшими весовыми значениями по CHVG. Из отобранных терминов строятся сети естественных иерархий терминов, в которых как узлы рассматриваются сами термины, а связи соответствуют непосредственным вхождениям одних терминов в другие. На последнем этапе формирования СЕИТ осуществляется ее визуализация графов. На рис. 1 приведена визуализация фрагментов рассматриваемой СЕИТ в виде радиальных диаграмм, получивших сегодня широкое распространение в области изучения иностранным языкам.

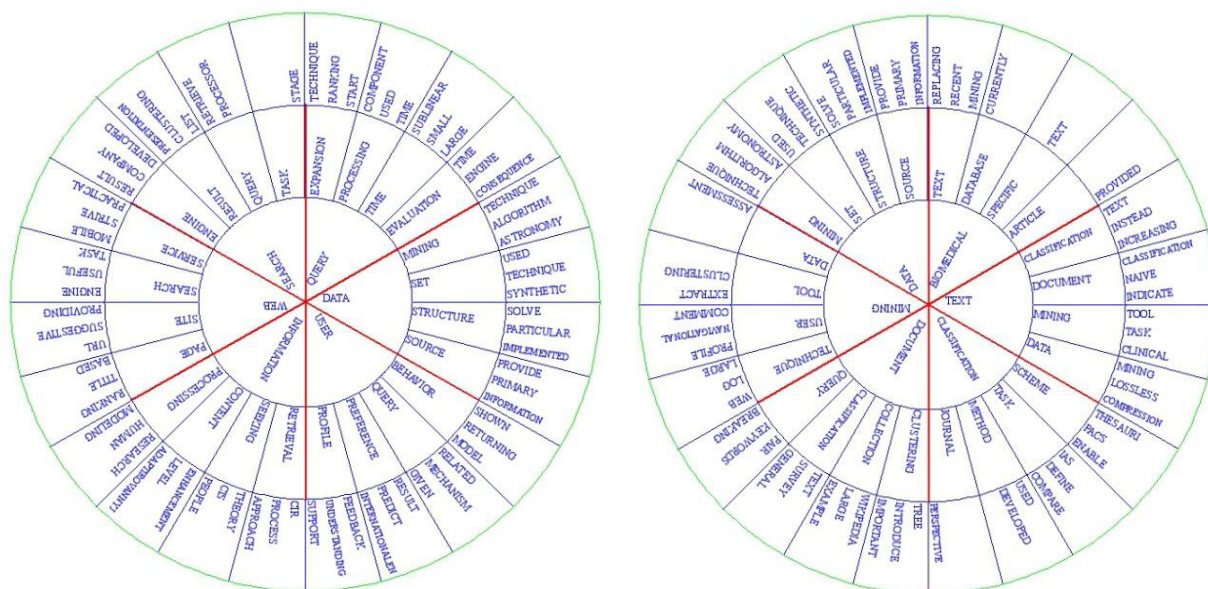


Рис. 1. Визуализация фрагментов СЕИТ в виде радиальных диаграмм

### Ранжирование узлов СЕИТ

Ранжирование узлов в СЕИТ возможно по свойствам, обуславливаемым сетевой структурой, ссылками. Например, для определения авторитетности узла как слова – источника порождения словосочетаний или как составного термина, состоящего из отдельных важных слов, можно анализировать СЕИТ, выбирая

при этом наиболее важных «авторов» или «хабов». Для решения этой задачи предлагается использовать известный алгоритм ранжирования веб-страниц, основанных на связях, HITS (hyperlink induced topic search), предложенный Дж. Клейнбергом [Kleinberg, 1998], который может применяться, например, наряду с алгоритмами PageRank (в этом случае – оценка единственная, интегрированная) или Salsa (этот метод идеально подходит для биграфов, а рассматриваемый нами граф – трехуровневый).

Алгоритм HITS обеспечивает выбор из информационного массива лучших «авторов» (узлов, на которые введут ссылки) и «посредников» (узлов, от которых идут ссылки включения). В рассматриваемом случае термин является хорошим посредником, если от него идут связи на важные словосочетания, и наоборот, термин (словосочетание) является хорошим автором, если на него ведут связи от важных авторов. В соответствии с алгоритмом HITS для каждого узла сети  $v_j$  рекурсивно вычисляется его значимость как автора  $a(v_j)$  и посредника  $h(v_j)$  по формулам:

$$a(v_j) = \sum_i h(v_i); h(v_j) = \sum_i a(v_i). \quad (1)$$

В данных формулах суммирование производится по всем узлам, которые ссылаются (или на которые ссылаются – во второй формуле) на данный узел.

Наиболее интересными с семантической точки зрения в рассматриваемой СЕИТ оказались узлы с наибольшим значением авторства и посредничества.

### Выявление ассоциативных связей

Рассматриваемые в предложенной модели СЕИТ связи являются направленными и могут рассматриваться как отношения «общее-частное» при построении общей онтологии. Вместе с тем, построенная сеть СЕИТ может рассматриваться как основа для формирования других связей между ее узлами. Если обозначить матрицу инцидентности СЕИТ буквой  $A$ , то матрицы  $AA^T$  и  $A^T A$  будут отражать связи вхождения таких типов: если два термина-узла данной сети  $a_i$  и  $a_j$  порождают третий термин  $a_k$ , то будем считать, что такие термины связаны ассоциативной связью, назовем ее ассоциативной связью первого рода (рис. 2 а); если два термина-узла данной сети  $a_i$  и  $a_j$  порождаются третьим термином  $a_k$ , который также входит в данную сеть, то будем считать, что такие термины связаны ассоциативной связью второго рода (рис. 2 б).

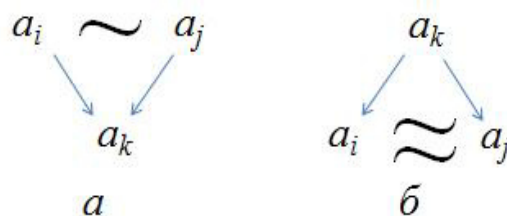


Рис. 2. Ассоциативные связи: а) первого рода «~»; б) второго рода «≈»

На рис. 3 приведен фрагмент сети СЕИТ, дополненной ассоциативными связями.

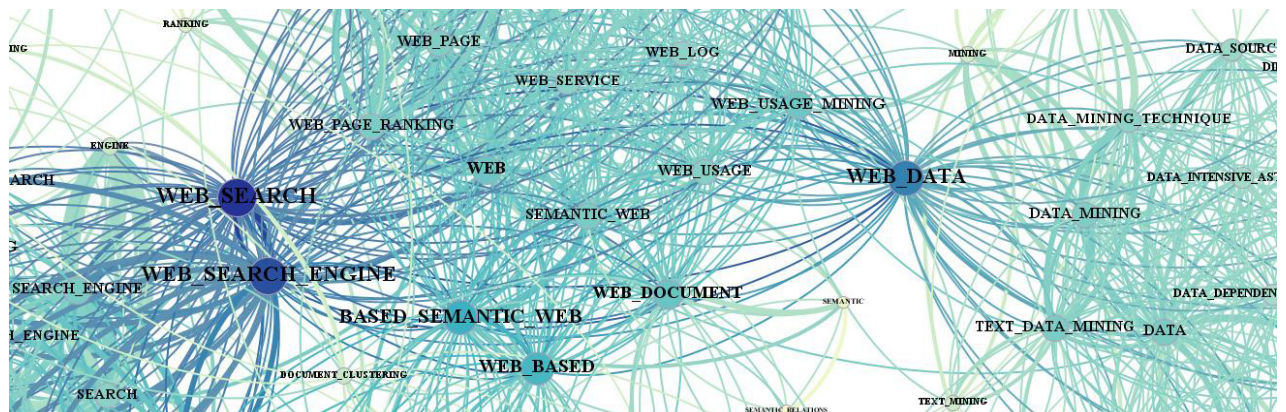


Рис. 3. Фрагмент СЕИТ с ассоциативными связями (визуализация с помощью программы Gephi)

## Выводы

Таким образом, предложено:

- алгоритм построения сетей естественных иерархий терминов на основе анализа текстов;
- метод визуализации фрагментов СЕИТ в виде радиальных диаграмм;
- алгоритм построения ассоциативных связей между терминами в СЕИТ;
- использование алгоритма HITS для выбора наиболее важных элементов в СЕИТ.

Сеть языка, построенную с помощью предложенной методики, можно использовать в качестве базы для построения онтологии предметной области (в рассмотренном примере – по проблематике живучести), использовать на практике в качестве готового к применению средства навигации в информационных массивах, а также для организации контекстных подсказок пользователям информационно-поисковых систем.

## Литература

- [Ландэ, 2014] Ландэ Д.В., Снарский А.А. Подход к созданию терминологических онтологий // Онтология проектирования, 2014. – № 2(12). – С. 83-91.
- [Salton, 1983] Salton G., McGill M.J. Introduction to Modern Information Retrieval. – New York : McGraw-Hill, 1983. – 448 p.
- [Luque, 2009] Luque B., Lacasa L., Ballesteros F., Luque J. Horizontal visibility graphs: Exact results for random time series // Phys. Review E, 2009. – P. 046103-1 – 046103-11.
- [Kleinberg, 1998] Kleinberg J. Authoritative sources in a hyperlinked environment // In Processing of ACM-SIAM Symposium on Discrete Algorithms, 1998, 46(5):604-632.

## Информация об авторах

**Дмитрий Владимирович Ландэ** – д.т.н., зав. отделом специализированных средств моделирования ИПРИ НАН Украины, ул. Шпака, 2, 03113 Киев, Украина, тел.(044) 4542163; e-mail: dwlande@gmail.com

**Андрей Александрович Снарский** – д.ф.-м.н., профессор кафедры общей и теоретической физики Национального технического университета Украины «КПИ», просп. Победы, 37, 03056 Київ-56, Украина; e-mail: asnarskii@gmail.com



## ДЕЯКІ ЗАДАЧІ ОЦІНКИ ПЛАТОСПРОМОЖНОСТІ ПІДПРИЄМСТВ У СФЕРІ КРЕДИТУВАННЯ

Микола Маляр, Володимир Поліщук

**Анотація:** Розглядається проблема побудови моделей і методів оцінки кредитоспроможності підприємств, ранжування інвестиційних проектів і підприємств для підтримки прийняття рішень у банківській сфері.

**Ключові слова:** багатокритеріальний вибір, кредитоспроможність, нечіткі множини.

### Вступ

Після фінансової кризи всі учасники ринку дуже обережно і обґрунтовано підходять до розподілу грошових коштів та перебувають у невизначеній ситуації по відношенню до правильного вибору позичальників щодо надання кредиту. Гостро виникає проблема якісного визначення платоспроможності позичальників, тобто вибору претендентів для інвестування. Сучасні методи і підходи до оцінки платоспроможності позичальників, вживані у вітчизняних банках, зводяться до загального поверхневого аналізу. Крім цього, банки, як правило, вибирають один-два підходи для оцінки платоспроможності клієнтів, на кінцевому етапі спираючись на суб'єктивні фактори ухвалення рішень. При цьому істотним недоліком є відсутність єдиного комплексного і автоматизованого підходу аналізу фінансового стану клієнта. В такому випадку необхідні нові, сучасні теорії та концепції, які враховують теперішні умови ведення бізнесу, реагують на кон'юнктуру ринку та, в тій чи іншій мірі, усувають невизначеність або суб'єктивізм його учасників. Постає актуальна проблема оцінки та вибору підприємств для надання кредиту, чи фінансування деякого інвестиційного проекту та побудови спеціальної інформаційно-аналітичної системи (ІАС) підтримки прийняття рішень.

### Основна частина

Для вирішення даної проблеми постає необхідність побудови моделей і методів, що дозволяють усунути суб'єктивізм експертів, дати можливість неупередженої обробки інформації і на їх основі автоматизувати процеси прийняття рішень у сфері кредитування. Приведемо формальні постановки деяких задач та ідеї методів їх розв'язання.

Задача ( $Z_1$ ) оцінювання і встановлення кредитного рейтингу підприємств. Нехай нам задано множину критеріїв  $K = (K_1, K_2, \dots, K_m)$ , за якими потрібно оцінити деякий суб'єкт (підприємство). На основі даних оцінок потрібно визначити для розглядуваного суб'єкта оцінку рівня кредитоспроможності  $D = (D_1, D_2, \dots, D_l)$ . Аналіз кредитоспроможності підприємств з використання апарату нечіткої логіки і лінгвістичних змінних [Malyar, 2012], у якому аналізується об'єкт із  $n$  входами та одним виходом –  $W = L(U_1, U_2, \dots, U_m)$ , де  $W$  – вихідна лінгвістична змінна,  $U_1, U_2, \dots, U_m$  – вхідні лінгвістичні оцінки, що будуються на відповідних критеріях  $K_1, K_2, \dots, K_m$ ,  $L$  – оператор, що ставить їм у відповідність вихідну змінну. Рішення щодо поточного рівня фінансового стану підприємства обирається таке, для якого функція належності вихідної змінної буде найбільшою по заданим значенням показників діяльності підприємства. Інший метод встановлює кредитний рейтинг підприємств на основі функцій належності критеріїв. Нехай для кожного показника  $K_i$  визначено функцію належності  $\mu(K_i) \in [0;1]$ ,  $i = \overline{1, m}$ . Далі, кожен такий показник буде описуватись множиною відповідних їм числових оцінок. Використовуючи згортку значень функцій належності критеріїв і їх вагових коефіцієнтів, отримуємо оцінку кредитоспроможності.

Обидва методи дозволяють визначити фінансовий стан і категорію якості підприємств у ринкових умовах функціонування, коли не можливо уникнути невизначеності, використовуючи апарат нечітких множин.

За умов, коли грошові ресурси обмежені, а попит на них зростає, постає задача ( $Z_2$ ) вибору підприємства для надання кредиту за багатьма критеріями та задача ( $Z_3$ ) відсіву неперспективних інвестиційних проектів. Ці задачі висвітлюють нові ідейні підходи для використання українськими банками. Для задачі  $Z_2$  запропоновано ряд різних методів розв'язання [Malyar, 2012b], які будуть відрізнятися представленням вхідної інформації та процедурами прийняття рішень. Формально ці всі методи представлені у вигляді задач багатокритеріального вибору альтернатив за їх оцінками за всіма критеріями. Необхідно, на основі таких оцінок, побудувати ранжувальний ряд, згідно якого визначається множина найкращих підприємств.

Формально представимо багатокритеріальну задачу  $Z_3$  оцінки та вибору інвестиційних проектів [Malyar, 2013] наступним чином. Нехай нам задано множину інвестиційних проектів  $X = \{x_1, x_2, \dots, x_n\}$ , які потрібно проранжувати. Множину проектів оцінюють експерти з різних галузей  $E^1, E^2, \dots, E^{k-1}$ , а також враховуються міркування інвестора, які позначимо –  $E^k$ . Кожен експерт та інвестор  $E^l$ , ( $l = \overline{1, k}$ ) використовує свою множину критеріїв  $I_1^l, I_2^l, \dots, I_m^l$  для оцінки інвестиційних проектів, на основі їх оцінюється та будується ранжувальний ряд інвестиційних проектів.

---

## Висновок

Запропоновані моделі і методи для задач оцінки та вибору підприємств для надання кредиту або фінансування інвестиційного проекту – є основою для створення ІАС, яка має актуальне значення для фінансових установ, що займаються активними операціями. Вирішенням проблеми автоматизування процесів прийняття рішень у сфері кредитування є створення підсистеми підтримки прийняття рішень, яка дозволяє розкривати невизначеність аналітиків у своїх судженнях і підвищувати їх об'єктивність в оцінюванні.

---

## Література

- [Malyar, 2012] Мальяр М.М. Модель оцінки кредитоспроможності підприємства в умовах невизначеності/ Мальяр М.М., Поліщук В.В.// Східно-Європейський журнал передових технологій. Сер. Математика і кібернетика – фундаментальні і прикладні аспекти. – Харків, 2012. - №1/4(55). – С.51-57.
- [Malyar, 2012b] M. Malyar, V.V. Polischuk Multicriterion choice problem for enterprises to crediting// ITHEA International Journal "Information Theories and Applications", Vol.19, Number 3, 2012. – P.241-248.
- [Malyar, 2013] M. Malyar, V. Polishchuk, Choice and evaluation methodics of investment projects// Košická bezpečnostná revue, Košice, 2013. – 1/2013/ - P.117-126. ISSN: 1338-4880

---

## Інформація про авторів

**Мальяр Микола Миколайович** – кандидат технічних наук, доцент, докторант Київського національного університету ім. Т.Шевченка, Україна, Ужгород, вул. Підгірна, 46, e-mail: malyarmm@gmail.com

**Поліщук Володимир Володимирович** – асистент кафедри інформаційних управляючих систем і технологій Ужгородського національного університету, Україна, Ужгород, вул. Корзо, 15/9, e-mail: v.polishchuk87@gmail.com

## ОТ ФИЗИЧЕСКИХ ОТРАЖЕНИЙ МАТЕРИАЛЬНОГО МИРА К ОСНОВАНИЯМ ИНФОРМАТИКИ И МАТЕМАТИКИ

Анатолий Мержвинский

**Аннотация.** Предложено использовать понятия статического и динамического икселей для представлений мгновенных состояний объектов-прототипов и объектов-отражений и рассматривать их в качестве элементной базы машинной математики, а идеальные объекты, формируемые сознанием человека, - в качестве элементной базы ментальной математики.

**Ключевые слова:** отражение, коммуникат, иксел, информатика и машинная математика.

### Введение

Преимущественная привязка базовых понятий теории информатики и кибернетики к деятельности субъекта, к языкам верхнего уровня, удаленность от физических взаимодействий материальных объектов (МО) осложняет идентификацию сути информационных машин (ИМ) и кибернетических систем, процесс интеграции их теории.

### 1. Физические принципы формирования отражений

**1.1 Свойства объекта.** Мысленный образ мира. Органы чувств человека, дополненные различными инструментами, позволяют представить реальный мир как пространство взаимодействующих объектов. Важнейшие свойства одиночных материальных объектов природы:

- *Свойство генерации* – способность создавать поля и волны различной природы (электрические, магнитные и др. поля) и потоки материальных частиц;
- *Свойство рассеяния* – способность объектов рассеивать волны и материальные потоки, в частности, преломлять и отражать их.
- *Свойство отражения* – способность отражать волны и материальные потоки, в результате чего изменяется направление их распространения;
- *Свойство поглощения* – способность объектов частично или полностью поглощать волны и материальные потоки

*О терминологии.* В литературе не однозначны понятия *сигнал*, *носитель* и *переносчик*. Понятие *сигнал* иногда трактуется как *носитель* сигнала. *Переносчик* сигнала (в виде колебаний электрического тока, радио- или световой волны) иногда именуется *носителем*. Термин *носитель* может относиться не только к сигналу, но и к элементу машинной памяти или рекламному листку бумаги. Отсутствует общепринятое фундаментальное определение понятия *информация*. *Информационный объект* принято определять как описание в той или иной форме некоторой сущности. Ниже будем исходить из общеизвестного определения "Материальный объект – неоднородность распределения материи" и определений автора: *Информация* - это отражения материальных объектов-прототипов или результатов деятельности человека в ИМ или на носителе в виде совокупности состояний объекта-отображения [Мержвинский, 2012]. *Информационный объект* (ИО) - это **отражение** материального объекта-прототипа или другого ИО в ИМ или на носителе в виде совокупности состояний их элементов (*икселёй*) [Мержвинский, 2013]. Наряду с термином *переносчик* будем использовать термин *коммуникат* [Мержвинский, 2009].

**1.2. Информационные процессы в неживых объектах.** Материальный объект может рассматриваться в ипостаси физического носителя отражения (философская категория), как результата процесса "отражения", означающего в этом случае воспроизведение признаков, свойств и отношений некоторой материальной неоднородности (объекта-прототипа) в совокупность элементов другой материальной неоднородности (объекта-отражения). Примеры физического процесса "отражения" в природе - воспроизведение изображения освещенного объекта на белой поверхности непрозрачной перегородкой с

узким отверстием (рис.1). При расстоянии от свечи до пластины много большего диаметра отверстия число пикселей сформированного изображения определяется отношением размеров сторон экрана к диаметру отверстия. Яркость пикселя определяется яркостью соответствующего элемента отображаемого объекта и может рассматриваться как индуцированное состояние пикселя в отличие от

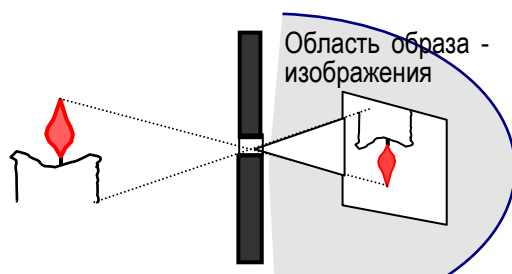


Рис.1. Эмиссия коммуниката и формирование изображения объекта отверстием в непрозрачной перегородке.

случая, когда состояние пикселя удерживается после прекращения воздействия.

Отметим такие стадии и средства формирования физических отображений:

- Эмиссия объектом *коммуниката* в свободное физическое пространство и формирование вокруг объекта-*прототипа* области информационного поля, и, при наличии других объектов изменяющих структуру поля (в данном случае пластины с отверстием), области образа-*изображения*.

- Восприятие некоторого физического параметра явления как отражения и преобразование его формы (*рецепция*), например, фотодиодом, преобразующим оптический поток  $\lambda$  от некоторого объекта в электрический ток  $i$ .

- Формирование в окружающих материальных объектах *образа* (*отражения* объекта-*прототипа*) в виде измененного состояния множества материальных носителей, например, в ячейках памяти.

## 2. Концепция статического и динамического икселей как носителей отражений объектов

**Статические и динамические иксели.** Материальные неоднородности и структуры конкретного объекта-*отражения* на нижнем уровне обычно представлены одномерными элементами: *пикселями*, *вокселями*, *периодическими колебаниями носителя*; в более сложных случаях - многозначными компонентами *hardware*. В литературе, к сожалению, отсутствует обобщающий термин, именующий наипростейший физический *отражающий элемент* независимо от того, что является носителем *отражения*: одно- или многозначный элемент памяти, некоторый аппаратный преобразователь или *коммуникат* [3]. По аналогии с терминами *пиксель* и *сенсель* (от *sensor element* - чувствительный элемент) для именованного конструктивного элемента отображающих неоднородностей введен обобщающий термин **иксел** (*ixel* – производное от *icon-element*) [2]. *Иксел* (рис.2) может быть реализован как:

- *Статический (H-иксел)* - элемент *hardware*, реагирующий на воздействие; например, элемент памяти, эмиттерный повторитель.

- *Динамический (D-иксел)* - элемент переносчика сигналов работающей функциональной цепи ИМ (импульс потока электронов, фотонов).

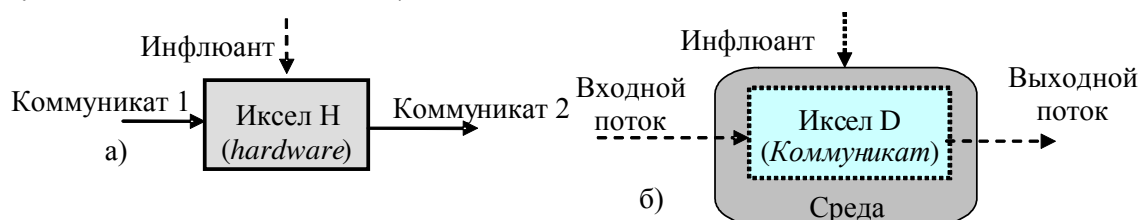


Рис.2. Концепт статического и динамического иксела

## 3. Первичные отражения реального мира – элементная база информатики и математики

Каждый объект материального мира, как объект-*прототип*, может быть подвергнут или не подвергнут влияниям *коммуникатов* - потоков фотонов, электронов и различных частиц из солнца, космоса, технологического оборудования. При этом может происходить процесс и собственной и индуцированной

эмиссии *коммуниката*, как носителя определенных свойств объекта-*прототипа*. Формирование ИО как результата эмиссии *коммуникатов* МО либо в результате ментальной деятельности человека позволяет информатику как науку связать по входам с материальными объектами и человеком, а по выходам - с математикой, техническими, гуманитарными и естественными науками. Характерные транзакции входных и выходных ИО отображены стрелками на рис. 3.

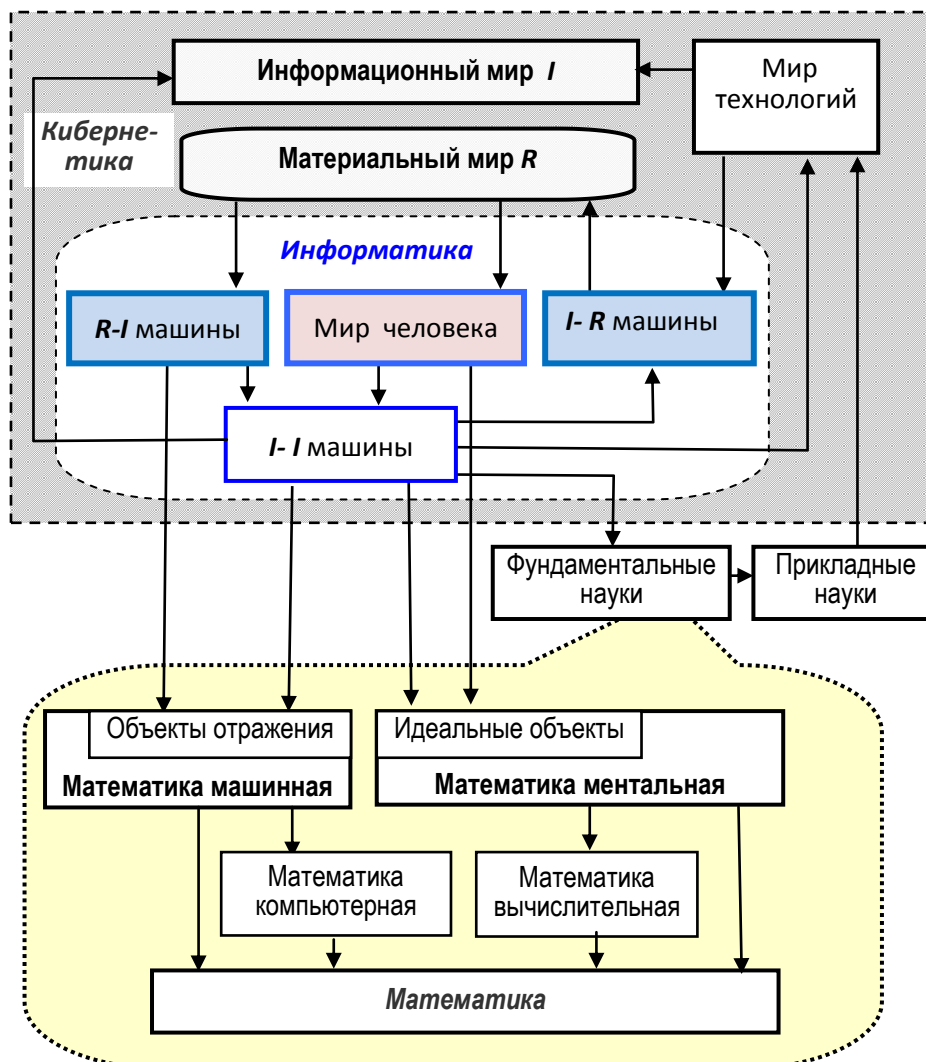


Рис. 3. Информационные объекты в структурах информатики и математики

Исходные ИО порождаются МО и человеком на основании коммуникатов излучаемых МО. Воспринятые сенсорами R-I машины и рецепторами человека отражения приводятся к форме, пригодной для обработки I-I машиной. I-I машины предназначены для обработки информации и решения задач. Выходные ИО поступают на исполнительные I-R машины и технологические линии. Очевиден факт расположения *иксельного* уровня *отражений* статики ПдО между уровнем *иксельного* представления реальных физических объектов, относящегося к экспериментальному миру, и уровнем отражений более высокого уровня (ИО), которые относятся к теории познаваемого мира. Легко видеть иерархию таких уровней отражений: *иксельный* (индивидуума или множества элементов); числовой, цифровой; образный, символичный (имена), атрибутивный; текстовый, предикатный, математический (в последнем случае математика рассматривается как средство отображения мира, а не решения задач). *Иксельное представление* позволяет по-новому подойти к некоторым аспектам математики. Суть в том, что вопрос об основаниях математики и о том, что представляет собой, в конечном счете, математика, остаётся открытым. Известно, что Герман Вейль пессимистически оценил возможность дать общепринятое

определение предмета математики.

Если исходить из определения «Математика (греч. *mathematike*, от *máthema* — знание, наука) - наука о количественных отношениях и пространственных формах действительного мира», то исходные отношения и пространственные формы могут исходить: а) из иксельных отражений воспринимаемых ИМ; б) из отражений продукта человеческого сознания. Это значит, что определения могут быть даны в зависимости от онтологического статуса исходных элементов математики и отдельно рассматриваться:

- *ментальная математика*, построенная на результатах процессов ментальной и практической деятельности человека согласно формальной логике и на основе аксиом об элементах, природа которых уже не определена (идеальные объекты);

- *машинная математика*, создаваемая на основе элементов из объектов-отражений иксельного уровня полученных техническими средствами и зафиксированных на неживых материальных носителях.

Существующие понятия *Вычислительной*, как и *Компьютерной*, *математики* можно связывать не только с инструментами подготовки и методами решений математических задач с помощью компьютеров, но и соотносить с первичными формами машинных отображений реального мира (*иксельными*) и, в общем случае, рассматривать их, как раздел *машинной* математики. Не лучше ли определения математики начинать с определений элементной базы *машинной математики* (с ее царствами икселей, чисел, символов, структур) и далее переходить к исторически сложившейся ментальной математике.

Что касается *информатики*, то очевидно, что в структуре взаимосвязей наук информатика находится между реальными материальными объектами с одной стороны и естественными науками и математикой с другой стороны. Поскольку процесс отражения имеет место при формировании понятий и естественных и гуманитарных наук, то применительно к дереву взаимосвязей наук информатика может претендовать на роль и рассматриваться как связующее звено объектов природы и понятий и языков естественных и гуманитарных наук. В кибернетике задачи информатики (формирования, транспортировки и реализации информации) дополнены задачами управления как информационными так материальными процессами. В ИИ эти задачи дополнены средствами автоматизации мыслительной деятельности и привлечением к решению задач баз знаний.

---

## Заключение

Введение понятий иксел и коммуникат (в отличие от понятия «сигнал») позволило перейти к элементной базе ИМ самого низкого уровня – коммуникатного, связать представление физических объектов, относящихся к экспериментальному миру, с уровнем отражений более высокого уровня (ИО), которые относятся к теории и более точно описать стадии и сущность информационных процессов.

---

## Литература

[Мержвинский, 2012] Мержвинский А.А., Мержвинский П.А. От структуры канала связи к знаковой модели мира Труды Северокавказского филиала московского технического университета связи и информатики, 2012, с.225-229).

[Мержвинский, 2013] Мержвинский А.А. Информационные машины: некоторые категории функций и компонент. International Journal "Information Theories and Applications", Vol.20, Number 1, 2013 p. 75-87

[Мержвинский, 2009] Мержвинский А.А. Модель универсума. Informations Models of Knowledge, XVI-th International Conference Knowledge – Dialogue – Solution, N. 15. – ITHEA, Sofia, 2009. p.31- 39.

---

## Информация об авторе

**Мержвинский Анатолий Александрович** – Институт кибернетики им. В.М.Глушкова НАНУ, 03680 ГСП Киев-187, пр-т Академика Глушкова,40, Украина; e-mail: merjv@mail.ru

Научные интересы: физико-технологические проблемы кибернетики, микро-оптоэлектроника, биосенсорика

## АЛГЕБРИ КВАЗІАРНИХ ВІДНОШЕНЬ ТА ЇХ ВЛАСТИВОСТІ

Нікітченко М.С., Шкільняк С.С., Матвіюк Д.А.

**Анотація:** вводяться поняття квазіарного відношення, операції над такими відношеннями та алгебри квазіарних відношень.

**Ключові слова:** квазіарні відношення, операції над відношеннями, алгебра квазіарних відношень.

Під  $n$ -арним відношенням традиційно розуміють множину кортежів довжини  $n$ . Скінченне  $n$ -арне відношення можна розглядати як певну таблицю, яка має  $n$  стовпчиків. Так визначене поняття відношення належить до найважливіших понять математики. Водночас низка задач інформатики та програмування вимагають певного узагальнення цього поняття. Наприклад, якщо розглядати екзаменаційну відомість як деяку таблицю, то не всі її клітинки можуть бути заповнені під час іспиту. Формально таку частково заповнену таблицю можна задати наступним чином. Нехай  $V$  – множина атрибутів,  $A$  – множина значень. Часткову функцію із  $V$  в  $A$  назвемо номінативною множиною. Клас всіх таких множин позначаємо  $\forall A$ . Квазіарним відношенням назвемо довільну підмножину  $R \subseteq \forall A$ . У цьому випадку номінативна множина, що входить у відношення  $R$ , може розглядати як частково заповнений рядок таблиці.

На множині всіх квазіарних відношень природним чином задаються булеві операції об'єднання, перетину, доповнення, а також спеціальні номінативні операції реномінації та квантифікації [Нікітченко, 2008].

Метою доповіді є визначення різних алгебр квазіарних відношень та дослідження їх зв'язків із алгебрами предикатів. Зокрема, доводиться ізоморфізм алгебри квазіарних відношень та першопорядкових алгебр тотальних предикатів.

Також будуються алгебри бі-квазіарних відношень, визначені на множинах пар квазіарних відношень. Визначаються різні підкласи таких алгебр та досліджуються їх зв'язки з недетермінованими, дуальними, насиченими та реляційними алгебрами часткових предикатів.

Отримані результати узагальнюються на випадок багатосортних відношень [Nikitchenko, 2013].

---

### Література

[Нікітченко, 2008] М.С. Нікітченко, С.С. Шкільняк. Математична логіка та теорія алгоритмів. Київ: ВПЦ Київський університет, 2008. 528 с.

[Nikitchenko, 2013] M.S.Nikitchenko, V. G. Tymofiev. Satisfiability and Validity Problems in Many-Sorted Composition-Nominative Pure Predicate Logics. Communications in Computer and Information Science, vol. 347, 89-110, Springer Berlin Heidelberg, 2013.

---

### Інформація про авторів

**Нікітченко Микола Степанович** – завідувач кафедри теорії та технології програмування, Київський національний університет імені Тараса Шевченка, 01601 МСП, Київ, вул. Володимирська, 64; e-mail: [nikitchenko@unicyb.kiev.ua](mailto:nikitchenko@unicyb.kiev.ua)

**Шкільняк Степан Степанович** – професор кафедри теорії та технології програмування, Київський національний університет імені Тараса Шевченка, 01601 МСП, Київ, вул. Володимирська, 64; e-mail: [sssh@unicyb.kiev.ua](mailto:sssh@unicyb.kiev.ua)

**Матвіюк Дмитро Анатолійович** – аспірант Східноєвропейського національного університету імені Лесі Українки, 43025, Луцьк, просп. Волі, 13; e-mail: [dimmat25@gmail.com](mailto:dimmat25@gmail.com)

## МОДЕЛИРОВАНИЕ РЫНОЧНЫХ МЕХАНИЗМОВ В АГЕНТНО-ОРИЕНТИРОВАННОЙ МОДЕЛИ ФУНКЦИОНИРОВАНИЯ ЭКОНОМИКИ

Диана Омелянчик

**Аннотация:** Предложена агентно-ориентированная модель функционирования экономики. Представлена реализация механизмов взаимодействия агентов на рынках труда и товаров.

**Ключевые слова:** агентно-ориентированное моделирование математическая экономика.

### Введение

Необходимость выявления предпосылок глобальных кризисов требует разработки качественно новых подходов к анализу рыночных механизмов и моделированию взаимодействий между экономическими агентами. Одним из современных методов является агентно-ориентированное моделирование. Основная идея этого подхода заключается в представлении об экономике как о сложной адаптивной системе, поведение которой формируется в результате множественных взаимодействий автономных гетерогенных экономических агентов, обладающих разным поведением и способностью к обучению.

### Основные рынки в агентно-ориентированной модели функционирования экономики

Разработанная модель содержит два типа активных агентов: фирмы-производители и домохозяйства [Гуляницкий, 2014]. Взаимодействие между ними происходит не прямо, а с помощью специальных сущностей-посредников – рынка товаров и рынка труда. На каждой итерации моделирования последовательно выполняются такие структурные блоки действий: взаимодействие фирм и домохозяйств на рынке труда, производство продукции, ценообразование, определение величины потребления, взаимодействие фирм и домохозяйств на рынке товаров, обучение фирм. Каждый структурный блок может быть реализован различными способами в зависимости от целей исследования и независимо от реализации других блоков.

Остановимся подробнее на реализации механизмов взаимодействия между агентами на двух рынках модели – рынке труда и рынке товаров.

*Рынок труда.* Рынок труда характеризует спрос на труд со стороны фирм. Запрос фирмы на заполнение вакансии содержит количество работников, которых требуется нанять, и предложенную заработную плату. Поиском работы на рынке труда занимаются безработные домохозяйства и недовольные своей зарплатой работники фирм. Взаимодействие фирм и домохозяйств на рынке труда реализовано с помощью следующего алгоритма подбора, моделирующего следующие этапы:

- Фирмы выставляют вакансии на рынок труда;
- Незанятые домохозяйства и домохозяйства, неудовлетворенные своей текущей заработной платой, рассматривают открытые вакансии и подают заявки на их заполнение;
- Фирмы рассматривают полученные заявки и рассылают приглашения избранным. Количество таких приглашений не превышает количества недостающих работников фирмы;
- Домохозяйства рассматривают полученные от фирм предложения и выбирают самый выгодный для них вариант;
- Фирмы обновляют информацию о своих работниках на основе решения домохозяйств.

Исследования показывают, что целесообразно выполнять цикл подбора ровно два раза и прерывать его, даже если фирмам не удалось полностью заполнить открытые вакансии.



*Рынок товаров.* Предполагается, что фирмы производят только один потребительский товар, который является гомогенным, т.е. товары разных фирм-производителей отличаются исключительно ценой. Произведенный товар отправляется на централизованный рынок товаров, где его могут приобрести домохозяйства. Предложение товара состоит из цены на единицу товару и соответствующего объема товара, выставленного на продажу.

В рамках базовой модели выбор товара домохозяйством считается случайным, причем вероятность приобретения того или иного товара обратно пропорциональна его цене. Тогда, если обозначить через  $v(p) = -\ln p$  ценность товара для домохозяйства, вероятность выбора  $prob_j$  товара фирмы  $j$  определяется формулой [Dawid, 2009]

$$prob_j = \frac{\text{Exp}[\lambda v(p_j)]}{\sum_j \text{Exp}[\lambda v(p_j)]},$$

где  $\lambda$  – интенсивность конкуренции на рынке товаров.

Выбрав таким образом товар, домохозяйство пытается полностью израсходовать бюджет потребления на его приобретение. Если для этого на рынке товара недостаточно, т.е. после покупки у домохозяйства остаются деньги, отведенные на потребление, то оно выбирает следующий товар и пытается потратить остаток бюджета на него. Так происходит до тех пор, пока бюджет не будет использован полностью или на рынке не закончатся товары.

Отметим, что важным для спецификации модели является предположение о бесконечной делимости товара или его отсутствие. В последнем случае домохозяйство может оказаться в ситуации, когда оно не сможет использовать ненулевой бюджет потребления на покупку товаров, поскольку установленные фирмами цены будут слишком высокими. Это может существенно повлиять на величину прибыли фирм и, соответственно, принимаемые ими решения.

---

## Заключение

Предложена агентно-ориентированная модель функционирования экономики. На первом этапе создания этой модели ставилась цель воссоздать рыночные механизмы взаимодействия между основными экономическими агентами, фирмами и домохозяйствами, на рынке товаров и труда. В дальнейшем планируется расширить возможности модели путем введения новых рынков (внешний рынок, финансовый рынок), новых типов агентов (банк, правительство), новых механизмов взаимодействия (уплата налогов, кредитование, открытие депозитов) и новых подсистем национальной экономики (регион, отрасль).

---

## Литература

[Dawid, 2009] Dawid H., Gemkow S., Harting P., Neugart M. On the effects of skill upgrading in the presence of spatial labor market frictions: An agent-based analysis of spatial policy design // J. of Artificial Societies and Social Simulation. – 2009. – N 12 (4).

[Гуляницкий, 2014] Гуляницкий Л.Ф., Омелянчик Д.А. Разработка и исследование базовой агентно-ориентированной модели функционирования экономики // Компьютерная математика. – № 1. – 2014. – С. 26-36.

---

## Информация об авторе

**Диана Омелянчик** – аспирант Института кибернетики имени В.М.Глушкова НАН Украины;  
e-mail: omelyanchikd@gmail.com

---

## АНАЛИЗ СУЩЕСТВУЮЩИХ КЛАССИФИКАЦИЙ КОМПЕТЕНЦИЙ

Юлия Панасовская, Максим Вороной

Научный руководитель – д.т.н., проф. Соловьева Е.А.

**Аннотация:** В докладе представлены разработки в предметной области «Компетенции», анализ существующих классификаций компетенций, оценка целесообразности применения метода системологического классификационного анализа, который позволит систематизировать знания в области управления человеческими ресурсами.

**Ключевые слова:** Компетенции, профессиональные компетенции, системологический классификационный анализ, системология, классификация.

---

### Введение

В современных условиях на рынке остаются только те компании, которые своевременно реагируют на изменение требований и внедряют новое, что способно серьезно улучшить результат работы. Но для того, чтобы нововведения работали, а фирма совершенствовалась и повышала эффективность работы – естественно, необходимы высококвалифицированные специалисты, имеющие талант, желание работать и высокий уровень компетентности в своем деле. Поэтому анализ компетенций сотрудников является актуальным.

---

### Анализ существующих классификаций компетенций

Управление человеческими ресурсами состоит в действиях по достижению организационных целей за счет обеспечения требуемого организацией производственного поведения каждого из ее сотрудников или реализации организационных компетенций, непосредственно связанных с персоналом [1]. Для описания компетенций и их свойств целесообразно использовать классификации компетенций.

Анализ источников показал, что «компетенции» – это сложное, многокомпонентное, междисциплинарное понятие, определяемое с помощью множества различных дефиниций. Описания содержания этого понятия имеют нестрогий характер и отличаются по объему, составу, семантической и логической структуре, и не все определения являются родовидовыми.

Предметная область «Компетенции» является слабо структурированной и не имеющей общепринятой классификации. Вместе с тем возможны классификации по разным основаниям, например: по уровню распространения компетенций, по уровню развития, по сущности и содержанию, по способу появления и др. Был проведен анализ классификаций на соответствие правилам формальной логики, критериям естественной классификации, который выявил, что каждая проверенная классификация построена с нарушениями этих правил.

Для примера приведен анализ существующей классификации профессиональных компетенций. В данной классификации профессиональные классификации были разделены на: коммуникативные компетенции, нравственно-социальные компетенции, организаторские компетенции, креативные компетенции, гностические компетенции, проектные компетенции, исследовательские компетенции, рефлексивные компетенции [2]. Данные понятия находятся на одном уровне и определяются через род «профессиональные компетенции». Но понятия «коммуникативные компетенции» и «нравственно-социальные компетенции» необходимо определять через род «социальные компетенции», а понятие «рефлексивные компетенции» необходимо определять через род «личностные компетенции». Понятия «организаторские компетенции», «креативные компетенции», «гностические компетенции», «проектные

---

компетенции», «исследовательские компетенции» необходимо определить через род «профессиональные компетенции».

Проанализировав данную классификацию с точки зрения формальной логики, можно сделать вывод, что классификация не соответствует правилам формальной логики, так как нарушается правило определения понятия через ближайший род.

Проведя анализ существующих классификаций компетенций, можно сделать вывод, что для данной предметной области не найдено классификации, отвечающей всем правилам формальной логики, и компетенции не проклассифицированы по функциональному назначению. Поэтому целесообразно разработать классификацию компетенций с использованием метода системологического классификационного анализа, соответствующую ранее упоминавшимся правилам и новым системным критериям [3].

Использование метода системологического классификационного анализа позволяет построить и оценить любую классификацию с точки зрения ее обоснованности (параметричности), отражения в ней существенных свойств объектов, возможности обнаружения и предсказания объектов по их месту в классификации, то есть с точки зрения возможности применения классификации в качестве инструмента теоретического анализа в соответствующей области [3]. Применение системологического классификационного анализа позволит разработать параметрическую классификацию предметной области, которую возможно расширить и углубить без ее изменения. Такая классификация будет соответствовать требованиям формальной логики и критериям естественной классификации.

---

## Выводы

В результате проведенной работы были получены следующие результаты:

- проведен анализ предметной области «Компетенции»;
- проведен анализ существующих классификаций компетенций и выявлены их недостатки.

Дальнейшие исследования в рассматриваемой предметной области позволят разработать параметрическую классификацию предметной области «Компетенции», которая будет максимально приближенной к естественной классификации.

---

## Список использованной литературы

1. Компетенции персонала в современной организации [Электронный ресурс] – Режим доступа: [www/URL: http://www.e-xecutive.ru/knowledge/announcement/694102/](http://www.e-xecutive.ru/knowledge/announcement/694102/) – 10.04.14 – Загл. с экрана.\
2. Основные профессиональные компетенции будущего специалиста экономического профиля [Электронный ресурс]. – Режим доступа: [www/URL: http://www.ssc.smr.ru/media/journals/izvestia/2010/2010\\_3\\_32\\_34.pdf](http://www.ssc.smr.ru/media/journals/izvestia/2010/2010_3_32_34.pdf) – 01.07.2014. – Загл. с экрана.
3. Соловьева Е.А. Естественная классификация: системологические основания / Е.А. Соловьева. [Текст] – Харьков: ХНУРЭ, 1999. – 222с.

---

## Сведения об авторах

**Юлия Панасовская** – соискатель кафедры Социальной информатики, Харьковский национальный университет радиоэлектроники, ул. Ленина, Харьков 61166, Украина; e-mail: [pannayulia@rambler.ru](mailto:pannayulia@rambler.ru)

**Максим Вороной** – к.т.н., кафедра Социальной информатики, Харьковский национальный университет радиоэлектроники, ул. Ленина, Харьков 61166, Украина; e-mail: [si@kture.kharkov.ua](mailto:si@kture.kharkov.ua).

---

## РАЗРАБОТКА МЕТОДОВ СЕМАНТИЧЕСКОГО ПОИСКА WEB-СЕРВИСОВ НА ОСНОВЕ ОНТОЛОГИЧЕСКОГО АННОТИРОВАНИЯ

Юлия Рогушина

**Аннотация:** предложены подходы к семантическому аннотированию Web-сервисов терминами онтологий предметных областей для их эффективного обнаружения и классификации

**Ключевые слова:** Web-сервис, онтология, семантическая разметка

---

### Семантическая разметка информационных объектов

---

В общем случае разметка текстового информационного ресурса (ИР) – это связывание фрагментов текста с тэгами из какого-либо множества. При семантической разметке между тэгами явным образом задаются семантические связи (например, иерархические). В случае если для задания семантики связей между тэгами используется онтология, то говорят об онтологической разметке.

Семантическая разметка позволяет использовать для распознавания информационных объектов (ИО) – как материальных, так и виртуальных – при анализе контента ИР априорные знания о предметной области (ПрО), в частности, о структуре ИО, сведения о которых содержатся в анализируемых ИР. Это позволяет значительно снизить вычислительную сложность распознавания ИО, т.к. вместо задачи кластеризации решается задача классификации. Задача распознавания состоит из двух подзадач, которые выполняются итеративно: 1) построение гипотезы о множестве классов, к которым могут относиться ИО; 2) отнесение конкретного ИО к одному из этих классов. Наличие разметки обеспечивает решение первой подзадачи. В частности, онтологическая разметка часто используется для поиска и классификации Web-сервисов. В Web семантическая разметка имеет свою специфику.

---

### Этапы распознавания ИО в Web

---

Этап 1. Определение класса искомого ИО (явно, через свойства ИО, через примеры ИО).

Этап 2. Формирование массива ИР, содержащих сведения об ИО.

Этап 3. Формирование ограничений на свойства ИО из выбранного класса.

Этап 4. Извлечение из ИР сведений о свойствах искомого ИО, которое сводится к семантической разметке набора ИР, сформированного на этапе 2, понятиями (свойствами), выделенными на этапе 3.

При этом могут применяться методы лингвистического анализа, распознавания речи и изображений (для анализа мультимедийных ИР), а также средства менеджмента знаний, направленные на автоматизированное использование знаний о структуре искомого ИО, представленных в интероперабельной форме, допускающих автоматизированную обработку.

Для интероперабельного представления знаний соответствующих ПрО целесообразно использовать онтологии, используемые в качестве их формальных моделей. Semantic Web является источником семантической метаинформации о распределенных ИО и ИР, аннотируя их контент и обеспечивая возможность автоматизированного извлечения из них сведений об искомым ИО.

Онтологию можно использовать как основу для представления структуры ИО (классы), а ИР – для создания экземпляров ИО. Это позволяет интегрировать информацию из различных ИР. При этом возникает ряд подзадач: 1) поиск онтологии, отражающей структуру ИО, знания о которых необходимы пользователю; 2) поиск ИР, содержащих сведения об этих ИО; 3) извлечение знаний об ИО из ИР; 4) представление извлеченных знаний в форме, понятной и удобной пользователю [Рогушина, 2014].

Следует отметить, что во многих случаях это процесс является итеративным и при повторном решении проблемы информацию об ИО необходимо обновлять, извлекая необходимые сведения из тех ИР, которые доступны пользователю (например, через Web или корпоративную сеть).

---

### **Поиск Web-сервисов на основе семантической разметки**

---

Рассмотрим в качестве примера распознавание сведений о таком ИО, как Web-сервис (подразумеваются решения на основе протоколов SOAP, UDDI и WSDL). Публикация сервисов в универсальном UDDI не обеспечивает ни семантического обнаружения, ни композирования Web-сервисов. Чтобы преодолеть эти ограничения, используются различные способы семантической разметки этого описания. Различные исследования предлагают для семантического аннотирования Web-сервисов такие средства, как: OWL-S, Web Service Modelling Language (WSML), Web Service Modelling Ontology (WSMO), Web Services Description Language Semantic (WSDL-S) и Semantic Annotations for Web Services Description Language (SAWSDL) [Talantikite, 2009]. OWL-S разработана W3C и базируется на языке представления онтологий OWL. OWL-S обеспечивает декларативные описания свойств Web-сервиса и возможности, которые могут использоваться для автоматического выявления сервиса. Для аннотирования Web-сервиса наиболее интересны понятия профиля OWL-S (входы и выходы) и ServiceGrounding.

Аннотирование Web-сервиса состоит из двух процессов – категоризации и сопоставления [Bouchiha, 2012]. В первом процессе, WSDL-описание сервиса классифицируется в соответствующий домен. Во втором процессе сущности WSDL сопоставляются с предварительно созданной онтологией. Поиск Web-сервисов сводится к сопоставлению онтологии запроса с онтологиями сервисов.

При использовании для именования параметров Web-сервисов многозначных терминов, применяемых в различных ПрО (например, «имя», «оценка», «адрес»), автоматическая классификация Web-сервисов либо крайне неточна, либо вообще невозможна. Поэтому целесообразно, чтобы разработчики Web-сервисов явным образом указывали пространство имен, которые используются для обозначения входных и выходных параметров их сервисов. Например, в качестве источников таких имен могут использоваться схемы RDF-S или OWL. В таком случае 1) Web-сервис однозначно может быть классифицирован с помощью соответствующей онтологии; 2) между терминами (и обозначенными этими терминами параметрами Web-сервисов) уже установлены семантические связи, которые могут обрабатываться автоматизированно не только для поиска необходимых пользователю сервисов, но и для композирования из них составных сервисов.

---

### **Литература**

---

- [Рогушина, 2014] Рогушина Ю.В. Знание-ориентированные средства поддержки семантического поиска в Web. LAP LAMBERT Academic Publishing. 2014.
- [Talantikite, 2009] Talantikite H. N., Aissani D., Boudjlida N. Semantic Annotations for Web Services Discovery and Composition. In: Computer Standards & Interfaces Journal, Elsevier, CSI, 31:1108-1117, 2009.
- [Bouchiha, 2012] Bouchiha D., Malki M. Semantic Annotation of Web Services In: Proc.ICWIT, 2012.
- 

### **Информация об авторе**

---

**Рогушина Юлия** – к.ф.-м.н., доцент, с.н.с. ИПС НАН Украины; e-mail: [ladamandraka2010@gmail.com](mailto:ladamandraka2010@gmail.com).

## КЛАСТЕРНЫЙ АНАЛИЗ: ПРОБЛЕМА АДЕКВАТНОСТИ

Ирина Рясная

**Аннотация:** Рассматриваются проблемы построения адекватных мер сходства на базе нечетких качественных характеристик. Формализовано понятие «нечеткая шкала сходства» и сформулированы условия адекватности такой шкалы с позиций репрезентативной теории измерений.

**Ключевые слова:** нечеткая мера сходства, адекватность, шкала измерения.

### Введение

При решении задач кластерного анализа на основе сходства объектов в многомерном пространстве характеристик (признаков), как правило, не исследуется проблема адекватности получаемых результатов с позиций репрезентативной теории измерений (РТИ) [Luce, 2007]. Для доказательства адекватности результатов необходимо в рамках РТИ не только доказать адекватность используемых мер сходства, но и построить соответствующую шкалу измерения. В докладе эти вопросы рассматриваются для случая измерения функций принадлежности нечетких качественных характеристик объектов в шкалах порядка, отношений, интервалов и абсолютных шкалах.

### Постановка и решение задачи

Пусть  $X$  – конечное множество элементов (объектов) эмпирической системы,  $W$  – конечное множество нечетких качественных характеристик элементов множества  $X$ . Качественной характеристике  $w \in W$  поставим в соответствие множество  $T = \{t_1, \dots, t_m\}$  вербальных значений этой характеристики,  $m = m(w)$ . Нечеткой мерой сходства на множестве  $X$  называется отображение

$$\mu: X \times X \rightarrow [0,1], \quad (1)$$

удовлетворяющее условиям рефлексивности –  $\mu(x, x) = 1$ ,  $x \in X$ , и симметричности –  $\mu(x, y) = \mu(y, x)$ ,  $x, y \in X$ . Отображение (1) задает на  $X$  нечеткое отношение сходства  $\tilde{r}$ . Матрицу, соответствующую этому отношению, обозначим  $R_{\tilde{r}}$ . Элементы этой матрицы  $\mu_{\tilde{r}}(x, y)$  представляют собой оценку сходства элементов множества  $X$ . Способ вычисления значений  $\mu_{\tilde{r}}(x, y)$  определяется как типом шкал измерения характеристик объектов эмпирической системы, так и теоретико-множественной интерпретацией семантических операторов И (and), ИЛИ (or) в форме нечетких логических операторов. Логическим эквивалентом семантического оператора И является оператор нормы, а оператора ИЛИ – оператор конормы [Поспелов, 1986]. В рамках нечеткой логики, связанной с теорией нечетких множеств, существует значительная свобода в выборе вида логических эквивалентов семантических операторов. В данной работе используются операторы, введенные Заде и Лукасевичем.

Системой с отношениями  $\mathfrak{M}$  называется кортеж  $\mathfrak{M} = \langle A, R_1, \dots, R_p \rangle$ , где  $A$  – непустое множество, называемое областью задания системы с отношениями,  $R_1, \dots, R_p$  – отношения на  $A$ . Если множество  $A$  состоит из эмпирических объектов и отношения на  $A$  определены эмпирически, то система  $\mathfrak{M}$  называется эмпирической системой с отношениями. Если  $A \subseteq \mathbb{R}^n$ , где  $\mathbb{R}$  – множество действительных чисел, система  $\mathfrak{M}$  называется  $n$ -мерной числовой системой с отношениями. Если  $A$  – множество нечисловых математических объектов (например, векторов, функций, матриц), то система  $\mathfrak{M}$  называется математической системой с отношениями.

Шкалой называется гомоморфное либо изоморфное отображение эмпирической системы с отношениями в числовую или математическую систему с отношениями. Нечеткая шкала определяется как гомоморфизм эмпирической системы с отношениями в систему с отношениями на множестве нечетких подмножеств [Блишун, 1988]. Значения функций принадлежности нечетких характеристик являются результатом прямых измерений и могут измеряться в порядковой шкале, а также шкалах отношений, интервалов и абсолютной шкале. Однако значения функции принадлежности нечеткой меры сходства являются результатом косвенных измерений, а тип и адекватность соответствующей нечеткой шкалы должны быть установлены в процессе исследования.

В докладе формализовано понятие нечеткой шкалы сходства и предложена методика построения такой шкалы, а также рассмотрены необходимые и достаточные условия её адекватности. Исследованы два способа построения нечетких мер сходства. Приведен пример построения адекватной нечеткой шкалы сходства в случае измерения значений функций принадлежности нечетких качественных характеристик объектов в шкалах порядка, отношений, интервалов и абсолютных шкалах.

---

### **Заключение**

Задачи кластерного анализа корректно решаются при использовании адекватных мер сходства и построении адекватной шкалы сходства. С позиций репрезентативной теории измерения формализовано понятие нечеткой шкалы сходства, предложена методика построения такой шкалы, рассмотрены условия её адекватности. Приведен пример построения адекватной нечеткой шкалы сходства в случае измерения значений функций принадлежности нечетких качественных характеристик объектов в шкалах порядка, отношений, интервалов и абсолютных шкалах.

---

### **Литература**

[Блишун, 1988] А. Ф. Блишун. Сравнительный анализ методов измерения нечеткости // Изв. АН СССР. Техническая кибернетика. – 1988. – № 5. – С. 152–175.

[Поспелов, 1986] Нечеткие множества в моделях управления и искусственного интеллекта / Под ред. Д.А.Поспелова. – М.: Наука, 1986. – 312 с.

[Luce, 2007] R. Luce, D. Krantz, Suppes and A. Tversky. Foundations of Measurement: Vol. III. – New York: Dover Publications, Inc., 2007. – 356 p.

---

### **Информация об авторе**

*Ирина Рясная* – научный сотрудник, Институт кибернетики им. В.М. Глушкова НАН Украины, Киев-187, проспект Академика Глушкова, 40; e-mail: riasnaia@gmail.com.

## АВТОМАТИЗИРОВАННАЯ СИСТЕМА АГРЕГАЦИИ, ПОИСКА, ВИЗУАЛИЗАЦИИ ДАННЫХ

Михаил Соколов, Сергей Николаев

**Аннотация:** Рассматривается масштабируемая автоматизированная информационная система агрегации, поиска, кластеризации и визуализации данных. Описывается архитектура системы, ее назначение, примеры использования. Исследуются методы автоматического извлечения графических данных, последующей их обработки и визуализации. Целью исследования является создание и внедрение системы агрегации, поиска и визуализации данных.

**Ключевые слова:** Агрегация, визуализация, кластеризация, датамайнинг, данные, информационные системы.

### 1. Введение

Объектом этих исследований являются архитектура системы, ее назначение, примеры использования. Здесь, предмет исследования – методы автоматического извлечения графических данных, последующей их обработки и визуализации. Цель исследования – создание и внедрение системы агрегации, поиска и визуализации данных.

В условиях экспоненциального роста интернета, создание масштабируемых систем агрегации данных позволяет упорядочить большие массивы фрагментарной информации. А системы их кластеризации и визуализации позволяют представить их в наиболее простом, для человеческого восприятия, виде. Одну из таких системы мы опишем в данной статье.

### 2. Агрегация данных

Рассмотрим агрегацию данных, как процесс сбора и извлечения соответствующей информации из различных источников, чтобы обеспечить удобство и добавить ценность, анализируя собранную информацию для конкретных целей с использованием интернет-технологий [1]. Процесс агрегации данных можно условно разделить на несколько этапов. Рассмотрим более подробно каждый из этапов.

- Первым этапом является определение конечных целей агрегации информации. Четко определенные цели позволяют ограничить категории источников.
- На втором этапе определяются и сортируются источники информации. Создается единый реестр всех источников информации.
- Третий этап – подготовка базы данных, инструментов для извлечения данных.
- Четвертый этап – непосредственно сбор данных.
- Пятый этап – последующая обработка и анализ полученных данных.

Распределенная и масштабируемая архитектура системы агрегации, которая будет описана в дальнейшем, позволяет значительно сократить время агрегации. Автоматизация на каждом из этапов является основой для создания автономной системы агрегации данных. В настоящий момент существуют автономные системы агрегации данных, работающие с минимальным участием человека на первом этапе, используемые, к примеру, в системах прогнозирования биржевых курсов. С другой стороны, существуют и системы консолидации информации, с формализацией процессов консолидации, информационной системой, однако без автоматизации. Такие системы часто используются для бизнес-аналитики и журналистики. В нашем случае, мы используем систему для агрегации визуальных программных компонентов для веб-приложений. Поэтому особое внимание мы уделяем подсистемам визуализации данных.

### 3. Архитектура масштабируемой системы

Рассмотрим кратко архитектуру масштабируемой системы агрегации визуальных программных компонентов для веб-приложений (Рис. 1).



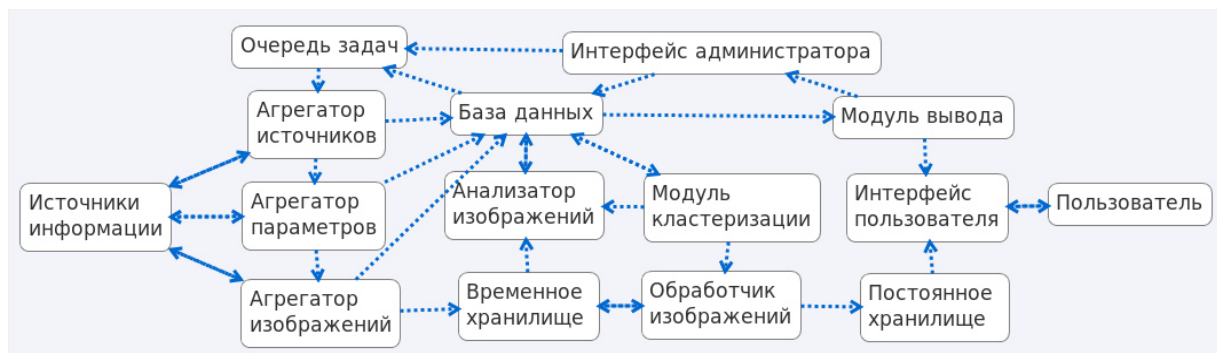


Рис. 1. Схема архитектуры масштабируемой системы агрегации данных.

Каждый из узлов системы является заменяемым. Очередь задач координирует работу компонентов системы и обеспечивает автономное выполнение заданий в строгой последовательности. Система достаточно гибка, чтобы делать возможным добавление в нее новых типов отношений, а также имеет распределенную архитектуру, чтобы не зависеть целиком от центрального компьютера и от центрального органа управления. Временное хранилище обеспечивает доступ к данным, полученным из источников. В постоянном хранилище находится уже обработанная информация, готовая к использованию. Распределенная архитектура обуславливает распределение не только вычислительных ресурсов, но и глобального управления системой и взаимодействия узлов. В распределенной архитектуре используются несколько параллельно работающих процессоров, подающих информацию в систему шин, из которой она может быть получена другими пользователями или процессорами. Шина может быть использована для запроса информации от других процессоров. Создаваемая шинами горизонтальная интеграция допускает обмен информацией между всеми процессорами и другими элементами системы и обеспечивает высокую степень гибкости и адаптивности.

#### 4. Кластеризация данных

Рассмотрим механизм кластеризации данных [2]. В нашем примере мы используем кластеризацию изображений и параметров. Кластеризация - задача разбиения заданной выборки объектов на непересекающиеся подмножества, называемые кластерами, так, чтобы каждый кластер состоял из схожих объектов, а объекты разных кластеров существенно отличались. Задача кластеризации относится к широкому классу задач обучения без учителя. В нашем случае кластеризация будет производиться по визуальным признакам и параметрам объектов. Из визуальных признаков выделим цвета спектра RGB и контрастность изображения. Из параметров: набор технических характеристик визуальных программных компонентов для веб-приложений, такие как совместимость с браузером, совместимость с устройством, совместимость с платформой веб-приложения и т.п. В результате кластеризации пользователь имеет возможность быстро находить визуальные программные компоненты для веб-приложений по определенным признакам.

#### 5. Поиск информации

Так как данные приходят в систему из различных источников, то у каждого элемента данных могут присутствовать свои дополнительные поля и характеристики, которых может не быть в основной системе. У каждого источника данных из единого реестра источников присутствует уникальный набор полей, который однозначно описывает все элементы данных, полученные с этого источника. Поэтому на последнем пятом этапе агрегации данных из всех источников производится их обработка с извлечением полезной информации и её унификацией в единый формат, по которому в дальнейшем и будет производиться поиск. При этом для каждого графического элемента данных сохраняется метainформация, которая включает как количественные (численные характеристики объектов), так и качественные (лингвистические) признаки, такие как:

- описания элементов данных;
- множества меток (тегов);
- названия и взаимосвязи с другими элементами.

Поисковый механизм системы использует лингвистическую и цветовую метаинформацию для нахождения допустимого множества ответов, а с помощью числовых признаков производится ранжирование поисковой выдачи. Следует отметить, что формула расчёта ранжирования результатов [5] учитывает известную метаинформацию о пользователях, которые производят запрос, таким образом, чтобы увеличить релевантность результатов. Так же в качестве части ранжирующего критерия может быть использован результат коллаборативной фильтрации [3,4], полученный путём отслеживания подмножеств взаимно интересных элементов данных похожим пользователям.

---

## 6. Визуализация данных

Альтернативой стандартному поиску выступает подсистема визуализации кластеров данных. Принцип данной системы основан на использовании проектирования кластеров графически заданных данных на плоскую карту. В качестве параметров кластеризации задаются как цветовые характеристики элементов данных, так и лингвистические признаки.

При проектировании данных на карту задаётся шаблон проекции, который и служит затравкой для топологического размещения элементов данных на карте. При использовании шаблонов проекции удаётся достичь эффекта близкого топологического расположения на карте близких по метрике кластеризации графических элементов, при этом сохраняя высокие эстетические свойства визуального отображения карты. Используя различные фильтры, возможно, так же как и в поиске, изменять отображаемую информацию и топологии карт.

---

## 7. Заключение

В данной статье была рассмотрена масштабируемая автоматизированная распределённая система агрегации, поиска и визуализации данных, заданных в графическом виде. Архитектура рассматриваемой системы является многоуровневой и адаптивной к различным источникам данных, что даёт возможность системе агрегировать данные с разнородных источников. После сбора из данных извлекается ценная информация, которая сортируется и сохраняется в единой базе. Применённые механизмы кластеризации и визуализации позволяют улучшить удобство и повысить скорость и качество поиска информации, интересной пользователям.

По нашему мнению, исследования в этом направлении позволяют улучшить доступность информации, скорость и качество поиска, что повышает удовлетворённость пользователей и удобство использования сервисов, построенных по принципу рассматриваемой архитектуры.

---

## Литература

1. Hongwei Zhu, Michael D. Siegel, Stuart E. Madnick. Information Aggregation – A Value-added E-Service, 2001
2. David Pinto, Paolo Rosso, Alfons Juan. A Comparative Study of Clustering Algorithms on Narrow-Domain Abstracts, [http://users.dsic.upv.es/~proso/resources/PintoEtAl\\_SEPLN06.pdf](http://users.dsic.upv.es/~proso/resources/PintoEtAl_SEPLN06.pdf)
3. Joonseok Lee, Mingxuan Sun, Guy Lebanon. A Comparative Study of Collaborative Filtering Algorithms. May 14, 2012. <http://arxiv.org/pdf/1205.3193.pdf>
4. Michael D. Ekstrand, John T. Riedl, Joseph A. Konstan. Collaborative Filtering Recommender Systems <http://files.grouplens.org/papers/FnT%20CF%20Recsys%20Survey.pdf>
5. Panagiotis Takis Metaxas. On The Evolution Of Search Engine Rankings [http://cs.wellesley.edu/~pmetaxas/Metaxas-Evolution\\_Search\\_Engine\\_Rankings.pdf](http://cs.wellesley.edu/~pmetaxas/Metaxas-Evolution_Search_Engine_Rankings.pdf)

---

## Информация об авторах

**Михаил Соколов** - Международный научно-учебный центр информационных технологий и систем НАН и МОН Украины, 03680 ГСП, г. Киев, проспект акад. Глушкова, 40; e-mail: [nvarius@gmail.com](mailto:nvarius@gmail.com)

**Сергей Николаев** – НТУУ «КПИ» УНК «ИПСА», e-mail: [sergiynicolaev@gmail.com](mailto:sergiynicolaev@gmail.com)

---

## О СЛЕДУЮЩЕМ ЭТАПЕ КОМПЬЮТЕРНОЙ СЕМАНТИКИ

Владимир Сторож

***Аннотация:** Для эффективного анализа текста с помощью компьютера необходима переработка текста под задачи пользователя, в том числе путем создания новых по отношению к анализируемому тексту интеллектуальных объектов в виде концептов и схем. Такие интеллектуальные объекты представляют собой процессы второго порядка и могут быть реализованы программным образом как динамические объекты*

***Ключевые слова:** анализ текста, компьютерная семантика, интеллектуальный объект*

---

### Введение

Количество информации, генерируемой человечеством, растет значительно быстрее, чем возможности отдельного человека воспринимать и анализировать эту информацию. Интернет сам по себе лишь незначительно решает задачу анализа информации. Является насущно необходимым создание компьютерных методов анализа текстовой информации, поиск в этой информации идей, необходимых пользователю и их предоставление в сжатом и адаптированном виде.

---

### Основное содержание

Современные компьютерные модели семантики текста и речи имеют значительно более бедный характер по сравнению с богатством содержания человеческого мышления, возникающего в процессе восприятия текста или речи. Это существенно ограничивает использование компьютеров при анализе текстов, особенно в научной деятельности и любой творческой работе.

Ограниченность компьютерного анализа текста вызвана тем, что компьютер «улавливает» лишь незначительную часть «релевантной» семантики, возникающей у человека при прочтении текста. В компьютере семантика текста сводится к сумме семантики отдельных слов текста и их связей. Человек же в результате интерпретации и понимания текста создает существенно отличающееся от исходного текста содержание, в результате чего может появиться ряд принципиально новых понятий, важных для понимания текста и отсутствующих в исходном варианте, новые связи, выводы и т.д. [1].

Вследствие этого, для перехода на следующий уровень компьютерной семантики необходимо реализовать компьютерное «понимание» текста и на этой основе построение новой (по отношению к исходному тексту) системы взаимосвязанных информационных, или правильнее сказать, интеллектуальных, объектов, содержащих некие модельные представления относительно текста и его восприятия субъектом.

Ключевым фактором в новой семантике являются создаваемые в процессе интерпретации и понимания текста интеллектуальные объекты. Возникает целый ряд вопросов об их природе, возможностях, инструментарии для построения и др. Без решения этих и других возникающих вопросов все попытки создания новой семантики будут носить несистематичный характер и приведут лишь к существенно ограниченным результатам. Поэтому сформулируем эти вопросы и возможные направления их решения.

Что есть интеллектуальный объект в искусственном интеллекте, да, собственно и сам искусственный интеллект? Модель естественного интеллекта? Частичная реализация, и что это такое? Каков вклад в интеллект физической и программной составляющей? И ряд других вопросов. Очевидно, в зависимости от ответов на эти вопросы для создания интеллектуальных объектов могут быть использованы совершенно разные принципы, дисциплины, инструментарий, да и способы физической реализации.

Специалисты по искусственному интеллекту используют понятия содержания, интерпретации и семантики как в значительной мере взаимозаменяемые. Не оспаривая общности в содержании этих понятий,

---

необходимо отметить также и существенные различия между ними. Указанные понятия должны нести различную смысловую нагрузку, участвовать в различных объяснительных схемах, и решать различные задачи. А неконтролируемое смешивание и подмена этих понятий приводит к путанице, в результате чего ни специалисты по ИИ, ни даже лингвисты не могут сказать, в чем же проблемы, и как их решать.

Существует целый ряд интеллектуальных и физических объектов, используемых при высокоуровневом анализе информации: фреймы, классы (в программировании), концепты, автоматы по фон Нейману, искусственные нейроны, реальные нейроны и др. Необходимо в большей степени понять степень общности между этими понятиями и объектами, и принципы их различия. Это позволит более эффективным образом строить на их основе интеллектуальные объекты новой семантики.

По своей сути интеллектуальные объекты новой семантики должны представлять собой некие устойчивые или квазиустойчивые процессы, возникающие и существующие в результате взаимодействия неких первичных процессов, т.е., образно говоря, процессы второго порядка. В мозгу живых существ такие процессы второго порядка реализуются за счет большого количества синапсов, приходящихся на один нейрон (порядка  $10^4$ ), системы иерархических связей, сложности синапсов. С точки зрения компьютерной идеологии интеллектуальные объекты семантики должны представлять собой некие динамические программные объекты, включающие в себя возможности: а) создания новых объектов типа экземпляра класса с учетом входной информации, знаний и целей субъекта; б) включения во вновь созданный объект уже существующих объектов, в частности, базовых описаний концептов, схем, фреймов, графов; в) рассмотрение в интеллектуальных операциях новых или модифицированных объектов как неких целостностей с точки зрения семантики и др.

Реализация изложенных выше идей возможна в виде программных продуктов для отдельных предметных областей, задач, пользователей. Такие продукты должны иметь развитые возможности диалога с пользователем, настройки и самосовершенствования в результате, как диалога, так и внутренних интеллектуальных процессов. В случае активного диалога с пользователем, настройки под его задачи и выполнения большого объема работ по анализу текстов и речи, такой программный продукт может приобрести некоторые черты индивидуальности и исполнять роль индивидуального аналитического помощника текстовой и схемной информации.

---

## **Заключение**

Следующий этап развития компьютерной семантики должен существенно основываться на переработке анализируемого текста и заменой его адаптированным под цели пользователя содержанием. Для этого необходимо решить проблему создания новых интеллектуальных объектов в виде сложных динамических процессов. Решение этой проблемы позволит переложить с пользователя на компьютер поиск нужных идей в чрезвычайно больших объемах информации, находящихся в Интернет.

---

## **Библиография**

[Сторож, 2014 ] В.В. Сторож Семантика и онтологии в естественных и искусственных системах В: Open Semantic Technologies for Intelligent Systems - OSTIS-2014 - Минск: БГУИР, 2014. – с.415-422.

---

## **Информация об авторе**

**Владимир Сторож** – к.ф.-н.м., доцент, докторант Донецкого национального технического университета, Донецк, Украина; e-mail: ws52@mail.ru

---

## О ПРИМЕНЕНИИ СИСТЕМОЛОГИИ В ТРАНСДИСЦИПЛИНАРНЫХ ИССЛЕДОВАНИЯХ

Екатерина Соловьева

**Аннотация:** Рассматривается целесообразность применения системологии и метода системологического классификационного анализа знаний в трансдисциплинарных исследованиях для обеспечения их целостности, эффективности и обоснованности.

**Ключевые слова:** трансдисциплинарность, знания, сущность, системология, системологический классификационный анализ, система, целостность, ноосфера.

---

### Введение

Трансдисциплинарность актуальна для науки и образования будущего; устойчивого развития цивилизации, государства и бизнеса; создания интеллектуального капитала, *ноосферы*, общества и экономики знаний. Трансдисциплинарность как исключительно важное явление в развитии науки, общества, человека можно рассматривать в различных аспектах. В тезисах невозможно охватить все аспекты, остановимся на одном из определяющих - системном аспекте, и очень кратко – на духовном.

---

### Анализ целесообразности применения системологии и системологического классификационного анализа в трансдисциплинарных исследованиях

Возникновение трансдисциплинарности свидетельствует о понимании необходимости *целостного* (эмерджентного) рассмотрения действительности. Трансдисциплинарность основывается на *содержательной качественной стороне информации*, работе со знаниями, трансдисциплинарной этике. Подобные идеи высказаны еще в Хартии Трансдисциплинарности, принятой в 1994 году. Трансдисциплинарность связана с проникновением в *сущность*, углублением на системной основе нашего понимания и знаний, с духовным развитием человека.

Ученые приходят к пониманию необходимости нового системного подхода для трансдисциплинарных исследований. Например, разработка трансдисциплинарного системного подхода включена в цели Института трансдисциплинарных технологий. Отмечается также сдвиг научного знания «в область объемлющего его знания жизненного мира» (Хартия Трансдисциплинарности 2). Для трансдисциплинарных исследований необходим новый системный подход, который рассматривает любой объект целостно, как систему; учитывает определяющую ее надсистему, поэтому целесообразно применение системологии в трансдисциплинарности.

*Системологический подход* ноосферного этапа развития науки исследует систему, начиная с определения надсистемы и ее запроса к системе; учитывает функциональную, структурную и субстанциональную целостность системы; обеспечивает понимание сущности системы, причин ее возникновения и динамики развития. *Системология* может быть успешно применена при исследовании *сложных качественных неформализованных (ill-structured) проблем* в произвольных, в том числе слабоструктурированных предметных областях. Системология позволяет изучать системы первой природы, не созданные человеком, и системы, включающие человека.

Обобщение понятия системы системологии Г.П. Мельникова позволяет эффективно применять системологию для исследования знаний [1]. Так, новый когнитивный *метод системологического классификационного анализа* знаний на основе конструктивных критериев естественной классификации, обобщающих и развивающих правила формальной логики, позволяет систематизировать понятия и получать новые глубинные знания обоснованно с учетом существенных и сущностных свойств [2].

Таким образом, системология может стать системной методологией трансдисциплинарных исследований, а системологический классификационный анализ и разработанная и обоснованная на его основе система категорий С.И. Маторина - базой для их успешного развития. Системологический подход как здравый смысл и целостный ноосферный взгляд «сверху» на систему, ее сущность, свойства, адаптацию, эволюцию, функционирование и развитие, позволяющий системно рассматривать материальный и нематериальный мир, в том числе - знания и менеджмент знаний, обеспечит обоснованные решения и прогнозы в любых сферах деятельности человека и в трансдисциплинарных исследованиях.

К трансдисциплинарному образованию можно отнести системологическое знаниеориентированное когнитивное обучение магистерской специальности «Консолидированная информация» на базовой в Украине кафедре *Социальной информатики* в Харьковском национальном университете радиозлектроники. Обучение включает формирование нового ноосферного системного мышления, основано на содержательной обработке информации и системной работе со знаниями, направлено на повышение конкурентоспособности и улучшение любой деятельности с учетом этических принципов.

Духовный аспект не входит в обучение и не всегда упоминается в трансдисциплинарных работах, но в связи с тем, что Хартия Трансдисциплинарности включает толерантность (как фундаментальный признак трансдисциплинарного подхода и видения, который «предполагает признание права на идеи и истины, противоположные нашим собственным»), хотелось бы отметить не общепринятую возможность применения в трансдисциплинарных исследованиях метода буддизма (не религиозного) работы с человеческим умом. Этот метод близок к представлениям современной науки, его поддерживали и лауреаты Нобелевской премии. Например, А. Эйнштейн утверждал, что буддизм представляет собой «последовательную логическую систему, основанную на опыте переживания реальности во всей ее полноте, и поэтому может считаться соответствующим научным стандартам» [3, с.23].

---

## Выводы

В результате проведенной работы можно сделать выводы о целесообразности и необходимости применения ноосферного системологического подхода и системологических методов для обеспечения целостности, качества, глубины и обоснованности трансдисциплинарных исследований, эффективной систематизации знаний, содержательной обработки информации, познания сущности, решения сложных качественных проблем и дальнейшего системного развития трансдисциплинарности.

Автор выражает искреннюю благодарность академику А.В. Палагину за обсуждение вопросов трансдисциплинарности и доктору философии Г.А. Брахману за обсуждение применений системологии в материальных и нематериальных сферах.

---

## Список использованной литературы

1. Бондаренко, М.Ф. Основы системологии / М.Ф. Бондаренко, Е.А. Соловьева, С.И. Маторин. – Харьков: ХТУРЭ, 1998. – 118 с.
2. Соловьева Е.А. Естественная классификация: системологические основания.– Харьков: ХНУРЭ, 1999. – 222с.
3. Форма и пустота. От буддизма к науке и обратно. Под ред. А. Пшибыславского.– Киев. Вогняне колесо, 2013.-231с.

---

## Сведения об авторе

**Екатерина Соловьева** – зав. кафедрой *Социальной информатики*, профессор, доктор технических наук, Харьковский национальный университет радиозлектроники, пр. Ленина, 14, Харьков 61166, Украина; e-mail: si@kture.kharkov.ua , gt\_ekasolo@yahoo.com

## КЛАССИФИКАЦИЯ БИОЭЛЕКТРИЧЕСКИХ СИГНАЛОВ НА ОСНОВЕ НЕЙРОПОДОБНЫХ МОДЕЛЕЙ

Наталья Н. Филатова, Дмитрий М. Ханеев

**Аннотация:** В статье описывается алгоритм классификации биоэлектрических сигналов. Данный классификатор основан на идее растущей пирамидальной сети, адаптированной для работы с нечеткими описаниями объектов. Рассматриваемый алгоритм решает задачу классификации, объединяя ее с задачей сегментации.

**Ключевые слова:** классификация, растущие пирамидальные сети, графы, обучающая выборка, нечеткая логика.

### Введение

Одно из современных направлений развития медицинской техники связано с созданием виртуальных приборов для функциональной диагностики (ВПФД) которые можно эффективно применять при скрининговых методиках обследования. Для автоматизации процедур их проектирования предлагается унифицировать структуру ВПФД, выделив базовое ядро, решающее задачу классификации биоэлектрических сигналов и формирования правил для их распознавания. В докладе рассматривается один из возможных путей решения этой задачи с помощью специальной нейроподобной иерархической структуры (НИС), основанной на идее растущих пирамидальных сетей [Гладун, 2004] и позволяющей генерировать правила классификации сигналов с нечетко проявляющимися свойствами.

### Автоматическая генерация правил классификации биоэлектрических сигналов

Биоэлектрические сигналы (БЭС), являющиеся объектами классификации, представлены обычно большим числом точек [Ifeachor, 2002]. Функциональные преобразования позволяют уменьшить размерность описания, например, для представления электроэнцефалограммы (ЭЭГ) по 9 отведениям рассчитываются спектры мощности, которые в совокупности позволяют представить этот объект вектором, включающим 558 составляющих.

Существующие индивидуальные особенности источников сигналов, позволяют выдвинуть гипотезу о целесообразности перехода к лингвистическим переменным при описании подобных графиков. Тогда координаты точек графика по оси абсцисс будут рассматриваться как список признаков, и значения каждого признака будет определяться как нечеткое множество, Supp которого задается на оси ординат. Таким образом, нечеткими становятся все оценки ординат точек графика. Для нелинейных графиков разброс значений каждого признака на множестве объектов одного класса будет различный. В связи с этим для фазификации каждого признака необходимо создание индивидуальной лингвистической шкалы. Так как все признаки характеризуют точки одного графика, то для построения всех нечетких шкал можно использовать одно и то же терм-множество, включающее три термина: HI, MID, LOW. Таким образом, обучающая выборка представляется нечеткими описаниями паттернов биоэлектрических сигналов или их функциональных преобразований.

Построение сети выполняется в два последовательных этапа. На первом этапе формируется структура сети, в сеть вводятся новые вершины и связи между ними, для этого используются правила из работы [Гладун, 2004], модифицированные для анализа объектов, заданных нечеткими признаками. Основная задача первого этапа заключается в выделении пересечений в признаковых описаниях объектов. На втором этапе осуществляется разметка контрольных вершин сети, которым соответствуют наиболее характерные для определенного класса сочетания признаков. Построенный классификатор БЭС локализует участки с одинаковыми значениями нечетких признаков, но это фактически приводит к сегментации паттернов на участки с близкими оценками их структурных свойств. Процедура сегментации

---

сигналов позволяет сформировать новое пространство укрупненных признаков и воспроизвести описания объектов с помощью иерархически взаимосвязанных признаков

Предложенные идеи реализованы в виде нейроподобного иерархического сетевого (НИС) классификатора, который позволяет автоматически формировать понятия о классах объектов по разнородным данным, с возможностью добавления новой информации о классах по мере ее поступления. Структура НИС легко конвертируется в набор логических высказываний, которые используются в дальнейшем как автономные правила, определяющие отдельные классы сигналов. В режиме классификации с помощью алгоритма нечеткого логического вывода осуществляется анализ описания объекта, и формируется суждение о степени его принадлежности к классам.

---

### Исследование НИС-классификатора

Программа-классификатор исследована при решении задач классификации знака эмоции на основе анализа спектральной плотности мощности (СПМ) записей сигналов электроэнцефалограмм (ЭЭГ) и речевого сигнала (РС), а так же классификации дыхательных шумов из классов «Норма» и «Патологии». Исследования показали возможность выделения интервалов наиболее информативных признаков, позволяющих обеспечить хороший уровень обобщения и уточнения описания объектов в классах. Для ДШ наиболее информативные признаки находятся в диапазоне 0-2100 Гц, для РС - в диапазоне 1100-5100 Гц. Для паттернов ЭЭГ признаки, получаемые из отведений только правого полушария головного мозга.

Обобщение признаков способствует улучшению результатов классификации, однако при этом НИС усложняется, что проявляется в увеличении количества ассоциативных элементов и связей между ними.

Классификационные правила, полученные с применением алгоритма объединения признаков, лучше описывают исследуемые классы и позволяют увеличить точность классификации. Максимально допустимый уровень шума, обеспечивающий разделение объектов не должен превышать 20%.

---

### Заключение

НИС классификатор прошел тестирование на искусственных и реальных клинических данных. Набор сгенерированных классификационных правил, отображающих закономерности в структуре НИС, и результаты классификации выборок сигналов (ДШ, РС и ЭЭГ) в большинстве случаев согласуются по форме с логическими выводами, сделанными экспертами при анализе этих же выборок.

---

### Библиография

- [Гладун, 2004] Гладун В.П. Растущие пирамидальные сети // Новости искусственного интеллекта. 2004. №1.
- [Филатова, 2013] Филатова Н.Н., Ханеев Д.М. Применение нейроподобных сетевых структур для генерации гипотез правил классификации // Нечеткие Системы и Мягкие Вычисления. 2013. Т. 8. №1. С. 43–58.
- [Ifeachor, 2002] Ifeachor E.C., Jervis B.W. Digital Signal Processing: A Practical Approach (2nd Edition). Pearson Education, Upper Saddle River, NJ, USA, 2002. 933 p.

---

### Информация об авторах

**Филатова Наталья Николаевна** – ТвГТУ, д.т.н., профессор; Россия, Тверь, пр. Ленина, 25, 170023;  
e-mail: [nfilatova99@mail.ru](mailto:nfilatova99@mail.ru)

**Ханеев Дмитрий Михайлович** – ТвГТУ, аспирант; Россия, Тверь, пр. Ленина, 25, 170023;  
e-mail: [t\\_ran@mail.ru](mailto:t_ran@mail.ru)



## ПРИМЕНЕНИЕ ШАБЛОНОВ ПРИ ПРОЕКТИРОВАНИИ ОНТОЛОГИЙ ДЛЯ ИЗВЛЕЧЕНИЯ ЗНАНИЙ ИЗ ПРАВОВЫХ ДОКУМЕНТОВ

Екатерина Хала

**Аннотация:** Проектирование правовой онтологии - трудоемкая задача, которая требует намного больше, чем просто экспертных знаний. Частично это обусловлено тонкостью юридических рассуждений, многозначностью логических утверждений в правовых приложениях, неполнотой и противоречивостью применяемой законодательной базы. В работе предлагаются некоторые проектные решения, основанные на выделении регулярных выражений в корпусе правовых текстов для построения шаблонов синтаксических структур знаний при синтезе семантических структур знаний, которые предназначены для разработки правовых онтологий. Представлен обзор типовых задач и услуг для представления правовых знаний, введено понятие шаблона онтологии, и показана эффективность его использования при работе с правовой областью.

**Ключевые слова:** правовая онтология, структура онтологии, анализ текста, регулярные выражения, текстовый контент.

### Введение

Для моделирования и описания правовых явлений и процессов актуальна задача создания систем, основанных на знаниях. Такие системы, включающие в себя элементы экспертной обработки и интеллектуального поиска, могут стать незаменимым инструментом в правотворческой и правоприменительной деятельности, стать основой различных семантических приложений.

Первым шагом в разработке такой системы является построение и принятие *онтологии* – в роли единого источника документированной совокупности терминов предметной области и правил, согласно которым эти термины могут быть использованы для построения достоверных утверждений о состоянии системы, а также для санкционированного логического вывода новых утверждений. Используя онтологии, пользователи и системы могут общаться друг с другом, поддерживая информационный обмен, базирующийся на семантике (а не только на синтаксисе) этого единой терминологической базы.

### Функциональность и методы правовой онтологии

Из-за автономии и зависимости правового знания и от физических, и от социальных знаний, правовые задачи рассуждения развивались специфическим образом, например *Требование* → *Последствие* (если *P* фактические знания, то правовые знания *Q*), *Обязательство* → *Право* (если у *A* есть обязательства по отношению к *B*, то *B* имеет право к *A*), *Норма* ↔ *Случай* (если ситуация выполняет условия для нарушения нормы, это становится правовым случаем – судебным делом), *Сценарий Преступления* (преступление совершено преступником и привлекает внимание властей, которые преследуют преступное действие) и др.

Разработчик онтологии имеет дело с проектированием, управлением и эксплуатацией онтологий в пределах информационных систем. *Преставление знаний с помощью онтологии* – наиболее выразительный способ, использующий все преимущества явного представления знаний. Достоинством данного способа является то, что онтологии можно объединять с другими компонентами, чтобы построить семантически-явные приложения, например, использовать совместно с: доказательством теорем, категоризацией и классификаторами экземпляра [Gangemi, 2004], вычислительными словарями [Gangemi, 2005], [Sagri, 2005], алгоритмами планирования [Gangem, 2004], [Sagri, 2005], случаями, основанными на рассуждениях [Sagri, 2005], системами основанными на правилах [Motta, 2003], [Sagri, 2005].

Для обеспечения работы с различными типами документов, в систему необходимо включить компоненты, обеспечивающие унифицированный доступ к документам в различных форматах. Реализация таких задач может быть основана на использовании шаблонов, образцов, позволяющих распознать структуру документа. Использование *регулярных выражений* [Friedl, 2003] позволяет легко учитывать различные формы слова и работать с большими объемами информации. Недостатком регулярных выражений является то, что при поиске они не позволяют учитывать местонахождение искомого слова/фразы.

Методы разработки онтологии используются в контексте «случаи общего применения», определенные для области применения. Некоторые методы, которые могут быть реализованы посредством семантически-явных приложений в правовой области, приведены ниже.

- Межсубъектные соглашения и смысловые договоренности, задача которых состоит в получении согласия (или разногласия) о подразумеваемом смысле правового термина, правового текстового модуля, и т.д.
- Извлечение знаний, процесс извлечения понятий, отношений, поименованных объектов, и сложных шаблонов знаний из базы данных, документа, или корпуса.
- Проверка соответствия – это задача, нацеленная на проверку того, что социальная ситуация (которая неким образом совместима с правовыми нормами) удовлетворяет правовому описанию. В обобщенном случае, когда ситуации, по некоторым причинам, уже известны, как правовые отношения (например, ситуация с преступлением), их можно проверить на соответствие дальнейшего правового описания (например, процедура решения по апелляции).
- Юридическая консультация подразумевает исследование отношений между, например, судебными делами и ситуациями со здравым смыслом.
- Сравнение норм – это проверка соответствия между различными нормами. Сравнение норм включает такие задачи, как:
  - нормативная проверка конфликта и обработка норм с тем же самым типом ситуации,
  - выявление неявных отношений между нормой и другими нормами из известного корпуса.
- Перефразирование норм, выражение содержания норм в различных терминах, которые могут быть или переводами на другие естественные языки, или разными формами в пределах одного и того же естественного языка.
- Службы соответствия и композиции включают операции, осуществляющие описание служб, например, для проверки того, что предложенная служба соответствует запрашиваемой, или компоновки нескольких простых служб для получения более сложной.

---

### Шаблоны проектирования онтологии

---

Семантически-явные приложения в правовой области требуют соответственно разработанных правовых формальных онтологий. Часть задач проектирования можно упростить путем создания или извлечения «шаблонов проектирования онтологий» (ШПО) для прикладной области [McGuinness, 2006]. Реальные преимущества от использования ШПО в жизненном цикле онтологии могут быть получены, если ШПО имеют следующую функциональность: классификация ШПО должна базироваться или на прецедентах, которые она поддерживают, или на понятиях, которые она кодирует; должны присутствовать алгоритмы сопоставления с шаблоном для получения ШПО, например, спецификации естественного языка или эскиза онтологии; поддерживаться специализации и структуры ШПО, взаимодействие между ШПО для визуализации, обсуждения и создания базы знаний; присутствовать богатый набор метаданных для манипулирования ШПО и использования в приложениях.

ШПО Норма ↔ Случай (рис. 1) используется в качестве образца для представления судебного дела, например, для исследования преступления и сопоставления с другими шаблонами; в качестве проверки конфликтов в нормах. В ШПО Норма ↔ Случай нормы используют задачи, роли и параметры. Судебные дела соответствуют нормам, когда действия, объекты и значения классифицируются задачами, ролями, и параметрами соответственно. Кроме того, отношения между правовыми ролями, задачами и параметрами соответствуют отношениям между объектами, действиями и значениями.

Интуитивно понятная характеристика ШПО обеспечивается следующим:

- ШПО является эталоном для представления, и, возможно, решения проблем моделирования;
- ШПО "извлекает" связный фрагмент эталонной онтологии;
- Отображение и композиция шаблонов требуют эталонной онтологии для того, чтобы проверить непротиворечивость композиций, или сравнить наборы аксиом, которые должны быть отображены;
- ШПО может быть представлен на языке представления онтологии, но его интуитивная и компактная визуализация является важным требованием;
- ШПО может быть элементом в частном порядке, где отношение упорядочивания требует, чтобы, по крайней мере, один из классов или отношений в шаблоне был специализирован. Иерархия ШПО может быть построена путем специализации или обобщения некоторых элементов (или классов, или отношений).

- ШПО должен интуитивно иллюстрироваться и улавливать соответствующие “базовые” понятия области.
- ШПО может часто создаваться из неформальных или упрощенных схем, используемых экспертами в этой области, вместе с поддержкой другого допускающего повторное использование ШПО или эталонной онтологии и методологий для анализа онтологии предметной области [McGuinness, 2006], [Moore, 2002], [Sagri, 2005].
- ШПО может походить на схемы баз данных, но шаблон определяется по отношению к эталонной онтологии, и имеет общий характер, независимо от конструкции системы [McGuinness, 2006], [Moore, 2002], [Gangemi, 2007].

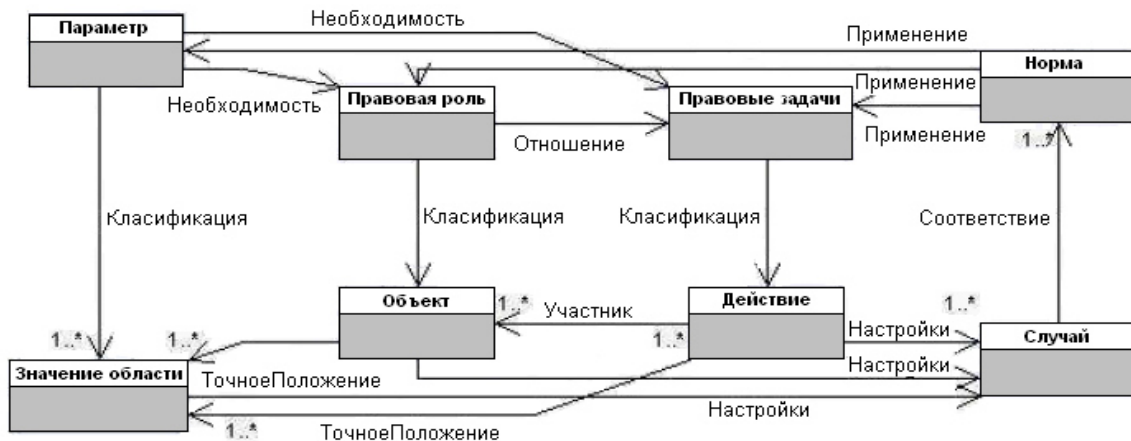


Рис. 1. ШПО Норма ↔ Случай

## Заключение

Была обоснована актуальность использования шаблонов проектирования онтологии для инженерии правовых знаний (ИПЗ). Для некоторых универсальных вариантов использования ШПО для ИПЗ, были приведены решения по обоснованию и моделированию. Приведен пример ШПО Норма ↔ Случай, показана эффективность использования ШПО при работе с правовой областью.

## Литература

- [Friedl, 2003] Jeffrey E. F. Friedl Mastering Regular Expressions. Third Edition, O'Reilly Media, 534p. 2006.
- [Gangemi, 2004] Gangemi, A., F. Fisseha, J. Keizer, J. Lehmann, A. Liang, I. Pettman, M. Sini, M. Taconet: A Core Ontology of Fishery and its Use in the FOS Project, in (Gangemi and Borgo, 2004) (2004).
- [Gangemi, 2005] Gangemi, A., Sagri M.T., Tiscornia D.: A Constructive Framework for Legal Ontologies. In R. Benjamins, P. Casanovas, J. Breuker, A. Gangemi (eds.): Law and the Semantic Web, Springer, Berlin (2005).
- [Gangemi, 2007] Gangemi A., Mika P. Understanding the Semantic Web through Descriptions and Situation, Meersman R, et al. (eds.), Springer, Berlin, 2007.
- [McGuinness, 2006] McGuinness D.L. and van Harmelen F. (editors), OWL Web Ontology Language Overview, W3C Recommendation, 10 February 2006.
- [Moore, 2002] Moore M.S. Legal Reality: A Naturalist Approach to Legal Ontology. Law and Philosophy 21: 619–705p, 2002.
- [Motta, 2003] Motta, E., Domingue, J., Cabral, L., Gaspari, M. (2003) IRS-II: A Framework and Infrastructure for Semantic Web Services. 2nd International Semantic Web Conference 20-23 October 2003.
- [Sagri, 2005] Sagri M.T., Tiscornia D., Gangemi A., An Ontology-based Approach for Representing “Bundle-of-rights”, M. Jarrar & A. Gangemi (eds.), in OTM Workshops, Springer, 2005.

## Информация об авторе

**Екатерина Хала** – аспирантка в Международном научно-учебном центре информационных технологий и систем НАН и МОН Украины, 03680 ГСП, г. Киев, проспект акад. Глушкова, 40; e-mail: [cecerongreat@ukr.net](mailto:cecerongreat@ukr.net)

## ПРОЕКТ СОЗДАНИЯ ЭЛЕКТРОННОГО МОЗГА ИНТЕЛЛЕКТУАЛЬНЫХ РОБОТОВ (ЭМИР)

Виталий Яценко

**Аннотация:** В работе рассматривается проект по развитию электронного мозга интеллектуальных роботов. Проект базируется на основе нового типа нейронных сетей - многомерных рецепторно - эффекторных нейронных растущих сетей.

**Ключевые слова:** электронный мозг, многомерные рецепторно - эффекторные нейронные растущие сети.

### Введение

В настоящее время в мире идет бурное развитие роботов и роботизированных устройств. Ян Пирсон, один из самых известных британских футурологов, утверждает: год 2015-й. В быту в качестве прислуги будут повсеместно использоваться домашние роботы; год 2018-й. Искусственный интеллект получит Нобелевскую премию; год 2020-й. Электронные формы жизни получают конституционные права; год 2025-й. В развивающихся странах будет больше роботов, чем людей; год 2030-й. Роботы и физически, и умственно превзойдут людей [1]. Для того чтобы не произошло умственного превосходства роботов над человеком, ученые работают над идеей симбиоза мозга человека с компьютером, а вернее с всемирной сетью компьютеров. Американские ученые Т.Гордон и О.Халмер предполагают, что эта идея симбиоза должна реализоваться около 2015 года. Виктор Михайлович Глушков прогнозировал 2020 год. На мой взгляд, этот прогноз более реален. Однако темпы развития науки опережают любые, даже самые смелые прогнозы.

Как стало известно, в Гонконге в совете директоров венчурного фонда Deep Knowledge Ventures (DKV) работает искусственный интеллект VITAL (Validating Investment Tool for Advancing Life Sciences). Это самообучающаяся программа, которая способна давать рекомендации по инвестициям в сектор life science [2].

В мире на решение проблемы создания полноценного искусственного интеллекта вкладываются огромные средства: **Human Brain Project** проект заказан правительством США. Объем финансирования - 1 млрд. евро на 2012-2023. Источник финансирования - Швейцарское правительство, гранты Евросоюза. Проект **Spinnaker** - Spiking Neural Network Architecture. Манчестерский университет, Университет Саутгемптона (Великобритания). Источник финансирования - EPSRC (Engineering and Physical Sciences Research Council). Объем финансирования £ 2.5m на 2005-2010 и еще £ 2.5m на 2010-2014. Проект **NeuroGrid** - Grid technology for neurosciences. IBM, Стэндфордский университет и др. (США). Источник финансирования - NSF (National Science Foundation), NIH (National Inst of Health). Объем финансирования \$ 4,9 млн. Проект **BrainScaleS** - Brain-inspired multiscale computation in neuromorphic hybrid systems, 18 исследовательских групп (в т.ч. 13 университетов) из 10 европейских стран. Источник финансирования грант Евросоюза. Объем финансирования € 8 500 000 сначала, плюс € 700 000 в продолжение. 2005-2010 - проект FACETS, с 2011 – 2014 BrainScaleS.

Проект **SyNAPSE** - Systems of Neuromorphic Adaptive Plastic Scalable Electronics. IBM, Hughes Research Labs, HP, ведущие университеты: Колумбийский, Корнельский и др. (США). Источник финансирования - DARPA (Defense Advanced Research Projects Agency), IBM. Объем финансирования 2009-2016: \$ 102 млн. IBM представила SyNapse — чип, который работает по тому же принципу, что и мозг человека. SyNapse стал возможным благодаря достижениям в области компьютерной электроники и неврологии. SyNapse — это попытка создать «полноценный искусственный интеллект», осуществляя имитацию в одном чипе одновременную работу левого и правого полушарий человеческого мозга. Другими словами, IBM

пытается создать компьютер, который думает в точности, как человек [3].

Этот подход к созданию полноценного искусственного интеллекта перекликается с нашим проектом разработки электронного мозга интеллектуальных роботов (**ЭМИР**), систем управления и вычислительных систем с однородной, многосвязной, многомерной, нейроподобной, активной, ассоциативной, растущей структурой (ОММААНРС). В проекте решение этой проблемы осуществляется за счет того, что создан новый тип нейронных сетей, которые не имеют аналога. Создана теория и методология разработки систем искусственного интеллекта, проведено моделирование и тестирование виртуального интеллектуального робота. Остаточная стоимость разработки составляет примерно \$ 50 000 на 1 год. Низкая стоимость реализации проекта объясняется высокой однородностью элементов и простой матричной структурой анализа, синтеза и накопления (запоминания) информации.

---

### **Проект ЭМИР**

---

В основу проекта положен новый тип нейронных сетей и теория искусственного интеллекта (ТИИ). ТИИ включает учение о нейроподобных элементах и многомерных нейроподобных растущих сетях, временной и долговременной памяти, учение о функциональной организации «мозга» систем с искусственным интеллектом, о сенсорной системе, модулирующей системе, моторной системе, условном и безусловном рефлексах, рефлексорной дуге, мотивации, целенаправленном поведении, о «мышлении», «сознании», «подсознании и искусственной личности, формируемой в результате обучения и воспитания». В основе теории искусственного интеллекта лежит аналогия с нервной системой человека [4,5].

Теория искусственного интеллекта включает *новую технологию разработки электронного мозга интеллектуальных систем и роботов*. Технология создана на базе *однородной, активной, многомерной, рецепторно-эффекторной нейроподобной растущей сети*, которая аналогична естественным нейронным сетям мозга человека. Технология предназначена для создания электронного мозга роботов различного применения и роботов андроидов с интеллектом равным интеллекту человека, интеллектуальных вычислительных систем с новой ассоциативной не фон-неймановской архитектурой. В перспективе получение искусственного интеллекта превышающего интеллект человека.

Новая технология позволяет объединить физический и виртуальный миры в интеллектуальных системах. Такой подход дает основание для развития и массового производства продвинутых живучих роботов интеллектуальных систем управления и вычислительных систем с однородной, многосвязной, многомерной, активной, ассоциативной, нейроподобной, растущей структурой, выполняющей операции в режиме массового параллелизма подобно работе мозга человека.

---

### **Цель проекта ЭМИР**

---

*Цель проекта:* Создание электронного мозга для интеллектуальных систем и роботов, которые могли бы понимать человека и полностью заменять его в рутинном труде или в опасных местах.

Электронный мозг может применяться в любой сфере робототехники, от игрушек до сложных мобильных объектов. Или просто быть виртуальным думающим помощником внутри обычного компьютера или самостоятельной интеллектуальной системой распознавания образов, системой поиска информации в сверхбольших информационных базах и пр. Электронный мозг потенциально в составе робота может решать все задачи, необходимые для робота-гуманоида, а именно: учиться на опыте, обнаружить закономерности и разрабатывать гипотезы, а также ориентироваться в пространстве; «понимать» человеческую речь; выполнять действия по просьбе пользователя; ходить, бегать, прыгать и т.д.

---

### **Главные элементы концепции проекта ЭМИР**

---

*Главные элементы концепции:* двоичное кодирование информации; управление действиями и приемом информации - обучение (отсутствие программного управления); принцип хранения информации -

активная, нейроподобная, ассоциативная память (аналог мозг человека); принцип параллельной организации обработки информации, согласно которому операции над информацией проводятся по всему объему памяти одновременно (аналог мозг человека).

### Функциональная схема электронного мозга роботов

В работах физиологов А.Р. Лурия, Е.Н. Соколова [6,7] и др. *с позиции системной организации функций деятельности мозга* выделяют три основных функциональных блока: блок приема и переработки сенсорной информации – сенсорные системы (анализаторы); блок модуляции, активации нервной системы – модулирующие системы (лимбико-ретикулярные системы) мозга; блок программирования, запуска и контроля поведенческих актов – моторные системы (двигательный анализатор).

*Система формирования искусственного интеллекта – «мозг» системы* представляет собой активную, ассоциативную, однородную структуру – многомерную рецепторно-эффекторную нейроподобную растущую сеть, состоящую из множества нейроподобных элементов, связанных синаптическими связями. Нейроподобные элементы воспринимают, анализируют, синтезируют и сохраняют информацию, позволяют системе познавать, обучаться, мыслить логически, систематизировать и классифицировать информацию, находить в ней связи, закономерности, отличия и вырабатывать сигналы управления внешними устройствами.

**Функциональная схема электронного мозга роботов** (рис.1) представляет собой ассоциативную, активную, матричную среду. По аналогии с функциональной организацией деятельности мозга состоит из: устройств восприятия информации (1); набора однородных, многосвязных, многомерных, ассоциативных, активных, нейроподобных (ОММААН) сенсорных матриц временной и долговременной памяти (2); набор ОММААН модулирующих матриц (3); набора ОММААН моторных матриц памяти действий (4); исполнительные механизмы (5).

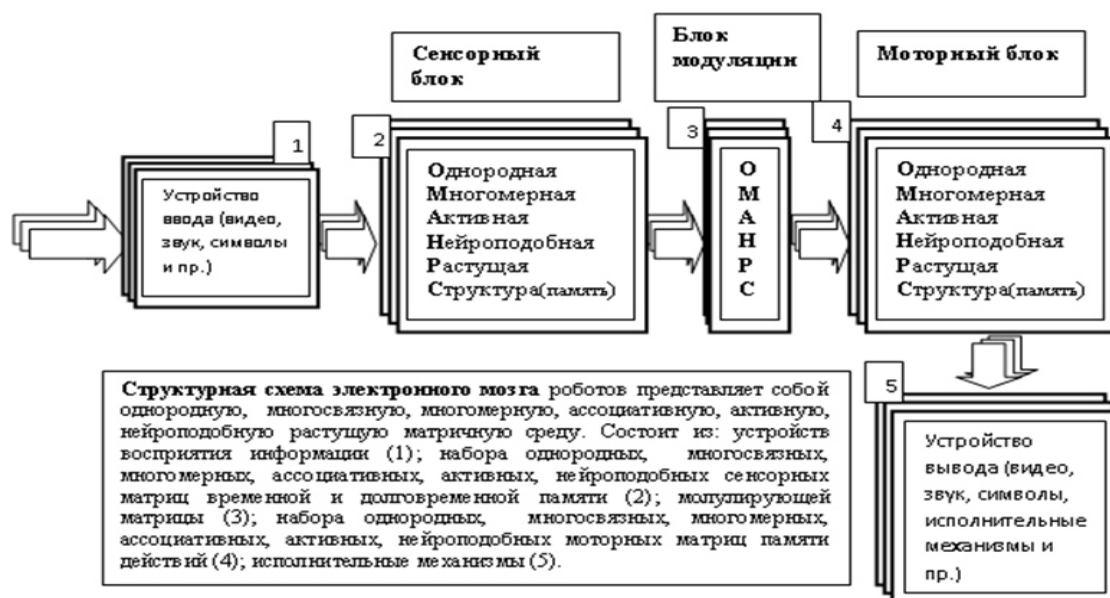
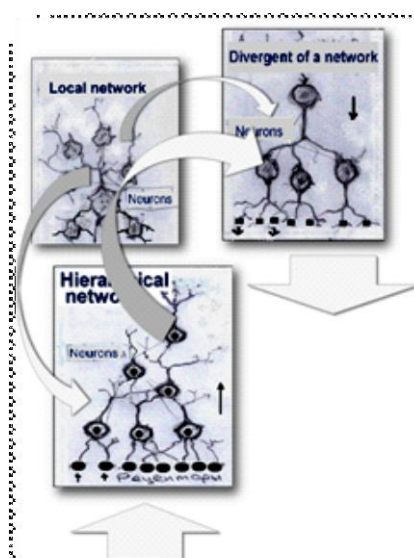


Рис.1. Функциональная схема организации «мозга» систем с искусственным интеллектом

*Сенсорные системы (анализаторы) мозга (2).* В сенсорной системе в каждом рецепторе, воздействующий физический фактор (свет, звук, тепло, давление), преобразуется в потенциал действия. Анализатор – это многоуровневая система с иерархическим принципом организации. Взаимоотношения между последовательными уровнями анализаторов построены по принципу «дивергенции –

конвергенции». *Модулирующие системы мозга (3)* являются аппаратом, выполняющим роль регулятора уровня бодрствования, также осуществляющим избирательную модуляцию и актуализацию приоритета той или иной функции. *Моторные (двигательные) системы мозга (4)*. Для двигательных систем характерна окончательная трансформация афферентных влияний, направленная на быстрейший выход афферентных возбуждений на периферию, т.е. на аппараты реализации конечной стадии поведения.

## Выводы



**Рис.2. Восходящая (ассоциативная), локальная (модулирующая) и нисходящая (моторная) нейроподобная растущая сеть**

От известных технологий предлагаемая технология отличается новым типом нейросети (рис.2) и новой нетрадиционной не фон-неймановской архитектурой системы. Преимущества предлагаемой технологии - абсолютная однородность структуры, массовый параллелизм, малое потребление энергии, сверхбыстродействие. Преимущества конечного продукта, производимого по предлагаемой технологии в сравнении с продуктами, имеющимися на рынке сегодня: Отсутствие программирования - система решает поставленные задачи с помощью обучения или самообучения, получая и синтезируя новые знания. Сверхбыстродействие восприятия и обработки информации за счет одновременного выполнения операций по всему объему активной структуры (памяти). Использование конечного продукта, разработанного на основе предлагаемой технологии, позволит заменить людей на работах, требующих напряженной интеллектуальной деятельности (напр., разработка новой техники и проектирование различных объектов, аналитическая деятельность, принятие решений), в управлении сверхсложными системами (атомные электростанции, авиатранспорт и т.д.).

## Библиография

1. <http://poslezavtra.org.ua/?p=263>
2. <https://mail.google.com/mail/u/0/#inbox/14813e82d21c8222>
3. <http://celest.bu.edu/outreach-and-impacts/the-synapse-project>
4. Yashchenko V.A. Receptor-effector neural-like growing network – an efficient tool for building intelligence systems / V.A. Yashchenko // Proc. of the second international conference on information fusion, (California, July 6–8 1999). – Sunnyvale Hilton Inn, Sunnyvale, California, USA, 1999. –Vol. II. – P. 1113 – 1118.
5. Яценко В.А. Искусственный интеллект. Теория. Моделирование. Применение. – К. Логос.2013 –289с. – Библ. с.283-289
6. Соколов Е.Н. Принцип векторного кодирования в психофизиологии / Е.Н. Соколов // Вестник Московского университета. Серия 14: Психология. – 1995. – № 4. – С. 3 – 13
7. Лурия А.Р. Основы нейропсихологии / Лурия А.Р. – М., 1973. – 173с.

## Сведения об авторе

**Виталий Яценко** – Ст. научн. сотрудник Института математических машин и систем НАН Украины. e-mail: [vitaly.yashchenko@gmail.com](mailto:vitaly.yashchenko@gmail.com)

## A FORMAL SEMANTICS OF ADDITIONAL OPERATIONS OF MULTISET TABLE ALGEBRA

**Dmitriy Buy, Iryna Glushko**

**Abstract:** Signature of previously known multiset table algebra is filled up with new operations such as inner and outer joins, semi join and aggregate operations (Sum, Min, Max, Count, Avg). A formal mathematical semantics of these operations is defined. The aim is to expand of possibility of relational databases due to use of multisets (bags).

**Keywords:** relational databases, multisets, multiset table algebra.

### Introduction

Nowadays the relational data model is widespread in use both in database scientific research and in practice. In its formal definition, originally proposed by E. Codd [Codd, 1970], the relational model is based on sets of tuples, i.e. it does not allow duplicate tuples in a relation. But many database languages and systems require a relational data model with multi-set semantics. There are two major reasons for this. In the first place, relations allowing duplicate tuples are useful in many application domains where duplicate entities can exist. For example, these are sociological polls of different population groups, calculations on DNA and others. In the second place, in the relational data model process of removing duplicates after implementation of projection (or union) operation assumes merging of identical elements or realizing some other labour-intensive actions. Naturally there is a need to expand of possibility of relational databases due to use of multisets. This problems were also considered in the works of Paul W.P.J. Grefen and Rolf A. de By [Grefen, de By, 1994], G. Lamperti, M. Melchiori, M. Zanella [Lamperti, Melchiori, Zanella, 2001], H. Garcia-Molina, J.D. Ullman, J. Widom [Garcia-Molina, Ullman, Widom, 2009], V. Redko, J. Brona, D. Buy, S. Polyakov [Redko, Brona, Buy, Poliakov, 2001]. However this question requires specification and extension as in one of the specified works the due attention isn't paid to operations of inner and outer joins, semijoin and aggregate operations of multiset table algebra.

### Multiset table algebra

Let's consider following sets:  $A$  is set of attributes and  $D$  is the universal domain. An arbitrary (finite) set of attributes  $R \subseteq A$  is named scheme. The tuple of scheme  $R$  is the nominal set on pair  $R, D$ . The projection of this nominal set for the first component equals  $R$ . The table is pair  $\langle \psi, R \rangle$ , where  $\psi$  is an arbitrary multiset ( $\Theta(\psi)$  is a basis of this multiset – an arbitrary set of tuples of the same scheme  $R$ ), and  $R$  is a scheme of table. Thus, a certain scheme is ascribed to every table. The multiset table algebra is an algebra  $\langle \Psi, \Omega_{P, \Xi} \rangle$ ,

where  $\Psi$  is a set of all tables,  $\Omega_{P, \Xi} = \left\{ \bigcup_{All}^{\Psi, R}, \bigcap_{All}^{\Psi, R}, \setminus_{All}^{\Psi, R}, \sigma_{p, R}, \pi_{X, R}, \otimes_{R_1, R_2}, Rt_{\xi, R}, \sim_R \right\}_{X, R, R_1, R_2 \subseteq A}^{p \in P, \xi \in \Xi}$  is signature, and

$P, \Xi$  are sets of parameters.

The operations of signature are defined in [Glushko, 2011]. These are multiset union, multiset intersection, multiset difference and also special operations (selection, projection, join, active complement and renaming). For each operation a basis of the resulting table and number of duplicates of every tuple are defined. It's shown how some operations can be expressed from other (main operations).

### Additional Operations Of Multiset Table Algebra

Signature of multiset table algebra is filled up with new operations such as inner and outer joins, semi join and aggregate operations (Sum, Min, Max, Count, Avg). A formal mathematical semantics of these operations is



defined. There are four kinds of inner joins: Cartesian Join  $Cj_{R_1, R_2}$ , Inner Natural Join  $\otimes_{R_1, R_2}$ , Inner Join using attributes  $A_1, \dots, A_n$   $\otimes_{A_1, \dots, A_n, R_1, R_2}$  and Inner Join on predicate  $p$   $\otimes_{p, R_1, R_2}$ . Moreover the following inclusions hold:

$$\langle \psi_1, R_1 \rangle_{Cj_{R_1, R_2}} \langle \psi_2, R_2 \rangle = \langle \psi_1, R_1 \rangle_{\otimes_{R_1, R_2}} \langle \psi_2, R_2 \rangle, \langle \psi_1, R_1 \rangle_{\otimes_{A_1, \dots, A_n, R_1, R_2}} \langle \psi_2, R_2 \rangle = \langle \psi_1, R_1 \rangle_{\otimes_{R_1, R_2}} \langle \psi_2, R_2 \rangle,$$

$$\left( \langle \psi_1, R_1 \rangle_{\otimes_{p, R_1, R_2}} \langle \psi_2, R_2 \rangle \right)_1 \preceq \left( \langle \psi_1, R_1 \rangle_{\otimes_{R_1, R_2}} \langle \psi_2, R_2 \rangle \right)_1, \text{ where } \preceq \text{ is a binary relation of inclusion over multisets}$$

[5]. The values of the operations in the left parts of these two equalities and inclusion must be defined. Notation  $\left( \langle \psi, R \rangle \right)_1$  denotes first component of pair  $\langle \psi, R \rangle$ , e.i. multiset  $\psi$ .

There are four kinds of outer joins: Outer Left Join, Outer Right Join, Outer Full Join and Outer Union Join. The special element NULL is inserted in the universal domain for to define outer joins. We use the logical scheme from [Redko, Brona, Buy, Poliakov, 2001] for definition outer joins. Four kinds of outer joins are defined as subordinated to one of inner joins operation.

Let's describe the scheme of define the aggregate operations. The functions of summation, taking of the minimum and maximal values, definitions of an arithmetic average and quantity of elements are defined on a finite multiset. Then these functions are transferred to the tables. It should also be noted that the parameter of aggregate operation can be not only a separate attribute, but also a function of the tuple.

---

## Conclusion

A formal mathematical semantics of the advanced operations of multiset table algebra is defined. The advanced operations include inner and outer join, semijoin and aggregate operations (Sum, Min, Max, Count, Avg). The aim is to expand of possibility of relational databases due to use of multisets (bags).

---

## Bibliography

- [Codd, 1970] E.F. Codd. A Relational Model for Large Shared Data Banks. In: Communications of the ACM. 13(6), 1970.
- [Garcia-Molina, Ullman, Widom, 2009] H. Garcia-Molina, J.D. Ullman, J. Widom. Database Systems: The Complete Book. Prentice Hall, Upper Saddle River, New Jersey, 2009.
- [Glushko, 2011] I.M. Glushko. Multiset Table Algebra. In: International Scientific Conference of Student and Young Scientists "Theoretical and Applied Aspects of Cybernetics", Kyiv, Ukraine, 2011.
- [Grefen, de By, 1994] Paul W.P.J. Grefen, Rolf A. de By. A Multi-Set Extended Relational Algebra. A Formal Approach to a Practical Issue. In: 10th International Conference on Data Engineering. Houston, TX, USA, 1994.
- [Lamperti, Melchiori, Zanella, 2001] G. Lamperti, M. Melchiori, M. Zanella. On Multisets in Database Systems. In: Multiset Processing: Mathematical, Computer Science, and Molecular Computing Points of View, number 2235 in Lecture Notes in Computing Science, 2001.
- [Redko, Brona, Buy, Poliakov, 2001] V. Redko, J. Brona, D. Buy, S. Poliakov. Relation database: relation algebras and SQL-similar languages, 2001.

---

## Authors' Information

**Dmitriy Buy** – Dsc., Physic and Mathematic Sciences, Full Professor, Taras Shevchenko National University of Kyiv, Academician Glushkov Avenue 4d, 03680 Kyiv, Ukraine; e-mail: [buy@unicyb.kiev.ua](mailto:buy@unicyb.kiev.ua).

**Iryna Glushko** – Ph.D., Physic and Mathematic Sciences, Senior Lecturer, Nizhyn Gogol State University, str. Kropyvianskoho 2, 16600 Nizhyn, Ukraine; e-mail: [glushkoim@gmail.com](mailto:glushkoim@gmail.com).

## GENERAL APPROACH FOR THE CLASSIFICATION PROBLEM SOLUTION

**Mykola Budnyk**

**Abstract:** A theoretic approach for classification problem is presented which include correct problem statement, definition of under-, over- and well-determined cases, features space dimension, complexity (multi-valuing) of decision-making rules, and number of learning groups. The principal condition of exact solution of the well-determined problem and relationships between feature space, decision rule and learning groups are obtained.

**Keywords:** classification, discrimination, feature space, multi-value decision-making rule.

### Introduction

Many researches are solving problem for distinguishing between different groups of objects, i.e. classification problem (CP). Importance of questions of features selection, syntheses of decision-making rule, building of robust, high-accuracy, and reliable classifier are clearly evident [Budnyk, 2006; Budnyk 2009; Liu, 2014]. That is why, in my opinion, some general approach should be developed in order to give researchers working in different areas by some unified viewpoint on CP solution. This paper is aimed to compare the performance of various CPs, and to develop a general view to achieve exact or general CP solution.

### What is well-defined general classification problem?

General CP is defined as problem in which quantity of classes  $Q_{CLASS}$  is unknown. Here for classification an analogy from system of linear algebraic equations is used so that 3 cases should be considered:

- a) Over-determined CP in which cells of feature space (FS) are more than classes. In this situation the CP can be solved but solution is non-unique and many variants may be obtained. In each solution some classes may cover more than one respective cell. In "bad" limit some cell cannot be interpreted, i.e. it is obtained undefined cell in which diagnosis is unknown. In result, it is impossible to obtain correct decision rule, that is why above solution is called as quasi-solution of CP.

$$Q_{CELL} > Q_{CLASS} \quad (1)$$

- b) Well-determined CP in which number of cells is equal number of classes. In this situation the CP can be solved in unique manner so that each cell relates only single class. In result, it is possible to obtain unique rule, that is why above solution is called as correct solution of CP.

$$Q_{CELL} = Q_{CLASS} \quad (2)$$

- c) Under-determined CP in which classes are more than cells. In this case the CP cannot be solved in narrow sense because we have deficit of cells. In wide sense solution is existed but non-unique and many variants may be obtained in which some cell may include more than one respective class, i.e. diagnosis. In result, it is impossible to obtain correct rule, that is why above solution is called as pseudo-solution of CP.

$$Q_{CELL} < Q_{CLASS} \quad (3)$$

Below we assume that well-defined in narrow or wide sense CP is well-determined CP according to (2) or (1-2).

### Solution of well-defined general classification problem at homogeneous space

It is well-known that number of subsets of set having  $n$  elements is equal to  $G=2^n$ . Here set is population involved in study consisting from  $N$  (minimum – 2) groups of objects. From the other hand, maximal quantity of classes  $Q_{CLASS}$  of set formed by  $N$  groups, which are intersected with each other, is also equal (4a)

$$Q_{CLASS} = 2^N, \quad Q_{CELL} = M^D \quad (4a,b)$$

Feature is defined as simple parameter, general parameter (e.g., additive or multiplicative its function) or combination some simple and general parameters. Number of cell of homogeneous FS  $Q_{CELL}$  is equal (4b) where  $D$  is space dimension and  $M$  is number of grades for each feature.

If number of classes is known one can obtain such 2 kinds of particular CP:

- 1) If number of classes is equal to (4) we have per se CP.
- 2) So-called discrimination problem (DP) is obtained if uncertainty class is removed from consideration so that number of classes is equal

$$Q_{DISCR} = Q_{CLASS} - 1 = 2^N - 1. \quad (5)$$

The CP is inverse problem, which is known as non-correct problem. That's why the CP solution is inverse problem solution, which in overall case is non-unique, i.e. includes many particular solutions. The CP and DP solution are well-defined problems in wide sense (so that full solution is achievable in principle) if such problem is well-defined in wide sense (1-2) and FS is homogeneous (4). Eq. (6) is main equation for CP and DP solution in wide sense including over-determined and well-determined cases according to (1) and (2).

$$M^D \geq Q_{CLASS}, \quad M^D \geq Q_{DISCR}. \quad (6)$$

---

### Solution of the well-defined classification problem

---

Combining (2) and (4) it will be obtained general condition for exact CP solution

$$M^D = 2^N. \quad (7)$$

Eq. (10) is main equation of the CP solution which can be considered as condition on FS dimension  $D$  if number of feature grades  $M$  is known (8a) or vice versa for calculating number of grades  $M$  if dimension  $D$  is known (8b)

$$D = \log_M 2^N = \frac{N}{\log_2 M}, \quad M = 2^{N/D}. \quad (8a,b)$$

Two limiting cases of the feature space (FS):

- 1) Binary space ( $M = 2$ ) - condition to the FS dimension is obtained from (8a)

$$D = N. \quad (9)$$

- 2) 1D space ( $D = 1$ ) - condition to the feature grades is also obtained from (8b)

$$M = 2^N. \quad (10)$$

The simplest problem (SP) of classification is defined if only 2 groups are involved into consideration  $N = 2$ . In result, from (4a) one can obtain 4 classes  $Q_{CLASS} = 4$ . Concerning to solution of SP at FS from (8) we can obtain such cases for various dimensions:

- a) 1D space  $D = 1$ , if number of grade is equal to  $M = 4$ ;
- b) 2D space  $D = 2$ , if number of grade is equal to  $M = 2$ ;
- c) 3D and higher spaces at the lowest number of grades  $M = 2$ , i.e. binary spaces have too many cells (8 and more) that lead to the over-determined problem.

So, solution into 1D and 2D FS space is achieved, but for 3D and higher spaces problem is over-determined and solution is non-unique. Figure 1a presents Euler's diagram reflecting principle of solution of the classification SP onto 4 classes into theoretic-set space and interpretation of classes. Figure 1b and 1c present general decision rule, i.e. proposed interpretation of cells into term of classes at 1D (Fig. 1b) and 2D (Fig. 1c) feature spaces.

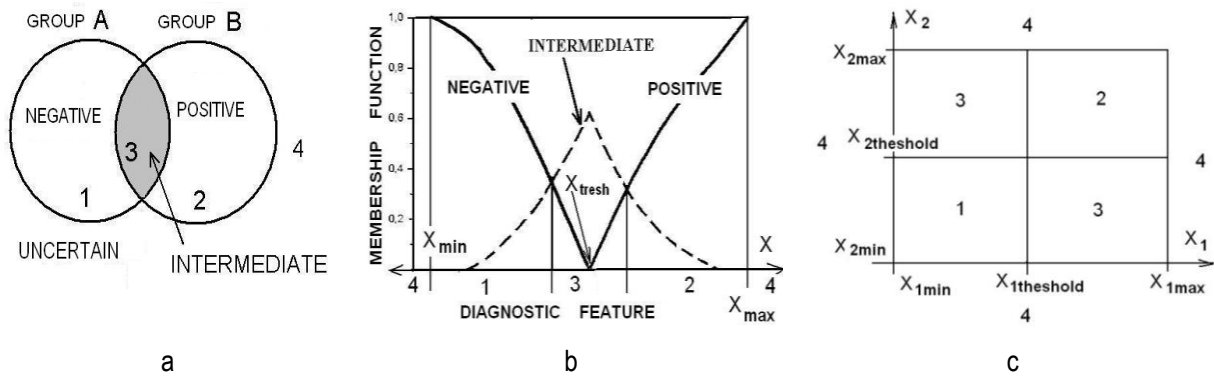


Figure 1. Solution of the SP of classification and discrimination (without uncertain class) at theoretic (a), 1D (b) and 2D (c) feature space: 1- negative class, 2 – positive class, 3- intermediate class, 4 – uncertain class.

### Solution of well-determined discrimination problem

Combining Eq. (5-6) it will be obtained main equation of the DP exact solution as condition

$$M^D = 2^N - 1. \quad (11)$$

Analogously to CP at above section, from (11) FS dimension  $D$  and number of grades  $M$  can be obtained

$$D = \log_M [2^N - 1], \quad M = \sqrt[D]{2^N - 1}. \quad (12a,b)$$

At special case of binary ( $M = 2$ ) and 1D ( $D = 1$ ) feature space we obtain from (12)

$$D = N + \log_2 [1 - 2^{-N}], \quad M = 2^N - 1. \quad (13a,b)$$

According to (13a) for many groups  $N \gg 1$  quantitative difference between discrimination and classification is neglected, but in fact threshold is equal to  $N = 3$  corresponding with 7 classes.

The simplest problem (SP) of discrimination: from (5) at  $N = 2$  one can obtain 3 classes  $Q_{CLASS} = 3$ , relative expressions for FS dimension  $D$  and number of grades  $M$  are obtained from (12)

$$D = \log_M 3, \quad M = \sqrt[3]{3}. \quad (14a,b)$$

Since  $M$  and  $D$  can be only integer number, Eqs. (14) has only 2 solutions:

- 1)  $D = 1$  if  $M = 3$ , according to (14a) – well-determined case (exact solution);
- 2)  $M = \text{ceil}(\sqrt[3]{3}) = 2$  if  $D = 2$ , according to (14b) – over-determined case.

Figure 1 relates to DP also if uncertain class 4 is omitted.

### Exact solution of well-defined CP in narrow sense

All above expressions include integer number  $N$ ,  $D$  and  $M$  and should be understanding as Diophantine-like equations. Solution in narrow sense assumes that only well-determined cases are involved according to (2). It is obvious that Eq. (7) in integer numbers has exact solution if next condition is only take place

$$M = 2^L, \quad (15)$$

where  $L$  is order of feature gradation. Substituting (15) to (7), condition for solution is found as

$$N = D * L, \quad L = \log_2 M. \quad (16)$$

Results about matching between  $N$ ,  $D$ , and  $L$  or  $M$  in order to achieve the exact CP solution are summarized below at Table 3-4 describing order of gradation  $L$ , if groups studied  $N$  and FS dimension  $D$  are known (Table 3), and describing the FS dimension  $D$  if number of groups  $N$  and grades  $M$  are known.

Table 3. Order of gradation  $L$ , in which CP is exactly solved, from groups studied  $N$  and FS dimension  $D$

D	Number of groups, N							
	2	3	4	5	6	7	8	
1	2	3	4	5	6	7	8	
2	1		2		3		4	
3		1			2			
4			1				2	
5				1				
6					1			
7						1		
8							1	

Table 4. Dimension of FS  $D$ , i.e. number of features needed, in which CP is exactly solved, depending from number of groups  $N$  and number of grades  $M$

M	Number of groups, N							
	2	3	4	5	6	7	8	
2	2	3	4	5	6	7	8	
4	1		2		3		4	
8		1			2			
16			1				2	

## Conclusion

Binary FS  $M = 2$  is the best classification space because exact solution is always at  $D = N$ , i.e. if space dimension is equal to number of groups. For 1D space synthesis of multi-value feature gradation is not clear, because need many learning groups, and for 2 groups at  $M > 4$  is under-determined problem. For more detail gradation, existence of exact CP solution is too rare events among all possible combinations of group quantity and space gradations. Therefore, proposed approach should be developed for: 1) over-determined case, 2) discrimination problem, 3) non-homogeneous feature space. Also, for differential CP analysis on "solubility" should be made. In this connection, developed previously [Budnyk, 2012] method for creating 3-8 gradation at  $N = 2$  is promising.

It should be noted that the well-known binary test (dividing 2 groups onto 2 classes by threshold rule at 1D space) is under-determined case for CP solution. Threshold rule is appropriate for solving the SP of classification or discrimination at only 2D space [Budnyk, 2012] and, strictly speaking, can be recommended only for feature selection. For application it is attractive advanced method proposed early [Budnyk, 2007; Budnyk 2008].

## Bibliography

- [Budnyk, 2006] Budnyk M., Voytovych I. Finding of informative parameters describing biomedical populations, Proc. 12<sup>th</sup> Int. Conf. "Knowledge-Dialogue-Solution", Varna (Bulgaria). – 2006. – p. 334-340.
- [Budnyk, 2009] Budnyk M. Development of biomedical informatic-measurement systems based on SQUID-magnetometers and technology for their application, DrSc Dissertation, Glushkov Institute of Cybernetics, Kyiv (Ukraine), 2009. – 404 p.
- [Liu, 2014] Liu F., Gramatikov B., Budnyk M., Ng E.Y.K. The System and Technology on Biological Information // Journal of Medical Imaging and Health Informatics. – 2014. – Vol.4, No.2. – p. 303-305.
- [Budnyk, 2012] Budnyk M., Zakorchenyi O. Development of gradation scale for 4-8-value decision rules and their interpretation for biomedical applications // Proc. Seminar "Biological & Medical Informatics and Cybernetics" (BMIC-2012), Glushkov Institute of Cybernetics, Kyiv (Ukraine). – 2012. – part 2. – p. 161-168.
- [Budnyk, 2012] Budnyk M., Zakorchenyi O. Method for determination of intermediate state of object, Patent UA 98838, 2012.
- [Budnyk, 2007] Budnyk M. Classification of groups based on normalized distribution functions at medical informatics // Control systems and Machines (USIM). – 2007. – No.3. – p. 57-64.
- [Budnyk, 2008] Budnyk M., Zakorchenyi O. Method for classification of patient group, Patent UA 84884, 2008.

## Authors' Information

**Mykola Budnyk** – Leading Researcher, Doctor of Computer Science, Glushkov Institute for Cybernetics of NAS of Ukraine, Glushkov ave. 40, P.O. Box: 03680, Kyiv-187, Ukraine; e-mail: budnyk@meta.ua

---

## ONTOLOGY DRIVEN CONTEXT BASED DECISION-MAKING

Yuriy Chaplinsky, Olena Subbotina

**Abstract:** Presented approach of support decision-making is based on usage of ontology and context. These realize activities and give possibilities of support of functional integration of decision-making processes; interdisciplinary integration of applied & scientific information and knowledge; access to the information and applications, outgoing from needs and competence of the user; reliable, scalable and safe architecture; integration of resources.

**Keywords:** decision-making, knowledge management, context, ontology

---

### Introduction

Comprehensive and systematic support of decision-making is the dominant dynamic business environment. Typical for humans is the ability to use only those features of reality that is important for a particular situation or decision-making problem. Therefore it is important to support the decision-making task in problem situations using decision support systems (DSS), which developed on the principles of knowledge engineering for a set of specific subject areas. The usage of the knowledge-oriented technologies for decision-making is one of basic conceptions and facilities of modern influence on efficiency and rationality of functioning of different management systems. Modern realization of such knowledge-oriented technologies is based on the organization of integrated information environment for the effective use of experience of specialists and modern technologies of management, analysis, modeling and design, decision-making, realization of advice and business processes, modern technologies of delivery of information and knowledge. In this way one of the effective tools of supporting of decision-making is ontology. The ontology is a system which describes the structure of a problem area, and consists of a set of classes of concepts, which are related by relations, their definitions and axioms, which define constraints on the interpretation of these concepts within the applied area or set of the applied areas [Staab et al, 2001]. Ontology give possibility consideration of factors, which can include the following: generic view: it provides the general understanding of the nature of the processes; contents view: it reveals the contents of the processes; contexts view: it reveals relevant process description; presentation view: the process is seen as a set of expressions presented in some language(s); physical view: it reveals the appearance(s) of the processes; structural view: it is defined as a modular structure of various parts.

---

### Ontology & Context for decision-making

As stated above ontologies should support multi-disciplinary advice/decision-making, assist formulate problem and select appropriated models using data, user information and problem solution requirements etc., select appropriate solving method and algorithm, choose solution and analyze the solution implementation.

Taking into account given requirements, we define the ontology of decision-making as the interrelated set of ontologies, which have an associative multi-level structure: meta-ontology, basic ontology, contextual ontology, multi-layer specified ontologies, realization ontology, and user presentation and interaction ontology.

These ontologies consider area of decision-making as a multi-level structure, which includes problems, models, methods and realizations. The description of decision-making defines the concepts and structures that describe the nature, structure and presentation of the process of decision-making and appropriated areas, which describe such a process. Levels interacted via relations: results and restrictions. The problem is a task, which is characterized significance, necessity, sufficient of the content, multiplicity of possible solving ways and variance results. Problem is based on a semantic basis, and defines the requirements for the development of the model. Models use a system of concepts and formulated to represent a problem situation or task to some specific language. The models are implemented and used in methods. Methods define the decision-making processes on the basis of model building and provide solving direction. Methods in the wide sense are represented through schemes and scenarios, methods in the narrow sense (defined methods) and the appropriated algorithms. Realization defines how and by what tools, in what environment models or methods can be implemented in the

DSS. This implementation consist three software based layers: the data source layer, the logic layer and the user interface layer. Integration on the data source layer provides a unified view on heterogeneous data sources. Integration on the logic layer unifies different implementations of logic of advice and decision making, each using its own data sources, under a common user interface. Integration on the user interface layer unifies different user interfaces in one common system.

Context is any information that can be used and describes the relevant problem area. [Dey et al, 2001]. Context system [Leppänen, 2007] assists to identify, understand, and provide the relevant elements of advice and decision-making both the contexts and within contexts. We define the context as an environment in which the task is performed. Context provides relevant, actual and accessible information to solve the problem or to understand the current situation of solving the problem. Context management approach should provide mechanisms that allow to process contextual information, as appropriate for the specific conditions for the defined purposes. We use three types of context: abstract, concrete and context of realization. Abstract context is an ontological model of decision-making tasks, which built on the basis of the integration of domain knowledge, which is relevant for this problem. Concrete context is an abstract specification of context in actual conditions by the data, which is derived from information resources. Context of the specification is a concrete context for real minds and taking into account the requirements of the hardware platform, software, and user competence.

---

## Conclusion

Presented approach for decision-making was implemented within USAID-funded project "Improving income of private Ukrainian agricultural producers through agricultural extension" (Ukraine), Ukrainian-Indian project "Development of integrated consulting and decision-making – oriented environment for agricultural extension services".

---

## Bibliography

- [Dey et al, 2001] Dey A.K., Salber D., Abowd G.D. A Conceptual Framework and a Toolkit for Supporting the Rapid Prototyping of Context-Aware Applications Context-Aware Computing. INT J HUM-COMPUT INT 2001; 16: 97 - 166.
- [Leppänen, 2007] Leppänen M. Towards an Ontology for Information Systems Development - A Contextual Approach. In: K. Siau, editor. Contemporary Issues in Database Design and Information Systems Development. PA:IGI Global; 2007. p. 1-36.
- [Staab et al, 2001] Staab S., Studer R., Schnurr H.-P., Sure Y. Knowledge Processes and Ontologies // IEEE Intelligent Systems. – 2001. – V. 16. – № 1. – P. 26-34.

---

## Authors' Information

**Yuriy Chaplinskyy** – senior researcher, Ph.D., V.M. Glushkov Institute of Cybernetics of National academy of sciences of Ukraine, prospekt Academician Glushkova, 40, Kyiv, 03680, Ukraine; e-mail: [cyuriy60@hotmail.com](mailto:cyuriy60@hotmail.com)

**Olena Subbotina** – researcher, V.M. Glushkov Institute of Cybernetics of National academy of sciences of Ukraine, prospekt Academician Glushkova, 40, Kyiv, 03680, Ukraine; e-mail: [olenas2011@gmail.com](mailto:olenas2011@gmail.com)

## FRAMEWORK TO MANAGE SCRUM MEETING ARTIFACTS

Elena Chebanyuk

**Abstract:** A framework and analytical apparatus for software model comparison are represented in this paper. The framework implements the sequence of operations that are necessary to perform in order to manage both software artifacts history and changes in architectural solutions. Analytical apparatus of software model comparison is based on Set-theory tool operations.

**Keywords:** Scrum meeting, software model comparison, Set-theory tool, software architecture, UML diagram.

### Topicality and challenges

Software development process according to Agile methodology becomes widespread today. Iterative and incremental software development lifecycles propose possibility to change requirements in every development iteration, namely during scrum meetings. That is why the task of designing approaches and frameworks to manage software artifacts that are appeared (changed or removed) after scrum meetings is important.

Most software artifacts that are changed during scrum meeting are software models (SM) that are represented as UML diagrams [UML 2.5]. In order to propose an approach to consider SM changes it is necessary to solve the next tasks:

- Design an analytical representation of SM.
- Propose a format for storing information about SM.
- Provide a tool of scrum artifacts matching.

Papers [Chebanyuk, 2013], [Chebanyuk, 2014] propose solution of the first task. To represent software models, namely Class and Use Case Diagrams, it is proposed to use Cartesian products. Elements of Cartesian products are constituents of these software models.

In order to design formats for storing data about software models XMI and XML formats can be used.

This paper proposes a framework for matching scrum artifacts changes that are represented as Use Case Diagrams. Use Case diagrams and their varieties, namely user stories, are central artifacts that are changed during scrum meetings. Also software prototypes and other artifacts for description of software behavior can be converted to Use Case diagrams too.

### Analytical apparatus for the Use Case diagrams matching

Consider the analytical representation of Use Case diagrams that was proposed in paper [Chebanyuk, 2014].

$$\left\{ \begin{array}{l} \omega \subseteq P^{\omega} \times K^{\omega} \times A^{\omega} \times P^{\omega}(\text{include}) \times P^{\omega}(\text{extends}) \times Y^{\omega} \times T^{\omega} \\ P^{\omega} \times K^{\omega} \times A^{\omega} \times P^{\omega}(\text{include}) \times P^{\omega}(\text{extends}) = \{ \langle p_1, p_2 \rangle, \langle a_1, p_1 \rangle, \langle a_1, k_1, p_1 \rangle, \langle p_1, k_1, p_2 \rangle, \\ \langle p_1, \nu_1, p_2 \rangle, \langle p_2, \tau_1, p_4 \rangle, \langle p_1(\text{include}), p_3 \rangle, \langle p_1(\text{extends}), p_4 \rangle, \langle p_5, p_6 \bullet p_5, p_7 \rangle, \\ \langle \text{sign } p_2^{\text{sign}}, p_4^{\text{sign}} \rangle \} \end{array} \right. , \quad (1)$$

where  $p_i \in P^{\omega}$ ,  $i=1, \dots, 7$ ,  $k_1 \in K^{\omega}$ ,  $a_1 \in A^{\omega}$ ,  $p_3(\text{include}) \in P^{\omega}(\text{include})$ ,  $p_3(\text{extends}) \in P^{\omega}(\text{extends})$ ,  $\nu_1 \in Y^{\omega}$ ,  $\tau_1 \in T^{\omega}$ .

More detailed description is introduced in the paper [Chebanyuk, 2014].

A reference Use Case diagram is denoted as follows:

$$\omega \subseteq P^{\omega} \times K^{\omega} \times A^{\omega} \times P^{\omega}(\text{include}) \times P^{\omega}(\text{extends}) \times Y^{\omega} \times T^{\omega}$$



And Use Case diagram that was changed after scrum meeting is denoted as follows:

$$\omega' \subseteq P^{\omega'} \times K^{\omega'} \times A^{\omega'} \times P^{\omega'}(\text{include}) \times P^{\omega'}(\text{extends}) \times Y^{\omega'} \times T^{\omega'}$$

Consider main operations that are executed when Use Case Diagram are proceeded and introduce their analytical representation. At our case it is necessary to define changes in  $\omega'$  software model.

#### Model comparison

1. Defining newly appeared SM constituents after scrum meeting

$$n\_a = \omega' \setminus \omega \quad (2)$$

where  $n\_a$  - newly appeared SM constituents.

*The note: When operation (2) is performed corresponding sets of models  $\omega$  and  $\omega'$ , namely  $P^{\omega}, K^{\omega}, A^{\omega}, P^{\omega}(\text{include}), P^{\omega}(\text{extends}), Y^{\omega}, T^{\omega}$ , are subtracted. Also this approach is used when further operations are introduced*

2. Defining model constituents that disappeared after the last scrum meeting

$$d = \omega \setminus \omega' \quad (3)$$

where  $d$  - disappeared SM constituents.

2. Defining components of an architecture solution that shouldn't be changed

$$s\_c = \omega' \cap \omega \quad (4)$$

where  $s\_c$  - stable components of an architectural solution

---

### Framework to manage scrum artifacts

---

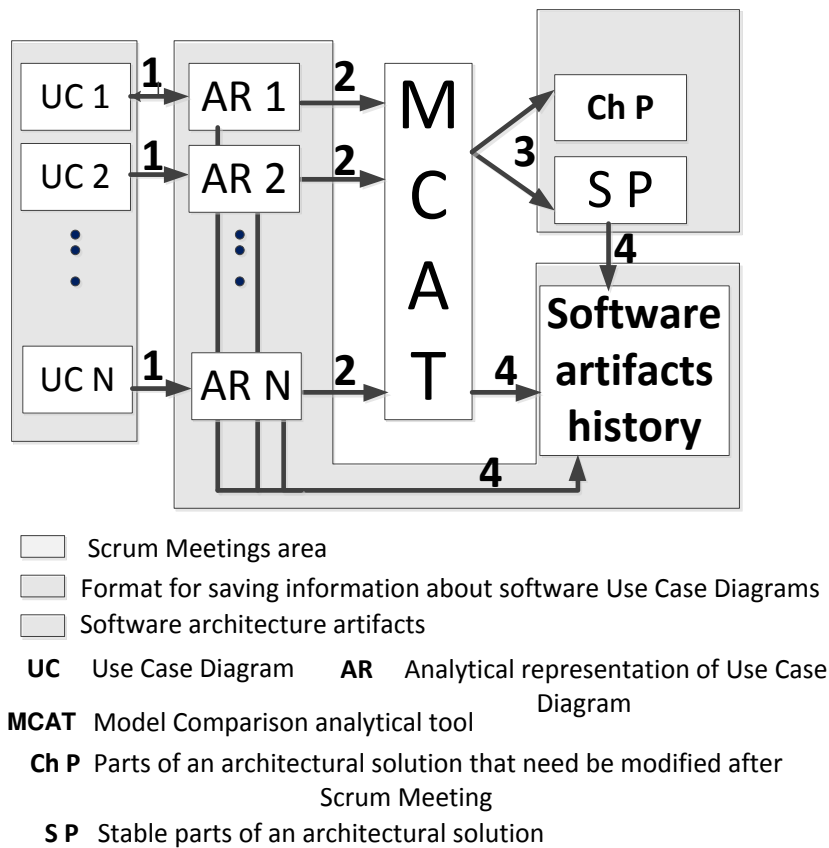
The process of scrum artifacts management contains from the following steps:

1. Obtaining updated Use Case (US Figure 1) diagrams after scrum meeting
2. Converting Use Case Diagrams to their analytical representation (AR Figure 1). In order to do this the format of file XMI or XML can be used.

Processing analytical representation of SM (MCAT Figure 1) using analytical apparatus for Use Case diagram matching, namely expressions (2)-(4).

3. Defining both parts of architectural solution that should be changed (Ch p Figure 1) and stay without modification (SP Figure 1) for its further effective processing.
4. Using an analytical representation software artifact history can be stored and managed.

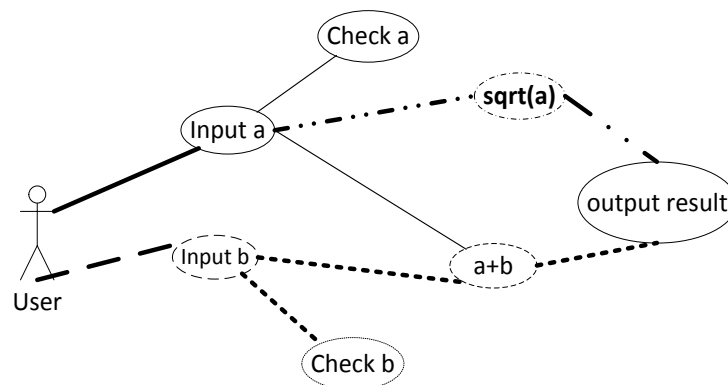
The numbers in the list described above, correspond to the activities in framework presented in Figure 1. Activities are marked as arrows.



**Figure 1.** Representation of the framework for scrum artifacts management

## Case tool

Consider a scrum artifacts history. Figure 2 represents a Use Case diagram, containing traces of two scrum meetings. Black colored parts of the Use Case Diagram are the parts that are stayed the same in reference and new diagram. Dashed lines and precedents of the Use Case diagram are the parts that are appeared after the last scrum meeting (Reference diagram was free from these parts). Dashed lines and precedents with dots of the Use Case diagrams are the parts that are disappeared after the last scrum meeting.



**Figure 2.** Use Case Diagram

Using an analytical tool for comparing Use Case diagrams define newly appeared and disappeared components of software architecture and also stable parts of an architectural solution.

1. Performing the same operations, as was described in [Chebanyuk, 2014] obtain an analytical representation of both Use Case Diagrams

## 1.1 Reference Use Case Diagram (Figure 2)

$$P^{\omega} = \{p_0^{\omega}, \dots, p_4^{\omega} \mid p_0^{\omega} = \text{input } a, p_1^{\omega} = \text{check } a, p_2^{\omega} = \text{sqrt}(a), p_4^{\omega} = \text{output rez}\}$$

$$A^{\omega} = \{a_0^{\omega} \mid a_0^{\omega} = \text{user}\}, P^{\omega}(\text{include}) = \{p_0^{\omega}(\text{include})p_1^{\omega}\}$$

$$\omega \subseteq \{< a_0^{\omega}, p_0^{\omega} >, < p_0^{\omega}, (\text{include})p_1^{\omega} >, < p_2^{\omega}, p_3^{\omega} >, < p_3^{\omega}, p_4^{\omega} >\}$$

## 1.2 Changed Use Case Diagram after the next scrum meeting (Figure 2)

$$P^{\omega'} = \{(P^{\omega} \setminus p_2^{\omega}) \cup \{p_5^{\omega'}, p_6^{\omega'}, p_7^{\omega'} \mid p_5^{\omega'} = \text{input } b, p_6^{\omega'} = \text{check } b, p_7^{\omega'} = a + b\}\}$$

$$A^{\omega'} = A^{\omega}, P^{\omega'}(\text{include}) = \{p_5^{\omega'}(\text{include})p_6^{\omega'}\}$$

$$\omega' \subseteq \{< a_0^{\omega'}, p_0^{\omega'} >, < p_0^{\omega'}, (\text{include})p_1^{\omega'} >, < p_5^{\omega'}, (\text{include})p_6^{\omega'} >, < p_5^{\omega'}, p_7^{\omega'} >, < p_7^{\omega'}, p_4^{\omega'} >\}$$

## 3. Defining newly appeared SM constituents after scrum meeting

$$n\_a = \omega' \setminus \omega = \{< p_5^{\omega'}, (\text{include})p_6^{\omega'} >, < p_5^{\omega'}, p_7^{\omega'} >, < p_7^{\omega'}, p_4^{\omega'} >\}$$

These components match of the Use Case diagram parts containing dashes. See Figure 2.

## 2. Defining model constituents that disappeared after the last scrum meeting

$$d = \omega \setminus \omega' = \{< p_2^{\omega}, p_3^{\omega} >, < p_3^{\omega}, p_4^{\omega} >\}$$

These components match of the Use Case diagram parts containing dashes and dots. See Figure 2.

## 4. Defining components of an architecture solution that shouldn't be changed

$$s\_c = \omega' \cap \omega = \{< a_0^{\omega'}, p_0^{\omega'} >, < p_0^{\omega'}, (\text{include})p_1^{\omega'} >\}$$

These components match with black colored parts of Use Case diagram. See Figure 2.

---

## Conclusion

The framework and analytical apparatus to manage scrum artifacts are represented in this paper. Proposed ideas are illustrated by means of case tool. Using both framework and analytical apparatus for processing Use Case diagrams allows designing effective tools for managing architectural solutions. Also it is possible to process software artifacts history. Effective solving of these tasks allows improving many activities such as code reuse, test cases designing, processing software prototypes, correcting algorithms for solving different task and others.

---

## Bibliography

[UML 2.5] <http://www.omg.org/spec/UML/2.5/Beta1/>

[Chebanyuk, 2013] E. Chebanyuk Algebra describing software static models. International Journal "Information Technologies & Knowledge" Vol.7, Number 1, 2013, P 83-93.

[Chebanyuk, 2014] E. Chebanyuk Method of behavioral software models synchronization. International journal "Information models & analyses" Volume 3, Number 2, 2014, P. 147-163.

---

## Authors' Information

**Elena Chebanyuk** – lecturer in National aviation university, associate professor of software engineering department, Ukraine; e-mail: [chebanyuk.elena@gmail.com](mailto:chebanyuk.elena@gmail.com)

Major Fields of Scientific Research: Model-Driven Architecture, Domain engineering, Code reuse.

## EXAMPLE OF MULTI-LAYER KNOWLEDGE REPRESENTATION BY MEANS OF NATURAL LANGUAGE ADDRESSING

**Krassimira Ivanova**

**Abstract:** An approach for knowledge representation based on post-relation type of information bases is outlined in the paper. The idea of Natural Language Addressing and based on it idea of Multi-layer Knowledge Representation are presented.

**Keywords:** Multi-layer Knowledge Representation; Natural Language Addressing

### Introduction

There are a lot of approaches for knowledge representation [Sowa, 2000]. In this work we show an example of knowledge representation by means of Natural Language Addressing (NLA) [Ivanova et al, 2012; Ivanova et al, 2013; Ivanova, 2014]. NLA is based on the Multi-domain Information Model [Markov, 1984; Markov, 2004]. In this paper we outline an approach for knowledge representation based on post-relation type of information bases.

We follow the proposition of Kr. Markov to use the computer encoding of name's (concept's) letters as logical address of connected to it information stored in a Multi-dimensional numbered information spaces [Markov, 2004; Markov, 2004a]. This way no indexes are needed and high speed direct access to the text elements is available. It is similar to the natural order addressing in a dictionary where no explicit index is used but the concept by itself locates the definition.

The idea of NL-addressing is to use encoding of the name both as relative address and as route in a Multi-dimensional information space and this way to speed the access to stored information. For instance, let have the next definition: "**Pirrin**: A mountain with co-ordinates (x, y)". In the computer memory, it may be stored in a file at relative address "50067328" and the index couple is: ("Pirrin", "50067328"). At the memory address "50067328" the main text, "A mountain ... (x,y)" will be stored. To read/write the main text, firstly we need to find name "Pirrin" in the index and after that to access memory address "50067328" to read/write the definition. If we assume that name "Pirrin" in the computer memory is encoded by six numbers (letter codes), for instance by using ASCII encoding system Pirrin is encoded as (80, 105, 114, 114, 105, 110), than we may use these codes for direct address to memory, i.e. ("Pirrin", "80, 105, 114, 114, 105, 110").

Above we have written two times the same name as letters and codes. Because of this we may omit this couple and index, and read/write directly to the address "80, 105, 114, 114, 105, 110". For human this address will be shown as "Pirrin", but for the computer it will be "80, 105, 114, 114, 105, 110".

### Models for knowledge representation

Maybe the simplest model for knowledge representation is one used for dictionaries and vocabularies. Representation as vocabulary is given in Table 1. It is good for using by humans but it is not appropriate for processing in the computer. The same knowledge may be represented in the most popular data format – the relational one (Table 2).

**Table 1.** Representation as vocabulary

<i>name</i>	<i>definition</i>
<b>Pirrin</b>	A mountain
<b>Aphla</b>	The South Side of Pirrin mountain
<b>Ateb</b>	The North Side of Pirrin mountain
<b>(x, y)</b>	Co-ordinates of Pirrin mountain

**Table 2.** Relational representation

object	<i>South_Side</i>	<i>North_Side</i>	Co-ordinates
<b>Pirrin</b>	<b>Aphla</b>	<b>Ateb</b>	<b>(x, y)</b>

Knowledge from Table 1 or Table 2 may be used for creating ontology with four concepts which may be represented by an ontology graph (Figure 1):

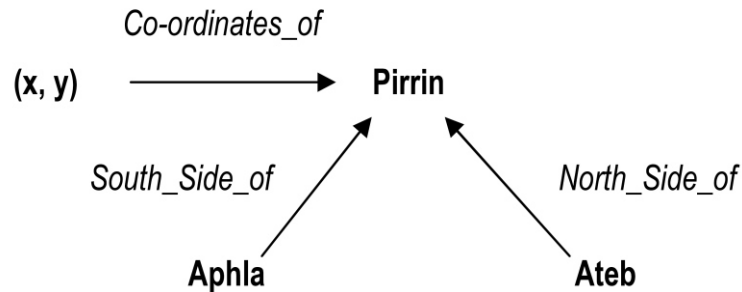


Figure 1. Ontology graph

RDF representation of ontology from Figure 1 is given in Table 3. Now we are ready to illustrate our approach for storing RDF-graphs by means of the NLA. Taking in account the interrelations between nodes and edges on Figure 1, a two-dimensional “multi-layer” representation of knowledge becomes possible (Table 4). It is usual for humans but it is not wide used in the computers. The layers from Table 4 may be stored in different files. If we will use the possibility for NLA, the name “Pirrin” of the column will define locations in files of layers.

Table 3. RDF representation of ontology

subject	relation	object
<b>Pirrin</b>	<i>South_Side</i>	<b>Aphla</b>
<b>Pirrin</b>	<i>North_Side</i>	<b>Ateb</b>
<b>Pirrin</b>	Co-ordinates	<b>(x, y)</b>

Table 4. Multi-layer representation of ontology

object	Pirrin
layer	
<i>South_Side</i>	<b>Aphla</b>
<i>North_Side</i>	<b>Ateb</b>
Co-ordinates	<b>(x, y)</b>

To receive all knowledge for given node, we have to take node (column) name as NLA, for instance “Pirrin”, and read all information stored at location determined by its encoding (“80, 105, 114, 114, 105, 110”) as NLA in different layers (rows “*South\_Side*”, “*North\_Side*”, and “Co-ordinates”). In this multi-layer knowledge representation we have important achievement – only cells from Table 4, given in bold, will be stored in computer memory. All other information is “virtual” address information used for access to real information. This causes avoiding of supporting indexes for speeding of information search and as a result reducing of used computer resources - memory and processing time.

## Conclusion

Concluding, let point on advantages and disadvantages of the illustrated above multi-layer knowledge representation by means of the NLA. The main advantages are reducing the number of tables, which represent the graph as well as reducing the number of filled cells. The main disadvantages are: the tables are sparse; avoiding pointers we receive a variety of names, which have different lengths and cause difficulties for the implementations in the data bases where the fixed length is preferable; the number of nodes may be very great and this way needs corresponded number of columns in the table (in any cases hundreds or thousands). The disadvantages may be avoided if we will use the Multi-Domain Information Model (MDIM) [Markov, 2004] and

---

corresponded Multi-Domain Access Method (MDAM) [Markov, 1984]. We upgraded MDAM to NLA. The experiments with NLA and multi-layer organization of the information have shown that they are good basis for implementing for real solutions.

---

### Bibliography

---

- [Ivanova et al, 2012] Krassimira Ivanova, Vitalii Velychko, Krassimir Markov. "About NL-addressing" (К вопросу о естество-языковой адресации) In: V. Velychko et al (ed.), Problems of Computer in Intellectualization. ITHEA@ 2012, Kiev, Ukraine - Sofia, Bulgaria, ISBN: 978-954-16-0061 0 (printed), ISBN: 978-954-16-0062-7 (online), pp. 77-83 (in Russian).
- [Ivanova et al, 2013] Krassimira B. Ivanova, Koen Vanhoof, Krassimir Markov, Vitalii Velychko, "Introduction to the Natural Language Addressing", International Journal "Information Technologies & Knowledge" Vol.7, Number 2, 2013, ISSN 1313-0455 (printed), 1313-048X (online), pp. 139–146.
- [Ivanova, 2014] Krasimira Ivanova, "Storing Data using Natural Language Addressing", PhD Thesis, Hasselt University, Belgium, 2014
- [Markov, 1984] Krassimir Markov, "A Multi-domain Access Method", Proceedings of the International Conference on Computer Based Scientific Research, PLOVDIV, 1984, pp. 558 - 563.
- [Markov, 2004] Krassimir Markov, "Multi-domain information model", Int. J. Information Theories and Applications, 11/4, 2004, pp. 303 - 308
- [Markov, 2004a] Krassimir Markov, "Co-ordinate based physical organization for computer representation of information spaces", (Координатно базирана физическа организация за компютърно представяне на информационни пространства) Proceedings of the Second International Conference "Information Research, Applications and Education" i.TECH 2004, Varna, Bulgaria, Sofia, FOI-COMMERCE – 2004, стр. 163 - 172 (in Bulgarian).
- [Sowa, 2000] John F. Sowa, Knowledge Representation: Logical, Philosophical, and Computational Foundations, Brooks Cole Publishing Co., Pacific Grove, CA, ©2000. 594 + xiv pages; ISBN 0-534-94965-7

---

### Authors' Information

---

**Ivanova Krassimira** – University of National and World Economy, Sofia, Bulgaria; Institute of Mathematics and Informatics, BAS, Bulgaria; e-mail: krazy78@mail.bg

*Major Fields of Scientific Research: Software Engineering, Business Informatics, Data Mining, Multidimensional multi-layer data structures in self-structured systems*

---

## ONTOLOGY USING IN GEOMETRICAL MODELS DATA PROCESSING OF COMPLEX TECHNICAL OBJECT

Dmytro Konotop

**Abstract:** *Geometrical models of complex technical object are created by means of different information technology selected by enterprise or departments of enterprise. However information technology formats of the geometrical models are not combined and consequently it is hard enough to process the different information concerning geometrical models. The approach offered in present work shows attempt to connect different information technology formats of the data which is working out of geometrical models, using ontology.*

**Keywords:** *ontology, geometrical models, complex technical object, information technology formats, CAx and PLM-systems.*

---

### Introduction

Different information technologies (IT) are widely used at the enterprises which are developing complex technical objects (CTO) for simplification of CTO design process and manufacturing, for the decision of the most important problems at all CTO design stages and the coordination of simultaneous work of designers over the project during CTO creation.

CAD/CAM/CAE (CAx) - systems under control of PLM-system are used at creation of geometrical models (GM). However there is a communication problem between GM created on previous and the final stages of CTO design and manufacture. There are also often the lacks connected with transfer of high-grade data concerning GM from different IT formats data. Present work offers use of ontology for the decision of the problems set forth above.

---

### Representation of CTO GM design process using of ontology approach

There are the following main GM of CTO design stages with the use of IT: master geometry, distribution of objects, and complete definition of product on which criteria of GM design are imposed and methods of experimental and numerical researches are selected. Actually, now at the enterprises these GM design stages are not connected and every GM from its design stage is designed separately from another stages. And the first problem of CTO GM design is GM connection for different design stages.

Geometrical models of CTO are created by means of different IT, which formats are not combined and consequently it is hard enough to process the information about GM. Loss of data during translation leads to that transferred GM between different systems becomes incomplete. Also accessible translators are created for concrete IT and are not universal and have a high cost. This is the second problem of CTO GM design is the IT formats exchange. One of the attempt to solve these problems is the ontology approach using.

Ontology is defined as specification of a representational vocabulary for a shared domain of discourse – definitions of classes, relations, functions, and other objects. [Gruber, T.R., 1993] For the coding of ontology, formal language of the ontology description is used. The most widespread language of ontology coding is OWL (Ontology Web Language). Among a considerable quantity of ontology editors easily expanded editor Protégé has been selected. This work presents the concept of control parametric model (CPM) which is a managing structure, which intended for support of initial data inheritance at all CTO design stages in CAx-system. The control parametric model consists from the following models: kinematic, construction, systems and equipment. On fig. 1 and 2 the ontology of knowledge-based CTO GM design with the use of CPM and activity diagram of ontology-based CTO GM design respectively are presented.

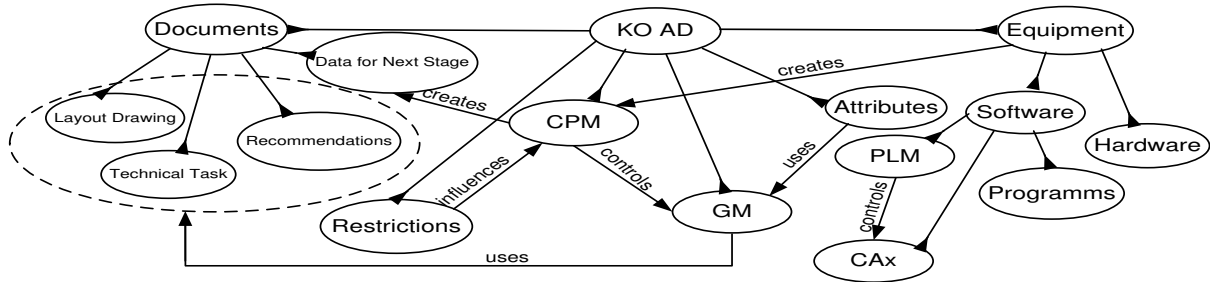


Fig. 1. Ontology of knowledge-based CTO GM design

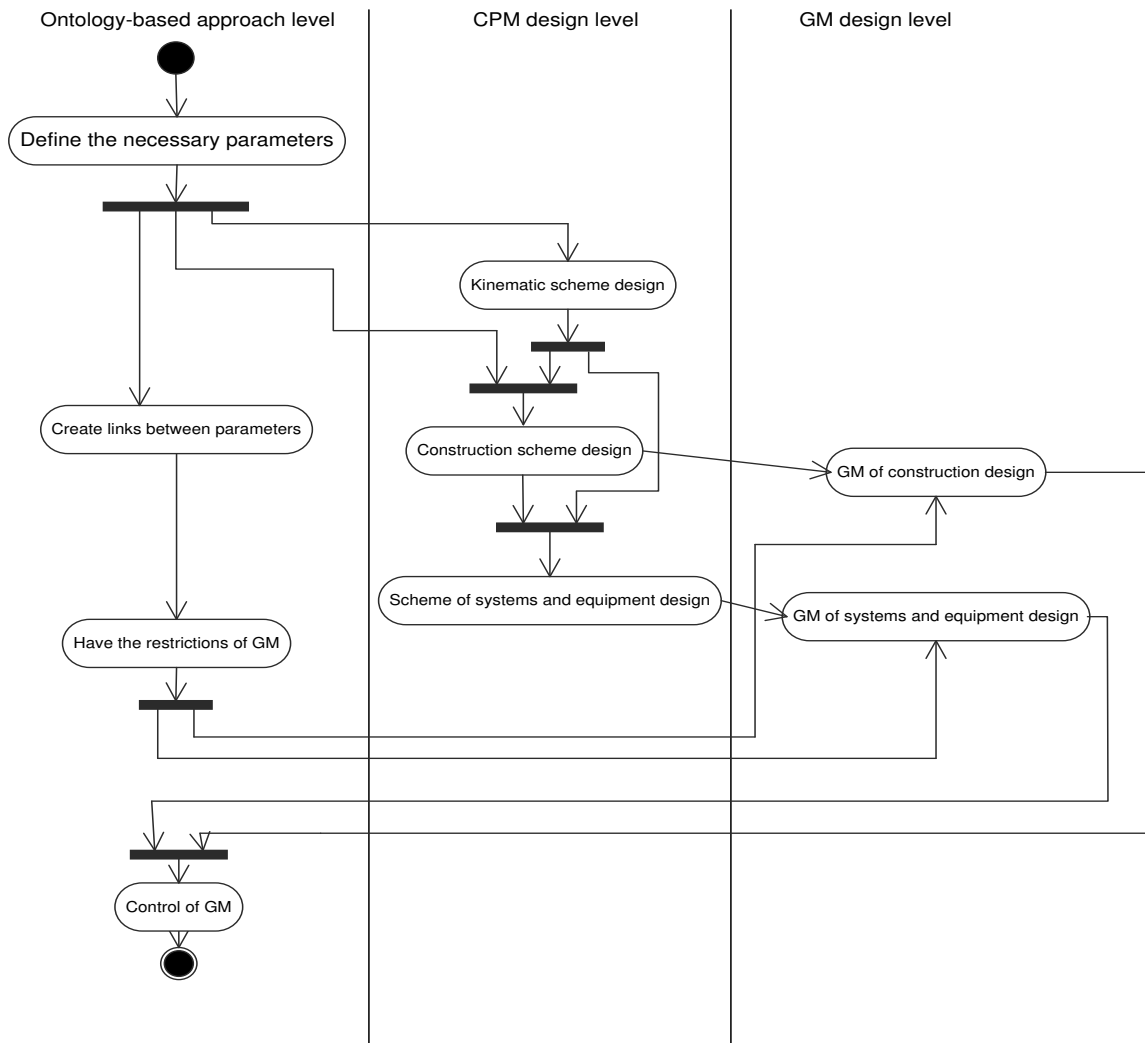


Fig. 2. Activity diagram of ontology-based CTO GM design

## Conclusion

This work describes the ontology-based approach which is used in geometrical modeling of complex technical object and helps to connect of all complex technical object design stages. Also proposed approach is possible to use for exchange of different formats of geometrical models.

## Bibliography

[Gruber, T.R., 1993] A Translation approach to portable ontology specifications. Knowledge Acquisition 5, pp.199-220, 1993.

## Authors' Information

**Dmytro Konotop** – National Technical University of Ukraine "Kyiv Polytechnic Institute", PhD student, engineer-designer, P.O. Box: 03056, Peremogy av., 37, Kyiv, Ukraine; e-mail: [konotop.dmitriy@gmail.com](mailto:konotop.dmitriy@gmail.com)



## GLOBAL SYSTEM FOR MONITORING, DIAGNOSTIC AND MANAGING OBJECTS, AS AN INTEGRAL PART OF THE INFORMATION SOCIETY OF THE FUTURE

**Alexander Kosianchuk**

**Abstract:** *Systems position monitoring of objects in space are an integral part of the modern information society. Extension of such systems can be called a system of monitoring, diagnosis and control. To audience it will be presented a prototype aircraft equipment monitoring and control system of flight data.*

**Keywords:** *building the information society, system of monitoring, diagnosis and control, aircraft equipment.*

---

### Introduction

---

Monitoring - continuous process monitoring and recording parameters of the object, as compared with the set criteria. Systems position monitoring (SPM) of objects in space is an integral part of the modern information society. It should be emphasized that a mandatory attribute of any SPM is mapping software that allows you to put the position on the map to determine the relative positioning. Extension of such systems can be called a system of monitoring, diagnosis and control. The basic principles of operation of such systems will be considered on the example of aviation equipment "Luch-14".

---

### The Main Theses of Manuscripts

---

An information society is a society where the creation, distribution, uses, integration and manipulation of information is a significant economic, political, and cultural activity. The aim of the information society is to gain competitive advantage internationally, through using information technology (IT) in a creative and productive way. SPM are an integral part of the modern information society.

Examples of monitoring systems are services such as Google locations, allowing the user to track the position of users through the provision of mobile devices; Flight Radar 24, visualizing information ADSB on the map of the planet; also this list includes a lot of the monitoring system of land transport, for example Naviset, Euromobile and others.

It should be emphasized that a mandatory attribute of any SPM is mapping software that allows you to put the position on the map to determine the relative positioning. Thus the foundation of the SPM is mapping services Google (Google Maps, Google Earth) or analogues of other companies (Yahoo! Maps, Microsoft Bing Maps, WikiMapia et al.).

For efficient and more economical operation of SPM can use available communication channels, such as WiFi, EDGE / GPRS, 3G, 4G, Iridium.

One of the most high-tech spheres of human activity today is aviation. Leaders of the air transportation system is widely used SPM, where they are called flight data monitoring system in real time (FDMS). Examples of these airlines are British Airways, Malaysia Airlines and others.

Ukraine also is providing researches in this area. Nowadays, the company "Aviatsiyni system" earned a prototype system FDMS "Luch-14", designed to replace obsolete aircraft emergency registrar "SARPP", which numbered around the world in the tens of thousands.

Hardware-software complex "Luch-14" for SARPP designed for the collection and analysis of flight data for the purpose of regular processing and statistical analysis, control, flight technical crew, troubleshooting avionics, proactive actions anticipation of aircraft accidents and incidents (AAI), and reactive actions after AAI to prevent recurrence of situations.

Components "Luch-14":

- 
- Protected onboard storage, storage systems equipped with parametric audio and video;
  - Coding unit (conversion of sensor signals in ARINC);
  - Additional built-in sensors barometric altitude dimensional overloads, three-dimensional magnetic field and GPS coordinates;
  - Communication tools, Iridium, Wi-Fi, GPRS / EDGE / 3G;
  - Lightweight onboard storage;
  - Software environment data processing and visualization.

"Luch-14" was tested on Mi-8 and proved calculated results about data transcribing speed.

Extension of such systems can be called a system of monitoring, diagnosis and managing (SMDM). The concept of SMDM is installing the program sequence (in the case of an automated system), monitoring of parameters of movement and interference in the control loop if necessary (for example in a critical situation).

---

## Conclusion

In the future, when SMDM is widely used, single center (aviation, transport, agriculture) management will be able to assume responsibility for the management of all subordinate means and do it more effectively and in a coordinated manner. Besides regular users will be able to control the movement of their vehicles or accurately calculate the arrival of the bus at a stop that will undoubtedly make life more comfortable and will free up more time for cultural, spiritual and physical development of the individual and society as a whole.

---

## Bibliography

[Soloviev, 2010] E.H. Soloviev. The information society. In: The new philosophic encyclopedia. V.S. Stepin. "Misl", Russian federation, 2010 ISBN 978-5-244-01115-9.

---

## Authors' Information

**Alexander Kosianchuk** – Ukrainian National aviation university student, P.O. Box: Kyiv Komarova ave 1, Ukraine; e-mail: [alexander.kosianchuk@gmail.com](mailto:alexander.kosianchuk@gmail.com)

## ENERGY VERSUS INFORMATION

Krassimir Markov

**Abstract:** The interrelation between concepts "Energy" and "Information" is considered in the paper. Using simple example, it is shown that the concept "Energy" has objective foundation but the concept "Information" has subjective one.

**Keywords:** General information theory, Energy, Information.

### Introduction

The interrelation between energy and information is fundamental and needs to be clearly explained.

Mark Burgin explained one of the common understandings of this relation as "Energy is a kind of information and information is a kind of energy". More concretely, in his monograph "Theory of information" [Burgin, 2010, p.117] he had written: If we take that matter is the name for all substances as opposed to energy and the vacuum, we have the relation that is represented by the following diagram called the Structure-Information-Matter-Energy (SIME) Square (Figure 1). In other words, we have the following principle: Information is related to structures as energy is related to matter.

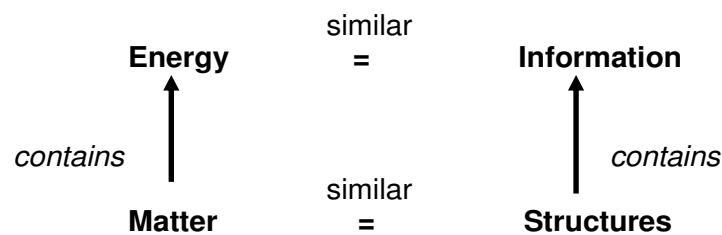


Figure 1. The Structure-Information-Matter-Energy (SIME) Square [Burgin, 2010, p.117]

From our point of view, this is not correct conclusion. Below we will present a simple example to show the main difference between "Energy" and "Information". Before the example, we will remember our definition of the concept "Information" and the common understanding of meaning of the concept "Energy".

### Information

Following "General Information Theory" [Markov et al, 2007] we may define concept "Information" as quadruple:

$$i = (\text{source}, \text{recipient}, \text{evidence}, \text{Infos}) = (s, r, e, I)$$

where:

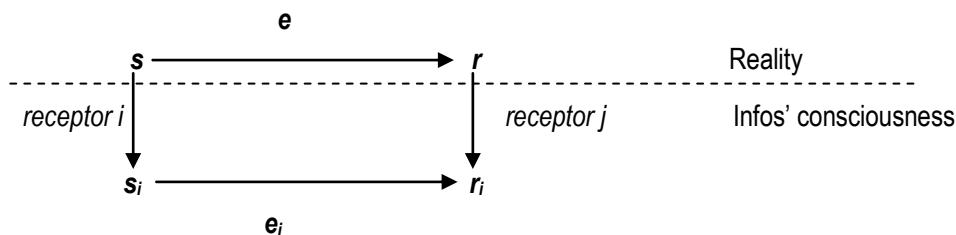
- Source "s" and Recipient "r" are structured sets;
- Infos "I" is an intelligent system;
- Evidence "e" is a mapping from s in r which preserves (all or partial) structure and resolves the Information Expectation of Infos.

In mathematics, a *structure on a set*, or more generally a *type*, consists of additional mathematical objects that, in some manner, attach (or relate) to the set, making it easier to visualize or work with, or endowing the collection with meaning or significance. A partial list of possible structures are measures, algebraic structures (groups, fields, etc.), topologies, metric structures (geometries), orders, events, equivalence relations, differential structures, and categories.

Sometimes, a set is endowed with more than one structure simultaneously; this enables mathematicians to study it more richly. For example, an order induces a topology. As another example, if a set both has a topology and is a group, and the two structures are related in a certain way, the set becomes a topological group.

Mappings between sets which preserve structures (so that structures in the domain are mapped to equivalent structures in the codomain) are of special interest in many fields of mathematics. Examples are homomorphisms, which preserve algebraic structures; homeomorphisms, which preserve topological structures; and diffeomorphisms, which preserve differential structures [Mathematical\_structure, 2014].

Resolving of Infos' Information expectation is constructing of commutative diagram:



where:

- Source “ $s$ ” and Recipient “ $r$ ” are structured sets;
- Infos is an intelligent system;
- Evidence “ $e$ ” is a mapping from  $s$  in  $r$  which preserves (all or partial) structure;
- Mental source “ $s_i$ ” and Mental recipient “ $r_i$ ” are structured sets;
- Mental evidence “ $e_i$ ” is a mapping from  $s_i$  in  $r_i$  which preserves (all or partial) structure.

We use the name “Information expectation” because of creation of the triple  $(s_i, r_i, e_i)$  in advance by Infos.

---

## Energy

---

In physics, “Energy” is a property of objects, transferable among them via fundamental interactions, which can be converted in form but not created or destroyed. There are many forms of energy, but all these types must meet certain conditions such as being convertible to other kinds of energy, obeying conservation of energy, and causing a proportional change in mass in objects that possess it. Common energy forms include the kinetic energy of a moving object, the radiant energy carried by light and other electromagnetic radiation, the potential energy stored by virtue of the position of an object in a force field such as a gravitational, electric or magnetic field, and the thermal energy comprising the microscopic kinetic and potential energies of the disordered motions of the particles making up matter. Some specific forms of potential energy include elastic energy due to the stretching or deformation of solid objects and chemical energy such as is released when a fuel burns. Any object that has mass when stationary, such as a piece of ordinary matter, is said to have rest mass, or an equivalent amount of energy whose form is called rest energy, though this isn't immediately apparent in everyday phenomena described by classical physics.

According to mass–energy equivalence, all forms of energy (not just rest energy) exhibit mass [Energy, 2014].

---

## Energy versus Information

---

From definition of Information given above follows that the information is a kind of reflection but not every reflection is information. Only the reflection, for which given Infos (subject, intelligent system) has evidence what the reflection reflects, is information only for this concrete Infos.

It is clear, to create a reflection, the energy is needed. Without energy no internal changes (reflections) in the entities may be realized. This means that energy is needed to realize reflection which may become information for given Infos.

*Without energy information is impossible.*

But the opposite correspondence does not exist. Energy does not depend on information. It exists in reality without subjective "decisions".

*Energy is objective phenomenon, Information is subjective phenomenon.*

---

### **Example**

---

Let we have two equal pieces of paper.

They contain some energy, let assume that its quantities are equal in both pieces. In other words, for instance, if we burn these pieces they will release practically the same quantities of energy.

If I have such piece of paper and you have another such one, we may exchange them as equivalent without any additional conditions.

Let now the pieces of paper are painted with some colors. The paint will add some additional energy to pieces. Let assume that, again, it is in equal quantities in both pieces.

Again, we may exchange pieces as equivalent without any additional conditions.

At the end, let pieces of paper are painted as follow:

- the first piece is painted as USD 100 (one hundred dollars);
- the second one is painted as RUB 100 (one hundred rubles),

i.e. let have two real banknotes.

Now, we will not agree to exchange these pieces of paper without additional conditions.

As it is shown by Bloomberg, on 08/29/2014, 12:59:59, (<http://www.bloomberg.com/quote/USDRUB:CUR>)

US DOLLAR-RUSSIAN RUBLE Exchange Rate is: Price of 1 USD in RUB is 37.1189,

i.e. now the first piece of paper is equivalent to more than 37 pieces of second one.

Because of information for the subjects, the pieces became different notwithstanding that the energy quantities are equal in both pieces.

*The subjective decisions have important role in this case.*

---

### **Conclusion**

---

Our understanding of the relation between energy and information differs from one in [Burgin, 2010]. After reading the example, presented in this paper, Mark Burgin had remarked: "Author very well explained the differences between energy and information in the sense of General Information Theory (GIT) [Markov et al, 2007]. These differences appear because GIT studies information on the higher level than the General Theory of Information (GTI) [Burgin, 2010]. If we look into mathematics, we see that group theory studies mathematical structures on the higher level than set theory. Although set theory is most basic, while group theory has more applications outside mathematics, both theories - set theory and group theory - are necessary for mathematics as a whole. Thus, we may compare GIT to group theory and GTI to set theory as groups have additional structure in comparison with sets as the information quadruple of GIT has additional structure in comparison with the information triad of GTI."

---

We agree that the "Energy" may be explained by triple: (source, recipient, transition)  $\Rightarrow (x, y, f) \Rightarrow y = f(x)$ , but "Information" has to be explained by quadruple (source, recipient, evidence, Infos) = (s,r,e,I). From our point of view, *the energy and information are different phenomena – objective and subjective, respectively.*

*The triples are object of study by Mathematics, quadruples – by Informatics.*

---

### **Bibliography**

---

[Burgin, 2010] Burgin M. Theory of Information - Fundamentality, Diversity and Unification. World Scientific Publishing Co. Pte. Ltd. 2010. ISBN-13 978-981-283-548-2

[Markov et al, 2007] Kr. Markov, Kr. Ivanova, I. Mitov. Basic Structure of the General Information Theory. IJ ITA, Vol.14, No.: 1, 2007. pp. 5-19.

[Mathematical\_structure, 2014] [http://en.wikipedia.org/wiki/Mathematical\\_structure](http://en.wikipedia.org/wiki/Mathematical_structure) (visited on 01.09.2014)

[Energy, 2014] <http://en.wikipedia.org/wiki/Energy> (visited on 01.09.2014)

---

### **Authors' Information**

---

**Krassimir Markov** – Institute of mathematics and informatics, BAS; e-mail: [markov@foibq.com](mailto:markov@foibq.com)

*Major Fields of Scientific Research: General theoretical information research, Multi-dimensional information systems*

## A COMPUTER DIALOG WITH WIND POWER GIVES THE NON-EMERGENCY WIND ENERGY

**Valery Pisarenko, Julia Pisarenko, Vladimir Malachinskij**

**Abstract:** *The variant of vibration diagnostics system for wind turbines, actual for remote diagnostics operating of commercial wind turbines, is proposed.*

**Keywords:** *vibration diagnostic, spectral analysis, vibroacoustic signal, spectrogram, wind turbines.*

---

### Introduction

In recent years, in most industrialized countries in the world are actively expanding network of industrial wind turbines. Technologically are very attractive the location of wind turbines industrial capacity in such areas of relief, where the average wind speed is maximum for the given area. At the same time, it is preferable to choose a place remote from the main local region's infrastructure (lack of roads, residential areas, industrial complexes). As a result of putting into operation new and powerful wind turbines in the majority of cases are removed from the control points and control of their work by people (operators).

---

### About the system of natural modeling of diagnostic equipment vibro windmill

The above circumstances (extreme modes of wind turbines operation, the distance of technical personnel placement, designed to monitor the performance of wind-energy complex) are forced to develop a highly reliable means of remote control of wind turbines monitoring. This is especially important for minimize of probability of accidents on wind turbines in extreme weather conditions [Bardik, 2012], the complicated conditions of staff access to wind turbines. In particular, in such difficult conditions are working offshore wind turbines in the North Sea, remoted is not uncommon for a distance 100 km or more. In the context of Ukraine, the actual of remote monitoring at any time in any weather high for existing and future wind turbines mountainous terrain of Carpathia and Transcarpathia. Especially given the fact that these regions, as known, have a high industrial wind potential, yet poorly developed. In our paper [Pisarenko, 2012], sets out a vision of an expert vibro diagnostics system of wind turbine using remote of current operation mode monitoring by the operator with the sensors mounted in the nacelle wind turbine. For the purpose of simulation modeling of called technique of remote control of pre-emergency modes of wind turbines operation in this paper are presents the results of full-scale experiments on the operating model of rotary mechanisms with a number of important features, highly informative for remote monitoring of rotating machines type wind turbine and industrial ventilators.

The goal of experiments, in this case, was to identify possible situations that resonant oscillations occur within the test speed range of the main shaft of the rotor installation. Technologically for solving the problem was used the following installation scheme, comprising: main shaft of rotary, sensors of vibration frequencies of the main shaft near the resonant frequencies of the individual components of the entire system (provides the ability to identify of the pre-emergency modes of vibration due to wearout of rolling bearings of the shaft defects, violating the normal mode of the fixation of through-hole assemblies of throughout installation, etc.).

The Table 1 shows the indications for each of the three cases of forced frequency selected for experiment. At first was extracted from the spectrum the four nearest competitor, then were calculated for each of the frequencies of the most dangerous competitors. According to the criterion

$$\min(\delta(f_{forced}) - \delta(f_i)), \quad (1)$$

where  $f_{forced}$  – frequency in emergency mode of wind turbine vibration,  $f_i$  – frequencies-satellites from the spectrum the four nearest competitor,

was selected the most dangerous mode with  $f_i(f_{forced})=70$  Hz.

Table 1 - Severity of accidents ( $f_i$  – frequencies-satellites of  $f_{forced}$ ) for three variants of the forced frequency

$f_{forced}$

No	$f_{forced}$ , Hz	$\delta(f_{forced})$ , dB	$f_i(f_{forced})$ , Hz	$\delta(f_i)$ , dB	$\min(\delta(f_{forced}) - \delta(f_i))$ , dB
1	42	-25,5	50	-31	5,5
			55	-35,5	
2	70	-16	52	-22,5	6,5
			87	-32	
			105	-33	
			125	-45	
3	73	-16,5	50	-26,5	6,5
			54	-23	
			125	-28	
			150	-28	

## Conclusion

In this paper were presents the results of full-scale experiments on the operating model of rotary mechanisms with a number of important features, highly informative for remote monitoring of rotating machines type wind turbine and industrial ventilators. In experiment were identified possible situations when resonant oscillations occur within the test speed range of the main shaft of the rotor installation.

## Bibliography

[Bardik, 2012] Bardik E.I. Simulation of a wind farm to assess the reliability of the changing weather conditions// The XIII International conference of renewable power (c.Mykolaivka). – 2012.

[Pisarenko, 2012] Pisarenko V.G., Shevtchenko O.A., Danilenko A.I., Prokopchuk V.V., Pisarenko Yu.V., Varava I.A. The concept of expert system of vibration diagnostics of VEU units on the basis of remotely operated multichannel vibrosensors // The XIII International conference of renewable power (c.Mykolaivka). – 2012.

## Authors' Information

**Valery Pisarenko** – chief of department, professor, Sc.D (doctor of physical and mathematical sciences), Glushkov Institute of Cybernetic of NAS of Ukraine, Ukraine, 03680, MSP, c.Kiev, Academishian Glushkov av., 40; e-mail: jvpisarenko@gmail.com

**Julia Pisarenko** – Senior Researcher, Ph.D. (candidate of technical sciences), Glushkov Institute of Cybernetic of NAS of Ukraine, Ukraine, 03680, MSP, c.Kiev, Academishian Glushkov av., 40; e-mail: jvpisarenko@gmail.com

**Vladimir Malachinskij** – engineer-programmer, Glushkov Institute of Cybernetic of NAS of Ukraine, Ukraine, 03680, MSP, c.Kiev, Academishian Glushkov av., 40



---

## WIRELESS SENSOR NETWORK FOR PRECISION FARMING AND ENVIRONMENTAL PROTECTION

**Volodymyr Romanov, Igor Galelyuka, Oleksandr Voronenko**

**Abstract:** *In the paper it is described the wireless sensor network for precision farming and environmental protection, which is developed in V.M.Glushkov Institute of Cybernetics. The main aim of the wireless sensor network is data acquisition in real-time mode of state of plants on large territories, on-line data processing and timely decision making according to specified criteria.*

**Keywords:** *wireless sensor network, biosensor, wireless data acquisition.*

---

### Introduction

---

In modern industrial farming it is necessary to obtain information in real-time about influence of climatic and anthropogenic stress factors on state of agricultural crops. For protection of green plantation in megalopolis, forest reserves and national parks it is reasonable to use similar tools for estimating impacts of stress factors of different origin on plantations of these territories. Up-to-date achievements of biosensorics, microelectronics and information technologies make it possible to successful solve this very complicated task for large territories.

---

### Wireless sensor network development

---

Several years ago portable computer devices of "Floratest" family were developed and prepared for full-scale production [Romanov, 2012]. The main purpose of these devices was diagnostics of plant state in real-time mode in laboratory and field conditions. Industrial precision farming and environmental protection are main application fields of these portable devices. But, the field testing in German agrarian farms showed some restrictions during active work with "Floratest" device in the field conditions. In particular, experienced specialist had to run measurements and the one measurement, including dark adaptation, took a long time. For speedup of measurements in general it is necessary to use a large number of autonomous portable devices. To avoid these restrictions it was started development of data acquisition systems on the base of portable computer devices of "Floratest" family [Palagin, 2011]. Such approach allowed to centralize running of measurements, made with many devices, and accumulate obtained information in the control center. This handling lets to cover by monitoring a large territory of agricultural lands, parklands or woodlands.

Modern achievements in microelectronics and information and communication technologies after detailed analysis show possibility of creating wireless sensor network for express-diagnostics of state of plants on large territories with using the same effect of chlorophyll fluorescence induction, on which the operation of devices of "Floratest" family is based on. Application of microsensors for measuring additional parameters of air and soil in sensor network gives possibility in indirect method to measure not only the general state of plant, but also such parameters, as pesticides content in soil, water and plants, level of soil corrosion, level of pollution in air etc.

Wireless sensor network for our applied problem has to satisfy requirements to typical sensor network for application in agriculture and environmental protection. Development of wireless sensor network depends on many factors, including failure-resistance, scaling, manufacture cost, type of operational environment, topology of sensor network, hardware restrictions, data transferring model, energy consumption and etc. Schemas for initial placing of nodes of sensor network have to satisfy requirements of cost reducing of installation, removing necessary of preliminary planning of placing schemas, improving flexibility of node placing, and assisting self-organization and failure-resistance. Nodes of our wireless sensor network are deployed in static mode. Since nodes of sensor network sometimes fail because of battery discharge or physical impacts, the structure of sensor network has to change its topology very often after the sensor network has been deployed. Changes of topology

cause modification of characteristics of network nodes, notably location, accessibility (because of noise, moving obstacles etc.), battery charge, failures, changing of node functions in network. It is specified possibility to include additional nodes in sensor network for replacement of non-working nodes or expansion of network or network functions.

For global data transfer in sensor network it is necessary, that communication environment was accessible in many countries. The best way to solve this task is to organize communication between nodes with using of industrial frequencies, which are accessible without licenses in majority countries. Hardware restrictions and compromise between antenna efficiency and power consumption put some restrictions on choosing of frequency for data transferring. The main advantage of frequencies for industrial application is wide frequency spectrum and their accessibility almost in whole world. For purpose of data transferring via wireless channel there were chosen tiny wireless modules (NXP company manufacture) with embedded multichannel ADC. Wireless modules work at frequency of 2,4 GHz.

During development of sensor network it is necessary to provide the possibility to integrate the sensor network with systems of top level, for example, decision-making and agrotechnological operation generating system and "cloud" technologies.

---

## Conclusion

Detailed analysis of modern literature and technical solutions shows, that development of wireless sensor network for industrial agriculture and environmental monitoring is very complicated task. Implementation of proposed wireless sensor networks to industrial agriculture and environmental protection makes it possible to increase efficiency and quality of end product, spare fertilizers, water and energy resources, timely protect plants of woodlands and parklands from infections and anthropogenic impact.

---

## Bibliography

- [Romanov, 2012] Romanov V., Artemenko D., Brayko Yu., Galelyuka I., Imamutdinova R., Kytayev O., Palagin O., Sarakhan Ye., Starodub M., Fedak V. Portable Biosensor: from Idea to Market // International Journal "Information Theories & Applications. Vol. 19, Number 2. – Sofia, Bulgaria. – 2012. – P. 126–131.
- [Palagin, 2011] Palagin O.V., Romanov V.O., Galelyuka I.B., Velichko V.Yu., Hrusha V.M. Data acquisition systems of plants' state in precision agriculture // Proceeding of the 6th IEEE International conference on "Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications", IDAACS'2011. – Prague, Czech Republic. – 2011, September 15–17. – P. 16–19.

---

## Authors' Information

**Volodymyr Romanov** – Head of department of V.M.Glushkov Institute of Cybernetics of National Academy of Sciences of Ukraine, Doctor of technical sciences, professor; Prospect Akademika Glushkova 40, Kiev–187, 03680, Ukraine; e-mail: [VRomanov@i.ua](mailto:VRomanov@i.ua); website: <http://www.dasd.com.ua>

**Igor Galelyuka** – Senior research fellow of V.M.Glushkov Institute of Cybernetics of National Academy of Sciences of Ukraine; Candidate of technical science; Prospect Akademika Glushkova 40, Kiev–187, 03680, Ukraine; e-mail: [galib@gala.net](mailto:galib@gala.net); website: <http://www.dasd.com.ua>

**Oleksandr Voronenko** – research fellow of V.M.Glushkov Institute of Cybernetics of National Academy of Sciences of Ukraine; Prospect Akademika Glushkova 40, Kiev–187, 03680, Ukraine; e-mail: [alexander.voronenko@dc.kiev.ua](mailto:alexander.voronenko@dc.kiev.ua); website: <http://www.dasd.com.ua>

---

## RESEARCH OF THE FLUOROMETER «FLORATEST» SENSITIVITY IMPACT OF STRESS FACTORS ON THE PLANTS

**Volodymyr Romanov, Volodymyr Hrusha, Oleksandra Kovyrova**

**Abstract:** Results of research of fluorometer "Floratest" sensitivity to impact of stress factors on state of plants on example of *datura* are considered in the article.

**Keywords:** chlorophyll fluorescence induction; fluorometer.

---

### Introduction

Portable computer devices of "Floratest" family were developed in V.M. Glushkov Institute of Cybernetics of NAS of Ukraine. Serial production of devices was started few years ago [Palagin, 2012]. The devices can provide express-diagnostics of changes of alive plant state and give opportunities for fast estimation of environment influence level and technological stresses on alive plants. The device principle based on measurement of chlorophyll fluorescence induction (ChFI) curve in real time without damaging alive plants [dasd.com.ua]. Application fields of the devices are quite wide and cover both scientific researches and industrial agriculture.

---

### Experiments description

In 2012-2013 in the Institute of Cybernetics some researches was completed for purpose of estimating the fluorometer sensitivity to impact of stress factors on alive plant state. Changes of curve form of chlorophyll fluorescence induction of weed *Datura* under influence of herbicide «Roundup» were studied.

Times of beginning measurement, parameters of environment (air and ground temperatures and humidity) were fixed. The dark adaptation of leaves was equal 3 minutes. The time of every measurement was 160 s. Leaves for measurement were being chosen from one level. It is known if leaves are under direct rays of the sun they need more than 10 minutes of dark adaptation before measurement. This fact also was confirmed with the fluorometer "Floratest" [Груша, 2012]. For the purpose of reduction of the dark adaptation period the plants were taken away from the sun and placed in the shadow for 30 minutes before the start of the measurements.

In 2012 experiment was conducted from July 07 till September 13. 16 measurement runs were completed on 32 plants. The plants of *Datura* were divided into equal groups: one control and two experimental. The two last groups were sprayed with two different doses of herbicide.

In addition in 2013 two experiments were completed too. The first experiment was conducted from April 12 till August 08. 15 measurement runs were completed on 44 plants. The plants were divided into three groups: one control and two experimental. Experimental groups were sprayed with two different doses of herbicide that differed from doses in 2012. The goal of the second experiment was to receive more large quantity of ChFI curves of plants that weren't sprayed with herbicide. Experiment was conducted from July 15 till September 10. Two measurement runs on 44 plants were completed. Some preliminary results of the analysis were presented in [Груша, 2013].

In 2014 two-factor experiment was completed. Plants of *Datura* were chosen again for experiment. Influence factors were different doses of watering and herbicide. In addition it was estimated the sensitivity of "Floratest" to different properties of water to artificial watering.

---

## Initial results

---

Analysis of obtained ChFI curves showed a depression of plants photosynthesis under stress factors. An effect of the herbicide expresses in gradual increase of ChFI level in groups of sprayed plants [Груша, 2012]. Besides it also showed an influence of weather conditions on beginning points of ChFI curves.

In 2012 in different days of experiment the parameters of temperature and humidity were varying from 13 to 30 °C and from 27 to 92% properly. Calculation of linear correlation coefficient between values of ChFI curves parameters ( $F_p$ ,  $F_s$ ,  $F_o$  [Roháček]) and values of environment parameters ( $t_{air}$ ,  $h_{air}$ ,  $t_{ground}$ ,  $h_{ground}$ ) shows a correlation presence. In 2013 the measurements were completed in days without sharp changes of the weather parameters. The air temperature was in range 23-30 °C, humidity – 28-44 %. Calculation of the linear correlation coefficient showed practically lack of the correlation between changing form of ChFI curve and weather conditions as opposed to results that we watched in 2012.

---

## Conclusion

---

The estimation method of plants state based on analysis of form of ChFI curves gives opportunities to diagnose the state of plants in latent period when the influence of spraying with herbicide aren't seen directly. It gives opportunities to evaluate an effect of herbicide in express mode and to make conclusion about efficiency of herbicide and thus its doses can be optimized.

---

## Bibliography

---

[Груша, 2012] Груша В.М., Ковирьова О.В. Дослідження чутливості флуориметра "Флоратест" до дії стресових факторів на стан рослин// Комп'ютерні засоби, мережі та системи. – 2012. – № 11. – С. 119 – 126.

[Груша, 2013] Груша В.М., Ковирьова О.В. Особливості обробки кривих індукції флуоресценції хлорофілу// Матеріали 18-го Ювілейного Міжнародного молодіжного форуму «Радиоелектроніка и молодіжь в XXI веке». Сб. матеріалів форуму. Т. 1. – Харків: ХНУРЭ. 2014. – с. 138-139.

[Roháček, 2002] Rohacek K. Chlorophyll Fluorescence Parameters: The Definitions, Photosynthetic Meaning, and Mutual Relationships// Photosynthetica - 2002, Volume 40, Issue 1, pp 13-29.

[Palagin, 2013] Palagin O., Romanov V., Galelyuka I., Voronenko O., Artemenko D., Kovyrova O., Sarakhan Y. Computer devices and mobile information technology for precision farming // Proceeding of the 7th IEEE International conference on "Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications", IDAACS'2013. – Berlin, Germany. – 2013, September 12–14. – P. 47–51.

[dasd.com.ua] <http://dasd.com.ua/>

---

## Authors' Information

---

**Volodymyr Romanov** – Head of department of V.M. Glushkov's Institute of Cybernetics of National Academy of Sciences of Ukraine, Doctor of technical sciences, professor; Prospect Akademika Glushkova 40, Kiev–187, 03680, Ukraine; e-mail: VRomanov@i.ua

**Volodymyr Hrusha** – research fellow of V.M. Glushkov's Institute of Cybernetics of National Academy of Sciences of Ukraine; Prospect Akademika Glushkova 40, Kiev–187, 03680, Ukraine; e-mail: vhrusha@gmail.com

**Oleksandra Kovyrova** – research fellow of V.M. Glushkov's Institute of Cybernetics of National Academy of Sciences of Ukraine; Prospect Akademika Glushkova 40, Kiev–187, 03680, Ukraine; e-mail: kovyrova.oleksandra@gmail.com

---

## ARCHITECTURE OF SMART SENSOR NETWORK FOR AGRICULTURAL PURPOSES

**Yevgenia Sarakhan**

**Abstract:** *Rapid and unpredictable change in climatic conditions and the consequent difficulty of using statistical calculations to predict yields showed an acute need to develop models predicting crop yields and the influence of climatic conditions, based on the use of non-statistical but operational data. Preliminary data also are essential for precision farming technology, which including operating models predict yields and is used to optimize the costs, such as water, fertilizer, etc., and to maximize the resulting crop.*

**Keywords:** *information technology, information system, Wireless Sensor Network, agriculture, ontology method.*

---

### Introduction

The most important agricultural purposes may include timely, accurate and prompt informing farmers about climatic changes, both at the individual sections of the field and at the level of administrative- territorial areas. Today one solution, which can be useful for automatic acquisition of knowledge about climate change indicators, is using a wireless sensor network. Distributed self-organizing network of many sensors that are combined with each other through the air, will be useful for the timely application of technological measures to obtain high yields of agriculture. Sensors convert environmental variables in the data that the output of the wireless sensor network through radio coming to the computer input, which with the help man turn into knowledge. For example, wireless sensor network the sensors measure the temperature and humidity, barometric pressure, wind speed and direction. With the release of the wireless sensor network computer input data entering the temperature and humidity, pressure, speed, wind direction, etc. Agronomist by processing these data on a computer receives the prediction, which help to obtain new knowledge, to get, to plan and carry out certain technical measures for the care of agriculture.

---

### Development of smart sensor network

Field data network consists of two levels: the first level is the transfer of data within the device. The second level is the transfer of data between devices for sending to the central unit for the purpose of delivery to the server data processing. These recommendations on the production network of the first level should be the result of further research and implementation.

The network of the second-level is based on the protocol of IEC 60870-5 (IEC 60870-5). This protocol allows you to transmit sensor readings on request, a timer and change the measured values, as well as monitor the health of the devices themselves. In addition to the functions of the telemetry and remote signaling protocol supports remote control. Thus, the device can, using a special module, pouring or lighting control command from the server.

Use of the standard protocol will allows building high compatibility system with third party software, as well as simplifying the development of the system. Also it will allow integrating easily them with RDF storages.

For replacement of addresses of information objects by predicates, special function which accepts the identifier of the sending device and the address of information object and returns the proper predicate in reply.

Structurally a network of the second level includes a self-organizing network of transceivers, such as ZigBee.

The central node has additional transceivers for communication with the server via two independent channels. For example, these channels can be – wired shedding data and GSM line. However, no wire line and Wi-Fi,

---

which eventually connected to the same wiring line.

The solution to this problem is achieved by a wireless sensor network has ontology-managed node (Ye. Sarakhan, A. Palagin, V. Romonov. Global network for the automatic acquisition of knowledge. Patent of Ukraine № 89025 of 04/10/14, Bull. № 7) [Sarakhan, 2014], the global Internet, computer and human-machine interface, the output of the wireless sensor network is connected to the input of the global Internet through ontology - controlled node, and the output of the global Internet through a computer linked to man- machine interface.

The introduction of new network nodes and connections can expand the functionality of the particular gain new knowledge automatically without human intervention.

Data about the environment come in ontology - controlled junction. Ontology -managed node collects information about the current state of the environment (declarative knowledge) and processes the data using procedural knowledge in the form of appropriate algorithms. As a result, processing of declarative knowledge that is running the agent (a computer program) is formed, such as weather or assessed the environment. This information comes from the output of user man- machine interface that is connected to ontology - controlled unit via a computer.

The proposed scheme of sensor network consisting of a global network to automatically acquire knowledge, can be implemented using serial tablet computers, notebooks, Smartphone, PDAs, etc., and does not require additional devices. This approach ensures the accuracy and completeness of the knowledge generated automatically without human intervention, and can be used to generate new knowledge about the objects or processes that occur, for example, in natural or artificially created environment for monitoring and control in agriculture.

---

## Conclusion

Successful testing independent serial device with wired sensor and dozens of orders of network sensor systems confirm appropriateness of implementation of proposed distributed smart sensors to industrial agriculture.

In addition to laboratory tests, this method is also used in industrial agriculture. It is especially important to use this method in terms of precision farming, which involves the integration of the monitoring of plant state in real time by means of application of fertilizers, pesticides and artificial irrigation controls, etc., and data collection and visualization require accurate binding to a digital map of the field. Especially effective is using precision agriculture with a compatible remote, such as space and ground-based monitoring of plant state.

The ultimate goal of this system is to maximize agricultural productivity while optimizing costs with the ability to make accurate predictions about future crop. It's worth to note that under this approach to creating and building a sensor network, biosensor devices can operate synchronously with space and other means of observation of the plants that provide the maximum requirements of precision farming.

---

## Bibliography

[Sarakhan, 2014] Ye.V. Sarakhan, A.V. Palagin , A.V. Romonov. Global network for the automatic acquisition of knowledge. Patenet Ukraine № 89025 of 04/10/14, Bull. № 7.

---

## Authors' Information

**Yevgenia Sarakhan** – *V.M. Glushkov Institute of Cybernetics of NAS of Ukraine, Senior Research Fellow, 40 Glushkova ave., 03680-Kiev, Ukraine; e-mail: [sarakhan2006@ukr.net](mailto:sarakhan2006@ukr.net)*

## JOINT STUDY OF VISUAL PERCEPTION MECHANISM AND COMPUTER VISION SYSTEMS THAT USE COARSE-TO-FINE APPROACH

**Anton Sharypanov, Alexandra Antoniouk, Vladimir Kalmykov**

**Abstract:** Another aspect of coarse-to-fine approach for technical vision tasks is presented. On the basis of systemological analysis of neurophysiology sources a new hypothesis about visual neuron's functioning is proposed. This hypothesis explains the mechanism and takes into account receptive fields excitatory zones resizing during visual act.

**Keywords:** coarse-to-fine, visual perception, intercellular processes, pattern recognition, variable resolution

Researches in the field of visual perception physiology and creation of information technologies for automatic processing of visual information are fairly interconnected domains of human activity. Indeed, the subject in both disciplines is the study of visual perception. For physiology of vision the subject is the visual perception of humans and animals, while one of the subjects for computer sciences is creation and testing of technical vision means. The progress in one of these domains may initiate the progress in the other.

One of the strategies that solves the problem of great computational complexity in various image processing tasks is so-called "coarse-to-fine approach", i.e. the technique of refining initial data that exclude inappropriate objects or irrelevant ranges of the image on earlier stages of processing in order to apply computationally intensive part of the algorithm to reduced volume of data.

Consider an image consisted of a line of arbitrary text on a one-color background and another one where the same line of text is placed over an arbitrary grating (Fig. 1). The text on (Fig. 1a) can be processed by both statistical and structural methods of recognition. The text on (Fig. 1b) can not be recognized by known methods. At the same time visual perception cope with similar tasks insensibly, seemingly on subconscious level. It is naturally to conjecture that the best resolution for decision making is selected in visual system in the sense of image processing, when the unnecessary details don't arise and the essential parts of objects don't disappear.



**Figure 1. Examples of image with arbitrary text**

It is possible to suggest that it is the processing of observed low resolution images at the beginning of visual act that makes possible consistent visual perception of symbols on different background texture. The text on Fig.1a is recognized successfully by standard OCR program. Processing the text on Fig.1b gives denial of recognition because of inability to determine an object on image. After digitizing these images with six times lower resolution (Fig. 2a, 2b) the text string is recognized successfully on both of them. In this case the coarse-to-fine approach is used to solve the problem that can not be addressed at all by traditional methods.

It was discovered that the sizes of receptive fields' excitatory zones of visual neurons change during visual act, which eventually mean dynamical changes in visual system's resolution. Even though this phenomenon is examined well, the mechanisms for generation of receptive field center width dynamics are unclear and further studying of intracellular processes in visual system is needed.

Proposed model of visual neuron's functioning during action potential generation is developed on the basis of systemological analysis of known ideas about neurons' functioning and results of other researchers. It is an

attempt to explain the mechanism and to take into account receptive fields excitatory zones resizing during visual act.

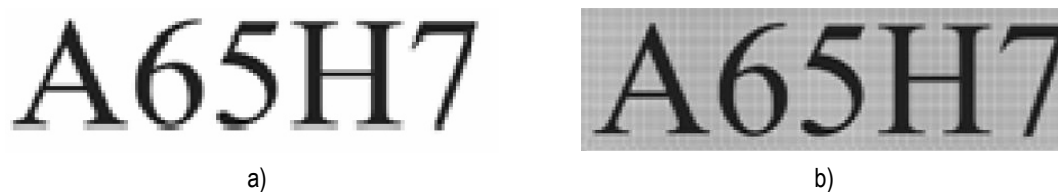


Figure 2. Images from Fig. 1 (a, b) having 6 times lower resolution

Under the influence of excitatory receptors post-synaptic potential and also with no impact of inhibitory receptors the potential-dependent sodium channels start to open and the spike begins. Sodium flow increases and depolarization increases too, resulting in opening more sodium channels, so the intracellular membrane potential increases up to the value of sodium equilibrium potential. Chlorine ions enter cell through membrane pores due to concentration gradient while counteractive electrical potential decreases. After spike ends, the cell membrane potential doesn't return to its initial value and resting potential shifts toward hyperpolarization due to chlorine ions ingress. It is known that the opening threshold of potential-dependent sodium channels increases with the distance growth from axon initial segment. At the beginning of each next pulse this results in potential-dependent sodium channels located outermost from axon hillock stops to open under the post-synaptic potential impact of excitatory receptors, located in the same area as these sodium channels. This means that potential of opening threshold for mentioned sodium channels is greater than initial resting potential plus shift value and they failed to participate in charge accumulation for the next pulse generation.

It is also possible to assume that the distance from excitatory receptor to axon hillock matches distance in the field of view from the point of given receptor to receptive field center. So the "nonparticipation" of some receptors in pulse generation matches lessening of receptive field excitatory zone.

---

## Conclusion

As for now we can state that a coarse-to-fine approach is used spontaneously by researchers in different fields of technical sciences. At the same time some studies are carried out in the domain of neurophysiology, showing the presence of such mechanism in living beings' visual system. A hypothesis explaining the functioning of that mechanism in retinal ganglion cells and LGN neurons was presented.

This hypothesis approval perhaps will enable systematization of coarse-to-fine approach in the field of technical vision; developing of general recommendations and best practices for its application to recognition tasks that can not be resolved at all by traditional methods.

---

## Authors' Information

**Anton Sharypanov** – Institute of cybernetics, prosp. akad. Glushkova 40, 03680, Kiev, Ukraine; e-mail: [\\_sha\\_@ukr.net](mailto:_sha_@ukr.net)

**Alexandra Antoniouk** – Institute of Mathematics NAS Ukraine, Tereshchenkivska, 3, 01 601, Kyiv, Ukraine; e-mail: [antoniouk@imath.kiev.ua](mailto:antoniouk@imath.kiev.ua)

**Vladimir Kalmykov** – Institute of problems of mathematical machines and systems, prosp. akad. Glushkova 42, 03680, Kiev 187, Ukraine; e-mail: [vl.kalmykov@gmail.com](mailto:vl.kalmykov@gmail.com), [kvg@immisp.kiev.ua](mailto:kvg@immisp.kiev.ua)



## MATRIX-EVOLUTIONARY METHOD FOR TIMETABLE FORMING

Vitaliy Snytyuk, Olena Sipko

**Abstract:** In this paper proposed a new method for automatic generation of timetables, based on the use of the objective function preferences of students and teachers, as well as matrix-slatted his presentation.

**Keywords:** timetable, matrix-evolutionary method.

### Introduction

Development of technologies for forming the timetable in universities has more than one decade. On the urgency of solving this problem indicates the number of stakeholders - lecturers, students and controllers. Today are used hundreds of software systems for scheduling, but the task of forming a truly effective timetable remains unresolved. The proposed models and methods are only partially closer to the achievement of the global goal - the creation of an optimal timetable. The main aspect is the need to meet the interests of stakeholders in terms of restrictions on the accompanying resource potential.

### Forming timetable as a matrix-lattice structure

To solve the problem of constructing an effective timetable by the authors earlier proposed the objective function

$$F(r) = \alpha_s \sum_{j=1}^I x_j \chi \{Z_j^v\} + \sum_{j=1}^K y_j \sum_{i=1}^M \chi \{L_i \in T_j\} \sum_{l=1}^{n_j} d_{il}^j \cdot \chi \{Z_{il}^{T_j}\} \rightarrow \max, \quad (1)$$

$$r \in \Omega(P, S, L, A),$$

where  $r$  - timetable,  $\alpha_s, \alpha_L$  - weights indicating the priorities of lecturers and students as subjects of the educational process,  $x_j$  - priorities requirements of students and lecturers,  $Z_j^v$  - demands from groups of students,  $L_i$  - lecturers,  $T_j$  - group of lecturers,  $Z_{il}^{T_j}$  - lecturers preferences,  $d_{il}^j$  - priorities such preferences,  $I$  - the number of students requests,  $K$  - the number of groups of lecturers defined their position, academic degree and academic rank,  $M$  - the number of lecturers,  $n_j$  - the number of lecturers in  $i^{\text{th}}$  group,  $\Omega$  - restrictions area,  $P, L, A$  - set of academic disciplines, lecturers and classrooms, respectively.

Solution of the problem (1) is requested to carry out using evolutionary techniques [Snytyuk, 2013]. To develop a structure of potential solutions, and using (1), we can estimate its prospects. But, unlike many optimization problems, the computation of the objective function as well as the verification of restrictions compliance are a time consuming processes.

In order to optimize the computational process is supposed to use a matrix-evolutionary method. Then the timetable can be represented as a table with some fields

$$Sh = \langle \text{Day}, \text{Time}, \text{Year}, \text{Group}, \text{Course}, \text{Lecturer}, \text{Type}, \text{Room} \rangle, \quad (2)$$

where  $\text{Day}$  - day of the week,  $\text{Time}$  - the number of class,  $\text{Year}$  - student learning year (course),  $\text{Group}$  - the group code (ID),  $\text{Course}$  - the name of academic discipline,  $\text{Lecturer}$  - a lecturer,  $\text{Type}$  - the type of class,  $\text{Room}$  - the number of classroom.

Then visual timetable looks like a three-dimensional lattice structure. Such a structure would be in the form of a rectangular parallelepiped with sides lie on the axes:

$$X_1 = \langle \text{Day} - \text{Time} \rangle, X_2 = \langle \text{Year} - \text{Group} \rangle, X_3 = \langle \text{Classroom} \rangle. \quad (3)$$

The nodes of the lattice parallelepiped will be the value

$$Z = \langle \text{Lecturer} - \text{Course} - \text{Type} \rangle. \quad (4)$$

---

Assuming that the number of school days per week is 6, and the possible number of lessons per day also 6, we find that  $X_1 \in \{1, 2, \dots, 36\}$ . A maximum of courses at universities is 6, the number of groups varies (suppose 200), then in this case, the total number of nodes in a three-dimensional lattice will be 2160000, at provided that the number of potential audiences is 300. Content of the lattice nodes may be different. As a first option there can be a number  $Z \in \{1, 2, \dots, 800000\}$ , considering that the number of teachers 200, items - 1000, and forms of training sessions - 4.

Thus, the schedule will be a collection of fragments  $\{X_1, X_2, X_3, Z\}$ . Since the definition of the structure of a problem potential solution of the creating schedules is important to optimize the computational process, we note that the transformation of the triad  $\langle \text{Lecturer-Course-Type} \rangle$  to an integer should be on some algorithm. Here are its features. In the first stage we write all forms of learning in one subject for specific lecturer. Further, the same for the other object, but the same lecturer. In the next step we write subjects and forms of learning for another lecturer, etc. Note that lecturers are ordered by position, degrees and titles. If these attributes are the same, the lecturers are written in alphabetical order. After receiving the list in the specified format we perform coding, so that the first record will be responsible one, and the last - 800000.

In the second variant, the lattice nodes can be coded only objects and forms of learning, as in the first two attributes, and the database can be uniquely set lecturer. However, this process probably will require more resource that will be installed additionally.

The proposed method of formation of the potential solutions fragment will optimize the computational process using the elements of a computer strategy by ensuring the continuity of obtained solutions. In particular, if it is found that some lecturers can not spend a learning session on the subject, it is more likely that the first thing the next step will be requested to change the shape of a class or subject.

Simultaneously with the construction of a timetables parallelepiped efficiently build another a parallelepiped of the same dimension, but in the lattice nodes which will be unity, indicating that in such day and such time for a group of students in the classroom is lesson and goes to zero if there are no classes. Such a parallelepiped is needed to accelerate the construction of the potential solutions-timetables and check their adequacy, as well as to calculate the values of the objective function.

In contrast to the traditional variant of evolutionary strategy using we offer an emphasis on ensuring the continuity of the search for new potential solutions, which corresponds to the logic of the process of timetable forming: "If the timetable is not satisfactory, then will be most effective its minimal changes."

---

## Conclusion

In the report will be shown a new method of creating timetables at the university, which is based on the development and use of the objective function, in the basis of which are the preferences of lecturers and students. We propose a new approach to the formation of potential solutions-timetables, based on the composite elements using evolutionary strategy, the fuzzy set theory and the analytic hierarchy process. The process of checking constraints and the adequacy of timetables is optimized.

---

## Bibliography

[Snytyuk, 2013] Снитюк В.Е.. Параметрическая оптимизация процесса эволюционного направленного поиска / Матеріали ІІ Міжн. наук.-техн. конф. «Обчислювальний інтелект». – Черкаси. – 2013. – С. 14-15.

---

## Authors' Information

**Vitaliy Snytyuk** – Taras Shevchenko National university of Kyiv, Professor, 03022, 81 Lomonosov Str., Kyiv, Ukraine; e-mail: [snytyuk@gmail.com](mailto:snytyuk@gmail.com);

**Olena Sipko** – Cherkassy State Technological University, Assistant, 18006, 460 Shevchenko Ave, Cherkassy, Ukraine; e-mail: [sipko888@gmail.com](mailto:sipko888@gmail.com)

## A HIERARCHICAL APPROACH TO MULTICRITERIA PROBLEMS

Albert Voronin, Yuriy Ziatdinov, Igor Varlamov

**Abstract:** It is shown, that any multicriteria problem can be represented by a hierarchical system of criteria. Individual properties of the object (alternative) are evaluated at the bottom level of the system, using a criteria vector; and a composition mechanism is used to evaluate the object as a whole at the top level. The problem is solved by the method of nested scalar convolutions of vector-valued criteria. The methodology of the problem solving is based on the complementarity principle by N. Bohr and the theorem of incompleteness by K. Gödel.

**Keywords:** hierarchical structure; nested scalar convolutions; multicriteria approach; decomposition; composition.

### Introduction

The problem of decision making in general view can be represented by the scheme

$$\{\{x\}, Y\} \rightarrow x^* ,$$

where  $\{x\}$  is a set of objects (alternatives);  $Y$  is the function of choice (rule establishing a prefer ability on a set of alternatives);  $x^*$  is the chosen alternatives (one or more).

The function  $Y$  is used to solve the problem of analysis and evaluation of alternatives. On results of estimation the choice of one or a few best alternatives from the given set follows. In decision theory, there are two different approaches to evaluating objects (alternatives) subject to choice. One of them is to evaluate an object as a *whole* and to choose an alternative by comparing objects as *gestalts* (holistic images of objects without detailing their properties). The second approach is detailed elaboration and assessment of various object vectors of properties and making decisions after comparing these properties. If a holistic approach implies choosing  $x^*$  directly using choice function  $Y$ , the *vector approach* requires a mechanism to carry out decomposition of  $Y$  into a set (vector) of the choice functions  $y$ . By decomposition of the choice function  $Y$  is understood its equivalent representation by a certain set of other functions  $y$  which composition is the initial choice function  $Y$ .

Separation of properties of alternatives on the basis of the analysis is the decomposition leading to the hierarchical structure of properties.

Properties, for which there exist objective numerical characteristics, are called *criteria*. The approach of comparison on separate properties, at all its attraction, derivates a serious problem of return transition to required comparison of alternatives as a whole [Voronin, 2013].

### Statement of the Problem

Quality of an alternative is determined by hierarchical system of vectors

$$y^{(j-1)} = \{y_i^{(j-1)}\}_{i=1}^{n^{(j-1)}} , j \in [2, m] ,$$

where  $y^{(j-1)}$  is the vector of criteria on the  $(j-1)$ -th level of the hierarchy, by the components of which the quality of properties of alternatives for the  $j$ -th level is assessed;  $m$  is the amount of levels of the hierarchy;  $n^{(j-1)}$  is the amount of estimated properties on  $(j-1)$ -th level of the hierarchy. The numerical values of  $n$  criteria  $y^{(1)} = y$  of the first level of the hierarchy for the alternative are given.

The same criterion on  $(j-1)$ -th level can participate in the evaluation of several properties of the  $j$ -th level, i.e. in

the hierarchy are possible cross-links. It is clear that  $n^{(1)} = \sum_{i=1}^{n_1} r_i = n$  and  $n^{(m)} = 1$ .

Importance (significance) of each of the components of the criterion of  $(j-1)$ -th level in the evaluation of properties of  $k$ -th level is characterized by a property coefficient of the priority, their set forming the priority vectors system

$$p_{ik}^{(j-1)} = \{p_{ik}^{(j-1)}\}_{k=1}^{n^{(j)}}, j \in [2, m] .$$

**It is required** to find an analytical evaluation  $y^*$  and qualitative evaluation of the effectiveness of this given alternative, and from the alternatives available to choose the best.

---

### The Method of Solution

---

At the study, the approach is used consisting in the creation and simultaneous co-existence of not one but many theoretical models of the same phenomenon, and some of them conceptually contradict each other. However, no one can be neglected, as each describes a property of the phenomenon and none can be taken as a single because it does not express the full range of its properties. Compare the said with the principle of complementarity, introduced into science by Niles Bohr: "... To reproduce the integrity of the phenomenon should be used mutually exclusive "complementary" classes of concepts, each of which can be used in its own, special conditions, but only when taken together, exhaust the definable information." It is the principle of complementarity that allows for separating and then linking these criteria in multicriteria evaluation. Only a full set of individual criteria (vector criterion) enables an adequate assessment of the functioning of a complex system as a manifestation of the contradictory unity of all its properties.

However, this possibility represents only a necessary but not a sufficient condition for the vector evaluation of the entire alternative as a whole.

For a complete evaluation it is necessary to go out from the lower level of the hierarchy and to rise on the following tier, i.e. to carry out an act of criteria composition. Let's compare this with the incompleteness theorem of Kurt Gödel "... In every complex enough not contradictory theory of the first order there is a statement, which by means of the theory is impossible neither to prove, nor to deny. But the self-consistency of a particular theory can be established by means of another, more powerful formal theory of the second order. But then the question of the self-consistency of this second theory arises, and so forth." We can say that Gödel's theorem is a methodological basis for the study of hierarchical structures.

With reference to our problem it means that for an adequate estimation of an alternative as a whole we should solve a task of the criteria composition on levels of hierarchy, consecutively passing from the bottom level up to top.

A scalar convolution of criteria can serve as a tool for the act of composition. The scalar convolution – it is a mathematical technique for data compressing and quantifying its integral properties by a single number.

A scalar convolution on **nonlinear compromise scheme** for the criteria subject to be minimized is proposed [Voronin, 2014]

$$Y[y(x)] = \sum_{k=1}^s p_k A_k [A_k - y_k(x)]^{-1},$$

applied in cases where the decision-maker considers as the preferred those solutions in which the values of individual criteria  $y_k(x)$  are farthest from their limit values,  $A_k$ . This convolution has a number of essential advantages, which include flexibility, universality and analyticity.

The choice of a compromises scheme is made by the decision-maker and appears as explicitly conceptual.

## Nested Scalar Convolutions

It is proposed for analytical evaluation of hierarchical structures to apply a method of nested scalar convolutions. The composition is performed on the "matryoshka principle": the scalar convolutions of the weighted components of vector criteria of lower level serve as the components of the vectors of higher level criteria. Scalar convolution of criteria obtained at the uppermost level is automatically considered as the expression for the analytical evaluation of effectiveness of the entire hierarchical system.

The algorithm for nested scalar convolutions is represented by an iterative sequence of operations of the weighed scalar convolutions of criteria for each level of the hierarchy from the bottom up, taking into account the priority vectors, based on the selected compromise scheme

$$\left\{ \left( y^{(j-1)}, p^{(j-1)} \right) \rightarrow y^{(j)} \right\} j \in [2, m] \quad (1)$$

and the searching and evaluating of effectiveness of the entire hierarchical system (alternative) as a whole is expressed by the problem of determining the scalar convolution of criteria on the top level of the hierarchy:

$$y^* = y^{(m)}.$$

When using the recurrent formula (1) important is the rational choice of the compromise scheme. For the method of nested scalar convolutions the adequate is a nonlinear compromise scheme. It is established that, without loss of generality, a premise for its use is that all the partial criteria were non-negative, were subject to minimization and were limited:

$$0 \leq y_i \leq A_i, A = \{A_i\}_{i=1}^n,$$

where  $A$  is the vector of restrictions on the criteria of the current level of the hierarchy;  $n$  is the amount of them.

Proceeding from (1) the expression to evaluate  $k$ -th property of an alternative for the  $j$ -th level of the hierarchy by using the nonlinear compromise scheme looks like

$$y_k^{(j)} = \sum_{i=1}^{n_k^{(j-1)}} p_{ik}^{(j-1)} \left[ 1 - y_{0ik}^{(j-1)} \right]^{-1}, k \in [1, n^{(j)}], \quad (2)$$

where criteria of the  $(j-1)$ -th level are normalized (reduced to unity). Thus,  $y_{0ik}^{(j-1)}$  are the normalized vector's  $y_0^{(j-1)}$  components involved in the evaluation of properties of the  $k$ -th alternative on the  $j$ -th level of the hierarchy;  $n_k^{(j-1)}$  is their amount;  $n^{(j)}$  is the amount of evaluated properties of the  $j$ -th level.

In the most simple and rather common case the multicriteria problem is formulated and solved without priorities, when decision-makers believe that all the importance parameters for all properties of alternatives are the same. In this case, a simple scalar convolution with the nonlinear trade-offs scheme in a unified form is used.

In order to formula (2) reflected the idea of the nested scalar convolutions method in accordance with the recurrent relation (1), this expression should be normalized, i.e., must be obtained a relative measure such that it were subject to be minimized, and it were the unit for it as the limit value.

The structure of the nonlinear compromise scheme enables normalizing the convolution (2) not to the maximum (which in this case is difficult), but to the minimum value of criteria convolution. Indeed, the ideal values for the criteria that are subject to be minimized are their zero points. Putting in (2)

$$y_{0ik}^{(j-1)} = 0, \forall i \in [1, n_k^{(j-1)}]$$

and taking into account the normalization  $\sum_{i=1}^n p_i = 1$ , we obtain  $y_{k\min}^{(j)} = 1$ .

After calculations and normalizing (reducing to unity), the final expression for the recurrent formula for calculating analytical assessments of the alternatives properties at all levels of the hierarchy becomes

$$y_{0k}^{(j)} = 1 - \left\{ \sum_{i=1}^{n_k^{(j-1)}} p_{ik}^{(j-1)} [1 - y_{0ik}^{(j-1)}]^{-1} \right\}^{-1}, k \in [1, n^{(j)}], j \in [2, m].$$

---

## Conclusion

The foregoing leads to the conclusion that any problem of vector assessment of an alternative can be represented by a hierarchical system of criteria, resulting from the decomposition of alternative properties. The lower level of the hierarchy is an object (alternative) assessment on selected properties, using initial criteria vector, and the upper level is obtained through the mechanism of the composition as a whole object evaluation. Central here is the problem of the composition of criteria for levels of the hierarchy to be solved by the method of nested scalar convolutions.

The methodological basis of an alternative properties decomposition to obtain the initial criteria vector is the Bohr's principle of complementarity. This is a *necessary* condition for vector estimation of alternatives.

The methodology of a criteria composition for levels of the hierarchy is based on the Gödel's theorem of incompleteness. This is a *sufficient* condition for vector estimation of alternatives.

We dare say that above inferences about notions of criteria decomposition and composition can be extended on the more general notions of analysis and synthesis.

---

## Bibliography

[Voronin, 2013] Albert Voronin and Yuri Ziatdinov. Theory and practice of multicriteria decisions: Models, methods, realization [in Russian]. Lambert Academic Publishing, 2013.

[Voronin, 2014] Albert Voronin. Multicriteria Decision-Making. Lambert Academic Publishing, 2014.

---

## Authors' Information

**Voronin Albert** – professor, DrSc(Eng), Professor of Chair of Computer Information Technologies of National Aviation University of Ukraine; e-mail: alnv@voliacable.com.

**Ziatdinov Yuri** – professor, DrSc(Eng), Head of Chair of Computer Information Technologies of National Aviation University of Ukraine; e-mail: oberst@nau.edu.ua.

**Varlamov Igor** – PhD, doctoral of National Defence University of Ukraine named after Ivan Cherniakhovsky; e-mail: igor\_varlamov0@rambler.ru.

## A HIBRID TECHNOLOGIES TO OPTIMIZE CONSTRAINED EOQ MODEL

Olha Yegorova

**Abstract:** in this paper the aspects of program realization of hybrid method of fuzzy simulation and evolutionary technologies of directed optimization to optimize constrained EOQ model are considered.

**Keywords:** EOQ, fuzzy simulation, evolutionary technologies of directed optimization

### Introduction

The economic order quantity models with constraints for deterioration items with two level of trade credit in one replenishment cycle, inspection errors, planned backorders, and sales returns are developed in both crisp and fuzzy environment in [Yegorova, 2014]. Here objectives are to maximize the return on sales. Demand rate, selling prices, interest paid rate, and interest earned rate are considered as fuzzy variables.

In this paper carried out the simulation technique to estimate the fuzzy parameters and solution methodology for fuzzy expected value model along with theoretical results to identify global optimal solution.

### Model formulation

Fuzzy expected value model constructed as follows

$$E[ROS_i(I_b, I_m, I_s)] = \frac{\begin{pmatrix} E[(R_s)_i] - K - E[(P_c)_i] - E[C_{si}] - \\ -E[C_{rg}] - E[C_h] - E[C_d] - E[C_s] - E[C_{un}] - \\ -E[C_{adj}] - E[C_{md}] - E[IP_i] + E[IE_i] \end{pmatrix}}{E[(R_s)_i]}, \quad i = 1, 2, \dots, 14, \quad (1)$$

subject to the constraints

$$E_{\min} \leq E[\tilde{P}_c] \leq E_{\max}, \quad b \cdot c \cdot d \cdot \lambda \cdot (E[\tilde{t}_1] + E[\tilde{t}_2]) \leq W, \quad E[\tilde{C}_h] < E[\tilde{P}_c],$$

$$P \left( a_2 \leq 1 - \frac{u+1}{\beta \cdot \lambda (E[\tilde{t}_1] + E[\tilde{t}_2])} \right) \leq p_s,$$

where  $R_s$  – sale revenue,  $K$  – the ordering cost,  $P_c$  – the purchase cost,  $C_{si}$  – the (inspection) screening cost,  $C_{rg}$  – the cost for return the rejection items to supplier,  $C_h$  – the inventory holding cost per cycle,  $C_d$  – the deterioration cost,  $C_s$  – the shortage cost per cycle due to backlog,  $C_{un}$  – the opportunity cost due to lost sale per cycle,  $C_{adj}$  – the cost of accepting a defective items,  $C_{md}$  – the cost of rejection a non-defective items,  $IP$  – interest paid,  $IE$  – interest earned.

### Fuzzy simulation technique

In model (1) it needs to find the appropriate vector  $(I_b, I_m, I_s)$  such that satisfies the constraints and reaches it maximal value. For the fixed value of  $(I_b, I_m, I_s)$  a fuzzy simulation technique [Taleizadeh, 2013] is employed to estimate the fuzzy parameters such as demand, selling prices, interest paid rate and interest earned rate.

Step 1. Set  $E=0$  and initialized  $G$  and  $O$ .

Step 2. We randomly generate sequences  $(D^g, p_v^g, p_b^g, i_p^g, i_w^g, i_e^g)$  from the  $\alpha$ -level sets of fuzzy variables  $\tilde{D}$ ,

$\tilde{p}_v, \tilde{p}_b, \tilde{i}_p, \tilde{i}_w, \tilde{i}_e, g = 1, 2, \dots, G$ , where  $\alpha$  is a sufficiently small positive number.

Step 3. Calculate  $ROS_i(I_b, I_m, I_s, D^g, p_v^g, p_b^g, i_p^g, i_w^g, i_e^g)$  for  $g = 1, 2, \dots, G, i = 1, 2, \dots, 14$ .

Step 4. Set  $a_i = ROS_i(I_b, I_m, I_s, D^1, p_v^1, p_b^1, i_p^1, i_w^1, i_e^1) \wedge \dots \wedge ROS_i(I_b, I_m, I_s, D^g, p_v^g, p_b^g, i_p^g, i_w^g, i_e^g)$ ,

$$b_i = ROS_i(I_b, I_m, I_s, D^1, p_v^1, p_b^1, i_p^1, i_w^1, i_e^1) \vee \dots \vee ROS_i(I_b, I_m, I_s, D^g, p_v^g, p_b^g, i_p^g, i_w^g, i_e^g)$$

Step 5. Randomly generate  $s_i$  from  $[a_i, b_i]$

Step 6. If  $s_i \geq 0$ , then  $E_i \leftarrow E_i + Cr\{ROS_i(I_b, I_m, I_s, \tilde{D}, \tilde{p}_v, \tilde{p}_b, \tilde{i}_p, \tilde{i}_w, \tilde{i}_e) \geq s_i\}$ , where

$$Cr\{ROS_i(I_b, I_m, I_s, D_j, p_{vj}, p_{bj}, i_{pj}, i_{wj}, i_{ej}) \geq s_i\} = \frac{1}{2} \left( \begin{array}{l} \text{Max}_{j=1,2,\dots,0} \{ \mu_{ij} | ROS_i(I_b, I_m, I_s, D_j, p_{vj}, p_{bj}, i_{pj}, i_{wj}, i_{ej}) \geq s_i \} + 1 - \\ - \text{Max}_{j=1,2,\dots,0} \{ \mu_{ij} | ROS_i(I_b, I_m, I_s, D_j, p_{vj}, p_{bj}, i_{pj}, i_{wj}, i_{ej}) < s_i \} \end{array} \right)$$

Otherwise,  $E_i \leftarrow E_i - Cr\{ROS_i(I_b, I_m, I_s, \tilde{D}, \tilde{p}_v, \tilde{p}_b, \tilde{i}_p, \tilde{i}_w, \tilde{i}_e) \leq s_i\}$ , where

$$Cr\{ROS_i(I_b, I_m, I_s, D_j, p_{vj}, p_{bj}, i_{pj}, i_{wj}, i_{ej}) \leq s_i\} = \frac{1}{2} \left( \begin{array}{l} \text{Max}_{j=1,2,\dots,0} \{ \mu_{ij} | ROS_i(I_b, I_m, I_s, D_j, p_{vj}, p_{bj}, i_{pj}, i_{wj}, i_{ej}) \leq s_i \} + 1 - \\ - \text{Max}_{j=1,2,\dots,0} \{ \mu_{ij} | ROS_i(I_b, I_m, I_s, D_j, p_{vj}, p_{bj}, i_{pj}, i_{wj}, i_{ej}) > s_i \} \end{array} \right)$$

Step 7. Repeat 5 and 6 for 0 times.

Step 8. Calculate  $E[ROS_i(I_b, I_m, I_s, \tilde{D}, \tilde{p}_v, \tilde{p}_b, \tilde{i}_p, \tilde{i}_w, \tilde{i}_e)] = a_i \vee 0 + b_i \wedge 0 + E_i \frac{b_i - a_i}{O}$ .

Here in order to solve the EVM, we choose Evolutionary Technologies of Directed Optimization [Snytyuk, 2004] as the foundation to design an algorithm which integrates fuzzy simulation and Evolutionary Technologies of Directed Optimization, where the fuzzy simulation is employed to estimate the maximal revenue on sales, and Evolutionary Technologies of Directed Optimization is used to find the optimal solution.

---

## Conclusion

The results of experiments and comparative analysis of classical genetic algorithm and the developed method were shown to reduce search time by 7-12% and increase the accuracy of the result.

---

## Bibliography

- [Yegorova, 2014] O. Yegorova. Fuzzy economic order quantity model. In: Proceedings 16-th International Conference SAIT 2014, Kyiv, Ukraine, May 26-30, 190-191.
- [Yegorova, 2014] O. Yegorova. Fuzzy expected value model with inspection errors and two level of trade credit in one replenishment cycle In: Information Models and Analyses, Vol. 3, № 1, 37-52.
- [Taleizadeh, 2013] A. A. Taleizadeh, S.T. Akhvan Niaka, M.-B. Aryanezhad. A hybrid method of fuzzy simulation and genetic algorithm to optimize constrained inventory control systems with stochastic replenishments and fuzzy demand. In: Information Sciences, 220, 425-441.
- [Snytyuk, 2004] V. E. Snytyuk. Compositional overcoming uncertainty in nonlinear multivariable optimization problems. In: Artificial Intelligence, 4, 207-210.

---

## Authors' Information

**Olha Yegorova** – assistant; Cherkasy State Technological University, Shevchenko Blvd., bl.460, Cherkasy, 18006, Ukraine; e-mail: [yegorovaov@gmail.com](mailto:yegorovaov@gmail.com)



## ПОСТРОЕНИЕ ПРОЦЕССА УДАЛЕННОГО ИНТЕЛЛЕКТУАЛЬНОГО ОБУЧЕНИЯ

Александр Бабенко

**Аннотация:** Интеллектуальное удаленное обучение строится на платформе с клиент-серверной архитектурой. Основной моделью обучения является функция знаний, которая строится при обработке входящей информации. Предлагается реализовать классификацию потенциального студента по типу основной модальности восприятия информации и использовать самообучение дистанционной системы интеллектуального обучения.

**Ключевые слова:** интеллектуальное обучение, обработка входящей информации, внутренние процессы, типы восприятия.

### Введение

Объектом исследования является процесс интеллектуализации дистанционного обучения. Предметом исследования является построение процесса удаленного интеллектуального обучения, самообучение системы дистанционного обучения.

Классический процесс обучения включает в себя дискретное распределение времени для лекций, практических занятий и прохождения экзаменов, что требует изменения нормального ритма жизни, дополнительных транспортных расходов. Кроме того, уроки усреднены и не принимают во внимание психологические особенности тех, кто обучается. Третий негативный аспект классического образования - субъективная оценка знаний. Выходом из этой ситуации является дистанционное интеллектуальное обучение. Удаленное интеллектуальное обучение позволит учиться в любое время и в любом месте, имея доступ к сети Интернет и приложения на рабочей станции.

### Обучаемый-система-обучаемый

Проектируемый процесс обучения является интеллектуальным, включая постоянную обратную связь – обучаемый-система-обучаемый. Чтобы охватить наибольшее количество желающих, в обучении наиболее эффективно использовать систему клиент-серверного типа. Таким образом, обработка данных осуществляется дистанционно. Во всех случаях, когда есть подключение к сети интернет клиента, установленного на рабочих станциях, обучаемые имеют доступ к обучающей информации и, собственно, к процессу обучения. В процессе обучения происходит формирование функции (модели) знаний обучаемого. Рисунок 1 иллюстрирует процесс интеллектуального обучения.

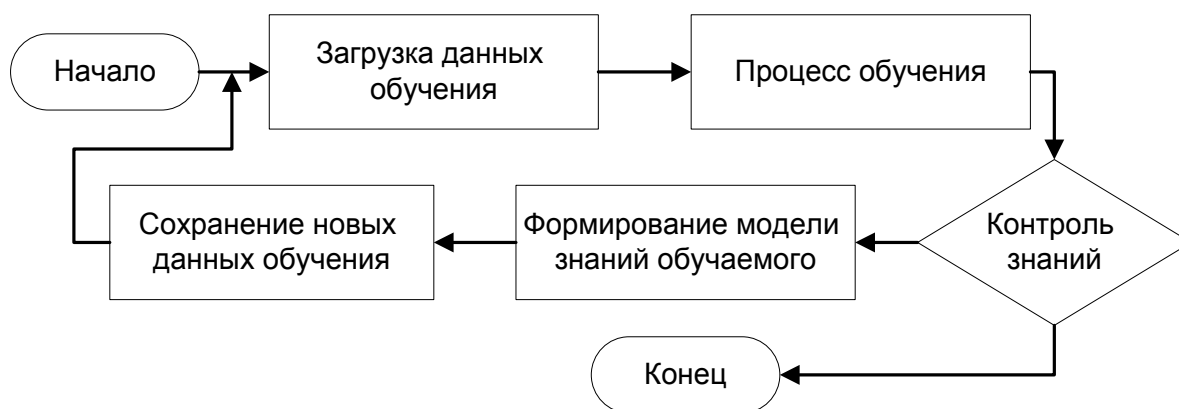


Рис. 1. Процесс интеллектуального обучения

Процесс обучения будет проходить более эффективно при представлении данных обучающимся в зависимости от типа основной модальности восприятия информации у обучаемого.

Всех людей можно разделить на четыре категории по типу основной модальности восприятия: визуалы, аудиалы, кинестетики и дискретны (табл.1), однако необходимо заметить, что очень трудно найти людей с ярковыраженным только одним доминирующим типом восприятия, чаще всего будет наблюдаться комбинация более выраженных нескольких типов восприятия информации.

Таблица 1. Типы восприятия

Тип Восприятия	Характеристика типа восприятия	Процент людей с основным типом восприятия
Визуальный (V)	Информация воспринимается через образы, картинки	35
Аудиальный (A)	Информация воспринимается через звук	20
Кинестетический (K)	Информация воспринимается через действие и физический контакт	40
Дискретный (D)	Информация воспринимается путем логического осмысления	5

Для достижения максимальной точности оценивания типа восприятия, переоценка основного типа происходит на каждом этапе обучения (рис. 2) [1].

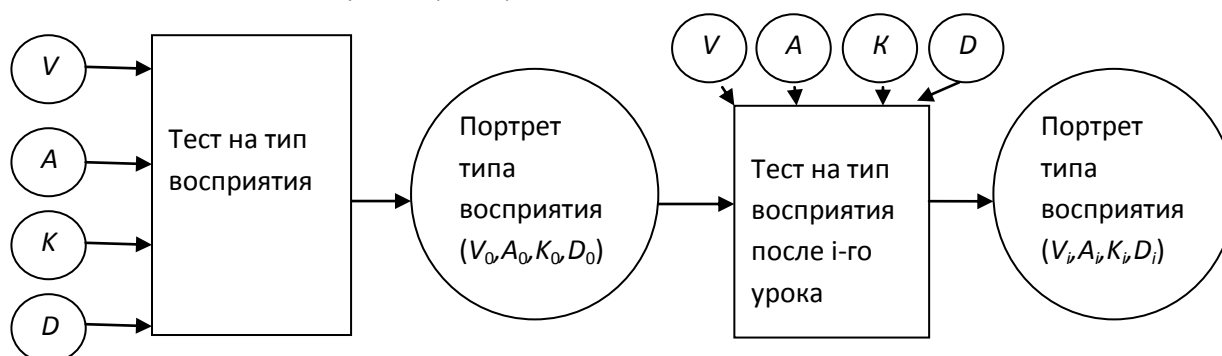


Рисунок 2. Динамика портрета типа восприятия

В систему вводятся параметры -  $X_{ij}$ ,  $X \in [V, A, K, D]$ ,  $i$  – шаг процесса обучения,  $i \in [0..n]$ , 0 характеризует начальный этап обучения;  $j$  – уровень достижений в определенном направлении,  $j \in [0..1]$ , 1 соответствует абсолютному знанию. Значения  $V, A, K, D$  связаны с типами восприятия. Перед началом процесса обучения устанавливается цель обучения (какие разделы необходимо изучить, какой уровень достигнуть) и время, которое можно посвятить обучению. Модель знаний включает в себя психологический портрет обучаемого, который обновляется с прохождением каждой стадии обучения, и выражаются функцией знаний [1].

### Обучаемый-система-эксперт

Следует рассмотреть сопровождение процесса обучения экспертом. Есть два действующих лица: учащийся, который нацелен на изучение необходимого материала, удаленная система интеллектуального обучения, основанная на базе знаний (например, в области иностранных языков), которая обеспечивает учебный материал, планирование обучения и принятия решения об успешном завершении прохождения данного материала. Третьим действующим лицом является эксперт (экспертная группа), которая осуществляет контроль функционирования системы интеллектуального обучения и принимает решения относительно совершенствования базы знаний системы [2,3].

На ранних стадиях работы системы риск принятия системой неправильного решения будет относительно высоким, так как изначально невозможно предусмотреть все возможные варианты поведения учащегося и для сведения к минимуму риска неправильного определения направления дальнейшего изучения материала необходимо иметь прямую и обратную связи с экспертом. Эксперт вносит необходимые

изменения базу знаний с тем, чтобы уменьшить риск принятия системой неправильного решения, а также, при необходимости, осуществляет коррекцию плана обучения студента, сгенерированного программным комплексом.

Для оценки текущего уровня качества интеллектуальной деятельности системы, эксперту необходимо анализировать отчеты, которые включают информацию о решениях, принятых системой интеллектуального обучения при генерации учебных планов. В базе знаний хранятся все сведения о всех инцидентах с неправильной оценкой, чтобы избежать принятия неправильных решений в последующей работе системы [4,5]. Рисунок 3 иллюстрирует взаимодействие студент-система-эксперт.

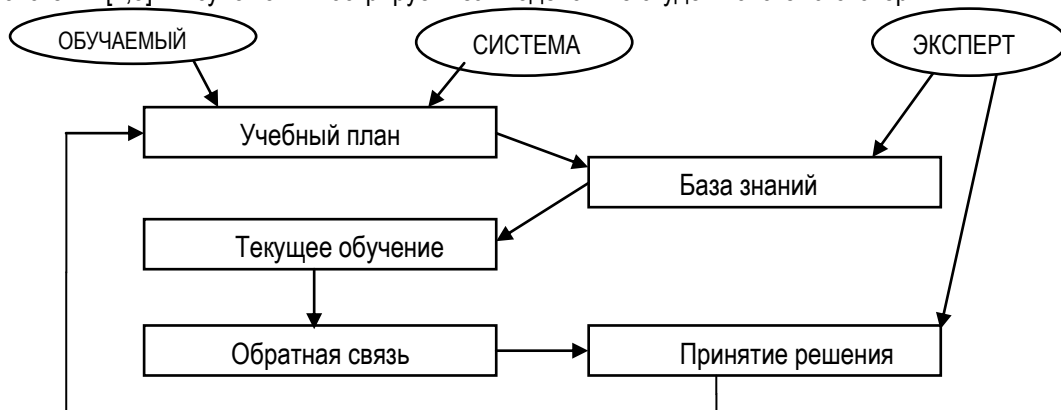


Рисунок 3. Взаимодействие обучаемый-система-эксперт

## Выводы

Клиент-серверный подход позволяет обучаться вне зависимости от местоположения учащегося и времени проведения занятий. Процесс построен на постоянной коррекции функции знаний на основе обратной связи обучаемый-система. В создаваемой удаленной системе интеллектуального обучения для интенсификации процесса обучения используется классификация студентов по психологическим признакам. Самообучение системы проходит под контролем эксперта для совершенствования процесса обучения.

## Библиография

1. Бабенко О.С. Интеллектуальная система когнитивного обучения иностранному языку и алгоритм ее работы // Интеллектуальний аналіз інформації: наук.праці XIII міжн.наук.конф. IAI-2013 ім. Т.А.Таран (15-17 травня 2013р., Київ).-К., 2013 с.4-12.
2. Клещев А.С., Черняховская М.Ю., Шалфеева Е.А. Парадигма автоматизации интеллектуальной профессиональной деятельности. Часть 1. Особенности интеллектуальной профессиональной деятельности
3. Клещев А.С., Черняховская М.Ю., Шалфеева Е.А. Парадигма автоматизации интеллектуальной профессиональной деятельности. Часть 2. Парадигма автоматизации отрасли.
4. Клещев А.С., Шалфеева Е.А. Системный анализ при автоматизации интеллектуальной профессиональной деятельности // XIII Национальная конференция по искусственному интеллекту с международным участием «КИИ-2012», 16–20 октября 2012 г. ISBN: 978-5-361-00182-8 Труды конференции, т.2. - Белгород: Изд-во БГТУ, 2012. - С.128–135.
5. Клещев А.С., Грибова В.В. Управление интеллектуальными системами. Известия РАН. // Теории и системы управления. 2010. № 6. - С. 122-137.

## Информация об авторе

**Александр Бабенко** – МНУЦ ИТС НАНУ, аспирант, Киев, Украина; e-mail: [ukritalia@bigmir.net](mailto:ukritalia@bigmir.net)



ISSN 1313-0087 (printed)  
ISSN 1313-1206 (online)