

# **СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ**

**Научный журнал Российской академии наук  
(издается под руководством Отделения нанотехнологий  
и информационных технологий РАН)**

Издается с 1989 года

Журнал выходит ежеквартально

**Учредитель:**

**Федеральный исследовательский центр  
«Информатика и управление» Российской академии наук**

## **РЕДАКЦИОННЫЙ СОВЕТ**

академик РАН И. А. Соколов — председатель Редакционного совета

академик РАН Г. И. Савин

академик РАН А. Л. Стемпковский

член-корреспондент РАН Ю. Б. Зубарев

профессор Ш. Долев (S. Dolev, Beer-Sheva, Israel)

профессор Ю. Кабанов (Yu. Kabanov, Besancon, France)

профессор В. Ротарь (V. Rotar, San-Diego, USA)

профессор М. Финкельштейн (M. Finkelstein, Bloemfontein, South Africa)

## **РЕДАКЦИОННАЯ КОЛЛЕГИЯ**

академик РАН И. А. Соколов — главный редактор

профессор, д.ф.-м.н. С. Я. Шоргин — заместитель главного редактора

д.т.н. В. Н. Захаров

д.ф.-м.н. В. И. Синицын

проф., д.ф.-м.н. А. И. Зейфман

проф., д.т.н. И. Н. Синицын

проф., д.т.н. В. Д. Ильин

проф., д.ф.-м.н. В. Г. Ушаков

проф., д.т.н. К. К. Колин

к.ф.-м.н. А. К. Горшенин — отв. секретарь

проф., д.ф.-м.н. В. Ю. Королев

к.ф.-м.н. С. А. Христочевский

к.ф.-м.н. Р. В. Разумчик

## **Редакция**

к.ф.-м.н. Е. Н. Арутюнов

к.ф.-м.н. Р. В. Разумчик

С. Н. Стригина

© Федеральный исследовательский центр «Информатика  
и управление» Российской академии наук, 2020

Журнал включен в базу данных Russian Science Citation Index (RSCI),  
интегрированную с Web of Science

Журнал входит в систему Российского индекса научного цитирования (РИНЦ)

Журнал включен в базу данных CrossRef (систему DOI — Digital Object Identifier),  
в базу данных Ulrich's periodicals directory

и в информационную систему «Общероссийский математический портал Math-Net.Ru»

Журнал реферируется в «Реферативном журнале» ВИНТИ  
и в системе Google Scholar

Журнал включен в сформированный Минобрнауки России Перечень рецензируемых научных  
изданий, в которых должны быть опубликованы основные научные результаты диссертаций  
на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук

<http://www.ipiran.ru/journal/collected>

# **СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ**

**Том 30 № 2 Год 2020**

## **СОДЕРЖАНИЕ**

Устойчивость самосинхронных комбинационных схем  
к кратковременным логическим сбоям

**Ю. А. Степченков, Ю. Г. Дьяченко,  
Ю. В. Рождественский, Н. В. Морозов,  
Д. Ю. Степченков, Д. Ю. Дьяченко**

**4**

Универсальный функциональный метод анализа больших  
самосинхронных схем

**Л. П. Плеханов, В. Н. Захаров**

**11**

Предотвращение массового вывода из строя слабозащищенных  
мобильных систем информационной поддержки цифровой  
экономики

**А. А. Грушо, Н. А. Грушо, В. В. Сенчило,  
Е. Е. Тимонина**

**21**

Определение ценности и защита инсайдерской информации

**А. С. Кабанов, А. А. Водолаженко**

**31**

Multidisciplinary neuroinformatics problems for execution  
in distributed computing infrastructures

**D. Y. Kovalev, I. A. Shanin, and E. M. Tirikov**

**43**

Архитектура платформы для проведения движимых гипотезами  
виртуальных экспериментов

**Д. Ю. Ковалев, Е. А. Тарасов, В. Н. Захаров,  
Н. М. Филимонов**

**56**

Подходы к интеграции прикладных концептуальных схем  
в составе унифицированной геонтологии

**С. К. Дулин, Д. А. Никишин**

**68**

Функциональная структура сплоченной гибридной  
интеллектуальной многоагентной системы

**С. В. Листопад**

**78**

Проблема задания коэффициентов готовности диагностических  
систем в проекте ИТЭР: постановка задачи и алгоритм решения

**Г. М. Коновалов**

**89**

# **СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ**

Том 30 № 2 Год 2020

## **СОДЕРЖАНИЕ**

Некоторые вопросы программного определения центров обработки данных

**В. Б. Егоров** **103**

Задачи базы данных фразеологического словаря и стадии ее проектирования

**В. В. Вакуленко, А. А. Гончаров, А. А. Дурново,  
И. М. Зацман** **113**

Нестабильность нейронного машинного перевода

**А. Ю. Егорова, И. М. Зацман, В. В. Косарик,  
В. А. Нуриев** **124**

Методический подход к оценке вклада информационных систем в эффективность организационных систем

**А. А. Зацаринный, Ю. С. Ионенков** **136**

Информационно-аналитический ситуационный центр стратегического планирования в области обеспечения национальной безопасности: перспективные архитектурные решения

**А. В. Босов, А. П. Сучков** **145**

Состав, структура и методы анализа данных систем управления научными сервисами

**А. П. Сучков** **163**

Модель функционирования системы автоматизированной оценки достоверности конкретно-исторических фактов

**И. М. Адамович, О. И. Волков** **177**

Персональные и корпоративные электронные банки в технологиях товарно-денежного обращения

**А. В. Ильин, В. Д. Ильин** **187**

## **ИСТОРИЯ. ПАМЯТНЫЕ ДАТЫ. СОБЫТИЯ**

У истоков российской информатики

**И. А. Соколов, А. А. Зацаринный, В. Н. Захаров** **195**

Об авторах **203**

Правила подготовки рукописей статей **207**

Requirements for manuscripts **211**

## УСТОЙЧИВОСТЬ САМОСИНХРОННЫХ КОМБИНАЦИОННЫХ СХЕМ К КРАТКОВРЕМЕННЫМ ЛОГИЧЕСКИМ СБОЯМ\*

*Ю. А. Степченков<sup>1</sup>, Ю. Г. Дьяченко<sup>2</sup>, Ю. В. Рождественский<sup>3</sup>,  
Н. В. Морозов<sup>4</sup>, Д. Ю. Степченков<sup>5</sup>, Д. Ю. Дьяченко<sup>6</sup>*

**Аннотация:** Рассматривается вопрос устойчивости самосинхронных (СС) комбинационных схем, изготовленных по технологии комплементарный металл – диэлектрик – полупроводник (КМДП), к кратковременным логическим сбоям (ЛС), вызываемым внешними причинами или внутренними помехами, не приводящими к разрушению полупроводниковых структур. Обсуждаются последствия воздействия физических причин, приводящих к ЛС в микросхеме, изготовленной по КМДП-технологии с проектными нормами 65 нм и ниже. Введена классификация ЛС в СС комбинационных КМДП-схемах в зависимости от времени их появления и типа сбоя. Самосинхронные схемы имеют более высокую степень устойчивости к кратковременным ЛС, чем их синхронные аналоги, благодаря двухфазной дисциплине работы, запрограммированному взаимодействию и парафазному кодированию информационных сигналов. Предложены схемотехнические и топологические методы, обеспечивающие снижение чувствительности СС комбинационных КМДП-схем к логическим сбоям за счет гарантированного отсутствия биполярного влияния источника ЛС на элементы, формирующие парафазные сигналы, и на их трассы в топологии схемы.

**Ключевые слова:** самосинхронная схема; логический сбой; сбоестойчивость; КМДП; рабочая фаза; спейсер; топология

**DOI:** 10.14357/08696527200201

### 1 Введение

Широкое использование интегральных КМДП-микросхем в аппаратуре, работающей в условиях воздействия неблагоприятных факторов и при ограничениях

\* Исследование выполнено за счет гранта Российского научного фонда (проект 19-11-00334).

<sup>1</sup> Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, YStepchenkov@ipiran.ru

<sup>2</sup> Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, diaura@mail.ru

<sup>3</sup> Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, YRogdest@ipiran.ru

<sup>4</sup> Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, NMorozov@ipiran.ru

<sup>5</sup> Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, stepchenkov@mail.ru

<sup>6</sup> Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, diaden87@gmail.com

ченном напряжении питания, ставит задачу обеспечения их надежности и сбоестойчивости. Неблагоприятные факторы способны вызвать критический ЛС в комбинационной схеме. Поэтому задача анализа появления и распространения ЛС в цифровых схемах является актуальной.

В синхронных схемах даже кратковременное изменение сигнала, произошедшее непосредственно перед активным фронтом тактового импульса, с большой вероятностью запишется в регистр и испортит данные. Самосинхронные схемы [1, 2] способны детектировать появление ЛС и в большинстве случаев либо замаскировать его, либо приостановить обработку данных до его окончания.

В данной статье рассматривается проблема сбоестойчивости комбинационных цифровых СС-схем, изготовленных по современной КМДП-технологии с проектными нормами 65 нм и ниже, и предлагаются способы их защиты от кратковременных ЛС.

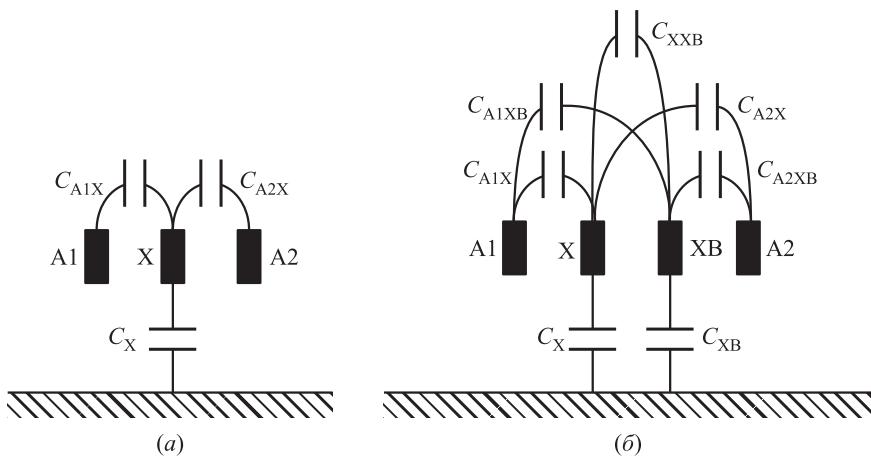
## 2 Типы логических сбоев в комбинационных самосинхронных схемах

Рассматриваемые СС-схемы используют парафазное со спейсером (ПФС) кодирование сигналов и строго соблюдают принцип запрос-ответного взаимодействия соседних в тракте обработки информации СС функциональных блоков. Любую СС-схему можно структурно разделить на две части: информационную, реализующую алгоритм обработки данных, и индикаторную, подтверждающую окончание переключения информационной части в текущую фазу работы. Появление ЛС в этих частях СС-схемы по-разному влияет на работоспособность СС-схемы.

### 2.1 Механизмы появления логических сбоев в КМДП-схемах

Существуют два основных механизма возникновения ЛС: ионизационный [3] и емкостной [4]. Ионизационный механизм запускается ядерной частицей или космическим лучом, проникающим в полупроводник. Пролет ядерной частицы через тело полупроводника вызывает импульсный ток ионизации длительностью  $10^{-11}\text{--}10^{-8}$  с [5] и может вызвать сбойное переключение элементов схемы.

Второй источник ЛС — наводки через паразитную емкостную связь трасс межсоединений и шин питания и «земли» в топологической реализации КМДП БИС (больших интегральных схем). Рисунок 1 демонстрирует упрощенную модель перекрестных наводок для унарного сигнала X [4] и ПФС-сигнала (X, XB). Трасса-«жертва» унарного сигнала X связана паразитными емкостями  $C_{A1X}$  и  $C_{A2X}$  с соседними трассами-«агрессорами» A1 и A2 и паразитной емкостью  $C_X$  — с подложкой (рис. 1, а). В наихудшем случае, при одновременном переключении трасс-«агрессоров» в одном направлении, на трассе-«жертве» индуцируется импульс напряжения, амплитуда которого пропорциональна отношению  $(C_{A1X} + C_{A2X})/(C_{A1X} + C_{A2X} + C_X)$ . Его длительность зависит от длительности фронта (спада) сигналов на трассах-«агрессорах» и нагрузочной



**Рис. 1** Модель перекрестных наводок для унарного сигнала (а) и ПФС-сигнала (б)

способности драйвера сигнала X. Обычно она меньше длительности ЛС, вызванного ионизационным механизмом.

Если компоненты одного ПФС-сигнала имеют разных соседей, то в наихудшем случае возможно одновременное появление на них импульсов напряжения разной полярности. Это может вызвать переключение атакованного ПФС-сигнала из корректного рабочего состояния (КРС) в противоположное — сбойное рабочее состояние (СРС).

Аналогичным образом действуют источники глобальных наводок: шумы по шинам питания и земли, шум по подложке, электромагнитный импульс от внешнего источника. Однако они, как правило, симметрично влияют на оба компонента ПФС-сигнала. В результате ПФС-сигнал из КРС может переключиться только в спейсер или его инверсию — антиспейсер (АС).

Из-за меньшей длительности наводки через паразитную емкостную связь менее опасны, чем сбои, вызванные ядерной частицей или космическими лучами. Но частота их появления сравнима с частотой работы схемы, поэтому они не менее важны.

## 2.2 Классификация логических сбоев в комбинационных самосинхронных схемах

Все возможные случаи ЛС в информационной части комбинационных СС-схем показаны в таблице. Классификация ЛС в комбинационных СС-схемах на основании их типа и времени появления показана на рис. 2. Однако вероятность появления событий, соответствующих случаям 8 и 9, можно свести к нулю с помощью следующих топологических методов:

## Возможные изменения ПФС-сигнала из-за ЛС

№ случая ЛС	Нулевой спейсер		Единичный спейсер	
	Исходное состояние	Состояние после ЛС	Исходное состояние	Состояние после ЛС
1	00	01	11	01
2	00	10	11	10
3	00	11	11	00
4	01	11	01	00
5	10	11	10	00
6	01	00	01	11
7	10	00	10	11
8	01	10	01	10
9	10	01	10	01

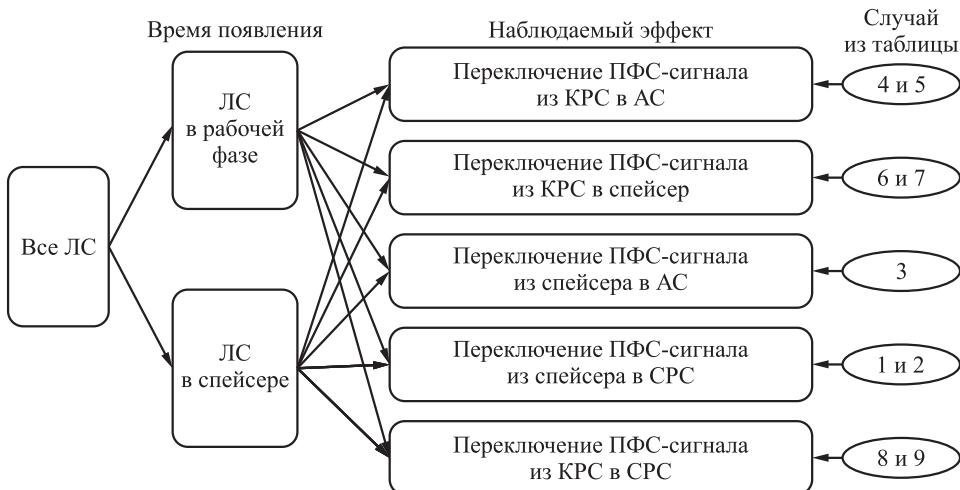


Рис. 2 Классификация ЛС в комбинационных СС-схемах

- принудительного расположения пары элементов, формирующих ПФС-сигнал, рядом друг с другом [3];
- принципа «витой пары» при разводке ПФС-сигнала (использование соседних трасс в каждом топологическом слое, см. рис. 1, б), тогда воздействие трасс-«агрессоров» на обе его части будет симметричным.

Таблица описывает ЛС только ПФС-сигналов. Однако на унарных сигналах индикаторной подсхемы СС-схемы также может появиться ЛС. Он ошибочно подтвердит завершение перехода индицируемой схемы в рабочую (спейсерную) фазу и спровоцирует преждевременное переключение ее входов в спейсер (рабочее состояние). Возникнет критический останов. Решает эту проблему DICE-реализация выходного гистерезисного триггера ( $\Gamma$ -триггера) [6] в каждой индикаторной

подсхеме и всех Г-триггеров схемы управления. Тогда ЛС в схеме управления приведет лишь к задержке в работе СС-схемы.

Представленная на рис. 2 классификация кратковременных ЛС служит основой для вычисления количественных оценок вероятности повреждения информации, обрабатываемой комбинационной СС-схемой, вследствие ЛС.

### 3 Заключение

Благодаря ПФС-кодированию информационных сигналов, строгой индикации окончания переключений всех элементов схемы и запрос-ответному взаимодействию СС-схемы обладают высокой устойчивостью к одиночным ЛС, вызванным внешними и внутренними источниками.

Принудительное размещение пары элементов, формирующих ПФС-сигнал, рядом друг с другом в топологии СС-схемы и использование соседних трасс в каждом топологическом слое для разводки ПФС-сигнала (по типу «витой пары») делают ЛС типа «переключение ПФС-сигнала из корректного рабочего состояния в инверсное рабочее состояние» принципиально невозможным.

Повышение устойчивости индикаторной подсхемы СС-схемы к одиночным ЛС обеспечивается DICE-реализацией выходного Г-триггера в каждой индикаторной подсхеме и всех Г-триггеров схемы управления.

### Литература

1. Kishinevsky M., Kondratyev A., Taubin A., Varshavsky V. Concurrent hardware: The theory and practice of self-timed design. — J. Wiley & Sons, 1994. 368 p.
2. Степченков Ю. А., Дьяченко Ю. Г., Горелкин Г. А. Самосинхронные схемы — будущее микроэлектроники // Вопросы радиоэлектроники, 2011. № 2. С. 153–184.
3. Stepchenkov Y. A., Kamenskih A. N., Diachenko Y. G., Rogdestvenski Y. V., Diachenko D. Y. Fault-tolerance of self-timed circuits // 10th Conference (International) on Dependable Systems, Services, and Technologies Proceedings. — IEEE, 2019. P. 41–44. doi: 10.1109/DESSERT.2019.8770047.
4. Taubin A., Kondratyev A., Cortadella J., Lavagno L. Behavioral transformations to increase noise immunity in asynchronous specifications // 5th Symposium (International) on Advanced Research in Asynchronous Circuits and Systems Proceedings. — IEEE, 1999. P. 36–47. doi: 10.1109/ASYNC.1999.761521.
5. Eaton P., Benedetto J., Mavis D., Avery K., Sibley M., Gadlage M., Turflinger T. Single event transient pulse width measurements using a variable temporal latch technique // IEEE T. Nucl. Sci., 2004. Vol. 51. Iss. 6. P. 3365–3368. doi: 10.1109/TNS.2004.840020.
6. Danilov I. A., Gorbunov M. S., Shnaider A. I., Balbekov A. O., Rogatkin Y. B., Bobkov S. G. DICE-based Muller C-elements for soft error tolerant asynchronous ICs // 16th European Conference on Radiation and Its Effects on Components and Systems. — IEEE, 2016. Art. ID: F4. P. 1–4. doi: 10.1109/RADECS.2016.8093145.

Поступила в редакцию 17.02.19

## SELF-TIMED COMBINATIONAL CIRCUIT TOLERANCE TO SHORT-TERM SOFT ERRORS

**Yu. A. Stepchenkov, Yu. G. Diachenko, Yu. V. Rogdestvenski, N. V. Morozov,  
D. Yu. Stepchenkov, and D. Yu. Diachenko**

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** The paper considers self-timed (ST) complementary metal–oxide–semiconductor (CMOS) combinational circuit tolerance to short-term soft errors caused by the external sources or internal noises that do not lead to semiconductor structure destruction. The paper discusses the consequences of physical causes impact, leading to soft errors in a chip manufactured by the 65-nanometer and below CMOS process. It introduces soft error classification in CMOS ST combinational circuits depending on their appearance time and the type of failure. Self-timed circuits have a higher degree of resistance to short-term soft errors than their synchronous counterparts due to the two-phase operation discipline, request-acknowledge interaction, and dual-rail information signal coding. The paper proposes circuitry and layout methods ensuring the lowering of CMOS ST combinational circuit sensitivity to soft errors due to the guaranteed absence of the bipolar influence of the soft error source on the cells forming dual-rail signals and on their wires in the circuit layout.

**Keywords:** self-timed circuit; soft error; fault tolerance; CMOS; working phase; spacer; layout

**DOI:** 10.14357/08696527200201

### Acknowledgments

The research was funded by a grant from the Russian Science Foundation (Project No. 19-11-00334).

### References

1. Kishinevsky, M., A. Kondratyev, A. Taubin, and V. Varshavsky. 1994. *Concurrent hardware: The theory and practice of self-timed design*. J. Wiley & Sons. 368 p.
2. Stepchenkov, Yu. A., Yu. G. Diachenko, and G. A. Gorelkin. 2011. Samosinkhronnye skhemy — budushchee mikroelektroniki [Self-timed circuits are microelectronics future]. *Voprosy radioelektroniki* [Issues of Radioelectronics] 2:153–184.
3. Stepchenkov, Y. A., A. N. Kamenskikh, Y. G. Diachenko, Y. V. Rogdestvenski, and D. Y. Diachenko. 2019. Fault-tolerance of self-timed circuits. *10th Conference (International) on Dependable Systems, Services, and Technologies Proceedings*. IEEE. 41–44. doi: 10.1109/DESSERT.2019.8770047.

4. Taubin, A., A. Kondratyev, J. Cortadella, and L. Lavagno. 1999. Behavioral transformations to increase noise immunity in asynchronous specifications. *5th Symposium (International) on Advanced Research in Asynchronous Circuits and Systems Proceedings*. IEEE. 36–47. doi: 10.1109/ASYNC.1999.761521.
5. Eaton, P., J. Benedetto, D. Mavis, K. Avery, M. Sibley, M. Gadlage, and T. Turflinger. 2004. Single event transient pulse width measurements using a variable temporal latch technique. *IEEE T. Nucl. Sci.* 51(6):3365–3368. doi: 10.1109/TNS.2004.840020.
6. Danilov, I. A., M. S. Gorbunov, A. I. Shnaider, A. O. Balbekov, Y. B. Rogatkin, and S. G. Bobkov. 2016. DICE-based Muller C-elements for soft error tolerant asynchronous ICs. *16th European Conference on Radiation and Its Effects on Components and Systems*. IEEE. Art. No. F4. doi: 10.1109/RADECS.2016.8093145.

*Received February 17, 2020*

## **Contributors**

**Stepchenkov Yuri A.** (b. 1951)— Candidate of Science (PhD) in technology, Head of Department, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; YStepchenkov@ipiran.ru

**Diachenko Yuri G.** (b. 1958)— Candidate of Science (PhD) in technology, senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; diaura@mail.ru

**Rogdestvenski Yuri V.** (b. 1952)— Candidate of Science (PhD) in technology, Head of Laboratory, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation, Moscow 119333, Russian Federation; YRogdest@ipiran.ru

**Morozov Nikolai V.** (b. 1956)— senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; NMorozov@ipiran.ru

**Stepchenkov Dmitri Yu.** (b. 1973) — senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; stepchenkov@mail.ru

**Diachenko Denis Y.** (b. 1987) — research-engineer, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; diaden87@gmail.com

## УНИВЕРСАЛЬНЫЙ ФУНКЦИОНАЛЬНЫЙ МЕТОД АНАЛИЗА БОЛЬШИХ САМОСИНХРОННЫХ СХЕМ\*

*Л. П. Плеханов<sup>1</sup>, В. Н. Захаров<sup>2</sup>*

**Аннотация:** Одна из главных задач создания самосинхронных (СС) электронных схем — анализ их самосинхронности. Известные событийные методы не обеспечивают полного анализа СС-схем реальной сложности из-за чрезмерного объема вычислений. В рамках функционального подхода предложен универсальный метод, основанный на автоматическом разделении схемы на минимальные СС-ячейки. Метод позволяет радикально уменьшить необходимые вычисления и анализировать самосинхронность схем любого размера.

**Ключевые слова:** самосинхронные схемы; анализ самосинхронных схем

**DOI:** 10.14357/08696527200202

### 1 Введение

Самосинхронные цифровые схемы имеют уникальные свойства по сравнению с традиционными типами схем, синхронными или асинхронными:

- полное и принципиальное отсутствие состязаний при любых задержках элементов схемы;
- отказобезопасность, проявляющаяся в остановке всех переключений при возникновении хотя бы одной константной неисправности типа залипания на 0 или 1 и некоторых других неисправностей.

Эти свойства обеспечивают правильность функционирования в максимально широком диапазоне внешних условий (температуры и напряжения питания), а также дают базис для создания высоконадежных устройств.

Впервые существование СС-схем было показано в работе Маллера и Бартки [1] и подтверждено прямыми экспериментами [2].

Практические принципы и методы создания СС-схем были разработаны В. И. Варшавским с сотр. [3]. Основные из этих принципов таковы:

- индикация изменений сигналов схемы на основе СС-кодирования — возможность на выходах всей схемы отследить такие изменения;

\*Исследование выполнено при финансовой поддержке по Программе фундаментальных исследований 2019 г. Президиума РАН (проект 2019-0054-2.2).

<sup>1</sup>Федеральный исследовательский центр «Информатика и управление» Российской академии наук, lplekhanov@inbox.ru

<sup>2</sup>Федеральный исследовательский центр «Информатика и управление» Российской академии наук, vzakharov@ipiran.ru

- порядок работы схемы — поочередная смена двух фаз: рабочей и спейсерной (промежуточной);
- обратная связь — после окончания текущей фазы (с помощью индикации выходов) инициирование следующей фазы на входах схемы.

При создании СС-схем принципиально важен анализ проектируемой схемы на самосинхронность. Классический подход состоит в анализе событий — изменений в нулях и единицах сигналов схемы. Подобные событийные методы в дальнейшем были созданы на основе сетей Петри и других представлений.

Главные недостатки событийных методов следующие:

- (1) экспоненциальная зависимость числа анализируемых состояний от количества элементов (функций) схемы. Если в схеме  $N$  элементов, число состояний может достигать  $2^N$ . Для не очень большой схемы из 100 элементов провести событийный анализ практически нереально;
- (2) событийный анализ проводится для замкнутого описания схемы. Это вынуждает разработчика делать дополнительную непростую работу по корректному замыканию выходов на входы;
- (3) отсутствие реального соблюдения полноты анализа [4]. Сеанс событийного анализа учитывает только одно начальное состояние и один набор значений входов. Схема, прошедшая неполный анализ, не может считаться СС. Необходимая полнота экспоненциально зависит от числа информационных входов схемы и переменных памяти. На практике полный событийный анализ никогда не делается.

Ранее был предложен функциональный подход (ФП) к проектированию СС-схем [5, 6]. Подход основан на описании СС-схем в логических функциях, а не в состояниях и событиях. Описание СС-схемы, таким образом, становится не динамическим, а статическим. Это позволяет, в частности, рассматривать разомкнутые схемы, что гораздо удобнее при проектировании. Для разомкнутых СС-схем сформулирован критерий самосинхронности на основе соблюдения приведенных выше свойств СС-схем. При этом разомкнутая схема, СС в смысле функционального подхода, при корректном замыкании остается СС и по определению Д. Маллера.

Функциональный подход позволяет делать анализ и синтез СС-схем иерархически, что обеспечивает создание таких схем любого размера. Отметим, что описанные в литературе событийные иерархические методы имеют скорее теоретический характер, так как их использование в практических случаях нереально.

Анализ СС-схем в функциональном иерархическом подходе состоит из двух частей: анализа на нижнем уровне иерархии и анализа на более высоких уровнях.

Анализ на нижнем уровне проводится при описании схемы в логических функциях. Такой анализ подробно изложен в статье [7].

Обозначим на этом уровне:

$N$  — число элементов (внутренних сигналов) схемы;

$M$  — число ее информационных входов;

$P$  — число переменных памяти (бистабильных ячеек).

В статье установлено, что сложность вычислений на этом уровне полиномиальна порядка 2 от  $N$  и экспоненциальна как  $2^{2(M+P)}$ .

На верхних уровнях иерархии СС-схема описывается структурно как вызовы СС-подсхем более низких уровней. Сложность вычислений на этих уровнях линейна от  $N$  — числа сигналов данной схемы, причем внутренние сигналы СС-подсхем в это число не входят.

В результате основная проблема объема вычислений оказывается на нижнем уровне иерархии, где есть экспоненциальная зависимость сложности.

Актуальность проблемы следует из того, что на практике уже давно разрабатываются СС-схемы с множеством входов. Например, в [8] представлен СС-вычислитель, имеющий три входных операнда по 64 разряда и один вход на 6 разрядов. Ясно, что анализ подобных схем представляет собой сложную задачу. Теоретически можно потребовать от разработчика разбивать схему на небольшие подсхемы. Однако разбиение по признаку анализа носит искусственный характер, значительно затрудняет проектирование и всегда нежелательно. Наилучшим решением могло бы стать создание метода, независимого от разбиения исходной схемы на СС-подсхемы.

Предлагаемый метод развивает иерархический метод в ФП. Его универсальность заключается в том, что анализируемая схема может быть представлена как простыми (несамосинхронными) элементами, так и СС-подсхемами, ранее прошедшими анализ, и эта схема может находиться на любом уровне иерархии.

Несамосинхронная по описанию часть схемы также должна быть проверена на самосинхронность. Идея состоит в автоматическом определении ее минимальных фрагментов, имеющих свойство самосинхронности, и анализе этих фрагментов. Далее схема уже будет состоять только из СС-подсхем и к ней можно будет применить анализ на верхнем уровне иерархии.

В результате в каждом фрагменте число  $N$  будет не более двух,  $M$  не превысит нескольких единиц из-за ограниченного числа входов фрагмента, а зависимость сложности вычислений от числа  $P$  удастся устраниТЬ. Анализ схемы в такой ситуации не представит вычислительной проблемы при любом ее размере.

## 2 Основные положения и определения

Сигналы в СС-схемах подразделяются по категориям: информационные и служебные, фазовые и нефазовые.

Информационные сигналы обеспечивают содержательные вычисления, служебные — нет. Фазовые сигналы несут информацию о текущей фазе, нефазо-

вые — нет. Всем фазовым сигналам приписывается *спейсер* 0 или 1 — значение в спейсерной фазе.

Информационные СС-сигналы следующие:

- *парафазные со спейсером* (ПФС) — два связанных фазовых сигнала в СС коде [3]. В этом коде каждый исходный информационный бит представляется двумя битами (сигналами): бит 0 — битами 01, бит 1 — битами 10, спейсерное (промежуточное) состояние кодируется одинаковыми битами 00 или 11, оставшееся состояние, противоположное спейсерному (11 или 00), запрещено;
- *бинарные нефазовые* (БНФ) — два выхода bistабильной ячейки триггера.

Другие разновидности информационных СС-сигналов: мультисигнальные со спейсером и унарные нефазовые — далее для простоты рассматриваться не будут. Их обработка несущественно отличается от обработки приведенных выше сигналов.

Основные служебные СС-сигналы следующие:

- *унарные фазовые* (УФ) — отдельные сигналы, несущие информацию о текущей фазе. Они выполняют роль индикаторных и управляющих сигналов.

Другие служебные сигналы далее рассматриваться не будут.

Сигналы в СС-схемах характеризуются *СС-атрибутами*:

- атрибут *назначения* со спейсером или без — один из приведенных выше типов сигналов;
- атрибут *транзитная фаза* — для БНФ-сигналов фаза, в которой эти сигналы меняются;
- *индикаторный* атрибут фазового выхода — список индицируемых на этом выходе сигналов в каждой из двух фаз по отдельности;
- *полные* атрибуты — все атрибуты выхода со спейсерами согласно типу этого выхода.

Элементы анализируемой схемы, кроме СС-подсхем, должны принадлежать одной библиотеке, называемой *основной*.

К основной библиотеке предъявляются следующие требования:

- она базируется на элементах КМДП-технологии;
- она должна содержать все необходимые для анализа элементы памяти — СС-триггеры. Это объясняется тем, что все практически значимые разновидности СС-триггеров к настоящему моменту созданы, проанализированы и запатентованы. Поэтому целесообразно их использовать целиком со своими СС-атрибутами и брать из библиотеки.

Библиотека может также содержать любые комбинационные несамосинхронные элементы, и любые СС-элементы со своими СС-атрибутами.

Пример такой библиотеки приведен в [9].

Анализ в предлагаемом методе связан с нахождением и контролем минимальных СС фрагментов. Самосинхронные подсхемы и СС-элементы библиотеки при этом не раскрываются и не анализируются.

*Назовем СС-ячейкой минимальный СС фрагмент схемы.*

В принятых условиях СС-ячейки могут быть следующих типов:

- (1) *ПФС-ячейка* — состоит из двух простых комбинационных элементов. Входами и выходами ячейки служат ПФС-сигналы;
- (2) *индикаторная ячейка (И-ячейка)* — комбинационный элемент с одним выходом. Входами и выходом ячейки будут фазовые сигналы. Эти ячейки предназначены для понижения числа индикаторных сигналов;
- (3) *триггерная ячейка (Т-ячейка)* — СС-триггер основной библиотеки.

На рисунке показаны примеры ПФС-ячейки, выполняющей исходную функцию 2И, и соответствующей И-ячейки. Примеры СС-триггеров приведены в [7, рис. 4.5].

### 3 Постановка задачи

Для анализа предъявляются: основная библиотека и анализируемая схема.

Схема должна быть представлена в структурной форме и может содержать:

- библиотечные элементы из основной библиотеки,
- СС-подсхемы из любых, возможно других, библиотек.

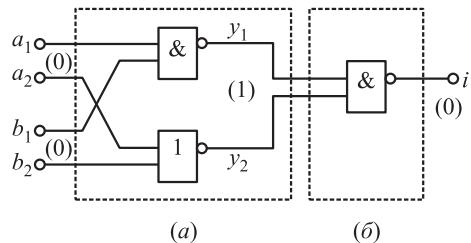
Все входы схемы должны быть описаны СС-атрибутами назначения со спейсерами, а выходы — такими же атрибутами без спейсеров.

Самосинхронные подсхемы должны быть снабжены полными СС-атрибутами входов и выходов.

Требуется определить: самосинхронность схемы и спейсерные значения, а также индикаторные атрибуты фазовых выходов схемы.

### 4 Процедура анализа

Для краткости будем называть элементами схемы как библиотечные элементы, так и СС-подсхемы.



Примеры СС-ячеек (в скобках указаны значения спейсеров): (a) ПФС-ячейка 2И; (б) И-ячейка

Назовем *ярусом* с номером  $i$  ( $i \geq 0$ ) множество элементов, все входы которых суть выходы элементов предыдущих ярусов. Выходы элементов яруса будем приписывать также к этому ярусу. Сигналами нулевого яруса будут считаться входы схемы. Сигналами последнего яруса будут выходы схемы.

Для двух соединенных между собой элементов схемы первый элемент назовем *источником* по отношению ко второму, а второй — *приемником* по отношению к первому.

Анализ начинается с разбиения схемы на ярусы и далее расчеты делаются последовательно с 1-го яруса. В первом ярусе все входные атрибуты известны от входов схемы. В каждом ярусе определяются и анализируются только имеющиеся в нем СС-ячейки. Самосинхронные подсхемы и библиотечные СС-элементы, заранее прошедшие анализ, в нем не нуждаются. В ходе анализа проверяется правильность соединений элементов по типам и спейсерам и вычисляются полные атрибуты выходов ячеек текущего яруса.

Списки индикации обрабатываются особым образом.

Эти списки приписываются фазовым выходам ячеек: ПФС-выходам — обоим сигналам, унарным выходам — каждому по отдельности.

В один список, приписанный одному или паре выходов, попадают:

- сам сигнал или оба сигнала, которым приписан список;
- индицируемые на данных выходах входы ячейки.

Здесь и далее в булевых выражениях знак  $\wedge$  обозначает отрицание, знак  $\vee$  — операцию ИЛИ, отсутствие знака — операцию И.

Индицируемые входы ячейки будут вычисляться путем имитации залипаний и сравнения с нормальными выходами [5].

При этом же расчете каждому индицируемому сигналу сопоставляется *условие индикации* вида  $U = 1$ , где  $U$  — логическое выражение от входных переменных ячейки [5]. Условие может быть *тождественным*, когда само выражение  $U$  тождественно равно единице.

Условия индикации необходимы для точного отслеживания индикации сигнала по мере его распространения в схеме. Если сигнал из списка последовательно прошел через ячейки  $i$  и  $j$ , то на выходе ячейки  $j$  для него будет условие  $U_i U_j = 1$ .

Если сигнал поступил на разные входы  $k$  и  $m$  одной ячейки, т. е. прошел по разным (параллельным) путям с условиями  $U_i = 1$  и  $U_j = 1$ , то его условие примет вид:

$$U_i U_k \vee U_j U_m = 1.$$

В таком порядке будут определены и все атрибуты выходов схемы. В соответствии с ФП для самосинхронности схемы должны быть выполнены два условия:

- (1) индицируемость всех внутренних сигналов схемы (не входящих в СС-подсхемы) на ее фазовых выходах в обеих фазах;
- (2) корректность соединений СС-триггеров между собой и окружением.

## 4.1 Распознавание и анализ индикаторных и парафазных со спейсером ячеек

По очередной паре ПФС-сигналов предыдущего яруса находится множество их приемников. Если в один приемник приходят оба сигнала пары, то этот приемник — И-ячейка. Каждая И-ячейка анализируется с условиями.

Например, все двухвходовые И-ячейки на элементах 2И, 2ИЛИ, 2И-НЕ и 2ИЛИ-НЕ тождественно индицируют все входы в обеих фазах, т. е. транслируют индикаторные атрибуты с входов на выходы.

После определения всех И-ячеек в данном множестве рассматриваются остальные приемники множества. Это ПФС-ячейки. Приемников должно быть четное число, и их необходимо разделить по парам.

Парафазные со спейсером ячейки распознаются по следующим признакам:

- функции двух элементов ПФС-ячейки взаимно двойственны;
- входы обоих элементов ПФС-ячейки принадлежат одним и тем же источникам.

Как следует из [3, с. 92], для самосинхронности найденных ячеек необходимо и достаточно только индицируемости их выходов. Эти выходы являются внутренними сигналами схемы (кроме выходов схемы), и их индицируемость будет отслежена в процессе дальнейшего анализа схемы.

Таким образом, для каждой ячейки остается определить индикаторные атрибуты выходов.

Например, анализ ПФС-ячейки на рисунке показывает на выходах такую индикацию:

- при переходе в спейсер сигналы  $a_1$  и  $b_1$  не индицируются вообще, сигналы  $a_2$  и  $b_2$  индицируются с тождественными условиями;
- при переходе в рабочую фазу сигналы  $a_1$  и  $b_1$  индицируются каждый с условием  $a_1 \wedge a_2$  и  $b_1 \wedge b_2 = 1$ , сигналы  $a_2$  и  $b_2$  не индицируются вообще.

Индикаторная ячейка на рисунке индицирует свои входы тождественно в обеих фазах.

## 4.2 Контроль соединений триггерных ячеек

По указанному выше алгоритму Т-ячейки заранее прошли анализ и снабжены полными атрибутами входов и выходов. Остается только проверка их включения в схеме.

Бинарные нефазовые входы Т-ячейки должны всегда сопровождаться управляющим УФ-сигналом от источника этих входов. Этот УФ-сигнал запрещает или разрешает переключение Т-ячейки.

Транзитные фазы ячеек источника и приемника должны быть разными.

Основное правило соединений Т-ячейки состоит в обеспечении дисциплины входов — порядка их изменений в разных фазах [7, рис. 1].

*В транзитной фазе источника изменение его БНФ-выходов должно начинаться позже запрета переключений приемника.*

Термин «позже» понимается в структурном смысле [5], т. е. обязательно через цепочку элементов.

В нетранзитной фазе источника дисциплина выполняется автоматически.

Правило показывает, что управляющие УФ-сигналы источника и приемника должны быть связаны друг с другом.

Проверка соединений проводится по трем признакам:

- (1) УФ-сигнал источника и приемника есть один и тот же сигнал [5, рис. 5.12];
- (2) УФ-сигнал источника соединен с УФ-сигналом приемника через цепочку элементов [5, рис. 5.13]. При этом сигнал источника должен индицировать сигнал приемника;
- (3) УФ-сигнал источника соединен по обратной связи с УФ-сигналом, индицирующим УФ-сигнал приемника или БНФ-сигналы приемника [5, рис. 5.15]. При этом сигнал источника должен индицировать сигнал обратной связи.

## 5 Финальные вычисления

На последнем ярусе схемы, т. е. на ее выходах, рассматриваются фазовые сигналы с индикаторными атрибутами. С учетом спейсеров они приводятся к одному УФ-сигналу (оценочному [5]). Все сигналы из индикаторных атрибутов этого УФ-сигнала проверяются по условиям индикации. Внутренние сигналы схемы должны иметь тождественные условия индикации. При необходимости формируются полные СС-атрибуты внешних сигналов схемы.

## 6 Заключение

Впервые предложен универсальный метод анализа больших СС-схем на основе функционального подхода.

Метод основан на автоматическом разделении схемы на минимальные СС-ячейки, что позволяет радикально сократить необходимый объем вычислений и обеспечить анализ СС-схем любого размера.

## Литература

1. Muller D. E., Bartky W. C. A theory of asynchronous circuits // Symposium (International) on the Theory of Switching Proceedings. — Harvard University Press, 1959. Part 1. P. 204–243.
2. Плеханов Л. П., Степченков Ю. А. Экспериментальная проверка некоторых свойств строго самосинхронных схем // Системы и средства информатики. — М.: Наука, 2006. Вып. 16. С. 476–485.

3. Автоматное управление асинхронными процессами в ЭВМ и дискретных системах / Под ред. В. И. Варшавского. — М.: Наука, 1986. 400 с.
4. Плеханов Л. П. Полнота анализа электронных схем на самосинхронность // Системы и средства информатики. — М.: Наука, 2010. Вып. 20. № 1. С. 48–58.
5. Плеханов Л. П. Основы самосинхронных электронных схем. — М.: Бином. Лаборатория знаний, 2013. 208 с.
6. Plekhanov L., Zakharov V., Stepchenkov Yu. Functional approach in self-timed circuits design // IEEE East-West Design & Test Symposium Proceedings. — IEEE, 2015. P. 282–285.
7. Плеханов Л. П. Анализ самосинхронности электронных схем на нижнем уровне иерархии // Системы и средства информатики, 2016. Т. 26. № 2. С. 23–42.
8. Соколов И. А., Рождественский Ю. В., Дьяченко Ю. Г. и др. Нечувствительный к задержкам блок умножения-сложения-вычитания с плавающей точкой // Проблемы разработки перспективных микро- и наноэлектронных систем: Сб. трудов. — М.: ИППМ РАН, 2018. Ч. II. С. 170–177.
9. Степченков Ю. А., Денисов А. Н., Дьяченко Ю. Г. и др. Библиотека функциональных ячеек для проектирования самосинхронных полузаизданных микросхем серий 5503 и 5507 // Полузаказные БИС на БМК серий 5503 и 5507. — М.: Техносфера, 2017. Кн. 4. 376 с.

Поступила в редакцию 13.03.20

---

## UNIVERSAL FUNCTIONAL METHOD FOR ANALYZING LARGE SELF-TIMED CIRCUITS

*L. P. Plekhanov and V. N. Zakharov*

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** One of the main tasks of creating self-timed circuits is to analyze their self-synchronicity. Known event-based methods do not provide a complete analysis of self-timed circuits of real complexity due to the excessive amount of calculations. Within the framework of the functional approach, a universal method based on the automatic division of the scheme into minimal self-timed cells is proposed. The method allows one to radically reduce the necessary calculations and analyze self-timed circuits of any size.

**Keywords:** self-timed circuits; analysis of self-timed circuits

**DOI:** 10.14357/08696527200202

### Acknowledgments

The research was performed under financial support of the Program of fundamental research 2019 of the Presidium of RAS (project 2019-0054-2.2).

## References

1. Muller, D. E., and W. C. Bartky. 1959. A theory of asynchronous circuits. *Symposium (International) on the Theory of Switching Proceedings*. Harvard University Press. 1:204–243.
2. Plekhanov, L. P., and Yu. A. Stepchenkov. 2006. Eksperimental'naya proverka nekotorykh svoystv strogo samosinkhronnykh skhem [Experimental verification of some properties of strictly self-timed circuits]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 16:476–485.
3. Varshavsky, V. I., eds. 1986. *Avtomatnoe upravlenie asinkhronnymi protsessami v EVM i diskretnykh sistemakh* [Automata control of asynchronous processes in computers and discrete systems]. Moscow: Nauka. 398 p.
4. Plekhanov, L. P. 2010. Polnota analiza elektronnykh skhem na samosinkhronnost' [Completeness of electronic circuit analysis for self-timing]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 20(1):48–58.
5. Plekhanov, L. P. 2013. *Osnovy samosinkhronnykh elektronnykh skhem* [The basics of self-timed electronic circuits]. Moscow: Binom. Laboratoriya znaniy. 208 p.
6. Plekhanov, L., V. Zakharov, and Yu. Stepchenkov. 2015. Functional approach in self-timed circuits design. *IEEE East-West Design & Test Symposium Proceedings*. IEEE. 282–285.
7. Plekhanov, L. P. 2016. Analiz samosinkhronnosti elektronnykh skhem na nizhnem urovne ierarkhii [Self-timing analysis of electronic circuits at the lower level of hierarchy]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 26(2):23–42.
8. Sokolov, I. A., Yu. V. Rozhdestvensky, Yu. G. Diachenko, et al. 2018. Nechuvstvitel'nyy k zaderzhkam blok umnozheniya-slozheniya-vychitaniya s plavayushchey tochkoj [Delay-insensitive floating point multiply-add-subtract unit]. *Problemy razrabotki perspektivnykh mikro- i nanoelektronnykh system: Sb. trudov* [Problems of developing promising micro- and nanoelectronic systems proceedings]. Moscow: IPPM RAN. II:170–177.
9. Stepchenkov, Yu. A., A. N. Denisov, Yu. G. Diachenko, et al. 2017. Biblioteka funktsional'nykh yacheek dlya proektirovaniya samosinkhronnykh poluzakaznykh mikroskhem seriy 5503 i 5507 [Library of functional cells for designing self-timed semicustom chips of the 5503 and 5507 series]. *Poluzakaznye BIS na BMK seriy 5503 i 5504* [Semicustom chips of the 5503 and 5507 series]. Moscow: Tekhnosfera. Vol. 4. 376 p.

Received March 13, 2020

## Contributor

**Plekhanov Leonid P.** (b. 1943)— Candidate of Science (PhD) in technology, senior scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; lplekhanov@inbox.ru

**Zakharov Victor N.** (b. 1948) — Doctor of Science in technology, associate professor, scientific secretary, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; vzakharov@ipiran.ru

## ПРЕДОТВРАЩЕНИЕ МАССОВОГО ВЫВОДА ИЗ СТРОЯ СЛАБОЗАЩИЩЕННЫХ МОБИЛЬНЫХ СИСТЕМ ИНФОРМАЦИОННОЙ ПОДДЕРЖКИ ЦИФРОВОЙ ЭКОНОМИКИ\*

*A. A. Грушо<sup>1</sup>, Н. А. Грушо<sup>2</sup>, В. В. Сенчило<sup>3</sup>, Е. Е. Тимонина<sup>4</sup>*

**Аннотация:** Статья посвящена разработке дешевых архитектурных мер по предотвращению массового вывода из строя мобильных систем поддержки цифровой экономики (ЦЭ). Рассмотрена возможность экономичного предотвращения массового вывода из строя слабозащищенных мобильных устройств (МУ) поддержки ЦЭ. Показано, что игнорирование проблемы массового вывода из строя МУ может нанести значительный экономический ущерб. Предложены методы организации адекватной защиты на основе доступных и дешевых средств. Предполагается, что массовый вывод из строя МУ возможен на основе создания бот-сети, реализующей в режиме С&С вредоносное воздействие в течение короткого промежутка времени. При этом подготовка атаки, реализующей внедрение ботов и создание бот-сети, может вестись длительное время в режиме P2P.

**Ключевые слова:** цифровая экономика; гетерогенные системы; информационная безопасность

**DOI:** 10.14357/08696527200203

### 1 Введение

Цифровая экономика [1] — это деятельность, которая появилась и постоянно совершенствуется с развитием цифровых технологий и в которую входят: сервисы различных онлайн-услуг, электронные платежи, интернет-торговля, интернет-реклама и даже интернет-игры и т. п. Параллельно с развитием ЦЭ в единую среду общения постепенно подключаются мобильные системы персональной автоматизации. И именно здесь смартфоны и планшеты выступают основными устройствами управления в решениях.

\*Работа частично поддержана РФФИ (проект 18-29-03081).

<sup>1</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, grusho@yandex.ru

<sup>2</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, info@itake.ru

<sup>3</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, volodias@mail.ru

<sup>4</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, eltimon@yandex.ru

Получив доступ к таким слабозащищенным устройствам, можно уже нанести большой ущерб в случае их массового вывода из строя. Многие сотрудники перемещаются по странам со своими МУ, пользуются недоверенными Wi-Fi-сетями и тем самым подвергают конфиденциальную корпоративную информацию большому риску.

Особый интерес злоумышленников представляют МУ на платформе Android. Главная опасность заключается в том, что это самая популярная мобильная платформа и что данная OS позволяет легко скачать и установить на свое персональное устройство любое недоверенное программное обеспечение (ПО). И как результат, любая работа с корпоративной информацией на таком устройстве находится под угрозой. Многие Android-устройства работают на устаревших версиях этой OS, что становится дополнительным риском, так как устаревшие системы содержат множество известных уязвимостей, которые исправлены в последующих версиях, но никто не вносит изменения в старые версии Andriod, хотя многие пользователи все еще ими активно пользуются.

До недавнего времени лучше обстояли дела с МУ на платформе iOS. Однако в последнее время эта система сдает позиции, появились в сети организации, которые торгуют информацией об уязвимостях в этой операционной системе (ОС). И даже наметилась тенденция к снижению стоимости таких услуг по сравнению с Android. В любом случае число обнаруженных уязвимостей быстро растет в этих популярных мобильных системах, и разработчики не всегда успевают «закрывать» их, что дает злоумышленнику хороший шанс для разнообразных атак на эти МУ с целью хищения как конфиденциальной информации, так и финансовых средств, либо вмешательство в архитектуру промышленной автоматизации, либо в решения интернета вещей.

На особом уровне находятся уязвимости на уровне «железа», когда вредоносное ПО находится прямо внутри чипов мобильных и стационарных устройств. Данный тип уязвимостей чрезвычайно сложно выявить, и поэтому в печати очень мало примеров, описывающих данный тип атак.

В этом случае под угрозой находятся уже сами протоколы обмена данными. Так, в 2017 и 2018 гг. были обнаружены аппаратные уязвимости на уровне процессоров Intel и ARM, которыми воспользовались известные атаки Meltdown и Spectre.

Наряду с аппаратными уязвимостями в мобильных и стационарных вычислительных устройствах особую опасность представляют устройства сетевой инфраструктуры (коммутаторы, роутеры и др.). Уже обнаружен особый класс «бэкдоров», которые позволяют осуществлять атаки на эти устройства, что ставит под угрозу уже работу всей компании, организации или предприятия и возможна утечка любой конфиденциальной информации.

Целью работы является разработка дешевых архитектурных мер по предотвращению массового вывода из строя мобильных систем поддержки ЦЭ. Как отмечалось выше, это смартфоны, планшеты, ноутбуки и др. Основное предположение этой работы состоит в том, что пользователи этих систем не обладают

глубокими знаниями по защите информации и используют слабозащищенные устройства для организации решения задач экономической деятельности (ЭД).

В работе используется классическая схема решения задач информационной безопасности (ИБ):

- рассмотрены вопросы работы МУ;
- построена модель угроз в проблеме массового вывода из строя МУ;
- построена модель нарушителя, способного реализовать массовый вывод из строя МУ;
- построены модели атак для массового вывода из строя МУ поддержки ЦЭ;
- разработаны методы защиты МУ поддержки ЦЭ, использующие доступные недорогие средства.

## 2 Мобильные сети

Основное отличие МУ от стационарных состоит в возможности изменять свою позицию и подключаться к сети в различных точках доступа. Мобильное устройство используется для управления деятельностью организации, ведения переговоров с удаленными корреспондентами, управления финансовыми потоками, разработки и создания продуктов вне офиса и производственных территорий и др. Для решения этих и других задач в МУ могут быть включены следующие функции [2]:

- телефонной связи GSM, LTE и др.;
- подключения к интернету;
- возможности загрузки и независимого выполнения приложений;
- Wi-Fi;
- определения местоположения (GPS, ГЛОНАСС и др.).

Кратко остановимся на двух базовых технологиях МУ: сотовой связи и Wi-Fi.

Сотовая связь основана на системе базовых станций (БС), делящих пространство на соты. В каждой соте МУ регистрируется и подключается к БС. Связь организуется между МУ и БС, далее обеспечивается виртуальное соединение абонентов через сеть БС. Мощность передачи автоматически регулируется так, чтобы обеспечить надежную связь внутри соты и не допустить влияния на соседние соты. Связь между БС обеспечивается высокоскоростными каналами. Для простого GSM возможно построение ложных БС, которые контролируют локальные МУ и обеспечивают транзит информации от МУ с БС провайдера.

Мобильные устройства могут локально взаимодействовать с помощью беспроводной связи на стандартах Wi-Fi. В интернет можно выйти через GSM-соединение с БС провайдера и далее через связь провайдера с интернетом. Другой

способ — это Wi-Fi-соединение с модемом, который через GSM-канал провайдера связан с интернетом. Также Wi-Fi-модем может работать через кабельный или спутниковый канал, связанный с интернетом. Сети Wi-Fi строятся на базе точки доступа, в качестве которой может использоваться любое устройство, реализующее соответствующие протоколы.

### **3 Модель угроз в проблеме массового вывода из строя мобильных устройств поддержки цифровой экономики**

Модель угроз, связанных с массовым выводом из строя МУ, отличается от изощренных атак на компьютерные системы [3]. Предположим, что злоумышленник (далее — противник) имеет доступ к персональным данным субъектов ЭД в регионе. В частности, противник имеет информацию об ЭД: адреса, номера телефонов, их типы, IP-адреса планшетов и мобильных компьютеров. Кроме того, он знает систему организации связи для управления и обсуждения важных для ЭД экономических решений. Для противника важна возможность реализации следующих угроз:

- нарушение связи в отдельных районах путем вывода из строя БС всех провайдеров сотовой связи;
- нарушение связи большого числа МУ, задействованных в ЭД и управлении субъектов ЦЭ, которые повлекут панику и хаос.

Реализовать эти угрозы могут нарушители ИБ, обладающие хорошей организацией, имеющие высокую квалификацию в области реализации компьютерных угроз, а также имеющие доступ к системам сотовой связи и интернету.

Высокие требования к возможностям противника связаны с гетерогенностью МУ и сложностью реализации массового нарушения связей этих МУ.

### **4 Уязвимости мобильных устройств**

Рассмотрим уязвимости и пути реализации указанных угроз.

Основной уязвимостью, которая может быть использована для реализации угрозы массового вывода из строя МУ поддержки ЦЭ, являются бот-сети. Бот предоставляет собой агента противника внутри МУ. Из-за слабости защиты персональных компьютеров и МУ бот-сети получили широкое распространение. В основном через бот-сети реализуются DDoS-атаки и спам. Однако бот-сети можно использовать для выполнения любой команды, которую агент противника получает от центра управления.

Различают два типа бот-сетей:

- C&C — это сети с централизованным управлением ботами;
- P2P — это сети, в которых команды передаются от бота к боту (от узла к узлу).

Боты внедряются в хосты гетерогенной сети постепенно [4], но могут быстро реализовать общую для всех команду. Незадействованный бот может долго «живь» в МУ, дожидаясь команды от системы управления. Такой командой может быть блокирование связи.

Следовательно, реализация атаки массового вывода из строя слабозащищенных МУ поддержки ЦЭ состоит в передаче команды боту, находящемся в МУ, через скрытый канал [5] или, например, через массовую рассылку SMS-сообщений. При этом атака уже предполагает наличие вредоносного кода на МУ [5] и интерпретатора команды, например интерпретатора SMS-сообщения. Для передачи такой рассылки создается система, состоящая из базы данных направления рассылок (адреса и телефоны) и сервера, формирующего SMS-сообщения, содержащие команду вывода из строя МУ. Эти сообщения по выбранным адресам и номерам телефонов передаются боту на МУ через провайдера GSM-связи. Если вредоносный код внедрен на прикладной уровень МУ, то время его «жизни» может быть ограничено периодом до перезагрузки системы.

Другой способ передачи команды ботам состоит в использовании робота обзванивания пользователей МУ. Организация обзыва аналогична системе массовой рассылки SMS-сообщений.

Необходимо упомянуть способ распространения ботов и команд через захват БС. Однако захват отдельной БС может привести только к блокированию ее соты, что в дальнейшем может быть скомпенсировано избыточностью системы БС. Функции выведенной из строя БС будут переданы другим БС с некоторой потерей качества связи.

## 5 Пути защиты от массового вывода из строя мобильных устройств поддержки цифровой экономики

Методы защиты от массового вывода из строя МУ основаны на изоляции МУ и избыточности систем использования МУ. Если ограничиться защитой МУ, используя только гетерогенность этих устройств, то, как показывает практика создания гигантских бот-сетей, с течением времени большинство доступных противнику МУ будут снабжены ботами.

Рассмотрим несколько вариантов защиты МУ от массового вывода из строя.

1. Предположим, что использование МУ в ЭД ЦЭ ограничено несколькими типами МУ, а именно: смартфонами, планшетами и переносными компьютерами (ПК). Сначала рассмотрим связь смартфонов. Большинство пользователей может позволить себе использовать пару смартфонов  $T_1$  и  $T_2$ . В таком случае предлагается использовать следующую схему использования связи. Смартфон  $T_1$  используется для связи вне решения проблем управления и работы организации, а  $T_2$  используется только для этих целей. В обычных условиях  $T_1$  и  $T_2$  могут подвергнуться атаке массового вывода из строя МУ. В связи с этим  $T_2$  будем защищать от атаки следующим образом.

Связь  $T_2$  с БС GSM-провайдера реализуется через модем GSM — Wi-Fi. Такие модемы выполняют функции взаимного преобразования информации в стандартах GSM и стандартах Wi-Fi. Эти функции относительно просты, поэтому такие модемы в настоящее время компактны и достаточно дешевы. Допустим, что со стороны GSM через БС возможно внедрение бота в модем. В связи с этим необходимо выбирать модемы с функцией перезагрузки (reset) и / или дублировать сами устройства. Тогда в случае атаки и вывода из строя модема возможен быстрый переход на другой модем и перезагрузка первого модема.

Связь Wi-Fi относится к беспроводному Ethernet, поэтому на  $T_2$  восстановление переданного сообщения не требует сложного преобразования. На  $T_2$  установлено специальное приложение, реализующее криптографические функции и имеющее интерфейсы с модулем Wi-Fi и, например, приложением Viber.

В задаче защиты от массового вывода из строя МУ можно использовать слабую криптографию [6], которая требует мало ресурсов и обеспечивает высокую скорость. Кроме того, слабая криптография не требует лицензирования в соответствии с нормативной базой РФ. Шифрование осуществляется с помощью симметричного алгоритма на ключах, соответствующих номерам абонентов корпоративной телефонной сети организации. Задачей криптографического преобразования является недопущение встраивания вредоносного кода в  $T_2$  транзитом через модем и недопущение взаимодействия с абонентами, не входящими в корпоративную сеть, т. е. с пользователями, не имеющими аналогичных  $T_2$  и ключей. Для организации связи целесообразно использовать  $T_1$ , при помощи которого вызывающий абонент предупреждает о необходимости защищенного соединения.

Такой подход позволяет в период отсутствия соединения помещать  $T_2$  в кожух, изолирующий от электромагнитного воздействия. Незащищенный работающий в режиме GSM смартфон автоматически регистрируется на БС и может быть подвергнут прямой атаке.

Более слабый вариант защиты не требует наличия  $T_1$  и основан на отключении GSM-связи, но сохранении работающей функции Wi-Fi и функции идентификации номера телефона.

Обоснование защищенности схемы телефонной связи основано на следующих принципах. При атаке с целью массового вывода из строя МУ основной удар принимают на себя  $T_1$  и модем. Как отмечено выше, модем можно перезапустить или сразу продублировать. Защита от атаки на  $T_2$  связана с тем, что простой функционал модема не позволяет внедрить сложный интеллектуальный бот. Поэтому использование протокола Wi-Fi вместо GSM и затрудненный доступ к оборудованию и ПО  $T_2$  ограничивают деструктивные возможности бота в модеме. Разумеется, такая возможность остается, но время внедрения вредоносного кода в  $T_2$  ограничено техническими средства-

ми защиты (кожух электромагнитной защиты) и сложностью вредоносного кода, необходимого для проникновения в  $T_2$ .

Слабая криптография ограничивает возможности реализации массовой атаки, так как требует индивидуального восстановления ключа криптографического алгоритма для каждой пары МУ в корпоративных сетях.

Предложенная схема защиты не предполагает вредоносного воздействия через Wi-Fi, поскольку при команде на массовый сбой модем как GSM-устройство должен первым выйти из строя.

2. Аналогичная схема может быть использована при защите МУ типа планшет и ПК. Отличие состоит в том, что GSM–Wi-Fi модем кроме связи с GSM-провайдером может замыкаться на кабельную или спутниковую связь с провайдером интернет-услуг. При появлении стека TCP IP-криптографию можно встраивать в приложение с протоколом IPSec.
3. Для мобильных ПК можно использовать схемы взаимодействия с интернетом на основе нескольких гостевых виртуальных машин на одном устройстве с гипервизором [7].

## 6 Заключение

В работе рассмотрена возможность экономичного предотвращения массового вывода из строя слабозащищенных МУ поддержки ЦЭ. Показано, что игнорирование проблемы массового вывода из строя МУ может нанести значительный экономический ущерб.

Предложены методы организации адекватной защиты на основе доступных и дешевых средств. Предполагается, что массовый вывод из строя МУ возможен на основе создания бот-сети, реализующей в режиме С&С вредоносное воздействие в течение короткого промежутка времени. При этом подготовка атаки, реализующей внедрение ботов и создание бот-сети, может вестись длительное время в режиме P2P.

Данный класс атак относится к кибератакам и служит не только для нанесения экономического ущерба, но также для создания паники и хаоса.

## Литература

1. О Стратегии развития информационного общества в Российской Федерации на 2017–2030 годы: Указ Президента Российской Федерации от 09.05.2017 г. № 203. <http://kremlin.ru/acts/bank/41919>.
2. Першин А. Ю. Прошлое, настоящее и будущее мобильных технологий. Продолжение // Безопасность информационных технологий: Мат-лы конф. — М., 2019. <https://www.bit.aciso.ru/bitmos2019>.
3. Грушо А. А., Забежайло М. И., Зацаринный А. А., Николаев А. В., Писковский В. О., Тимонина Е. Е. Классификация ошибочных состояний в распределенных

- вычислительных системах и источники их возникновения // Системы и средства информатики, 2017. Т. 27. № 2. С.29–40.
4. Грушо А.А., Грушо Н.А., Тимонина Е.Е. Методы защиты информации от атак с помощью скрытых каналов и враждебных программно-аппаратных агентов в распределенных системах // Вестник РГГУ. Сер. Документоведение и архивоведение. Информатика. Защита информации и информационная безопасность, 2009. Т. 10. С. 33–45.
  5. Грушо А. А. Скрытые каналы и безопасность информации в компьютерных системах // Дискретная математика, 1998. Т. 10. № 1. С. 3–9.
  6. ISO / IEC 29192-1:2012. Information technology — Security techniques — Lightweight cryptography — Part 1: General. — ISO / IEC JTC 1 / SC 27, 2012. 20 p. <https://www.iso.org/ru/standard/56425.html>.
  7. Грушо А. А., Грушо Н. А., Тимонина. Е. Е. Методы защиты от массового вывода из строя слабозащищенных компьютеров цифровой экономики // Системы и средства информатики, 2019. Т. 29. № 1. С. 4–11.

*Поступила в редакцию 15.03.20*

---

## **PREVENTION OF MASS FAILURE OF LOW-SECURITY MOBILE SYSTEMS OF INFORMATION SUPPORT FOR DIGITAL ECONOMY**

***A. A. Grusho, N. A. Grusho, V. V. Senchilo, and E. E. Timonina***

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** The aim of the paper is to develop cheap architectural measures to prevent mass failure of mobile systems supporting the digital economy. The possibility of cost-effective prevention of mass failure of weak protected mobile devices supporting the digital economy has been considered. It has been shown that ignoring the problem of mass disruption of mobile devices can cause significant economic damage. The methods of organization of adequate protection on the basis of affordable and cheap tools are proposed. It is assumed that mass failure of mobile devices is possible by creating a botnet that implements a malicious impact in C&C mode in a short period of time. At the same time, the preparation of the attack, which implements the introduction of bots and the creation of a botnet, can be carried out for a long time in the P2P mode.

**Keywords:** digital economy; heterogeneous systems; information security

**DOI:** 10.14357/08696527200203

## Acknowledgments

The paper was partially supported by the Russian Foundation for Basic Research (project 18-29-03081).

## References

1. Указ Президента Российской Федерации № 203. 2017. О Стратегии развития информационного общества в Российской Федерации на 2017–2030 годы [On the Strategy for the development of the information society in the Russian Federation for 2017–2030 years]. Available at: <http://kremlin.ru/acts/bank/41919> (accessed March 16, 2020).
2. Pershin, A. Yu. 2019. Proshloe, nastoyashchee i budushchee mobil'nykh tekhnologiy. Prodolzhenie [The past, present and future of mobile technology. Continuation]. *Bezopasnost' informatsionnykh tekhnologiy: Mat-ly konf.* [Information Technology Security Proceedings]. Moscow. Available at: <https://www.bit.aciso.ru/bitmos2019> (accessed March 16, 2020).
3. Grusho, A. A., M. I. Zabeshailo, A. A. Zatsarinny, A. V. Nikolaev, V. O. Piskovski, V. V. Senchilo, and E. E. Timonina. 2017. Klassifikatsiya oshibochnykh sostoyaniy v raspredelennykh vychislitel'nykh sistemakh i istochniki ikh vozniknoveniya [Erroneous states classifications in distributed computing systems and sources of their occurrences]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 27(2):29–40.
4. Grusho, A., N. Grusho, and E. Timonina. 2009. Metody zashchity informatsii ot atak s pomoshch'yu skrytykh kanalov i vrazhdebnykh programmno-apparatnykh agentov v raspredelennykh sistemakh [Methods of information protection against covert channels attacks and malicious software/hardware agents in distributed systems]. *Vestnik RGGU. Ser. Dokumentovedenie i arkhivovedenie. Informatika. Zashchita informatsii i informatsionnaya bezopasnost'* [RGGU Bulletin. Document science and archive science. Informatics. Information security and information security ser.] 10:33–45.
5. Grusho, A. A. 1998. Hidden channels and information security in computer systems. *Discrete Math. Appl.* 10(1):3–9.
6. ISO / IEC 29192-1:2012. 2012. Information technology — Security techniques — Lightweight cryptography — Part 1: General. Available at: <https://www.iso.org/ru/standard/56425.html> (accessed March 16, 2020).
7. Grusho, A. A., N. A. Grusho, and E. E. Timonina. 2019. Metody zashchity ot massovogo vyyoda iz stroya slabozashchishchennykh kompyuterov tsifrovoy ekonomiki [Methods of protection against mass inactivation of low protected computers of digital economy]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(1):4–11.

Received March 15, 2020

## Contributors

**Grusho Alexander A.** (b. 1946) — Doctor of Science in physics and mathematics, professor, principal scientist, Institute of Informatics Problems, Federal Research Center “Computer Sciences and Control” of the Russian Academy of Sciences; 44-2 Vavilov Str., Moscow 119133, Russian Federation; grusho@yandex.ru

**Grusho Nikolai A.** (b. 1982) — Candidate of Science (PhD) in physics and mathematics, senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Sciences and Control” of the Russian Academy of Sciences; 44-2 Vavilov Str., Moscow 119133, Russian Federation; info@itake.ru

**Senchilo Vladimir V.** (b. 1963) — scientist, Institute of Informatics Problems, Federal Research Center “Computer Sciences and Control” of the Russian Academy of Sciences; 44-2 Vavilov Str., Moscow 119133, Russian Federation; volodias@mail.ru

**Timonina Elena E.** (b. 1952) — Doctor of Science in technology, professor, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Sciences and Control” of the Russian Academy of Sciences; 44-2 Vavilov Str., Moscow 119133, Russian Federation; eltimon@yandex.ru

## ОПРЕДЕЛЕНИЕ ЦЕННОСТИ И ЗАЩИТА ИНСАЙДЕРСКОЙ ИНФОРМАЦИИ

*А. С. Кабанов<sup>1</sup>, А. А. Водолаженко<sup>1</sup>*

**Аннотация:** Рассматриваются различные аспекты внедрения средств противодействия инсайдерской деятельности, показан прагматичный подход к оценке необходимости и стоимости внедрения средств противодействия инсайдерской деятельности конкретной организации. Приведено условие экономической целесообразности внедрения средств противодействия инсайдерской деятельности, а также предложены подходы к оценке инсайдерской информации и сформулированы основные выводы. Статья носит аналитический характер и может быть полезна руководителям служб информационной безопасности, преподавателям, аспирантам и студентам вузов.

**Ключевые слова:** инсайдер; инсайдерская деятельность; ценность инсайдерской информации

**DOI:** 10.14357/08696527200204

### 1 Введение

О методах поиска инсайдеров написано довольно много научных трудов, изобилующих формулами, моделями, принципами, теориями вероятностей, полезности и прочими теориями. Специфический язык этих трудов, а также значительная оторванность от практики и отсутствие параллелей с теми обстоятельствами, в которых вынужден функционировать современный бизнес, приводят к тому, что их аудитория ограничивается очень узким кругом специалистов, имеющих специализированное образование и владеющих соответствующим математическим аппаратом. Другая часть трудов уделяет чрезмерное внимание различным практическим рекомендациям и реализациям (техническим, нетехническим и т. д.) без какого-либо технико-экономического обоснования и оценки целесообразности применения. Авторы надеются, что данная статья поможет читателям перейти к систематическому контролю инсайдерской деятельности, используя простой и прагматичный подход, основанный на доступном каждому человеку здравом смысле.

Под **инсайдерской информацией** организации будем понимать сведения, представляющие интерес для сторонних организаций (далее — организаций-конкурентов), которые могут нанести материальный, финансовый, репутационный и прочий ущерб организации при получении их организацией-конкурентом от инсайдера.

---

<sup>1</sup>Московский институт электроники и математики имени А. Н. Тихонова Национального исследовательского университета «Высшая школа экономики», kabanov\_as@mail.ru

## **2 Построение системы противодействия инсайдерской деятельности**

По мнению авторов, первостепенной задачей лиц, ответственных за противодействие инсайдерской деятельности, должно стать создание реестра соответствующей информации с оценкой ее стоимости, а также выбор средств противодействия инсайдерской деятельности и определение стоимости их внедрения. Следует отметить, что тип информации (ограниченного и неограниченного распространения) достаточно часто не влияет на ценность информации. Зачастую открытая информация представляет собой значительно больший «интерес» и ценность для инсайдера, чем информация ограниченного распространения. Например, открытая информация о компаниях-партнерах и условиях сделок для организации-конкурента намного «интереснее», чем защищаемые законом и обрабатываемые специальным образом с использованием средств защиты персональные данные.

После создания соответствующего реестра для организации возможны две стратегии, выбор которых зависит от внутренней структуры организации, а также специфики реализации технологических и бизнес-процессов, а именно:

- (1) для небольших организаций либо организаций, в которых инсайдерская информация достаточно компактно сосредоточена (в силу внутренней структуры и реализации технологических и бизнес-процессов), защита от инсайдерской деятельности может быть сведена к оптимизации структуры в части обработки данной информации. Например, если инсайдерской информации немного и ее концентрация в одном месте не приводит к нарушению технологических и бизнес-процессов организации, то целесообразно разместить ее в одном помещении одного подразделения с доступом узкого круга доверенных лиц, а в случае необходимости реализовать достаточно недорогие организационно-технические меры защиты. Кроме того, необходимы разработка и утверждение порядка доступа к инсайдерской информации, правил соблюдения ее конфиденциальности, а также создание (определение, назначение) структурного подразделения (должностного лица), в обязанности которого входит осуществление контроля. Данная стратегия, как правило, не требует значительных финансовых затрат (в основном организационные) и реализация систем противодействия инсайдерской деятельности достаточно широко описана в различных трудах (см., например, [1–4]);
- (2) для больших, инфраструктурно сложных, разветвленных организаций, в которых инсайдерская информация сильно рассредоточена (в силу специфики реализации технологических и бизнес-процессов), для обоснования целесообразности внедрения различных средств противодействия инсайдерской деятельности необходимо предварительно сопоставить стоимость утечки инсайдерской информации и стоимость средств противодействия инсайдерской деятельности. Внедрение средств противодействия инсайдерской деятельности имеет экономический смысл при выполнении принципа разумной достаточности [5], а именно следующего условия:

$$C_{\text{СПИ}} < C_{\text{РИСК}},$$

где  $C_{\text{СПИ}}$  — стоимость внедрения средств противодействия инсайдерской деятельности;  $C_{\text{РИСК}}$  — стоимость утечки инсайдерской информации для организации.

Следует отметить, что в процессе расчета  $C_{\text{РИСК}}$  в его стоимость зачастую необходимо включать так называемые репутационные риски организации.

При условии  $C_{\text{СПИ}} \geq C_{\text{РИСК}}$  внедрение средств противодействия инсайдерской деятельности бессмысленно с точки зрения экономического эффекта. В этом случае, как вариант, возможно построение графа передачи инсайдерской информации между средствами обработки с целью проведения его оптимизации для уменьшения стоимости  $C_{\text{СПИ}}$  либо приведения к организации первого типа.

Алгоритм выбора средств противодействия отображен на рисунке.

Далее рассмотрим подход к определению сложности организации. Для сложных организаций характерна разветвленная структура с большим числом элементов (средств обработки информации, филиалов и т. д.). Точечные решения для таких организаций не всегда применимы и эффективны. При этом сетевые программно-аппаратные средства защиты (в первую очередь SIEM (security information and event management) системы, а также DLP (data loss prevention) и т. д.) являются более дорогими в части закупки и внедрения. Поскольку интерес представляет только инсайдерская информация, которая хранится, обрабатывается и передается в/между средствами обработки, следовательно, оценка сложности организации тождественна инфраструктурной сложности совокупности данных систем и связей между ними. Таким образом, сложность организации (инфраструктуры) можно оценить, используя теорию графов, где вершины соответствуют отдельным средствам обработки инсайдерской информации, а ребра определяют информационное взаимодействие между ними [6]. Веса ребер графа определяют объем и частоту передаваемой инсайдерской информации. Наличие разных типов взаимодействия определяет множество инфраструктурных графов  $G_k = \langle V_k, E_k, W_k \rangle$ , где  $V_k$  — множество вершин графа;  $E_k$  — множество ребер графа;  $W_k$  — вес ребер графа;  $k$  — тип взаимодействия.

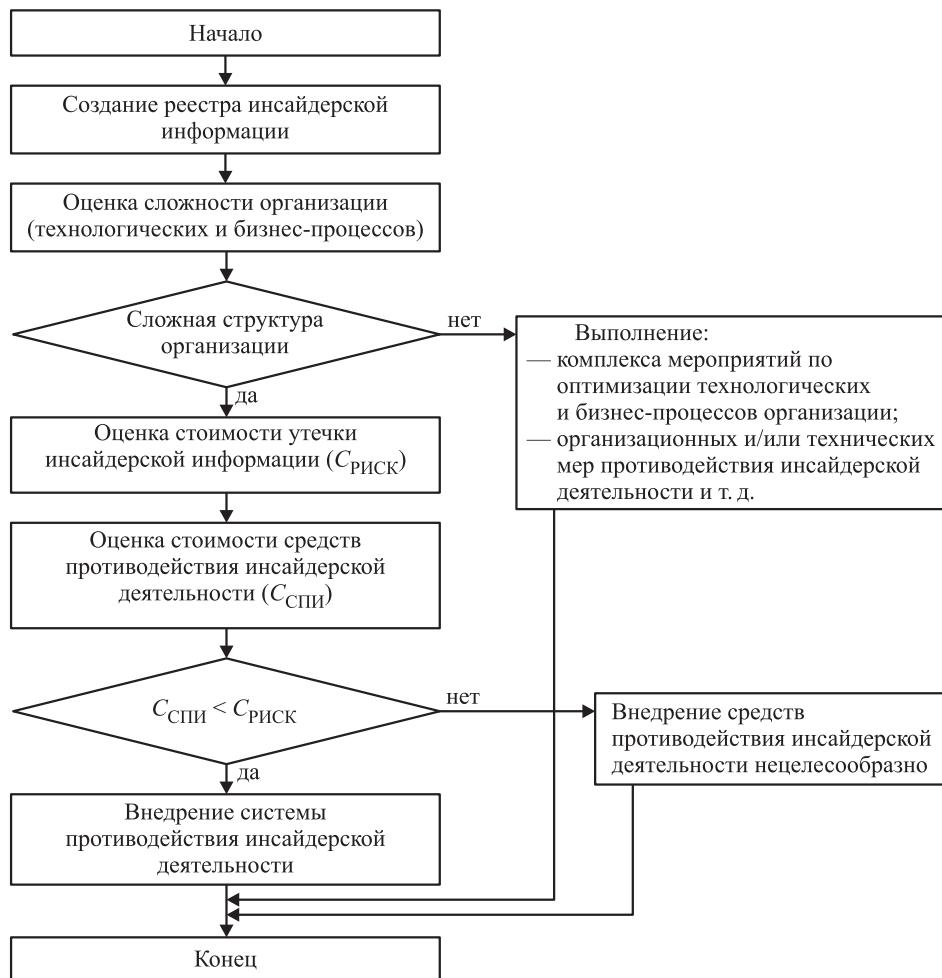
Объединенный график инфраструктуры имеет вид:

$$G = \bigcup_{k=1}^K G_k = \langle V, E, W \rangle.$$

Здесь

$$V = \bigcup_{k=1}^K V_k; \quad E = \bigcup_{k=1}^K E_k,$$

где  $K$  — число типов взаимодействия. Вес объединенного графа определяется суммой весов исходных графов.



Алгоритм внедрения средств противодействия инсайдерской деятельности

Таким образом, получив график средств обработки инсайдерской информации и связей между ними, необходимо провести оценку сложности организации. Создание указанного графа для некоторых небольших организаций первого типа позволит решить оптимизационные задачи, например для сосредоточения инсайдерской информации в одном месте и т. д. Под структурной сложностью инфраструктуры организации будем понимать свойство, оценивающее размерность такого объединенного графа, многообразие маршрутов между его вершинами, число циклов, близость между вершинами и др.

Необходимо сформулировать требования, которым должен удовлетворять объединенный граф инфраструктуры, чтобы организацию можно было отнести к первому или второму типу.

На этапе создания реестра инсайдерской информации необходима идентификация всех вершин графа, относящихся к средствам обработки инсайдерской информации ( $N_{\text{СОИ}}^i$ ). Также должна быть определена максимальная вместимость кластеров средств обработки инсайдерской информации, требующих выполнения недорогостоящих организационно-технических мер защиты информации ( $N_{\text{МЕСТ}}^i$ ), а также число кластеров в конкретной организации ( $L$ ).

Таким образом, для идентификации организации необходимо проверить три условия:

- (1)  $N_{\text{СОИ}} \leq \sum_{i=1}^L N_{\text{МЕСТ}}^i$ ;
- (2) размещение всех  $N_{\text{СОИ}}^i$  во всех  $\sum_{i=1}^L N_{\text{МЕСТ}}^i$  не влечет нарушения технологических и бизнес-процессов организации;
- (3) отсутствие удаленных частей организации с каналами связи, по которым циркулирует инсайдерская информация.

По мнению авторов, выполнение всех трех условий характерно для предприятий первого типа (как правило, небольших). Невыполнение хотя бы одного условия влечет идентификацию как предприятия второго типа. Наличие третьего условия обусловлено относительной дороговизной сетевых средств защиты (особенно сертифицированных) и более разветвленной структурой процесса обработки инсайдерской информации со всеми вытекающими последствиями.

Следует отметить, что предложенный подход к построению графа сложности организации позволит наглядно увидеть все информационные потоки инсайдерской информации, оценить объем и частоту передаваемой инсайдерской информации для принятия решений по внедрению систем сетевой защиты, а также объективно оценить  $C_{\text{СПИ}}$  конкретной организации.

### 3 Определение ценности инсайдерской информации

Далее рассмотрим возможные подходы к оценке стоимости инсайдерской информации, например для создания реестра инсайдерской информации.

Под ценностью инсайдерской информации, как правило, понимается широкий круг свойств информации, которые определяют степень влияния и воздействия на деятельность организации в случае ее утечки к организации-конкуренту. Если информация касается незначительных вопросов и утечка такой информации не ведет к значительному ущербу, то фактом утечки такой информации можно в определенной мере пренебречь. В то же время информация о ситуации, которая чревата банкротством организации, требует повышенного внимания и привлечения значительных ресурсов даже в том случае, если утечка маловероятна.

**Таблица 1** Шкала оценки инсайдерской информации

Рейтинг	Оценка инсайдерской информации	Последствия утечки для организации
5	Имеет решающее значение для деятельности организации	Катастрофические
4	Высокая значимость	Значительные
3	Средняя значимость	Умеренные
2	Низкая значимость	Малозначительные
1	Незначительный риск	Несущественные

Несмотря на то что ценность инсайдерской информации включает в себя комплекс показателей, главным критерием значимости должен служить потенциальный ущерб для организации, измеряемый в натуральных или денежных показателях.

Ценность инсайдерской информации может быть проранжирована в соответствии с определенным заданным критерием. Как правило, это уровень потенциальных негативных эффектов для организации. Рейтинг может присваиваться как самой инсайдерской информации, так и ее источникам [7]. Для оценки инсайдерской информации может быть применена шкала, представленная в табл. 1.

Представленная шкала не претендует на истину в последней инстанции, но, по мнению авторов, может успешно применяться для ранжирования ценности инсайдерской информации (например, в реестре инсайдерской информации организации). Наиболее очевидным для ранжирования ценности инсайдерской информации и вероятности ее утечки в соответствии с табл. 1 представляется использование экспертного оценивания. Однако для оценки выполнения критерия  $C_{СПИ} < C_{Риск}$  кроме ранжирования необходима стоимостная оценка инсайдерской информации. Следует отметить, что не всю инсайдерскую информацию можно достаточно легко оценить в денежном выражении (например, репутационный ущерб). В этом случае, по мнению авторов, целесообразно воспользоваться методом непосредственной оценки (балльным методом). Он позволяет определить, насколько один фактор более значим, чем другие. В этом случае диапазон изменения характеристик объекта разбивается на отдельные интервалы, каждому из которых приписывается определенная оценка (балл), например от 0 до 10. Следует отметить, что экспертное оценивание выходит за рамки данной статьи, поэтому определение степени согласованности мнений экспертов и другие вопросы оценки адекватности и качества экспертного оценивания здесь не рассматриваются.

Для определения стоимости инсайдерской информации (на основе полученных балльных оценок) можно воспользоваться аддитивной моделью. При использовании этой модели определение ценности базируется на экспертных

оценках компонент данной информации и при объективности стоимостных оценок ее компонент подсчитывается искомая величина — их сумма в стоимостном эквиваленте [8]. Основная проблема заключается в том, что количественная оценка компонент информации часто оказывается необъективной, даже у высококвалифицированных специалистов — причина заключается в неоднородности компонент в целом. Для решения этой проблемы принято использовать иерархическую относительную шкалу, которая представляет собой линейный порядок, с помощью которого сравниваются отдельные компоненты по ценности одна относительно другой. Случай единой шкалы равносителен тому, что все компоненты, имеющие равную порядковую оценку, равнозначны между собой.

Рассмотрим следующий пример. Пусть даны  $n$  объектов инсайдерской информации  $O_1, O_2, \dots, O_n$ , оценка производится по десятибалльной шкале; результат оценки экспертами — вектор ценностей объектов каждого относительно другого:  $(3, 5, \dots, 8)$ . Предположим, что изначально определена цена одного из объектов, например  $O = 150$  тыс. руб.

Вычисляем стоимость одного балла:  $O_2/k = 150/5 = 30$  тыс. руб., где  $k$  — оценка объекта в балах.

Аналогичным образом производится оценка других объектов. Сумма стоимостей объектов инсайдерской информации дает полную стоимость всей инсайдерской информации.

Рассмотрим обратную ситуацию. Если известна конечная стоимость инсайдерской информации, то по ней можно обратным преобразованием найти стоимость каждого объекта инсайдерской информации.

С учетом применения метода непосредственного оценивания и аддитивной модели табл. 1 примет числовой вид, который можно с успехом применять для проверки критерия  $C_{\text{СПИ}} < C_{\text{Риск}}$ .

**Рассмотрим два подхода к оценке стоимости инсайдерской информации с точки зрения ее пользы для организации-конкурента.** Следует отметить, что при зеркальном рассмотрении (с позиции оценки инсайдерской информации организацией) данные подходы справедливы для определения  $C_{\text{Риск}}$ .

Прагматичный подход к оценке информации (в данном случае — инсайдерской информации) предложен А. А. Харкевичем [9]. Мерой ценности информации выступает изменение вероятности достижения цели при получении этой информации. В данном случае количественная мера ценности инсайдерской информации  $I_{\text{Ц}}$  выражена следующим образом:

$$I_{\text{Ц}} = \log P_1 - \log P_0 = \log \frac{P_1}{P_0},$$

где  $P_0$  — начальная, до получения инсайдерской информации, вероятность достижения цели организацией-конкурентом;  $P_1$  — вероятность достижения цели организацией-конкурентом после получения информации.

Возможны следующие три случая. В первом случае полученная инсайдерская информация оказывается ценной, увеличивающей вероятность достижения цели

организацией-конкурентом, т. е.  $P_1 > P_0$ . Следовательно, информация признается ценной, полезной и количественная мера ценности информации  $I_{\Pi} > 0$ .

Во втором случае информация не изменяет вероятность достижения цели организацией-конкурентом. Она оказывается бесполезной. При этом  $P_1 = P_0$  и  $I_{\Pi} = 0$ .

В третьем случае вероятность достижения цели уменьшается так как полученная информация оказывается ложной, ошибочной. При этом  $P_1 < P_0$  и  $I_{\Pi} < 0$ .

Данный подход можно успешно применять совместно с методами экспертного оценивания, рассмотренными ранее.

Рассмотрим подход, **основанный на изменении экономической эффективности принятых решений** организацией-конкурентом после получения инсайдерской информации.

Например, некоторая организация-конкурент входит в число лидеров отрасли. Ей необходимо провести модернизацию своих производственных мощностей. На рынке также имеется другая организация, которая в силу своей значительности обладает информацией о перспективах развития рынка. Данная информация, несомненно, представляет интерес для организации-конкурента, поскольку позволяет выработать оптимальную стратегию развития. Цель организации-конкурента — определить ценность инсайдерской информации организации.

Допустим, руководство организации-конкурента рассматривает три варианта действий:

- (1) вложить средства в собственную разработку производственных мощностей на новом технологическом уровне, не имеющем аналогов у конкурентов;
- (2) вложить средства в закупку имеющихся на рынке технологий;
- (3) не проводить модернизацию производства (бездействие).

Размер выигрыша или потерь организации-конкурента зависит от благоприятного или неблагоприятного состояния рынка [10].

Первоначально информация о состоянии рынка отсутствует, но экономисты могут рассчитать выигрыш/потери организации-конкурента от реализации стратегий при благоприятных и неблагоприятных условиях. В соответствии с принципом Байеса состояние рынка (благоприятное или неблагоприятное) принимают равновероятным и равным 0,5.

Средний ожидаемый выигрыш рассчитывается по формуле:

$$W = P_b W_b + P_n W_n ,$$

где  $P_b$  и  $P_n$  — вероятности благоприятных и неблагоприятных состояний соответственно;  $W_b$  и  $W_n$  — выигрыш и потери организации-конкурента соответственно.

Результаты расчетов до и после получения инсайдерской информации от организации представлены в табл. 2 и 3.

**Таблица 2** Средний выигрыш до получения инсайдерской информации

Вариант	Действия организации-конкурента	Выигрыш/потери в зависимости от конъюнктуры рынка, млн руб.		Средний выигрыш до получения инсайдерской информации, млн руб.
		Благоприятная $P_B = 0,5$	Неблагоприятная $P_H = 0,5$	
1	Собственная разработка	130	-120	5
2	Покупка технологий	30	-10	10
3	Бездействие	10	-20	-5

**Таблица 3** Средний выигрыш после получения инсайдерской информации

Вариант	Действия организации-конкурента	Выигрыш/потери в зависимости от конъюнктуры рынка, млн руб.		Средний выигрыш после получения инсайдерской информации, млн руб.
		Благоприятная $P_B = 0,7$	Неблагоприятная $P_H = 0,3$	
1	Собственная разработка	130	-120	55
2	Покупка технологий	30	-10	18
3	Бездействие	10	-20	1

Результаты расчетов показывают, что наилучшим решением при отсутствии инсайдерской информации оказывается закупка технологий, так как при этом средний выигрыш максимален и равен  $W_0 = 10$  млн руб.

Допустим, покупка инсайдерской информации обойдется организации-конкуренту в 1 млн руб. Инсайдерская информация указывает на то, что прогнозируется благоприятное состояние рынка с вероятностью 0,7 и неблагоприятное с вероятностью 0,3. В этом случае максимальный средний выигрыш ( $W_1 = 55$  млн руб.) организация-конкурент получит при выборе стратегии по собственной разработке производственных мощностей на новом технологическом уровне.

Таким образом, полученная дополнительная информация изменяет представление о целесообразности стратегии организации-конкурента и, соответственно, повышает экономический эффект с 10 млн до 55 млн руб.

Ценность дополнительной информации равна разности максимальных средних выигрышей организации-конкурента до и после получения инсайдерской информации с учетом стоимости самой информации:

$$I_{\Pi} = W_1 - W_0 - S_{\Pi},$$

где  $S_{\Pi}$  — стоимость получения инсайдерской информации.

Для рассмотренного примера  $I_{\Pi} = (55 - 10) - 1 = 44$  млн руб.

Таким образом, ценность инсайдерской информации определяется выгодой от ее использования. Ценность инсайдерской информации может существенно

превысить стоимость ее получения. Платить за инсайдерскую информацию имеет смысл, если  $W_1 - W_0 > S_{\pi}$ , т. е.  $I_{\Pi} > 0$ .

В рассмотренном примере инсайдерская информация указывает на прогноз благоприятного и неблагоприятного состояния рынка, что на практике, например, можно определить по объему выпускаемых в продажу средств производства с определенными характеристиками, которые влияют на конкуренцию в определенном сегменте рынка.

Очевидно, что  $I_{\Pi}$  входит, а в некоторых случаях полностью тождественно  $C_{\text{Риск}}$ .

## **4 Заключение**

Описанный экономический критерий целесообразности внедрения средств противодействия инсайдерской деятельности (взятый из классической теории информационной безопасности) во многих трудах незаслуженно отсутствует, что с точки зрения практической реализации некорректно ввиду его ключевой роли (например, для руководителей организации).

Рассмотренные подходы к внедрению средств противодействия инсайдерской деятельности для различных организаций, по мнению авторов, наиболее логичны и позволяют быстро определиться с методами, формами и средствами противодействия инсайдерской деятельности сквозь призму экономической целесообразности. Дальнейшим направлением развития может стать разработка типовых моделей с детализацией условий классификации и характеристик для сложных организаций.

Представленные подходы к оценке и ранжированию инсайдерской информации могут с успехом применяться на практике, поскольку достаточно просты и интуитивно понятны. Выбор используемого подхода должен определяться удобством и наличием исходных данных для конкретной организации.

## **Литература**

1. *Дьяченко О.* Как защититься от инсайдера, 13.01.2012. [http://www.prostobankir.com.ua/it/stati/kak\\_zaschititsya\\_ot\\_insaydera](http://www.prostobankir.com.ua/it/stati/kak_zaschititsya_ot_insaydera).
2. *Cohen F.* Нетехнические методы противодействия инсайдерским угрозам, 23.06.2015. <https://www.securitylab.ru/analytics/473402.php>.
3. *Кабанов А. С., Суроев А. В., Лось А. Б.* Методы социальной инженерии в сфере информационной безопасности и противодействие им // Российский следователь, 2015. № 18. С. 32–37.
4. *Кабанов А. С., Лось А. Б.* Причины, профилактика и методы противодействия инсайдерской деятельности // Безопасность бизнеса, 2016. № 3. С. 28–35.
5. *Гайкович В. Ю., Ершов Д. В.* Основы безопасности информационных технологий. — М.: МИФИ, 1995. 96 с.

6. Наумов В. Н., Кучеренко Д. В. Исследование структурной сложности инфраструктуры государственных информационных систем методами анализа социальных графов // Современные научноемкие технологии, 2019. № 2. С. 114–122.
7. Шарамко М. М. Внутренний контроль: методология, система и процессы. — М.: Русайнс, 2016. 228 с.
8. Гришо А. А., Тимонина Е. Е. Теоретические основы защиты информации. — М.: Яхсмен, 1996. 192 с.
9. Волкова В. Н. Теория информационных процессов и систем. — М.: Юрайт, 2016. 502 с.
10. Лазебник В. М. Экономическая кибернетика. <https://studfiles.net/preview/1100363>.

Поступила в редакцию 03.04.19

---

## VALUATION AND PROTECTION OF INSIDER INFORMATION

*A. S. Kabanov and A. A. Vodolazhenko*

National Research University Higher School of Economics, 34 Tallinskaya Str., Moscow 123458, Russian Federation

**Abstract:** Different aspects of introduction of means to counteract insider activity are considered. The pragmatic approach to assessing the necessity and cost of implementing countermeasures to insider activity of a certain organization is shown. The condition of the economic feasibility of introducing measures to counteract insider activities is given, the approaches to evaluating insider information are proposed, and the basic conclusions are formulated. The article is analytical and can be useful for heads of information security services, teachers, and students.

**Keywords:** insider; insider activity; value of insider information

**DOI:** 10.14357/08696527200204

## References

1. Дьяченко О. 13.01.2012. Kak zashchitit'sya ot insaydera [How to protect yourself from an insider]. Available at: [http://www.prostobankir.com.ua/it/stati/kak\\_zaschititsya\\_ot\\_insaydera](http://www.prostobankir.com.ua/it/stati/kak_zaschititsya_ot_insaydera) (accessed March 20, 2020).
2. Cohen, F. 23.06.2015. Netekhnicheskie metody protivodeystviya insayderskim ugrozam [Nontechnical methods of countering insider threats]. Available at: <https://www.securitylab.ru/analytics/473402.php> (accessed March 20, 2020).
3. Kabanov, A. S., A. B. Los, and A. V. Suroev. 2015. Metody sotsial'noy inzhenerii v sfere informatsionnoy bezopasnosti i protivodeystvie im [Methods of social engineering in the sphere of information safety and counteraction thereto]. *Rossiyskiy sledovatel'* [Russian Investigator] 18:32–37.

4. Kabanov, A. S., and A. B. Los. 2016. Prichiny, profilaktika i metody protivodeystviya insayderskoy deyatel'nosti [Reasons, prevention and methods of counteraction of insider activity]. *Bezopasnost' biznesa* [Business Security] 3:28–35.
5. Gajkovich, V. Y., and D. V. Ershov. 1995. *Osnovy bezopasnosti informatsionnykh tekhnologiy* [Fundamentals of information technology security]. Moscow: MIFI. 96 p.
6. Naumov, V. N., and D. V. Kucherenko. 2019. Issledovanie strukturnoy slozhnosti infrastruktury gosudarstvennykh informatsionnykh sistem metodami analiza sotsial'nykh grafov [Study of the structural complexity of the infrastructure of state information systems by the analysis of social graphs]. *Sovremennye naukoemkie tekhnologii* [Modern High Technologies] 2:114–122.
7. Sharamko, M. M. 2016. *Vnutrenniy kontrol': metodologiya, sistema i protsessy* [Internal control: Methodology, system and processes]. Moscow: Rusayns. 228 p.
8. Grusho, A. A., and E. E. Timonina. 1996. *Teoreticheskie osnovy zashchity informatsii* [Theoretical bases of information security]. Moscow: Yakhsmen. 192 p.
9. Volkova, V. N. 2016. *Teoriya informatsionnykh protsessov i sistem* [Theory of information processes and systems]. Moscow: Yurayt. 502 p.
10. Lazebnik, V. M. 2008. Ekonomicheskaya kibernetika [Economic cybernetics]. Available at: <https://studfile.net/preview/1100363/> (accessed March 20, 2020).

*Received April 3, 2019*

## **Contributors**

**Kabanov Artem S.** (b. 1979) — Candidate of Science (PhD) in technology, associate professor, Moscow Institute of Electronics and Mathematics, National Research University Higher School of Economics, 34 Tallinskaya Str., Moscow 123458, Russian Federation; kabanov\_as@mail.ru

**Vodolazhenko Anatoly A.** (b. 1943) — employee, Moscow Institute of Electronics and Mathematics, National Research University Higher School of Economics, 34 Tallinskaya Str., Moscow 123458, Russian Federation; kabanov\_as@mail.ru

## MULTIDISCIPLINARY NEUROINFORMATICS PROBLEMS FOR EXECUTION IN DISTRIBUTED COMPUTING INFRASTRUCTURES

*D. Y. Kovalev<sup>1</sup>, I. A. Shanin<sup>2</sup>, and E. M. Tirikov<sup>3</sup>*

**Abstract:** Neuroinformatics lies at the intersection of computer science and neuroscience, making it possible to use methods and tools from one domain for accumulating, processing, analyzing, and managing data and modeling techniques from another. Nowadays, neuroinformatics is evolving very fast, this leads to a rapid expansion of the range of scientific problems that need to be solved. This article deals with a number of urgent problems in the area of cognitive functions modeling of the neurophysiology domain. Problems are analyzed from the point of view of neuroinformatics. Common pitfalls, methods, processing tools, and implementation issues are examined. In total, four problem statements are discussed with data sets residing in the resting state and task functional magnetic resonance imaging as well as in electroencephalograms. The methods vary from simple linear models to highly sophisticated deep neural networks. Justifications for using distributed computing infrastructures are discussed for each problem, including high dimensionality in data that requires, on the one hand, distributed implementation and, on the other hand, using computationally extensive methods that require low-level GPU-based parallelization.

**Keywords:** data intensive research; neuroinformatics; distributed computing infrastructures

**DOI:** 10.14357/08696527200205

### 1 Introduction

Neurophysiology is one of the prominent examples of data intensive scientific domain. Being a combination of anatomy, physiology, genetics, biochemistry, neuropathology, and psychology, it remains the forefront of brain research and modeling. The amount of data generated in a typical laboratory conducting research in neuroscience is growing with astonishing speed. Reaching petabytes in volume, it makes data integration, processing, and analysis complicated problems. To solve it,

---

<sup>1</sup>Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation, dkovalev@ipiran.ru

<sup>2</sup>Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation, ishanin@frccsc.ru

<sup>3</sup>Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation, em.tirikov@gmail.com

neuroinformatics is required, so that both new and previously known data become more accessible to the research community to accelerate the understanding of the brain.

Neuroinformatics arises as an interdisciplinary field with the collaboration of computer scientists and neuroscientists, making it possible to use methods and tools from one domain for accumulating, processing, analyzing, and managing data and modeling techniques from another. Neuroinformatics is evolving very fast, this leads to a rapid expansion of the range of scientific problems that need to be solved on the basis of the data obtained and enables accumulation of experience in solving such problems and the possibility of its interdisciplinary use.

Most of existing brain projects and initiatives explicitly state the importance of cognitive models of the human brain. Examples include the task of comparing visual, speech and auditory presentation in the human brain. Additional complexity is brought by using different imaging techniques for collecting data sets, including functional Magnetic Resonance Imaging (fMRI), Magnetoencephalography (MEG), Electroencephalography (EEG), etc. Currently, there is no single universally accepted distributed platform for solving problems in the field of neurophysiology. The need to develop methods and tools for solving data analysis problems in distributed computing infrastructures remains high.

A project 18-29-22096 financially supported by the Russian Foundation for Basic Research and performed by Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences and Research Center of Neurology is aimed at development of methods and tools for solving data analysis problems in distributed computing infrastructures in the field of neurophysiology. The project is largely interdisciplinary, since it is positioned within the intersection of the areas of development of distributed computing infrastructures and applying data analysis methods for the neurophysiology subject domain. This particular article is concentrated on a number of urgent problems in the area of cognitive functions modeling of neurophysiology domain and analyzing them from neuroinformatics point of view, including the highlight of common pitfalls, methods, and processing tools, implementation problems, which is an initial step for the problem solving over distributed computing infrastructures.

The article is structured as follows. Section 2 considers problems from neurophysiology domain and states where data come from and why the problems are data intensive. Way forward for each problem statement is discussed at the end of each subsection. Section 3 depicts common pitfalls and other comments for the problem statements. Section 4 concludes the article.

Research is carried out in the frame of Research Data Alliance activity<sup>1</sup>.

---

<sup>1</sup>“The Research Data Alliance (RDA) was launched as a community-driven organization in 2013 by the European Commission, the United States National Science Foundation and National Institute of Standards and Technology, and the Australian Government’s Department of Innovation with the goal of building the social and technical infrastructure . . . that promotes data-sharing and data-driven research. . . .” (<https://www.rd-alliance.org/about-rda>).

## 2 Neuroinformatics Problem Statements

### 2.1 Nonlinear functional connectivity analysis

Functional connectivity search is a fundamental part in understanding brain cognitive activities [1]. It consists in finding significant dependencies between anatomically separated areas of the human brain. The study of functional connectivity provides additional knowledge about the large-scale communication of neurons in the human brain. In particular, resting state fMRI is applied to several neuropsychiatric diseases, including attention deficit disorder, depression, nicotine decontamination, schizophrenia, Alzheimer's disease, etc. [2].

Nonlinear functional connectivity analysis is intended to construct temporal nonlinear functions for the brain regions and to compare them afterwards. Initial step is to extract regions of interest (ROIs). Region of interest is a set of anatomically close voxels, which are built-in units of fMRI image, representing small cubes of the brain. As fMRI image captures time changes in blood oxygen level, each voxel is associated with respective time series. Regions of interest can also be related to time series, e.g., by averaging all its voxels' time series. Regions of interest can be selected using atlases of human brain [3] that are mappings from voxels into regions. Human brain atlases are sometimes inaccurate and controversial. Drawbacks of using atlases can be overcome by using Principal Component Analysis (PCA) and Independent Component Analysis (ICA) [4] to automatically extract ROIs from small datasets, and MELODIC's Incremental Group-PCA (MIGP) [5] method to extract ROIs from huge datasets. The MIGP computation time varies linearly with the number of objects.

The idea of nonlinear approaches is to build a nonlinear function  $x_1 = f(x_2, \dots, x_n)$ , e.g.,  $x_1 = x_2x_3 + x_4^3$ , for each ROI  $x_1$  and all other ROIs  $x_2, \dots, x_n$  that explicitly defines dependencies between regions. Well-known procedure to compute such equations is General Linear Model (GLM) [6]. General linear model has low computational complexity. Nonlinearity is achieved by using nonlinear combination of input variables. The disadvantage of this approach is that the number of input nonlinear combinations grows exponentially with regard to number of ROIs. General linear model combines relatively small number of nonlinearities into linear combinations and selects the best-fitting combination. In such way, some meaningful nonlinear combinations can be missed.

Another way to construct such nonlinear functions is to use genetic programming which does not require any *a priori* assumptions [7]. Complexity of genetic algorithms is exponential, increasing with fMRI dataset dimensionality. It is possible to overcome the issue with search space decrease by PCA/ICA methods.

Sometimes, it is feasible to find relationships in implicit form, e.g., in parametric form. It means that there exists a parameter variable  $t$ , so that ROI becomes not a function of another ROIs, but a function of  $t$ . In [8], dynamic causal modeling

with some modifications is used to approximate ROIs with hidden variable in form of ordinary differential equation.

Another method of nonstandard approximation is to use neural ordinary differential equations (NODE) neural network [9]. Hidden state derivative is parameterized with neural network, hidden states variable becomes continuous (in contrast to discrete hidden state variables of classical neural networks). Next, differential equation solver is used to compute output of the network. The advantage of the approach is less memory consumption comparing to other methods.

Search for the functional connectivity of the regions of the human brain in a linear form (significant pairwise correlations for time series) and nonlinear form (in the form of formulaic relationships between time series for brain regions) consists of several tasks:

- (1) choosing algorithm appropriate to data volume and computational complexity for finding linear / nonlinear differences in functional relationships;
- (2) preprocessing incoming data sets, including methods for isolating ROIs;
- (3) computing pairwise correlations or functions;
- (4) testing significance hypotheses; and
- (5) developing distributed implementation of chosen algorithm.

As huge resting state fMRI data sets become available (Human Connectome Project (HCP) [10] and 1000 connectomes [11]), the problem becomes data and computationally intensive. The use of large datasets can improve statistical power of research leading to more accurate results.

## 2.2 Study of gender difference patterns in brain connectivity

One of crucial areas of cognitive brain analysis is the study of gender difference patterns in brain connectivity. Understanding the role of gender-related effects can provide insight into the specialized treatment which can differ for men and women.

There are multiple different studies using different datasets which come to controversial outcomes. In [12], the authors present neuroimaging experiment where differences between men and women using language and emotional processing dataset with 26 subjects are examined showing that there is no significant difference between men and women. Time-series analysis is performed with ANCOVA (analysis of covariance) [13] to test null hypothesis for each subject and voxel that signal did not significantly correlate with stimulus epoch of interest.

Other studies [14, 15] investigate gender difference patterns of the brain at rest. In [14], gender difference in the cerebral cortex using resting-state fMRI dataset is analyzed. In total, 112 subjects are analyzed. Kendall's coefficient of concordance (KCC) [15] is calculated to measure the similarity of a given voxel to its nearest neighbors, after  $z$ -normalization is applied to it and significance hypothesis test is computed. The result is that men and women have significant difference in

several regions, such as primary visual network, left attention network, default mode network, sensorimotor network, executive network, and dorsal medial prefrontal network.

In [16], the authors investigate gender difference in functional connectivity of the human brain using resting state fMRI. One hundred and four healthy subjects are examined. Preprocessing includes application of ICA for data reduction. The ANCOVA algorithm is applied to estimate significant gender effects on regional level. Significant difference between men and women resting state fMRI is found in frontal and temporal regions and in the cerebellum.

The related task of finding gender differences is to classify sex using fMRI. In [17], 820 subjects from HCP are used to build a sex classifier. Authors preprocess data with predefined functional template to decrease number of features, which becomes the input to partial least squares (PLS) regression. Ten-fold cross-validation is used to check that the classification accuracy score of 87% and AUC (average under the curve) score from 0.88 to 0.93 is robust. Feature importance is analyzed showing that there are seven features associated with default mode network in top twenty.

In [18], instead of using whole-brain connectivity, classifiers are trained on separate brain regions. The HCP dataset containing data on 1200 subjects is used. The authors use brain atlas to extract regions of interest. Pairwise correlations and Fisher  $z$ -scores are computed as additional features. Support Vector Machine (SVM) with radial basis function is used as the classifier. Ten-fold cross-validation is used to test robustness of model prediction. Significance test is then applied to check if prediction is significant. In total, prediction accuracy reached 75%.

Study of gender difference patterns consists of several tasks:

- (1) selecting if whole-brain or single region is used;
- (2) split dataset into two separate datasets and compute group averaged statistics;
- (3) preprocessing incoming data sets including methods for isolating ROIs;
- (4) engineering of additional features, e.g., entropy, time-related, model-related;
- (5) select appropriate to data volume and computational complexity statistical methods;
- (6) testing significance hypotheses;
- (7) developing distributed implementation of chosen algorithm; and
- (8) cross-validating results, including validation on different datasets.

Usage of more subjects, e.g., from HCP and 1000 functional connectomes, can increase robustness of such studies. As more data are analyzed, additional deep and machine learning models should be brought into the research. Deep neural networks can improve classification accuracy significantly. Another way to increase accuracy score is to engineer more features, such as entropy and time related features.

### 2.3 Automatic artefact detection in electroencephalography

Every electroencephalographic recording is inevitably contaminated with artefacts of various origins. Most commonly, the signal is affected by electrical muscle activity produced by eye movements, blinks, jaw muscles (teeth clenching and chewing), neck muscles, and cardiac activity. Accurate detection and removal of artefacts significantly degrades the performance of automatic EEG analysis. The task of artefact removal can be decomposed in two parts: artefact detection and signal reconstruction.

Most of the automatic artefact detection pipelines identify bad channels, filtering, artefact removal based on ICA, bad channel interpolation, and quality assessment. The following classifiers are used to study ICA-based feature representations: SVM, LDA, hierarchical clustering, decision trees. The ICA-based solution may suffer from correction error and often requires visual inspection which can be a major drawback if the dataset is large enough. In [19], the following hand-crafted features are used to classify independent components (ICs): alpha-theta band power, eye movement template correlation, entropy, kurtosis, variance, and spatial features. Set of listed features comprises a polar diagram that is called the IC “fingerprint.”

Recent achievements in artificial neural networks (ANN) show their superiority in automatic feature extraction, pattern recognition, and noise removal tasks in various domains compared with classical machine learning methods. A common approach is to incorporate ANN into EEG preprocessing pipeline as a classifier or feature extractor. At the same time, recent approaches follow so-called end-to-end approach when a single neural architecture is used to solve a given problem and all the parameters are learned through back propagation. In [20], a combination of sparse autoencoders (SAE) and recursive least squares (RLS) was used to remove artefacts, associated with eye movements (based on electrooculography signal) and proved to be more efficient than standard approach. In [21], a similar approach that included deep autoencoder (DAE) was used to remove artefacts associated with eye movements and jaw clenching.

There are several datasets available that contain an artefact markup. The authors of [19] introduced a dataset that includes EEG recordings of externally cued artefact-related events (eye blinks, eye movements, and jaw muscle contractions). In [22], the crowd labeled ICA dataset is introduced. By the time of publication, it consisted of more than 200 000 ICs from more than 6000 EEG recording, over 8000 ICs were provided with one of six artefact labels: muscle, heart, eye, line noise, channel noise, and “other.”

The following problems remain unsolved and require further research: end-to-end artefact correction method and GAN-based EEG reconstruction. Although several studies are devoted to the topic of GAN-based EEG interpolation [23, 24], this method has not been applied to artefact correction problem.

### 3 Emotion Recognition

Accurate human emotion recognition is an important problem in the area of affective brain–computer interaction. The problem of emotion recognition can be solved with various modalities such as face image, voice, and eye tracking patterns but the neuroimaging-based approaches (such as fMRI and EEG) can possibly improve the quality of recognition and provide valuable insights to affective brain–computer interaction development.

The first openly available dataset for EEG-based emotion recognition was introduced in [25] as a part of eINTERFACE-2006 project. The DEAP dataset [26] is one of the most widely-used publicly available EEG-based emotion recognition benchmark data with biosignals of 32 participants. The dataset presented in [27] is comprised of the data from the commercial-level brain–computer interface devices which gained noticeable popularity over recent years.

As deep learning methods gained their popularity due to their transcendent performance on computer vision tasks, intensive application of deep neural network-based feature extraction and emotion recognition became an important research topic. In [28], stacked autoencoder (SAE) performs an emotion classification over power spectral features extracted from a DEAP dataset. In [29], the Long Short-Term Memory (LSTM) neural network performs regression to arousal valence coordinates. In [30], these methods are used together, resulting in a proposed SAE+LSTMframework, which gave a significant boost on a DEAP dataset.

The following problems and ideas require further research. One of the approaches to increase available training data would be to use a teacher–student neural architectures from multimodal studies of different purposes. Within this approach, a large collection of emotion unrelated datasets has been used to perform cross-domain data labeling. An accurate model training requires an algorithm that is robust to between-participant and between-device variations. The compensation of this variation should be included to a neural architecture as one of the domain adaptation technique.

### 4 Data Intensive Research of Huge Functional Magnetic Resonance Imaging Datasets

Statistically underpowered research remains a major problem in neurophysiology [31]. A lot of studies infer their outcomes based on analysis of 20–30 subjects. Combined with large number of analyzed variables, this frequently leads to low statistical power and, consequently, to high rate of missed brain connections and misleading inference of brain organization. Also, neuroscience community is less likely to achieve consensus using controversial underpowered studies.

The appearance of hugeneuroscience datasets allows to overcome some of the mentioned problems. It brings in a need for (*i*) scalable distributed computing tools

able to process large amount of incoming data and (ii) parallelizable algorithms to work with computationally intensive models, e. g., deep neural networks.

Addressing the first issue, Apache Spark [32] has been widely used for solving data-intensive problems. It allows to effectively handle and process large datasets. There are several successful examples of applying Spark to bio- and neuroscience problems [33].

Modeling brain cognitive functionally requires not only large amounts of incoming data, but efficient computing engines as well. In [34], the authors present a combination of Apache Spark and Tensorflow open source platform for implementing deep neural networks [35]. The works on combination of Apache Spark and other deep learning engines and using other types of incoming data, e. g., resting state fMRI and EEG data sets, are not known yet.

## 5 Concluding Remarks and Future Work

In total, four urgent problem statements in the area of cognitive functions modeling of neurophysiology domain were collected and analyzed in the article. Common pitfalls, methods and processing tools, and implementation problems are discussed. Way forward for each problem is presented. A common issue of all discussed problems is low statistical power unless huge datasets are used. This results as a set of requirements and suggestions of the tools to be used for problem solving.

Discussed problems are going to be solved over a distributed computing infrastructure. Architecture of the distributed infrastructure for problem solving in the field of neurophysiology is being developed within the mentioned RFBR project 18-29-22096. When designing and implementing the distributed computing infrastructure, modern methods of distributed computing and free software tools such as Spark are planned to be used. It is also supposed to use a combination of distributed tools with platforms for implementing deep neural networks for problem solving over different types of available neuroscience datasets.

## Acknowledgments

This research was partially supported by the Russian Foundation for Basic Research (project 18-29-22096).

## References

1. Friston, K. 2011. Functional and effective connectivity: A review. *Brain Connectivity* 1(1):13–36.
2. Greicius, M. 2008. Resting-state functional connectivity in neuro-psychiatric disorders. *Curr. Opin. Neurol.* 21(4):424–430.
3. Allgaier, N., T. Banaschewski, G. Barker, et al. 2015. Nonlinear functional mapping of the human brain. 21 p. Available at: <https://arxiv.org/abs/1510.03765> (accessed March 23, 2020).

4. Beckmann, C., and S. Smith. 2004. Probabilistic independent component analysis for functional magnetic resonance imaging. *Med. Imaging* 23(2):137–152.
5. Rachakonda, S., R. Silva, J. Liu, and V. Calhoun. 2016. Memory efficient PCA methods for large group ICA. *Front. Neurosci.* 10:738–749.
6. Soch, J., A. Meyer, and J. Haynes. 2017. How to improve parameter estimates in GLM-based fMRI data analysis: Cross-validated Bayesian model averaging. *Neuroimage* 158:186–195.
7. Ilknur, I., A. Nicholas, C. Allgaier, *et al.* 2014. A deterministic and symbolic regression hybrid applied to resting-state fMRI data. *Genetic programming theory and practice XI. Genetic and evolutionary computation*. New York, NY: Springer. 155–173.
8. Cao, X., B. Sandstede, and X. Luo. 2019. A functional data method for causal dynamic network modeling of task-related fMRI. *Front. Neurosci.* 13:127. 19 p.
9. Chen, T., Y. Rubanova, J. Bettencourt, and D. Duvenaud. 2018. Neural ordinary differential equations. *Advances in Neural Information Processing Systems Conference Proceedings*. Eds. S. Bengio, H. Wallach, H. Larochelle, *et al.* Montréal, Canada. 6571–6583.
10. Human Connectome Project. Available at: <http://humanconnectome.org/1000> (accessed March 23, 2020).
11. Functional Connectomes Project. Available at: [http://fcon\\_1000.projects.nitrc.org/](http://fcon_1000.projects.nitrc.org/) (accessed March 23, 2020).
12. Garn, C., M. Allen, and J. Larsen. 2009. An fMRI study of sex differences in brain activation during object naming. *Cortex* 49(5):610–618.
13. Rutherford, A. 2001. *Introducing ANOVA and ANCOVA: A GLM approach*. Thousand Oaks, CA: Sage Publications. 182 p.
14. McRae, K., K. Ochsner, I. Mauss, J. Gabrieli, and J. Gross. 2008. Gender differences in emotion regulation: An fMRI study of cognitive reappraisal. *Group Process. Interg.* 11(2):143–162.
15. Xu, C., C. Li, H. Wu, *et al.* 2014. Gender differences in cerebral regional homogeneity of adult healthy volunteers: A resting-state fMRI study. *BioMed Res. Int.* 2015:183074. 8 p.
16. Filippi, M., P. Valsasina, P. Misci, *et al.* 2013. The organization of intrinsic brain activity differs between genders: A resting-state FMRI study in a large cohort of young healthy subjects. *Hum. Brain Mapp.* 34(6):1330–1343.
17. Zhang, C., C. C. Dougherty, S. A. Baum, T. White, and A. M. Michael. 2018. Functional connectivity predicts gender: Evidence for gender differences in resting brain connectivity. *Hum. Brain Mapp.* 39(4):1765–1776.
18. Weis, S., K. R. Patil, F. Hoffstaedter, A. Nostro, B. T. Thomas Yeo, and S. B. Eickhoff. 2019. Sex classification by resting state brain connectivity. *Cereb. Cortex. Art. ID: bhz129*. Available at: <https://academic.oup.com/cercor/advance-article/doi/10.1093/cercor/bhz129/5524764> (accessed March 23, 2020).
19. Tamburro, G. , P. Fiedler, D. Stone, J. Haueisen, and S. Comani. 2018. A new ICA-based fingerprint method for the automatic removal of physiological artifacts from EEG recordings. *Peer J.* 6:e4380.
20. Yang, B., K. Duan, and T. Zhang. 2016. Removal of EOG artifacts from EEG using a cascade of sparse autoencoder and recursive least squares adaptive filter. *Neurocomputing* 214:1053–1060.

21. Nobre Leite, N. M., E. T. Pereira, E. C. Gurjao, and L. R. Veloso. 2018. Deep convolutional autoencoder for EEG noise filtering. *Conference (International) on Bioinformatics and Biomedicine (BIBM)*. IEEE. 2605–2612.
22. Pion-Tonachini, L., K. Kreutz-Delgado, and S. Makeig. 2019. ICLLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage* 198:181–197.
23. Hartmann, K. G., R. T. Schirrmeister, and T. Ball. 2018. EEG-GAN: Generative adversarial networks for electroencephalographic (EEG) brain signals. Available at: <https://arxiv.org/abs/1806.01875> (accessed March 23, 2020).
24. Abdelfattah, S. M., G. M. Abdelrahman, and M. Wang. 2018. Augmenting the size of EEG datasets using generative adversarial networks. *Joint Conference (International) on Neural Networks*. IEEE. 1–6.
25. Savran, A., K. Ciftci, G. Chanel, et al. 2007. Emotion detection in the loop from brain signals and facial images. *eINTERFACE Workshop Proceedings*. Presses universitaires de Louvain. 47926.
26. Koelstra S., C. Muhl, M. Soleymani, et al. 2012. Deap: A database for emotion analysis; using physiological signals. *IEEE T. Affect. Comput.* 3(1):18–31.
27. Katsigiannis, S. and N. Ramzan. 2018. DREAMER: A database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices. *IEEE J. Biomed. Health* 22(1):98–107.
28. Jirayucharoensak, S., S. Pan-Ngum, and P. Israsena. 2014. EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation. *Sci. World J.* 2014:627892. 10 p.
29. Yan, J., S. Chen, and S. Deng. 2019. A EEG-based emotion recognition model with rhythm and time characteristics. *Brain Informatics* 6(1):7.
30. Xing, X., Z. Li, T. Xu, L. Shu, B. Hu, and X. Xu. 2019. SAE+LSTM: A new framework for emotion recognition from multi-channel EEG. *Front. Neurorobotics* 13:37.
31. Cremers, H., T. Wager, and T. Yarkoni. 2017. The relation between statistical power and inference in fMRI. *PLoS One* 12(11):e0184923.
32. Zaharia, M., R. S. Xin, P. Wendell, et al. 2016. Apache spark: A unified engine for big data processing. *Commun. ACM* 59(11):56–65.
33. Guo, R., Y. Zhao, Q. Zou, X. Fang, and S. Peng. 2018. Bioinformatics applications on Apache Spark. *Gigascience* 7(8):giy098. 10 p.
34. Makkie, M., X. Li, S. Quinn, et al. 2018. A distributed computing platform for fMRI big data analytics. *IEEE T. Big Data* 5(2):109–119.
35. TensorFlow Project. Available at: <https://www.tensorflow.org/> (accessed March 23, 2020).

Received November 15, 2019

## Contributors

**Kovalev Dmitry Y.** (b. 1988)—scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; dkovalev@ipiran.ru

**Shanin Ivan A.** (b. 1991) — junior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; ishanin@frccsc.ru

**Tirikov Egor M.** (b. 1996) — master student, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; em.tirikov@gmail.com

---

---

## МУЛЬТИДИСЦИПЛИНАРНЫЕ ЗАДАЧИ НЕЙРОФИЗИОЛОГИИ ДЛЯ ИСПОЛНЕНИЯ НА РАСПРЕДЕЛЕННЫХ ВЫЧИСЛИТЕЛЬНЫХ ИНФРАСТРУКТУРАХ\*

**Д. Ю. Ковалев<sup>1</sup>, И. А. Шанин<sup>1</sup>, Е. М. Тириков<sup>2</sup>**

<sup>1</sup> Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

<sup>2</sup> Московский государственный университет им. М. В. Ломоносова

**Аннотация:** Нейроинформатика находится на пересечении информатики и нейробиологии, что позволяет использовать методы и инструменты из одного источника для сбора, обработки, анализа и управления данными и методов моделирования из другого. В настоящее время нейроинформатика быстро развивается, что приводит к существенному расширению круга решаемых научных задач. Данная статья посвящена ряду актуальных задач моделирования когнитивных функций в области нейрофизиологии. Проведен анализ задач с точки зрения нейроинформатики, рассматриваются общие методы и средства обработки, вопросы реализации. В общей сложности обсуждаются шесть постановок задач над наборами данных функциональной магнитно-резонансной томографии покоя и действия, а также над электроэнцефалограммами. Методы решения задач варьируются от простых линейных моделей до сложных глубоких нейронных сетей. Для каждой задачи обсуждается обоснование использования распределенных вычислительных инфраструктур, включая высокую размерность данных и необходимость использования вычислительно сложных методов, которые требуют низкоуровневого распараллеливания на основе графических ускорителей.

**Ключевые слова:** исследования с интенсивным использованием данных; нейроинформатика; распределенные вычислительные инфраструктуры

**DOI:** 10.14357/08696527200205

---

\*Работа выполнена при финансовой поддержке РФФИ (проект 18-29-22096).

## Литература

1. Friston K. Functional and effective connectivity: A review // Brain Connectivity, 2011. Vol. 1. Iss. 1. P. 13–36.
2. Greicius M. Resting-state functional connectivity in neuro-psychiatric disorders // Curr. Opin. Neurol., 2008. Vol. 21. Iss. 4. P. 424–430.
3. Allgaier N., Banaschewski T., Barker G., et al. Nonlinear functional mapping of the human brain. arXiv:1510.03765 [q-bio.NC], 2015. 21 p.
4. Beckmann C., Smith S. Probabilistic independent component analysis for functional magnetic resonance imaging // Med. Imaging, 2004. Vol. 23. Iss. 2. P. 137–152.
5. Rachakonda S., Silva R., Liu J., Calhoun V. Memory efficient PCA methods for large group ICA // Front. Neurosci., 2016. Vol. 10. P. 738–749.
6. Soch J., Meyer A., Haynes J. How to improve parameter estimates in GLM-based fMRI data analysis: Cross-validated Bayesian model averaging // Neuroimage, 2017. Vol. 158. P. 186–195.
7. Ilknur I., Nicholas A., Allgaier C., et al. A deterministic and symbolic regression hybrid applied to resting-state fMRI data // Genetic programming theory and practice XI. Genetic and evolutionary computation. — New York, NY, USA: Springer, 2014. P. 155–173.
8. Cao X., Sandstede B., Luo X. A functional data method for causal dynamic network modeling of task-related fMRI // Front. Neurosci., 2019. Vol. 13. Art. No. 127. 19 p.
9. Chen T., Rubanova Y., Bettencourt J., Duvenaud D. Neural ordinary differential equations // Advances Neural Information Processing Systems Conference Proceedings / Eds. S. Bengio, H. Wallach, H. Larochelle, et al. — Montréal, Canada, 2018. P. 6571–6583.
10. Human Connectome Projects. <http://humanconnectome.org>.
11. Functional Connectomes Project. [http://fcon\\_1000.projects.nitrc.org](http://fcon_1000.projects.nitrc.org).
12. Garn C., Allen M., Larsen J. An fMRI study of sex differences in brain activation during object naming // Cortex, 2019. Vol. 49. Iss. 5. P. 610–618.
13. Rutherford A. Introducing ANOVA and ANCOVA: A GLM approach. — Thousand Oaks, CA, USA: Sage Publications, 2001. 182 p.
14. McRae K., Ochsner K., Mauss I., Gabrieli J., Gross J. Gender differences in emotion regulation: An fMRI study of cognitive reappraisal // Group Process. Intergr., 2008. Vol. 11. Iss. 2. P. 143–162.
15. Xu C., Li C., Wu H., et al. Gender differences in cerebral regional homogeneity of adult healthy volunteers: A resting-state fMRI study // BioMed Res. Int., 2015. Vol. 2015. Art. 183074. 8 p.
16. Filippi M., Valsasina P., Misci P., et al. The organization of intrinsic brain activity differs between genders: A resting-state FMRI study in a large cohort of young healthy subjects // Hum. Brain Mapp., 2013. Vol. 34. Iss. 6. P. 1330–1343.
17. Zhang C., Dougherty C. C., Baum S. A., White T., Michael A. M. Functional connectivity predicts gender: Evidence for gender differences in resting brain connectivity // Hum. Brain Mapp., 2018. Vol. 39. Iss. 4. P. 1765–1776.
18. Weis S., Patil K. R., Hoffstaedter F., Nostro A., Yeo B. T. T., Eickhoff S. B. Sex classification by resting state brain connectivity // Cereb. Cortex, 2019. Art. ID: bhz129.

19. Tamburro G., Fiedler P., Stone D., Haueisen J., Comani S. A new ICA-based fingerprint method for the automatic removal of physiological artifacts from EEG recordings // Peer J., 2019. Vol. 6. Art. ID: e4380.
20. Yang B., Duan K., Zhang T. Removal of EOG artifacts from EEG using a cascade of sparse autoencoder and recursive least squares adaptive filter // Neurocomputing, 2016. Vol. 214. P. 1053–1060.
21. Nobre Leite N. M., Pereira E. T., Gurjao E. C., Veloso L. R. Deep convolutional autoencoder for EEG noise filtering // 2018 IEEE Conference (International) on Bioinformatics and Biomedicine (BIBM). — IEEE, 2018. P. 2605–2612.
22. Pion-Tonachini L., Kreutz-Delgado K., Makeig S. ICLLabel: An automated electroencephalographic independent component classifier, dataset, and website // NeuroImage, 2019. Vol. 198. P. 181–197.
23. Hartmann K. G., Schirrmeister R. T., Ball T. EEG-GAN: Generative adversarial networks for electroencephalographic (EEG) brain signals // arXiv:1806.01875 [eess.SP], 2018. <https://arxiv.org/abs/1806.01875>.
24. Abdelfattah S. M., Abdelrahman G. M., Wang M. Augmenting the size of EEG datasets using generative adversarial networks // Joint Conference (International) on Neural Networks. — IEEE, 2018. P. 1–6.
25. Savran A., Ciftci K., Chanel G., et al. Emotion detection in the loop from brain signals and facial images // eINTERFACE 2006 Workshop Proceedings. — Presses universitaires de Louvain, 2007. Art. ID: 47926. P. 1–12.
26. Koelstra S., Muhl C., Soleymani M., et al. Deap: A database for emotion analysis; using physiological signals // IEEE T. Affect. Comput., 2012. Vol. 3. Iss. 1. P. 18–31.
27. Katsigiannis S., Ramzan N. DREAMER: A database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices // IEEE J. Biomed. Health, 2018. Vol. 22. Iss. 1. P. 98–107.
28. Jirayucharoensak S., Pan-Ngum S., Israsena P. EEG-based emotion recognition using deep learning network with principal component-based covariate shift adaptation // Sci. World J., 2014. Vol. 2014. Art. ID: 627892. 10 p.
29. Yan J., Chen S., Deng S. A EEG-based emotion recognition model with rhythm and time characteristics // Brain Informatics, 2019. Vol. 6. Iss. 1. Art. No. 7.
30. Xing X., Li Z., Xu T., Shu L., Hu B., Xu X. SAE+LSTM: A new framework for emotion recognition from multi-channel EEG // Front. Neurorobotics, 2019. Vol. 13. Art. No. 37.
31. Cremers H., Wager T., Yarkoni T. The relation between statistical power and inference in fMRI // PLoS One, 2017. Vol. 12. Iss. 11. Art. ID: e0184923.
32. Zaharia M., Xin R. S., Wendell P., et al. Apache spark: A unified engine for big data processing // Commun. ACM, 2016. Vol. 59. Iss. 11. P. 56–65.
33. Guo R., Zhao Y., Zou Q., Fang X., Peng S. Bioinformatics applications on Apache Spark // Gigascience, 2018. Vol. 7. Iss. 8. Art. ID: giy098. 10 p.
34. Makkie M., Li X., Quinn S., et al. A distributed computing platform for fMRI big data analytics // IEEE T. Big Data, 2018. Vol. 5. Iss. 2. P. 109–119.
35. TensorFlow Project. <https://www.tensorflow.org>.

Поступила в редакцию 15.11.2019

## АРХИТЕКТУРА ПЛАТФОРМЫ ДЛЯ ПРОВЕДЕНИЯ ДВИЖИМЫХ ГИПОТЕЗАМИ ВИРТУАЛЬНЫХ ЭКСПЕРИМЕНТОВ\*

*Д. Ю. Ковалев<sup>1</sup>, Е. А. Тарасов<sup>2</sup>, В. Н. Захаров<sup>3</sup>, Н. М. Филимонов<sup>4</sup>*

**Аннотация:** Рассматривается проблема проведения движимых гипотезами виртуальных экспериментов. Проведен анализ методов управления виртуальными экспериментами в существующих системах работы с экспериментами, по результатам которого сформирован жизненный цикл виртуального эксперимента. Для этапов жизненного цикла представлены основные операции управления виртуальными экспериментами, а также основные этапы по работе эксперта с платформой исполнения виртуального эксперимента. Предложена программная архитектура платформы управления виртуальными экспериментами и гипотезами с описанием основных компонентов платформы и их функций, которые реализуют основные операции жизненного цикла.

**Ключевые слова:** жизненный цикл виртуального эксперимента; гипотеза; интенсивное использование данных

**DOI:** 10.14357/08696527200206

### 1 Введение

В настоящее время в научных исследованиях необходимо решать задачи по получению новых знаний на основе разноструктурированных распределенных данных. Наборы данных получают в результате наблюдений, экспериментов или компьютерного моделирования, а также построения математических моделей для интерпретации наблюдаемых явлений. Такие исследования опираются на явное использование гипотез и реализующих их вычислительных моделей, используемых для преодоления разрыва между концептуальным описанием изучаемого явления и его симуляцией.

Проект 18-07-01434, выполняемый Федеральным научным центром «Информатика и управление» Российской академии наук при финансовой поддержке

\*Работа выполнена при частичной финансовой поддержке РФФИ (проект 18-07-01434 А).

<sup>1</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, dkovalev@ipiran.ru

<sup>2</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, etarasov@outlook.com

<sup>3</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, vzakharov@ipiran.ru

<sup>4</sup>Институт информационных и вычислительных технологий Национального исследовательского института «Московский энергетический институт», filimonovn160@gmail.com

Российского фонда фундаментальных исследований, направлен на разработку методов и средств проведения виртуальных экспериментов, основанных на проверке гипотез, в исследованиях с интенсивным использованием данных для их ускорения.

Виртуальный эксперимент в рамках исследований определяется как кортеж  $\langle O, H, M, R, W, C \rangle$  [1], где  $O$  — это онтология предметной области, которая представляет собой набор понятий и отношений, определенный на некотором формальном языке (например, OWL 2 [2]);  $H$  — набор спецификаций гипотез и взаимосвязей между ними (описание гипотез в работе ограничивается системами уравнений), являющийся частью онтологии  $O$  и использующий ее понятия;  $M$  — набор моделей, в котором каждая модель представляет собой набор программных функций, реализующих спецификацию некоторой гипотезы;  $R : M \rightarrow H$  — сюръективное отображение из множества моделей в множество гипотез, соответствующее реализации гипотез моделями;  $W$  — поток работ, т. е. набор задач, организованный с помощью определенных конструкций (образцов потоков работ, таких как разбиение, объединение и т. д. [3]) и реализующий эксперимент (каждая задача представляет собой вызов функций некоторой модели из  $M$ );  $C$  — конфигурация эксперимента, представляющая собой отображение задач потока работ в набор значений параметров функций. При проведении нового виртуального эксперимента требуется определить полностью кортеж  $\langle O, H, M, R, W, C \rangle$ . Пример определения элементов кортежа виртуального эксперимента для задачи моделирования масс звезд Безансонской модели Галактики приведен в работе [1].

Анализ возможностей и ограничений существующих систем управления виртуальными экспериментами показывает необходимость разработки и реализации методов и инструментов, имеющих дело с наборами зависимых гипотез, называемых далее *решетками гипотез*. В работе [4] приведен алгоритм построения решеток гипотез для виртуального эксперимента, а также дана оценка сложности алгоритма. Построение решеток гипотез вносит ряд улучшений в проведение виртуального эксперимента, например: использование промежуточных результатов вычисления независимых гипотез приводит к ускорению вычислений. Порождение гипотез представляет дополнительную трудность, а потому при необходимости гипотезы могут автоматически генерироваться из данных. Так, в работе [5] рассмотрены подходы к извлечению гипотез в виде нелинейных функциональных зависимостей между несколькими переменными.

Данная статья посвящена разработке архитектуры платформы управления виртуальными экспериментами, включающей средства концептуального представления гипотез в процессе проведения виртуальных экспериментов, сохранения эволюции гипотез и моделей, построения связей между несколькими гипотезами в одном эксперименте, воспроизведимости экспериментов. Платформа предполагает исследование взаимосвязи между гипотезами в форме математических уравнений, предоставляя информацию не только о зависимостях переменных в одной гипотезе/модели, но также и о нескольких зависимых гипотезах в одном виртуальном эксперименте. Одна из важнейших целей платформы —

повышение скорости проведения исследований за счет повторного использования уже вычисленных фрагментов виртуальных экспериментов.

Работа структурирована следующим образом. В разд. 2 приводится анализ существующих методов управления экспериментами, по результатам которого выделяются жизненный цикл и основные операции формального манипулирования виртуальными экспериментами и гипотезами. В разд. 3 представлена архитектура платформы для проведения движимых гипотезами виртуальных экспериментов, реализующих операции всех этапов жизненного цикла эксперимента.

## 2 Методы и операции управления виртуальными экспериментами и гипотезами

В данном разделе рассматриваются методы манипулирования гипотезами и экспериментами в ряде существующих систем работы с виртуальными экспериментами. По результатам анализа методов манипулирования виртуальными экспериментами выделены операции для последующей реализации в рамках платформы управления виртуальными экспериментами.

В системе Hephaestus [6] основными методами для манипулирования гипотезами и виртуальными экспериментами являются:

- (1) метод автоматического порождения гипотез из данных;
- (2) метод интервенций при исполнении виртуального эксперимента;
- (3) метод корректного разбиения данных наблюдений при проверке статистических гипотез;
- (4) метод ранжирования нескольких конкурирующих гипотез;
- (5) метод построения графа причинно-следственных и вероятностных зависимостей.

В качестве *метода автоматического порождения гипотез* используется символьная регрессия [7] при подборе аппроксимирующей формулы для выбранного подмножества переменных на наборе данных. Сущность *метода интервенций* при исполнении виртуального эксперимента заключается в установлении существующих и возможной генерации новых гипотез, которые значимо влияют на выбранную экспертом гипотезу [8]. *Метод корректного разбиения данных наблюдений* при проверке статистических гипотез требуется для обеспечения несмещанных оценок при дальнейшем тестировании гипотез [9]. *Метод ранжирования* нескольких конкурирующих гипотез позволяет по выбранной метрике (например, достижаемому уровню значимости) или нескольким метрикам выбрать одну или несколько гипотез, наиболее соответствующих наблюдениям. *Построение графа причинно-следственных и вероятностных зависимостей* применяется для оценки соответствия полученных зависимостей общепринятым в данной области, а также

для потенциального исследования значимых корреляций и их последующего преобразования в причинно-следственные зависимости. Кроме того, построенные графы зависимостей могут в явном виде передаваться другим экспертам в данной области.

В системе Upsilon-DB [10] в качестве основных методов можно выделить:

- (1) формальное кодирование гипотез в качестве кортежей баз данных;
- (2) оценку соответствия гипотез наборам данных;
- (3) построение причинно-следственных зависимостей для параметров одной гипотезы.

Сущность метода *формального кодирования гипотез* заключается в представлении гипотез в виде систем уравнений над кортежами базы данных. При этом осуществляется проверка корректности гипотез, строится схема базы данных и набор функциональных зависимостей для переменных, извлеченных из системы уравнений. *Метод оценки соответствия гипотез наборам данных* использует правило Байеса [11]. Дополнительно в схему базы данных добавляется мера соответствия гипотез исследуемому явлению. В процессе оценки соответствия используется *метод построения графа причинно-следственных зависимостей* для параметров одной гипотезы.

В системе FCCE [12] ключевым методом является построение попарных значимых корреляционных зависимостей для всех переменных множества разрозненных наборов данных. Для метода существует распределенная реализация. Метод хорошо масштабируется при сборе, извлечении и запросах к географически распределенным большим наборам данных.

В системе Robot Scientist [13] используются:

- (1) метод логического вывода новых гипотез при проведении экспериментов;
- (2) метод статической оценки корректности соответствия гипотез полученным данным.

*Метод логического вывода* использует абдуктивное логическое программирование [14] для генерации достоверных гипотез, объясняющих наблюдения, а затем использует эти гипотезы для определения следующего наиболее информативного эксперимента. В начале исследования все потенциальные гипотезы считаются в равной степени верными. После этого устанавливается значимость оставшихся гипотез при помощи, например, теста Стьюдента. В работе [15] демонстрируется подход к оценке значимости гипотез при помощи нескольких статистических тестов.

Основываясь на вышеприведенном анализе методов управления виртуальными экспериментами, можно определить *основные этапы жизненного цикла виртуального эксперимента*.

На **первом этапе** происходит инициализация виртуального эксперимента: определение онтологии предметной области, определение гипотез и соответству-

ющих им моделей, определение потока работ для исполнения виртуального эксперимента, задание конфигурации эксперимента. Используется представление гипотез как специальных структур данных, сочетающих уравнения, входящие в них переменные и причинно-следственные отображения над ними. После загрузки виртуального эксперимента в систему происходит автоматическое построение решетки гипотез виртуального эксперимента. Решетка гипотез определяется как направленный ациклический граф с вершинами, отвечающими гипотезам. Ребра графа соответствуют отношениям зависимости между гипотезами, когда результат вычислений одной гипотезы используется в вычислениях другой гипотезы.

На **втором этапе** по желанию пользователя происходит автоматическое порождение дополнительных гипотез из данных.

На **третьем этапе** происходит построение оптимального плана исполнения эксперимента.

На **четвертом этапе** происходит непосредственное проведение эксперимента, при этом система отслеживает, существуют ли сохраненные результаты исполнения отдельных функций эксперимента при заданных параметрах, и подставляет их, если совпадение найдено. После исполнения виртуального эксперимента результат возвращается эксперту.

В процессе определения жизненного цикла выделены следующие *операции в работе над виртуальным экспериментом*. При реализации первого этапа к доступным эксперту операциям над виртуальными экспериментами относятся добавление, модификация и удаление виртуальных экспериментов, гипотез, потоков работ, а также конфигураций экспериментов. Этап также включает автоматические операции: построение графа причинно-следственных зависимостей между переменными одной гипотезы, специфицированной системой уравнений, а также построение структуры для описания взаимодействия между несколькими гипотезами — решетки гипотез. Операция построения графа причинно-следственных зависимостей строит граф на основании системы уравнений, соответствующей гипотезе. Решетка гипотез для виртуального эксперимента строится на основании упомянутого графа и потока работ виртуального эксперимента.

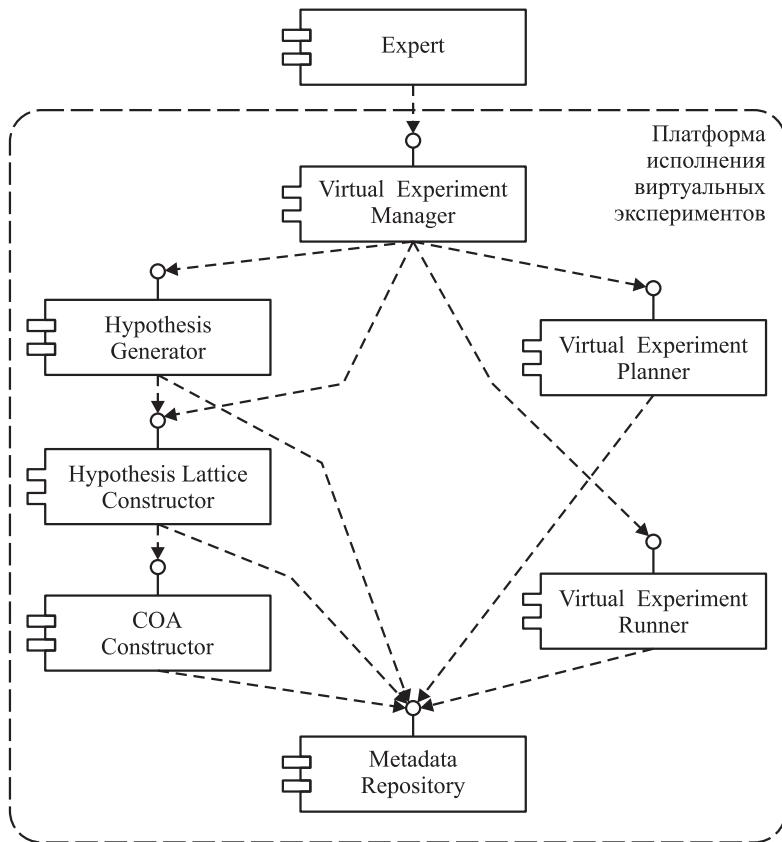
Для реализации второго этапа используются операции автоматического порождения новых гипотез из данных по желанию эксперта и поиска корреляционных зависимостей между параметрами различных независимых гипотез.

Для реализации третьего этапа используется операция построения оптимального плана исполнения виртуального эксперимента.

Для реализации четвертого этапа используется операция исполнения виртуального эксперимента. Промежуточные результаты расчетов, соответствующих вызовам отдельных функций моделей, и оценка их соответствия экспериментальным данным сохраняются в базе данных для последующего использования и отслеживания эволюции гипотез и моделей виртуального эксперимента.

### 3 Программная архитектура платформы исполнения виртуальных экспериментов

В данном разделе представлена архитектура платформы (см. рисунок) и рассмотрено, каким образом компоненты архитектуры поддерживают различные этапы деятельности эксперта по проведению виртуальных экспериментов. На рисунке штриховые стрелки обозначают доступ из одного компонента к интерфейсу другого компонента. Программная архитектура системы исполнения виртуального эксперимента включает в себя компоненты, поддерживающие различные этапы проведения виртуальных экспериментов. В настоящее время архитектура находится в стадии программной реализации.



Архитектура платформы исполнения виртуальных экспериментов

Компонент *Metadata Repository* представляет собой репозиторий метаинформации (базу данных), предназначенный для хранения результатов выполнения виртуального эксперимента с учетом накопления данных о добавлении, модификации и удалении гипотез, потоков работ, онтологий и виртуальных экспериментов. Дополнительно в репозитории сохраняются построенные зависимости между гипотезами (полученными либо от эксперта, либо автоматически извлеченные) и имеющейся информацией о корреляции между параметрами как в рамках одной, так и параметров нескольких гипотез.

Компонент *Virtual Experiment Manager* представляет пользователю интерфейс с операциями формального манипулирования виртуальными экспериментами и гипотезами. После загрузки виртуального эксперимента в систему его спецификация сохраняется в базе данных. Далее по запросу эксперта происходит вызов компонента автоматического порождения гипотез из данных *Hypothesis Generator*. После этого автоматически происходит вызов компонента *Hypothesis Lattices Constructor* для построения решетки гипотез виртуального эксперимента. Далее вызывается компонент *Virtual Experiment Planner*, отвечающий за построение плана исполнения виртуального эксперимента. Построенный план сохраняется в базе данных. Наконец, происходит вызов компонента *Virtual Experiment Runner*, ответственного за непосредственное исполнение виртуального эксперимента. Компонент поддерживает одновременное управление несколькими экспериментами.

Компонент *Hypothesis Generator* по запросу эксперта обеспечивает автоматическое построение гипотез по данным в виде нелинейных функциональных зависимостей. В качестве основного метода используется символьная регрессия, обеспечивающая интерпретируемость полученных формул. Эксперту вместе с системой уравнений возвращается оценка качества соответствия данным, по которой принимается решение об использовании построенной гипотезы. Компонент реализован в виде модуля на языке Python с использованием библиотек pandas, numpy, sympy и deap [5]. Для построения функциональной зависимости между переменными в данных используется реализация символьной регрессии на основе генетического программирования с арифметическими и тригонометрическими операциями [7].

Компонент *Hypothesis Lattices Constructor* обеспечивает автоматическое построение решеток гипотез по соответствующим им системам уравнений. Компонент загружает из базы данных метаинформации спецификацию потока работ и гипотез. Далее системы уравнений, соответствующих гипотезам, из спецификации переводятся во внутреннее представление, после чего вызывается компонент *COA Constructor* для построения причинно-следственного графа зависимостей переменных. Дальнейшая работа компонента состоит в следующем. Задачи, составляющие поток работ, рассматриваются по очереди в соответствии с порядком, заданным направлением ребер потока работ. Для каждой пары гипотез  $I$  и  $J$ , где  $I$  используется в текущей задаче, а  $J$  — в некоторой задаче, достижимой из текущей задачи, конструируется транзитивное замыкание объединения структур двух

гипотез (состоящих из уравнений и переменных) [4]. После этого производится объединение всех таких транзитивных замыканий. Структура результирующей решетки соответствует структуре объединения замыканий. Решетка гипотез сохраняется в базе данных метаинформации. Компонент реализован в виде модуля на языке Python с использованием библиотек pandas, numpy и itertools.

Компонент *COA Constructor* позволяет построить причинно-следственный граф зависимостей переменных одной гипотезы. Компонент включает механизм вероятностного вывода, позволяющего отследить не только причинно-следственные связи между гипотезами, но и неявные зависимости, например значимые корреляции между параметрами гипотез, не связанных причинно-следственными зависимостями. Входными данными компонента служат пары гипотез и решетка гипотез, а выходными — пары значимо коррелирующих переменных, соответствующих параметрам различных независимых гипотез. Использование пар значимо коррелирующих переменных позволяет системе указывать эксперту, например, на необходимость их совместного изменения. Компонент реализован в виде модуля на языке Python с использованием библиотек pandas, numpy и scipy. При реализации используется информация о вхождении переменных в уравнения в виде булевой матрицы. При этом установление причинно-следственных связей сводится к поиску подматриц с ненулевым рангом.

Компонент *Virtual Experiment Planner* вычисляет оптимальный план исполнения виртуального эксперимента, при этом используются решетки гипотез, содержащиеся в репозитории. Это позволяет до проведения виртуального эксперимента спланировать оптимальным образом подбор параметров гипотез и отсеять заранее непригодные комбинации параметров. Например, уже вычисленные результаты вызова функции модели могут быть повторно использованы в другом виртуальном эксперименте, а также могут помочь в исключении виртуального эксперимента, производящего данные, не совпадающие с реальными наблюдениями.

Компонент *Virtual Experiment Runner* запускается для непосредственного исполнения виртуального эксперимента. Из репозитория метаинформации загружаются спецификация виртуального эксперимента и дополнительные построенные решетка гипотез и план проведения виртуального эксперимента. Исполнение происходит в соответствии с планом, при этом система использует сохраненные результаты исполнения части виртуального эксперимента (при их наличии в репозитории метаинформации). Используя спецификацию моделей виртуального эксперимента, компонент осуществляет вызов соответствующих функций, а затем собирает полученные данные от них. Результаты проведения виртуального эксперимента сохраняются в базе данных метаинформации.

## 4 Заключение

В работе проведен анализ методов управления экспериментами, реализованных в существующих системах. Рассмотрены основные этапы жизненного цикла

виртуального эксперимента; выделены операции, используемые на каждом этапе жизненного цикла. Операции разделены на автоматические и вызываемые экспертом. Предложена программная архитектура платформы управления виртуальными экспериментами, выделены ее основные компоненты, реализующие операции жизненного цикла виртуального эксперимента.

Платформа находится в стадии программной реализации и экспериментальных исследований. Реализованы компоненты построения причинно-следственного графа зависимостей переменных одной гипотезы, компонент автоматического построения решеток гипотез по соответствующим им системам уравнений, компонент автоматического порождения гипотез из данных, компонент непосредственного исполнения виртуального эксперимента. Остальные компоненты платформы находятся в стадии разработки. Планируются экспериментальные исследования применения платформы для проведения виртуальных экспериментов в нескольких областях с интенсивным использованием данных, в том числе в астрономии и нейрофизиологии.

## Литература

1. Kovalev D., Kalinichenko L., Stupnikov S. Organization of virtual experiments in data-intensive domains: Hypotheses and workflow specification // CEUR Workshop Proceedings: Selected Papers of the 19th Conference (International) on Data Analytics and Management in Data Intensive Domains / Eds. L. Kalinichenko, Ya. Manolopoulos, N. Skvortsov, V. Sukhomlin. — Moscow, 2017. Vol. 2022. P. 293–300. <http://ceur-ws.org/Vol-2022/paper46.pdf>.
2. Grau B. C., Horrocks I., Motik B., Parsia B., Patel-Schneider P., Sattler U. OWL 2: The next step for OWL // J. Web Semant., 2008. Vol. 6. Iss. 4. P. 309–322.
3. Van Der Aalst W. M., Ter Hofstede A. H., Kiepuszewski B., Barros A. P. Workflow patterns // Distrib. Parallel Dat., 2003. Vol. 14. Iss. 1. P. 5–51.
4. Kovalev D., Stupnikov S. Constructing hypothesis lattices for virtual experiments in data intensive research // Ivannikov Memorial Workshop. — IEEE, 2019. P. 14–17.
5. Tirikov E. Comparison of male and female nonlinear brain functional connectivity // CEUR Workshop Proceedings: Selected Papers of the 21st Conference (International) on Data Analytics and Management in Data Intensive Domains / Eds. A. Elizarov, B. Novikov, S. Stupnikov. — Kazan, 2019. Vol. 2523. P. 404–412. <http://ceur-ws.org/Vol-2523/paper39.pdf>
6. Duggan J., Brodie M. Hephaestus: Data reuse for accelerating scientific discovery // 7th Biennial Conference on Innovative Data Systems Research: Online Proceedings, 2015. Paper 29. 12 p. [http://users.eecs.northwestern.edu/~jennie/pubs/hephaestus\\_full.pdf](http://users.eecs.northwestern.edu/~jennie/pubs/hephaestus_full.pdf).
7. Uy N. Q., Hoai N. X., O'Neill M., McKay R. I., Galván-López E. Semantically-based crossover in genetic programming: Application to real-valued symbolic regression // Genet. Program. Evol. M., 2011. Vol. 12. Iss. 2. P. 91–119.
8. Triantafillou S., Tsamardinos I. Constraint-based causal discovery from multiple interventions over overlapping variable sets // J. Mach. Learn. Res., 2015. Vol. 16. Iss. 1. P. 2147–2205.

9. Neal R. M. Slice sampling // Ann. Stat., 2003. Vol. 31. Iss. 3. P. 705–741.
10. Goncalves B., Silva F., Porto F. Upsilon-DB: A system for data-driven hypothesis management and analytics // arXiv.org, 2014. arXiv:1411.7419 [cs.DB]. 6 p.
11. Darwiche A. Modeling and reasoning with Bayesian networks. — Cambridge University Press, 2009. 548 p.
12. Schales D., Hu X., Jang J., et al. FCCE: Highly scalable distributed feature collection and correlation engine for low latency big data analytics // IEEE 31st Conference (International) on Data Engineering. — IEEE, 2015. P. 1316–1327.
13. King R. D., Whelan K. E., Jones F. M., Reiser P. G., Bryant C. H., Muggleton S. H., Kell D. B., Oliver S. G. Functional genomic hypothesis generation and experimentation by a robot scientist // Nature, 2004. Vol. 427. Iss. 6971. P. 247–252.
14. Kakas A. C., Kowalski R. A., Toni F. Abductive logic programming // J. Logic Comput., 1992. Vol. 2. Iss. 6. P. 719–770.
15. Тарасов Е., Ковалев Д. Оценка качества научных гипотез в виртуальных экспериментах в областях с интенсивным использованием данных // CEUR Workshop Proceedings: Selected Papers of the 19th Conference (International) on Data Analytics and Management in Data Intensive Domains / Eds. L. Kalinichenko, Ya. Manolopoulos, N. Skvortsov, V. Sukhomlin. — Moscow, 2017. Vol. 2022. P. 281–292. <http://ceur-ws.org/Vol-2022/paper46.pdf>.

Поступила в редакцию 15.03.20

---

## ARCHITECTURE OF THE PLATFORM FOR MANAGING HYPOTHESES-DRIVEN VIRTUAL EXPERIMENTS

**D. Y. Kovalev<sup>1</sup>, E. A. Tarasov<sup>1</sup>, V. N. Zakharov<sup>1</sup>, and N. M. Filimonov<sup>2</sup>**

<sup>1</sup>Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

<sup>2</sup>Institute of Automatics and Computer Engineering, National Research University “Moscow Power Engineering Institute,” 14 Krasnokazarmennaya Str., Moscow 111250, Russian Federation

**Abstract:** The problem of carrying out virtual experiments driven by hypotheses is considered. The analysis of methods for managing virtual experiments in existing systems for working with experiments is carried out. According to the results of the analysis, a life cycle of a virtual experiment is formed. Basic operations of managing virtual experiments are presented for the stages of an experiment life cycle, as well as the main stages of the expert’s work with the platform for executing a virtual experiment. The program architecture of the platform for managing virtual experiments and hypotheses is proposed with

a description of the main components of the platform and their functions that implement the basic operations of the life cycle.

**Keywords:** virtual experiment life cycle; hypotheses; data intensive research

**DOI:** 10.14357/08696527200206

## Acknowledgments

This research was partially supported by the Russian Foundation for Basic Research (project 18-07-01434 A).

## References

1. Kovalev, D., L. Kalinichenko, and S. Stupnikov. 2017. Organization of virtual experiments in data-intensive domains: Hypotheses and workflow specification. *CEUR Workshop Proceedings: Selected Papers of the 19th Conference (International) on Data Analytics and Management in Data Intensive Domains*. Eds. L. Kalinichenko, Ya. Manolopoulos, N. Skvortsov, and V. Sukhomlin. Moscow. 2022:293–300. Available at: <http://ceur-ws.org/Vol-2022/paper46.pdf> (accessed April 9, 2020).
2. Grau, B. C., I. Horrocks, B. Motik, B. Parsia, P. Patel-Schneider, and U. Sattler. 2008. OWL 2: The next step for OWL. *J. Web Semant.* 6(4):309–322.
3. Van Der Aalst, W. M., A. H. Ter Hofstede, B. Kiepuszewski, and A. P. Barros. 2003. Workflow patterns. *Distrib. Parallel Dat.* 14(1):5–51.
4. Kovalev, D., and S. Stupnikov. 2019. Constructing hypothesis lattices for virtual experiments in data intensive research. *Ivanov Memorial Workshop*. IEEE. 14–17.
5. Tirikov, E. 2019. Comparison of male and female nonlinear brain functional connectivity. *CEUR Workshop Proceedings: Selected Papers of the 21st Conference (International) on Data Analytics and Management in Data Intensive Domains*. Eds. A. Elizarov, B. Novikov, and S. Stupnikov. Kazan. 2523:404–412. Available at: <http://ceur-ws.org/Vol-2523/paper39.pdf> (accessed April 9, 2020).
6. Duggan, J., and M. Brodie. 2015. Hephaestus: Data reuse for accelerating scientific discovery. *7th Biennial Conference on Innovative Data Systems Research: Online Proceedings*. Papre 29. 12 p. Available at: [http://users.eecs.northwestern.edu/~jennie/pubs/hephaestus\\_full.pdf](http://users.eecs.northwestern.edu/~jennie/pubs/hephaestus_full.pdf) (accessed April 9, 2020).
7. Uy, N. Q., N. X. Hoai, M. O'Neill, R. I. McKay, and E. Galván-López. 2011. Semantically-based crossover in genetic programming: Application to real-valued symbolic regression. *Genet. Program. Evolv. M.* 12(2):91–119.
8. Triantafillou, S., and I. Tsamardinos. 2015. Constraint-based causal discovery from multiple interventions over overlapping variable sets. *J. Mach. Learn. Res.* 16(1):2147–2205.
9. Neal, R. M. 2003. Slice sampling. *Ann. Stat.* 31(3):705–741.
10. Goncalves, B., F. Silva, and F. Porto. Upsilon-DB: A system for data-driven hypothesis management and analytics. *arXiv.org*. 6 p. Available at: <https://arxiv.org/abs/1411.7419> (accessed April 9, 2020).
11. Darwiche, A. 2009. *Modeling and reasoning with Bayesian networks*. Cambridge University Press. 548 p.

12. Schales, D., X. Hu, J. Jang, et al. 2015. FCCE: Highly scalable distributed feature collection and correlation engine for low latency big data analytics. *31st Conference (International) on Data Engineering*. IEEE. 1316–1327.
13. King R. D., K. E. Whelan, F. M. Jones, P. G. Reiser, C. H. Bryant, S. H. Muggleton, D. B. Kell, and S. G. Oliver. 2004. Functional genomic hypothesis generation and experimentation by a robot scientist. *Nature* 427(6971):247–252.
14. Kakas, A. C., R. A. Kowalski, and F. Toni. 1992. Abductive logic programming. *J. Logic Comput.* 2(6):719–770.
15. Tarasov E., and D. Kovalev. 2017. Otsenka kachestva nauchnykh gipotez v virtual'nykh eksperimentakh v oblastyakh s intensivnym ispol'zovaniem dannykh [Estimation of scientific hypotheses quality in virtual experiments in data intensive domains]. *CEUR Workshop Proceedings: Selected Papers of the 19th Conference (International) on Data Analytics and Management in Data Intensive Domains*. Eds. L. Kalinichenko, Ya. Manolopoulos, N. Skvortsov, and V. Sukhomlin. Moscow. 2022:281–292. Available at: <http://ceur-ws.org/Vol-2022/paper46.pdf> (accessed April 9, 2020).

Received March 15, 2020

## Contributors

**Kovalev Dmitry Y.** (b. 1988)— scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; dkovalev@ipiran.ru

**Tarasov Evgeny A.** (b. 1987) — programmer, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; e.tarasov@outlook.com

**Zakharov Victor N.** (b. 1948) — Doctor of Science in technology, associate professor, scientific secretary, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; vzakharov@ipiran.ru

**Filimonov Nikita M.** (b. 1998) — student, Institute of Automatics and Computer Engineering, National Research University “Moscow Power Engineering Institute,” 14 Krasnokazarmennaya Str., Moscow 111250, Russian Federation; filimonovn160@gmail.com

## ПОДХОДЫ К ИНТЕГРАЦИИ ПРИКЛАДНЫХ КОНЦЕПТУАЛЬНЫХ СХЕМ В СОСТАВЕ УНИФИЦИРОВАННОЙ ГЕООНТОЛОГИИ

*С. К. Дулин<sup>1</sup>, Д. А. Никишин<sup>2</sup>*

**Аннотация:** Представлен подход к обеспечению согласования разнородных прикладных концептуальных схем (КС), с достаточной полнотой описывающих отдельные предметные области, на основе их унификации и интеграции в форме универсальной, единой геоонтологии. Рассмотрены возможные факторы неоднородности частных КС, представлены требования и основные проектные решения для построения унифицированной геоонтологии, предназначенный для использования в контексте многоаспектной базы геоданных (БГД). Показана структура данных для построения такой унифицированной геоонтологии.

**Ключевые слова:** согласование прикладных концептуальных схем; унифицированная геоонтология

**DOI:** 10.14357/08696527200207

### 1 Введение

Статья посвящена продолжению исследований особенностей информационных трансформаций геоданных в географических информационных системах (ГИС) в контексте парадигмы полиадического компьютеринга [1–3]. Конечной целью ставится создание структуры БГД, обеспечивающей потенциал для совершенствования методологии анализа и отображения информации о местности, приуроченной к различным уровням детализации и специфическим прикладным моделям геоданных [4].

Задача данной публикации — анализ факторов неоднородности частных КС и разработка подхода к их согласованию посредством унификации и интеграции в виде единой КС (ЕКС) классификации топографической информации, которая будет служить основой для создания единой унифицированной геоонтологии.

### 2 Онтологии и пространственные типы данных

Характерной особенностью современных ГИС становится использование различных онтологий, поддерживающих семантическое моделирование при синтезе информации.

<sup>1</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, skdulin@mail.ru

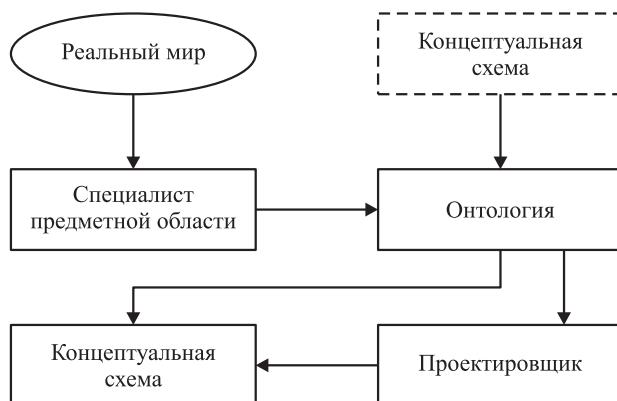
<sup>2</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, dmnikishin@mail.ru

Онтология в информатике, согласно современным толкованиям [5], является «точной спецификацией концептуализации предметной области», которая должна содержать тезаурус терминов и определенные спецификации. Онтологии содержат взаимосвязанные факты, объединенные в классы объектов и отношений, соответствующие выбранной проблемной области. По сути, онтология — это наиболее полная концептуализация проблемной области [5]. Онтологии способны улучшить работу ГИС, устранивая противоречия и конфликты между концепциями и реализациями программного обеспечения.

Можно классифицировать онтологии согласно их зависимости от специфики задачи [6]:

- онтологии верхнего уровня, описывающие обобщенные понятия;
- онтологии типов геоданных, описывающие тезаурус, связанный с доменом концепций геообъектов (типа дистанционного восприятия перегона или железнодорожной станции);
- онтологии задачи, описывающие задачу или действие типа интерпретации изображения или оценки вероятности;
- прикладные онтологии, описывающие концепции, зависящие как от типов геоданных, так и от задачи; они определяют требования пользователя с учетом специфики приложения.

Концептуальные схемы БГД могут способствовать разработке онтологий, представляя собой описание специалистом части проблемной области, пригодное для решения конкретных задач. С другой стороны, разработанные онтологии могут использоваться для создания КС геоданных (рис. 1). Обычно выделяют три уровня, на которых находятся как онтологии, так и КС (рис. 2).



**Рис. 1** Порядок создания КС

	Онтологии	Концептуальные модели
Формальный уровень	Абстракция формальных особенностей научных предметов: геометрия, география, время, пространство	Понятия концептуального моделирования: объекты {цели}, области{поля}, отношения
Уровень логических типов данных прикладной области	Онтология географических видов: представление, местоположение, топология, географическая информатика	Концептуальная модель и нотация: классы, пространственные отношения, пространственные ограничения целостности
Прикладной уровень	Онтология инфраструктуры объектов ОАО «РЖД»: станции, перегоны и др.	Концептуальная схема БГД, атрибутивное описание сущностей ОАО «РЖД»

Рис. 2 Три уровня абстракций для онтологий и КС

Первый уровень — формальный, на котором онтологии и КС содержат абстрактные концепции. Онтологии на этом уровне содержат формальные особенности геоданных. Концептуальные схемы на этом уровне содержат концепции и отношения между ними.

При сопоставлении содержания формального уровня с геоданными можно выделить второй уровень — типов геоданных. Здесь онтологии представляют тезаурус географических видов, который отражает знания о геоданных. Этот уровень содержит географические понятия, географические сущности и процессы. Конкретизация онтологий на прикладном уровне приводит к конкретизации предыдущего уровня и задает конкретные сущности, определенные приложением, которые допустимы для совместного использования легитимными пользователями. Три уровня абстракции необходимы, чтобы верифицировать КС, переходя от концептуального описания к уровню типов геоданных и далее к прикладному уровню.

Онтологии прикладного уровня фиксируют конкретизацию концепций геоинформационного контента. В [6, 7] вводится понятие географической информационной системы, управляемой онтологиями (ODGIS — ontology-driven geographic information system). Идея управления с помощью онтологий позволяет фиксировать географические описания специальными процедурами, функционирующими на основе наборов геоданных. Следовательно, необходимы онтологии не только для сущностей, но и для действий над наборами геоданных.

В [8] предполагается, что формальное описание геоданных могло бы привести к целостной онтологии географических описаний. Становится неоспоримой

важность онтологии как практического средства совместного использования геоконтента.

### **3 Проблема согласования частных концептуальных схем и подход к ее решению**

Даже в контексте одной БГД, не говоря уже об интеграции пространственных данных из разных источников, может иметь место неоднородность в используемых КС<sup>1</sup>.

Одним из факторов неоднородности выступают различия в КС, предназначенных для представления геоданных применительно к определенным задачам их использования. Характерным примером может служить, например, КС кадастра, использующая концепт полигона кадастрового участка, и КС для описания транспортной сети, основанная на концепте графа транспортных маршрутов. При этом концепты обеих схем могут быть топологически сопряжены между собой (например, дорога и границы кадастровых участков).

Кроме того, в рамках даже одной КС «может понадобиться как площадная метрика дороги для отображения объекта на карте, так и ее осевая линия — для решения навигационных задач» [9, с. 48]. В первом случае речь идет о площадной метрике дороги для визуализации ситуации, а во втором — об осевой линии дороги, используемой для построения маршрута в контексте графа дорожной сети.

Таким образом, первый фактор обуславливает различия в моделях описания геообъектов в рамках одной или нескольких КС.

Вторым источником различий становится использование отдельных КС (классификаторов топографической информации) на разных уровнях генерализации данных (масштабных уровнях их обобщенного представления). Примером может служить ряд отдельных классификаторов для каждого масштабного уровня карты, например представленные ресурсом [10].

Можно выделить еще три фактора различия КС:

- различная номенклатура свойств у сопоставляемых эквивалентных по смыслу моделей (классов) геообъектов;
- различия в номенклатуре номинаций у сопоставляемых эквивалентных свойств;
- различия области значений, покрываемых аналогичными по смыслу номинациями.

---

<sup>1</sup> Концептуальная схема здесь понимается как система взаимосвязанных понятий, необходимая и достаточная для описания требуемого аспекта моделируемого объекта, включающая модели данных (классы геообъектов), сопряженные с ними ограничения и методы их обработки.

Таким образом, возможно одновременное существование ряда наборов геоданных, представляющих ситуацию на местности для различных прикладных задач, масштабов и т. п. на основе соответствующих КС.

Учитывая, что актуализация данных в БГД предполагает также каскадное обновление на уровнях их генерализированных представлений, а также необходимость согласования между собой различных прикладных версий представления этих данных, наличие неоднородностей в используемых частных КС обусловливает проблематику реализации каскадного согласования вариантов представления геоданных. Эта проблема заключается в необходимости идентификации и согласования различных определений одного и того же свойства как в разных КС, так и в пределах одной КС.

Одним из возможных подходов к решению этой проблемы может быть унификация и интеграция нескольких частных КС в виде ЕКС классификации топографической информации. Такой подход должен обеспечить формализацию процессов генерализации и согласования различных представлений ситуации, в том числе с целью их автоматизации. Обеспечение автоматизации этих процессов жизненно необходимо для современных ГИС, к которым предъявляются требования оперативной актуализации данных, их анализа и визуализации, в том числе в реальном масштабе времени. Кроме того, ЕКС также может способствовать совершенствованию методов анализа геоданных и решения по ним различных задач.

## **4 Требования к единой концептуальной схеме**

К концепции создаваемой ЕКС можно выдвинуть следующие требования.

1. Существующие традиционные КС обычно оперируют в качестве модели описания геообъектов не их аналитическим представлением (аналитическими моделями, АМ), а моделями их картографического изображения — условными знаками (УЗ). Между АМ и УЗ есть ряд различий, которые приводят к определенным следствиям:
  - (а) отсутствие в большинстве существующих систем УЗ полноценной, однозначной связи как между геообъектом и УЗ, так и различных УЗ между собой. Для первого случая примером могут служить геообъекты типа «овраг», «насыпь», «выемка», которые передаются несколькими структурными элементами (брюка, подошва), при этом явное связывание этих компонентов в более сложные объекты (овраг, вал и т. п.) не предусматривается, эта их связь может быть восстановлена лишь на ментальном уровне или путем сложных алгоритмов/процедур пространственного анализа. Примером второго случая служат дороги, которые могут претерпевать разрывы в населенных пунктах, на мостах/эстакадах, насыпях и т. п. Процедуры восстановления их конфигурации также не являются тривиальными. И, в-третьих, в большинстве существующих систем

УЗ не предусматривается связывание отдельных геообъектов компонентов в более сложные структуры (по функциональному, топологическому и другим признакам), например объединение дорог и других элементов в дорожную сеть. Это также ограничивает возможности анализа;

- (б) для некоторых геообъектов при определенных значениях их свойств могут применяться разные способы их картографического изображения (т. е. использоваться разные УЗ). Это приводит к тому, что при определенных изменениях значений свойств геообъекта вместо соответствующей модификации данных его модели, по сути, происходит удаление этого геообъекта (как УЗ, соответствующего предшествующему состоянию) и создание вместо него нового геообъекта (соответствующего УЗ нового состояния). Отсутствие связи между старым и новым УЗ геообъекта делает невозможным установление их темпоральной преемственности [4];
- (в) в контексте генерализации УЗ на мелкомасштабных уровнях обобщения могут подвергаться приемам смещения, утрирования и т. п., что ограничивает их использование для целей анализа.

Резюмируя, можно отметить, что ЕКС целесообразно строить на основе АМ геообъектов в БГД. Элементы данных, относящиеся к визуализации геообъектов посредством УЗ, должны играть подчиненную роль по отношению к основной КС.

2. Единая КС должна поддерживать однозначность и обратимость взаимосвязей между представлениями геообъектов на смежных уровнях обобщения [4, 11]. Использование АМ позволит формализовать правила традиционной генерализации и способствовать их автоматизации. Первая очередь ЕКС должна охватывать основной ряд масштабных уровней: 1:10 000, 1:25 000, 1:50 000, 1:100 000, 1:200 000, 1:500 000, 1:1 000 000. Вторая очередь ЕКС предполагает обеспечение крупномасштабных геоданных (масштабы 1:500–1:5 000). В перспективе целесообразен охват более мелких масштабов обзорно-географических карт (до 1:10 000 000 и мельче).
3. Единая КС должна поддерживать различные версии представления геоданных в необходимых тематических аспектах, т. е. интегрировать в себе частные КС. Необходимость этого и круг этих аспектов определяются кругом задач, стоящих перед такой БГД и ГИС.
4. Опционально ЕКС может выступать в качестве основы для интеграции геоинформационных ресурсов с разнородными КС.

## 5 Структура единой концептуальной схемы

Рассмотрим подходы к структуре и содержанию ЕКС и использующей ее БГД. Структура ЕКС включает следующие концепты:

- класс геообъектов  $C_i$ ;
- атрибутивное свойство класса геообъектов  $P_{ij}$ ;
- значение атрибутивного свойства  $V_{ijk}$ .

Такое разделение концептов в общем соответствует подходу к структурированию пространственных данных в документе [12].

Основным элементом ЕКС выступает *свойство* (атрибут). По своей сути оно отражает некоторый количественный или качественный признак объектов местности и предполагает область возможных значений, характеризующих степень обладания этим признаком. Формально, с точки зрения организации данных в БГД каждый объект «атрибутивное свойство» объединяет в себе множество возможных «значений» — номинаций, представляющих собой отдельные позиции или интервалы из области значений свойства. Концепт *класса* служит прототипом для АМ геообъектов и образует «корень» КС.

Описание свойств и их значений должно соответствовать правилам классификации, которые применительно к данному случаю можно сформулировать так:

- отсутствие смыслового перекрытия между свойствами;
- отсутствие в одном свойстве разных оснований;
- отсутствие у свойства неохватываемой области значений (полнота охвата множеством значений области значений этого свойства);
- отсутствие у свойства смыслового перекрытия номинаций его значений.

Структура ЕКС образует концептуальную составляющую БГД, содержательная часть этой БГД дополняется концептом геообъекта  $G_n$  — моделью объекта местности, которая характеризуется определенными значениями для каждого из множества атрибутивных свойств, входящих в состав класса; с этой целью в БГД предусмотрены соответствующие взаимосвязи. Структура БГД, использующей ЕКС, проиллюстрирована на рис. 3.

На рис. 3 левый блок содержит три таблицы («Классы», «Свойства» и «Значения»), которые связаны между собой основными отношениями «один ко многим» и образуют ЕКС. Справа расположен блок содержательной части БГД,

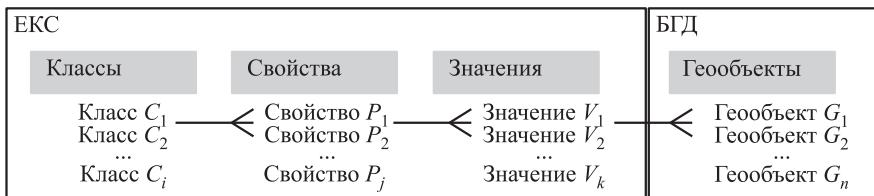


Рис. 3 Структура БГД с использованием ЕКС

в котором показана одна таблица «Геообъекты» (технические детали на рисунке исключены).

## 6 Заключение

Единая КС позиционируется как основа для информационных трансформаций геоданных. Рассмотрены факторы неоднородности частных КС, подход к их согласованию на основе унифицированной КС и требования к ее построению. Представлена общая схема организации БГД, использующей ЕКС, в которой выделены основные концепты и взаимосвязи между ними.

В качестве образцов прикладных КС для формирования фрагмента экспериментальной унифицированной КС предполагается использовать структуры классов объектов открытого краудсорсингового ресурса OpenStreetMap и классификации топографической информации, принятые для отечественных топографических карт.

## Литература

1. Rosenbloom P. S. On computing: The fourth great scientific domain. — Cambridge, MA, USA: MIT Press, 2013. 307 p.
2. Зацман И. М. Методология обратимой генерализации в контексте классификации информационных трансформаций // Системы и средства информатики, 2018. Т. 28. № 2. С. 128–144.
3. Гончаров А. А., Зацман И. М. Информационные трансформации параллельных текстов в задачах извлечения знаний // Системы и средства информатики, 2019. Т. 29. № 1. С. 180–193.
4. Никишин Д. А. Структура и особенности генерализации в контексте функционирования темпоральной базы геоданных // Системы и средства информатики, 2020. Т. 30. № 1. С. 147–159.
5. McDaniel M., Storey V. C. Evaluating domain ontologies: Clarification, classification, and challenges // ACM Comput. Surv., 2019. Vol. 52. Iss. 4. Art. No. 70. 44 p.
6. Li N., Raskin R., Goodchild M., Janowicz K. An ontology-driven framework and Web Portal for Spatial Decision Support // T. GIS, 2012. Vol. 16. Iss. 3. P. 313–329.
7. Jingya Y., Guilbert E., Saux E. An ontology-driven multi-agent system for nautical chart generalization // Cartogr. Geogr. Inf. Sc., 2016. Vol. 44. Iss. 3. P. 1–15.
8. Hornsby K., Egenhofer M. Identity-based change: A foundation for spatio-temporal knowledge representation // Int. J. Geogr. Inf. Sci., 2000. Vol. 14. Iss. 3. P. 207–224.
9. Важенин И. А., Гусачев М. С., Карманов Д. В. Технология хранения мультимасштабных пространственных данных в объектно-ориентированной базе геоданных ArcGIS // Информация и космос, 2013. № 1. С. 48–52.
10. ГИС «Панорама». Классификаторы. [www.gisinfo.ru/classifiers/classifiers.htm](http://www.gisinfo.ru/classifiers/classifiers.htm).

11. Никишин Д. А. Процессы генерализации в аналоговой и цифровой картографии // Системы и средства информатики, 2018. Т. 28. № 3. С. 204–216.
12. ГОСТ Р 52439-2005. Модели местности цифровые. Каталог объектов местности. Требования к составу. — М: Стандартинформ, 2006. 58 с.

*Поступила в редакцию 21.02.20*

---

## **APPROACHES TO THE INTEGRATION OF THE APPLICATION CONCEPTUAL SCHEMAS IN THE UNIFIED GEOONTOLOGY**

***S. K. Dulin and D. A. Nikishin***

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** An approach is presented to ensure the coordination of heterogeneous applied conceptual schemes that describe individual subject areas with sufficient completeness, based on their unification and integration in the form of a universal, unified geoontology. Possible factors of heterogeneity of particular conceptual schemes are considered, requirements and main design solutions for building an unified geoontology intended for use in the context of a multiaspect geodatabase are presented. The data structure for building such an unified geoontology is shown.

**Keywords:** harmonization of the application conceptual schema; unified geoontology

**DOI:** 10.14357/08696527200207

## **References**

1. Rosenbloom, P. S. 2013. *On computing: The fourth great scientific domain*. Cambridge, MA: MIT Press. 307 p.
2. Zatsman, I. M. 2018. Metodologiya obratimoy generalizatsii v kontekste klassifikatsii informatsionnykh transformatsiy [Methodology of reversible generalization in the context of classification of information transformations]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(2):128–144.
3. Goncharov, A. A., and I. M. Zatsman. 2019. Informatsionnye transformatsii parallel'nykh tekstov v zadachakh izvlecheniya znanii [Information transformations of parallel texts in problems of knowledge extraction]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(1):180–193.
4. Nikishin, D. A. 2020. Struktura i osobennosti generalizatsii v kontekste funktsionirovaniya temporal'noy bazy geodannyykh [Structure and features of generalization in the context of functioning of a temporal geodatabase]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(1):147–159.

5. McDaniel, M., and V. C. Storey. 2019. Evaluating domain ontologies: Clarification, classification, and challenges. *ACM Comput. Surv.* 52(4):70. 44 p.
6. Li, N., R. Raskin, M. Goodchild, and K. Janowicz. 2012. An ontology-driven framework and Web Portal for Spatial Decision Support. *T. GIS* 16(3): 313–329.
7. Jingya, Y., E. Guibert, and E. Saux. 2016. An ontology-driven multi-agent system for nautical chart generalization. *Cartogr. Geogr. Inf. Sc.* 44(3):1–15.
8. Hornsby, K., and M. Egenhofer. 2000. Identity-based change: A foundation for spatio-temporal knowledge representation. *Int. J. Geogr. Inf. Sci.* 14(3):207–224.
9. Vazhenin, I. A., M. S. Gusachev, and D. V. Karmanov. 2013. Tekhnologiya khra-neniya mul'timasshtabnykh prostranstvennykh dannykh v ob"ektno-orientirovannoy baze geodannykh ArcGIS [ArcGIS object-oriented database multiscale spatial data storage technology]. *Informatsiya i kosmos* [Information and Space] 1:48–52.
10. GIS “Panorama.” Klassifikatory [Classifiers]. Available at: <https://gisinfo.ru/classifiers/classifiers.htm> (accessed March 19, 2020).
11. Nikishin, D. A. 2018. Protsessy generalizatsii v analogovoy i tsifrovoy kartografii [A generalization processes in analog and digital cartography]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(3):204–216.
12. GOST R 52439-2005. 2006. *Modeli mestnosti tsifrovye. Katalog ob"ektov mestnosti. Trebovaniya k sostavu* [Digital Terrain models. Catalog of locality objects. Requirements for the composition]. Moscow: Standardinform Publs. 58 p.

Received February 21, 2020

## Contributors

**Nikishin Dmitry A.** (b. 1976)— Candidate of Science (PhD) in technology, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; dmnikishin@mail.ru

**Dulin Sergey K.** (b. 1950)— Doctor of Science in technology, professor; leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; skdulin@mail.ru

## ФУНКЦИОНАЛЬНАЯ СТРУКТУРА СПЛОЧЕННОЙ ГИБРИДНОЙ ИНТЕЛЛЕКТУАЛЬНОЙ МНОГОАГЕНТНОЙ СИСТЕМЫ\*

*С. В. Листопад<sup>1</sup>*

**Аннотация:** Эффективность работы систем гибридного и синергетического искусственного интеллекта достигается преимущественно путем правильной организации взаимодействия их элементов, а не повышения сложности или интеллектуальности последних. В случае с гибридными интеллектуальными многоагентными системами (ГиИМАС), моделирующими коллективы экспертов, решающих проблемы «за круглым столом», для эффективной организации взаимодействия агентов должны выполняться условия непротиворечивости их целей и моделей предметной области, единства протокола решения проблемы, совместимости языков передачи сообщений. Особенно актуальны данные условия, когда ГиИМАС строится из агентов, разработанных различными независимыми коллективами. В этом случае агенты должны быть способны самостоятельно согласовать цели, модели предметной области и выработать протокол для решения поставленной перед ними проблемы в рамках интеллектуальной системы нового класса — сплоченной ГиИМАС (СГИМАС). В данной статье рассматривается функциональная структура такой системы.

**Ключевые слова:** сплоченность; гибридная интеллектуальная многоагентная система; коллектив экспертов

**DOI:** 10.14357/08696527200208

### 1 Введение

Как показано А. С. Нариньянни, проблемы, возникающие в практике принятия решений, характеризуются множеством НЕ-факторов: неоднородностью, недоопределенностью, неточностью, нечеткостью, неполнотой и др. [1]. У В. Ф. Спирионова понятие «проблема» характеризуется слабой формализацией, комплексным строением, сетевым характером условий и целей (политичностью), непрозрачностью (неопределенностью), а также субъективностью и динамичностью [2]. Практическое исследование проблем, возникающих при управлении объектами транспортной логистики, электроэнергетики, медицины и других областей, показало релевантность данных оценок [3–5]. При разработке интеллектуальных систем для таких условий следует моделировать успешные

\* Исследование выполнено при финансовой поддержке РФФИ (проект 20-07-00104а).

<sup>1</sup> Калининградский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, ser-list-post@yandex.ru

практики решения проблем, применяющиеся в группах экспертов под руководством лица, принимающего решения. Успех работы таких групп зависит от многих факторов: профессионализма участников, их целей, эффективности коммуникации и управления, справедливости распределения обязанностей и ресурсов, организации процесса коллективного решения проблем и др. Во многом успех работы таких групп зависит от способности лидера (формального или неформального) создать из малой группы экспертов-индивидуалистов коллектив, т. е. объединение разделяющих общую цель людей, обладающее целостностью, организацией, распределением функций, структурой руководства и управления.

Вопросы интеграции знаний и рассуждений малой группы экспертов под руководством лица, принимающего решения, исследованы в рамках ГиИМАС, позволяющих динамически и релевантно проблемной ситуации синтезировать над гетерогенным модельным полем метод решения проблемы и имитировать групповую работу. Однако из поля зрения разработчиков ГиИМАС выпадал существенный аспект — формирование сплоченного коллектива агентов, которое особенно актуально при создании агентов различными автономными коллективами разработчиков. В этом случае агенты должны не только «говорить» на одном языке и быть совместимыми с реализованными в системе базовыми протоколами, но и иметь механизмы, обеспечивающие их интеграцию в единый коллектив, такие как обмен информацией о своих целях, интересах, знаниях, опыта с другими агентами, выработка и согласование единых для системы целей, ценностей, норм, средств и способов деятельности, а также их интериоризация каждым агентом. В этой связи актуально моделирование групповой сплоченности коллектива экспертов средствами нового класса ГиИМАС — СГИМАС.

## 2 Групповая сплоченность малого коллектива экспертов

При формировании коллектива экспертов для решения поставленной практической проблемы недостаточно подбирать их исключительно по функциональным требованиям: группа рискует не договориться, погрязнуть в конфликтах или, чрезмерно увлекшись деталями, так и не решить проблему в целом в отведенные для этого сроки [6]. Набор знаний, умений и навыков не гарантирует решения поставленной проблемы, и на первый план выходят вопросы коммуникации, психологической совместимости, конфликтов, сплоченности, т. е. групповые, а не индивидуальные механизмы.

Групповые механизмы в малых группах изучаются в рамках групповой динамики [7] — одного из направлений социальной психологии. В рамках исследований по групповой динамике одно из ведущих мест занимают исследования групповой сплоченности. Групповая сплоченность проявляется в создании единой социально-психологической общности людей, входящих в группу, и предполагает возникновение системы свойств группы, препятствующих нарушению ее психологической целостности [8]. Феномен групповой сплоченности повышает удовлетворенность участников от работы в группе, интенсивность взаимодействия

между ними, а также продуктивность группы в целом. Групповая сплоченность — мера взаимосвязанности членов команды, которая определяется степенью позитивности и интенсивности эмоциональных межличностных отношений всех со всеми, совпадения ориентаций на основные ценности, касающиеся процесса совместной деятельности, а также разделения целей существования группы [6]. Фундаментальной моделью в понимании сплочения групп и коллективов стала стратометрическая концепция А. В. Петровского [9], согласно которой сплоченность группы рассматривается на трех уровнях (стратах):

- (1) внешний уровень (эмоциональные межличностные отношения);
- (2) ценностно-ориентационное единство (отношения опосредованы совместной деятельностью, на основе чего возникает единство основных ценностей);
- (3) ядро (члены группы разделяют цели групповой деятельности, поэтому здесь могут выявляться мотивы выбора членами группы друг друга, которые, в свою очередь, могут опосредоваться общими ценностями: отношением к миру, обществу, труду).

Три слоя групповых структур одновременно могут рассматриваться как три уровня развития группы, в том числе и три уровня развития групповой сплоченности. С целью исследования зависимости эффективности работы коллектива от степени сплоченности, а также влияния норм группового поведения на сплоченность коллектива предлагается выполнить компьютерное моделирование эффекта сплоченности средствами ГиИМАС [3].

### **3 Модель сплоченной гибридной интеллектуальной многоагентной системы**

Гибридные интеллектуальные многоагентные системы — это гибридные интеллектуальные системы (ГиИС), реализующие многоагентный подход [3]. Элементы таких ГиИС реализуются в виде агентов, обладающих свойством автономности [10]. Как и многоагентные системы (МАС), они моделируют взаимодействие автономных агентов между собой и с внешней средой, в результате чего архитектура системы может динамически перестраиваться в соответствии с конкретными функциями (ролями) агентов и установившимися отношениями между ними. Гибридные интеллектуальные многоагентные системы сочетают в себе положительные стороны ГиИС и МАС: благодаря сочетанию нескольких методов искусственного интеллекта они релевантны задачам с высокой сложностью моделирования [3]; за счет имитации взаимодействия экспертов и возникающих при этом коллективных процессов они способны менять свою архитектуру для достижения синергетического эффекта, т. е. получения более качественных решений, чем могли бы найти агенты, работающие порознь [3].

Необходимое условие для реализации СГИМАС в соответствии со стратометрической концепцией А. В. Петровского — разработка модели агента, действующего в соответствии с собственной целью и моделью предметной области,

а не целью и моделью предметной области системы в целом. При этом нужно отметить, что ввиду отсутствия эмоциональной составляющей у агентов, применяющихся для моделирования коллективов экспертов, страта эмоциональных межличностных отношений не рассматривается. Таким образом, на основе модели ГиИМАС [3] может быть сформулирована модель СГИМАС:

$$\text{chimas} = \langle \text{AG}, \text{env}, \text{INT}, \text{ORG}, \{\text{gln}, \text{ontng}, \text{protng}\} \rangle,$$

Здесь AG — множество агентов:

$$\text{AG} = \left\{ \text{ag}^{\text{dm}}, \text{ag}^{\text{fc}}, \text{ag}^{\text{med}}, \text{ag}^{\text{int}}, \text{ag}^{\text{pc}} \right\} \cup \text{AG}^{\text{ex}} \cup \text{AG}^{\text{tr}} \cup \text{AG}^{\text{it}}, \quad (1)$$

где  $\text{ag}^{\text{dm}}$  — агент, принимающий решения,  $\text{ag}^{\text{fc}}$  — агент-фасилитатор,  $\text{ag}^{\text{med}}$  — агент-посредник,  $\text{ag}^{\text{int}}$  — интерфейсный агент,  $\text{ag}^{\text{pc}}$  — агент контроля протокола,  $\text{AG}^{\text{ex}} = \{\text{ag}_1^{\text{ex}}, \dots, \text{ag}_{n_{\text{ex}}}^{\text{ex}}\}$  — подмножество агентов-экспертов ( $n_{\text{ex}}$  — число агентов-экспертов),  $\text{AG}^{\text{tr}} = \{\text{ag}_1^{\text{tr}}, \dots, \text{ag}_{n_{\text{tr}}}^{\text{tr}}\}$  — подмножество агентов-переводчиков ( $n_{\text{tr}}$  — число агентов-переводчиков),  $\text{AG}^{\text{it}} = \{\text{ag}_{\text{an}}^{\text{it}}, \text{ag}_{\text{st}}^{\text{it}}, \text{ag}_{\text{lo}}^{\text{it}}, \text{ag}_{\text{fu}}^{\text{it}}, \text{ag}_{\text{sy}}^{\text{it}}, \text{ag}_{\text{cnv}}^{\text{it}}\}$  — подмножество агентов интеллектуальных технологий ( $\text{ag}_{\text{an}}^{\text{it}}$  — аналитический агент;  $\text{ag}_{\text{st}}^{\text{it}}$  — стохастический агент;  $\text{ag}_{\text{lo}}^{\text{it}}$  — логический агент;  $\text{ag}_{\text{fu}}^{\text{it}}$  — нечеткий агент;  $\text{ag}_{\text{sy}}^{\text{it}}$  — символьный агент;  $\text{ag}_{\text{cnv}}^{\text{it}}$  — агент-преобразователь); env — концептуальная модель внешней среды СГИМАС; INT — элементы структурирования взаимодействий агентов:

$$\text{INT} = \{\text{prot}_{\text{bsc}}, \text{PRC}, \text{LANG}, \text{ont}_{\text{bsc}}\},$$

где  $\text{prot}_{\text{bsc}}$  — базовый протокол, обеспечивающий взаимодействие агентов по формированию протокола сплоченного взаимодействия для решения поставленных перед СГИМАС проблем, PRC — множество элементов для конструирования протокола решения проблем агентами-экспертами и агентом, принимающим решения, LANG — множество языков передачи сообщений, которые используют агенты СГИМАС,  $\text{ont}_{\text{bsc}}$  — базовая, общая для всех агентов СГИМАС онтология, обеспечивающая понимание агентами смысла передаваемых сообщений по согласованию собственных моделей предметной области, целей, формированию протокола сплоченного взаимодействия; ORG — множество архитектур СГИМАС;  $\{\text{gln}, \text{ontng}, \text{protng}\}$  — множество концептуальных моделей макроуровневых процессов в СГИМАС: gln — модель процесса согласования целей агентов между собой, обеспечивающая моделирование сплоченности на уровне ядра согласно концепции А. В. Петровского; ontng — модель процесса согласования моделей предметной области агентов, соответствующая обмену знаниями, опытом и убеждениями между экспертами и моделированию страты ценностно-ориентационного единства; protng — модель процесса формирования агентами протокола сплоченного взаимодействия, обеспечивающая согласование норм взаимодействия и моделирование страты ценностно-ориентационного единства.

Агент  $ag \in AG$  из формулы (1) описывается выражением:

$$ag = \langle id^{ag}, gl^{ag}, LANG^{ag}, ont^{ag}, ACT^{ag} \rangle,$$

где  $id^{ag}$  — идентификатор агента;  $gl^{ag}$  — цель агента;  $LANG^{ag} \subseteq LANG$  — множество языков, сообщения на которых могут быть записаны или прочитаны агентом;  $ont^{ag}$  — модель предметной области агента;  $ACT^{ag}$  — множество действий, реализуемых агентом, среди которых для агентов-экспертов и агента, принимающего решения, выделяются согласование целей  $act_{glng}^{ag}$ , согласование моделей предметной области  $act_{ontng}^{ag}$ , выработка протокола решения проблемы  $act_{protng}^{ag}$ , т. е.  $\forall ag \in (AG^{ex} \cup \{ag^{dm}\}) \left( \{act_{glng}^{ag}, act_{ontng}^{ag}, act_{protng}^{ag}\} \subset ACT^{ag} \right)$ .

Действие агента из множества  $ACT^{ag}$  описывается выражением:

$$act^{ag} = \langle met_{act}^{ag}, it_{act}^{ag} \rangle,$$

где  $met_{act}^{ag}$  — метод решения проблемы;  $it_{act}^{ag}$  — интеллектуальная технология, в рамках которой реализован метод  $met_{act}^{ag}$ .

Таким образом, функция СГИМАС описывается выражением:

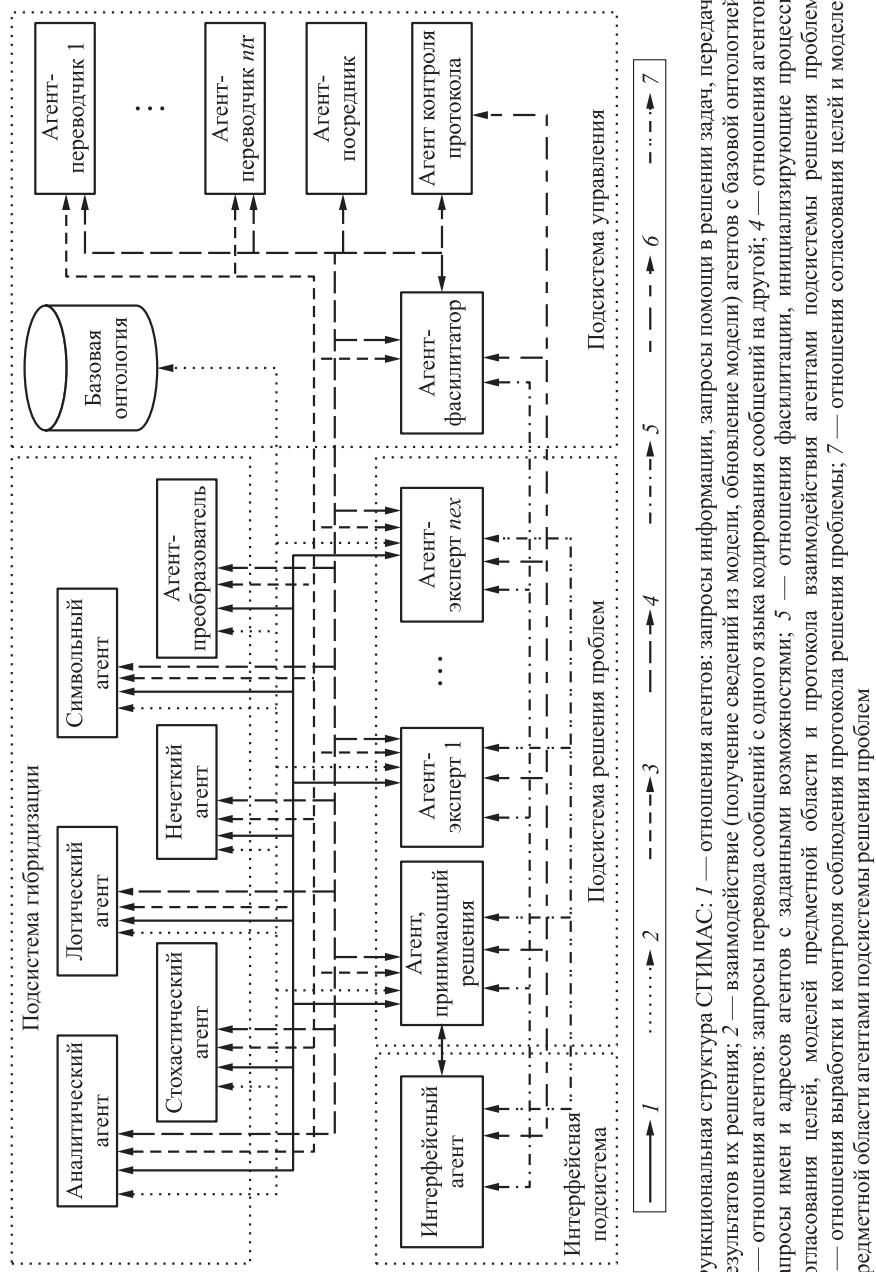
$$act_{chimas} = \left( \bigcup_{ag \in AG^*} ACT_{ag} \right) \cup act_{col}, \quad \left| \bigcup_{ag \in AG} \bigcup_{act \in ACT^{ag}} \right| \geq 2,$$

где  $act_{col}$  — коллективная функция СГИМАС, конструируемая агентами динамически в соответствии с выработанным протоколом решения проблемы; накладываемое ограничение требует, чтобы в составе СГИМАС использовалось не менее двух интеллектуальных технологий [3].

#### 4 Функциональная структура гибридной интеллектуальной многоагентной системы, моделирующей сплоченность агентов

Для компьютерного моделирования групповой сплоченности малого коллектива экспертов предлагается функциональная структура СГИМАС, представленная на рисунке.

Как показано на рисунке, в данной структуре СГИМАС выделяются четыре подсистемы: интерфейсная, гибридизации, решения проблем, управления. Интерфейсная подсистема состоит из единственного агента. Интерфейсный агент отвечает за взаимодействие с пользователем: запрашивает входные данные и выдает результат, а также визуализирует процессы, происходящие в СГИМАС, в частности динамику сплоченности агентов, выработки протокола решения проблемы, согласования целей и моделей предметной области агентами подсистемы решения проблем.



Функциональная структура СГИМАС. 1 — отношения агентов: запросы информации, запросы помощи в решении задач, передача результатов их решения; 2 — взаимодействие (получение сведений из модели, обновление модели) агентов с базовой онтологией; 3 — отношения агентов: запросы перевода сообщений с одного языка кодирования на другой; 4 — отношения агентов: запросы имен и адресов агентов с заданными возможностями; 5 — отношения фасилитации, инициализирующие процессы согласования целей, моделей предметной области и протокола взаимодействия агентами подсистемы решения проблем; 6 — отношения выработки и контроля соблюдения протокола решения проблемы; 7 — отношения согласования целей и моделей предметной области агентами подсистемы решения проблем

Подсистема управления содержит базовую онтологию, множество агентов-переводчиков, агента-посредника, агента-фасilitатора и агента контроля протокола. Базовая онтология — семантическая сеть, основа взаимодействия агентов, обеспечивающая понимание агентами смысла передаваемых сообщений по согласованию собственных моделей предметной области, целей, формированию протокола сплоченного взаимодействия. Агенты-переводчики предназначены для трансляции сообщений, передаваемых между агентами, с одного языка на другой. Необходимость таких агентов обусловлена тем, что агенты СГИМАС могут создаваться разными коллективами разработчиков и использовать разные языки передачи сообщений из множества LANG, при этом они способны общаться между собой напрямую, только если пересечение множеств языков, которыми они «владеют», не пусто. В противном случае им необходимо прибегать к услугам агентов-переводчиков. Агент-посредник выполняет служебную работу по отслеживанию имен, моделей и возможностей зарегистрированных агентов. Агент-фасilitатор отвечает за организацию эффективной коллективной работы агентов: идентифицирует этапы процесса решения проблемы, состав агентов-экспертов, согласованность их целей, моделей предметной области, сложившуюся ситуацию в СГИМАС, возникающие положительные и отрицательные групповые эффекты, воздействует на агентов-экспертов, чтобы активировать процессы переговоров по согласованию целей, моделей предметной области или корректировке протокола решения проблемы. Агент контроля протокола фиксирует изменения протокола решения проблемы, принятые агентами подсистемы решения проблемы, а также контролирует его соблюдение.

Подсистема гибридизации представлена аналитическим, логическим, символным, стохастическим и нечетким агентами, которые вместе с агентом-преобразователем реализуют гибридную составляющую СГИМАС, комбинируя разнородные знания, и предоставляют «услуги» агентам с использованием следующих моделей и алгоритмов: алгебраических уравнений для описания причинно-следственных связей концептов предметной области; метода Монте-Карло; продукционной экспертной системы с рассуждениями в прямом направлении; алгоритма нечеткого вывода Мамдани и др. Агенты данной подсистемы рефлекторные, т. е. не имеют развитой модели предметной области или целеполагания и выполняют поручения других агентов в соответствии с заложенными алгоритмами, в этом смысле они схожи с объектами в парадигме объектно-ориентированного программирования.

Агенты подсистемы решения проблем моделируют рассуждения коллектива экспертов под руководством лица, принимающего решения, «за круглым столом» по решению поставленной проблемы. Агент, принимающий решения, моделирует работу лица, принимающего решения: ставит задачи агентам-экспертам, собирает результаты работы, определяет, достигнут ли критерий останова, и либо принимает итоговое решение, либо запускает новую итерацию процесса решения проблемы. Агенты-эксперты моделируют рассуждения экспертов и с использованием собственных моделей предметной области и алгоритмов решения проблем

генерируют решения частей проблемы или альтернативные решения проблемы в целом в зависимости от задачи, поставленной им агентом, принимающим решения.

Агенты подсистемы решения проблем в общем случае могут создаваться разными коллективами разработчиков, а потому их модели предметной области и цели могут различаться и даже противоречить друг другу. В этой связи для эффективного решения поставленных проблем агенты подсистемы решения проблем согласуют свои модели предметной области, цели или вырабатывают протокол решения проблемы по запросу агента-фасилитатора. При этом они жестко ему не подчиняются, а «договариваются» между собой. Интерфейсный агент отслеживает эти процессы и отображает пользователю.

Таким образом, благодаря наличию механизмов согласования целей и моделей предметной области агентов подсистемы решения проблем, а также выработки согласованного протокола решения проблем в интеллектуальной системе обеспечивается сплоченное поведение агентов системы, позволяющее преодолевать разногласия и избегать конфликтов, вызванных различиями в моделях проблемы и целях ее решения. Агент-фасилитатор предохраняет систему от возможных негативных последствий чрезмерного сплочения агентов, таких как конформизм. В результате СГИМАС динамически перестраивает алгоритм своего функционирования, каждый раз при работе над проблемой вырабатывая релевантный ей гибридный интеллектуальный метод решения. Моделирование сплоченности обеспечивает развитие интеллектуальной системы и ее самоорганизацию в сильном смысле [11], т. е. возникающую за счет распределенного взаимодействия агентов без явного централизованного управления этим процессом одним из них, что релевантно малым коллективам экспертов, решающим проблемы «за круглым столом».

## 5 Заключение

Рассмотрены подходы к определению понятия сплоченность в социальной психологии и, в частности, в динамике групп, показана актуальность моделирования этого состояния коллектива и процессов, приводящих к нему, в системах, построенных на принципах синергетического искусственного интеллекта. Представлена модель СГИМАС, содержащая механизмы согласования целей и моделей предметной области, а также выработки протокола решения проблемы агентами системы без внешнего управления этим процессом. На основе данной модели разработана функциональная структура такой системы, иллюстрирующая состав, функционал и отношения, возникающие между агентами в процессе решения проблем. Моделирование сплоченности агентов обеспечивает развитие интеллектуальной системы и ее самоорганизацию в сильном смысле, в результате чего она динамически перестраивает алгоритм своего функционирования, каждый раз при работе над проблемой вырабатывая релевантный ей гибридный интеллектуальный метод решения.

## Литература

1. Нариньянин А. С. Инженерия знаний и НЕ-факторы: краткий обзор-08 // Вопросы искусственного интеллекта, 2008. № 1. С. 61–77.
2. Спиридовонов В. Ф. Психология мышления: Решение задач и проблем. — М.: Генезис, 2006. 319 с.
3. Колесников А. В., Кириков И. А., Листопад С. В. Гибридные интеллектуальные системы с самоорганизацией: координация, согласованность, спор. — М.: ИПИ РАН, 2014. 189 с.
4. Кириков И. А., Колесников А. В., Листопад С. В., Румовская С. Б. «Виртуальный консилиум» — инструментальная среда поддержки принятия сложных диагностических решений // Информатика и её применения, 2016. Т. 10. Вып. 3. С. 81–90.
5. Колесников А. В., Листопад С. В. Модель гибридной интеллектуальной многоагентной системы гетерогенного мышления для информационной подготовки оперативных решений в региональных электрических сетях // Системы и средства информатики, 2018. Т. 28. № 4. С. 31–41.
6. Жуков Ю. М., Журавлев А. В., Павлова Е. Н. Технологии командообразования. — М.: Аспект Пресс, 2008. 320 с.
7. Lewin K. Field theory in social science: Selected theoretical papers. — New York, NY, USA: Harper, 1976. 346 р.
8. Васильченко М. В., Литвиненко М. П., Мансурова И. С., Подопригора С. Я., Суроедова Е. А. Психология личности и группы. — Ростов-на-Дону: ДГТУ, 2011. 174 с.
9. Петровский А. В. Опыт построения социально-психологической концепции групповой активности // Вопросы психологии, 1973. № 5. С. 3–17.
10. Тарасов В. Б. От многоагентных систем к интеллектуальным организациям: философия, психология, информатика. — М.: Эдиториал УРСС, 2002. 352 с.
11. Serugendo G. D. M., Gleizes M.-P., Karageorgos A. Self-organization in multiagent systems // Knowl. Eng. Rev., 2005. Vol. 20. Iss. 2. P. 165–189.

Поступила в редакцию 01.03.20

---

## FUNCTIONAL STRUCTURE OF THE COHESIVE HYBRID INTELLIGENT MULTIAGENT SYSTEM

*S. V. Listopad*

Kaliningrad Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 5 Gostinaya Str., Kaliningrad 236000, Russian Federation

**Abstract:** The effectiveness of the hybrid and synergetic artificial intelligence systems is achieved, mainly, by the proper organization of the interaction of their elements and not by increasing the complexity or intelligence of the latter. In the

case of hybrid intelligent multiagent systems, simulating teams of experts solving problems “at round-table,” for the effective organization of interaction between agents, the conditions for the consistency of their goals and domain models, the unity of the protocol for solving the problem, and the compatibility of message transfer languages should be fulfilled. These conditions are especially relevant when a hybrid intelligent multiagent system is built from agents developed by various independent teams. In this case, agents should be able to independently coordinate goals, domain models and develop a protocol to solve the problem posed to them within the framework of a new class of intelligent systems, namely, a cohesive hybrid intelligent multiagent system. This article discusses the functional structure of such a system.

**Keywords:** cohesion; hybrid intelligent multiagent system; expert team

**DOI:** 10.14357/08696527200208

## Acknowledgments

The reported study was funded by the Russian Foundation for Basic Research, project No. 20-07-00104a.

## References

1. Narinyni, A. S. 2008. Inzheneriya znaniy i NE-faktory: kratkiy obzor-08 [Knowledge engineering and non-factors: A brief overview-08]. *Voprosy iskusstvennogo intellekta* [Artificial Intelligence Issues] 1:61–77.
2. Spiridonov, V. F. 2006. *Psichologiya myshleniya: Reshenie zadach i problem* [Psychology of thinking: Solving tasks and problems]. Moscow: Genezis. 319 p.
3. Kolesnikov, A. V., I. A. Kirikov, and S. V. Listopad. 2014. *Gibridnye intellektual'nye sistemy s samoorganizatsiyey: koordinatsiya, soglasovannost', spor* [Hybrid intelligent systems with self-organization: Coordination, consistency, dispute]. Moscow: IPI RAN. 189 p.
4. Kirikov, I. A., A. V. Kolesnikov, S. V. Listopad, and S. B. Rumovskaya. 2016. “Virtual'nyy konsilium” — instrumental'naya sreda podderzhki prinyatiya slozhnykh diagnosticheskikh resheniy [“Virtual council” — source environment supporting complex diagnostic decision making]. *Informatika i ee Primeneniya — Inform. Appl.* 10(3):81–90.
5. Kolesnikov, A. V., and S. V. Listopad. 2018. Model' gibridnoy intellektual'noy mnogaagentnoy sistemy heterogenного myshleniya dlya informatsionnoy podgotovki operativnykh resheniy v regional'nykh elektricheskikh setyakh [Model of a hybrid intelligent multiagent system of heterogeneous thinking for preparation of information about operational decisions in a regional power system]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(4):31–41.
6. Zhukov, Yu. M., A. V. Zhuravlev, and E. N. Pavlova. 2008. *Tekhnologii komandoobrazovaniya* [Team building technologies]. Moscow: Aspekt Press. 320 p.
7. Lewin, K. 1976. *Field theory in social science: Selected theoretical papers*. New York, NY: Harper. 346 p.

8. Vasil'chenko, M. V., M. P. Litvinenko, I. S. Mansurova, S. Ya. Podoprígora, and E. A. Suroedova. 2011. *Psichologiya lichnosti i gruppy* [Psychology of personality and group]. Rostov-on-Don: DGTU. 174 p.
9. Petrovskiy, A. V. 1973. Opyt postroeniya sotsial'no-psikhologicheskoy kontseptsii gruppovoy aktivnosti [The experience of building a socio-psychological concept of group activity]. *Voprosy psichologii* [Psychology Issues] 5:3–17.
10. Tarasov, V. B. 2002. *Ot mnogoagentnykh sistem k intellektualnym organizatsiyam: filosofiya, psichologiya, informatika* [From multiagent systems to intelligent organizations: Philosophy, psychology, and informatics]. Moscow: Editorial URSS. 352 p.
11. Serugendo, G. D. M., M.-P. Gleizes, and A. Karageorgos. 2005. Self-organization in multiagent systems. *Knowl. Eng. Rev.* 20(2):165–189.

Received March 1, 2020

## Contributor

**Listopad Sergey V.** (b. 1984)— Candidate of Science (PhD) in technology, senior scientist, Kaliningrad Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 5 Gostinaya Str., Kaliningrad 236000, Russian Federation; ser-list-post@yandex.ru

## ПРОБЛЕМА ЗАДАНИЯ КОЭФФИЦИЕНТОВ ГОТОВНОСТИ ДИАГНОСТИЧЕСКИХ СИСТЕМ В ПРОЕКТЕ ИТЭР: ПОСТАНОВКА ЗАДАЧИ И АЛГОРИТМ РЕШЕНИЯ

Г. М. Коновалов<sup>1</sup>

**Аннотация:** Изучается проблема задания коэффициентов готовности для диагностических систем (диагностик) в комплексе термоядерного экспериментального реактора ИТЭР. Такие системы должны обеспечивать требуемый уровень надежности, с которым измеряются многочисленные параметры комплекса. Возникает ряд сложностей на пути постановки задачи и ее решения. Это, в частности, уникальность объекта, многообразие и многочисленность параметров и диагностик, сложная взаимосвязь между ними, отсутствие аналогов в литературе по надежности. Описана математическая модель надежности измерений, которая учитывает как строгие ограничения, так и нечетко выраженные пожелания разработчиков. Разработан эффективный алгоритм расчета коэффициентов готовности диагностик. Приведен иллюстрационный пример расчета, выполненный с помощью специально созданной программной системы, в которую наряду с разработанным алгоритмом внедрена возможность волевого принятия решений.

**Ключевые слова:** ИТЭР; математические модели надежности; анализ надежности больших систем; коэффициент готовности

**DOI:** 10.14357/08696527200209

### 1 Введение

В настоящее время во Франции силами семи участников (ЕС, Япония, США, Россия, Китай, Индия, Южная Корея) под общим руководством международной организации ITER (International Thermonuclear Experimental Reactor) ведется сооружение термоядерного реактора ИТЭР [1, 2]. ИТЭР является ядерным объектом, и к нему применяются повышенные требования в плане надежности и безопасности.

Для обеспечения надежности ИТЭР был разработан комплекс мер, называемый RAMI-анализом [3, 4]. RAMI — аббревиатура от reliability, availability, maintainability, inspectability.

Одна из основных задач RAMI-анализа — определение коэффициентов готовности различных систем реактора ИТЭР, для которых проектной документацией определены минимальные значения. В частности, к важнейшим системам

<sup>1</sup>Частное учреждение Государственной корпорации по атомной энергии «Росатом» «Проектный центр ИТЭР», g.konovalov@iterrf.ru

ИТЭР относится диагностический комплекс, позволяющий осуществлять измерение и контроль параметров плазмы и токамака<sup>1</sup>, которые разделены на четыре группы в зависимости от их назначения.

Для диагностического комплекса проектная документация ИТЭР определяет только требования к минимальным значениям коэффициентов готовности, причем целиком для группы параметров. Тем не менее эти «глобальные» коэффициенты готовности не дают целевых значений для коэффициентов готовности отдельных диагностик. В настоящий момент при определении коэффициентов готовности диагностик разработчики предполагают, что глобальные требования распространяются на подсистему без учета того факта, что параметр может измеряться одновременно несколькими диагностиками и что вклады каждой диагностики в измерение параметра неравнозначны. Такой подход дает завышенные требования к коэффициентам готовности конечных подсистем, так как не учитывает имеющееся дублирование измерений.

Для разработчиков диагностик наибольший интерес представляет как раз коэффициент готовности для конечной подсистемы, так как его уменьшение позволяет упростить конструкцию и, как следствие, существенно сократить затраты на ее разработку и изготовление.

Предложенная в статье модель, а также алгоритм, разработанный на ее основе, позволяют решить задачу определения требований к коэффициентам готовности отдельных диагностик на основе заданных коэффициентов готовности групп измеряемых параметров и известной структурно-функциональной взаимосвязи между диагностиками и измеряемыми параметрами.

## 2 Математическая модель надежности измерения параметров

### 2.1 Параметры и диагностики

Исходными объектами модели служат два конечных множества:  $P = \{p\}$  — набор из  $M$  подлежащих измерению *параметров*,  $|P| = M$ , и  $D = \{d\}$  — совокупность из  $N$  *диагностик*, осуществляющих эти измерения,  $|D| = N$ .

Каждый параметр  $p \in P$  характеризуется:

- коэффициентом готовности  $b_p$ ;
- минимально допустимым значением коэффициента готовности  $b_p^*$ ;
- набором диагностик  $D_p \subset D$ , способных осуществлять измерение данного параметра.

Таким образом, необходимым требованием к надежности измерения параметров являются неравенства:

$$b_p \geq b_p^*, \quad p \in P. \quad (1)$$

---

<sup>1</sup>Токамак (тороидальная камера с магнитными катушками) — тороидальная установка для магнитного удержания плазмы с целью достижения условий, необходимых для протекания управляемого термоядерного синтеза. Реактор ИТЭР по своей конструкции относится к токамакам.

Если ввести обозначения для вектора коэффициентов готовности параметров  $b = (b_p, p \in P)$ , а также для вектора требований на коэффициенты готовности измерения параметров  $b^* = (b_p^*, p \in P)$ , то неравенство (1) можно переписать в виде:

$$b \geq b^*. \quad (2)$$

Каждый из параметров имеет вполне определенное *назначение* (*operation role*) в системе. Имеются четыре типа назначений с условными названиями Machine Protection (MP), Basic Control (BC), Control (AC) и Physics (Ph). Таким образом, множество параметров  $P$  разбивается на четыре непересекающихся подмножества:

$$P = P^{\text{MP}} + P^{\text{BC}} + P^{\text{AC}} + P^{\text{Ph}}.$$

Каждому типу назначения сопоставлено число, характеризующее общие требования по коэффициенту готовности для параметров, принадлежащих одному и тому же подмножеству:

$$b^{\text{MP}} = 0,985; \quad b^{\text{BC}} = 0,98; \quad b^{\text{AC}} = 0,91; \quad b^{\text{Ph}} = 0,9.$$

С помощью этих чисел определяются нижние границы для коэффициентов готовности параметров. Это делается следующим образом. Предполагается, что если группа параметров с одинаковым назначением находится в критическом состоянии (т. е. имеет минимально допустимый коэффициент готовности), то она образует в смысле надежности «последовательную схему» независимых одинаково надежных элементов. Из этого предположения следует, что

$$\begin{aligned} b_p^* &= \left( b^{\text{MP}} \right)^{1/|P^{\text{MP}}|}, \quad p \in P^{\text{MP}}; \quad b_p^* = \left( b^{\text{BC}} \right)^{1/|P^{\text{BC}}|}, \quad p \in P^{\text{BC}}; \\ b_p^* &= \left( b^{\text{AC}} \right)^{1/|P^{\text{AC}}|}, \quad p \in P^{\text{AC}}; \quad b_p^* = \left( b^{\text{Ph}} \right)^{1/|P^{\text{Ph}}|}, \quad p \in P^{\text{Ph}}. \end{aligned}$$

Каждая диагностика  $d \in D$  характеризуется:

- собственным коэффициентом готовности  $a_d$ ;
- ограничениями на минимально и максимально возможные значения собственного коэффициента готовности:

$$\underline{a}_d \leq a_d \leq \bar{a}_d; \quad (3)$$

- набором параметров  $P_d \in P$ , доступных для измерения с помощью данной диагностики.

Введем векторные обозначения

$$a = (a_d, d \in D); \quad \underline{a} = (\underline{a}_d, d \in D); \quad \bar{a} = (\bar{a}_d, d \in D).$$

Неравенство (3) можно записать также в виде:

$$\underline{a} \leq a \leq \bar{a}. \quad (4)$$

Предполагается, что любая диагностика  $d \in D$  функционирует как единое техническое устройство, которое, однако, не равноценно по отношению к измерению разных параметров из набора  $P_d$ . Более точно это предположение означает следующее:

- (1) измерение любого параметра  $P_d$  происходит независимо от остальных измерений;
- (2) отказ диагностики  $d$  означает ее неспособность к измерению любого параметра  $p \in P_d$ .
- (3) собственный коэффициент готовности  $a_d$  диагностики  $d$  по отношению к измеряемому параметру  $p \in P_d$  умножается на *понижающий фактор*, который обозначается  $k_{p,d}$ ,  $k_{p,d} \leq 1$ .

Таким образом, одна и та же диагностика  $d$  обладает, вообще говоря, разными коэффициентами готовности при измерении параметров из набора  $P_d$ . Эти коэффициенты готовности будем называть *фактическими*. Они равны

$$\tilde{a}_{p,d} = a_d k_{p,d}, \quad p \in P_d. \quad (5)$$

Выбор коэффициентов  $k_{p,d}$  является экспертной задачей. В качестве варианта предлагается однозначно увязать значение коэффициента  $k_{p,d}$  с *ролью* (*contribution*), которая устанавливается техническим заданием для диагностики  $d$  при измерении параметра  $p$ . В этом случае имеются всего три различных допустимых значения для понижающих факторов (по числу типов ролей — Primary, Backup, Supplementary). Однако сами значения определяются эксперты путем.

Поясним физический смысл ограничений (3). Верхняя граница для коэффициента готовности диагностики отражает инженерно-технические и экономические возможности в изготовлении «безотказных» устройств. Для нижней границы диагностики, проектирование и изготовление которой только предстоит, естественным значением представляется 0. Для диагностики, находящейся в стадии изготовления, нижней границей служит проектный уровень коэффициента готовности. Для полностью готовой диагностики левая и правая части в неравенстве (3) равны и совпадают с собственным коэффициентом готовности.

Резюмируя этот пункт, перечислим основные входные и выходные данные модели.

**Входные данные** модели:

- множества параметров  $P$  и диагностик  $D$ ;
- *структура* системы, заданная с помощью наборов  $\{D_p, p \in P\}$  и  $\{P_d, d \in D\}$ ;

- вектор требований к коэффициентам готовности измерения параметров  $b^*$ ;
- векторы ограничений на допустимые значения коэффициентов готовности диагностик  $\underline{a}$  и  $\bar{a}$ ;
- набор понижающих факторов  $K = \{k_{p,d} \leq 1, p \in P, d \in D\}$  (полагаем  $k_{p,d} = 0$ , если диагностика  $d$  не участвует в измерении параметра  $p$ ).

**Выходными данными** модели выступают векторы коэффициентов готовности:

- параметров,  $b = (b_1, \dots, b_M)$ ;
- диагностик,  $a = (a_1, \dots, a_N)$ ,

причем векторы  $a$  и  $b$  должны удовлетворять условиям (2) и (4).

## 2.2 Вычисление коэффициентов готовности параметров

При вычислении коэффициентов готовности параметров по заданной структуре и заданным коэффициентам готовности диагностик будем считать, что справедливо следующее предположение.

*Отказы всех диагностик независимы в совокупности.*

Это предположение означает, в частности, что процесс измерения какого-либо параметра с помощью  $n \geq 1$  диагностик может трактоваться как работа параллельной системы, состоящей из  $n$  независимых ненадежных элементов. Отсюда следует, что фактический коэффициент готовности  $b_p$  параметра  $p$  определяется совокупностью коэффициентов готовности  $A_p = \{\tilde{a}_{p,d}, d \in D_p\}$ , где  $\tilde{a}_{p,i}$  — определенные формулой (5) «фактические» коэффициенты готовности диагностик из множества  $D_p$  по отношению к параметру  $p$  (с учетом понижающих факторов).

Вычисление коэффициентов готовности  $b_p$  осуществляется следующим образом. Если  $|D_p| = 0$ , то полагаем  $b_p = 0$ . Пусть  $|D_p| > 0$ . Поскольку диагностики работают параллельно, то отказ диагностической системы в измерении параметра  $p$  наступает тогда и только тогда, когда отказали одновременно все диагностик из набора  $D_p$ . В силу независимости диагностик вероятность такого события равна  $\prod_{d \in D_p} (1 - \tilde{a}_{p,d})$ ; следовательно,

$$b_p = 1 - \prod_{d \in D_p} (1 - \tilde{a}_{p,d}) = 1 - \prod_{d \in D_p} (1 - a_d k_{p,d}). \quad (6)$$

Отметим, что при фиксированных значениях коэффициентов готовности диагностик  $a$  коэффициенты готовности параметров  $b = b(a)$  вычисляются однозначно.

### 3 Постановка задачи для определения коэффициентов готовности диагностик

В предыдущем разделе была formalизована зависимость, которая существует между надежностью измерения параметров и надежностью измеряющих диагностик. Однако постановка на этой основе математической задачи, решение которой позволило бы определить коэффициенты готовности диагностик, наталкивается на определенные трудности, которые связаны с отсутствием каких-либо априорных ограничений или критерииев, кроме неравенств (1) и (3).

Если ограничиться только этими двумя неравенствами, то задача должна быть сформулирована следующим образом:

*A. Определить вектор  $a$  так, чтобы выполнялись неравенства (1) и (3).*

Легко понять, что решение такой задачи тривиально. Достаточно рассчитать по формуле (6) коэффициенты готовности параметров, задав технологически максимальные коэффициенты готовности диагностик, т. е. положить

$$a = \bar{a},$$

и вычислить

$$\bar{b} = b(\bar{a}).$$

В этом случае неравенство (3) выполняется автоматически, а неравенство (1), которое имеет вид

$$\bar{b} \geq b^*, \quad (7)$$

нуждается в дополнительной проверке.

Заметим, что выполнение неравенства (7) является необходимым условием для решения задачи определения вектора  $a$  в любой постановке. В самом деле, если это неравенство не выполняется, то требования к коэффициентам готовности невозможно выполнить ни при каких параметрах диагностик. Последнее соображение вытекает из так называемого принципа монотонности [5, гл. 8], который (полагаем без обсуждения) имеет место для рассматриваемой системы: увеличение надежности какого-либо элемента системы не уменьшает надежность всей системы.

Указанное тривиальное решение задачи  $A$  не является единственно возможным, и оно также не является удовлетворительным по ряду причин, о которых ниже. Поэтому требуется рассмотреть дополнительные факторы, которые хотя бы косвенно отражали бы естественные инженерно-экономические ограничения. Основные неформальные интуитивные предпосылки при разработке алгоритма для определения коэффициентов готовности диагностик заключаются в следующем. Желательно найти такое решение, которое

- (а) обеспечивает выполнение неравенств (1) и (3);
- (б) является в каком-то смысле наиболее экономичным;

- (в) обеспечивает более или менее «равномерные» значения коэффициентов готовности для различных диагностик.

Необходимость требования (а) очевидна. Условия (б) и (в) нуждаются в пояснении. Дело в том, что, формулируя проблему, разработчики проекта ИТЭР не указали каких-либо дополнительных по сравнению с ограничениями (1) и (3) требований, которые ограничивали бы либо значения коэффициентов готовности диагностик, либо их техническую сложность, либо стоимость их разработки. Поэтому, кажется, единственный способ попытаться «сэкономить» — это принять предположение, что изготовить диагностику тем легче и тем дешевле, чем меньше ее коэффициент готовности. Кроме того, довольно естественным представляется желание избежать «перекосов», т. е. таких решений, при которых некоторые диагностики имеют предельно высокие значения коэффициента готовности, а другие, наоборот, допускают частые отказы.

Поскольку минимальное значение коэффициента готовности диагностики задано значением  $a_d$ , то естественно рассмотреть величину

$$\delta_d = a_d - \underline{a}_d, \quad (8)$$

которую назовем *избыточностью диагностики*  $d$ . Она характеризует превышение коэффициента готовности данной диагностики над минимально возможным (технологическим) уровнем.

Из сказанного следует, что желательно максимально уменьшить избыточность диагностик. В то же время ясно, что, вообще говоря, нельзя добиться того, чтобы избыточность всех диагностик равнялась нулю, поскольку в этом случае могут нарушаться ограничения (1) на коэффициенты готовности для параметров.

Обобщим понятие избыточности диагностики и назовем *запасом готовности диагностики*  $d$  максимальную величину  $\rho_d$ , на которую можно уменьшить коэффициент готовности этой диагностики, не нарушая ни одно из неравенств (1) и (3), при условии, что коэффициенты готовности остальных диагностик не меняются.

Очевидно, что уменьшение коэффициента готовности диагностики  $d$  может привести либо к нарушению левой части неравенства (3), либо к нарушению каких-либо из неравенств (1) для  $p \in P_d$ , т. е. для тех и только тех параметров, которые измеряются данной диагностикой. Отсюда получаем выражение запаса готовности диагностики  $d$  в виде минимума из двух величин:

$$\rho_d = \min \left\{ \delta_d; \min_{p \in P_d} s_{p,d} \right\}, \quad (9)$$

где  $\delta_d = a_d - \underline{a}_d$  — определенная формулой (8) избыточность диагностики  $d$ , а  $s_{p,d}$  — максимальная величина, на которую можно уменьшить коэффициент готовности  $a_d$  без нарушения неравенства (1) для параметра  $p$ .

Получим выражение для величины  $s_p$ . Согласно формуле (6)

$$b_p = 1 - (1 - a_d k_{p,d}) B_d, \quad (10)$$

где

$$B_d = \prod_{f \in D_p \setminus d} (1 - a_f k_{p,f}) .$$

Если уменьшить коэффициент готовности  $a_d$  на величину  $s < 0$ , то коэффициент готовности  $b_p$  уменьшится до значения

$$b_p(s) = 1 - (1 - (a_d - s) k_{p,d}) B_d .$$

Максимально возможное, без нарушения неравенства (1), уменьшение  $s = s_p$  получается очевидно, из соотношения:

$$b_p(s_{p,d}) = b_p^* = 1 - (1 - (a_d - s_{p,d}) k_{p,d}) B_d . \quad (11)$$

Из равенств (10) и (11) находим:

$$s_{p,d} = \left( \frac{1}{k_{p,d}} - 1 \right) \frac{b_p - b_p^*}{1 - b_p} . \quad (12)$$

Таким образом, соотношениями (9) и (12) запас готовности диагностики  $d$  выражается как

$$\rho_d = \rho_d(a) = \min \left\{ a_d - \underline{a}_d; \min_{p \in P_d} \left( \frac{1}{k_{p,d}} - 1 \right) \frac{b_p - b_p^*}{1 - b_p} \right\} .$$

Ясно, что в терминах запаса готовности можно эквивалентным образом записать условие (7), необходимое для разрешимости задачи о нахождении коэффициентов готовности диагностик, а именно:

$$\bar{b} \geq b^* \Leftrightarrow \bar{\rho} \geq 0 , \quad (13)$$

где  $\bar{\rho} = (\rho_d(\bar{a}), d \in D)$ .

Заметим также, что если условие (13) выполнено, то, в отличие от избыточности  $\delta_d$ , всегда можно добиться выполнения всех равенств

$$\rho_d = 0 , \quad d \in D . \quad (14)$$

В самом деле, неравенство  $\rho_d > 0$  по определению означает, что коэффициент готовности диагностики  $d$  можно уменьшить, уменьшив при этом значение  $\rho_d$ , не выходя за ограничения (1) и (3). В то же время если имеет место (14), то никакие дальнейшие изменения коэффициентов готовности диагностик в сторону уменьшения невозможны.

Высказанные соображения позволяют переформулировать постановку задачи:

*Б. Определить вектор  $a$  так, чтобы выполнялись соотношения (14).*

Самый простой способ решения задачи *Б* вытекает из определения запаса готовности. Достаточно последовательно уменьшать коэффициент готовности любой из диагностик  $d$ , у которой  $\rho_d > 0$ , на величину  $\rho_d$ , пересчитывая каждый раз значения запаса готовности. Процесс заканчивается, когда выполнены все соотношения (14).

Указанный алгоритм ориентирован на «минимизацию» коэффициентов готовности диагностик и в этом смысле отвечает высказанному выше требованию экономичности (б). Однако получаемое с его помощью решение плохо соответствует неформальному критерию (в), который заключается в стремлении минимизировать разброс между значениями коэффициентов готовности диагностик.

В следующем разделе излагается алгоритм, решающий задачу *Б*, который, используя понятие запаса готовности, обеспечивает более равномерный спектр коэффициентов готовности диагностик.

## 4 Алгоритм определения коэффициентов готовности диагностик

Введем в рассмотрение еще один числовой коэффициент, определяемый экспертыным путем — *значимость параметра  $p$* . Обозначим его через  $r_p$ . Будем считать, что значимость параметра однозначно определяется его назначением (см. разд. 2).

*Значимость диагностики* определим как

$$r_d = r_d(a) = \rho_d(a) \sum_{p \in P_d} r_p k_{p,d}, \quad d \in D,$$

где  $\rho_d(a)$  — запас готовности диагностики  $d$ , а  $k_{p,d}$  — понижающий фактор, который определяет фактический коэффициент готовности диагностики  $d$  при измерении параметра  $d$  (5). Отметим, что сумма в этой формуле не зависит от вектора  $a$ .

Пусть  $D^+ \subseteq D$  — подмножество всех диагностик с положительным запасом готовности:

$$D^+ = D^+(a) = \{d \in D : \rho_d(a) > 0\}.$$

Пусть также  $E \subseteq D^+$  — диагностики из множества  $D^+$ , которые имеют наибольшую значимость среди элементов  $D^+$ :

$$E = E(a) = \{d \in D^+ : r_d \geq r_e, e \in D^+\}.$$

Общая схема алгоритма расчета допустимых значений для коэффициентов готовности диагностик (без учета частных деталей) выглядит следующим образом.

**Шаг 0.** Присваиваем коэффициентам готовности диагностик наибольшие допустимые значения:

$$a_d := \bar{a}_d, \quad d \in D.$$

Если при этих значениях неравенства (1) выполняются, то переходим к следующему шагу. В противном случае задача определения допустимых значений для коэффициентов готовности диагностик не имеет решения.

**Шаг 1.** Определяем множество  $D^+$ . Если  $D^+ \neq \emptyset$ , то полагаем  $n = 1$  и переходим к шагу 2. В противном случае алгоритм заканчивает работу.

**Шаг 2.** Находим множество  $E$  и преобразовываем (уменьшаем) коэффициенты готовности диагностик из множества  $E$ :

$$a_d := \max \left\{ a_d - \frac{\alpha}{n}, \underline{a}_d - \varepsilon \right\}, \quad d \in E,$$

где  $\alpha > 0$  — фиксированное число;  $\varepsilon$  — точность вычислений ( $\alpha$  и  $\varepsilon$  — параметры алгоритма).

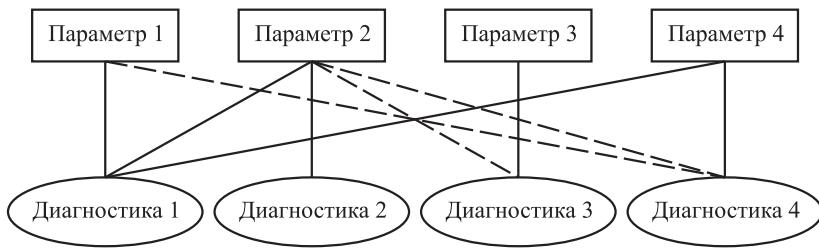
**Шаг 3.** Осуществляем проверку условий  $b_p \geq b_p^* - \varepsilon$  для тех параметров  $p$ , для которых коэффициент готовности мог измениться на шаге 2, т. е. для  $p \in \bigcup_{d \in E} P_d$ . Если все неравенства выполнены, то повторяем шаг 1. В противном случае возвращаем коэффициентам готовности  $a_d$  те значения, которые они имели до выполнения шага 2, увеличиваем  $n$ , полагая  $n := 2n$ , и повторяем шаг 2.

Алгоритм завершает работу за конечное число шагов в тот момент, когда выполнены все равенства (14) (с точностью до малого параметра  $\varepsilon$ ). На каждом шаге алгоритма уменьшаются коэффициенты готовности только тех диагностик, для которых уменьшение возможно и которые находятся в наиболее удаленных от минимальных значений положениях. Это минимизирует разброс итоговых значений коэффициентов готовности диагностик.

## 5 Программная система и пример расчета

Для применения на практике описанного выше алгоритма обнаружился ряд препятствий. Алгоритм не учитывает в явном виде такие сложно формализуемые характеристики диагностик, как стоимость, сложность монтажа, возможность оперативного ремонта и т. п. Это привело к созданию на основе алгоритма программной системы с развитым пользовательским интерфейсом.

Программа позволяет создавать модели, состоящие из произвольного числа измеряемых параметров и диагностик. Каждый элемент модели — параметр или диагностика — может быть сконфигурирован произвольным образом. Параметры можно разбивать на группы в зависимости от их назначения и для каждой группы задавать требуемый коэффициент готовности. Вклад каждой диагностики в измерение параметров задается пользователем произвольно. Для моделирования



Модель диагностической системы: сплошные линии — измерения с понижающим фактором, равным 1; штриховые — измерения с понижающим фактором, равным 0,3

ситуации, когда диагностика считается уже разработанной, программа предоставляет возможность зафиксировать коэффициент готовности на произвольном уровне. Структурно-функциональная связь между измеряемыми параметрами и диагностиками также определяется пользователем.

В качестве небольшой иллюстрации работы алгоритма рассмотрим предельно упрощенную модель, состоящую из четырех параметров и четырех диагностик. Функциональная связь между параметрами и диагностиками показана на рисунке. Минимальный коэффициент готовности для параметров задан равным 99,6%.

Очевидно, что коэффициент готовности у *Диагностики 3* не может быть ниже 99,6%, так как она единственная обеспечивает измерение *Параметра 3*. Достаточно высокие требования к коэффициенту готовности будут также и для *Диагностики 1*: она вносит основной вклад в измерение *Параметра 1*. Измерения *Параметров 2* и *4* дублируются *Диагностиками 1, 2* и *3* и *Диагностиками 1* и *4* соответственно. Это позволяет предположить, что требования к минимальным коэффициентам готовности для *Диагностик 2* и *4* будут существенно ниже, чем для *Диагностик 1* и *3*.

В таблице представлены результаты расчета, выполненные с помощью разработанной программы.

Значения минимальных коэффициентов готовности для *Диагностик 2* и *4* определяются минимальным ограничением, установленным в программе. Как видно из таблицы, результаты не противоречат очевидным соображениям, приведенным выше.

Результаты расчета минимальных коэффициентов готовности

№ диагностики	Минимальный коэффициент готовности, %
1	99,56
2	50
3	99,61
4	50

Эффективность алгоритма была проверена и подтверждена результатами, которые были получены при моделировании и расчете реального диагностического комплекса ИТЭР, состоящего из 78 диагностик, призванных измерять 98 параметров. Изложение и обсуждение этих результатов, однако, выходит за рамки данной статьи.

## **6 Заключение**

Изложена своеобразная проблематика определения коэффициентов готовности диагностических систем, возникшая в процессе разработки и строительства термоядерного экспериментального реактора ИТЭР. Обсуждены инженерные предпосылки проблемы, построена математическая модель, предложена концепция постановки задачи, а также алгоритм решения. Создана программная система, в которой реализованы математические и алгоритмические разработки, а также интерфейсные возможности для самостоятельных действий оператора. Совокупный результат исследования может быть охарактеризован как новая информационная технология в области RAMI-анализа — специфического комплекса обеспечения надежности ИТЭР. На данном этапе она стала рабочим инструментом для задания коэффициентов готовности диагностик. Однако в дальнейшем планируется продолжение работ по модернизации алгоритма и программной системы с целью динамического определения ограничений на коэффициенты готовности каждого параметра в зависимости от ограничений на коэффициент готовности группы параметров и от набора диагностик, измеряющих этот параметр.

## **Литература**

1. International Thermonuclear Experimental Reactor. <http://www.iter.org/proj/iterhistory>.
2. Holtkamp J. An overview of the ITER project // Fusion Eng. Des., 2007. doi: 10.1016/j.fusengdes.2007.03.029. Vol. 82. P. 427–434.
3. Van Houtte D., Okayama K., Sagot F. RAMI approach for ITER // Fusion Eng. Des., 2010. Vol. 85. P. 1220–1224. doi: 10.1016/j.fusengdes.2010.03.007.
4. Коновалов Г. М. Методика расчета надежности международного термоядерного экспериментального реактора и оптимизация действий по снижению рисков // Системы и средства информатики, 2016. Т. 26. № 1. С. 166–181. doi: 10.14357/08696527160111.
5. Байхельт Ф., Franken П. Надежность и техническое обслуживание. Математический подход / Пер. с нем. — М.: Радио и связь, 1988. 392 с. (Beichelt F., Franken P. Zuverlässigkeit und Instandhaltung. Mathematische Methoden. — Berlin: VEB Verlag Technik, 1983. 315 p.)

*Поступила в редакцию 01.10.19*

# THE PROBLEM OF SETTING AVAILABILITY TARGETS FOR ITER DIAGNOSTICS: PROBLEM STATEMENT AND SOLUTION ALGORITHM

**G. M. Konovalov**

Institution “Project Center ITER” of the Russian Federation National Nuclear Corporation “ROSATOM,” 1-3 Kurchatova Sq., Moscow 123182, Russian Federation

**Abstract:** The problem of setting availability targets for diagnostic systems (diagnostics) in the ITER project is studied. Such systems must provide the required level of availability, with which numerous parameters of ITER are measured. There are a number of difficulties on the way of problem statement and its solution. These are, in particular, the uniqueness of the object, the diversity and plurality of parameters and diagnostics, the complex interrelation between them, and the absence of analogues in the literature on reliability. The article describes a mathematical model of measurement reliability which takes into account both strict limitations and vaguely expressed wishes of developers. An effective algorithm for calculating diagnostic availability targets is developed. The illustration example of calculation was prepared with the help of a specially created software which uses along with the developed algorithm the possibility of volitional decision-making.

**Keywords:** ITER; mathematical models of reliability; large system reliability analysis; availability

**DOI:** 10.14357/08696527200209

## References

1. International Thermonuclear Experimental Reactor. Available at: <http://www.iter.org/proj/iterhistory> (accessed April 6, 2020).
2. Holtkamp, J. 2007. An overview of the ITER project. *Fusion Eng. Des.* 82:427–434. doi: 10.1016/j.fusengdes.2007.03.029.
3. Van Houtte, D., K. Okayama, and F. Sagot. 2010. RAMI approach for ITER. *Fusion Eng. Des.* 85:1220–1224. doi: 10.1016/j.fusengdes.2010.03.007.
4. Konovalov, G. M. 2016. Metodika rascheta nadezhnosti mezhdunarodnogo termoyader-nogo eksperimental'nogo reaktora i optimizatsiya deystviy po snizheniyu riskov [Methodology of reliability calculation of international thermonuclear experimental reactor and optimization of risk mitigation actions]. *Systemy i Sredstva Informatiki — Systems and Means of Informatics* 26(1):166–181. doi: 10.14357/08696527160111.
5. Beichelt, F., and P. Franken. 1983. *Zuverlässigkeit und Instandhaltung. Mathematische Methoden*. Berlin: VEB Verlag Technik. 315 p.

*Received October 1, 2019*

## **Contributor**

**Konovalov Grigory M.** (b. 1980) — Candidate of Science (PhD) in physics and mathematics, Deputy Project Manager for Quality, Institution “Project Center ITER” of the Russian Federation National Nuclear Corporation “ROSATOM,” 1-3 Kurchatova Sq., Moscow 123182, Russian Federation; g.konovalov@iterrf.ru

## НЕКОТОРЫЕ ВОПРОСЫ ПРОГРАММНОГО ОПРЕДЕЛЕНИЯ ЦЕНТРОВ ОБРАБОТКИ ДАННЫХ

*В. Б. Егоров<sup>1</sup>*

**Аннотация:** Очередным этапом развития центров обработки данных (ЦОД) обещает стать их программное определение. Появление публикаций по программно определяемым ЦОД показывает интерес к этой теме, однако их содержание выявляет существенные расхождения в оценке предмета и трудности в понимании сущности дела. Для того чтобы ЦОД стал программно определяемым, ему недостаточно программно определяемых компонентов: вычислительной платформы, сети и хранилищ. Также принципиально необходима централизация управления ЦОД и его автоматизация на основании установленных правил специальным программным обеспечением (ПО), так называемым «оркестровщиком». Стремление повысить эффективность ЦОД и снизить расходы на обслуживание стимулирует их владельцев двигаться в сторону программного определения. Однако для этого в настоящее время на рынке нет готовых универсальных решений. Поэтому примеры программно определяемых ЦОД демонстрируют лишь немногие крупные компании, обладающие современной инфраструктурой и возможностями ее «оркестровки». Прочие владельцы ЦОД вынуждены либо отказываться от их программного определения, либо искать палиативные решения в виде частично программно определяемых («двухскоростных») ЦОД.

**Ключевые слова:** автоматизация управления; датацентр; оркестровщик; программно определяемый центр обработки данных

**DOI:** 10.14357/08696527200210

### 1 Введение

Бурное развитие в последнее десятилетие мобильной связи породило такой феномен информационных технологий (ИТ), как «облачные» вычисления — принципиально новую модель предоставления и потребления ИТ-сервисов. Но «облако» — это удобная абстракция и привлекательное рекламно-маркетинговое представление для пользователей, в первую очередь — многочисленных рядовых обладателей разнообразных гаджетов. Для провайдеров же «облачных» услуг это сложная многослойная и многокомпонентная инфраструктура аппаратных и программных средств, определяющие позиции в которой занимают ЦОД.

За более чем полвека своего существования ЦОД прошли долгий путь развития от отдельных изолированных вычислительных центров с громоздкими мейнфреймами до современных компактных конвергентных систем, включающих

<sup>1</sup> Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, VEgorov@ipiran.ru

множество плотно упакованных многопроцессорных серверов и дисковых хранилищ информации, а также интеллектуальное коммуникационное оборудование, объединяющее это множество локальными сетями и «вплетающее» его в региональные и глобальные сети. Современные ЦОД в большинстве своем виртуализированы, причем виртуализации подверглись все их основные функциональные компоненты. Имеющая уже более чем полувековую историю виртуализация вычислительных ресурсов развивается параллельно на двух конкурирующих и в то же время дополняющих друг друга уровнях: виртуальных машин (virtual machines) и контейнеров (containers). Сетевая виртуализация также имеет достаточно долгую историю и тоже проявляется в двух формах: собственно виртуальных сетей (VLAN — virtual local area network, VXLAN — virtual extensible local area network, VPN — virtual private network) и виртуализации сетевых функций (network functions virtualization, NFV). В области хранения информации виртуализация тоже прошла впечатляющий путь развития от простых массивов RAID до практически безразмерных и безотказных, повсеместно доступных «облачных» хранилищ.

Следующим за виртуализацией важным этапом развития ЦОД стало программное определение основных его компонентов. Первыми программно определились сети, и понятие программно определяемых сетей (software-defined networking, SDN) прочно закрепилось в ИТ-лексиконе. За SDN последовали программно определяемые хранилища (software-defined storage, SDS). Логика процесса естественно поставила на повестку дня программное определение ЦОД в целом.

Несмотря на историческую естественность возникновения понятия программно определяемого ЦОД (software-defined data-center, SDDC), оно пока не общепризнано. Сам акроним SDDC был введен в обращение компанией VMware, и в настоящее время только эта компания предлагает продукты, явно содержащие его в названии, хотя конкуренты тоже имеют предложения, решающие схожие задачи. Тем не менее понятие SDDC постепенно приживается в сфере ИТ, и ему уже посвящено немало публикаций. Эти публикации, будучи индикатором растущего интереса к SDDC, не только выявляют наличие ряда связанных с ними проблем, но и, к сожалению, нередко вводят читателя в заблуждение, демонстрируя непонимание их авторами сущности предмета.

Далее сделана попытка выявления некоторых актуальных проблем программного определения ЦОД с анализом ряда сомнительных суждений об этом непростом предмете.

## **2 Спорные вопросы программно определяемого центра обработки данных**

В отношении возможности создания SDDC до сих пор высказываются разные, порой полярные мнения: от утверждения, что SDDC — это реальность

сегодняшнего дня, до отрицания самой возможности программного определения ЦОД. Первое из них подтверждается, по крайней мере на уровне деклараций и рекламы, предложением программных продуктов для SDDC и «облаков» от компаний VMware и Citrix. В противовес им разработчики инфраструктуры «облака» OpenStack в компании Cloudscaling полагают SDDC «букетом нелепиц», а все происходящее вокруг него «программно определяемым хайпом». По их мнению, с которым трудно не согласиться, создание SDDC могут себе позволить компании Yahoo или Google с их полностью гомогенными ИТ-системами, но не абсолютное большинство рядовых бизнес-пользователей, которых в прошлом заверяли, что все их головные боли устранит конвергентная инфраструктура, и которым теперь в качестве лекарства от тех же головных болей предлагают SDDC. Однако на деле это не более чем реклама поставщиков с целью продать уже имеющиеся у них ИТ-продукты [1]. Действительно, обычному потребителю ИТ-сервисов даже трудно уловить смысл термина «программное определение», и его появление действительно можно рассматривать как чисто маркетинговый ход продавцов ИТ-продуктов. Сегодня непросто сыскать не только реальные примеры действующих SDDC, но и, похоже, общепринятое понимание понятия программного определения [2]. Отсюда и нередко встречающиеся в литературе сомнительные, а то и откровенно ложные утверждения относительно SDDC. Вот лишь некоторые характерные примеры.

**Пример первый:** «Двигателем развития SDDC и облачных вычислений является постоянное удешевление аппаратуры серверов и памяти. Чем они становятся дешевле, тем более выгодным становится создание больших ЦОД и продажа их ресурсов в качестве инфраструктуры как сервиса IaaS (infrastructure-as-a-service)» [3]. Но на самом деле по элементарной логике все должно происходить наоборот и в плане создания ЦОД, и с точки зрения продажи ресурсов. Действительно, чем дешевле серверы и дисковые хранилища, тем доступнее они массовому потребителю, и такая доступность должна стимулировать у этого массового потребителя создание своих собственных ЦОД. С другой стороны, любой товар продается, если на него находится покупатель. Поэтому «двигателем развития SDDC и облачных вычислений» становится не удешевление аппаратуры, а скорее бурное развитие мобильной связи и, как следствие, растущие запросы пользователей многочисленных гаджетов на получаемые из «облаков» современные ИТ-услуги, включая IaaS. Как известно из основ экономики, именно спрос, в данном случае массовый спрос, на «облачные» сервисы порождает предложение — создание предоставляющих эти сервисы «облачных» ЦОД. А массовому распространению «персональных» ЦОД препятствует не стоимость оборудования, а дороговизна ПО и, главное, сложности развертывания, содержания и менеджмента даже небольших ЦОД, подразумевающие в числе прочего немалый штат высококвалифицированных специалистов, которые обходятся во много раз дороже обслуживаемой ими аппаратуры. Поэтому все большее число организаций и предприятий с радостью перепоручает такого рода заботы

«облакам». Но от такого перепоручения проблемы развертывания, содержания и менеджмента ЦОД не исчезают, а лишь переносятся в «облачные» выси, там аккумулируются и множатся по мере усложнения «облачных» ЦОД из-за роста объемов предоставляемых ими сервисов, порождая объективную необходимость в автоматизации управления ими.

**Пример второй:** «Программно определяемый ЦОД обеспечивает централизацию управления с помощью облачных инструментов и средств виртуализации» [4]. На самом деле и тут все обстоит ровным счетом наоборот: не программное определение обеспечивает централизацию управления, а именно централизация управления инфраструктурой ЦОД создает условия для последующей автоматизации управления и менеджмента как основы программного определения. Для превращения обычного ЦОД в SDDC необходима в первую очередь централизация управления ЦОД и лишь потом его программное определение, т. е. автоматизация централизованного управления в соответствии с наперед заданными «политиками» (policies) посредством специального ПО, часто называемого «оркестровщиком» (orchestrator). Первична здесь именно централизация управления, без нее невозможно никакое программное определение [2, 5] и не спасут никакие «облачные инструменты и средства виртуализации».

**Пример третий:** «Централизация аппаратуры в ЦОД и управления в SDDC дает ряд дополнительных возможностей, таких как улучшение безопасности или экономия суммарной потребляемой мощности» [3]. Вероятно, программное определение ЦОД действительно позволит экономить энергию. Однако в отношении безопасности данное утверждение не столь однозначно. С точки зрения операционной безопасности и доступности сервисов централизация управления позволяет, с одной стороны, отслеживать состояние всей системы в целом и гибко ее реконфигурировать в случае локальных сбоев или отказов. Но, с другой стороны, отказ в самом централизованном управлении, не важно аппаратной его части или ПО, может повлечь катастрофические последствия для всей системы. Аналогичная ситуация и с информационной безопасностью. Централизованное управление позволяет оперативно развертывать непосредственно в критических местах ЦОД средства защиты (брандмауэры, средства глубокой инспекции пакетов и пр.) для локализации последствий своевременно выявленных вредоносных атак (вирусы, DDOS (distributed denial of service) и др.). Но если так или иначе пораженным оказывается, например, «оркестровщик», то губительные последствия могут грозить не только самому ЦОД, но и его многочисленным клиентам.

Также к массовым заблуждениям можно отнести часто повторяемое утверждение об обязательности виртуализации всех компонентов SDDC. На самом деле программное определение ЦОД может быть с большим или меньшим успехом реализовано и в невиртуализированных ЦОД. Другое дело, что для таковых вряд ли будет иметь смысл тратить усилия и немалые средства на программное определение устаревающей инфраструктуры с неадекватным набором сервисов.

### 3 Трудности на пути создания программно определяемого центра обработки данных

В посвященных SDDC публикациях оказывается непростым делом вычленить его адекватное определение. Одно из самых удачных развернутых определений SDDC дает Webopedia: «SDDC — это ЦОД, в котором вся инфраструктура виртуализирована и предоставляется пользователям как сервис IaaS. Управление таким ЦОД, включая конфигурацию его аппаратуры, полностью автоматизировано и реализуется интеллигентным ПО. Часто SDDC рассматриваются как очередной шаг в развитии виртуализации и «облачных» вычислений, обеспечивающий возможность поддержки как старых промышленных приложений, так и новых «облачных» сервисов. SDDC характеризуется виртуализацией и программным определением трех базовых компонентов: серверов, памяти и сети. Также в SDDC предполагается слой управления и менеджмента для удовлетворения требований приложений, соблюдения соглашений об уровне предоставления услуг (service level agreements), следования принятым правилам поведения и учета ограничений по стоимости обслуживания» [6].

Таким образом, не виртуализация и даже не программное определение отдельных компонентов ЦОД, а централизация и автоматизация управления ЦОД как единым целым в конечном счете порождает SDDC. Тогда в очищенном от рекламных заявлений и ошибочных суждений виде SDDC предстает как ЦОД, предпочтительно полностью виртуализированный, в котором менеджмент и управление предоставлены сервисов строго централизовано, насквозь автоматизировано и осуществляется на основе заранее установленных «политик» с помощью специального ПО, часто называемого «оркестровщиком».

Найти живые примеры функционирующих SDDC столь же непросто, как и хорошее их определение. Сегодня если не эталонные SDDC, то нечто приближающееся к ним являются миру в основном интернет-гиганты, такие как Google, Amazon или Facebook. Стать первопроходцами на пути программного определения своих ЦОД их вынуждает быстро растущий объем запросов услуг и ресурсов от многочисленных пользователей по всему миру. При этом следует отметить, что у крупных компаний диапазон предоставляемых сервисов может быть относительно ограниченным. Между тем организации с меньшими возможностями порой вынуждены обслуживать более широкий круг запросов на устаревшем оборудовании, которое они не в состоянии одномоментно обновить. Кроме того, лишь крупные компании обладают ресурсами для разработки собственных «оркестровщиков», в то время как прочие вынуждены довольствоваться либо дорогими готовыми, но, как правило, мало пригодными в каждом конкретном случае коммерческими продуктами, либо более дешевыми, но требующими серьезной высокопрофессиональной доработки по месту полуфабрикатами с открытыми кодами. В итоге на деле построение SDDC оказывается сегодня недоступным абсолютному большинству владельцев ЦОД. Поэтому, вероятно, в обозримом будущем они будут отказываться от преобразования собственных

ЦОД в SDDC, отдавая предпочтение аренде сервисов, в том числе «облачных».

На сегодняшний день наиболее широко обсуждаемым (вследствие чего, надо полагать, и наиболее спорным) компонентом SDDC остается SDN. Действительно, здесь пересекаются интересы многих поставщиков оборудования, ПО и услуг. Хотя в наибольшей степени SDN могут проявить свои достоинства на специальной, соответствующей их особенностям инфраструктуре из унифицированных централизованно управляемых коммутаторов, принципиально создание таких сетей возможно на любом оборудовании, в том числе имеющемся и даже устаревшем. Однако, несмотря на широкое обсуждение, отсутствие видимых ограничений и заманчивые перспективы, реальные примеры SDN пока ограничились ЦОД нескольких крупных телекоммуникационных компаний, все тех же Google, Amazon или Facebook. По-видимому, это является следствием как концептуальных ограничений применимости сетей SDN сугубо в рамках ЦОД [5], так и экономических факторов, связанных с необходимостью для реализации и поддержки SDN-проектов содержать команду высококвалифицированных программистов, к тому же дефицитных вследствие отсутствия мирового опыта проектирования и построения сетей SDN [7].

В еще большей степени потребность унификации справедлива для SDS. Если в SDN разномастность сетевого оборудования в значительной степени нивелируется унифицированным протоколом взаимодействия с ним из центра управления (контроллера сети), в частности протоколом OpenFlow, то SDS лишены такого преимущества из-за исторически слишком большого разнообразия форм организации физических хранилищ. В этом плане выглядит весьма обещающим быстрое развитие конвергентных, а в последнее время и гиперконвергентных инфраструктур. Именно они, отвечая объективному запросу на унификацию оборудования, в том числе систем хранения данных, дают реальный шанс на более активное внедрение SDS и, в более отдаленной перспективе, рост числа действующих SDDC.

Хотя SDN и SDS худо-бедно внедряются в ЦОД, самым главным препятствием для увеличения числа реально действующих SDDC остается сложность и дороговизна создания их «оркестровщиков». Так, по некоторым оценкам, в ближайшие годы в сфере программного определения наибольшие траты ожидаются, в порядке уменьшения их доли, на SDS, SDN, «облачный» менеджмент и «оркестровку» [8]. «Оркестровка» здесь стоит на последнем месте, при том что это главный определяющий компонент SDDC. А поскольку это еще и самый дорогостоящий элемент приведенного выше перечня, не приходится ожидать резкого увеличения числа потенциальных вкладчиков в «оркестровку» своих ЦОД. Таким образом, приведенные оценки в целом демонстрируют движение к программному определению, но их порядок — скорее к частному, касающемуся хранилищ и сетей, чем к глобальному в масштабах целого ЦОД.

Для владельцев ЦОД, которым по разным причинам недоступно их одновременное преображение в SDDC и коих абсолютное большинство, компания

«Инфосистемы Джет» рекомендует двигаться пошагово, постепенно внедряя отдельные решения по программному определению и встраивая их в общую концепцию будущих SDDC [9]. Хотя такая рекомендация весьма смахивает на рекламу собственных фирменных услуг во внедрении такого рода частных решений, в ней может быть свой резон. Паллиативное решение частично программно определяемых или «двуухскоростных» ЦОД также допускают для организаций, не способных мгновенно обновить устаревшее оборудование, и некоторые теоретики SDDC [8].

Отдельная важная, требующая специального изучения проблема создания SDDC — обеспечение их безопасности в условиях централизации управления и препоручения решения всех оперативных задач программному «оркестровщику». Так, по утверждению экспертов компании Gartner, SDDC «требуют стратегию защиты, отличную от той, которая применялась на устройствах конечных пользователей» [10]. Представляется весьма вероятным, что для эффективной защиты SDDC без ухудшения его характеристик потребуется новое поколение решений по безопасности.

## 4 Заключение

Очередным этапом развития ЦОД обещает стать их программное определение. Дискуссии на эту тему показывают растущий интерес к SDDC, но также выявляют существенные расхождения в оценке предмета и демонстрируют трудности в понимании существа вопроса.

Понятие SDDC по определению подразумевает программно определяемыми все основные составляющие ЦОД: вычислительную платформу, хранилища информации и сеть. Но этого недостаточно. Для того чтобы ЦОД превратился в SDDC, ему принципиально необходимы, во-первых, централизация управления инфраструктурой и предоставлением сервисов и, во-вторых, автоматизация этого управления программным «оркестровщиком» на основании заранее установленных правил поведения. Таким образом, «в теленомическом пределе концепция SDDC должна дополнить программно определяемые системы специальным слоем ПО для «оркестровки» ЦОД, т. е. автоматизации его менеджмента и администрации» [11].

С одной стороны, в стремлении минимизировать расходы и сложности, связанные с получением ИТ-услуг, небольшие, средние и даже крупные организации отказываются от создания собственных ЦОД в пользу аренды сервисов, в том числе «облачных». С другой стороны, то же самое стремление к уменьшению расходов и забот с предоставлением этих сервисов заставляет владельцев ЦОД двигаться в сторону SDDC. Однако в настоящее время на рынке нет готовых решений SDDC, вряд ли их следует ожидать и в ближайшей перспективе. Предлагаемые варианты представляют собой некие «полуфабрикаты», для превращения которых в готовый продукт, развертывания и дальнейшей поддержки в каждом конкретном случае нужны дорогостоящие профессиональные услуги.

Поэтому сегодня живые примеры SDDC демонстрируют только крупные интернет-компании, обладающие ЦОД с современной инфраструктурой и средствами для содержания высококвалифицированного обслуживающего ЦОД персонала, способного осуществить его «оркестровку».

Прочие владельцы ЦОД, которым недоступно их одномоментное преобразование в SDDC, могут выбрать постепенное, поэтапное внедрение отдельных имеющихся на рынке готовых решений по частичному программному определению, прибегая к помощи предоставляющих такого рода услуги посредников, и довольствоваться паллиативными частично программно определяемыми ЦОД («двуухскоростными» SDDC).

## **Литература**

1. Kovar J. F. Software-defined data centers: Should you jump on the bandwagon? // CRN, 2013. <https://www.crn.com/news/data-center/240154576/software-defined-data-centers-should-you-jump-on-the-bandwagon.htm>.
2. Егоров В. Б. Практическое определение «программно определяемого» // Системы и средства информатики, 2019. Т. 29. № 2. С. 85–94.
3. What's a software defined data center? // SDxCentral, 2015. <https://www.sdxcentral.com/sdn/definitions/software-defined-data-center>.
4. 5 фактов о программно-определяемых ЦОД // Cloud4Y, 2016. <https://www.cloud4y.ru/about/news/5-faktov-o-programmno-opredelyaemykh-tsod-/>.
5. Егоров В. Б. Некоторые вопросы практической реализации концепции SDN // Системы и средства информатики, 2016. Т. 26. № 1. С. 109–120.
6. Beal V. SDDC — software-defined data center // Webopedia. [https://www.webopedia.com/TERM/S/software\\_defined\\_data\\_center\\_SDDC.html](https://www.webopedia.com/TERM/S/software_defined_data_center_SDDC.html).
7. Притула П. Программно-определяемый ЦОД: маркетологи передали эстафету практикам // CNews, 2016. [http://www.cnews.ru/articles/2016-12-22-programmnoopredelyaemyj\\_tsod\\_marketologi\\_peredali\\_estafetu\\_praktikam](http://www.cnews.ru/articles/2016-12-22-programmnoopredelyaemyj_tsod_marketologi_peredali_estafetu_praktikam).
8. Robinson D. SDDC is the answer — but how do we get there? // Data Center Dynamics, 2018. <https://www.datacenterdynamics.com/analysis/sddc-is-the-answer-but-how-do-we-get-there/>.
9. Программно-определяемый ЦОД сегодня и завтра // Инфосистемы Джет, 2015. <https://jet.su/about/news/18464>.
10. Securing the software-defined datacenter // Bitdefender, 1997. <https://www.bitdefender.com/business/datacenter-security-solutions/sddc.html>.
11. Sanders J. Software defined data centers: A cheat sheet // TechRepublic, 2017. <https://www.techrepublic.com/article/software-defined-data-centers-the-smart-persons-guide>.

*Поступила в редакцию 28.11.19*

## SOME ISSUES OF SOFTWARE-DEFINED DATA CENTERS

*V. B. Egorov*

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** A next stage in the data centers perfection process develops as software-defined data centers (SDDC). First publications on the SDDC show increasing interest to this topic revealing at the same time considerable discrepancies in estimations of the subject and difficulties with understanding the essence of the matter. To turn a data center into an SDDC, availability of both the software-defined networking and the software-defined storage is insufficient. It is also conceptually needed that the data center control would be definitely centralized and automated by means of special software (a so-called “orchestrator”) on the policy-driven basis. Aspiration for expenditures reduction and troubles abatement would impel data center owners to move toward the SDDC. At present, however, real SDDC examples are demonstrated only by few big internet companies possessing the up-to-date infrastructure and resources to create an adequate orchestrator, since no universal commercial orchestration products are available on the market. Nevertheless, other organizations may possibly choose a compromise palliative solution of partially software-defined or “double-speed” data centers.

**Keywords:** control automation; infrastructure as a service (IaaS); orchestrator; policy-driven basis; SDDC; SDN; SDS; software-defined

**DOI:** 10.14357/08696527200210

## References

1. Kovar, J. F. 2013. Software-defined data centers: Should you jump on the bandwagon? *CRN*. Available at: <https://www.crn.com/news/data-center/240154576/software-defined-data-centers-should-you-jump-on-the-bandwagon.htm> (accessed November 28, 2019).
2. Egorov, V. B. 2019. Praktichnoe opredelenie “programmno opredelyaemogo” [A practical definition of “software-defined”]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(2):85–94.
3. What’s a software defined data center? Available at: <https://www.sdxcentral.com/data-center/definitions/software-defined-data-center> (accessed November 28, 2019).
4. 5 faktov o programmno-opredelyaemykh TSOD [5 facts on software-defined data centers]. Available at: <https://www.cloud4y.ru/about/news/5-faktov-o-programmno-opredelyaemykh-tsod/> (accessed November 28, 2019).

5. *Egorov, V. B.* 2016. Nekotorye voprosy prakticheskoy realizatsii kontseptsii SDN [Some issues of the SDN concept practical implementation]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 26(1):109–120.
6. *Beal, V.* SDDC — software-defined data center. *Webopedia*. Available at: [https://www.webopedia.com/TERM/S/software\\_defined\\_data\\_center\\_SDDC.html](https://www.webopedia.com/TERM/S/software_defined_data_center_SDDC.html) (accessed November 28, 2019).
7. *Pritula, P.* 2016. Programmno-opredelyaemyy TSOD: marketologi peredali estafetu praktikam [The software-defined data center: Marketologists handed the baton to practitioners]. *CNews*. Available at: [http://www.cnews.ru/articles/2016-12-22-programmnoopredelyaemyj\\_tsod\\_marketologi\\_peredali\\_estafetu\\_praktikam](http://www.cnews.ru/articles/2016-12-22-programmnoopredelyaemyj_tsod_marketologi_peredali_estafetu_praktikam) (accessed November 28, 2019).
8. *Robinson, D.* 2018. SDDC is the answer — but how do we get there? *Data Center Dynamics*. Available at: <https://www.datacenterdynamics.com/analysis/sddc-is-the-answer-but-how-do-we-get-there/> (accessed November 28, 2019).
9. Jet Infosystems. 2015. Programmno-opredelyaemyy TSOD segodnya i zavtra [The software-defined data center today and tomorrow]. Available at: <https://jet.su/about/news/18464/> (accessed November 28, 2019).
10. Bitdefender. 1997. Securing the software-defined datacenter. Available at: <https://www.bitdefender.com/business/datacenter-security-solutions/sddc.html> (accessed November 28, 2019).
11. *Sanders, J.* 2017. Software defined data centers: A cheat sheet. *TechRepublic*. Available at: <https://www.techrepublic.com/article/software-defined-data-centers-the-smart-persons-guide/> (accessed November 28, 2019).

Received November 28, 2019

## Contributor

**Egorov Vladimir B.** (b. 1948)— Candidate of Science (PhD) in technology, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; VEGorov@ipiran.ru

## ЗАДАЧИ БАЗЫ ДАННЫХ ФРАЗЕОЛОГИЧЕСКОГО СЛОВАРЯ И СТАДИИ ЕЕ ПРОЕКТИРОВАНИЯ

*B. B. Вакуленко<sup>1</sup>, A. A. Гончаров<sup>2</sup>, A. A. Дурново<sup>3</sup>, И. М. Зацман<sup>4</sup>*

**Аннотация:** Проектирование базы данных (БД) двуязычного фразеологического словаря, поддерживающей генерацию его бумажных и электронных версий, во многом обусловлено структурой словарной статьи. Так, ее структура во многом определяет архитектуру БД в том, что касается обеспечения: (1) связанных компонентов одной словарной статьи (для двуязычного фразеологического словаря к ним относятся идиома, ее значения, представленные через различные способы перевода, примеры, комментарии и т. д.); (2) возможности добавления гиперссылок между двумя и более словарными статьями или их компонентами; (3) поиска, визуализации и редактирования словарных статей. При этом одним из условий проектирования лексикографических БД ставится раздельное описание логической структуры данных и форм представления этих данных при их визуализации в соответствии с запросами пользователей. Цель статьи состоит в описании подхода к расширению спектра форм представления словарных статей за счет их группировки на основе фасетной классификации, а также стадий реализации этого подхода.

**Ключевые слова:** структура словарной статьи; лексикографическая база данных; структура данных; визуализация словарных статей; проектирование базы данных

**DOI:** 10.14357/08696527200211

### 1 Введение

Проектируемая БД двуязычного фразеологического словаря предназначена для решения трех основных задач:

- (1) формирования, поиска и редактирования словарных статей;
- (2) генерации словаря для его издания в бумажной форме;
- (3) генерации словаря в электронной форме.

<sup>1</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, vvak@pm.me

<sup>2</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, a.gonch48@gmail.com

<sup>3</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, duralex49@mail.ru

<sup>4</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, izatsman@yandex.ru

Исходными информационными ресурсами для создания БД служат около 2000 файлов словарных статей. Проектируемая БД должна обеспечивать не только формирование статей с нуля, но также импорт текстов статей, созданных до формирования БД, структурирование наследуемых лексикографических ресурсов и рубрикацию словарных статей с использованием признаков, отражающих морфологию, синтаксис и семантику идиом, а также тематику контекстов их употребления [1, 2].

Основным условием проектирования лексикографических БД ставится раздельное описание логической структуры данных и форм представления данных при их визуализации в соответствии с запросами пользователей, а также описание формата словарных статей, обеспечивающее их экспорт из БД и генерацию текста словаря для издания в бумажной форме. Традиционно к основным задачам, которые решаются при помощи лексикографических БД, относятся [3–5]:

- (1) импорт и создание новых словарных статей для формирования словарей;
- (2) редактирование словарных статей;
- (3) добавление гиперссылок между двумя и более словарными статьями или их компонентами;
- (4) поиск и визуализация словарных статей;
- (5) формирование словарей для издания в бумажной и электронной форме.

Для решения этих задач к настоящему времени разработан широкий спектр средств проектирования лексикографических БД и формирования электронных словарей, краткий обзор которых приведен в работе [6]. Одним из наиболее известных средств проектирования является продукт TLex (или TshwaneLex), который предназначен как для профессиональных лексикографов, так и для обычных пользователей. Он поддерживает формированиеmonoязычных, двуязычных и многоязычных словарей как в бумажной, так и в электронной форме [7].

С учетом существования готовых инструментов для формирования лексикографических БД и электронных словарей закономерным становится вопрос о необходимости проектирования новой БД для двуязычного фразеологического словаря. Цель статьи — обоснование необходимости создания такой БД, которая концептуально отличается от существующих продуктов, описание подхода к расширению спектра форм представления словарных статей за счет их группировки на основе рубрик фасетной классификации семантических, грамматических и других признаков, а также описание стадий реализации этого подхода.

## **2 Группировка словарных статей**

Сама идея группировки словарных статей по тому или иному признаку не нова. Проиллюстрируем ее реализацию, сравнив два французско-русских сло-

варя [8, 9]. Если в первом из них открыть словарную статью, например, про книжный формат *in-folio*, то это не поможет найти статьи о других форматах, поскольку для этого в бумажной версии словаря нужно заранее знать их названия (*in-quarto*, *in-octavo*, *in-seize*, *in-trente-deux*)<sup>1</sup>. Во втором словаре используется группировка отдельных словарных статей по 85 семантическим классам (времена года, страны света, дни недели, месяцы, книжные форматы, степени родства, возрастные периоды, знаки препинания и др.). Поэтому, выбирая семантический класс «*le format* — книжный формат», читатель получает список всех выше-перечисленных названий форматов на французском языке с их переводами на русский язык.

Другой пример реализации этой идеи — Тезаурус [10]. Он включает около 6000 идиом русского языка, сгруппированных по смысловым категориям, которые образуют многоуровневую иерархию. Верхний уровень этой иерархии включает 87 категорий («1. Время», «2. Пространство», «3. Движение», ... «87. Вещество»). При этом если в словаре [8] семантический класс «воздрастные периоды» включает 9 периодов и входит в список из 85 разнородных семантических классов, то в Тезаурусе идиомы про три возрастных периода («1.3.4.1. Детство, молодость»<sup>2</sup>, «1.3.4.2. Зрелость»<sup>3</sup>, «1.3.4.3. Старость»<sup>4</sup>) находятся на четвертом уровне иерархии ниже следующих трех уровней: «1. Время», «1.3. Выделенный временной интервал»<sup>5</sup>, «1.3.4. Возраст»<sup>6</sup>.

Таким образом, в первом примере для группировки словарных статей использовался линейный список семантических классов, а во втором примере основой для группировки идиом была иерархия семантических категорий. Однако кроме семантических для группировки могут использоваться и другие признаки (морфологические, синтаксические, функциональные и т. д.), при использовании которых могут образовываться те или иные значимые для пользователей лексикографические категории. Отметим, что группировка на основании разнородных признаков и их сочетаний, как правило, предполагает использование фасетных классификаций.

Важно отметить, что формирование лексикографических БД и электронных словарей с использованием фасетных классификаций стало новой и во многом нерешенной проблемой из-за отсутствия релевантных классификаций лексикографических категорий и средств автоматизации их формирования. В предла-

<sup>1</sup> Цитируется по словарю [9], включенному в ABBYY Lingvo x3. В навигационной панели его электронной версии перечисленные книжные форматы находятся рядом или недалеко друг от друга, но это далеко не всегда так для языковых единиц, близких по значению.

<sup>2</sup> В нежном возрасте, от горшка два вершка, молоко на губах не обсохло, под стол пешком ходить, молодо-зелено [10].

<sup>3</sup> В летах, не первой молодости, в отцы годиться, в самом соку, во цвете лет [10].

<sup>4</sup> Старая гвардия, на закате лет, последний из могикан, на старости лет [10].

<sup>5</sup> Золотой век, бабье лето, медовый месяц, мертвый сезон, черная суббота, тихий час [10].

<sup>6</sup> ... (какой-либо) десяток разменять, не по годам (например, оказался не по годам умным мальчиком) [10].

гаемом подходе эта проблема включает постановку и решение следующих задач методами компьютерной лингвистики:

- (1) сформировать представительный массив структурированных словарных статей;
- (2) создать БД словарных статей, обеспечивающую формирование и использование фасетной классификации [11];
- (3) присвоить словарным статьям релевантные рубрики сформированной фасетной классификации.

Новизна концепции проектируемой БД двуязычного фразеологического словаря состоит в том, что она должна обеспечивать решение всех трех задач.

### 3 Структура словарной статьи

Исходными данными для решения первой из перечисленных задач — формирования массива словарных статей — служат около 2000 статей, описывающих немецкие идиомы и их переводы на русский язык. Каждая статья начинается с вокабулы, т. е. с опорного слова (см. подробнее об определении вокабулы [12, с. 595–597]), которое позволяет упорядочить идиомы по алфавиту, осуществлять их поиск и группировку. Вокабула, таким образом, представляет собой средство алфавитизации идиом, необходимое для их представления в бумажных версиях словаря. В электронных словарях поиск и группировку статей можно выполнить, как правило, по любому слову леммы.

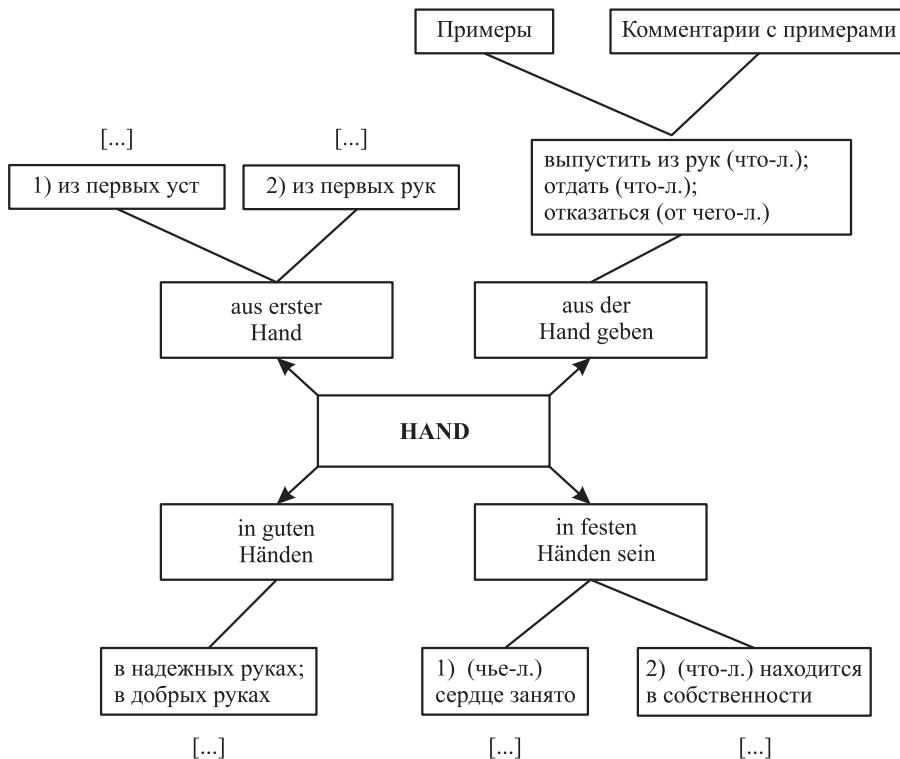
За вокабулой следует лемма — идиома в традиционной словарной форме, значение которое описывается через варианты перевода. В словаре каждой лемме соответствует одна словарная статья. В качестве иллюстрации рассмотрим четыре леммы на немецком языке со словом *Hand* и их переводы на русский язык<sup>1</sup> (примеры использования идиом и комментарии здесь не приводятся):

- aus der Hand geben — выпустить из рук (что-л.); отдать (что-л.); отказаться (от чего-л.);
- aus erster Hand — 1) из первых уст; 2) из первых рук;
- in guten Händen — в надежных руках; в добрых руках;
- in festen Händen sein — 1) (чье-л.) сердце занято; 2) (что-л.) находится в собственности.

Слово *Hand* для этих четырех лемм служит вокабулой. Первая лемма имеет одно значение и три варианта перевода, вторая — два значения с одним вариантом перевода каждое, третья — одно значение и два варианта перевода, четвертая — два значения с одним вариантом перевода каждое. В словарной статье описание

---

<sup>1</sup>Леммы и их переводы взяты из рукописи немецко-русского фразеологического словаря, который готовится к изданию под редакцией Д. О. Добровольского.



Связи между вокабулой и леммами

значения леммы содержит как минимум один пример использования идиомы в этом значении и может включать комментарии. В конце статьи указывается ее автор.

На рисунке графически представлены связи между вокабулой *Hand*, вышеперечисленными леммами, их значениями, примерами и комментариями из четырех словарных статей. Структура словарной статьи создаваемого немецко-русского фразеологического словаря представлена в таблице. Эта структура является общей для всех статей, но наборы компонентов отдельных статей могут отличаться.

Серым цветом выделен минимальный набор обязательных компонентов. В таблице указано 8 таких компонентов: (1) вокабула; (2) лемма; (3) зона значения; (4) значение ( $\min^1 1$ ); (5) перевод ( $\min 1$ ); (6) зона примеров ( $\min 1$ ); (7) пример ( $\min 1$ ); (8) автор статьи ( $\min 1$ ).

<sup>1</sup>Минимальные значения указаны для тех элементов, которых в принципе может быть больше одного; в таблице возможности повторения структурных элементов отмечены многоточием.

## Структура словарной статьи

Вока-була	Лемма	Пропозициональная форма леммы					
		Вариант (грамм.) леммы					
		Варианты (лекс.) леммы					
		Стилистические пометы					
		Зона значе-ния	Значе-ние 1	Вариант (грамм.) леммы в данном значении ( <i>возможно, если значений &gt; 1</i> )			
				Пропозициональная форма леммы в данном значении ( <i>возможно, если значений &gt; 1</i> )			
		Зона приме-ров	Перевод леммы в данном значении	Перевод леммы в данном значении			
				Варианты (лекс.) леммы в данном значении			
				Пример 1	Текст примера 1 на нем. языке		
					Источник примера 1		
					Перевод примера 1 на русский язык		
			Пример M	...			
				Текст примера M на нем. языке			
				Источник примера M			
				Перевод примера M на русский язык			
		Значе-ние N	Особенности употребления леммы в данном значении (так называемый «комментарий»)				
			...				
			...				
Гиперссылки на другие леммы							
Автор словарной статьи							

Все прочие элементы статьи являются факультативными и могут относиться либо к лемме в целом, либо к отдельным ее значениям. Это пропозициональная форма леммы (личная форма, противопоставляемая словарной), грамматические варианты леммы (например, отражающие вариативность модели управления или использования artikelей), лексические варианты леммы (отражающие вариативность лексического состава), стилистические пометы, гиперссылки.

Комментарии, раскрывающие особенности употребления леммы (в том числе с точки зрения межъязыковых различий), обычно относятся к зоне примеров,

однако они могут относиться и к другим зонам статьи. Более детальное рассмотрение структуры словарной статьи на примерах конкретных идиом немецкого языка и их описания в проектируемом немецко-русском фразеологическом словаре см. в [12, с. 597–602].

## 4 Заключение

Проблема формирования БД и электронных словарей с использованием фасетных классификаций включает три основные задачи, перечисленные в конце разд. 2, которые существенно различаются по их сложности из-за отсутствия релевантных фасетных классификаций лексикографических категорий (ФКЛК), значимых для пользователей словарей. Поэтому проектирование БД планируется выполнить в несколько стадий. При этом наиболее сложные задачи предполагается перенести на завершающую стадию.

На первой стадии формируется массив структурированных словарных статей на основе данных неструктурированных исходных файлов в формате \*.doc (один файл содержит текст одной статьи). Для структурирования текстов и разметки словарных статей в соответствии с таблицей была разработана система тегов и соответствующих им макросов. После разметки словарных статей с помощью макросов появляется возможность программной загрузки структурированных статей в БД, создаваемую на второй стадии.

Отметим, что на второй стадии задача создания БД словарных статей с возможностью формирования и использования ФКЛК решается только частично: сначала проектируется БД, обеспечивающая загрузку наследуемых статей, создание новых словарных статей, их поиск и редактирование без использования ФКЛК, а также экспорт структурированных статей. Это позволяет уже по завершении этой стадии опубликовать словарь в бумажной и/или электронной форме, но пока без возможности поиска и группировки словарных статей при помощи рубрик ФКЛК.

Наиболее трудоемкая и длительная третья стадия необходима для формирования ФКЛК и приписывания ее рубрик словарным статьям. На этой стадии одновременно завершается решение второй задачи и полностью решается третья задача. Для достижения этой цели планируется адаптировать метод лингвистического аннотирования параллельных текстов в надкорпусных базах данных [13–16] и применить его к двуязычным словарным статьям, уже хранящимся в БД. Предлагаемый подход даст возможность по завершении этой стадии сформировать словарь в электронной форме с функцией группировки словарных статей по рубрикам ФКЛК.

Возможность группировки словарных статей существенно расширяет спектр видов поиска и форм их визуализации за счет использования рубрик ФКЛК. К числу новых видов относится поиск словарных статей по рубрикам лексикографических категорий, таких, например, как семантические классы (времена года, страны света, дни недели и др.).

## Литература

1. Phraseologie / Eds. H. Burger, D. Dobrovolskij, P. Kühn, N. R. Norrick. — Berlin – New York: Walter de Gruyter, 2007. Vol. 1. 613 p.
2. Dobrovolskij D. The notion of “inner form” and idiom semantics // Proverbes et stéréotypes: Formes, forme et contextes / Ed. S. Viellard. — Paris: Université Paris-Sorbonne – CNRS, 2016. P. 21–35.
3. De Schryver G.-M. Lexicographers’ dreams in the electronic-dictionary age // Int. J. Lexicogr., 2003. Vol. 16. No. 2. P. 143–199.
4. Tarp S. Lexicography in the borderland between knowledge and non-knowledge: General lexicographical theory with particular focus on learner’s lexicography. — Tübingen: Max Niemeyer Verlag, 2008. 308 p.
5. Müller-Spitzer C. Textual structures in electronic dictionaries compared with printed dictionaries: A short general survey // Dictionaries: An international encyclopedia of lexicography. Supplementary Volume: Recent developments with focus on electronic and computational lexicography. — Berlin: De Gruyter Mouton, 2013. P. 367–381.
6. Suhardijanto T., Dinakaramani A. Building a collaborative workspace for lexicography works in Indonesia // Electronic Lexicography in the 21st Century Conference Proceedings. — Brno: Lexical Computing, 2017. P. 299–308.
7. TLex Suite User Guide. <https://www.tshwanedje.com/docs/TLex%20Suite%20User%20Guide.pdf>.
8. Гак В. Г., Триомф Ж. Французско-русский словарь активного типа. — М.: Русский язык, 1991. 1056 с.
9. Гак В. Г., Ганишина К. А. Новый французско-русский словарь. — М.: Дрофа, 2010. 1160 с.
10. Баранов А. Н., Добровольский Д. О. Тезаурус русских идиом: Семантические группы и контексты. — М.: Лекрус, 2018. 888 с.
11. Зацман И. М., Инькова О. Ю., Нуриев В. А. Построение классификационных схем: методы и технологии экспертного формирования // Научно-техническая информация. Сер. 2: Информационные процессы и системы, 2017. № 1. С. 8–22.
12. Добровольский Д. О. Беседы о немецком слове. — М.: Языки славянской культуры, 2013. 744 с.
13. Зализняк Анна А., Зацман И. М., Инькова О. Ю., Кружков М. Г. Надкорпусные базы данных как лингвистический ресурс // Корпусная лингвистика: Тр. 7-й Междунар. конф. — СПб.: СПбГУ, 2015. С. 211–218.
14. Дурново А. А., Зацман И. М., Лощилова Е. Ю. Кросслингвистическая база данных для аннотирования логико-семантических отношений в тексте // Системы и средства информатики, 2016. Т. 26. № 4. С. 124–137.
15. Зализняк А. А., Зацман И. М., Инькова О. Ю. Надкорпусная база данных коннекторов: построение системы терминов // Информатика и её применения, 2017. Т. 11. Вып. 1. С. 100–108.
16. Егорова А. Ю., Зацман И. М., Мамонова О. С. Надкорпусные базы данных в лингвистических проектах // Системы и средства информатики, 2019. Т. 29. № 3. С. 77–91.

*Поступила в редакцию 11.03.20*

## TASKS OF THE PHRASEOLOGICAL DICTIONARY DATABASE AND STAGES OF ITS DESIGN

**V. V. Vakulenko, A. A. Goncharov, A. A. Durnovo, and I. M. Zatsman**

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** The design of the database (DB) for a bilingual phraseologic dictionary which supports the generation of paper and digital versions depends largely on the structure of the dictionary entry. This structure defines the DB's architecture in the following aspects: (i) the connectedness of the components of a single dictionary entry (for a bilingual phraseologic dictionary, the components are: idiom and its meanings represented by translation variants, examples, commentary, etc.); (ii) the ability to add hyperlinks between two and more entries and their components; and (iii) searching, visualizing, and editing entries. Additionally, when designing lexicographical DBs, one of the conditions includes a separate description of the logical data structure and the representation forms of this data during its visualization in accordance with the user's query. The goal of this paper is to describe the approach to expanding the range of dictionary entry representation forms by clustering the entries according to the faceted classification and to display the stages of implementing this approach.

**Keywords:** dictionary entry structure; lexicographical database; data structure; dictionary entry visualization; database design

**DOI:** 10.14357/08696527200211

### References

1. Burger, H., D. Dobrovolskiy, P. Kuehn, and N. R. Norrick, eds. 2007. *Phraseologie*. Berlin–New York: Walter de Gruyter. Vol. 1. 613 p.
2. Dobrovolskiy, D. 2016. The notion of “inner form” and idiom semantics. *Proverbes et stéréotypes: Formes, forme et contextes*. Ed. S. Viillard. Paris: Université Paris-Sorbonne–CNRS. 21–35.
3. De Schryver, G.-M. 2003. Lexicographers’ dreams in the electronic-dictionary age. *Int. J. Lexicogr.* 16(2):143–199.
4. Tarp, S. 2008. *Lexicography in the borderland between knowledge and non-knowledge. General lexicographical theory with particular focus on learner’s lexicography*. Tübingen: Max Niemeyer Verlag. 308 p.
5. Müller-Spitzer, C. 2013. Textual structures in electronic dictionaries compared with printed dictionaries: A short general survey. *Dictionaries. An international encyclopedia of lexicography: Supplementary Volume: Recent developments with focus on electronic and computational lexicography*. Berlin: De Gruyter Mouton. 367–381.

6. Suhardijanto, T., and A. Dinakaramani. 2017. Building a collaborative workspace for lexicography works in Indonesia. *Lexicography in the 21st Century Conference Proceedings*. Eds. I. Kosem and C. Tiberius. Brno: Lexical Computing. 299–308.
7. TLex Suite User Guide. Available at: <https://www.tshwanedje.com/docs/TLex%20Suite%20User%20Guide.pdf> (accessed March 8, 2020).
8. Gak, V. G., and Zh. Triomf. 1991. *Frantsuzsko-russkiy slovar' aktivnogo tipa* [French-Russian dictionary of the active type]. Moscow: Russkiy yazyk. 1056 p.
9. Gak, V. G., and K. A. Ganshina. 2010. *Novyy frantsuzsko-russkiy slovar'* [New French-Russian dictionary]. Moscow: Drofa. 1160 p.
10. Baranov, A. N., and D. O. Dobrovolskiy. 2018. *Tezaurus russkikh idiom. Semantic-skiye gruppy i konteksty* [Thesaurus of Russian idioms. Semantic groups and contexts]. Moscow: Lekrus. 888 p.
11. Zatsman, I. M., O. Yu. Inkova, and V. A. Nuriev. 2017. The construction of classification schemes: Methods and technologies of expert formation. *Automatic Documentation Mathematical Linguistics* 51(1):27–41.
12. Dobrovolskiy, D. O. 2013. *Besedy o nemetskom slove* [Studies on German lexis]. Moscow: Yazyki slavyanskoy kul'tury. 744 p.
13. Zaliznyak, A. A., I. M. Zatsman, O. Yu. Inkova, and M. G. Krushkov. 2015. Nadkorpusnye bazy dannykh kak lingvisticheskiy resurs [Supracorpora databases as linguistic resource]. *Corpus Linguistics: 7th Conference (International) Proceedings*. St. Petersburg. 211–218.
14. Durnovo, A. A., I. M. Zatsman, and E. Yu. Loshilova. 2016. Krosslingvisticheskaya baza dannykh dlya annotirovaniya logiko-semanticeskikh otnosheniy v tekste [Cross-lingual database for annotating logical-semantic relations in the text]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 26(4):124–137.
15. Zaliznyak, A. A., I. M. Zatsman, and O. Yu. Inkova. 2017. Nadkorpusnaya baza dannykh konnektorov: postroenie sistemy terminov [Supracorpora database on connectives: Term system development]. *Informatika i ee Primeneniya — Inform. Appl.* 11(1):100–108.
16. Egorova, A. Yu., I. M. Zatsman, and O. S. Mamanova. 2019. Nadkorpusnye bazy dannykh v lingvisticheskikh proektakh [Supracorpora databases in linguistic projects]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(3):77–91.

Received March 11, 2020

## Contributors

**Vakulenko Vasily V.** (b. 1995) — engineer-researcher, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; vvvak@pm.me

**Goncharov Alexander A.** (b. 1994) — junior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian

Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation;  
a.gonch48@gmail.com

**Durnovo Aleksandr A.** (b. 1949) — leading programmer, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; duralex49@mail.ru

**Zatsman Igor M.** (b. 1952) — Doctor of Science in technology, Head of Department, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; izatsman@yandex.ru

## НЕСТАБИЛЬНОСТЬ НЕЙРОННОГО МАШИННОГО ПЕРЕВОДА\*

*А. Ю. Егорова<sup>1</sup>, И. М. Зацман<sup>2</sup>, В. В. Косарик<sup>3</sup>, В. А. Нуриев<sup>4</sup>*

**Аннотация:** Дано описание эксперимента по исследованию нестабильности нейронного машинного перевода (НМП). Сформированный массив текстовых фрагментов на русском языке повторно и многократно переводился на французский язык в течение одного года с шагом в один месяц посредством системы НМП Google. Это дало возможность зафиксировать нестабильность перевода одних и тех же фрагментов текстов. Затем было выполнено лингвистическое аннотирование полученных переводов и определено несколько видов нестабильности НМП. При этом использовалась ранее разработанная типология ошибок перевода. В процессе аннотирования в нее был внесен ряд изменений, описание которых приведено в статье. Конечная цель эксперимента — получить частотную оценку видов нестабильности НМП. На первом его этапе, описанию результатов которого и посвящена статья, ставилась задача провести категоризацию нестабильности. Эмпирическим материалом для эксперимента служат русско-французские аннотации, сформированные в надкорпусной базе данных. Каждая аннотация содержит одно или несколько предложений оригинального текста на русском языке, их перевод и описание встретившихся ошибок перевода.

**Ключевые слова:** машинный перевод; нестабильность; мониторинг результатов перевода; лингвистическое аннотирование; виды нестабильности

**DOI:** 10.14357/08696527200212

### 1 Введение

В 2018 г. в издательстве «Шпрингер» началась публикация серии книг, посвященных технологиям и применению систем машинного перевода [1, 2]. Первый том этой серии полностью посвящен анализу результатов машинного

---

\*Работа выполнена в Институте проблем информатики ФИЦ ИУ РАН при поддержке РФФИ (проект по гранту № 18-07-00192).

<sup>1</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, ann.shurova@gmail.com

<sup>2</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, izatsman@yandex.ru

<sup>3</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, valery\_lek@mail.ru

<sup>4</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, nurieff.v@gmail.com

перевода и включает обзор экспертного и автоматизированного подходов к классификации и аннотированию его ошибок, что говорит об актуальности этой проблематики. Сравнение двух указанных подходов позволило сделать вывод, что первый из них может дать более подробное описание ошибок, поскольку эксперты могут различать большее число их видов, чем автоматизированный подход. При этом отмечается, что выбор релевантной типологии ошибок для анализа результатов машинного перевода представляет собой самостоятельную и довольно сложную задачу. Автоматизированный анализ ошибок выполняется быстрее и требует меньше финансовых ресурсов, но он, как правило, не дает подробного описания ошибок и не позволяет провести тонкую дифференциацию разных их видов, хотя это не всегда является легким и для экспертов [3]. Отметим, что использование автоматизированного анализа может упростить реализацию и экспертного подхода, например, если ввести этап компьютерного сравнения повторных / многократных переводов, чтобы пометить переводные варианты, не изменившиеся в период проведения эксперимента.

В статье описывается эксперимент по мониторингу НМП на основе экспертного подхода. Сама идея такого эксперимента возникла в ходе российско-швейцарского проекта, поддержанного РФФИ (грант № 16-24-41002) и Швейцарским национальным научным фондом (грант № IZLRZ1\_164059), когда было показано, что НМП одних и тех же фрагментов текста может изменяться со временем, т. е. быть нестабильным.

В работе [4] приведен пример двукратного машинного перевода одного и того же предложения из пьесы А. П. Чехова «Дядя Ваня», который был выполнен в феврале и октябре 2018 г. (табл. 1). При этом второй перевод явно уступает первому, так как в нем слово «impudent» повторено два раза, что существенно ухудшило качество перевода.

Поэтому кроме оценки качества машинного перевода в статике [5–7] целесообразно разработать подход к описанию нестабильности НМП с использованием классификации ошибок перевода и качественной оценки перевода во времени. Отметим, что задача исследования нестабильности отличается от задачи повышения устойчивости НМП, когда небольшие возмущения на входе могут серьезно исказить промежуточные представления и, таким образом, повлиять на качество перевода [8]. В интересах разработки такого подхода был сформирован массив

**Таблица 1** Пример результата машинного перевода

Когда бываю в таком состоянии, то становлюсь нахальным и наглым до крайности.	Quand je suis dans cet état, je deviens impudent et insolent à l'extrême. (translate.google.com, дата обращения 05.02.2018 19:48)
-------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------

**Примечание.** С момента записи в надкорпусную базу данных коннекторов 05.02.2018 перевод, предлагаемый системой Google's Neural Machine Translation, изменился следующим образом: «Quand je suis dans un tel état, je deviens impudent et impudent à l'extrême» (11.10.2018 19:35).

русскоязычных текстовых фрагментов, полученных из Национального корпуса русского языка [9], и затем был проведен эксперимент по машинному переводу, описание результатов которого и стало основной целью статьи.

Для сформированного массива текстовых фрагментов сначала был выполнен их перевод на французский язык посредством системы НМП Google (Google's Neural Machine Translation system, <https://translate.google.com>). Архитектура этой системы включает в себя две рекуррентные нейронные сети с долгой краткосрочной памятью (long short-term memory). Одна — кодирующая — обрабатывает входную текстовую последовательность, другая — декодирующая — отвечает за формирование выходного перевода. Сети имеют восьмислойную структуру, где связь между слоями помогают осуществлять остаточные соединения (residual connections), обеспечивающие градиентный поток (gradient flow). Кодирующий и декодирующий сегменты соединяются модулем внимания (attention module), который позволяет оптимизировать работу с протяженными входными последовательностями. Чтобы обеспечить параллелизм при обработке данных, внимание нижнего уровня декодирующей сети подключается к верхнему уровню кодирующей сети. Для уменьшения времени вывода применяются вычисления с низкой точностью (low-precision arithmetic), а для дополнительного ускорения — специальное оборудование (тензорный процессор Google) (подробнее см. в [10]).

Перевод в системе Google выполнялся многократно в течение одного года с временным шагом в один месяц (с марта 2019 г. по февраль 2020 г.). Это дало возможность получить серию переводов для каждого фрагмента. В настоящей статье проанализированы переводы, полученные на протяжении первых 10 месяцев эксперимента. В результате программного анализа были помечены и исключены из рассмотрения те фрагменты, переводы которых оставались неизменными в период проведения эксперимента. Таким образом, предметом последующей экспертизы стали серии переводов только тех текстовых фрагментов, в отношении которых фиксировалась нестабильность перевода.

На стадии экспертизы было выполнено лингвистическое аннотирование каждого отобранного фрагмента текста на русском языке и серии его переводов на французский язык, включающее описание встретившихся в переводе ошибок. Метод аннотирования и виды аннотаций описаны в работе [11]. Аннотирование каждого фрагмента и соответствующей серии переводов дало возможность зафиксировать нестабильность НМП и ее виды. На первом этапе экспертизы ставилась задача провести категоризацию нестабильности НМП.

## **2 Экспериментальные данные**

Для оценки изменения НМП был проведен эксперимент по ежемесячному аннотированию результатов машинного перевода с помощью надкорпусной базы данных (НБД) [12, 13]. Источником текстов (объем которых составляет около 4 млн словоупотреблений), используемых в НБД для поиска и анно-

**Таблица 2** Примеры экспорттированных из НБД строк

№*	Дата изменения	Р** в оригинале	Контекст РР в оригинале	Контекст РР в переводе
28365	04.03.2019 14:06	не  а	Гость направился не за город, а в город.	L'invité n'est pas allé en ville, mais en ville.
28860	05.04.2019 15:03	не  а	Гость направился не за город, а в город.	-"-
29043	02.05.2019 14:52	не  а	Гость направился не за город, а в город.	-"-
29978	06.06.2019 14:33	не  а	Гость направился не за город, а в город.	-"-
30286	04.07.2019 14:53	не  а	Гость направился не за город, а в город.	-"-
30513	04.08.2019 15:35	не  а	Гость направился не за город, а в город.	L'invité ne s'est pas dirigé hors de la ville mais dans la ville.
30873	07.09.2019 21:48	не  а	Гость направился не за город, а в город.	L'invité ne s'est pas dirigé hors de la ville, mais dans la ville.
31137	04.10.2019 15:00	не  а	Гость направился не за город, а в город.	-"-
31701	07.11.2019 16:18	не  а	Гость направился не за город, а в город.	-"-
32319	06.12.2019 16:34	не  а	Гость направился не за город, а в город.	L'invité n'est pas sorti de la ville, mais dans la ville.

\*Номер двуязычной аннотации в НБД, включающей фрагмент текста оригинала и его перевод.

\*\*РР (речевая реализация) — та форма, в которой коннектор встречается в конкретном высказывании [15].

тирования языковых единиц — в данном случае коннекторов,<sup>1</sup> — послужил русско-французский подкорпус Национального корпуса русского языка. Из текстов НБД коннекторов было отобрано для эксперимента 250 текстовых фрагментов на русском языке (см. их примеры в 4-м столбце в табл. 2), каждый из которых содержит один из 10 двухкомпонентных коннекторов (таких как *не (расстояние) а*; *хотя (расстояние) но* и т. п.). На каждый коннектор при этом приходится в среднем 25 фрагментов (от 9 до 51 фрагмента). Для эксперимента были выбраны двухкомпонентные коннекторы, которые в силу своей структуры чаще вызывают трудности при переводе.

<sup>1</sup> Коннектор — языковая единица, функция которой состоит в выражении логико-семантических отношений, существующих между двумя соединенными с ее помощью текстовыми компонентами, имеющими предикативный характер (см. [14, с. 17]).

В ходе эксперимента выполнялось ежемесячное (в течение первых 7 дней каждого месяца) формирование 250 аннотаций, включающих фрагмент на русском языке, его перевод на французский, выполненный с помощью системы НМП Google, и описание ошибок, если таковые имели место.

Для удобства сопоставления и анализа полученные экспериментальные данные были экспортированы из НБД в табличной форме (см. табл. 2). Что касается машинных переводов, оставшихся неизменными на протяжении двух и более месяцев подряд, то они были заменены на знак повтора. Общее число строк после экспорта составило 2500. При этом в течение первых 10 месяцев проведения эксперимента переводы 10 из 250 фрагментов оставались неизменными (что составляет 4% от общего числа фрагментов).

### **3 Классификация ошибок в машинных переводах**

В процессе лингвистического аннотирования переводов применялась классификация ошибок, которая специально предназначена для анализа машинных переводов. Она является развитием ранее созданной классификации. Ее первая версия появилась в результате анализа переводов, сделанных посредством системы Яндекс.Переводчик в 2016–2017 гг. Первоначально она включала 8 рубрик для обозначения ошибок в переводе как текстовых фрагментов с коннектором, так и самих коннекторов (подробнее см. в [5]), а также рубрику «Ошибка нет» (NoError), применявшуюся для маркировки допустимых вариантов перевода.

Работа над второй версией классификации [6] началась в 2018 г. Она проводилась по результатам анализа переводов, полученных с помощью системы НМП Google. Вторая версия классификации включала уже 15 рубрик (помимо рубрики «Ошибка нет»). Отличие использованной в эксперименте третьей версии классификации ошибок состоит в том, что она была разработана для оценки качества НМП не в статике, а в динамике. По сравнению с работами [4–6] в эксперименте фокус переместился с анализа перевода коннектора на перевод всего содержащего коннектор фрагмента целиком, так как только детальный анализ всех ошибок позволяет оценить качество перевода и проследить нестабильность НМП во времени.

В табл. 3 представлено сопоставление второй и третьей версий классификации. Не считая рубрики «Ошибка нет», в третью версию вошло 20 рубрик (для наглядности новые рубрики выделены полужирным шрифтом). Классификация была уточнена следующим образом. Два признака из второй версии, указывающие на «локальную аграмматичность фрагмента» (AgramLocal и AgramPostCNT), были специфицированы по грамматическим уровням. Также были добавлены три новых признака, учитывающие ранее не затронутые аспекты (пунктуацию и семантику всего текстового фрагмента). В ходе экспертизы НМП также была выявлена необходимость ввести новые рубрики, фиксирующие вариативность допустимых версий перевода.

**Таблица 3** Сопоставление второй и третьей версий классификации

	Название рубрики		Код рубрики	
	2-я версия	3-я версия	2-я версия	3-я версия
Рубрики, относящиеся к переводу текстового фрагмента	Все предложение аграмматично		AgramTotal	
	Локальная аграмматичность во фрагменте текста, не вводимом коннектором	<b>Синтаксическая ошибка во фрагменте текста, не вводимом коннектором</b>	AgramLocal	<b>ErrorSyntax</b>
		<b>Морфологическая ошибка во фрагменте текста, не вводимом коннектором</b>		<b>ErrorMorph</b>
	Локальная аграмматичность во фрагменте текста, вводимом коннектором	<b>Синтаксическая ошибка во фрагменте текста, вводимом коннектором</b>	AgramPostCNT	<b>ErrorSyntax-PostCNT</b>
		<b>Морфологическая ошибка во фрагменте текста, вводимом коннектором</b>		<b>ErrorMorph-PostCNT</b>
	Лексическая ошибка во фрагменте текста		ErrorLex	
	—	<b>Пунктуационная ошибка</b>	—	<b>ErrorPunct</b>
	—	<b>Семантическая ошибка, искажение смысла</b>	—	<b>ErrorSemant</b>
	—	<b>Избыточный перевод*</b>	—	<b>Pleonasm</b>
	Слова кириллицей		Cyrillic	
Рубрики, относящиеся к переводу коннектора	Русское слово латинским шрифтом		Latin	
	Пропуск фрагмента текста		Lacuna	
	Переведена первая часть неоднословного коннектора		TrPartI	TrPart <sub>1</sub> CNT
	Переведена вторая часть неоднословного коннектора		TrPartII	TrPart <sub>2</sub> CNT
	Первая часть неоднословного коннектора переведена ошибочно		ErrorPartI	ErrorPart <sub>1</sub> CNT
	Вторая часть неоднословного коннектора переведена ошибочно		ErrorPartII	ErrorPart <sub>2</sub> CNT
	Коннектор переведен несуществующей единицей		ErrorTotal	<b>ErrorTotalCNT</b>
	Коннектор ошибочно заменен языковой единицей, не являющейся коннектором		ErrorMorph	<b>ErrorMorphCNT</b>
Орфографическая ошибка в форме коннектора		ErrorOrth	<b>ErrorOrthCNT</b>	
Семантическая ошибка в выборе коннектора		ErrorCNT		

\* Имеется в виду дублирование при переводе элемента смысла.

## 4 Категоризация результатов нейронного машинного перевода

Третья версия классификации ошибок позволила снабдить сформированные аннотации дополнительными рубриками-признаками, сигнализирующими о появлении и/или исчезновении ошибок НМП, и указать в аннотациях типы ошибок. Это дало возможность описать виды нестабильности НМП с использованием рубрик классификации ошибок, а также оценить частотность тех или иных типов ошибок для каждого вида нестабильности.

В результате экспертизы эмпирического массива была проведена категоризация нестабильности НМП. Было выявлено пять видов нестабильности, которые иллюстрируются примерами.

1. Повышение качества НМП: число ошибок в последующих переводах уменьшается. Так, в переводе от 05.08.19 исправлена морфологическая ошибка (неверное местоимение), допущенная в переводе от 05.03.19<sup>1</sup>.

Существуют обязанности не только по отношению к Богу и другим людям, но и к самому себе.	Il y a des obligations <i>non seulement</i> envers Dieu et envers les autres, <i>mais aussi</i> envers <b><u>lui-même</u></b> . (аннотация сформирована 05.03.19, ErrorMorph-PostCNT)	Il y a des obligations <i>non seulement</i> envers Dieu et envers les autres, <i>mais aussi</i> envers <b><u>soi-même</u></b> . (аннотация сформирована 05.08.19, NoError)
------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------

2. Снижение качества НМП: число ошибок в НМП увеличивается с течением времени. Например, вариант перевода от 04.03.19 приемлемый, а в переводе от 06.12.19 появляется лексическая ошибка.

Не жизнь, а санаторий для партийных работников.	Pas la vie, <i>mais un sanatorium pour les travailleurs du parti.</i> (аннотация сформирована 04.03.19, NoError)	Pas la vie, <i>mais un sanatorium pour les fêtards.</i> (аннотация сформирована 06.12.19, ErrorLex)
-------------------------------------------------	------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------

3. Колебание качества НМП между повышением и снижением: качество перевода сначала повышается (в переводе от 04.07.19 нет ошибок в отличие от перевода от 04.03.19, где выбран семантически неверный предлог), а затем понижается (в переводе от 06.12.19 есть лексическая ошибка: в данном контексте предлог *dans* не сочетается с глаголом *sortir*, во второй части

<sup>1</sup>За первые 10 месяцев эксперимента (с марта по декабрь 2019 г.) было зафиксировано 3 варианта перевода данного фрагмента. Принимая во внимание ограниченный объем статьи, здесь и далее будут представлены только те варианты перевода, которым при аннотировании присваивались разные наборы дополнительных признаков-рубрик, маркирующих ошибки.

предложения следовало добавить другой глагол). Изменения качества могут происходить неоднократно и в любой последовательности.

Гость направился <i>не</i> за город, <i>а</i> в город.	L'invité <i>n'est pas</i> allé <b>en</b> ville, <i>mais</i> en ville. (аннотация сформирована 04.03.19, ErrorSemant)	L'invité <i>ne s'est pas dirigé</i> hors de la ville <i>mais</i> dans la ville. (аннотация сформирована 04.07.19, NoError)	L'invité <i>n'est pas sorti</i> de la ville, <i>mais</i> <b>dans</b> la ville. (аннотация сформирована 06.12.19, ErrorLex)
--------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------

4. Чередование ошибок в НМП без динамики его качества: одна ошибка со временем исчезает, но появляется другая. В переводе от 07.08.19 нет морфологической ошибки во фрагменте с коннектором, которая была в переводе от 03.03.19 (неверное наклонение глагола *chercher*), однако остается та же лексическая ошибка (устойчивое выражение «искать руки» переведено словно), к тому же возникает другая морфологическая ошибка (неверное согласование местоимения в лице).

И <i>хотя</i> он, точно, искал руки моей дочери, будучи сам хорошего, трезвого поведения и великой учености, но я никогда не подавала ему никакой надежды.	Et <i>bien que</i> , certes, il <b>cherchait les mains</b> de ma fille, ayant lui-même un bon comportement sobre et une grande érudition, <i>mais</i> je ne lui ai jamais donné aucun espoir. (аннотация сформирована 03.03.19, ErrorMorphPostCNT, ErrorLex)	Et <i>bien qu'il ait certainement cherché les mains</i> de ma fille, étant <b>moi-même</b> d'un bon comportement sobre и d'une grande érudition, <i>mais</i> je ne lui ai jamais donné aucun espoir. (аннотация сформирована 07.08.19, ErrorLex, возникает другая ErrorMorphPostCNT)
------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

5. Изменение НМП без динамики его качества: средства перевода меняются, но качество остается на том же уровне. Так, перевод от 02.08.19 изменился по сравнению с переводом от 01.03.19 (одновременность действий передана с помощью придаточного времени вместо предложно-именной группы), но его качество осталось прежним — оба варианта допустимы. К этой категории относятся также переводы с ошибками, число и состав которых не меняются от перевода к переводу (при этом меняются средства перевода, т. е. применяются другие переводные эквиваленты).

Как она проснется, так он уже тут.	<u>À son réveil</u> , il est déjà là. (аннотация сформирована 01.03.19, NoError)	<u>Alors qu'elle se réveille</u> , il est déjà là. (аннотация сформирована 02.08.19, NoError)
------------------------------------	-------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------

В целом, качество перевода применительно к пяти описанным видам нестабильности может изменяться (как в 1–3) или оставаться на одном уровне, хотя перевод и претерпевает ряд изменений (виды 4–5). Также отметим, что категории 3 и 4 схожи в том, что переводы, где наблюдаются эти виды нестабильности, характеризует чередование ошибок, но категория 3 при этом фиксирует повышение/снижение качества перевода, а в категории 4 качество почти не меняется.

## 5 Заключение

В статье приведено описание первого этапа эксперимента по многократному фиксированию машинных переводов, полученных с помощью системы НМП Google. На этом этапе ставилась задача провести категоризацию нестабильности по видам. В ходе эксперимента была создана третья версия классификации ошибок в МП. В результате экспертизы полученных данных были выделены и описаны пять видов нестабильности НМП.

В дальнейшем планируется разработать методику анализа ошибок, включенных в третью версию классификации (см. табл. 2), и соотнести виды нестабильности НМП с рубриками ошибок. При этом планируется определить частотности рубрик ошибок, релевантных для каждого вида нестабильности НМП.

## Литература

1. Machine translation: Technologies and applications. Vol. 1: Translation quality assessment / Eds. J. Moorkens, S. Castilho, F. Gaspari, S. Doherty. — Cham: Springer International Publishing, 2018. 287 p.
2. Scott B. Machine translation: Technologies and applications. Vol. 2: Translation, brains and the computer: A neurolinguistic solution to ambiguity and complexity in machine translation. — Cham: Springer International Publishing, 2018. 241 p.
3. Popović M. Error classification and analysis for machine translation quality assessment // Machine translation: Technologies and applications. Vol. 1: Translation quality assessment. — Cham: Springer International Publishing, 2018. P. 129–158.
4. Бунтман Н. В., Гончаров А. А., Зацман И. М., Нуриев В. А. Количественный анализ результатов машинного перевода с использованием надкорпусных баз данных // Информатика и её применения, 2018. Т. 12. Вып. 4. С. 96–105.
5. Nuriev V., Buntman N., Inkova O. Machine translation of Russian connectives into French: Errors and quality failures // Информатика и её применения, 2018. Т. 12. Вып. 2. С. 105–113.
6. Гончаров А. А., Бунтман Н. В., Нуриев В. А. Ошибки в машинном переводе: проблемы классификации // Системы и средства информатики, 2019. Т. 29. № 3. С. 92–103.

7. Рычухин А. К. О методах оценки качества машинного перевода // Системы и средства информатики, 2019. Т. 29. № 4. С. 106–118.
8. Cheng Y., Tu Z., Meng F., Zhai J., Liu Y. Towards robust neural machine translation // 56th Annual Meeting of the Association for Computational Linguistics Proceedings. — Melbourne: Association for Computational Linguistics, 2018. Vol. 1. P. 1756–1766. <https://www.aclweb.org/anthology/P18-1163>.
9. Национальный корпус русского языка. <http://www.ruscorpora.ru>.
10. Нуриев В. А. Архитектура системы нейронного машинного перевода // Информатика и её применения, 2019. Т. 13. Вып. 3. С. 90–96.
11. Зализняк А. А., Зацман И. М., Инькова О. Ю. Надкорпусная база данных коннекторов: построение системы терминов // Информатика и её применения, 2017. Т. 11. Вып. 1. С. 100–108.
12. Зацман И. М., Кружков М. Г. Надкорпусная база данных коннекторов: развитие системы терминов проектирования // Системы и средства информатики, 2018. Т. 28. № 4. С. 156–167.
13. Егорова А. Ю., Зацман И. М., Мамонова О. С. Надкорпусные базы данных в лингвистических проектах // Системы и средства информатики, 2019. Т. 29. № 3. С. 77–91.
14. Инькова-Манзотти О. Ю. Коннекторы противопоставления во французском и русском языках. Сопоставительное исследование. — М.: Информэлектро, 2001. 429 с.
15. Инькова О. Ю. Надкорпусная база данных как инструмент формальной вариативности коннекторов // Компьютерная лингвистика и интеллектуальные технологии: по мат-лам ежегодной Международ. конф. «Диалог». — М.: РГГУ, 2018. Вып. 17(24). С. 240–253.

*Поступила в редакцию 03.03.20*

---

## INSTABILITY OF NEURAL MACHINE TRANSLATION

**A. Yu. Egorova, I. M. Zatsman, V. V. Kosarik, and V. A. Nuriev**

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** The paper describes an experiment focused on studying the instability of neural machine translation (NMT). In the course of a year, an array of text fragments in Russian was repeatedly translated into French. The time step was one month. To produce translations, the Google’s NMT system was used. The experiment helps reveal the instability of NMT, i. e., it shows that translations of a given text fragment tend to change with time but not always improving the quality. The generated translations were linguistically annotated, which led to uncovering several different types of the NMT instability. While annotating,

a previously designed classification of machine translation errors was employed. It was altered to meet the objectives of the experiment, the ultimate goal of which was to obtain a frequency distribution of different types of the NMT instability. Yet, the first step of the experiment limited itself to only categorizing the NMT instability, and it is this very step that the paper describes. As the empirical data, the experiment uses Russian–French annotations generated in a supracorpora database. Each annotation contains a fragment of the source Russian text, its translation into French, and the description of translation errors occurring there.

**Keywords:** machine translation; instability; translation monitoring; linguistic annotation; instability types

**DOI:** 10.14357/08696527200212

## Acknowledgments

The study was funded by the Russian Foundation for Basic Research, project 18-07-00192.

## References

1. Moorkens, J., S. Castilho, F. Gaspari, and S. Doherty, eds. 2018. *Machine translation: Technologies and applications. Vol. 1: Translation quality assessment*. Cham: Springer International Publishing. 287 p.
2. Scott, B. 2018. *Machine translation: Technologies and applications. Vol. 2: Translation, brains and the computer: A neurolinguistic solution to ambiguity and complexity in machine translation*. Cham: Springer International Publishing. 241 p.
3. Popović, M. 2018. Error classification and analysis for machine translation quality assessment. *Machine translation: Technologies and applications. Vol. 1: Translation quality assessment*. Cham: Springer International Publishing. 129–158.
4. Buntman, N. V., A. A. Goncharov, I. M. Zatsman, and V. A. Nuriev. 2018. Kolichestvennyy analiz rezul'tatov mashinnogo perevoda s ispol'zovaniem nadkorpusnykh baz dannykh [Using supracorpora databases for quantitative analysis of machine translations]. *Informatika i ee Primeneniya — Inform. Appl.* 12(4):96–105.
5. Nuriev, V., N. Buntman, and O. Inkova. 2018. Machine translation of Russian connectives into French: Errors and quality failures. *Informatika i ee Primeneniya — Inform. Appl.* 12(2):105–113.
6. Goncharov, A. A., N. V. Buntman, and V. A. Nuriev. 2019. Oshibki v mashinnom perevode: Problemy klassifikatsii [Machine translation errors: Problems of classification]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(3):92–103.
7. Rychikhin, A. K. 2019. O metodakh otsenki kachestva mashinnogo perevoda [On methods of machine translation quality assessment]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(4):106–118.
8. Cheng, Y., Z. Tu, F. Meng, J. Zhai, and Y. Liu. 2018. Towards robust neural machine translation. *56th Annual Meeting of the Association for Computational Linguistics Proceedings*. Melbourne: Association for Computational Linguistics. 1:1756–1766. Available at: <https://www.aclweb.org/anthology/P18-1163> (accessed June 1, 2020).

9. Natsional'nyy korpus russkogo jazyka [Russian National corpus]. Available at: <http://www.ruscorpora.ru/> (accessed February 25, 2020).
10. Nuriev, V. A. 2019. Arkhitektura sistemy neyronnogo mashinnogo perevoda [Architecture of a machine translation system]. *Informatika i ee Primeneniya — Inform. Appl.* 13(3):90–96.
11. Zaliznyak, A. A., I. M. Zatsman, and O. Yu. Inkova. 2017. Nadkorpusnaya baza dannykh konnektorov: postroenie sistemy terminov [Supracorpora database on connectives: Term system development]. *Informatika i ee Primeneniya — Inform. Appl.* 11(1):100–108.
12. Zatsman, I. M., and M. G. Kruzhkov. 2018. Nadkorpusnaya baza dannykh konnektorov: razvitiye sistemy terminov proektirovaniya [Supracorpora database of connectives: Design-oriented evolution of the term system]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(4):156–167.
13. Egorova, A. Yu., I. M. Zatsman, and O. S. Mamonova. 2019. Nadkorpusnye bazy dannykh v lingvisticheskikh proektakh [Supracorpora databases in linguistic projects]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(3):77–91.
14. Inkova-Manzotti, O. Yu. 2001. *Konnektory protivopostavleniya vo frantsuzskom i russkom jazykakh. Sopostavitel'noe issledovanie* [Connectors of opposition in French and Russian: A comparative study]. Moscow: Informelektro. 429 p.
15. Inkova, O. Yu. 2018. Nadkorpusnaya baza dannykh kak instrument formal'noy variativnosti konnektorov [Supracorpora database as an instrument of the study of the formal variability of connectives]. *Computer Linguistic and Intellectual Technologies: Conference (International) “Dialog” Proceedings*. Moscow. 17(24):240–253.

Received March 3, 2020

## Contributors

**Egorova Anna Yu.** (b. 1991) — junior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; ann.shurova@gmail.com

**Zatsman Igor M.** (b. 1952) — Doctor of Science in technology, Head of Department, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; izatsman@yandex.ru

**Kosarik Valery V.** (b. 1970) — scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44- 2 Vavilov Str., Moscow 119333, Russian Federation; valery\_lek@mail.ru

**Nuriev Vitaly A.** (b. 1980) — Candidate of Science (PhD) in philology, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; nurieff.v@gmail.com

## МЕТОДИЧЕСКИЙ ПОДХОД К ОЦЕНКЕ ВКЛАДА ИНФОРМАЦИОННЫХ СИСТЕМ В ЭФФЕКТИВНОСТЬ ОРГАНИЗАЦИОННЫХ СИСТЕМ\*

*А. А. Зацаринный<sup>1</sup>, Ю. С. Ионенков<sup>2</sup>*

**Аннотация:** Статья посвящена оценке вклада информационных систем (ИС) в эффективность соответствующих организационных систем (министерств, ведомств). Рассмотрен общий методический подход к оценке вклада ИС в эффективность организационных систем, учитывающий особенности, принципы и условия построения соответствующих организационных систем. Представлен перечень частных показателей эффективности для каждой из трех групп обобщенных показателей эффективности ИС (показатели рациональности организационной структуры ИС, показатели эффективности функционирования ИС и показатели организационно-технического уровня ИС). Предложена методика оценки вклада ИС в эффективность организационных систем, основанная на методе анализа иерархий.

**Ключевые слова:** организационная система; информационная система; эффективность; показатель; критерий; технология

**DOI:** 10.14357/08696527200213

### 1 Введение

Сегодня Россия находится на этапе важнейших перемен. Информационные технологии уверенно входят во все сферы деятельности общества и конкретного человека. Современные реалии таковы, что динамичное развитие информационных технологий, повсеместное усложнение бизнес-процессов, а также накопление значительных объемов данных приводят к объективному появлению такого понятия, как цифровая экономика. Известно, что старт процессам цифровой экономики официально был дан в Послании Президента РФ Федеральному Собранию 1 декабря 2016 г., после чего были утверждены Стратегия научно-технологического развития РФ, Программа цифровой экономики, План мероприятий, а также Национальный проект «Цифровая экономика». Ключевые цели нацпроекта — увеличение внутренних затрат на развитие цифровой экономики, создание устойчивой и безопасной информационно-телекоммуникационной инфраструктуры высокоскоростной передачи, обработки и хранения больших объемов данных, доступной для всех организаций.

\* Работа выполнена при частичной финансовой поддержке РФФИ (проект 18-29-03091).

<sup>1</sup>Федеральный исследовательский центр «Информатика и управление» Российской академии наук, AZatsarinny@ipiran.ru

<sup>2</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, UIonenkov@ipiran.ru

Национальный проект «Цифровая экономика» включает в себя 6 федеральных проектов. Один из них — «Цифровое государственное управление» — нацелен на предоставление гражданам и организациям доступа к приоритетным государственным услугам и сервисам в цифровом виде, создание национальной системы управления данными, развитие инфраструктуры электронного правительства, внедрение сквозных платформенных решений в государственное управление [1].

Из целей и задач этого проекта видна возрастающая роль ИС в создании информационно-телекоммуникационной инфраструктуры, внедрении в государственное управление современных информационных технологий, включая технологии искусственного интеллекта, больших данных, машинного обучения и др. При этом вопросы эффективности применения этих ИС должны рассматриваться с позиций вклада в повышение эффективности функционирования соответствующей организационной системы (министерства, ведомства, корпоративной структуры и т. п.) [2].

Вместе с тем необходимо отметить, что при реализации этих проектов отсутствует системность в комплексе выполняемых работ, в том числе в части подготовки комплекса нормативных документов, обучения и подготовки необходимых кадров, обоснования профиля необходимых информационных технологий, создания необходимой инфраструктуры, прежде всего в регионах, а также обеспечения информационной безопасности [3].

В связи с этим представляет интерес разработка методической базы для оценки влияния ИС на эффективность функционирования организационной системы, которая определяется системой целевых показателей.

В настоящей статье рассматриваются требования и общие подходы к формированию перечня показателей эффективности, учитывающих вклад ИС в общую эффективность организационных систем, а также предложения по методам расчета обобщенных и частных показателей эффективности ИС.

## **2 Общий подход к оценке вклада информационных систем в эффективность организационных систем**

Роль современных ИС заключается в повышении эффективности функционирования соответствующей организационной системы за счет улучшения качества и своевременности принятия управленческих решений, повышения уровня доступности информации, сокращения сроков обработки различных документов, расширения внутриведомственного и межведомственного взаимодействия, внедрения систем электронного документооборота.

При оценке вклада ИС в эффективность организационных систем необходимо учитывать особенности, условия и ограничения, характеризующие соответствующую организационную систему, ее функции, структуру и состав, связи между объектами системы и с взаимодействующими системами, потребности в автоматизации функций, применении современных информационно-аналитических

технологий. При этом целесообразно оценить прирост эффективности организационной системы за счет внедрения новых ИС либо модернизации существующих для различных вариантов реализации ИС.

Оценка вклада ИС в эффективность организационной системы представляет собой процесс, состоящий из нескольких этапов:

- (1) анализ особенностей, целей развития и задач организационной системы;
- (2) обоснование и выбор групп показателей (обобщенных показателей) эффективности ИС, учитывающих различные аспекты деятельности организационной системы и ее особенности;
- (3) определение частных показателей эффективности для каждой из групп;
- (4) разработка методов расчета частных и обобщенных показателей эффективности, а также интегрального показателя эффективности системы;
- (5) оценка влияния ИС на эффективность функционирования организационной системы, изменение целевых показателей системы.

Представляется целесообразным выделить несколько групп показателей, характеризующих различные аспекты функционирования ИС в организационной системе, в частности: показатели рациональности организационной структуры ИС; показатели эффективности функционирования ИС; показатели организационно-технического уровня ИС.

Возможный вариант с тремя группами показателей эффективности сведен в таблицу.

Необходимо отметить, что некоторые из указанных показателей могут быть комплексными, т. е. включать несколько частных показателей, что увеличивает число уровней иерархии показателей в системе. Например, уровень экономичности структуры управления зависит от удельного веса численности руководителей в общей численности персонала, удельного веса затрат на содержание руководителей, а также соотношения управленческих расходов и объема выпускаемой продукции. Уровень надежности управления характеризуется степенью контроля функций и операций, удельным временем ожидания документа либо решения и уровнем надежности принятых решений.

Порядок выбора показателей эффективности ИС применительно к различным типам систем и стадиям их жизненного цикла достаточно подробно рассмотрен в [4].

Представляется, что предлагаемый набор показателей достаточно универсален. Естественно, что для каждой конкретной организационной системы этот набор показателей должен уточняться с учетом специфики ее построения и функционирования.

В качестве источников информации для расчета показателей могут использоваться: принятая структура организационной системы; положение о соответствующей организационной системе; штатное расписание; нормативы и стандарты, принятые в системе; различные виды отчетности; показатели деятельности за соответствующие периоды и т. п.

Показатели, характеризующие влияние ИС на эффективность организационной системы

Обобщенные показатели	Частные показатели
Показатели рациональности организационной структуры ИС	<ul style="list-style-type: none"> <li>– Соответствие числа и организационной структуры ИС задачам организационной системы и перспективам ее развития</li> <li>– Соответствие численности обслуживающего персонала объему задач системы</li> <li>– Уровень использования современных инфокоммуникационных технологий в ИС</li> <li>– Полнота обеспечения ИС требуемой информацией</li> <li>– Стоимостные показатели действующих и внедряемых ИС</li> <li>– Экономичность структуры управления</li> </ul>
Показатели эффективности функционирования ИС	<ul style="list-style-type: none"> <li>– Уровень автоматизации функций (процессов) организационной системы</li> <li>– Уровень реализации задач (услуг, проектов) организационной системы с помощью ИС</li> <li>– Сокращение времени на принятие управленческих решений</li> <li>– Сокращение времени на сбор, обработку и доступ к информации</li> <li>– Повышение доступности информации</li> <li>– Повышение надежности управления</li> <li>– Сокращение времени представления отчетности и трудозатрат на ее подготовку</li> <li>– Прямые доходы либо экономический эффект, получаемые при внедрении системы</li> </ul>
Показатели организационно-технического уровня ИС	<ul style="list-style-type: none"> <li>– Наличие и укомплектованность подразделений эксплуатации ИС</li> <li>– Уровень организации рабочих мест</li> <li>– Уровень профессиональной подготовки персонала</li> <li>– Уровень готовности сотрудников организации к внедрению и применению новых технологий</li> <li>– Способность сотрудников к обучению</li> </ul>

### 3 Предложения по методике оценки вклада информационных систем в эффективность организационных систем

Методический аппарат для оценки вклада ИС в эффективность организационных систем должен, с одной стороны, позволять производить оценку эффективности с достаточной точностью, а с другой — быть пригодным для практического использования и понятным для лиц, производящих такую оценку. Исходя из этого, представляется целесообразным использовать методы, обеспечивающие получение необходимых количественных оценок в условиях неопределенности и нечеткости исходной информации.

Одним из наиболее апробированных и широко используемых в практике методов такого рода является метод анализа иерархий, разработанный Т. Саати [5, 6]. Для этого метода характерны простота, апробированность, возможность ра-

боты с большой размерностью данных. Кроме того, он ориентирован на нечеткие оценки, что более просто для экспертов, чем точные количественные оценки. Основное содержание данного метода и его применение для оценки эффективности информационно-телеkomмуникационных систем достаточно подробно представлены в [7].

С учетом этого предлагается методика оценки вклада ИС в эффективность организационных систем, включающая следующие этапы (шаги).

**Шаг 1.** Выбор и обоснование номенклатуры показателей эффективности.

Для каждой из трех групп обобщенных показателей эффективности (показатели рациональности организационной структуры ИС, показатели эффективности функционирования ИС, показатели организационно-технического уровня ИС) предлагается несколько частных показателей с учетом специфики организационной системы.

Частные показатели эффективности определяются следующим образом:

- по типовым формулам (например, коэффициент готовности и др.);
- путем нормирования показателей относительно их максимальных и минимальных значений;
- как отношение реального значения показателя к требуемому либо максимально (минимально) возможному;
- экспертным путем по соответствующей шкале и приведением к значениям от 0 до 1.

**Шаг 2.** Выбор метода расчета обобщенного показателя эффективности.

Для расчета обобщенного показателя эффективности предлагается использовать метод анализа иерархий. На этом шаге осуществляется построение трехуровневой иерархической структуры:

- (1) уровень цели — оценка вклада ИС в эффективность организационной системы;
- (2) первый уровень критериев: рациональность организационной структуры ИС; эффективность функционирования ИС; организационно-технический уровень ИС;
- (3) второй уровень критериев — показатели по каждой группе критериев первого уровня, представленных выше.

Общий вид иерархической структуры представлен на рисунке.

**Шаг 3.** Сбор исходных данных и расчет частных показателей эффективности 2-го уровня. Определяется сбор данных, характеризующих представленные выше три группы обобщенных показателей, и их обработка и нормирование экспертами. Ряд показателей определяется экспертным путем.

**Шаг 4.** Сравнительный анализ элементов процесса. Для этого с привлечением экспертов производится заполнение обратно-симметричных матриц парных сравнений для критериев первого и второго уровня.



**Шаг 5.** Определение значений весовых коэффициентов для критериев первого и второго уровня из заполненных матриц.

**Шаг 6.** Оценка степени согласованности экспертов.

**Шаг 7.** Определение обобщенных показателей эффективности по критериям первого уровня, а также интегрального показателя эффективности организационной системы.

**Шаг 8.** Оценка полученных результатов и выработка рекомендаций.

Представленная методика применима к различным типам ИС и обеспечивает возможность оценки их вклада в эффективность соответствующих организационных систем с учетом их характеристик, структуры и т. д.

## 4 Заключение

Информационное обеспечение управлеченческих решений приобретает все большее значение. Вследствие интенсивного развития и внедрения информационных технологий и ИС проблематика оценки их вклада в эффективность соответствующих организационных систем становится все более актуальной.

В статье рассмотрены требования и общие подходы к формированию перечня показателей, характеризующих влияние ИС на эффективность организационных систем, а также предложена методика расчета обобщенных и частных показателей эффективности ИС, позволяющая оценить вклад ИС в эффективность соответствующей организационной системы, что имеет большое практическое значение в совершенствовании механизма управления соответствующими системами.

## Литература

1. Цифровая экономика: Национальный проект Российской Федерации. <https://strategy24.ru/rf/management/projects/natsional-nyy-proyekt-tsifrova-ekonomika>.
2. Зацаринный А. А., Шабанов А. П. Технология информационной поддержки деятельности организационных систем на основе ситуационных центров. — М.: ТОРУС ПРЕСС, 2015. 232 с.
3. Зацаринный А. А., Киселев Э. В., Козлов С. В., Колин К. К. Информационное пространство цифровой экономики России. Концептуальные основы и проблемы формирования / Под общ. ред. А. А. Зацаринного. — М.: ФИЦ ИУ РАН, 2018. 236 с.
4. Зацаринный А. А., Ионенков Ю. С. Некоторые методические аспекты выбора показателей эффективности информационных систем // Системы высокой доступности, 2019. Т. 15. № 4. С. 19–26.
5. Саати Т. Принятие решений. Метод анализа иерархий. — М.: Радио и связь, 1993. 278 с.
6. Саати Т. Принятие решений при зависимостях и обратных связях. Аналитические сети. — М.: ЛКИ, 2008. 360 с.
7. Зацаринный А. А., Ионенков Ю. С. К вопросу оценки эффективности автоматизированных систем с использованием метода анализа иерархий // Системы и средства информатики, 2015. Т. 25. № 3. С. 161–178.

Поступила в редакцию 15.03.20

---

## METHODOLOGICAL APPROACH TO ASSESSING THE CONTRIBUTION OF INFORMATION SYSTEMS TO THE EFFECTIVENESS OF ORGANIZATIONAL SYSTEMS

*A. A. Zatsarinny<sup>1</sup> and Yu. S. Ionenkov<sup>2</sup>*

<sup>1</sup>Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

<sup>2</sup>Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** The article is devoted to evaluating the contribution of information systems (IS) to the effectiveness of relevant organizational systems (ministries and departments). A general methodological approach to assessing the contribution of IS to the effectiveness of organizational systems is considered taking into account the features, principles, and conditions for building appropriate organizational systems. The authors present the list of performance indicators for each of the three groups of generalized indicators of the effectiveness of the IS (the index of rationality of organizational structures, the performance indicators for IS, and the indicators of organizational and technical level). A method for evaluating

the contribution of IS to the effectiveness of organizational systems based on the hierarchy analysis method is proposed.

**Keywords:** organizational system; information system; efficiency; indicator; criterion; technology

**DOI:** 10.14357/08696527200213

## Acknowledgments

The paper was partially supported by the Russian Foundation for Basic Research (project 18-29-03091).

## References

1. Tsifrovaya ekonomika: Natsional'nyy proekt Rossiyskoy Federatsii [Digital Economy: National Project of the Russian Federation]. Available at: <https://strategy24.ru/rf/management/projects/natsional-nyy-proyekt-tsifrova-ekonomika> (accessed February 28, 2020).
2. Zatsarinnyy, A. A., and A. P. Shabanov. 2015. *Tekhnologiya informatsionnoy podderzhki deyatel'nosti organizatsionnykh sistem na osnove situatsionnykh tsentrov* [A technology of information support of organizational systems activity based on situational centers]. Moscow: TORUS PRESS. 232 p.
3. Zatsarinnyy, A. A., E. V. Kiselev, S. V. Kozlov, and K. K. Kolin. 2018. *Informatsionnoe prostranstvo tsifrovoy ekonomiki Rossii. Kontseptual'nye osnovy i problemy formirovaniya* [Information space of digital economy. Conceptual framework and problems of formation]. Moscow: FRC CSC RAS. 236 p.
4. Zatsarinnyy, A. A., and Yu. S. Ionenkov. 2019. Nekotorye metodicheskie aspekty vybora pokazateley effektivnosti informatsionnykh system [Some methodological of the choice of performance indicators of information systems]. *Sistemy Vysokoy Dostupnosti* [Highly Available Systems] 15(4):19–26.
5. Saati, T. 1993. *Prinyatie resheniy. Metod analiza ierarkhiy* [Decision making. Analytic hierarchy process]. Moscow: Radio i svyaz'. 278 p.
6. Saati, T. 2008. *Prinyatie resheniy pri zavisimostyakh i obratnykh svyazyakh. Analiticheskie seti* [Decision making in dependences and feedbacks. Analytic networks]. Moscow: LKI. 360 p.
7. Zatsarinnyy, A. A., and Yu. S. Ionenkov. 2015. K voprosu otsenki effektivnosti avtomatizirovannykh sistem s ispol'zovaniem metoda analiza ierarkhiy [Regarding automated systems efficiency evaluation using analytic hierarchy process]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 25 (3):161–178.

Received March 15, 2020

## Contributors

**Zatsarinny Alexander A.** (b. 1951) — Doctor of Science in technology, professor, Deputy Director, Federal Research Center “Computer Science and Control” of

the Russian Academy of Sciences (FRC CSC RAS); principal scientist, Institute of Informatics Problems, FRC CSC RAS; 44-2 Vavilov Str., Moscow 119333, Russian Federation; AZatsarinny@ipiran.ru

**Ionenkov Yurij S.** (b. 1956) — Candidate of Science (PhD) in technology, senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; uionenkov@ipiran.ru

# ИНФОРМАЦИОННО-АНАЛИТИЧЕСКИЙ СИТУАЦИОННЫЙ ЦЕНТР СТРАТЕГИЧЕСКОГО ПЛАНРИРОВАНИЯ В ОБЛАСТИ ОБЕСПЕЧЕНИЯ НАЦИОНАЛЬНОЙ БЕЗОПАСНОСТИ: ПЕРСПЕКТИВНЫЕ АРХИТЕКТУРНЫЕ РЕШЕНИЯ\*

*A. V. Bosov<sup>1</sup>, A. P. Suchkov<sup>2</sup>*

**Аннотация:** Предложен вариант перспективной архитектуры информационно-аналитического ситуационного центра (ИАСЦ) для управления процессами стратегического планирования в области национальной безопасности. Решение базируется на ранее опубликованных концептуальных подходах и анализе технических, технологических, организационных и правовых факторов, оказывающих принципиальное влияние на проект ИАСЦ. Предлагаемое архитектурное решение обеспечивает как учет выделенных факторов влияния, так и разработанную ранее функциональную схему ИАСЦ, базирующуюся на принципах ситуационного анализа в процессах стратегического планирования в области национальной безопасности. Техническую основу решения обеспечивает группа современных технологий, составляющих элементы построения программно определяемых центров обработки данных (ЦОД). Подлежащую учету специфику в рамках такого ЦОД составляет предметная область ИАСЦ, предполагающая применение специальных методов ситуационного анализа, широкий охват разнородных информационных источников, гибкое взаимодействие с ведомственными системами, в том числе обрабатывающими информацию ограниченного доступа.

**Ключевые слова:** стратегическое планирование; ситуационный центр; архитектурные решения

**DOI:** 10.14357/08696527200214

## 1 Введение

В соответствии с основополагающим актом [1] стратегическое планирование — это деятельность участников стратегического планирования по целеполаганию, прогнозированию, планированию и программированию социально-экономического развития РФ, субъектов РФ и муниципальных образований, отраслей экономики и сфер государственного и муниципального управления, обеспечения

---

\*Работа выполнена при поддержке РФФИ (проект 18-00-00297).

<sup>1</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, AVBosov@ipiran.ru

<sup>2</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, ASuchkov@ipiran.ru

национальной безопасности РФ, направленная на решение задач устойчивого социально-экономического развития РФ, субъектов РФ и муниципальных образований и обеспечение национальной безопасности РФ. Как показано в [2–4], эффективной формой управления процессами подготовки, реализации и сопровождения документов стратегического планирования, а также координации деятельности в области обеспечения национальной безопасности могут стать ИАСЦ. Обсуждаемый проект такого специализированного под цели стратегического планирования ИАСЦ обеспечивает создание национальной системы мониторинга, прогнозирования, выявления, анализа и оценки угроз национальной безопасности. В соответствии с этим ИАСЦ позиционируется как составная часть системы управления, обеспечивающей информационно-аналитическую поддержку процессов принятия решений на основе автоматизации функций мониторинга и ситуационного анализа обстановки в контролируемом информационном пространстве, включая сбор, обобщение, аналитическую обработку, хранение, передачу, визуализацию и защиту информации.

Перечисленные положения обеспечивают концептуальное и функционально-алгоритмическое наполнение проекта ИАСЦ, а также частично и информационное. В данной статье предполагается обсудить последующие шаги в направлении реализации проекта и выработать подходящее поставленным задачам архитектурное решение, предложив модель программно-аппаратной системы — платформы ИАСЦ. Окончательная реализация ИАСЦ в дальнейшем сведется к конкретизации представленной модели вариантами технических решений.

Предлагаемое архитектурное решение базируется не только на привлечении современных технических концепций и технологий, но и на детальном анализе особенностей — факторов влияния [5], определяющих наиболее важные аспекты создания и последующего функционирования ИАСЦ.

## **2 Варианты функциональных моделей информационно-аналитического ситуационного центра**

В качестве потенциально возможных моделей функционирования ИАСЦ рассматривались:

- модель децентрализации по признаку категории информации (модель 1);
- модель полной централизации (модель 2);
- модель динамического ЦОД (модель 3).

Модель 1 характерна для многих действующих систем автоматизации в федеральных органах исполнительной власти (ФОИВ) и заключается в выделении в отдельные автономные контуры подсистем, обеспечивающих работу с информацией одной категории защищенности или даже одной тематической направленности. Простой пример: контур «Отдел» для работы со сведениями, составляющими гостайну, контур «Производство» для работы с внутренними

ведомственными документами, контур «Интернет» для работы с открытыми источниками. В лучшем случае между такими контурами поддерживается односторонний информационный обмен, при котором данные передаются в направлении сегмента с более высоким уровнем защищенности обрабатываемой информации. Модель 1 предполагает использование на высоких уровнях «тяжелых» средств защиты информации, аттестованных и сертифицированных уполномоченными государственными органами.

Основные недостатки данной модели — это избыточные накладные расходы, связанные с наличием нескольких изолированных контуров; чрезвычайно высокие накладные расходы на обеспечение информационной безопасности, а также сложность процессов, связанных с расширением функциональных возможностей.

Модель 2 полной централизации основана на популярной, проверенной многими годами применения и получившей в последние годы новые технологические импульсы парадигме единого централизованного хранилища. Системы на основе этой модели создавались не одно десятилетие и продолжают создаваться. В настоящий момент успешно функционируют как относительно старые системы, так и новые, созданные совсем недавно с использованием самых современных технологий. Наиболее известные крупные решения данного класса — это государственные федеральные информационные системы и хранилища. Такие, можно сказать, образцовые системы имеются, например, в Федеральной налоговой службе России (АИС «Налог-3» [6]), в Пенсионном фонде Российской Федерации (АИС «ПФР-2» [7]), есть похожие решения и у других ФОИВ, и у Центрального банка Российской Федерации.

И ранее созданные, и современные решения такого класса базируются на проприетарных технологиях мировых лидеров ИТ-индустрии, обеспечивающих высокую производительность для всех типов приложений баз данных, включая системы оперативной обработки транзакций (OLTP — online transaction processing), хранилища данных, системы аналитической обработки (в том числе поддерживающие технологию BigData). Таким образом, основными составными элементами в рамках данной модели функционирования выступают комплексы из высокопроизводительного оборудования и программного обеспечения (Teradata Enterprise Data Warehouse, Oracle Exadata Database Machine, IBM Z-series), технологии больших данных, глобально распределенные базы данных. Все эти практики очень хорошо себя зарекомендовали в течение многолетней эксплуатации.

Применительно к модели 2 обязательно нужно выделить еще один основополагающий элемент — единую модель данных, и даже более того — единый бизнес-процесс. Понятно, что эффективное размещение большого объема данных в центральном хранилище возможно лишь в том случае, когда данные типизированы и структурированы, а в центральном хранилище реализована функциональная поддержка соответствующих типов и структур. Более того, для эффективной работы с данными центрального хранилища необходимо, чтобы все требуемые процедуры обработки быть детально прописаны, надлежащим

образом автоматизированы, реализованы, оптимизированы и отложены. Таким образом, применение модели полной централизации в качестве модели функционирования целесообразно в тех случаях, когда еще до начала разработки системы известны как модель данных, так и бизнес-логика (по меньшей мере предметная область позволяет с самого начала определить и зафиксировать и модель данных, и бизнес-логику). Именно поэтому решения на основе модели полной централизации успешно применяются в хорошо изученных и формализованных областях, таких как таможня, финансы, налоги и т. п. (например, National Information Exchange Model (NIEM) — национальная система обмена данными США).

Главный недостаток систем, основанных на модели полной централизации, заключается в том, что, после того как такая система создана, инкорпорация в нее данных, не отвечающих действующей модели данных, представляет собой весьма значительную проблему, преодоление которой требует полной ревизии большого числа ранее принятых и реализованных решений. Таким системам можно, конечно, придать минимальную гибкость, но наполнить ее реальными инструментами интеграции разнородных информационных ресурсов не представляется возможным. И в этом видится достаточное основание, чтобы отказаться от модели 2 в качестве архитектурного решения для ИАСЦ. Также необходимо отметить, что для реализации модели 2 требуются огромные первоначальные финансовые вложения в создание централизованной платформы и инфраструктуры для ее функционирования. Неслучайно успешное внедрение и эксплуатация таких систем в первую очередь осуществляется в ведомствах, контролирующих огромные финансовые потоки.

Отрицательные и положительные стороны централизованного подхода можно было бы обсуждать и далее, например в рамках дискуссии о возможности реализации централизованного хранилища всей тематически значимой с точки зрения стратегического планирования информации. Однако указанных выше проблем достаточно, чтобы отказаться от рассмотрения данной модели в качестве перспективной.

Наконец, модель 3 предполагает использование динамического ЦОД. Главный концепт модели — изначально заложенные в нее возможности эволюционного изменения ЦОД в процессе функционирования. Модель предполагает последовательное создание ИАСЦ, исходя из текущей инфраструктуры с использованием современных стандартных элементов вычислительной инфраструктуры. В процессе развития идет постепенное наращивание аппаратных и программных мощностей ЦОД, внедрение новых технологий происходит параллельно с поддержкой уже существующих и т. д. Можно показать, что данный подход является оптимальным с точки зрения минимизации затрат на интеграцию разнородных систем [8].

Технологическим фундаментом, который позволяет создать такой динамический ЦОД и обеспечить его дальнейшее функционирование, с одной стороны, и модификацию, с другой, служит полнофункциональная система виртуализации. Главная задача системы виртуализации в контексте обсуждаемой

модели — предоставить в рамках единой аппаратно-программной платформы ЦОД возможности по созданию и динамической модификации обособленных, программно-изолированных сегментов, в каждом из которых поддерживается функционирование определенного набора сервисов и работа определенных групп пользователей.

Модель 3 не фиксирует жестко состав и содержимое входящих в ЦОД сегментов, но допускает возможность динамического изменения как содержимого сегментов, так и их общего числа. Существование каждого конкретного сегмента (в определенный период времени) должно быть функционально обусловлено, т. е. связано с наличием конкретной задачи (группы задач), требующей выделения ресурсов для ее решения, и наличием заинтересованной группы пользователей. Нормальным представляется удаление ненужных сегментов, т. е. сегментов, востребованных ранее для решения задач, потерявших актуальность к настоящему моменту времени. Удаление сегментов сопровождается, разумеется, освобождением ресурсов. Напротив, для решения вновь появившихся задач создаются новые сегменты.

Именно модель 3 выглядит перспективной моделью функционирования ИАСЦ. Далее эта модель детализируется и обосновывается во всех принципиальных аспектах.

### **3 Перспективная модель функционирования информационно-аналитического ситуационного центра**

В состав перспективной модели функционирования ИАСЦ в первую очередь входит динамический ЦОД, который служит центральным узлом (центральным элементом) модели и представляет собой общий пул вычислительных средств, гибко распределемых для решения всего множества текущих задач. Динамический ЦОД включает набор сегментов, каждый из которых предназначен для решения конкретной функциональной задачи или группы таких задач и гарантированно (легитимно) изолирован от остальных сегментов. Вторая часть общей модели — ряд частных моделей и обеспечивающих решений, а именно:

- модель информационной безопасности;
- модель взаимодействия субъектов стратегического планирования;
- модель сопровождения и эксплуатации;
- комплекс технических решений по общесистемной программной платформе, аппаратной платформе, коммуникационному оборудованию, прикладным программным средствам.

Логическими структурными элементами модели, составляющими динамический ЦОД, выступают сегменты. Характеристическими признаками, определяющими сегмент, являются функциональность, состав пользователей, категория обрабатываемой информации, а также его изолированность (обособленность) от

других сегментов. В сегменте циркулирует (обрабатывается) информация одной категории защищенности. Наконец, изолированность (обособленность) сегмента означает, что

- доступ пользователей из данного сегмента к другим сегментам (к информационным, программным и аппаратным ресурсам других сегментов) исключен;
- обмен информацией между данным сегментом и другими сегментами, а также внешними системами может осуществляться только путем контролируемого обмена данными посредством специальных программных или программно-аппаратных шлюзов.

Можно предложить ряд функционально значимых сегментов, которые должны создаваться в составе динамического ЦОД в первую очередь.

**Сегменты обмена документами федерального/муниципального уровня.**

Задача данных сегментов — автоматизировать документооборот, возникающий при разработке документов стратегического планирования на соответствующих уровнях субъектов. Основная ценность сегмента — очень простой и оперативный обмен и совместная разработка документов всех без исключения субъектов стратегического планирования.

**Сегмент внутренних банков** ИАСЦ предназначен для размещения внутренних информационных ресурсов ИАСЦ и аналитических систем для работы с этими ресурсами. Информационные ресурсы формируются собственными силами ключевого субъекта ИАСЦ, координирующего, согласно [1], всю государственную деятельность в области стратегического планирования, — Совета национальной безопасности РФ — и не разделяются с другими субъектами. Доступ в сегмент обеспечивается только сотрудникам аппарата Совбеза.

**Сегмент банков субъектов** должен создаваться в интересах всех субъектов и наполняться типизированными информационными ресурсами — *тематическими банками субъектов*. Каждый тематический банк представляет собой, по существу, витрину данных некоторого собственного информационного ресурса субъекта, содержащего тематические данные. Пользователи сегмента, во-первых, имеют возможность выполнять поиск интересующих их сведений в массиве данных, размещенных в тематических банках. Во-вторых, пользователи получат возможность обратиться к информационным ресурсам субъекта за детализацией сведений, найденных ими в эталонных банках.

**Резервный сегмент.** Резервный сегмент дает возможность сотрудникам аппарата Совбеза быстро размещать в ИАСЦ специальные информационные ресурсы на временной основе, для чего в сегменте должна быть развернута вся необходимая программно-техническая инфраструктура.

Следует еще раз явно сформулировать положение, согласно которому в каждом изолированном сегменте ЦОД циркулирует информация одной категории защищенности и/или тематической направленности, а вопросы разграничения доступа пользователей к этой информации и выбора механизмов его реализации решаются исходя из функциональных задач сегмента и, следовательно, предпо-

лагаемого состава его пользователей. При этом внутри большинства сегментов от разграничения доступа можно отказаться вообще, обеспечив защиту от несанкционированного доступа к самим сегментам. Соответственно, в разных сегментах могут использоваться разные средства защиты. В этой связи еще раз подчеркнем важную роль системы виртуализации как главного инструмента реализации перспективной модели функционирования ИАСЦ, который обеспечивает изоляцию сегментов легитимными средствами.

## 4 Обеспечивающие модели

Фундаментальный элемент рассмотренной модели функционирования ИАСЦ — модель взаимодействия субъектов стратегического планирования. Под взаимодействием субъектов, естественно, понимается информационное взаимодействие, сама модель описывает информационный обмен между субъектами, принадлежащими им информационными ресурсами и ресурсами ИАСЦ. Цель такого обмена — наполнение банков ИАСЦ информацией, которой обладают субъекты стратегического планирования. Можно предполагать, что она размещена в собственных тематически значимых для стратегического планирования информационных ресурсах, которые далее также именуются *внешними информационными ресурсами*. Предлагаемая модель взаимодействия относит каждый внешний информационный ресурс к одному из следующих трех типов: *делегируемый ресурс*, *представляемый ресурс* и *закрытый ресурс* — и определяет типы участия: *делегирование, представление и информирование* соответственно.

Делегируемым ресурсом считается тот внешний ресурс, информация которого передается субъектом в ИАСЦ в полном объеме. Делегирование ресурса осуществляется путем формирования его информационной копии в сегменте внутренних банков ИАСЦ. Первичное наполнение и последующее поддержание такого внутреннего банка данных в актуальном состоянии предполагается осуществлять в дальнейшем в автоматическом режиме, направление информации при этом одностороннее: от субъекта к ИАСЦ.

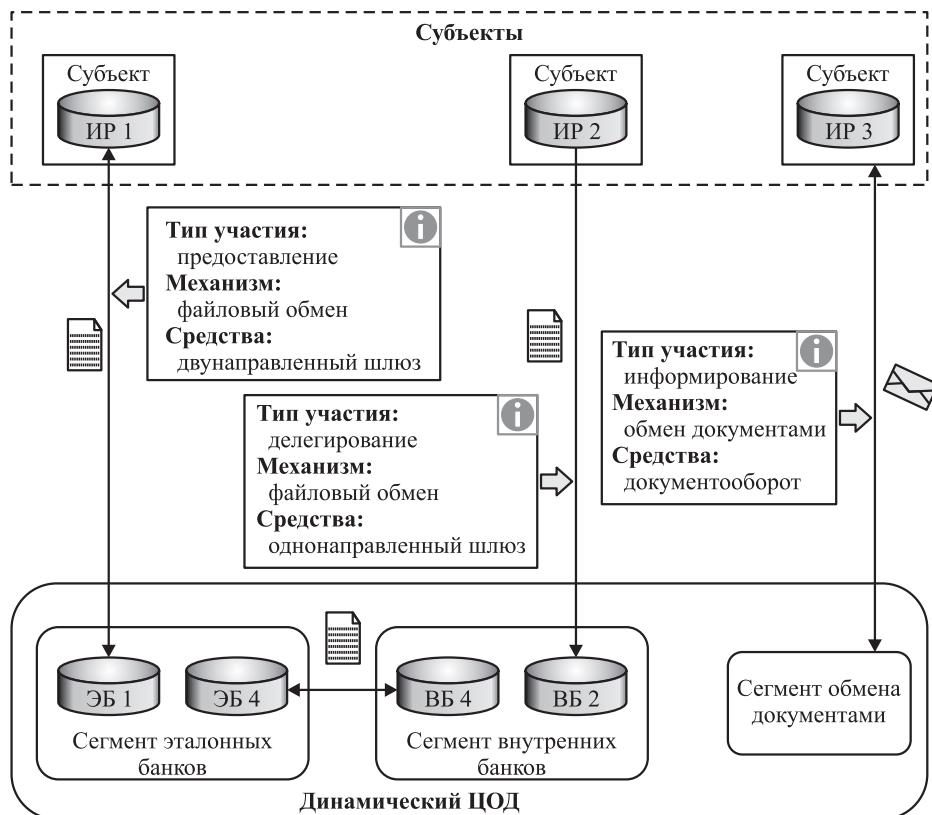
Внешний ресурс становится представляемым ресурсом в том случае, если из всего объема содержащейся в нем информации в ИАСЦ передается лишь часть, составляющая витрину для всего объема данных ресурса. Такой тип участия используется, когда нецелесообразно дублирование внешнего ресурса в составе ИАСЦ, например по причине излишней детализации и чрезмерных объемов, или субъект не готов полностью передать свой информационный ресурс, например в связи с ограничениями на распространение сведений определенного типа. Представление ресурса осуществляется путем формирования на его основе соответствующего *эталонного банка*, т. е., по существу, специальной витрины данных внешнего ресурса в сегменте эталонных банков ИАСЦ. Первичное наполнение и последующее поддержание эталонного банка (витрины данных) в актуальном состоянии осуществляется в автоматическом режиме. Функционал сегмента эталонных банков должен обеспечить наличие для пользователей

возможности запросить и получить из оригинального информационного ресурса дополнительные уточняющие сведения по учетной информации, представленной в витрине данных. Поэтому для данного типа участия требуется реализация двунаправленного информационного обмена.

Наконец, внешний ресурс, информационное взаимодействие с которым в автоматическом / автоматизированном режиме по тем или иным причинам невозможно, является закрытым ресурсом. Получение сведений из такого внешнего ресурса возможно путем обмена сообщениями; соответствующий функционал предоставляется сегментом обмена документами соответствующего уровня.

В рамках этой модели один субъект может обладать сразу несколькими внешними ресурсами, при этом разные ресурсы могут относиться к разным типам по классификации модели взаимодействия.

Представленная на рис. 1 схема перспективной модели взаимодействия субъектов иллюстрирует сказанное. На схеме изображены три субъекта, по одному

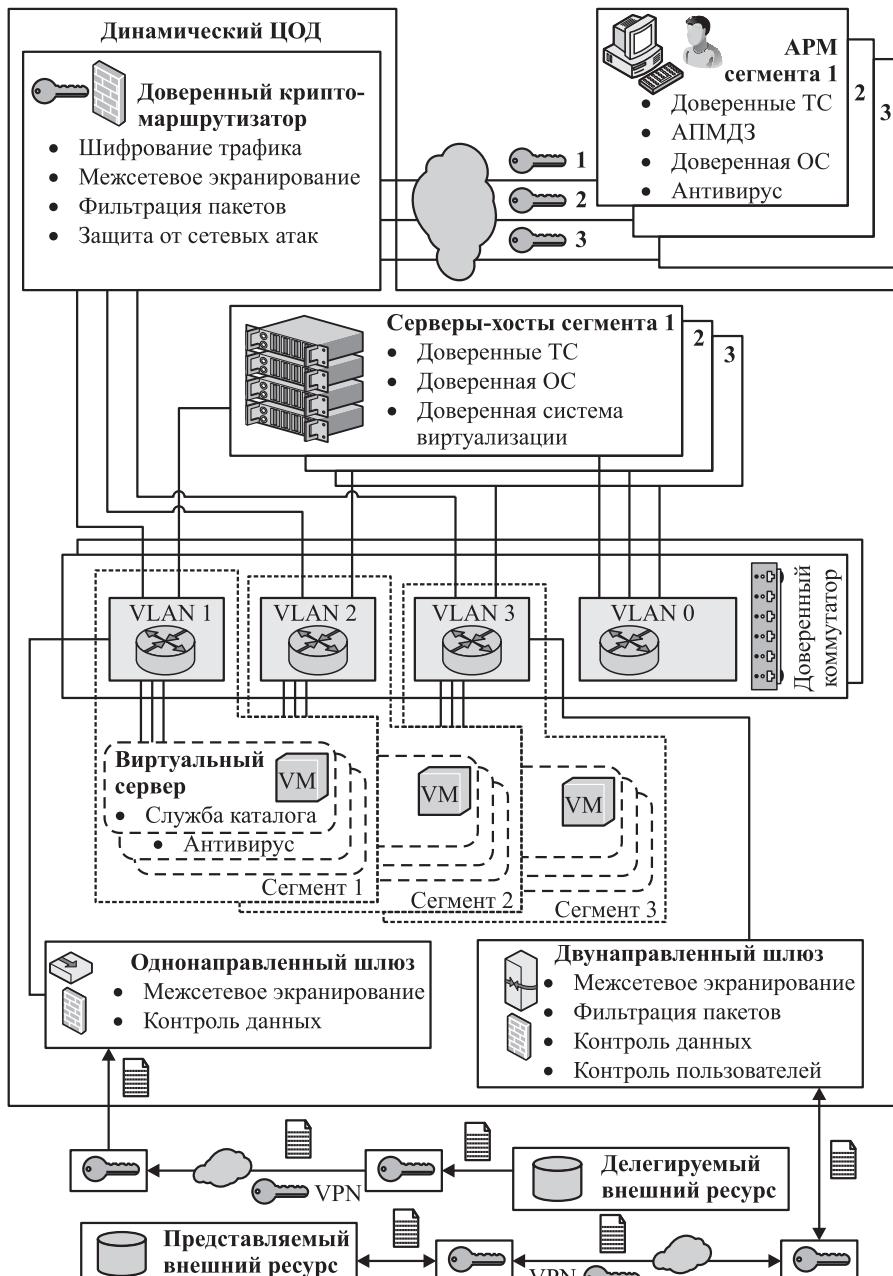


**Рис. 1** Модель взаимодействия субъектов ИАСЦ

внешнему информационному ресурсу каждого (ИР 1, ИР 2 и ИР 3), динамический ЦОД и три его сегмента. ИР 1 выступает представляемым ресурсом, которому соответствует эталонный банк (витрина данных) ЭБ 1 в сегменте эталонных банков (тип участия — представление). ИР 2 служит делегируемым ресурсом, которому соответствует внутренний банк ВБ 2 в сегменте внутренних банков (тип участия — делегирование). ИР 3 является закрытым ресурсом (тип участия — информирование).

Следующий важный элемент обсуждаемой модели функционирования ИАСЦ — *модель информационной безопасности*. Ключевой тезис по данной модели — необходимость отказа от *принципа тотальной защиты*, который часто и не всегда уместно применяется ко всем системам в защищенном исполнении, к которым относится и ИАСЦ. Этот принцип иллюстрируют следующие типовые решения. Во-первых, объектами защиты являются все минимальные «атомарные» объекты: отдельный файл файловой системы, отдельная запись реестра операционной системы, отдельная таблица базы данных и т. д. Во-вторых, защита обеспечивается на каждом вычислительном узле, т. е. на каждом сервере и автоматизированном рабочем месте (АРМ). В-третьих, весь сетевой трафик подлежит программному шифрованию. Применение такой тотальной защиты приводит к очень значительным накладным расходам, связанным с обеспечением информационной безопасности. Действительно, выполнение любой операции обработки данных на любом вычислительном средстве сопровождается действиями по антивирусной проверке данных, проверке прав и полномочий пользователя, инициировавшего операцию, регистрации операции в протоколе безопасности и т. д. Кроме того, передача данных между вычислительными средствами сопровождается как минимум однократно выполняемыми процедурами шифрования-десифрования. В результате заметное время и большой объем имеющихся вычислительных ресурсов тратятся на решение хотя и важной, но нецелевой задачи, что существенно снижает общую производительность системы. Масштаб ИАСЦ и информационное многообразие его ресурсов, объясняемое назначением системы, не допускают централизованной полнофункциональной модели информационной безопасности и предполагают отказ от принципа тотальной защиты.

Перспективной для ИАСЦ представляется *распределенная избирательная модель информационной безопасности*. Эта модель, схематически изображенная на рис. 2, предполагает отказ от принципа тотальной защиты в пользу *принципа достаточной защиты*. Можно выделить основные факторы, которые позволяют существенно изменить и облегчить традиционные тотальные модели обеспечения информационной безопасности. Во-первых, аппаратная платформа ИАСЦ изначально должна создаваться на основе доверенных технических средств (ТС). Во-вторых, доверенными должны быть основные общесистемные программные средства, формирующие программную платформу, в частности операционные системы (ОС) серверов-хостов и АРМ. Отдельно здесь следует выделить доверенную систему виртуализации, которая играет важнейшую роль в формировании программной платформы ИАСЦ вообще и динамического ЦОД в частности.



**Рис. 2** Модель информационной безопасности ИАСЦ (АПМДЗ — аппаратно-программный модуль доверенной загрузки)

Наконец, ключевым фактором является архитектура динамического ЦОД, который технологически представляет собой единый пул вычислительных средств, а с функциональной точки зрения состоит из набора сегментов, предназначенных для решения конкретной целевой задачи или группы таких задач. Вычислительные средства сегмента ЦОД представляют собой совокупность развернутых на аппаратных серверах-хостах виртуальных серверов, которые связаны посредством виртуальной локальной сети (VLAN, Virtual Local Area Network). Таким образом, необходимое свойство сегмента — *изолированность* — может быть гарантировано, т. е. обеспечено легитимными средствами (при использовании доверенных средств).

Каждый сегмент создается для решения конкретной функциональной задачи или группы родственных функциональных задач. Функционал сегмента определяет состав его пользователей: возможность работы в сегменте имеют только предопределенные доверенные пользователи, заинтересованные в решении функциональных задач сегмента и допущенные к решению этих задач. Это позволяет считать, что все пользователи сегмента имеют равные права доступа к информационным / программным ресурсам сегмента, и отказаться от использования внутри сегментов дополнительных ресурсоемких легитимных средств гарантированного дискреционного контроля. Задача идентификации / аутентификации пользователей сегмента решается системой виртуализации и / или службой каталога сегмента.

Применяемый в модели принцип достаточности означает следующее. Объектами защиты являются сегменты динамического ЦОД, т. е. защите подлежит вся информация каждого из сегментов в целом. Защита сегмента реализуется на его внешнем периметре путем контроля входных и выходных информационных потоков. Действительно, поскольку каждый сегмент изолирован, а обмен информацией между данным сегментом и внешней по отношению к этому сегменту средой осуществляется по наперечет известным каналам, то для защиты информации сегмента достаточно контролировать входные / выходные данные в соответствующих точках входа / выхода. Повторный контроль при обработке и / или передаче данных внутри изолированного сегмента становится излишним, как и шифрование внутрисегментного сетевого трафика.

Работа пользователей в сегментах осуществляется с использованием АРМ и, естественно, сопровождается соответствующим информационным обменом. В простейшем случае АРМ — это отдельный компьютер, размещенный, вообще говоря, на удаленной площадке и обеспечивающий пользователю доступ в определенный сегмент динамического ЦОД. В таком виде АРМ представляет собой доверенное ТС, оснащенное аппаратными средствами защиты и функционирующее под управлением доверенной ОС. На АРМ должно быть установлено полнофункциональное средство антивирусной защиты, что обеспечит антивирусный контроль входных / выходных данных. Для защиты магистральных сетевых каналов АРМ-ЦОД должны использоваться криптомаршрутизаторы, которые обеспечивают шифрование трафика и, кроме того, межсетевое экранирование,

фильтрацию сетевых пакетов и защиту от сетевых атак. Отметим, что АРМ для доступа пользователей в сегменты может быть реализовано и как виртуальная машина. Решение вопроса организации защиты входных/выходных потоков данных в ходе взаимодействия сегментов ЦОД и внешних источников и межсегментного взаимодействия обеспечивается набором шлюзов, включая односторонний шлюз, разграничающий информационную систему (ИС) субъекта и целевой сегмент ЦОД; двухсторонний шлюз, контролирующий ограниченный трафик, порождаемый в процессе обмена данными в режиме запрос–ответ, возникающего при использовании делегированного ресурса; программный двухсторонний шлюз, обслуживающий обмен данными между сегментами ЦОД.

Финальный элемент в ряду обеспечивающих моделей функционирования ИАСЦ — **модель сопровождения и эксплуатации**. С точки зрения субъектов стратегического планирования — пользователей ИАСЦ — целью функционирования ИАСЦ ставится максимально оперативное массовое предоставление некоторого конечного спектра стандартных информационных услуг определенного качества потребителям этих услуг (сервисов). Ранее авторами была разработана процессная модель модернизации и развития ИС на всех стадиях жизненного цикла [9], обеспечивающая возможность непрерывного развития и модернизации ИАСЦ в связи с возникновением новых функциональных требований, выявлением ошибок, с учетом новых предложений по техническим, организационным, кадровым и нормативно-правовым изменениям. Исходя из этого, задачи сопровождения и эксплуатации ИАСЦ заключаются в поддержании всех ключевых характеристик полезности и гарантии информационных сервисов, а также обновлении и расширении их спектра. В связи с этим ИАСЦ и эксплуатирующие его субъекты должны рассматриваться в рамках парадигмы сервис-провайдинга, эффективная реализация которой достигается внедрением в ИАСЦ методологии ITIL/ITSM (IT Infrastructure Library/IT Service Management). Согласно методологии ITIL, модель сопровождения и эксплуатации ИАСЦ должна включать:

- адаптацию сервисной методологии функционирования к условиям ИАСЦ и модели его функционирования (процесс функционирования ИАСЦ следует рассматривать с точки зрения предоставления сервисов различным группам потребителей, а сам ИАСЦ — как сервис-провайдера; реализуется повторяемый жизненный цикл сервисов из пяти этапов: (1) стратегия сервиса; (2) проектирование сервиса; (3) преобразование (ввод в эксплуатацию) сервиса; (4) эксплуатация сервиса; (5) непрерывное улучшение сервиса);
- принципы взаимодействия поставщиков и потребителей сервисов (разделение потребителей сервисов на конечных потребителей (пользователи, обеспечивающие решение целевых задач ИАСЦ) и промежуточных потребителей; разделение поставщиков сервисов на внешних (субъектов стратегического планирования), внутренних и общих (служба эксплуатации ИАСЦ, экспертные организации и др.);

- организационно-управляющую структуру (единая сервисная служба эксплуатации, включающая в себя сотрудников всех подразделений, участвующих в разработке, создании, эксплуатации и модернизации);
- принципы формирования метрик модели функционирования (построение ИАСЦ как саморегулирующейся ИТ-системы с эффективной обратной связью на основе системы метрик; отправной точкой формирования метрик должен служить набор целевых показателей реализации ИАСЦ в рамках государственного процесса стратегического планирования);
- портфель сервисов первой очереди реализации сервисного подхода (предложенная модель функционирования позволяет предложить такой начальный набор: сервис аутентификации пользователей, сервис доступа к ресурсам локальной вычислительной сети, сервис электронного документооборота, сервис резервного копирования и восстановления, сервис создания банков данных (эталонных и внутренних), сервис поддержания банков данных);
- требования к специальному программному обеспечению (СПО) ITIL/ITSM, планируемому к внедрению в ИАСЦ (СПО минимально должно обеспечивать информационную поддержку сотрудников на этапах проектирования, преобразования, эксплуатации и непрерывного улучшения услуги и автоматизировать выполнение базовых процессов ITIL/ITSM: управление уровнем сервиса; управление каталогом услуг; управление событиями; управление инцидентами и запросами на обслуживание (обращениями); управление проблемами; управление активами и конфигурациями; управление изменениями; управление релизами; управление знаниями; учет затрат).

Сказанное здесь в отношении модели сопровождения и эксплуатации ИАСЦ по методологии ITIL/ITSM носит обзорный характер. За деталями следует обратиться к традиционным источникам по данному направлению [10–14].

## 5 Архитектурные решения — принципы построения программно определяемого центра обработки данных и его основные элементы

Ключевое архитектурное решение в рамках предложенной модели функционирования ИАСЦ — реализация динамического ЦОД — будет осуществляться в сложных условиях, обусловленных действием экономических и внешнеполитических факторов [5], которые необходимо учитывать, в особенности при выборе аппаратно-программной платформы. Очевидным современным трендом в архитектуре построения ЦОД является движение в сторону программно определяемых ЦОД (SDDC, Software-Defined Data Center).

Термин программно-конфигурируемая сеть (Software-Defined Networking) изначально использовался для описания работы вокруг OpenFlow в Стэнфордском университете [15]. Всемирно известная аналитическая компания Forrester Research дает следующее формальное определение: SDDC — это интегрированный уровень абстракции, который полностью определяет ЦОД посредством

программного уровня, представляющего ресурсы ЦОД в виде пулов виртуальных и физических ресурсов и позволяющего их комбинировать для поддержки произвольных определяемых пользователем сервисов [16]. В SDDC используется абстрактное представление вычислительных ресурсов, сетей передачи данных и систем хранения данных, что позволяет легко и гибко распределять пулы ресурсов между прикладными задачами/системами, что, в свою очередь, существенно повышает гибкость и возможности масштабирования ИТ-инфраструктуры.

**Первый элемент SDDC** — технология виртуализации вычислительных ресурсов — широко распространена и в настоящее время де-факто стала стандартом при создании практически любого современного ЦОД. Наиболее распространенным представляется стандарт, используемый в открытых проектах на базе ОС GNU Linux, которым является гипервизор KVM (kernel-based virtual machine) с эмулятором QEMU (quick emulator).

**Вторым важным элементом SDDC** выступают технологии программно-конфигурируемых сетей SDN (Software-Defined Networking), которые позволяют перевести сетевые элементы под контроль настраиваемого программного обеспечения, сделать их более интеллектуальными, облегчить управление ими. Основная суть SDN состоит в физическом разделении уровня управления сетью и уровня передачи данных за счет переноса функций управления (маршрутизаторами, коммутаторами и т. п.) в приложения, работающие на отдельном сервере (контроллере). В результате должна получаться гибкая, управляемая, адаптивная и экономичная архитектура, которая способна эффективно адаптироваться под передачу больших потоков разнородного трафика.

**Третьим важнейшим элементом SDDC** стало использование решений SDS (Software-Defined Storage) — программно определяемых хранилищ. Компания RAIDIX, один из реальных ведущих отечественных разработчиков программных решений в области хранения данных, следующим образом определяет основные характеристики программно определяемых систем хранения данных (СХД) [17]:

- SDS служат интеллектуальным компонентом, т. е. «мозгами» СХД, и они отделены от аппаратного обеспечения, которое становится исполнителем;
- SDS значительно упрощают администрирование СХД; физическое расположение данных, их перемещения со временем, организация их защиты — все решает программа-«мозг», а не администратор, но администратор задает политики, на основе которых решения принимаются;
- SDS по своей природе имеют масштабируемую архитектуру и позволяют масштабировать производительность и объем на лету; при этом добавление в кластер очередного узла не отличается по своей сложности как в начале эксплуатации, так и после проведения нескольких тысяч расширений.

Наконец, из возможности программного определения всех основных элементов ЦОД и обеспечения их функционирования на общедоступных аппаратных средствах логичным образом следуют решения по гиперконвергенции — программно определяемому подходу к созданию ИТ-инфраструктуры, при котором

объединение виртуализированных средств хранения, вычислительных мощностей и, возможно, сетевых ресурсов происходит в одном физическом устройстве. Управление таким устройством осуществляется как единым целым, и все компоненты в нем тесно интегрированы. При этом ЦОД состоит из однотипных аппаратно-программных блоков, а горизонтальное масштабирование происходит путем добавления в ЦОД очередного подобного устройства.

## 6 Заключение

В рамках представленной работы окончательно сформирован проектный облик создаваемого ИАСЦ — концепции функционального и информационного обеспечения дополнены моделью функционирования системы и обеспечивающей ее архитектурой. За рамками обсуждения остались варианты конкретных технических решений. Для подкрепления сформулированных положений, касающихся архитектурно-технических решений, требуется дополнить исследование анализом конкретных производителей и выпускаемой продукции. Этот предмет выходит за рамки данной статьи, но важно подчеркнуть, что все сформулированные предложения подкреплены более или менее широким спектром вариантов реализации. Так, в части принципиальных элементов SDDC богатый материал дает инициатива OpenStack, решения для SDN реализуются консорциумом Open Networking Foundation (ONF) и структурами OpenDaylight.org. Весьма широк рынок решений SDS, и на нем есть представители российской промышленности (например, уже упомянутая компания RAIDIX, компания Yadro). Имеется российское представительство и в поставках гиперконвергентных решений («Рос-платформа», IBS). Отдельного обсуждения заслуживает и анализ традиционных аппаратно-программных средств: АРМ и серверов, коммуникационных средств и системного ПО, средств разработки, особенно с учетом фактора импортозамещения. Эти рабочие вопросы реализации проекта ИАСЦ подразумевают вывод проекта уже в реализационную стадию, этапы научного поиска в этом проекте оформлены в серии цитируемых статей в целом законченной картиной.

## Литература

1. О стратегическом планировании в Российской Федерации: Федеральный закон от 28 июня 2014 г. № 172-ФЗ.
2. Сучков А. П., Босов А. В. Концептуальные подходы к созданию информационно-аналитического ситуационного центра в интересах стратегического планирования в области обеспечения национальной безопасности // Системы и средства информатики, 2018. Т. 28. № 4. С. 73–85.
3. Сучков А. П., Босов А. В., Макоско А. А. Ситуационный анализ в процессах стратегического планирования в области обеспечения национальной безопасности // Инновации, 2018. Т. 238. Вып. 8. С. 33–39.
4. Сучков А. П., Зацаринный А. А. Проблемно-ориентированная модель угрозы национальной безопасности // Информационные войны, 2017. Т. 44. № 4. С. 74–80.

5. Босов А. В., Сучков А. П. Информационно-аналитический ситуационный центр стратегического планирования в области обеспечения национальной безопасности: факторы влияния // Системы и средства информатики, 2020. Т. 30. № 1. С. 115–125.
6. О вводе в промышленную эксплуатацию программного обеспечения, реализующего автоматизацию технологических процессов ФНС России в составе подсистем и компонент АИС «Налог-3». <https://www.garant.ru/products/ipo/prime/doc/72184530>.
7. Завершение внедрения АИС ПФР-2. [http://www.pfrf.ru/press\\_center~2016/12/20/126789](http://www.pfrf.ru/press_center~2016/12/20/126789).
8. Сучков А. П. Анализ процессов межведомственного информационного взаимодействия // Системы и средства информатики, 2018. Т. 28. № 3. С. 118–130.
9. Сучков А. П. Процессная модель модернизации и развития информационных систем на всех стадиях жизненного цикла // Системы и средства информатики, 2018. Т. 28. № 4. С. 84–95.
10. ITIL Service Strategy. — London: TSO, 2007. 266 p.
11. ITIL Service Design. — London: TSO, 2007. 338 p.
12. ITIL Service Transition. — London: TSO, 2007. 266 p.
13. ITIL Service Operations. — London: TSO, 2007. 266 p.
14. ITIL Continual Service Improvement. — London: TSO, 2007. 226 p.
15. Greene K. Breakthrough technologies: Software-defined networking, 2009. MIT Technical Review 10. <http://www2.technologyreview.com/article/412194/tr10-software-defined-networking>.
16. Ганьжа Д. На пути к программно определяемому центру обработки данных // Журнал сетевых решений // LAN, 2016. № 07-08.
17. Платонов С. Под капотом у SDS: поедет или нет? Специально для журнала Connect // Connect WIT, 2015. № 3.

*Поступила в редакцию 17.10.19*

---

---

## **INFORMATION AND ANALYTICAL SITUATIONAL CENTER FOR STRATEGIC PLANNING IN THE FIELD OF NATIONAL SECURITY: PERSPECTIVE ARCHITECTURAL SOLUTIONS**

*A. V. Bosov and A. P. Suchkov*

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** The article proposes a variant of perspective architecture, information and analytical situation center (IASC) for management of strategic planning processes in the field of national security. The decision is based on previously published conceptual approaches and analysis of technical, technological, organizational, and legal factors that have a fundamental impact on the IASC

project. The proposed architectural solution provides both the account of the selected factors of influence and the previously developed functional scheme of the IASC, based on the principles of situational analysis in the processes of strategic planning in the field of national security. The technical basis of the solution is provided by a group of modern technologies that make up the elements of building software-defined data processing centers. The subject area of the IASC, which involves the use of special methods of situational analysis, a wide coverage of heterogeneous information sources, flexible interaction with departmental systems, including those processing information of limited access, is subject to accounting within such a data center.

**Keywords:** strategic planning; situation center; architectural solutions

**DOI:** 10.14357/08696527200214

## Acknowledgments

The study was funded by the Russian Foundation for Basic Research, project No. 18-00-00297.

## References

1. 172-FZ. 2014. O strategicheskem planirovaniyu v Rossiyskoy Federatsii: Federal'nyy zakon [On strategic planning in the Russian Federation: Federal law No. 172-FZ dated June 28, 2014].
2. Suchkov, A. P., and A. V. Bosov. 2018. Kontseptual'nye podkhody k sozdaniyu informatsionno-analiticheskogo situatsionnogo tsentra v interesakh strategicheskogo planirovaniya v oblasti obespecheniya natsional'noy bezopasnosti [Conceptual approaches to the creation of information and analytical situation center in the interests of strategic planning in the field of national security]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(4):73–85.
3. Suchkov, A. P., A. V. Bosov, and A. A. Makosko. 2018. Situationsanaliz v protsessakh strategicheskogo planirovaniya v oblasti obespecheniya natsional'noy bezopasnosti [Situation analysis in the processes of strategic planning in the field of national security]. *Innovatsii [Innovations]* 238(8):33–39.
4. Suchkov, A. P., and A. A. Zatsarinnyy. 2017. Problemno-oriyentirovannaya model' ugrozy natsional'noy bezopasnosti [Problem-oriented model of national security threat]. *Informatsionnye voyny [Information Wars]* 44(4):74–80.
5. Bosov, A. V., and Suchkov. 2020. Informatsionno-analiticheskiy situatsionnyy tsentr strategicheskogo planirovaniya v oblasti obespecheniya natsional'noy bezopasnosti: faktory vliyaniya [Information and analytical situational center for strategic planning in the field of national security: Factors of influence]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(1):115–125.
6. O vvode v promyshlennuyu ekspluatatsiyu programmnogo obespecheniya, realizuyushchego avtomatizatsiyu tekhnologicheskikh protsessov FNS Rossii v sostave podsistemy komponent AIS “Nalog-3” [About commissioning of the software realizing automation of technological processes of FTS of Russia as a part of subsystems and AIS

- component “Nalog-3”]. Available at: <https://www.garant.ru/products/ipo/prime/doc/72184530/> (accessed October 17, 2019).
- 7. Zavershenie vnedreniya AIS PFR-2 [Completion of implementation of AIS PFR-2]. Available at: [http://www.pfrf.ru/press\\_center~2016/12/20/126789](http://www.pfrf.ru/press_center~2016/12/20/126789) (accessed October 17, 2019).
  - 8. Suchkov, A. P. 2018. Analiz protsessov mezhvedomstvennogo informatsionnogo vzaimodeystviya [Analysis of processes of interdepartmental information interaction]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(3):118–130.
  - 9. Suchkov, A. P. 2018. Protsessnaya model' modernizatsii i razvitiya informatsionnykh sistem na vsekh stadiyakh zhiznennogo tsikla [Process model of modernization and development of information systems at all stages of the life cycle]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(4):84–95.
  - 10. TSO. 2007. ITIL service strategy. London. 266 p.
  - 11. TSO. 2007. ITIL service design. London. 338 p.
  - 12. TSO. 2007. ITIL service transition. London. 266 p.
  - 13. TSO. 2007. ITIL service operations. London. 266 p.
  - 14. TSO. 2007. ITIL continual service improvement. London. 226 p.
  - 15. Greene, K. 2009. Breakthrough technologies: Software-defined networking. MIT Technical Review 10. Available at: <http://www2.technologyreview.com/article/412194/tr10-software-defined-networking> (accessed October 17, 2019).
  - 16. Gan'zha, D. 2016. Na puti k programmno-opredelyaemomu tsentru obrabotki dannykh [Towards a software-defined data center]. *LAN* 07-08. Available at: <https://www.osp.ru/lan/2016/07-08/13050118/> (accessed October 17, 2019).
  - 17. Platonov, S. 2015. Pod kapotom u SDS: poedet ili net? Spetsial'no dlya zhurnala Connect [Under the hood at SDS: Will go or not? Specifically for Connect magazine]. *Connect WIT* 3. Available at: <http://www.connect-wit.ru/pod-kapotom-u-sds-poedet-ili-net.html> (accessed October 17, 2019).

Received October 17, 2019

## Contributors

**Bosov Alexey V.** (b. 1969) — Doctor of Science in technology, principal scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; AVBosov@ipiran.ru

**Suchkov Alexander P.** (b. 1954) — Doctor of Science in technology, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; ASuchkov@frccsc.ru

## СОСТАВ, СТРУКТУРА И МЕТОДЫ АНАЛИЗА ДАННЫХ СИСТЕМЫ УПРАВЛЕНИЯ НАУЧНЫМИ СЕРВИСАМИ\*

*A. П. Сучков<sup>1</sup>*

**Аннотация:** Включение научных исследований в цифровую сферу во многом обусловлено возможностью создания эффективной системы широкодоступных научных сервисов. Система научных сервисов (СНС) должна обеспечивать управление процессами автоматизированного подбора и предоставления релевантных сервисов и осуществления различных формальных и неформальных коммуникаций исследователя и государства, научного сообщества и бизнеса. Эффективность такой системы напрямую зависит от полноты и целостности информационной модели данных, ориентированной на релевантные методы анализа, что, в свою очередь, обеспечивается правильным выбором методологии управления научными сервисами с учетом всей совокупности обеспечивающих процессов. Данная модель опирается на методы ситуационного управления и ситуационного анализа с использованием средств цифровой платформы, обеспечивающих процессы поддержки принятия решений в режиме реального времени.

**Ключевые слова:** научные сервисы; система научных сервисов; информационная модель; ситуационный анализ

**DOI:** 10.14357/08696527200215

### 1 Введение

Включение научных исследований в цифровую сферу во многом обусловлено возможностью создания эффективной системы широкодоступных научных сервисов. Научный сервис — совокупность процессов и ресурсов по обслуживанию проектов научно-исследовательского и прикладного характера путем сервисного предоставления потребителю (исследователям, специалистам или организациям) продуктов интеллектуальной научной деятельности, оборудования, информационно-коммуникационных ресурсов, расходных материалов и услуг обслуживающего персонала. Результатом сервисной деятельности является услуга [1].

Очевидно, что научный сервис опирается как на цифровые технологии (автоматическое и автоматизированное предоставление услуг), так и на использование интеллектуальных и обслуживающих человеческих ресурсов. Система научных сервисов должна обеспечивать управление процессами автоматизированного подбора и предоставления релевантных сервисов и осуществления различных

---

\*Работа выполнена при частичной поддержке РФФИ (проект 18-29-03091).

<sup>1</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, ASuchkov@ipiran.ru

формальных и неформальных коммуникаций исследователя и государства, научного сообщества и бизнеса.

Система научных сервисов должна обеспечивать «безбарьерный» доступ исследователей к заказу и получению услуг с использованием существующей научной инфраструктуры (Центры коллективного пользования (ЦКП), уникальные научные установки (УНУ), оцифрованные коллекции и банки данных организаций, системы поиска публикаций и др.), что должно обеспечить технологическую основу для цифровой трансформации научной деятельности. Система управления научной инфраструктурой должна быть обеспечена достоверной и актуальной информацией об использовании ресурсов научной инфраструктуры, а также аналитическими инструментами, обеспечивающими повышение эффективности их использования при выполнении научно-исследовательских и опытно-конструкторских работ.

Эффективность такой системы напрямую зависит от полноты и целостности информационной модели данных, ориентированной на релевантные методы анализа, что, в свою очередь, обеспечивается правильным выбором методологии управления научными сервисами с учетом всей совокупности обеспечивающих процессов [2].

## 2 Общая методология

Исходя из целей и задач СНС должна обеспечивать возможность непрерывного управления инфраструктурой с целью обеспечения эффективного доступа к научным сервисам, выработки планов эффективной загрузки ЦКП и оборудования, обеспечения их жизненного цикла, мониторинга процессов коллективного пользования инфраструктурой и текущего контроля исполнения принятых решений [3]. Данная постановка предполагает использование методов ситуационного управления и ситуационного анализа с использованием средств цифровой платформы, обеспечивающих процессы поддержки принятия решений в режиме реального времени и опирающихся на такие понятия, как событие, ситуация, угроза, управление. *Событием* называется изменение состояния объектов контролируемого пространства. *Ситуация* определяется состоянием взаимосвязанных элементов обстановки в контролируемом пространстве; изменения обстановки определяются событиями, образующими некоторые разворачивающиеся во времени, наблюдаемые и регистрируемые потоки. При этом под управлением понимается *целенаправленное* воздействие органа управления на подчиненные ему или взаимодействующие элементы обстановки (ресурсы).

Совокупность ситуаций в системе управления распадается на текущие, прогнозируемые и целевые ситуации. При этом текущие ситуации являются результатом наблюдения и регистрации событий, прогнозируемые определяются методами ситуационного анализа, а целевые отражают краткосрочные, среднесрочные и долгосрочные цели управления. Процесс установления и формализации целей называется *целеполаганием*. Последнее немаловажно, так как

зачастую ситуационный анализ понимается как обеспечение реакций системы управления на чрезвычайные ситуации после того, как они сложились. Однако теория ситуационного подхода предполагает учет «планируемой и прогнозируемой обстановки», отражающей стратегические, тактические и оперативные цели управления, а также учет факторов самоорганизации управляющего сегмента системы, определяющих стимулы для достижения этих целей.

Под *угрозой* в процессах управления понимается ситуация или совокупность ситуаций, развитие которых противоречит целям управления и отдаляет текущее состояние от целевого.

Понятие ситуационного анализа используется в процессах управления, когда возникает необходимость выработки управленческих решений в режиме реального времени в условиях сложной обстановки, складывающейся в отношении контролируемых объектов. Ситуационный анализ обеспечивает на основе событийного мониторинга обстановки идентификацию ситуаций (угроз), слежение за состоянием ситуаций, прогноз сценариев их развития, а также поддержку процесса выработки управляющих воздействий для достижения целевой ситуации (нормализация ситуации).

На рис. 1 представлена общая схема формирования информационной модели СНС исходя из методологии ситуационного управления, состава процессов, обеспечивающих функционирование этой системы, и совокупности информационных объектов, поддерживающих эти процессы.

### 3 Целеполагание

Совокупность объектов мониторинга СНС формируется исходя из результатов процессов целеполагания. Согласно общим принципам целеполагания в системах управления формулируемые цели должны быть конкретными, измеримыми, достижимыми, ресурсообеспеченными и привязанными ко времени [4]. Выполнение этих требований обычно осуществляется путем формирования системы целевых показателей и индикаторов (ЦПИ), представляющих собой функции от определенных параметров, имеющих числовое значение и единицу измерения, а также критерии их достижения (см. рис. 1).

Так, Постановлением Правительства Российской Федерации от 17 мая 2016 г. № 429 установлены следующие целевые показатели федерального уровня для ЦКП и УНУ:

- отношение фактического времени работы оборудования центра и (или) уникальной установки к максимально возможному времени работы оборудования центра и (или) уникальной установки за год (значение указанного показателя должно быть установлено в размере не менее 70%);
- отношение фактического времени работы оборудования центра и (или) уникальной установки в интересах третьих лиц к фактическому времени работы оборудования центра и (или) уникальной установки за год;

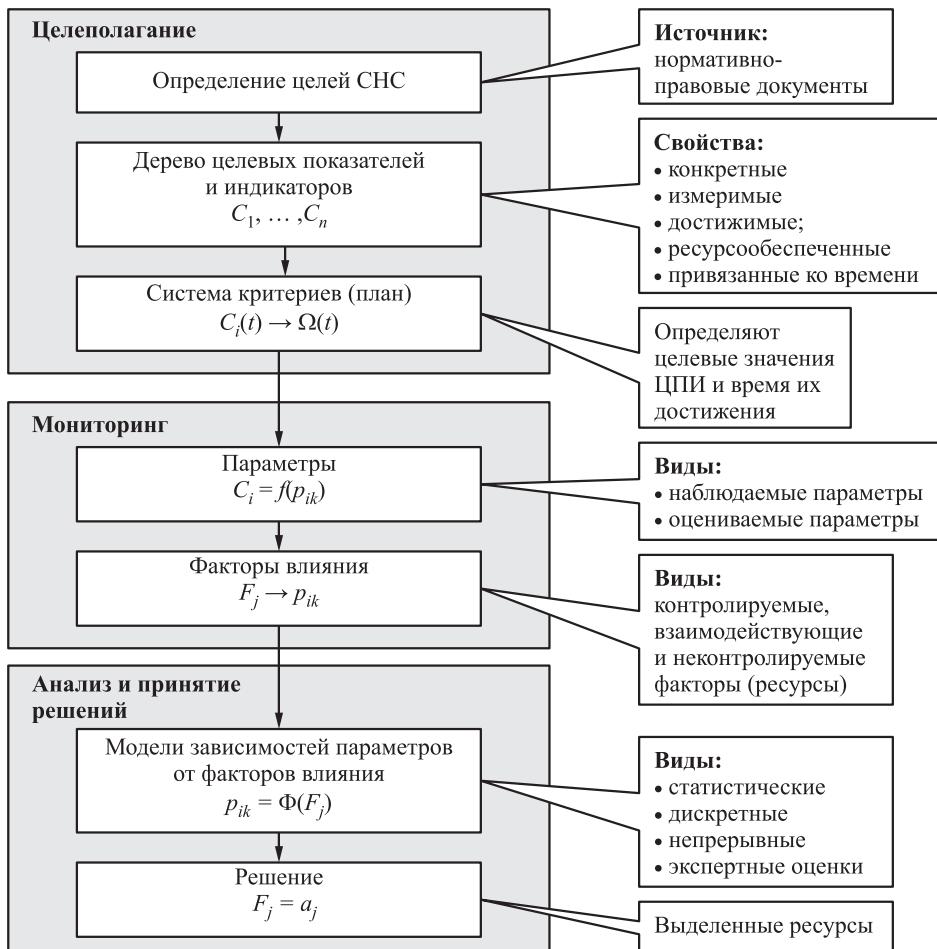


Рис. 1 Формирование информационной модели СТЧ

- количество организаций-пользователей и (или) организаций, участвующих в проведении исследований (экспериментов) с использованием уникальной установки, в год;
- количество публикаций в российских и иностранных научных журналах, индексируемых в информационно-аналитических системах научного цитирования «Сеть науки» (WEB of Science Core Collection) и «Scopus», а также иных результатов интеллектуальной деятельности, полученных с использованием уникальной установки, в год.

Значения ЦПИ устанавливаются на правительственноом, ведомственном, подведомственном и внутрисистемном уровнях, образуя дерево целей системы.

В качестве примеров ЦПИ различных уровней можно привести следующие:  
*Ведомственный и подведомственный уровень:*

- средняя загрузка оборудования ЦКП (УНУ) в интересах бюджетных организаций за год;
- средняя загрузка оборудования ЦКП (УНУ) в интересах третьих лиц за год;
- число подведомственных организаций, участвующих в проведении исследований на ЦКП (УНУ), в год;
- число подведомственных ЦКП (УНУ);
- число обслуженных исследовательских бюджетных программ;
- число выполненных заявок;
- число поступивших заявок;
- среднее время исполнения одной заявки;
- индикатор выполнения регламентных работ ЦКП (УНУ);
- индикатор своевременного обеспечения ЦКП / УНУ ресурсами;
- среднее время ремонта ЦКП / УНУ.

*Внутрисистемный уровень:*

- процент готовности системы;
- среднее время наработки на отказ;
- индикатор выполнения плана регламентных работ.

В табл. 1 приведен пример формализации понятия ЦПИ.

В рамках целеполагания осуществляются процессы планирования ЦПИ — формирование системы критериев достижения требуемых значений целевых показателей в заданные периоды времени. Атрибуты объекта «Планирование»

**Таблица 1** Пример структуры информационного объекта ЦПИ

Наименование	Вид	Параметры	Формула	Ед. изм.	Критерий	Критическое отклонение
Загрузка оборудования ЦКП и (или) уникальной установки в интересах бюджетных работ за год	ЦПИ федерального уровня	– время суммарной загрузки оборудования ЦКП ( $t_c$ ); – максимально возможное время работы оборудования ( $t_m$ )	$\frac{100t_c}{t_m}$	%	не менее 70% за год	20%

в информационной модели СНС: ЦПИ; критерий ( $\ll=$ ,  $\ll\leq$ ,  $\ll\geq$ ); плановое значение; текущее значение; единица измерения; срок (временная градация); критическое отклонение. Планирование ЦПИ должно учитывать объективные «сезонные колебания», обусловливающие неравномерность их изменения в планируемом периоде (1 год) для получения возможности текущего контроля ЦПИ. Планирование обычно осуществляется с учетом показателей предыдущих плановых периодов.

## 4 Мониторинг

Целью мониторинга каждой системы ситуационного управления является сбор данных, необходимых для принятия решений, при этом она взаимодействует с внутренними и внешними контролируемыми объектами с помощью своей сенсорной системы. Исходя из вышеизложенной структуры информационной модели, осуществляются следующие процессы:

- учет объектов мониторинга;
- мониторинг контролируемого пространства с использованием технических средств, платформы обмена документами, средств контроля и управления функционированием систем;
- мониторинг ЦПИ;
- мониторинг реализации плановых показателей;
- мониторинг параметров и факторов влияния;
- регламентный и событийный мониторинг научных сервисов;
- мониторинг заявок на предоставление научных сервисов;
- регламентный и событийный мониторинг элементов СНС, учета информации о наличии, состоянии, загрузке оборудования ЦКП и УНУ;
- мониторинг исполнения принятых решений;
- мониторинг неструктурированной информации (нормативно-справочная информация (НСИ), научные результаты, целевые публикации, новости из внешних источников, элементы инфраструктуры, включая сайты ЦКП);
- визуализация данных мониторинга с индикацией событий и ситуаций в режиме реального времени.

Учету подлежат все контролируемые объекты СНС. Прежде всего это ЦПИ, а также параметры и факторы влияния, обеспечивающие возможность вычислять значения показателей. Осуществляется мониторинг научных сервисов, ЦКП, УНУ, оцифрованных коллекций и банков данных организаций, заявок, принимаемых решений, новостей из внешних источников, целевых публикаций, элементов инфраструктуры, включая сайты ЦКП. Учет объектов мониторинга

осуществляется с целью обеспечения возможности ввода, хранения, корректировки перечня объектов мониторинга.

Особое место занимает мониторинг научных результатов, которые являются квинтэссенцией всех научных исследований и основным информационным объектом для процессов постановки научной проблемы, формулировки научных гипотез, мониторинга научно-технической информации (НТИ). Формализованное понятие «научный результат», интегрирующее всю информационную структуру научного исследования, используется при организации значительного числа научных сервисов: сервисов извлечения фактов и знаний (извлечение фактов, понятий, связей и формализация фактографических данных на основе лингвистического анализа слабоструктурированной информации); интеллектуального поиска информации; тематического индексирования; наукометрического анализа; анализа фронта исследований; сервисов коммуникаций (сервисов работы с НТИ и коммуникаций в научном сообществе; интеллектуального анализа специализированных социальных сетей и других средств научных коммуникаций); проверки заимствований [5]. Особенно это актуально для междисциплинарных исследований, где обеспечивается поиск и подбор релевантных научных положений и инструментария в смежных областях научного знания.

Мониторинг инфраструктуры СНС осуществляется с целью сбора данных о функционировании информационной, программной и технической составляющих СНС, а также состоянии информационной безопасности.

Мониторинг исполнения принятых решений осуществляется с использованием платформы работы с документами; в его задачи входит контроль срока исполнения и автоматическая индикация просрочки, о чем извещается ответственный исполнитель.

В рамках мониторинга неструктурированной информации (новостные ленты, сайты научных организаций, сайты взаимодействующих ведомств) обеспечивается возможность анализа интернет-ресурсов с целью отбора данных по тематике научных сервисов.

Данные мониторинга заносятся в базу данных (БД) «События», где каждое событие имеет следующие атрибуты: объект мониторинга; дата / время; состояние; параметр; условия, где состояние характеризуется списком допустимых состояний объекта мониторинга (например, для заявки: подана, принята к исполнению, выполняется, выполнена, закрыта и т. д.), параметр характеризует градацию состояния, условия — текстовое пояснение состояния.

Событийный мониторинг научных сервисов СНС и обеспечение сбора данных мониторинга научных сервисов осуществляются в режиме реального времени по запросу или по установленному регламенту с целью установления работоспособности сервиса. Под событием (инцидентом) понимается изменение состояния объекта мониторинга. Источники данных — агенты и ручной ввод. Данные мониторинга помещаются в соответствующие разделы БД «Событие».

По результатам мониторинга осуществляется ведение следующих информационных ресурсов:

- БД НСИ;
- каталога ЦКП, УНУ, коллекций и БД;
- каталога подведомственных организаций;
- каталога научных сервисов;
- реестра научных результатов;
- БД «События»;
- БД «Обстановка»;
- БД «Новостная лента»;
- БД «Контроль функционирования»;
- БД «Типовые решения»;
- перечня решений на контроле.

Визуализация данных мониторинга осуществляется с индикацией событий и ситуаций в режиме реального времени:

- отображение текстовых, графических, звуковых индикаторов;
- настройка системы визуализации;
- автоматическая индикация.

## 5 Анализ данных мониторинга

Так как функционирование СНС является целесообразным, а цели системы formalизованы в виде совокупности ЦПИ и документов планирования, основной задачей анализа данных мониторинга становится анализ текущего состояния и прогноз степени реализации плановых показателей ЦПИ. Анализ текущего состояния заключается в выявлении отклонений от плановых показателей, их причин и путей нормализации ситуации. Прогнозирование ситуации на основе ретроспективных данных осуществляется на ближнюю и дальнюю перспективу с учетом реализации различных сценариев развития ситуации, включая принимаемые решения по применению управляющих воздействий.

Если имеется набор ЦПИ  $C_i, i = 1, \dots, n$ , то планирование осуществляется путем задания их целевых значений в определенные временные отрезки, на которые разбит планируемый период:

$$C_i(t_s) \in \Omega_i(t_s), \quad s = 1, \dots, S,$$

где  $\Omega_i(t_s)$  — целевое множество для ЦПИ  $C_i$  в моменты времени  $t_s$ . Целевое множество определяются либо в виде конкретного числа, либо интервалом чисел. По определению ЦПИ должны быть измеримыми, т. е. функциями от совокупности наблюдаемых и оцениваемых в процессе мониторинга параметров  $p_{ik}$ ,  $i = 1, \dots, n, k = 1, \dots, K_i$ :

$$C_i = f_i(p_{ik}).$$

Обычно функции  $f_i$  задаются в процессе целеполагания. Например, в табл. 1 целевой показатель выражается формулой

$$C_1 = \frac{100t_c}{t_m},$$

где параметрами выступают время суммарной загрузки оборудования ЦКП ( $t_c$ ) и максимально возможное времена работы оборудования ( $t_m$ ). Целевое множество  $\Omega$  представляет собой полуинтервал:  $C_1 \geq 70\%$  за годовой период.

Целесообразное управление СНС осуществляется с использованием ресурсов системы (финансовых, технических, трудовых и пр.), с помощью которых оказывается воздействие на значения ЦПИ. Управляющие воздействия в виде целенаправленно выделяемых ресурсов можно рассматривать как факторы влияния  $F$  на значения параметров ЦПИ. Очевидно, что одна из основных задач аналитической СНС — выявление зависимостей параметров  $p$  от факторов влияния  $F$  и формирование моделей вида

$$p_{ik} = \Phi(F_{ik}).$$

Такие модели могут быть получены в результате изучения предыдущего опыта эксплуатации СНС, а также на основе имитационного моделирования процессов управления либо на основе экспертных оценок. Так, в рассматриваемом случае параметр суммарной загрузки оборудования  $t_c$  зависит от следующих факторов: числа заявок на предоставление услуг  $Z$ , среднего времени обслуживания заявки  $t_o$  и времени неработоспособности оборудования  $t_u$  (ремонт и техническое обслуживание). Зависимость описывается уравнением:

$$\left[ \frac{t_m - t_u}{t_o} \right] = Z, \quad t_u = t_m - t_c \Rightarrow \left[ \frac{t_c}{t_o} \right] = Z,$$

откуда, с учетом целевого множества, следует

$$C_1 = \frac{t_o}{t_c} Z \geq 0,7.$$

Полученная модель зависимости целевого показателя от факторов влияния позволяет распределять ресурсы с целью увеличения загрузки оборудования и числа исполняемых заявок для достижения установленных нормативов. В общем виде задачу управления ЦПИ можно сформулировать следующим образом:

$$\begin{aligned} \forall s \quad 1 \leq s \leq S \quad \exists \vec{F}(t_s) = \vec{a}_s \in A : \text{dist} \left( \vec{C}(t_s), \vec{\Omega}(t_s) \right) = 0, \\ \vec{C}(t_s) = \vec{f}(\vec{p}(t_s)), \quad \vec{p}(t_s) = \vec{\Phi} \left( \vec{F}(t_s) \right), \end{aligned}$$

где  $A$  — множество ресурсов;  $\tilde{\Omega}$  — совокупность целевых множеств;  $\vec{F}(t_s) = \vec{a}_s$  — принимаемые решения по применению управляющих воздействий. Так как система целей СНС является множеством, применяются методы многокритериального выбора и сценарного прогнозирования.

Как правило, для выбора альтернатив из множества вариантов решений в условиях ограничения на ресурсы применяются технологии, обеспечивающие снижение размерности признакового пространства и построение интегрального показателя качества, при этом агрегируются исходные признаки с использованием знаний эксперта и/или предпочтений лица, принимающего решение (ЛПР) [6]. Например, можно использовать алгоритм сужения множества Парето на основе нечеткой информации об отношениях предпочтения ЛПР [7].

Оценка качества выбранных альтернатив решения может осуществляться с использованием метода сценарного прогнозирования, который позволяет учесть управляющие воздействия при прогнозировании развития обстановки. Для анализа обстановки и поддержки процессов принятия решений применяются математические методы статистического анализа:

- анализ временных рядов, характеризующих изменение количественных и качественных атрибутов ЦПИ (анализ трендов, сезонных колебаний, тенденций и аномалий);
- прогнозирование изменения ЦПИ с учетом выявленных трендов и анализируемых сценариев развития обстановки;
- динамическое моделирование ситуаций;
- статистическая оценка количественных и качественных характеристик потоков событий.

Объектами контроля при прогнозировании состояния являются следующие элементы обеспечивающей части СНС:

- оборудование ЦКП, УНУ;
- программное обеспечение;
- информационные ресурсы (включая коллекции и БД);
- кадровые ресурсы.

При этом управление ресурсами может осуществляться в следующих областях:

- инвентаризация ресурсов (например, контроль состава ЦКП и УНУ, управление программным обеспечением, управление уровнем обновлений);
- оперативный мониторинг (например, какие сервисы или какое сетевое оборудование в данный момент работает или не работает);
- мониторинг производительности (например, на сколько процентов загружены вычислительные мощности и установки, что позволяет планировать развитие инфраструктуры и предотвращать сбои, вызванные избыточной нагрузкой);

- управление идентификационными данными сотрудников доступом (системы однократного ввода логина/пароля, системы управления доступа к информационным ресурсам на основе учетной записи — логина, системы управления учетными записями, соответствие информационных ресурсов корпоративным политикам безопасности);
- мониторинг процессов обработки информации (например, анализ взаимосвязи информационных ресурсов, реализующих единую функциональную задачу).

В процессе нормального функционирования объекта контроля решаются задачи планирования ремонта и модернизации, планирования необходимых трудовых ресурсов. В основе планирования лежит прогнозирование сроков отказа элементов контроля.

Для решения задачи прогнозирования сроков отказов используются различные модели, отличающиеся видом исходной информации, методами ее анализа и формой представления результатов. Анализ подобных моделей позволяет выделить четыре типа исходной информации: экспертные оценки ресурсов; статистика обслуживания ресурса; статистика диагностик; данные об эксплуатационных нагрузках. Это позволяет все имеющиеся модели прогнозирования также разделить на четыре вида, приведенных в табл. 2.

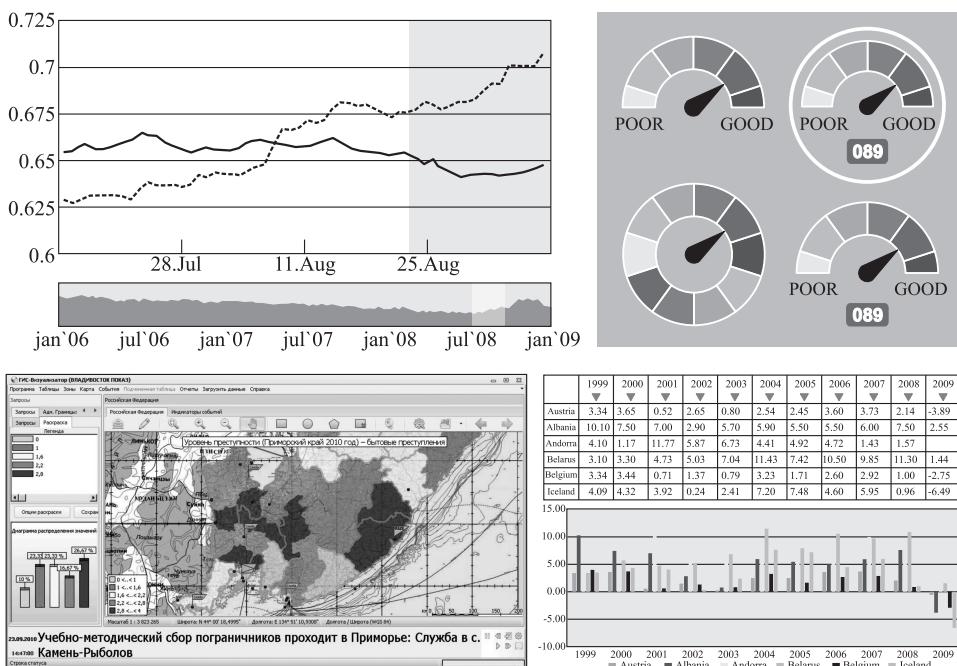
Визуализация аналитических данных с использование деловой графики (пример см. на рис. 2) предполагает реализацию следующих процедур:

- отображение инфографики аналитических отчетов, возможно на фоне цифровой карты местности для территориально распределенных систем;
- настройка системы визуализации;
- просмотр аналитики по ЦКП и интегрально по СНС.

Информационно-аналитическая поддержка выработки вариантов решений по планированию и по ситуации предоставления научных сервисов осуществляется с использованием системы управления инцидентами (ситуациями). Обеспечивается возможность идентификации ситуации и подбор типового решения. Под

**Таблица 2** Виды моделей прогнозирования

Тип модели	Исходная информация	Срок реализации прогноза
Ресурсная	Сроки обслуживания	Начало эксплуатации
Нагрузочная	Характеристики элемента контроля, характеристики эксплуатационных нагрузок	Окончание периода наблюдения
Диагностическая	Данные диагностики	Окончание последней диагностики
Экспертная	Экспертные оценки ресурсов элемента контроля	Начало эксплуатации элемента контроля



**Рис. 2** Примеры визуализации ЦПИ

ситуацией понимается достижение ЦПИ заданного критического значения, характеризующего его существенное отклонение от нормального или планового значения. Обеспечивается ведение каталога «Решения», содержащего текстовое описание алгоритма действий для нормализации ситуации.Осуществляется идентификация ситуаций по планированию, выбор типовых решений, выработка новых решений и пополнение каталога.

В процессе информационно-аналитической поддержки выработки вариантов решений по планированию и по штатным и нештатным ситуациям с использованием системы управления событиями может осуществляться:

- прогнозирование изменения обстановки: без управляющего воздействия, с управляющим воздействием, сценарное прогнозирование с учетом внешних факторов;
- поддержка процессов принятия управленческих решений: адаптация типовых решений и выработка нетиповых решений (с учетом прогнозирования).

В случае штатной ситуации применяются наработанные в системе типовые решения с привязкой их к конкретной обстановке, а в случае нештатных ситуаций осуществляется информационно-аналитическая поддержка экспертных групп для выработки новых решений.

## Литература

1. Исследование вопросов управления результатами научно-исследовательской деятельности организаций, подведомственных ФАНО России, и научными сервисами сети ЦКП ФАНО: Отчет о НИР «Сервис-У». — М.: ФИЦ ИУ РАН, 2016. 437 с.
2. Зацаринный А. А., Кондрашев В. А., Сучков А. П. Система научных сервисов как актуальный компонент научных исследований // Системы и средства информатики, 2019. Т. 29. № 1. С. 23–38.
3. Сучков А. П. Полнфункциональный процессный подход к реализации систем ситуационного управления // Системы и средства информатики, 2017. Т. 27. № 1. С. 86–100.
4. Сучков А. П. Формирование системы целей для ситуационного управления // Системы и средства информатики, 2013. Т. 23. № 2. С. 171–182.
5. Сучков А. П. Научный результат как информационный объект в контексте системы управления научными сервисами // Информатика и её применения, 2019. Т. 13. Вып. 10. С. 137–144.
6. Петровский А. Б., Лобанов В. Н. Многокритериальный выбор в пространстве признаков большой размерности: мультиметодная технология ПАКС-М // Искусственный интеллект и принятие решений, 2014. № 3. С. 92–104.
7. Басков О. В. Сужение множества Парето на основе нечеткой информации об отношении предпочтения ЛПР // Искусственный интеллект и принятие решений, 2014. № 1. С. 57–65.

Поступила в редакцию 26.11.19

---

## CONTENT, STRUCTURE, AND DATA ANALYSIS METHODS OF SCIENTIFIC SERVICES MANAGEMENT SYSTEM

*A. P. Suchkov*

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** The inclusion of scientific research in the digital sphere is largely due to the possibility of creating an effective system of widely available scientific services. The system of scientific services should provide management of processes of the automated selection and provision of relevant services and implementation of various formal and informal communications of the researcher and the state, scientific community and business. The effectiveness of such a system directly depends on the completeness and integrity of the information data model focused on relevant methods of analysis, which, in turn, is provided by the correct choice of methodology for managing scientific services, taking into account the totality of supporting processes. This model is based on the methods of situational management and situational analysis using digital platform tools that provide real-time decision support processes.

**Keywords:** scientific services; system of scientific services; information model; situational analysis

**DOI:** 10.14357/08696527200215

## Acknowledgments

The study was partially funded by the Russian Foundation for Basic Research, project No. 18-29-03091.

## References

1. FRC CSC RAS. 2016. Otchet o NIR “Servis-U”: Issledovanie voprosov upravleniya rezul’tatami nauchno- issledovatel’skoy deyatel’nosti organizatsiy, podvedomstvennykh FANO Rossii, i nauchnymi servisami seti TSKP FANO [Study of the issues of managing the results of research activities of organizations subordinate to the FANO of Russia, and the scientific services of the network of the Central Scientific and Educational Center of the FANO: Report on the research “U-Service”]. Moscow. 437 p.
2. Zatsarinny, A. A., V. A. Kondrashev, and A. P. Suchkov. 2017. Sistema nauchnykh servisov kak aktual’nyy komponent nauchnykh issledovanii [The system of scientific services as an actual component of scientific research]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(1):23–38.
3. Suchkov, A. P. 2017. Polnofunktional’nyy protsessnyy podkhod k realizatsii sistem situatsionnogo upravleniya [A fully functional process-based approach to the implementation of systems of situational management]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 27(1):86–100.
4. Suchkov, A. P. 2013. Formirovanie sistemy tseley dlya situatsionnogo upravleniya [The formation of the objective system to situational management]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 23(2):171–182.
5. Suchkov, A. P. 2019. Nauchnyy rezul’tat kak informatsionnyy ob”ekt v kontekste sistemy upravleniya nauchnymi servisami [Scientific result as an information object in the context of scientific services management system]. *Informatika i ee Primeneniya — Inform. Appl.* 10(13):137–144.
6. Petrovsky, A. B., and V. N. Lobanov. 2014. Mnogokriterial’nyy vybor v prostranstve priznakov bol’shoj razmernosti: mul’timetodnaya tekhnologiya PAKS M [Multiple criteria choice in the attribute space of large dimension: Multi-method technology PAKS-M]. *Iskusstvennyy intellekt i prinyatie resheniy* [Artificial Intelligence and Decision Making] 3:92–104.
7. Baskov, O. V. 2014. Suzhenie mnozhestva Pareto na osnove nechetkoy informatsii ob otnoshenii predpochteniya LPR [Algorithm for Pareto set reduction using fuzzy information on the DM preference relation]. *Iskusstvennyy intellekt i prinyatie resheniy* [Artificial Intelligence and Decision Making] 1:57–65.

Received November 26, 2019

## Contributor

**Suchkov Alexander P.** (b. 1954) — Doctor of Science in technology, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; ASuchkov@frccsc.ru

## МОДЕЛЬ ФУНКЦИОНИРОВАНИЯ СИСТЕМЫ АВТОМАТИЗИРОВАННОЙ ОЦЕНКИ ДОСТОВЕРНОСТИ КОНКРЕТНО-ИСТОРИЧЕСКИХ ФАКТОВ

*И. М. Адамович<sup>1</sup>, О. И. Волков<sup>2</sup>*

**Аннотация:** Статья продолжает серию работ, посвященных технологии поддержки конкретно-исторических исследований (ПКИ). Технология построена на принципах створчества и краудсорсинга и ориентирована на широкий круг не являющихся профессиональными историками и биографами пользователей. Данная статья посвящена описанию и обоснованию подхода к моделированию механизма автоматизированной оценки достоверности обрабатываемой информации, включенного в технологию с целью ее дальнейшего развития. Предложенный подход заключается в соответствующей модификации модели семантической сети, построенной с опорой на принципы графодинамики, модель Барабаша–Альберт и поддержку механизмов имитации активности пользователей. С помощью данной модели были проведены экспериментальные проверки эффективности итерационных алгоритмов реализации нового механизма автоматизированной оценки достоверности конкретно-исторической информации и оценено влияние ограничения глубины итерации на среднюю достоверность вычисленных оценок в целом.

**Ключевые слова:** моделирование; распределенная технология; достоверность информации; историко-биографический факт; автоматизированная процедура

**DOI:** 10.14357/08696527200216

### 1 Введение

В [1–3] описана распределенная технология ПКИИ, основанная на принципах краудсорсинга (мобилизации ресурсов широкого круга добровольцев посредством информационных технологий).

Модель данных, положенная в основу технологии ПКИИ, опирается на иерархическое представление историко-биографического факта (ИБФ) [4].

Ресурсом для ИБФ служит фрагмент текста историко-биографической направленности. Ссылкой на ресурс в технологии служит указатель фрагмента текста (УФТ). В совокупности ИБФ образуют семантическую сеть, состоящую из двух подсетей — сети понятий и сети экземпляров.

<sup>1</sup> Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, Adam@amsd.com

<sup>2</sup> Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, Volkov@amsd.com

Узлы сети понятий представляют собой именованные универсальные классы объектов. Единственным обязательным свойством класса является его имя. Из распределенности технологии вытекает, что сеть создается независимо разными исследователями, а поэтому она формируется из автономных именованных подсетей. Для «склейки» этих подсетей между собой предусмотрена связь эквивалентности. Помимо нее предусмотрены направленные связи типа «подпонятие», «атрибут» и «значение».

Каждая связь может иметь признак условности. Условная связь для успешной интерпретации пользователем снабжена описанием причин условности, т. е. границ ее действия. Описание состоит из текстового комментария и при необходимости набора ссылок на источники в форме УФТ.

Также каждая связь типа «атрибут» может иметь признак уникальности значения, необходимый для задач поиска противоречий в данных, и признак наследуемости значения, использующийся только для связей типа «часть».

Факты задаются значениями экземпляров классов и отношениями между ними. Отношения наследуются из сети классов. Между экземплярами одного класса может быть дополнительно установлена связь эквивалентности. Также может быть добавлено дополнительное отношение типа «часть», необходимое для наследования атрибутов. Для разделения атрибутов на наследуемые и ненаследуемые служит признак наследуемости. Каждый экземпляр имеет значение, которое может быть пустым. Непустое значение сопровождается либо текстовым комментарием, либо УФТ, связывающим его с документом-источником. Поскольку документы-источники предполагают автоматизированную обработку, они хранятся в базе данных (БД) технологии в простом текстовом формате вместе с набором своих атрибутов.

В [5] обоснована необходимость включения в состав технологии ПКИИ средств автоматизированной оценки достоверности конкретно-исторических фактов. Приведено подробное описание соответствующего механизма, определены процедуры автоматического формирования оценок достоверности информации.

Намечены пути оценки эффективности предложенного механизма. Предложен количественный показатель эффективности  $\eta$ , представляющий собой отношение объемов наследованных и пустых оценок при заданном объеме пользовательских оценок. Обосновано применение методов моделирования для его определения.

Цель статьи — определение динамики показателя  $\eta(t)$  в процессе развития семантической сети технологии ПКИИ и определение предельного уровня оценок  $\rho$ , при превышении которого автоматизированное оценивание нецелесообразно.

## 2 Описание подхода

В основе системы оценок  $O$  лежат оценки пользователей вида  $U = \langle s, d, \text{Id} \rangle$ , где  $s$  — оценка достоверности;  $d$  — степень обоснованности оценки;  $\text{Id}$  — уникальный идентификатор пользователя. Пользовательских оценок  $U$  у одного

объекта может быть несколько. Единая оценка  $O$  объекта имеет вид  $\langle \bar{s}, \bar{d}.r \rangle$ , где  $\bar{s}$  — сформированная системой оценка достоверности;  $\bar{d}$  — сформированная системой степень ее обоснованности;  $r$  — уровень оценки, определяющий ее место в иерархии. Оценка может формироваться одним из следующих способов:

- (1) *назначенная оценка* формируется на основании имеющихся пользовательских оценок данного объекта. Уровень  $r$  сформированной таким способом оценки считается наивысшим и обозначается  $r_0$ ;
- (2) *наследованная оценка* формируется для объекта, не имеющего пользовательских оценок, на основании уже сформированных оценок объектов, напрямую связанных с данным. Наследование происходит от объектов с наивысшим уровнем  $r_n$ , и сформированная оценка будет иметь уровень  $r_{n+1}$  на одну ступень ниже уровня  $r_n$ ;
- (3) *пустая оценка*  $\langle 0, 0, r_\infty \rangle$  с самым низким уровнем назначается объекту в случае отсутствия у него пользовательских оценок и связанных оцененных объектов. Для пустой оценки принимается  $s = 0$  и  $d = 0$ .

Назначенная оценка  $O = \langle s, d, r_0 \rangle$  складывается из пользовательских оценок  $\{(s_i, d_i, \text{Id}_i) | i = 1, \dots, k, k > 0\}$ , заданных данному объекту  $k$  пользователями по следующим правилам:

$$\bar{s} = \frac{\sum s_i d_i}{\sum d_i}; \quad (1)$$

$$\bar{d} = \frac{\sum d_i (1 - |\bar{s} - s_i|^\gamma)}{k}, \quad \gamma > 1. \quad (2)$$

Наследованная оценка формируется по следующим правилам: пусть  $M$  — множество объектов, непосредственно связанных с текущим. Через  $M^+$  обозначим подмножество  $M$  с максимальным уровнем их оценок. Тогда

$$\bar{s} = \frac{\sum_{i=1}^{|M^+|} s_i d_i}{\sum_{i=1}^{|M^+|} d_i}; \quad (3)$$

$$\bar{d} = \frac{\sum_{i=1}^{|M^+|} d_i (1 - |\bar{s} - s_i|^\gamma)}{|M|}, \quad \gamma > 1. \quad (4)$$

Уровень оценки  $\bar{r}$  будет на одну ступень ниже уровня оценки объектов множества  $M^+$ .

Для произвольного объекта  $m$  с наследуемой или пустой оценкой определена итерационная процедура коррекции оценки  $\text{Kor}(m)$ :

- (1) в соответствии с правилами формирования наследованной оценки [формулы (3) и (4)] формируется оценка  $O(m) = \langle s, d, r \rangle$ ;
- (2) для каждого объекта  $m_i$ , непосредственно связанного с  $m$  и имеющего уровень оценки ниже  $r$ , выполняется  $\text{Kor}(m_i)$ ;
- (3) при достижении предельного уровня оценки  $\rho$  процедура прерывается, что позволяет не создавать оценок с заведомо низкой достоверностью.

Изменение оценок узлов семантической сети может быть инициировано введением в сеть нового объекта или назначением новой пользовательской оценки существующему объекту. В первом случае новый объект изначально получает пустую оценку. Далее для всех связанных объектов, имеющих наследованные оценки, выполняется процедура  $\text{Kor}$ . Во втором случае для оцененного объекта формируется назначенная оценка в соответствии с формулами (1) и (2). Далее для всех связанных объектов, имеющих наследованные и пустые оценки, выполняется процедура  $\text{Kor}$ .

### 3 Описание модели

Для изучения динамики семантической сети технологии ПКИИ была разработана и обоснована имитационная модель [6], опирающаяся на принципы графодинамики, позволяющие строить граф на основе строго определенных правил, соответствующих практике, и на модель Барабаши–Альберт (алгоритм генерации случайных безмасштабных сетей).

Графодинамический подход к имитационному моделированию процессов развития сетевых структур [7] заключается во введении набора стандартных операций над графами, отражающих изменения конфигурации реальных сетевых структур, и построении конечного автомата трансформирования графов на заданном временному горизонте. Состояниям автомата соответствуют действия, а переходы обусловлены событиями, описываемыми индикаторными логическими формулами. События определяются истечением конкретного времени с момента запуска предыдущего действия и внешними воздействиями на сетевую структуру.

Поскольку построение конечного автомата трансформирования графов для распределенной технологии ПКИИ не представляется возможным, была использована дополнительная концепция, определяющая вероятности стандартных изменений и способы соединения добавляемых блоков, отражающая неравномерность топологической структуры сети, соответствующей структуре семантики предметной области, образом которой является сеть. Характер этой неравномерности определяется тем, что семантическая сеть технологии ПКИИ по своей природе относится к когнитивным сетям [8], которые, как показывают исследования [9], являются безмасштабными. Такие сети имеют медленно спадающее распределение узлов по числу связей, и в этих сетях хабы (узлы с большим числом связей) составляют заметную долю от всех узлов. Для них характерна асимптотическая зависимость распределения узлов по числу связей в виде

степенного закона  $P(q) \sim q^{-\gamma}$  [10]. Для моделирования такой структуры сети была использована модель Барабаши–Альберт — алгоритм генерации случайных безмасштабных сетей. В основу построения графа Барабаши–Альберт положен принцип предпочтительного присоединения: чем больше вершина имеет связей, тем выше вероятность ее выбора для присоединения к ней новой вершины. Реалистичность моделей, построенных по этому принципу, была проверена на практике [11].

Далее модель была расширена [12, 13] за счет включения в нее механизмов реализации ролей пользователей и различных форм их активности. Как показано в [13] на основе анализа примеров сообществ с разными формами активности, распределение участников по формам активности соответствует усеченному нормальному закону с модой в первой половине всего интервала форм активности. В [5] обосновано отнесение назначения фактам оценки их достоверности к третьей, наиболее редкой форме активности пользователей. Это, в принципе, позволяет использовать модель для исследования функционирования механизмов автоматизированной атрибуции фактов признаком достоверности, но в ее модели данных отсутствует поддержка сущностей «документ» и «автор документа» и их связи с конкретно-историческим фактом.

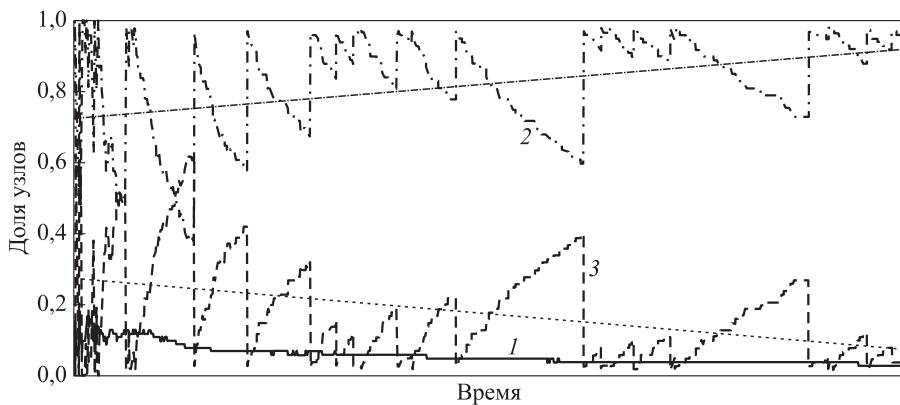
Соответствующая доработка модели была выполнена. Новые сущности были включены в модель данных, а для построения алгоритма работы с ними были проанализированы популярные сетевые библиотеки «Библиотека Мошкова» и «Флибустан». Анализ показал, что документы по объему и авторы по числу документов распределены в соответствии со степенным законом с показателем в диапазоне 2–3, что и следовало ожидать для когнитивных безмасштабных сетей. В целях моделирования было принято допущение, что количество фактов в документе примерно пропорционально его объему.

## 4 Результаты моделирования

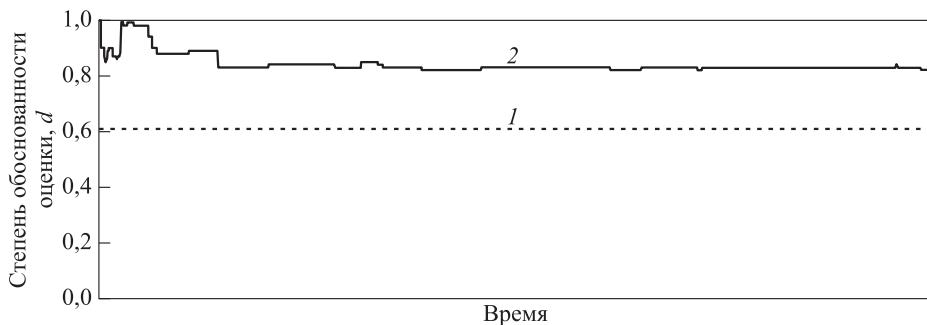
Было проведено несколько прогонов модели для различных значений предельного уровня оценки  $\rho$ . Типичные результаты приведены на рис. 1 и 2.

На рис. 1 видно, что количество узлов с назначенной оценкой падает, несмотря на то что скорость добавления пользовательских оценок постоянна. Это объясняется тем, что в безмасштабных сетях вероятность добавления оценки к узлу, уже имеющему оценки, выше, чем к узлу без оценок.

Графики количества узлов с наследованной оценкой и, соответственно, неоцененных узлов испытывают значительные колебания. Это объясняется тем, что в безмасштабных сетях хабы (узлы с большим числом связей) редки, но связаны при этом с огромным числом узлов сети. Соответственно, добавление пользовательской оценки к обычному узлу не приводит к значительному изменению количества наследованных оценок. Но добавление пользовательской оценки к хабу приводит к резкому изменению количества наследованных оценок. Линии тренда позволяют оценить общую тенденцию динамики оценок. Количество узлов



**Рис. 1** Изменение количества узлов с назначенной оценкой (1), наследованной оценкой (2) и пустой оценкой (3) по мере развития семантической сети



**Рис. 2** Снижение степени обоснованности оценки  $d$  для наследованных оценок: 1 —  $d$  минимальное; 2 —  $d$  среднее

с наследованной оценкой имеет тенденцию к росту, а количество неоцененных узлов падает.

Рисунок 2 показывает динамику снижения степени обоснованности  $d$  для наследованных оценок по отношению к исходным пользовательским оценкам. Моделирование показало, что при задании для пользовательских оценок степени обоснованности  $d \equiv 1,0$  для различных значений  $\rho$  ни для одного узла с наследованной оценкой степень обоснованности не падала ниже 0,61, а средняя степень обоснованности для всех узлов с наследованной оценкой установилась на уровне 0,82 для всего периода моделирования. При этом вне зависимости от величины предельного уровня оценок  $\rho$  уровень оценки  $r$  всегда оставался в пределах 6, а значит, глубина наследования определяется исключительно топологией семантической сети и ее дополнительное ограничение не требуется.

Расчет показателя эффективности  $\eta$ , представляющего собой отношение объемов наследованных и пустых оценок при заданном объеме пользовательских оценок, показал, что  $\eta(t) > 5,1713$  и имеет устойчивую тенденцию к росту.

## 5 Выводы

Моделирование показало: несмотря на то что в семантической сети только незначительное число фактов (несколько процентов) будет снабжено пользовательскими оценками, предложенный механизм позволит автоматически оценить большую часть фактов при сохранении высокой степени обоснованности оценок.

Эффективность механизма высока, и ограничение предельного уровня оценок не требуется.

Предложенный механизм автоматизации оценки достоверности фактов существенно дополняет и развивает технологию ПКИИ, ориентированной на широкий круг не являющихся профессиональными историками и биографами пользователей, что очень актуально в связи со все возрастающим общественным интересом к «частной», семейной истории.

## Литература

1. Адамович И. М., Волков О. И. Технология распределенного автоматизированного анализа исторических текстов // Системы и средства информатики, 2016. Т. 26. № 3. С. 148–161. doi: 10.14357/08696527160311.
2. Адамович И. М., Волков О. И. Единая технология поддержки конкретно-исторических исследований // Системы и средства информатики, 2019. Т. 29. № 1. С. 194–205. doi: 10.14357/08696527190116.
3. Адамович И. М., Волков О. И. Принципы организации данных для технологии поддержки конкретно-исторических исследований // Системы и средства информатики, 2019. Т. 29. № 2. С. 161–171. doi: 10.14357/08696527190214.
4. Адамович И. М., Волков О. И. Иерархическая форма представления биографического факта // Системы и средства информатики, 2016. Т. 26. № 2. С. 108–122. doi: 10.14357/08696527160207.
5. Адамович И. М., Волков О. И. Автоматизированная оценка достоверности конкретно-исторических фактов // Системы и средства информатики, 2020. Т. 30. № 1. С. 160–170. doi: 10.14357/08696527200114.
6. Адамович И. М., Волков О. И. Об одном подходе к моделированию процесса развития семантической сети // Системы и средства информатики, 2017. Т. 27. № 2. С. 143–154. doi: 10.14357/08696527170212.
7. Юдицкий С. А. Графодинамическое имитационное моделирование развития сетевых структур // Управление большими системами, 2011. Вып. 33. С. 21–34.
8. Bak P. How nature works. The science of self-organized criticality. — Oxford: Oxford University Press, 1997. 62 р.
9. Евин И. А., Кобляков А. А., Савриков Д. В., Шувалов Н. Д. Когнитивные сети // Компьютерные исследования и моделирование, 2011. Т. 3. № 3. С. 231–239.

10. Евин И. А. Введение в теорию сложных сетей // Компьютерные исследования и моделирование, 2010. Т. 2. № 2. С. 121–141.
11. Бадрызлов В. А. Принципы генерации случайных графов для моделирования сети Интернет // Омский научный вестник, 2014. № 3(133). С. 204–208.
12. Адамович И. М., Волков О. И. Влияние ошибок пользователей на динамику качества семантической сети // Системы и средства информатики, 2017. Т. 27. № 4. С. 150–163. doi: 10.14357/08696527170412.
13. Адамович И. М., Волков О. И. Модель процесса коррекции ошибок в семантической сети // Системы и средства информатики, 2018. Т. 28. № 1. С. 65–76. doi: 10.14357/08696527180105.

*Поступила в редакцию 13.01.20*

---

## MODEL OF AUTOMATED ESTIMATING THE TRUSTWORTHINESS OF CONCRETE-HISTORICAL FACTS SYSTEM FUNCTIONING

*I. M. Adamovich and O. I. Volkov*

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** The article continues the series of works devoted to the technology of concrete historical research supporting. The technology is based on the principles of co-creation and crowdsourcing and is designed for a wide range of users which are not professional historians and biographers. This article is devoted to the description and rationale of the approach to the modeling of the mechanism of automated estimating the trustworthiness of processed information which is included in the technology for its further development. The proposed approach is to appropriately modify the semantic net model which is based on the concepts of graphodynamics and the Barabasi–Albert model and the support for user activity simulation mechanism. Using this model, the experimental tests of the effectiveness of iterative algorithms for the implementation of a new mechanism for automated estimating the trustworthiness of concrete-historical information were carried out and the influence of the iteration depth constraint on the average trustworthiness of the calculated estimates was generally evaluated.

**Keywords:** modeling; distributed technology; reliability of information; historical-biographical fact; automated procedure

**DOI:** 10.14357/08696527200216

### References

1. Adamovich, I. M., and O. I. Volkov. 2016. Tekhnologiya raspredelennogo avtomatizirovannogo analiza istoricheskikh tekstov [The distributed automated technology of

- historical texts analysis]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 3(26):148–161. doi: 10.14357/08696527160311.
- 2. Adamovich, I. M., and O. I. Volkov. 2019. Edinaya tekhnologiya podderzhki konkretno-istoricheskikh issledovaniy [Unified technology of concrete historical research support]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 1(29):194–205. doi: 10.14357/08696527190116.
  - 3. Adamovich, I. M., and O. I. Volkov. 2019. Printsipy organizatsii dannykh dlya tekhnologii podderzhki konkretno-istoricheskikh issledovaniy [The principles of data organization for the technology of concrete historical research support]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 2(29):161–171. doi: 10.14357/08696527190214.
  - 4. Adamovich, I. M., and O. I. Volkov. 2016. Ierarkhicheskaya forma predstavleniya biograficheskogo fakta [Hierarchical format of biographical fact]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 2(26):108–122. doi: 10.14357/08696527160207.
  - 5. Adamovich, I. M., and O. I. Volkov. 2020. Avtomatizirovannaya otsenka dostovernosti konkretno-istoricheskikh faktov [Automated estimating the trustworthiness of concrete-historical facts]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(1):160–170. doi: 10.14357/08696527200114.
  - 6. Adamovich, I. M., and O. I. Volkov. 2017. Ob odnom podkhode k modelirovaniyu protsessa razvitiya semanticheskoy seti [On some approach to modeling of semantic net evolution]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 2(27):143–154. doi: 10.14357/08696527170212.
  - 7. Yuditskiy, S. A. 2011. Grafodinamicheskoe imitatsionnoe modelirovanie razvitiya setevykh struktur [Graphodynamic simulation modeling of network structures evolution]. *Upravlenie bol'shimi sistemami* [Large-Scale Systems Control] 33:21–34.
  - 8. Bak, P. 1997. *How nature works. The science of self-organized criticality*. Oxford: Oxford University Press. 62 p.
  - 9. Yevin, I. A., A. A. Koblyakov, D. V. Savricov, and N. D. Shuvalov. 2011. Kognitivnye seti [Cognitive networks]. *Komp'yuternye issledovaniya i modelirovanie* [Computer Research and Modeling] 3(3):231–239.
  - 10. Yevin, I. A. 2010. Vvedenie v teoriyu slozhnykh setey [Introduction to the theory of complex networks]. *Komp'yuternye issledovaniya i modelirovanie* [Computer Research and Modeling] 2(2):121–141.
  - 11. Badryzlov, V. A. 2014. Printsipy generatsii sluchaynykh grafov dlya modelirovaniya seti Internet [The principles of generation of random graphs for simulation of the Internet]. *Omskiy nauchnyy vestnik* [Omsk Scientific Bulletin] 3(133):204–208.
  - 12. Adamovich, I. M., and O. I. Volkov. 2017. Vliyanie oshibok pol'zovateley na dinamiku kachestva semanticheskoy seti [The influence of user errors on the semantic network quality dynamics]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 4(27):150–163. doi: 10.14357/08696527170412.
  - 13. Adamovich, I. M., and O. I. Volkov. 2018. Model' protsessa korrektsii oshibok v semanticheskoy seti [The model of semantic net error correction process]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 1(28):65–76. doi: 10.14357/08696527180105.

Received January 13, 2020

## **Contributors**

**Adamovich Igor M.** (b. 1934)— Candidate of Science (PhD) in technology, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; Adam@amsd.com

**Volkov Oleg I.** (b. 1964) — leading programmer, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; Volkov@amsd.com

## ПЕРСОНАЛЬНЫЕ И КОРПОРАТИВНЫЕ ЭЛЕКТРОННЫЕ БАНКИ В ТЕХНОЛОГИЯХ ТОВАРНО-ДЕНЕЖНОГО ОБРАЩЕНИЯ

*A. V. Ильин<sup>1</sup>, В. Д. Ильин<sup>2</sup>*

**Аннотация:** Представлены основы методологического обеспечения технологий товарно-денежного обращения в цифровой среде (ТДО-технологий), реализуемых посредством персональных электронных банков (ПЭБов), принадлежащих физическим лицам, и корпоративных электронных банков (КЭБов), владельцами которых выступают юридические лица. Персональные и корпоративные электронные банки представляют собой специализированные системы искусственного интеллекта, реализуемые на основе переносных компьютерных устройств (смартфонов, планшетов) и стационарных компьютеров. Прикладной смысл каждой ТДО-технологии определен системой правил, направляющих и контролирующих действия участников договора о товарно-денежном обращении в цифровой среде (ТДО-договора). Правила, реализуемые в ТДО-технологиях, основаны на предположении, что владельцам ПЭБов и КЭБов законом предоставлено право кредитования, реализация которого контролируется банком-регулятором. Технологии товарно-денежного обращения позволяют кредиторам использовать только собственные денежные средства. Правила исполнения каждого ТДО-договора контролируются программно во взаимодействии с цифровыми двойниками, отслеживающими события, определенные в договоре. Технологии товарно-денежного обращения ориентированы на повышение экономической безопасности сделок и уменьшение влияния факторов, вызывающих снижение товарной емкости денег.

**Ключевые слова:** технология товарно-денежного обращения в цифровой среде (ТДО-технология); персональный электронный банк (ПЭБ); корпоративный электронный банк (КЭБ)

**DOI:** 10.14357/08696527200217

### 1 Введение

В нынешней кредитной системе, концептуальные основы которой были определены в [1–3], центральный банк выполняет функции кредитора последней инстанции для кредитных организаций банковского сектора (включающего коммерческие, сберегательные и ипотечные банки), страхового сектора (страховые

---

<sup>1</sup>Государственный научно-исследовательский институт авиационных систем,  
ilyin@res-plan.com

<sup>2</sup>Вычислительный центр им. А. А. Дородницына Федерального исследовательского центра «Информатика и управление» Российской академии наук, vdilyin@yandex.ru

компании и пенсионные фонды) и специализированных небанковских учреждений. Банкам законом (о банках и банковской деятельности<sup>1</sup>) предоставлено исключительное право кредитной деятельности, в которой используются денежные средства физических и юридических лиц.

Современные банки, продолжая погоню за прибылью, упорно стремятся сохранить возможность слабо ограниченного распоряжения денежными средствами физических и юридических лиц, используя их в процессах кредитования, торговли валютой и другой деятельности, уменьшающей товарную емкость денег. Несмотря на существующие технологические возможности обеспечить экономическую безопасность банковских операций, продолжается способствующее банковскому мошенничеству выполнение операций над банковскими счетами без обязательных онлайн-подтверждений допустимости операций, предварительно полученных от владельцев счетов [4, 5]. Но логика прогрессирующей цифровизации различных видов деятельности неотвратимо приближает дни, когда привычной станет увеличивающая экономическую безопасность сделок *технология назначенных платежей* [6], а смартфоны и планшеты со специальным программным обеспечением будут выполнять роль *персональных электронных банков* (впервые предложенных в [7, с. 126]), и *корпоративных электронных банков* [4, 5].

**Запись формул и выделение фрагментов текста.** Для выделения определений, замечаний и примеров используются средства языка TSM-комплекса (TSM: textual symbolic modeling), разработанного для формализованного описания текстовых моделей<sup>2</sup>.

В статье применены следующие средства выделения фрагментов текста:

$\square$  (фрагмент описания)  $\square \approx$  утверждение (определение, аксиома и др.) (здесь и далее символ  $\approx$  заменяет слово «означает»);

$\diamond$  (фрагмент описания)  $\diamond \approx$  замечание;

$\circlearrowleft$  (фрагмент описания)  $\circlearrowleft \approx$  пример.

Курсивом выделены первые вхождения названий понятий и фрагменты описания, к которым авторы хотят привлечь внимание.

**Обсуждаемые результаты.** В статье представлена часть результатов методологического обеспечения технологий товарно-денежного обращения в цифровой среде с использованием персональных и корпоративных электронных банков. Результаты получены при выполнении научно-исследовательской работы «Моделирование социальных, экономических и экологических процессов» (№ 0063-2016-0005), выполняемой в соответствии с государственным заданием ФАНО России для Федерального исследовательского центра «Информатика и управление» РАН.

---

<sup>1</sup>Закон РФ «О банках и банковской деятельности». [https://cbr.ru/content/document/file/36339/law\\_banks.pdf](https://cbr.ru/content/document/file/36339/law_banks.pdf).

<sup>2</sup>Ильин В. Д. Символьное моделирование // Большая российская энциклопедия, 2019. [Электронный ресурс]. [http://dev.bigenc.ru/technology\\_and\\_technique/4010980](http://dev.bigenc.ru/technology_and_technique/4010980).

◊ Материал статьи адресован прежде всего занимающимся методологическим обеспечением разработок *информационных технологий цифровой экономики* [8–12]. ◊

## 2 Основы методологического обеспечения технологий товарно-денежного обращения в цифровой среде

Рассматриваемые здесь ТДО-технологии представлены *технологией назначенных платежей в среде цифровых двойников* [6], *технологией платежного кредитования и технологией товарного кредитования*.

Технологии кредитования рассчитаны на кредитные отношения двух видов: заказчик кредитует поставщика товара (*платежный кредит*) и поставщик товара кредитует заказчика (*товарный кредит*). Технологии товарных кредитов основаны на *технологии электронной долговой торговли* [13].

Платежные средства, предоставляемые в кредит, могут быть любыми [из списка допустимых в кредитных отношениях конкретного типа (○ электронные российские рубли, электронные юани и др. ○)]. ◊ Обязательным требованием является использование платежных средств, принадлежащих кредитору (*собственных платежных средств*). ◊

Товары, продаваемые в долг, тоже могут быть любыми [из списка допустимых в кредитных отношениях конкретного типа (○ транспортные средства, образовательные услуги и др. ○)].

◊ Участники ТДО-сделки (ТДО-участники) — владельцы ПЭБов и / или КЭБов, наделенные правом совершать ТДО-сделки. ◊

Программные средства ТДО-технологий входят в состав *программного обеспечения нормализованного банкинга* (ПО НБ), ядром которого является специальная система искусственного интеллекта, которая посредством цифровых двойников (англ. digital twins) [14] отслеживает события, определенные в ТДО-договоре. Программное обеспечение нормализованного банкинга установлено на ПЭБах и КЭБах участников договора и на компьютерах банков-провайдеров, которые обслуживают участников ТДО-сделок.

### 2.1 Договор о товарно-денежном обращении в цифровой среде

В реализации ТДО-договор — это *интеллектуальная система, содержащая описание ТДО-сделки и правила ее исполнения ТДО-участниками*. В процессе контроля исполнения договора система обменивается данными с созданными ею обучаемыми цифровыми двойниками, отслеживающими события, определенные в ТДО-договоре. Среди опциональных функций системы — *ситуационное планирование расходов* [15] в ходе выполнения сделок.

Формально правило — это выражение одного из двух видов: либо  $D_1 \rightarrow D_2$ , либо  $D \rightarrow I$ . Здесь  $D$ ,  $D_1$  и  $D_2$  — описания произвольных событий, принад-

лежащих множеству контролируемых событий, определенных в ТДО-договоре, а  $I$  — инструкция, определяющая действия в случае наступления события  $D$ .

## 2.2 Представление имущественных статусов участников ТДО-сделки

□ Система имущественных статусов участников сделки (ИС-система) — это система цифрового представления денежной и неденежной составляющих имущественных статусов участников ТДО-сделки. Денежная составляющая выражена значениями сумм в разделах *поливалютных счетов участников сделок* (ИС-счетов). Неденежная — электронными документами, подтверждающими право собственности на недвижимость, транспорт и другое имущество (которое при необходимости может рассматриваться как залоговое). □

**ИС-счет.** □ ИС-счет — уникальный поливалютный счет, размещенный в ПЭБе физического лица или КЭБе юридического лица. Представляет собой унифицированный электронный документ, состоящий из *валютных частей* [4, 5]. □ ◊ Программно реализуемые записи в файлах ИС-счетов имеют право инициировать только владельцы ИС-счетов, совершающие сделку ◊. Запросы на обслуживание сделки (удостоверение состояния ИС-счета и сохранение обновленной копии ИС-счета) получают банки-провайдеры (см. подразд. 2.3), сервисами которых пользуются совершающие сделку.

**Аксиома допустимости операции над ИС-счетом.** □ Никакое изменение состояния ИС-счета не может быть осуществлено без документируемого зашифрованного подтверждения владельца ИС-счета<sup>1</sup>. □

## 2.3 Банки в реализации договоров о товарно-денежном обращении в цифровой среде

Для реализации ТДО-договоров применяются ПЭБы, КЭБы, банки-провайдеры и банк-регулятор (центральный банк).

Предполагается, что *банк-регулятор* располагает сетью серверов, размещенных на территории страны, под юрисдикцией которой находится экономическая система. *Банки-провайдеры* — предприятия, учрежденные юридическими лицами (или объединениями юридических лиц, объединениями физических лиц, объединениями юридических и физических лиц), занимающимися производством товаров и/или их реализацией. Банк-провайдер располагает объединенными в сеть серверами, предназначенными для обслуживания запросов от ПЭБов и КЭБов клиентов и взаимодействия с серверами банка-регулятора.

**Персональные и корпоративные электронные банки.** Как правило, это достаточно производительные переносные компьютерные устройства (○ планшеты, смартфоны ○), наделенные надежно защищенными коммуникационными средствами. В ПЭБах и КЭБах в зашифрованной форме хранятся оригиналы ИС-счетов и ТДО-договоры.

---

<sup>1</sup> Эта аксиома нарушается нынешними банками.

## 2.4 Платежное и товарное кредитование с помощью персональных и корпоративных электронных банков

□ *Платежное кредитование* — программно контролируемая частичная или полная предоплата заказанных товаров (○ оборудования, образовательных услуг и др. ○), правила реализации которой жестко связаны с правилами исполнения заказа. График перечисления платежного кредита содержится в ТДО-договоре и контролируется в процессе реализации договора. □

□ *Товарное кредитование* — программно контролируемая поставка заказанных товаров с частичной или полной отсрочкой оплаты, правила реализации которой жестко связаны с правилами поставки. □ Основано на технологии электронной долговой торговли [13]. Обеспечивает оформление отсроченной части оплаты товара как *долга покупателя (заказчика) продавцу (поставщику), имеющему КЭБ или ПЭБ*. Неоплаченная часть стоимости товара оформляется как долг, график выплаты которого фиксируется в ТДО-договоре (где указываются штрафные санкции за нарушение графика выплат долга и ненадлежащее качество проданного (поставленного) товара).

Применительно к долгам, образовавшимся в результате продажи (поставки) ряда товаров, могут действовать правила досрочного возврата долгов продавцам (поставщикам) из средств банка-регулятора. Делается это по запросам кредиторов в *долговой отдел банка-регулятора*, направляемых с использованием онлайн-сервиса. При положительном решении на ИС-счет кредитора переводится сумма долга (полностью или частично), а после этого покупатель (заказчик) выплачивает эту сумму не поставщику, а долговому отделу банка-регулятора [13]. Возможность такой схемы возврата долга указывается в ТДО-договоре.

## 2.5 Технология товарно-денежного обращения долговой денежной эмиссии

□ *Денежная эмиссия* выполняется банком-регулятором только тогда, когда сумма долгов, возвращенных долговому отделу, меньше очередной долговой суммы, запрошенной для возврата поставщику-кредитору. Эмитируемая сумма равна разности запрошенной долговой суммы и суммы на счету долгового отдела [13]. □

## 3 Заключение

1. На современном этапе развития ТДО в цифровой среде значительная роль принадлежит передовой *технологии назначенных платежей* [6] и *технологиям платежного и товарного кредита*, реализуемым посредством ПЭБов и КЭБов. Эти технологии обеспечивают высокий уровень экономической безопасности и реализуемости сделок в соответствии с ТДО-договорами.

2. Поскольку технологии платежного и товарного кредита позволяют кредиторам, имеющим ПЭБы и / или КЭБы, использовать только собственные средства, реализация этих технологий не связана с уменьшением товарной емкости денег [4, 5].
3. Долговая денежная эмиссия служит средством регулирования общей суммы денег в экономической системе [5, 6, 13].
4. В реализации обсуждаемых в статье технологий ПЭБы и КЭБы рассматриваются как системы искусственного интеллекта, наделенные способностью обучаться в процессах выполнения ТДО-договоров. Данные о контролируемых событиях, описанных в договорах, поставляют цифровые двойники.

## **Литература**

1. Fisher I. The purchasing power of money: Its determination and relation to credit, interest, and crises. — New York, NY, USA: The Macmillan Co., 1922. 515 p.
2. Keynes J. The general theory of employment, interest and money. — Macmillan Cambridge University Press, 1935. 472 p.
3. Friedman M., Bordo M. The optimum quantity of money. — New Brunswick, NJ, USA: Transaction Publishers, 2006. 296 p.
4. Ilyin A. V., Ilyin V. D. Towards a normalized economic mechanism based on E-services // Agris on-line Papers in Economics and Informatics, 2014. Vol. 6. Iss. 3. P. 39–49.
5. Ilyin A. V., Ilyin V. D. The normalized economic mechanism in the digital environment // Int. J. Open Information Technologies, 2019. Vol. 7. Iss. 12. P. 77–83.
6. Ильин В. Д. Технология назначенных платежей в среде цифровых двойников // Системы и средства информатики, 2018. Т. 28. № 3. С. 221–229.
7. Ильин В. Д. Модель нормализованной экономики (НЭк-модель): основы концепции // Управление большими системами, 2009. Вып. 25. С. 116–138.
8. Tapscott D. The digital economy: promise and peril in the age of networked intelligence. — New York, NY, USA: McGraw-Hill, 1996. 342 p.
9. Christensen C. M. The innovator's dilemma: When new technologies cause great firms to fail. — Boston, MA, USA: Harvard Business School Press, 1997. 288 p.
10. The new digital economy: How it will transform business. — Oxford Economics, 2015. 34 p.
11. G20 digital economy development and cooperation initiative // G20 Summit, 2016. <http://en.kremlin.ru/supplement/5111>.
12. Цифровая экономика Российской Федерации: Программа, утвержденная распоряжением Правительства Российской Федерации от 28 июля 2017 г. № 1632-р.
13. Ilyin A. V., Ilyin V. D. E-trade with direct lending and normalized money // Agris on-line Papers in Economics and Informatics, 2015. Vol. 7. Iss. 4. P. 57–64.
14. The Digital Twin // General Electric, 2018. [https://www.ge.com/digital/sites/default/files/The-Digital-Twin\\_Compressing-Time-to-Value-for-Digital-Industrial-Companies.pdf](https://www.ge.com/digital/sites/default/files/The-Digital-Twin_Compressing-Time-to-Value-for-Digital-Industrial-Companies.pdf).

15. Ilyin A. V., Ilyin V. D. Solving situationally definable linear problems of resource planning: A review of updated technology // Информационные технологии и вычислительные системы, 2019. № 3. С. 99–106.

Поступила в редакцию 05.03.20

---

## PERSONAL AND CORPORATE E-BANKS IN COMMODITY-MONEY CIRCULATION TECHNOLOGIES

A. V. Ilyin<sup>1</sup> and V. D. Ilyin<sup>2</sup>

<sup>1</sup>State Research Institute of Aviation Systems, 7 Viktorenko Str., Moscow 125319, Russian Federation

<sup>2</sup>A. A. Dorodnicyn Computing Center, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 40 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** The review presents the basics of methodological support for technologies of commodity-money circulation in the digital environment (CMC-technologies) implemented by means of personal e-banks (PEBs) owned by individuals and corporate e-banks (CEBs) owned by legal entities. Personal and corporate e-banks are the specialized artificial intelligence systems implemented on the basis of portable computer devices (smartphones, tablets) and stationary computers. The applied meaning of each CMC-technology is defined by a system of rules that direct and control actions of the parties to the contract on commodity-money circulation in the digital environment (CMC-agreement). The rules implemented in CMC-technologies are based on the assumption that the owners of PEBs and CEBs are granted the right to credit by law, the implementation of which is controlled by the bank-regulator. The CMC-technologies allow the lenders to use only their own funds. The rules for execution of each CMC-agreement are controlled by software interacting with the digital twins that track events defined in the agreement. The CMC-technologies are aimed at improving the economic security of deals and reducing the influence of factors causing a decrease in the commodity capacity of money.

**Keywords:** technology of commodity-money circulation in the digital environment; personal e-bank; corporate e-bank; contract on commodity-money circulation in the digital environment

**DOI:** 10.14357/08696527200217

## References

1. Fisher, I. 1922. *The purchasing power of money. Its determination and relation to credit, interest, and crises*. New York, NY: The Macmillan Co. 515 p.
2. Keynes, J. 1935. *The general theory of employment, interest and money*. Macmillan Cambridge University Press. 472 p.

3. Friedman, M., and M. Bordo. 2006. *The optimum quantity of money*. New Brunswick, NJ: Transaction Publs. 296 p.
4. Ilyin, A. V., and V. D. Ilyin. 2014. Towards a normalized economic mechanism based on E-services. *Agris on-line Papers in Economics and Informatics* 6(3):39–49.
5. Ilyin, A. V., and V. D. Ilyin. 2019. The normalized economic mechanism in the digital environment. *Int. J. Open Information Technologies* 7(12):77–83.
6. Ilyin, V. D. 2018. Tekhnologiya naznachennykh platezhey v srede tsifrovych dvoynikov [Designated payments technology in digital twins environment]. *Systemy i Sredstva Informatiki — Systems and Means of Informatics* 28(3):221–229.
7. Ilyin, V. D. 2009. Model' normalizovannoy ekonomiki (NEk-model'): osnovy kontseptsii [The model of normalized economics (NEc-model): Basics of framework]. *Upravleniye bol'shimi sistemami* [Large-Scale Systems Control] 25:116–138.
8. Tapscott, D. 1996. *The digital economy: Promise and peril in the age of networked intelligence*. New York, NY: McGraw-Hill. 342 p.
9. Christensen, C. M. 1997. *The innovator's dilemma: When new technologies cause great firms to fail*. Boston, MA: Harvard Business School Press. 288 p.
10. The new digital economy: How it will transform business. 2015. Oxford Economics. 34 p. Available at: <http://www.pwc.com/mt/en/publications/assets/the-new-digital-economy.pdf> (accessed March 5, 2020).
11. G20 digital economy development and cooperation initiative. 2016. G20 Summit. Available at: <http://en.kremlin.ru/supplement/5111> (accessed March 5, 2020).
12. Tsifrovaya ekonomika Rossiyskoy Federatsii: Programma, utverzhдennaya rasporyazheniem Pravitel'stva Rossiyskoy Federatsii [Digital economy of the Russian Federation: Program Approved by Order No. 1632-r dated July 28, 2017 of the Government of the Russian Federation]. Available at: <http://d-russia.ru/wpcontent/uploads/2017/07/programma-tsifrov-econ.pdf> (accessed March 5, 2020).
13. Ilyin, A. V., and V. D. Ilyin. 2015. E-trade with direct lending and normalized money. *Agris on-line Papers in Economics and Informatics* 7(4):57–64.
14. The Digital Twin. Available at: [https://www.ge.com/digital/sites/default/files/The-Digital-Twin\\_Compressing-Time-to-Value-for-Digital-Industrial-Companies.pdf](https://www.ge.com/digital/sites/default/files/The-Digital-Twin_Compressing-Time-to-Value-for-Digital-Industrial-Companies.pdf) (accessed March 5, 2020).
15. Ilyin, A. V., and V. D. Ilyin. 2019. Solving situationally definable linear problems of resource planning: A review of updated technology. *Informatsionnye tekhnologii i vychislitel'nyye sistemy* [J. Information Technologies Computing Systems] 3:99–106.

Received March 5, 2020

## Contributors

**Ilyin Alexander V.** (b. 1975) — Candidate of Science (PhD) in technology, leading scientist, State Research Institute of Aviation Systems, 7 Viktorenko Str., Moscow 125319, Russian Federation; [ilyin@res-plan.com](mailto:ilyin@res-plan.com)

**Ilyin Vladimir D.** (b. 1937) — Doctor of Science in technology, professor, leading scientist, A. A. Dorodnicyn Computing Center, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences, 40 Vavilov Str., Moscow 119333, Russian Federation; [vdilyin@yandex.ru](mailto:vdilyin@yandex.ru)

---

# ИСТОРИЯ. ПАМЯТНЫЕ ДАТЫ. СОБЫТИЯ

---

## У ИСТОКОВ РОССИЙСКОЙ ИНФОРМАТИКИ



Игорь Александрович Мизин  
(12.04.1935–08.09.1999)

Данный очерк посвящен 85-летнему юбилею академика И. А. Мизина, выдающегося ученого в области телекоммуникаций, информатики и информационных технологий. Рассмотрены основные научные и практические результаты, полученные под руководством И. А. Мизина на протяжении 40 лет трудовой деятельности: 30 лет в НИИ автоматической аппаратуры и 10 лет директором Института проблем информатики РАН. Показана роль И. А. Мизина в обосновании информационных технологий для создания и развития информационных сетей в интересах органов государственной власти с учетом требований по защите информации.

### 1 Введение

Двенадцатого апреля 2020 г. исполнилось 85 лет со дня рождения выдающегося советского и российского ученого академика РАН И. А. Мизина. Он ушел из жизни в возрасте всего 64 лет, но внес огромный вклад как в теорию

отечественных информационно-телекоммуникационных систем, так и в практику их разработки и внедрения.

В 1952 г. И. А. Мизин поступил в Военно-воздушную академию имени Н. Е. Жуковского, после окончания которой в течение 30 лет (1959–1989) трудился в НИИ автоматической аппаратуры (ныне — АО НИИ АА им. академика В. С. Семенихина). Последовательно прошел ступени карьерного роста от инженера до начальника научно-тематического центра систем передачи данных — заместителя директора по научной работе [1]. В 1966 г. защитил кандидатскую, а в 1972 г. — докторскую диссертацию по проблематике информационных сетей территориальных автоматизированных систем управления. В 1984 г. избран членом-корреспондентом АН СССР, в 1997 г.— академиком РАН.

На протяжении последних 10 лет жизни (1989–1999) И. А. Мизин — директор Института проблем информатики РАН (ИПИ РАН), генеральный конструктор работ по модернизации информационных сетей органов государственной власти, генеральный конструктор АСУ Вооруженных Сил РФ.

И. А. Мизин — автор более 190 научных работ, в том числе 12 монографий и 12 изобретений. Научные работы И. А. Мизина отличались сочетанием активного использования математического аппарата с инженерной интуицией и практической направленностью. Наиболее известными работами являются [2–9].

## 2 Создатель системы обмена данными

В первые годы в НИИ АА Игорь Александрович Мизин занимался исследованиями и разработкой комплексов повышения достоверности передачи информации по каналам связи различной физической природы для АСУ разного назначения [10, 11]. В 1967 г., когда НИИ АА, который возглавлял В. С. Семенихин (с 1972 г. академик АН СССР), был определен головной организацией по созданию автоматизированной командной системы боевого управления (КСБУ), молодой инженер И. А. Мизин назначается главным конструктором системы обмена данными (СОД) [10, 12]. Особенность этой разработки заключалась в необходимости обеспечения очень высоких вероятностно-временных характеристик доведения информации до объектов, распределенных по территории всей страны. Генеральным конструктором В. С. Семенихиным совместно с И. А. Мизиным было обосновано решение создавать СОД на основе метода коммутации сообщений [2, 10–12]. Такое решение в дальнейшем полностью себя оправдало, так как позволило создать СОД как базовую систему для интеграции многих автоматизированных систем управления (АСУ) военного назначения на единых принципах.

И. А. Мизин совместно с Л. С. Уринсоном и Г. К. Храмешиним обосновал принципиально новые технические решения по построению СОД на основе центров коммутации сообщений (ЦКС), размещаемых по всей территории СССР [2, 4, 5]. Для передачи информации через СОД с требуемыми характе-



Ближайшие соратники Главного конструктора И. А. Мизина, 1977 г.: сидят — Г. К. Храмешин, Н. Я. Матюхин, И. А. Мизин, Ш. Ш. Чипашвили, П. Д. Чеботарев; стоят — С. В. Назаров, С. К. Муравьев, С. И. Кононов, В. Н. Березин, А. В. Тамошинский

ристиками была разработана единая система протоколов обмена данными (ЕС-ПИО). Эта разработка на несколько лет опередила известные рекомендации по архитектуре открытых вычислительных систем, которые появились в 1980-е гг.: ISO / OSI, SNA (IBM), DEC и др. [5].

Решения, положенные в основу построения СОД, оказались эффективными и перспективными, обеспечившими возможность ее эволюционного развития и превращения в последующем в базовую СОД для АСУ Вооруженных Сил, включая АСУ видов и родов войск.

В процессе создания СОД коллектив, возглавляемый Игорем Александровичем Мизиным, работал в тесном контакте с широкой кооперацией предприятий (НПО «Красная заря» (г. Ленинград), НИИССУ (г. Москва), НИИ автоматики (г. Москва), Ереванский НИИ математических машин, Пензенский НИИЭИ, Тамбовский завод «Тамбоваппарат», Рязанский завод «Красное знамя», Горьковский НИИРС, КБ Завода «Сигнал» (г. Кишинев), Электромеханический завод (г. Свердловск), монтажные предприятия «Каскад» и др. [10]).

В своей работе И. А. Мизин и его коллектив постоянно и конструктивно взаимодействовали с Генеральным заказчиком АСУ (начальник связи — маршал А. И. Белов) и с головными военными НИИ, прежде всего с 16 ЦНИИ связи и 27 ЦНИИ [10, 12–15].

Созданная под руководством И. А. Мизина система стала крупным научно-техническим достижением отечественной системотехники. Она более четверти века успешно выполняла задачи по информационному обмену в интересах сотен



В день 50-летнего юбилея с друзьями и соратниками, 12 апреля 1985 г.

территориально удаленных объектов АСУ [10]. За эту разработку И. А. Мизин как главный конструктор в 1981 г. был удостоен Ленинской премии СССР, а ряд его сотрудников — Государственной премии СССР и правительственные наград.

### **3 Информационные технологии для создания крупных телекоммуникационных сетей**

Даже в период интенсивной работы над созданием СОД И. А. Мизин стремился в какой-то мере влиять и на развитие общегосударственной сети передачи данных [3, 6]. Однако условия для наиболее полного проявления широты своего научного кругозора, системного видения, глубоких знаний и огромного накопленного опыта И. А. Мизин получил в конце 1980-х гг., когда стал директором ИПИ РАН. Он уделял огромное внимание исследованиям, направленным на изыскание наиболее рациональных системотехнических решений по адаптации зарубежных технологий к российским условиям.

К наиболее значимым достижениям в этой области следует отнести [1, 6–9, 16] развитие теоретических основ информационно-телекоммуникационных технологий, разработку ряда технологий для крупномасштабных информационно-вычислительных сетей, создание комплекса базовых аппаратно-программных комплексов. Среди них следует назвать:

- обоснование основных этапов развития телекоммуникационных технологий (доклад на II Международном конгрессе ЮНЕСКО «Образование и информатика» в 1996 г.);
- обоснование комплекса работ по реализации крупномасштабных общегосударственных и региональных телекоммуникационных проектов;
- разработку методов и средств моделирования телекоммуникационных систем;
- постановку задач по описанию класса крупномасштабных телекоммуникационных систем двойного применения общенационального масштаба, совмещающих функции специальных систем и систем общего пользования [8, 9, 17].

Игорь Александрович много сделал для практической реализации полученных научных результатов. Так, в начале 1990-х гг. специалистами ИПИ РАН под его руководством была разработана и введена в действие региональная сеть передачи данных в интересах администрации Псковской области [18]. В этот период И. А. Мизин активно сотрудничал в области телекоммуникаций с ведущими научными организациями: МТУСИ (В. В. Шахгильдян), ЦНИИС МС (Л. Е. Варакин), НИИ радио (Ю. Б. Зубарев) и др.

#### **4 Генеральный конструктор комплекса работ по созданию информационных систем органов государственной власти**

На начало 1990-х гг. пришелся переломный момент в становлении и использовании информационных технологий в России в различных сферах: промышленности, науке, образовании, а также в государственном управлении. В сентябре 1994 г. указом Президента РФ была поставлена актуальная задача модернизации информационных и телекоммуникационных систем органов государственной власти на базе современных достижений. Этим же указом И. А. Мизин был назначен генеральным конструктором, а ИПИ РАН определен головной организацией [15, 16].

В процессе этой работы были сформулированы фундаментальные принципы автоматизации процессов государственного управления, разработаны системно-технические решения построения автоматизированных систем, которые долгое время остаются категориями первого выбора для многих разработчиков в нашей стране.



Генерал-майор И. А. Мизин



Березин В. Н., Копелев Ю. М., Мизин И. А. и Зацаринный А. А., 1996 г.

Под руководством И. А. Мизина коллективом ИПИ РАН совместно с ФАПСИ, ФСБ России и Минобороны РФ были разработаны методические подходы и системотехнические решения по защите информации, содержащей государственную тайну, в автоматизированных системах. Эти решения применяются при создании систем в защищенном исполнении для органов государственной власти и в настоящее время.

## 5 Заключение

Игорь Александрович Мизин постоянно вел огромную научно-организационную работу: председатель секции «Теория передачи и обработки информации» Научного совета по комплексной проблеме «Кибернетика» АН СССР, член комиссии по вычислительным центрам коллективного пользования и сетям ЭВМ при Президиуме АН СССР, председатель Научного совета РАН по информационно-вычислительным сетям, член президиума научно-технического совета по программе «Информатизация России», член Координационного совета по информатизации при Администрации Президента РФ, член научных советов при Совете Безопасности РФ, Совете обороны РФ, Минэкономики РФ, член

президиума НТС при правительстве г. Москвы, главный редактор ежегодника «Системы и средства информатики».

С 1996 г. И. А. Мизин активно работал в бюро Отделения информатики, вычислительной техники и автоматизации РАН.

И. А. Мизин был заведующим базовыми кафедрами в МИРЭА и в МТУСИ, председателем двух специализированных диссертационных советов. Подготовил значительное число кандидатов и докторов технических наук.

После прихода в ИПИ РАН И. А. Мизин установил и поддерживал активные контакты с зарубежными коллегами, получил международную известность и снискал глубокое уважение в научных кругах США, Германии, Италии, стран Юго-Восточной Азии.

Научная, научно-организационная и педагогическая деятельность И. А. Мизина была по достоинству оценена нашей Родиной. Лауреат Ленинской и Государственной премий, награжден орденом Трудового Красного Знамени, орденом «За заслуги перед Отечеством» 4 степени, многими медалями.

Игорь Александрович Мизин умер в расцвете творческих сил 8 сентября 1999 г. после тяжелой болезни. Коллеги и ученики И. А. Мизина продолжают начатые им работы, реализуя и развивая его научные идеи в институтах РАН, на предприятиях промышленности и в ведущих вузах.

## Литература

1. *Зацаринный А. А., Захаров В. Н., Синицын И. Н.* ИПИ РАН и создание информационно-телекоммуникационных сетей в интересах органов государственной власти Российской Федерации (к 25-летию ИПИ РАН) // Ведомственные корпоративные сети, 2008. № 2(47). С. 118–128.
2. *Мизин И. А., Уринсон Л. С., Храмешин Г. К.* Передача информации в сетях с коммутацией сообщений. — М.: Связь, 1972. 319 с.; М.: Связь, 1977. 328 с. (2-е изд., перераб. и доп.)
3. *Мизин И. А., Шварцман В. О.* Система передачи данных общего пользования на основе центров коммутации сообщений // Вестник АН СССР, 1975. № 3. С. 32–36.
4. *Мизин И. А., Богатырев В. А., Кулешов А. П.* Сети коммутации пакетов. — М.: Радио и связь, 1986. 408 с.
5. *Аничкин С. А., Белов С. А., Бернштейн А. В. и др.* Протоколы информационно-вычислительных сетей: Справочник / Под ред. И. А. Мизина, А. П. Кулешова. — М.: Радио и связь, 1990. 503 с.
6. *Мизин И. А.* К проблеме создания общегосударственной интегрированной сети передачи информации // Системы и средства информатики. — М.: Наука, 1993. Вып. 4.
7. *Мизин И. А.* О некоторых теоретических предпосылках совершенствования международных рекомендаций с учетом специфики первичной сети каналов связи России // Системы и средства информатики. — М.: Наука, 1995. Вып. 6.
8. *Мизин И. А.* Состояние и перспективы развития информационных телекоммуникационных технологий для сферы образования и науки // Образование и информатика: Мат-лы II Междунар. конгресса ЮНЕСКО. — М., 1996.

9. *Мизин И. А.*. Современное состояние проблематики интегрированных информационно-телекоммуникационных систем и сетей // Системы и средства информатики. — М.: Наука. Физматлит, 1999. Вып. 9. С. 11–23.
10. Автоматизация управления. Наш путь. К 50-летию НИИ АА им. академика В. С. Семенихина. — М.: НИИ АА, 2006. 210 с.
11. *Зацаринный А. А.*. Аппаратура систем и комплексов обмена данными // История отечественных средств связи. — М.: Столичная энциклопедия, 2013. С. 501–506.
12. *Зацаринный А. А.*. Академик В. С. Семенихин — выдающийся руководитель работ по созданию АСУ Вооруженных Сил Российской Федерации // Стратегические приоритеты, 2018. № 1(17). С. 78–93.
13. *Соколов И. А., Зацаринный А. А., Захаров В. Н.* У истоков создания отечественных информационно-телекоммуникационных систем: К 80-летию академика РАН И. А. Мизина // Электросвязь, 2015. № 3. С. 18–21.
14. *Зацаринный А. А.*. Академик Мизин И. А. — военный ученый, инженер и конструктор // Стратегическая стабильность, 2015. № 2. С. 62–66.
15. *Соколов И. А., Зацаринный А. А.*. Игорь Александрович Мизин // Военно-промышленная комиссия. 60 лет на службе Родине. — М.: Оружие и технологии, 2017. С. 284–289.
16. *Соколов И. А.*. Вклад академика И. А. Мизина в развитие отечественных информационных технологий и систем // И. А. Мизин — ученый, конструктор, человек. — М.: ИПИ РАН, 2010. С. 85–95.
17. *Соколов И. А., Зацаринный А. А., Захаров В. Н.* О вкладе академика И. А. Мизина в теорию и практику создания отечественных информационно-телекоммуникационных систем: К 80-летию со дня рождения // Системы и средства информатики, 2015. Т. 25. № 1. С. 213–229.
18. *Гайкович Ю. В.*. Региональные телекоммуникации от Мизина // И. А. Мизин — ученый, конструктор, человек. — М.: ИПИ РАН, 2010. С. 193–196.

*И. А. Соколов, А. А. Зацаринный, В. Н. Захаров*

---

## О Б А В Т О Р АХ

---

**Адамович Игорь Михайлович** (р. 1934) — кандидат технических наук, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Босов Алексей Вячеславович** (р. 1969) — доктор технических наук, главный научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Вакуленко Василий Васильевич** (р. 1995) — инженер-исследователь Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Водолаженко Анатолий Андреевич** (р. 1943) — сотрудник Московского института электроники и математики имени А. Н. Тихонова Национального исследовательского университета «Высшая школа экономики»

**Волков Олег Игоревич** (р. 1964) — ведущий программист Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Гончаров Александр Анатольевич** (р. 1994) — младший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Грушо Александр Александрович** (р. 1946) — доктор физико-математических наук, профессор, главный научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Грушо Николай Александрович** (р. 1982) — кандидат физико-математических наук, старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Дулин Сергей Константинович** (р. 1950) — доктор технических наук, профессор, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Дурново Александр Адрианович** (р. 1949) — ведущий программист Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Дьяченко Денис Юрьевич** (р. 1987) — инженер-исследователь Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Дьяченко Юрий Георгиевич** (р. 1958) — кандидат технических наук, старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Егоров Владимир Борисович** (р. 1948) — кандидат технических наук, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Егорова Анна Юрьевна** (р. 1991) — младший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Захаров Виктор Николаевич** (р. 1948) — доктор технических наук, доцент, научный секретарь Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Зацаринный Александр Алексеевич** (р. 1951) — доктор технических наук, заместитель директора Федерального исследовательского центра «Информатика и управление» Российской академии наук (ФИЦ ИУ РАН), главный научный сотрудник Института проблем информатики ФИЦ ИУ РАН

**Зацман Игорь Моисеевич** (р. 1952) — доктор технических наук, заведующий отделом Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Ильин Александр Владимирович** (р. 1975) — кандидат технических наук, ведущий научный сотрудник Государственного научно-исследовательского института авиационных систем

**Ильин Владимир Дмитриевич** (р. 1937) — доктор технических наук, профессор, ведущий научный сотрудник Вычислительного центра им. А. А. Дородницына Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Ионенков Юрий Сергеевич** (р. 1956) — кандидат технических наук, старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Кабанов Артем Сергеевич** (р. 1979) — кандидат технических наук, доцент Московского института электроники и математики имени А. Н. Тихонова Национального исследовательского университета «Высшая школа экономики»

**Ковалёв Дмитрий Юрьевич** (р. 1988) — научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Коновалов Григорий Михайлович** (р. 1980) — кандидат физико-математических наук, заместитель руководителя проекта по качеству Частного учреждения

Государственной корпорации по атомной энергии «Росатом» «Проектный центр ИТЭР»

**Косарик Валерий Валентинович** (р. 1970) — научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Листопад Сергей Викторович** (р. 1984) — кандидат технических наук, старший научный сотрудник Калининградского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Морозов Николай Викторович** (р. 1956) — старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Никишин Дмитрий Александрович** (р. 1976) — кандидат технических наук, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Нуриев Виталий Александрович** (р. 1980) — кандидат филологических наук, старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Плеханов Леонид Петрович** (р. 1943) — кандидат технических наук, старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Рождественский Юрий Владимирович** (р. 1952) — кандидат технических наук, заведующий сектором Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Сенчило Владимир Викторович** (р. 1963) — научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Соколов Игорь Анатольевич** (р. 1954) — академик Российской академии наук, доктор технических наук, директор Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Степченков Дмитрий Юрьевич** (р. 1973) — старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Степченков Юрий Афанасьевич** (р. 1951) — кандидат технических наук, руководитель отдела, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Сучков Александр Павлович** (р. 1954) — доктор технических наук, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Тарасов Евгений Александрович** (р. 1987) — программист Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Тимонина Елена Евгеньевна** (р. 1952) — доктор технических наук, профессор, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Тириков Егор Михайлович** (р. 1996) — студент магистратуры факультета вычислительной математики и кибернетики Московского государственного университета имени М. В. Ломоносова

**Филимонов Никита Максимович** (р. 1998) — студент Института информационных и вычислительных технологий Национального исследовательского института «Московский энергетический институт»

**Шанин Иван Андреевич** (р. 1991) — младший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

---

## **Правила подготовки рукописей статей для публикации в журнале «Системы и средства информатики»**

---

Журнал «Системы и средства информатики» публикует теоретические, обзорные и дискуссионные статьи, посвященные научным исследованиям и разработкам в области информационных технологий.

Журнал издается на русском языке. По специальному решению редколлегии отдельные статьи могут печататься на английском языке.

Тематика журнала охватывает следующие направления:

- информационно-телекоммуникационные системы и средства их построения;
- архитектура и программное обеспечение вычислительных машин, комплексов и сетей;
- методы и средства защиты информации.

1. В журнале печатаются статьи, содержащие результаты, ранее не опубликованные и не предназначенные к одновременной публикации в других изданиях.

Публикация предоставленной автором(ами) рукописи не должна нарушать положений глав 69, 70 раздела VII части IV Гражданского кодекса, которые определяют права на результаты интеллектуальной деятельности и средства индивидуализации, в том числе авторские права, в РФ.

Ответственность за нарушение авторских прав, в случае предъявления претензий к редакции журнала, несут авторы статей.

Направляя рукопись в редакцию, авторы сохраняют свои права на данную рукопись и при этом передают учредителям и редколлегии журнала неисключительные права на издание статьи на русском языке (или на языке статьи, если он отличен от русского) и на перевод ее на английский язык, а также на ее распространение в России и за рубежом. Каждый автор должен представить в редакцию подписанный с его стороны «Лицензионный договор о передаче неисключительных прав на использование произведения», текст которого размещен по адресу <http://www.ipiran.ru/publications/licence.doc>. Этот договор может быть представлен в бумажном (в 2-х экз.) или в электронном виде (отсканированная копия заполненного и подписанного документа).

Редколлегия вправе запросить у авторов экспертное заключение о возможности публикации представленной статьи в открытой печати.

2. К статье прилагаются данные автора (авторов) (см. п. 8). При наличии нескольких авторов указывается фамилия автора, ответственного за переписку с редакцией.

3. Редакция журнала осуществляет экспертизу присланных статей в соответствии с принятой в журнале процедурой рецензирования.

Возвращение рукописи на доработку не означает ее принятия к печати.

Доработанный вариант с ответом на замечания рецензента необходимо прислать в редакцию.

4. Решение редколлегии о публикации статьи или ее отклонении сообщается авторам.

Редколлегия может также направить авторам текст рецензии на их статью. Дискуссия по поводу отклоненных статей не ведется.

5. Редактура статей высылается авторам для просмотра. Замечания к редактуре должны быть присланы авторами в кратчайшие сроки.
6. Рукопись предоставляется в электронном виде в форматах MS WORD (.doc или .docx) или L<sup>A</sup>T<sub>E</sub>X (.tex), дополнительно — в формате .pdf, на дискете, лазерном диске или электронной почтой. Предоставление бумажной рукописи необязательно.
7. При подготовке рукописи в MS Word рекомендуется использовать следующие настройки.

Параметры страницы: формат — А4; ориентация — книжная; поля (см): внутри — 2,5, снаружи — 1,5, сверху и снизу — 2, от края до нижнего колонтитула — 1,3.

Основной текст: стиль — «Обычный», шрифт — Times New Roman, размер — 14 пунктов, абзацный отступ — 0,5 см, 1,5 интервала, выравнивание — по ширине. Рекомендуемый объем рукописи — не свыше 15 страниц указанного формата. При превышении указанного объема редколлегия вправе потребовать от автора сокращения объема рукописи.

Сокращения слов, помимо стандартных, не допускаются. Допускается минимальное количество аббревиатур.

Все страницы рукописи нумеруются.

Шаблоны примеров оформления представлены в Интернете:

<http://www.ipiran.ru/publications/collected/template.doc>

8. Статья должна содержать следующую информацию на **русском и английском языках**:

- название статьи;
- Ф.И.О. авторов, на английском можно только имя и фамилию;
- место работы, с указанием города и страны и электронного адреса каждого автора;
- сведения об авторах, в соответствии с форматом, образцы которого представлены на страницах:  
[http://www.ipiran.ru/journal/collected/2019\\_29\\_03\\_rus/authors.asp](http://www.ipiran.ru/journal/collected/2019_29_03_rus/authors.asp) и  
[http://www.ipiran.ru/journal/collected/2019\\_29\\_03\\_eng/authors.asp](http://www.ipiran.ru/journal/collected/2019_29_03_eng/authors.asp);
- аннотация (не менее 100 слов на каждом из языков). Аннотация — это краткое резюме работы, которое может публиковаться отдельно. Она является основным источником информации в информационных системах и базах данных. Английская аннотация должна быть оригинальной, может не быть дословным переводом русского текста и должна быть написана хорошим английским языком. В аннотации не должно быть ссылок на литературу и, по возможности, формул;
- ключевые слова — желательно из принятых в мировой научно-технической литературе тематических тезаурусов. Предложения не могут быть ключевыми словами.
- источники финансирования работы (ссылки на гранты, проекты, поддерживающие организации и т. п.).

9. Требования к спискам литературы.

Ссылки на литературу в тексте статьи нумеруются (в квадратных скобках) и располагаются в каждом из списков литературы в порядке первых упоминаний.

Списки литературы представляются в двух вариантах:

- (1) **Список литературы к русскоязычной части.** Русские и английские работы — на языке и в алфавите оригинала.
- (2) **References.** Русские работы и работы на других языках — в латинской транслитерации с переводом на английский язык; английские работы и работы на других языках — на языке оригинала.

Необходимо для составления списка “References” пользоваться размещенной на сайте <http://www.translit.net/ru/bgn/> бесплатной программой транслитерации русского текста в латиницу.

Список литературы “References” приводится полностью отдельным блоком, повторяя все позиции из списка литературы к русскоязычной части, независимо от того, имеются или нет в нем иностранные источники. Если в списке литературы к русскоязычной части есть ссылки на иностранные публикации, набранные латиницей, они полностью повторяются в списке “References”.

Примеры ссылок на различные виды публикаций в списке “References”:

**Описание статьи из журнала:**

Zhang, Z., and D. Zhu. 2008. Experimental research on the localized electrochemical micromachining. *Rus. J. Electrochem.* 44(8):926–930. doi:10.1134/S1023193508080077.

**Описание статьи из электронного журнала:**

Swaminathan, V., E. Lepkoswka-White, and B. P. Rao. 1999. Browsers or buyers in cyberspace? An investigation of electronic factors influencing electronic exchange. *JCMC* 5(2). Available at: <http://www.ascusc.org/jcmc/vol5/issue2/> (accessed April 28, 2011).

**Описание материалов конференций:**

Usmanov, T. S., A. A. Gusmanov, I. Z. Mullagalin, R. Ju. Muhametshina, A. N. Chervyakova, and A. V. Sveshnikov. 2007. Особенности проектирования разработки месторождений с применением гидравлического трещинообразования [Features of the design of field development with the use of hydraulic fracturing]. *Trudy 6-go Mezhdunarodnogo Simpoziuma “Novye resursosberegayushchie tekhnologii nedropol’zovaniya i povышeniya neftegazootdachi”* [6th Symposium (International) “New Energy Saving Subsoil Technologies and the Increasing of the Oil and Gas Impact” Proceedings]. Moscow. 267–272.

**Описание книги (монографии, сборники):**

Lindorf, L. S., and L. G. Mamikonants, eds. 1972. *Ekspluatatsiya turbogeneratorov s neposredstvennym okhlazhdeniem* [Operation of turbine generators with direct cooling]. Moscow: Energy Publs. 352 p.

**Описание переводной книги** (в списке литературы к русскоязычной части необходимо указать: / Пер. с англ. — после названия книги, а в конце ссылки указать оригинал книги в круглых скобках):

1. В русскоязычной части:

Тимошенко С. П., Янг Д. Х., Уивер У. Колебания в инженерном деле / Пер. с англ. — М.: Машиностроение, 1985. 472 с. (Timoshenko S. P., Young D. H., Weaver W. Vibration problems in engineering. — 4th ed. — N.Y.: Wiley, 1974. 521 p.)

2. В англоязычной части:

Timoshenko, S. P., D. H. Young, and W. Weaver. 1974. *Vibration problems in engineering*. 4th ed. N.Y.: Wiley. 521 p.

**Описание неопубликованного документа:**

Latypov, A. R., M. M. Khasanov, and V. A. Baikov. 2004. Geology and production (NGT GiD). Certificate on official registration of the computer program No. 2004611198. (In Russian, unpubl.)

**Описание интернет-ресурса:**

Pravila tsitirovaniya istochnikov [Rules for the citing of sources]. Available at: <http://www.scribd.com/doc/1034528/> (accessed February 7, 2011).

**Описание диссертации или автореферата диссертации:**

Semenov, V. I. 2003. Matematicheskoe modelirovanie plazmy v sisteme kompaktnyy tor [Mathematical modeling of the plasma in the compact torus]. D.Sc. Diss. Moscow. 272 p.

Kozhunova, O. S. 2009. Tekhnologiya razrabotki semanticheskogo slovarya informacionnogo monitoringa [Technology of development of semantic dictionary of information monitoring system]. PhD Thesis. Moscow: IPI RAN. 23 p.

**Описание ГОСТа:**

GOST 8.586.5-2005. 2007. Metodika vypolneniya izmereniy. Izmerenie raskhoda i kolichestva zhidkostey i gazov s pomoshch'yu standartnykh suzhayushchikh ustroystv [Method of measurement. Measurement of flow rate and volume of liquids and gases by means of orifice devices]. Moscow: Standardinform Publs. 10 p.

**Описание патента:**

Bolshakov, M. V., A. V. Kulakov, A. N. Lavrenov, and M. V. Palkin. 2006. Sposob orientirovaniya po krenu letatel'nogo apparata s opticheskoy golovkoj samonavedeniya [The way to orient on the roll of aircraft with optical homing head]. Patent RF No. 2280590.

10. Присланные в редакцию материалы авторам не возвращаются.
11. При отправке файлов по электронной почте просим придерживаться следующих правил:
  - указывать в поле subject (тема) название журнала и фамилию автора;
  - использовать attach (присоединение);
  - в состав электронной версии статьи должны входить: файл, содержащий текст статьи, и файл(ы), содержащий(е) иллюстрации.
12. Журнал «Системы и средства информатики» является некоммерческим изданием. Плата за публикацию не взимается, гонорар авторам не выплачивается.

**Адрес редакции журнала «Системы и средства информатики»:**

Москва 119333, ул. Вавилова, д. 44, корп. 2, ФИЦ ИУ РАН

Тел.: +7 (499) 135-86-92 Факс: +7 (495) 930-45-05

e-mail: [rust@ipiran.ru](mailto:rust@ipiran.ru) (Стригина Светлана Николаевна)

<http://www.ipiran.ru/journal/collected>

---

## Requirements for manuscripts submitted to Journal “Systems and Means of Informatics”

---

Journal “Systems and Means of Informatics” publishes theoretical, review, and discussion articles on the research and development in the field of information technology.

The journal is published in Russian. By a special decision of the editorial board, some articles can be published in English.

Topics covered include the following areas:

- information and communication systems and tools of their design;
- architecture and software of computational complexes and networks; and
- methods and tools of information protection.

1. The Journal publishes original articles which have not been published before and are not intended for simultaneous publication in other editions. An article submitted to the Journal must not violate the Copyright law. Sending the manuscript to the Editorial Board, the authors retain all rights of the owners of the manuscript and transfer the nonexclusive rights to publish the article in Russian (or the language of the article, if not Russian) and its distribution in Russia and abroad to the Founders and the Editorial Board. Authors should submit a letter to the Editorial Board in the following form:

***Agreement on the transfer of rights to publish:***

*“We, the undersigned authors of the manuscript “. . .,” pass to the Founder and the Editorial Board of the Journal “Systems and Means of Informatics” the nonexclusive right to publish the manuscript of the article in Russian (or in English) in both print and electronic versions of the Journal. We affirm that this publication does not violate the Copyright of other persons or organizations.”*

*Author(s) signature(s): (name(s), address(es), date).”*

This agreement should be submitted in paper form or in the form of a scanned copy (signed by the authors).

The Editorial Board has the right to request from the authors an official expert conclusion that the submitted article has no classified data prohibited for publication.

2. A submitted article should be attached with **the data on the author(s)** (see item 8). If there are several authors, the contact person should be indicated who is responsible for correspondence with the Editorial Board and other authors about revisions and final approval of the proofs.
3. The Editorial Board of the Journal examines the article according to the established reviewing procedure. If authors receive their article for correction after reviewing, it does not mean that the article is approved to be published. The corrected article should be sent to the Editorial Board for the subsequent review and approval.
4. The decision on the article publication or its rejection is communicated to the authors. The Editorial Board may also send the reviews on the submitted articles to the authors. Any discussion upon the rejected articles is not possible.
5. The edited articles will be sent to the authors for proofread. The comments of the authors to the edited text of the article should be sent to the Editorial Board as soon as possible.
6. The manuscript of the article should be presented electronically in the MS WORD (.doc or .docx) or L<sup>A</sup>T<sub>E</sub>X (.tex) formats, and additionally in the .pdf format. All documents

may be sent by e-mail or provided on a CD or diskette. A hard copy submission is not necessary.

7. The recommended typesetting instructions for manuscript.

Pages parameters: format A4, portrait orientation, document margins (cm): left — 2.5, right — 1.5, above — 2.0, below — 2.0, footer 1.3.

Text: font —Times New Roman, font size — 14, paragraph indent — 0.5, line spacing — 1.5, justified alignment.

The recommended manuscript size: not more than 10 pages of the specified format. If the specified size exceeded, the editorial board is entitled to require the author to reduce the manuscript.

Use only standard abbreviations. Avoid abbreviations in the title and abstract. The full term for which an abbreviation stands should precede its first use in the text unless it is a standard unit of measurement.

All pages of the manuscript should be numbered.

The templates for the manuscript typesetting are presented on site:

<http://www.ipiran.ru/publication/collected/template.doc>

8. Articles should enclose data both in **Russian and English**:

- title;
  - author's name and surname;
  - affiliation — organization, its address with ZIP code, city, country, and official e-mail address;
  - data on authors according to the format (see site):  
[http://www.ipiran.ru/journal/collected/2019\\_29\\_03\\_rus/authors.asp](http://www.ipiran.ru/journal/collected/2019_29_03_rus/authors.asp) and  
[http://www.ipiran.ru/journal/collected/2019\\_29\\_03\\_eng/authors.asp](http://www.ipiran.ru/journal/collected/2019_29_03_eng/authors.asp);
  - abstract (not less than 100 words) both in Russian and in English. Abstract is a short summary of the article that can be published separately. The abstract is the main source of information on the article and it could be included in leading information systems and data bases. The abstract in English has to be an original text and should not be an exact translation of the Russian one. Good English is required. In abstracts, avoid references and formulae.
  - Indexing is performed on the basis of keywords. The use of keywords from the internationally accepted thematic Thesauri is recommended.
- Important! Keywords must not be sentences.
- Acknowledgments.

9. References. Russian references have to be presented both in English translation and in Latin transliteration (refer <http://www.translit.net/ru/bgn/>).

Please take into account the following examples of Russian references appearance:

**Article in journal:**

Zhang, Z., and D. Zhu. 2008. Experimental research on the localized electrochemical micromachining. *Rus. J. Electrochem.* 44(8):926–930. doi:10.1134/S1023193508080077.

**Journal article in electronic format:**

Swaminathan, V., E. Lepkoswka-White, and B. P. Rao. 1999. Browsers or buyers in cyberspace? An investigation of electronic factors influencing electronic

exchange. *JCMC* 5(2). Available at: <http://www.ascusc.org/jcmc/vol5/issue2/> (accessed April 28, 2011).

**Conference proceedings:**

Usmanov, T.S., A.A. Gusmanov, I.Z. Mullagalin, R.Ju. Muhametshina, A.N. Chervyakova, and A.V. Sveshnikov. 2007. Osobennosti proektirovaniya razrabotki mestorozhdeniy s primeneniem gidrorazryva plasta [Features of the design of field development with the use of hydraulic fracturing]. *Trudy 6-go Mezhdunarodnogo Simpoziuma "Novye resursosberegayushchie tekhnologii nedropol'zovaniya i povysheniya neftegazootdachi"* [6th Symposium (International) "New Energy Saving Subsoil Technologies and the Increasing of the Oil and Gas Impact" Proceedings]. Moscow. 267–272.

**Books and other monographs:**

Lindorf, L. S., and L. G. Mamikonants, eds. 1972. *Ekspluatatsiya turbogeneratorov s neposredstvennym okhlazhdeniem* [Operation of turbine generators with direct cooling]. Moscow: Energy Publs. 352 p.

**Dissertation and Thesis:**

Kozhunova, O. S. 2009. Tekhnologiya razrabotki semanticheskogo slovarya informatsionnogo monitoringa [Technology of development of semantic dictionary of information monitoring system]. PhD Thesis. Moscow: IPI RAN. 23 p.

**State standards and patents:**

GOST 8.586.5-2005. 2007. Metodika vypolneniya izmereniy. Izmerenie raskhoda i kolichestva zhidkostey i gazov s pomoshch'yu standartnykh suzhayushchikh ustroystv [Method of measurement. Measurement of flow rate and volume of liquids and gases by means of orifice devices]. M.: Standardinform Publs. 10 p.

Bolshakov, M. V., A. V. Kulakov, A. N. Lavrenov, and M. V. Palkin. 2006. Sposob orientirovaniya po krenu letatel'nogo apparata s opticheskoy golovkoj samonavedeniya [The way to orient on the roll of aircraft with optical homing head]. Patent RF No. 2280590.

References in Latin transcription are presented in the original language.

References in the text are numbered according to the order of their first appearance; the number is placed in square brackets. All items from the reference list should be cited.

10. Manuscripts and additional materials are not returned to Authors by the Editorial Board.
11. Submissions of files by e-mail must include:
  - the journal title and author's name in the "Subject" field;
  - an article and additional materials have to be attached using the "attach" function;
  - an electronic version of the article should contain the file with the text and a separate file with figures.
12. "System and Means of Informatics" journal is not a profit publication. There are no charges for the authors as well as there are no royalties.

**Editorial Board address:**

FRC CSC RAS, 44, block 2, Vavilov Str., Moscow 119333, Russia

Ph.: +7 (499)135 86 92, Fax: +7 (495)930 45 05

e-mail: [rust@ipiran.ru](mailto:rust@ipiran.ru) (to Svetlana Strigina)

[http://www.ipiran.ru/english/journal\\_systems.asp](http://www.ipiran.ru/english/journal_systems.asp)

# **SYSTEMS AND MEANS OF INFORMATICS (СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ)**

**SCIENTIFIC JOURNAL**

**Volume 30 No.2 Year 2020**

**Editor-in-Chief and Chair of Editorial Council**

Academician I. A. Sokolov

## **I N T H I S I S S U E:**

SELF-TIMED COMBINATIONAL CIRCUIT TOLERANCE TO SHORT-TERM SOFT ERRORS

*Yu. A. Stepchenkov, Yu. G. Diachenko, Yu. V. Rogdestvenski,  
N. V. Morozov, D. Yu. Stepchenkov, and D. Yu. Diachenko*

4

UNIVERSAL FUNCTIONAL METHOD FOR ANALYZING LARGE SELF-TIMED CIRCUITS

*L. P. Plekhanov and V. N. Zakharov*

11

PREVENTION OF MASS FAILURE OF LOW-SECURITY MOBILE SYSTEMS  
OF INFORMATION SUPPORT FOR DIGITAL ECONOMY

*A. A. Grusho, N. A. Grusho, V. V. Senchilo, and E. E. Timonina*

21

VALUATION AND PROTECTION OF INSIDER INFORMATION

*A. S. Kabanov and A. A. Vodolazhenko*

31

MULTIDISCIPLINARY NEUROINFORMATICS PROBLEMS FOR EXECUTION  
IN DISTRIBUTED COMPUTING INFRASTRUCTURES

*D. Y. Kovalev, I. A. Shanin, and E. M. Tirikov*

43

ARCHITECTURE OF THE PLATFORM FOR MANAGING HYPOTHESES-DRIVEN  
VIRTUAL EXPERIMENTS

*D. Y. Kovalev, E. A. Tarasov, V. N. Zakharov, and N. M. Filimonov*

56

APPROACHES TO THE INTEGRATION OF THE APPLICATION CONCEPTUAL  
SCHEMAS IN THE UNIFIED GEOONTOLOGY

*S. K. Dulin and D. A. Nikishin*

68

FUNCTIONAL STRUCTURE OF THE COHESIVE HYBRID INTELLIGENT  
MULTIAGENT SYSTEM

*S. V. Listopad*

78

THE PROBLEM OF SETTING AVAILABILITY TARGETS FOR ITER DIAGNOSTICS:  
PROBLEM STATEMENT AND SOLUTION ALGORITHM

*G. M. Konovalov*

89