

# **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. Стемповский  
член-корреспондент РАН Ю. Б. Зубарев  
профессор Ш. Долев (S. Dolev, Beer-Sheva, Israel)  
профессор Ю. Кабанов (Yu. Kabanov, Besancon, France)  
профессор М. Никулин (M. Nikulin, Bordeaux, 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  
проф., д.ф.-м.н. А. И. Зейфман  
проф., д.т.н. В. Д. Ильин  
проф., д.т.н. К. К. Колин  
проф., д.ф.-м.н. В. Ю. Королев  
к.ф.-м.н. Р. В. Разумчик  
проф., д.г.-м.н. Р. Б. Сейфуль-Мулюков  
д.ф.-м.н. В. И. Сеницын  
проф., д.т.н. И. Н. Сеницын  
проф., д.ф.-м.н. В. Г. Ушаков  
к.ф.-м.н. А. К. Горшенин — отв. секретарь  
к.ф.-м.н. С. А. Христочевский

## **Редакция**

профессор, д.г.-м.н. Р. Б. Сейфуль-Мулюков  
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, 2019

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 29 № 4 Год 2019

## СОДЕРЖАНИЕ

Снижение размерности для смеси вероятностных анализаторов главных компонент применительно к задачам медицинской диагностики	
<b><i>М. П. Кривенко</i></b>	<b>4</b>
Оптимизация индикации многоразрядных самосинхронных схем	
<b><i>Ю. А. Степченков, Ю. Г. Дьяченко, Ю. В. Рождественский, Н. В. Морозов, Д. Ю. Степченков, Д. Ю. Дьяченко</i></b>	<b>14</b>
Поиск эмпирических причин сбоев и ошибок в компьютерных системах и сетях с использованием метаданных	
<b><i>А. А. Грушо, М. И. Забейайло, Н. А. Грушо, Е. Е. Тимонина</i></b>	<b>28</b>
Визуальный метаязык описания решения диагностических проблем	
<b><i>С. Б. Румовская, А. В. Колесников, А. А. Литвин</i></b>	<b>39</b>
Стационарные характеристики обслуживания в системе GI/MSP/n/∞ сообщенным обновлением	
<b><i>И. С. Зарядов, Л. А. Мейханаджян, Т. А. Милованова</i></b>	<b>50</b>
Метод случайного отбора при прогнозировании временных рядов рынка криптовалют	
<b><i>О. Е. Сороколетова, Т. В. Захарова</i></b>	<b>65</b>
Анализ производительности систем «новое радио» сети 5G с помощью системы массового обслуживания с переменными требованиями к ресурсу	
<b><i>В. А. Бесчастный, Д. Ю. Острикова, Ю. В. Гайдамака</i></b>	<b>73</b>
Метод извлечения экспертных оценок из социальных сетей при проведении процесса группового принятия решений	
<b><i>О. В. Чухно, Н. В. Чухно, Ю. В. Гайдамака, К. Е. Самуйлов</i></b>	<b>84</b>

# СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ

Том 29 № 4 Год 2019

## СОДЕРЖАНИЕ

Принципы создания и функции интеллектуальной словарной системы <b>В. В. Вакуленко, И. М. Зацман</b>	96
О методах оценки качества машинного перевода <b>А. К. Рычихин</b>	106
Исследование процесса формирования виртуального сообщества как нелинейной динамической системы <b>И. М. Адамович, О. И. Волков</b>	119
Функциональная модель системы управления процессами стратегического планирования <b>А. П. Сучков</b>	129
Символы, коды, сигналы <b>В. Д. Ильин</b>	140
Об авторах	149
Авторский указатель за 2019 г.	153
2019 Author Index	159
Правила подготовки рукописей статей	165
Requirements for manuscripts	169

# СНИЖЕНИЕ РАЗМЕРНОСТИ ДЛЯ СМЕСИ ВЕРОЯТНОСТНЫХ АНАЛИЗАТОРОВ ГЛАВНЫХ КОМПОНЕНТ ПРИМЕНИТЕЛЬНО К ЗАДАЧАМ МЕДИЦИНСКОЙ ДИАГНОСТИКИ

*М. П. Кривенко*<sup>1</sup>

**Аннотация:** Рассматриваются алгоритмы выбора структурных параметров смеси вероятностных анализаторов главных компонент применительно к задачам медицинской диагностики. Решение предлагается искать, комбинируя применение информационных критериев для формирования начальных приближений с последующим уточнением получающихся решений. Описанные подходы и алгоритмы приводят к результатам, которые в общем случае не гарантируют наилучшего решения. Но они позволяют прояснить, возможно ли снижение размерности, приводящее к повышению качества классификации. Кроме того, формируется новая информация об объектах исследования. На примере экспериментов по диагностированию болезней печени и прогнозированию химического состава мочевых камней демонстрируются возможности описанных процедур анализа данных. Предлагаемые решения являются источником повышения точности классификации, дают толчок специалистам в предметной области для прояснения сути протекающих процессов.

**Ключевые слова:** анализ главных компонент; смеси нормальных распределений; снижение размерности; информационные критерии; перепроверка; медицинская диагностика

**DOI:** 10.14357/08696527190401

## 1 Введение

Одной из интенсивно развивающихся областей применения методов анализа данных стала медицина. Много сделано в применении технологии машинного обучения для медицинской диагностики. Полученный на основе накопленных данных о правильных диагнозах классификатор может быть затем использован либо для оказания помощи врачу, когда обследуются новые пациенты, с целью повышения скорости, точности и/или надежности диагностики, либо для проведения углубленных медицинских исследований с целью получения новых знаний.

Чтобы система машинного обучения была полезна при решении задач медицинской диагностики, желательно наличие следующих функций: умение справляться с отсутствующими и/или аномальными данными; прозрачность диагностических решений и возможность их объяснять; высокая производительность, достаточная для обучения и переобучения; способность методов обработки данных сокращать число тестов, необходимых для получения достоверного диагноза.

---

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

Прозрачность диагностических решений означает, что сформированные знания и объяснения должны быть понятными для врача. Он должен быть в состоянии проанализировать формируемые знания. В идеале автоматически полученные выводы и решения должны обеспечить врачу новую точку зрения на решаемую проблему, помочь выявить новые взаимосвязи и закономерности, которые раньше не были видны явно. Способность объяснения должна обеспечивать поддержку при диагностике новых пациентов.

Сокращение числа тестов проистекает из того, что в медицинской практике сбор данных о пациентах — часто дорогой, трудоемкий и иногда вредный для пациентов процесс. Кроме того, подчас только в результате сокращения признаков классификации становится доступным всеобъемлющий анализ предлагаемых решений.

Есть много и других общих факторов, которые мотивируют включение уменьшения размерности как этапа в различных системах обработки данных. Если данные суть наборы вещественных векторов и эти векторы оказываются большой размерности, то обработка становится невозможной. По этой причине часто полезно, а иногда и необходимо, уменьшить размерность данных до более управляемых размеров с минимально возможной потерей информации. Это может быть случай, когда явления высокой размерности порождаются существенно более простыми по составу переменными. Здесь процесс уменьшения размерности касается параметров модели данных, он будет выступать в качестве инструмента для моделирования соответствующих явлений, повышения их ясности. Часто присутствует значительный объем избыточной или вводящей в заблуждение информации; такая информация должна быть удалена перед дальнейшей обработкой данных. Это может не только сократить время обработки и повысить качество классификации, но и сделать полученные правила более понятными.

Обычно выделяют методы отбора (селекции) признаков и методы преобразования признаков [1]. При селекции исследователь считает, что среди всех доступных переменных только часть действительно необходимы, все остальные излишни и реально не имеют значения для решения поставленной задачи. Второй подход к уменьшению числа переменных предполагает, что только несколько комбинаций (обычно линейных) исходных переменных значимы для формирования диагноза, поэтому цель снижения размерности — это идентифицировать эти комбинации.

К наиболее простым и широко используемым методам преобразования относятся линейные, а среди них — анализ главных компонент (РСА — principal component analysis). Но полезность РСА не проявляется для данных, демонстрирующих нелинейность, что привело к построению методов, которые могут эффективно справляться с ней. Понятно, что первые попытки были связаны с расширением имеющихся подходов, в частности путем кластеризации данных и выполнения РСА внутри кластеров.

В [2] рассматривался выбор структурных параметров (число элементов смеси и размерности этих элементов), характеризующих модель типа смеси вероят-

ностных анализаторов главных компонент. Способы оценивания структурных параметров выбранной модели данных формируются в непростых условиях: необходимость привлечения сложных априорных предположений и, как следствие, проблемы полного их задания, отсутствие конкретных наработок в области селекции моделей типа смеси анализаторов главных компонент, сложность решения сопутствующих аналитических и вычислительных задач. Поэтому приходится обращаться к отдельным приемам, реализующим общие идеи и позволяющим получать гарантированные по качеству решения. Для реализации выбора размерностей предлагается использовать комбинацию известных методов выбора размерностей принятой модели, а для оценивания качества обучаемой классификации данных — методы управления выборкой.

## 2 Алгоритмы

Базовая вероятностная модель анализа главных компонент (PPCA — probabilistic PCA) для сниженной размерности  $k$  основывается на представлении

$$\mathbf{y} = \mathbf{W}\mathbf{x} + \boldsymbol{\mu} + \boldsymbol{\varepsilon},$$

где  $\mathbf{y}$  —  $(d \times 1)$  наблюдаемая переменная,  $\mathbf{y} \sim N(\boldsymbol{\mu}, \mathbf{C}(k))$ ;  $\mathbf{W}$  —  $(d \times k)$ -матрица преобразования;  $\mathbf{x}$  —  $(k \times 1)$  латентная переменная,  $\mathbf{x} \sim N(\mathbf{0}, \mathbf{I})$ ;  $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$ ;  $\mathbf{C}(k) = \mathbf{W}\mathbf{W}^T + \sigma^2 \mathbf{I}$ . Здесь  $d$  — исходная размерность данных;  $k$  — сниженная размерность данных;  $\boldsymbol{\mu}$ ,  $\mathbf{W}$  и  $\sigma^2$  суть параметры модели, число скалярных параметров при  $k < d$  равно  $dk + d + 1$ , а при  $k = d$  равно  $(d^2 + 3d)/2$ . При переходе от  $d > 2$  к сниженной размерности  $k$  количество параметров модели уменьшается только для  $k = 1, \dots, [d/2]$ .

Связь вероятностной модели со стандартным PCA позволяет моделировать сложные структуры данных с помощью комбинации локальных подмоделей PPCA и реализации смеси вероятностных анализаторов главных компонент (PPCAM — PPCA mixtures). Этот подход позволяет определять все параметры модели путем максимизации правдоподобия. Логарифм правдоподобия для такой модели смеси есть

$$L = \sum_{i=1}^n \ln \{p(\mathbf{y}_i)\} = \sum_{i=1}^n \ln \left\{ \sum_{j=1}^m \pi_j p(\mathbf{y}_i | j) \right\},$$

где  $p(\mathbf{y}_i | j)$  отвечает элементарной PPCA( $j$ )-модели и  $\pi_j$  — соответствующий вес элемента смеси с  $\pi_j \geq 0$  и  $\sum_{j=1}^m \pi_j = 1$ ; при этом с каждым  $j$ -м элементом смеси связаны свои параметры  $\boldsymbol{\mu}_j$ ,  $\mathbf{W}_j$  и  $\sigma_j^2$ .

Для оценивания параметров PPCAM существуют два основных варианта:

- (1) поэтапный, когда сначала находятся оценки параметров смеси нормальных распределений, затем осуществляется факторизация ковариационных матриц элементов смеси и, наконец, строятся оценки параметров  $\mathbf{W}_j$  и  $\sigma_j^2$ ;

- (2) одновременный для оценивания всех параметров  $\pi_j$ ,  $\mu_j$ ,  $\mathbf{W}_j$  и  $\sigma_j^2$ ,  $j = 1, \dots, m$ , в совокупности [3].

Второй вариант из-за специфической организации итерационного процесса обладает вычислительными преимуществами, что может стать существенным при разовом оценивании. Но снижение размерности на основе выбора конкретной модели РРСАМ подразумевает многократное оценивание ее параметров при различных предположениях. Поэтому поэтапный вариант оценивания становится более привлекательным, так как позволяет однократно получать представления ковариационных матриц элементов смеси, а затем многократно формировать из них нужные оценки. Таким образом, при относительно небольших значениях  $d$  и необходимости перебора вариантов различных значений  $k_j$  более приемлемым оказывается поэтапный метод: найти оценки параметров смеси для  $k_j = d$ , а затем использовать эти оценки для формирования оценок для  $k_j < d$ .

Указанный принцип перехода от задачи с большей размерности к меньшей можно эффективно использовать и при переборе различных значений числа элементов смеси  $m$ . Дело в том, что одним из сложных элементов реализации EM (expectation-maximization) алгоритма оценивания параметров смеси является формирование начальных приближений. Автор данной статьи для этого эффективно использует следующие шаги:

- сначала оценить параметры смеси для «заведомо большого» значения  $\tilde{m}$  (получая начальное приближение, например, с помощью многократного случайного заполнения матрицы апостериорных вероятностей принадлежности наблюдаемого значения к элементу смеси и сравнения получающихся результатов оценивания);
- затем найти оценку для меньшего значения  $\tilde{m} - 1$  путем выбора наилучшей оценки из  $\tilde{m}$  оценок, построенных при начальных приближениях, полученных исключением одного из элементов смеси в оценке для  $\tilde{m}$ ;
- повторять описанные действия до того, как будет получена оценка для требуемого значения  $m < \tilde{m}$ .

Алгоритм выбора размерностей согласно АИС (Akaike information criterion) и ВИС (Bayesian information criterion) принимает вид:

- в отдельности для каждого класса данных обучающей выборки в диапазоне возможных значений числа элементов смеси  $m = 1, \dots, m_{\max}$  и размерности пространства признаков  $k = 1, \dots, d$  найти оценки максимального правдоподобия для параметров модели РРСАМ, что позволит совместно с соответствующими значениями числа параметров модели найти значения информационных критериев;
- выбрать из полученных значений минимальные по величине и соответствующие им «лучшие»  $k^*$  и  $m^*$ ;
- оценить при полученных  $k^*$  и  $m^*$  методом перепроверки точность байесовского классификатора.

В этом алгоритме принято некоторое упрощение: для всех элементов смеси действует одно и то же значение  $k$ .

При относительно небольших размерностях и объемах данных можно пошаговым улучшением получать более реалистичные результаты. Для этого достаточно последовательно реализовать следующие этапы выбора значений параметров размерности, не ухудшающих точность классификации с точки зрения метода перепроверки:

- перебрать при  $k = d$  допустимые значения  $m$ , одинаковые для всех классов, что даст набор  $(m_0^*, \dots, m_0^*)$  соответственно для классов  $1, \dots, G$ ;
- уточнить значения  $m$  для каждого класса, т. е. перебрать при  $k = d$  всевозможные комбинации вида  $(m_1, \dots, m_G)$ , где  $m_i = 1, \dots, m_0^*$  для  $i = 1, \dots, G$ , что приведет к набору  $(m_1^*, \dots, m_G^*)$ ;
- уточнить размерности пространства признаков для каждого элемента смеси и каждого класса, т. е. перебрать всевозможные комбинации вида  $\{(k_{11}, \dots, k_{1m_1^*}), \dots, (k_{11}, \dots, k_{1m_G^*})\}$ , где  $k_{ij} = 1, \dots, d$  для  $i = 1, \dots, G$  и  $j = 1, \dots, m_i^*$ , результат которого принять за искомую оценку размерностей.

Последние два этапа суть реализация циклического покоординатного спуска. Все этапы перебора сопровождаются контролем значения оценки точности классификации.

### 3 Эксперименты

**Диагностирование заболеваний печени.** Анализировались результаты реальных обследований (см. [4, разд. 5]) с целью дискриминации данных о ферментах, полученных для 218 пациентов с четырьмя заболеваниями печени.

В данной статье для описания данных каждого из четырех классов принималась модель РРСАМ и ставилась задача выбора эффективной размерности модели, т. е. значений параметров  $m_j$  для каждого  $j$ -го класса и значений  $k_{ji}$  для каждого  $j$ -го класса и  $i$ -го элемента смеси нормальных распределений в этом классе.

Результаты применения информационных критериев приведены в табл. 1, в которой  $q$  — число параметров модели. Оценки точности классификации  $A^*$  для найденных значений размерности следующие: для АИС — 83%, для ВИС — 73%. Для сравнения, для модели нормального распределения данных из каждого класса, рассмотренной в [4], точность классификации составляла 67%.

Последовательное уточнение оценок размерности дало следующие результаты:

- перебор в диапазоне  $m = 1, \dots, 10$  единых для всех классов значений числа элементов смеси при  $k_{ji} \equiv 4$  привел к оценкам  $m^* = (6, 6, 6, 6)$  и  $A^* = 84\%$ ,  $q = 356$ ;



**Таблица 1** Оценки размерности модели

№ класса (число наблюдений)	AIC			BIC		
	$m^*$	$k^*$	$q$	$m^*$	$k^*$	$q$
1 (57)	10	4	149	3	4	44
2 (44)	7	4	104	3	4	44
3 (40)	10	4	149	4	4	59
4 (77)	10	4	149	3	4	44

- уменьшение отдельных значений  $m^*$  привело к оценкам  $m^* = (5, 5, 4, 5)$  и  $A^* = 85\%$ ,  $q = 281$ ;
- уменьшение отдельных значений  $k = \{(4, 4, 4, 4, 4), (4, 4, 4, 4, 4), (4, 4, 4, 4), (5, 5, 5, 5, 5)\}$  привело к оценкам  $k^* = \{(1, 3, 1, 1, 1), (1, 1, 1, 1, 3), (1, 3, 1, 2), (2, 1, 2, 2, 2)\}$  и  $A^* = 85\%$ ,  $q = 238$ .

Анализ данных показал:

- из информационных критериев лучшим оказывается AIC, что соответствует его нацеленности на аппроксимацию участвующих в классификации плотностей распределения данных;
- при относительно небольших значениях структурных размерностей удается найти решения, приемлемые по сложности и превосходящие по точности привычный подход [4] и более сложный подход, но не дополненный исчерпывающим решением задачи выбора модели [5];
- полученные оценки в виде смесей не только говорят о непростой структуре данных в классах, но и позволяют описать существующие кластеры как через их параметры, так и с помощью соответствующих наборов наблюдаемых значений.

**Прогноз химического состава мочевых камней.** Задача прогнозирования состава камня по набору показателей была сформулирована как задача обучаемой классификации типов камней (о данных более подробно см. [6]). При этом предполагалось, что исследователем задана классификация типов камней по их составу. В группу признаков, характеризующих химический состав камня пациента, входили 9 компонентов, в частности: вевеллит (WH), веделлит (WD), даллит, или карбонатапатит (Dh). В данной работе рассматривалось деление камней на 2 класса («кальциевые» и «остальные») по правилам: 1-й класс, если по составу камня  $WH + WD + Dh > 50\%$  и 2-й класс в противном случае. В задаче обучаемой классификации использовались 14 показателей, включающих метаболические и антропологические показатели 386 пациентов. Полученный классификатор использовался для прогноза типа камня по представляемым данным о показателях. Для описания этих данных каждого из двух классов принималась модель PPCAM и решалась задача выбора эффективной размерности модели.

**Таблица 2** Оценки размерности модели

№ класса (число наблюдений)	AIC			BIC		
	$m^*$	$k^*$	$q$	$m^*$	$k^*$	$q$
1 (261)	6	14	719	3	14	359
2 (125)	3	14	359	2	14	239

Результаты применения информационных критериев приведены в табл. 2. Оценки точности классификации для найденных значений размерности следующие: для AIC — 52%; для BIC — 40%.

Последовательное уточнение оценок размерности дало следующие результаты:

- перебор в диапазоне  $m = 1, \dots, 10$  единых для всех классов значений числа элементов смеси при  $k_{ji} \equiv 14$  привел к оценкам  $m^* = (1, 1)$  и  $A^* = 68\%$ ,  $q = 238$  (вариант 1) или к оценкам  $m^* = (10, 10)$  и  $A^* = 68\%$ ,  $q = 2398$  (вариант 2).

Для варианта 1 в уменьшении отдельных значений  $m^*$  необходимость отпала; а уменьшение отдельных значений  $k = \{(14), (14)\}$  привело к оценкам  $k^* = \{(13), (13)\}$  и  $A^* = 68\%$ ,  $q = 394$ .

В случае варианта 2 последовательно было получено:

- уменьшение отдельных значений  $m^*$  привело к оценкам  $m^* = (1, 10)$  и  $A^* = 72\%$ ,  $q = 1318$ ;
- уменьшение отдельных значений  $k = \{(14), (14, 14, 14, 14, 14, 14, 14, 14, 14, 14)\}$  привело к оценкам  $k^* = \{(2), (1, 1, 1, 1, 1, 1, 1, 1, 1, 3)\}$  и  $A^* = 73\%$ ,  $q = 370$ .

Анализ данных показал:

- информационные критерии «угадывают» далеко не лучшие решения;
- уточнение структуры модели привело по сравнению со 2-м вариантом и с [6] к повышению точности классификации;
- более сложный по структуре 2-й вариант оценки параметров модели привел в итоге процесса оптимизации к меньшим значениям числа параметров и, самое важное, подтвердил сложную структуру данных из второго класса «прочие», включающего камни нескольких типов.

## 4 Заключение

Описанные подходы и алгоритмы не гарантируют наилучшего решения, но позволяют прояснить, возможно ли снижение размерности, приводящее к повышению качества классификации. Кроме того, формируется новая информация об

объектах исследования. Проведенные эксперименты продемонстрировали, что предлагаемые решения являются источником роста точности классификации, дают толчок специалистам в предметной области для прояснения сути протекающих процессов.

История развития машинного обучения и его применения в медицинской диагностике показывает, что происходит переход от простых данных и алгоритмов их обработки, а также соответствующих систем и методологий к решению проблем, сложных по структуре и содержанию. Интеллектуализация анализа данных играет все более важную роль благодаря огромному объему информации, производимой и хранимой с помощью современных технологий. Современные алгоритмы машинного обучения должны предоставлять инструменты, которые помогут врачам выявить интересные свойства и неожиданные связи в своих данных.

Среди причин медленного восприятия практикой достаточно сложных подходов, пожалуй, наиболее разумной является то, что внедрение технологии машинного обучения еще больше увеличит число приемов, доступных для врачей. Любой новый инструмент имеет нежелательный побочный эффект дальнейшего усложнения работы врача, которая и так не проста. Поэтому технология машинного обучения должна быть интегрирована в существующие технические средства, что сделает ее использование максимально простым и естественным.

## Литература

1. *Jensen R., Shen Q.* Computational intelligence and feature selection. Rough and fuzzy approaches. — Hoboken, NJ, USA: John Wiley & Sons, 2008. 340 p.
2. *Кривенко М. П.* Выбор размерностей для смеси вероятностных анализаторов главных компонент // Системы и средства информатики, 2019. Т. 29. № 3. С. 4–15.
3. *Tipping M. E., Bishop C. M.* Mixtures of probabilistic principal component analyzers // Neural Comput., 1999. Vol. 11. Iss. 2. P. 443–482.
4. *Alber A.* Multivariate interpretation of clinical laboratory data. — New York, NY, USA: CRC Press, 1987. 386 p.
5. *Кривенко М. П.* Обучаемая классификация неполных клинических данных // Информатика и её применения, 2017. Т. 11. Вып. 3. С. 27–33.
6. *Кривенко М. П.* Критерии значимости отбора признаков классификации // Информатика и её применения, 2016. Т. 10. Вып. 3. С. 34–40.

*Поступила в редакцию 16.07.19*

## DIMENSIONALITY REDUCTION FOR MIXTURE OF PROBABILISTIC PRINCIPAL COMPONENT ANALYZERS IN RELATION TO THE TASKS OF MEDICAL DIAGNOSTICS

M. P. Krivenko

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 considers algorithms of choosing structural parameters characterizing the mixture of probabilistic principal component analyzers model in relation to the tasks of medical diagnostics. The solution is proposed to search by combining the application of information criteria for the formation of initial approximations, followed by refinement of the resulting solutions. The described approaches and algorithms lead to results that generally do not guarantee the best solution. But they make it possible to clarify whether it is possible to reduce the dimensionality which leads to an increase in the quality of classification. In addition, new information about the objects of study is being formed. Using the example of experiments to diagnose liver diseases and predict the chemical composition of urinary stones, the capabilities of the described data analysis procedures are demonstrated. The proposed solutions are a source of improving the accuracy of classification and give impetus to experts in the subject area to clarify the essence of the processes.

**Keywords:** principal components analysis; Gaussian mixture model; dimensionality reduction; information criterions; cross-validation; medical diagnostics

**DOI:** 10.14357/08696527190401

### References

1. Jensen, R., and O. Shen. 2008. *Computational intelligence and feature selection. Rough and fuzzy approaches*. Hoboken, NJ: John Wiley & Sons. 340 p.
2. Krivenko, M. P. 2019. Vybór razmernostey dlya smesi veroyatnostnykh analizatorov glavnykh komponent [Selecting the dimensionality for mixture of probabilistic principal component analyzers]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(3):4–15.
3. Tipping, M. E., and C. M. Bishop. 1999. Mixtures of probabilistic principal component analyzers. *Neural Comput.* 11(2):443–482.
4. Alber, A. 1987. *Multivariate interpretation of clinical laboratory data*. New York, NY: CRC Press. 386 p.
5. Krivenko, M. P. 2017. Obuchaemaya klassifikatsiya nepolnykh klinicheskikh dannyykh [Supervised learning classification of incomplete clinical data]. *Informatika i ee Primeneniya — Inform. Appl.* 11(3):27–33.

6. Krivenko, M. P. 2016. Kriterii znachimosti otbora priznakov klassifikatsii [Significance tests of feature selection for classification]. *Informatika i ee Primeneniya — Inform. Appl.* 10(3):34–40.

*Received July 16, 2019*

### **Contributor**

**Krivenko Michail P.** (b. 1946) — 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; mkrivenko@ipiran.ru

## ОПТИМИЗАЦИЯ ИНДИКАЦИИ МНОГОРАЗРЯДНЫХ САМОСИНХРОННЫХ СХЕМ\*

*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:** Индикаторная подсхема в самосинхронных (СС) схемах обеспечивает контроль окончания переключения всех элементов в текущую фазу и управление взаимодействием функциональных блоков. С увеличением разрядности СС-схемы растет и вклад индикаторной подсхемы в задержку переключения схемы. В статье рассматриваются вопросы оптимизации индикаторной подсхемы и организации СС-конвейера для обеспечения более высокого быстродействия всей схемы. Для реализации разряда регистров хранения промежуточных данных в ступенях конвейера предлагается использовать вместо обычного СС-RS-триггера гистерезисный триггер, имеющий меньшую сложность и обеспечивающий хранение как рабочего, так и спейсерного состояния парафазных данных. Дисциплина управления фазами ступеней конвейера с помощью общих индикаторных выходов соседних ступеней конвейера заменяется принципом поразрядной индикации и управления фазами, использующей параллельность вычислений в многоразрядных СС-схемах. Предлагаемые решения обеспечивают существенное повышение быстродействия СС-схем за счет незначительного усложнения индикаторной подсхемы.

**Ключевые слова:** самосинхронная схема; конвейер; гистерезисный триггер

**DOI:** 10.14357/08696527190402

### 1 Введение

Теоретически СС-схемы [1–3] обеспечивают максимально возможное быстродействие в любых конкретных условиях окружающей среды. Они используют

---

\* 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

запрос-ответное взаимодействие между задатчиком данных и их приемником. Благодаря этому быстродействие СС-схем определяется лишь реальными в данных условиях эксплуатации задержками элементов, из которых они состоят. Этим они выгодно отличаются от синхронных схем, которые в отсутствие адаптации источника тактового сигнала к изменяющимся условиям окружающей среды вынуждены ориентироваться на «наихудший случай»: минимально возможное напряжение питания, максимально допустимую температуру окружающей среды, «медленные» транзисторы и т. д. В результате в ряде приложений СС-схемы, несмотря на свою аппаратную избыточность, оказываются более быстродействующими, чем их синхронные аналоги.

Основные факторы ограничения быстродействия СС-схем:

- двухфазная дисциплина работы — рабочая фаза, в которой реализуется заложенный в схему алгоритм обработки входных данных, и спейсерная фаза (спейсер), служащая для разделения соседних рабочих фаз;
- наличие индикаторной подсхемы.

Индикаторная подсхема обеспечивает контроль окончания переключения СС-схемы в текущую фазу. Для подтверждения окончания переключения СС-схемы требуется проиндицировать выходы всех ее элементов и собрать их в один индикаторный сигнал, служащий сигналом управления задатчиком.

В относительно несложных схемах с небольшой разрядностью индикаторная подсхема работает в «фоновом» режиме. В многоразрядных СС-схемах большое число индицирующих сигналов приводит к «разбуханию» индикаторной подсхемы, которая начинает вносить существенный вклад в задержку СС-схемы. Поэтому разработка методов ускорения индикации многоразрядных СС-схем является актуальной задачей. Данная работа посвящена исследованию возможностей ускорения и распараллеливания индикаторной подсхемы КМОП-СС-схем и анализу вариантов ее реализации в типовых устройствах.

## 2 Самосинхронный конвейер

Сложные СС-схемы, как и их синхронные аналоги, строятся в виде конвейера. Ступень СС-конвейера содержит комбинационную часть, реализующую алгоритм обработки данных, и входной или выходной регистр, хранящий результаты обработки данных. В синхронных схемах все ступени конвейера, как правило, синхронизируются одним тактовым сигналом. Самосинхронный конвейер управляется сигналами, индицирующими окончание переключения соседних ступеней конвейера в очередную фазу работы.

Базовый принцип взаимодействия ступеней СС-конвейера заключается в следующем: текущая ступень конвейера начинает переключаться в рабочую (спейсерную) фазу только после того, как предыдущая ступень конвейера переключилась в рабочую (спейсерную) фазу, а следующая ступень конвейера

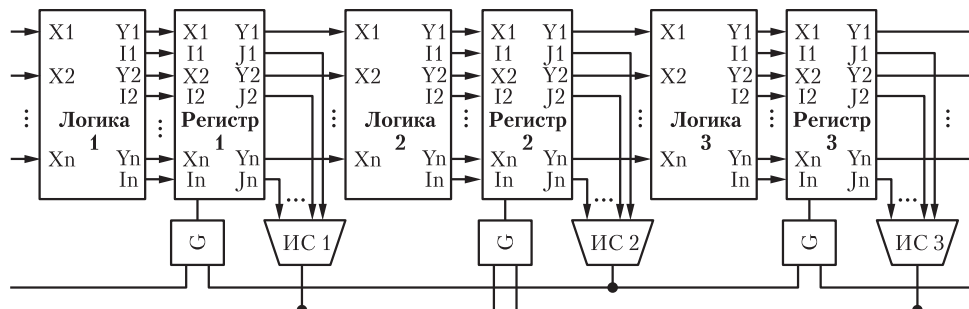


Рис. 1 Схема классического СС-конвейера

переключилась в спейсерную (рабочую) фазу. На рис. 1 приведена схема классического СС-конвейера. Здесь «Логика» — комбинационная часть ступени конвейера; «Регистр» — выходной регистр ступени; «ИС» — индикаторная схема ступени конвейера; «Г» — гистерезисный триггер (Г-триггер) [4], формирующий сигнал управления фазой регистра. Индикаторная схема является общей для комбинационной части и регистра ступени конвейера.

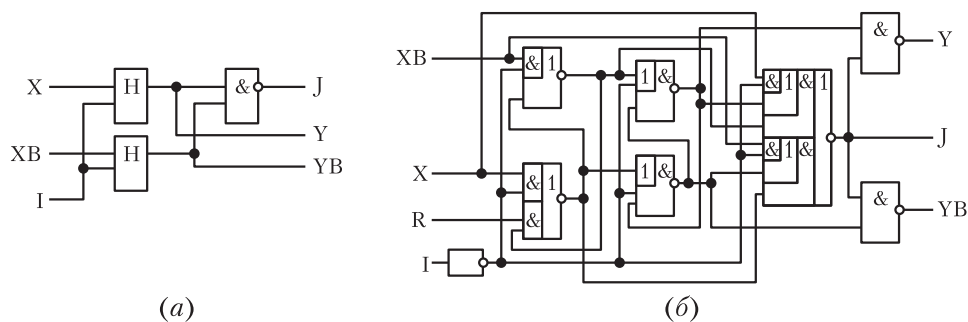
Такая дисциплина управления СС-конвейером привязывает начало переключения текущей ступени конвейера к общим индикаторным выходам соседних ступеней конвейера и существенно замедляет его работу. Действительно, пусть все ступени конвейера характеризуются примерно одинаковыми задержками переключения (ЗП) информационных выходов в рабочую фазу и в спейсер и одинаковыми задержками формирования индикаторных выходов ступени после готовности информационных выходов:

- $T_{РФЛ}$  — ЗП комбинационной части ступени конвейера в рабочую фазу;
- $T_{СФЛ}$  — ЗП комбинационной части ступени конвейера в спейсер;
- $T_{РФР}$  — ЗП регистра ступени конвейера в рабочую фазу;
- $T_{СФР}$  — ЗП регистра ступени конвейера в спейсер;
- $T_{РФИ}$  — ЗП индикаторной схемы ступени конвейера в рабочую фазу;
- $T_{СФИ}$  — ЗП индикаторной схемы ступени конвейера в спейсер.

Тогда при строгом соблюдении базового принципа взаимодействия ступеней СС-конвейера длительность цикла работы (рабочая фаза плюс спейсерная фаза) каждой ступени классического СС-конвейера будет определяться формулой:

$$T_{РЦК} = T_{РФР} + T_{РФИ} + T_{СФР} + T_{СФИ} + 2T_{G2} + 2 \max \{ T_{СФР} + T_{СФЛ} + T_{СФИ}, T_{РФЛ} + T_{РФР} + T_{РФИ} \}, \quad (1)$$





**Рис. 2** Схема разряда регистра ступени конвейера: (а) оптимизированная; (б) классическая

где  $T_{G2}$  — усредненная задержка переключения двухвходового Г-триггера (G2). Здесь задержка переключения индикаторной схемы ступени конвейера представляет собой интегрированную задержку формирования общего индикаторного выхода ступени конвейера по отношению к максимально задержанному выходу из числа выходов комбинационной части и регистра. Однако выбор парафазной со спейсером (ПФС) дисциплины формирования входных данных комбинационной части ступени конвейера и специальная реализация регистра позволяют ускорить взаимодействие ступеней конвейера.

Действительно, ПФС-дисциплина сигналов отображает не только бит передаваемой информации, но и фазу, в которой находится сигнал в любой момент времени. Таким образом, выходы комбинационной части ступени конвейера своим состоянием сигнализируют об окончании своего переключения в следующую фазу работы.

В случае многоразрядной СС-схемы с отсутствующей или слабой взаимосвязью разных разрядов это свойство ПФС-сигналов может быть использовано для поразрядной индикации ступени конвейера за счет реализации разряда регистра на основе Г-триггеров [5]. На рис. 2, а показана схема разряда регистра хранения ПФС-сигнала с единичным спейсером на Г-триггерах, имеющая общий сигнал управления I и индикаторный выход J. Она хранит как рабочее, так и спейсерное состояние ПФС-сигнала, и ее выход (Y, YB) является также ПФС-сигналом. Она выгодно отличается от традиционного СС-триггера на бистабильных ячейках с информационным ПФС-входом, пример которого показан на рис. 2, б.

При использовании схемы рис. 2, а в качестве разряда регистра ступени конвейера структурная схема СС-конвейера преобразуется к виду, изображенному на рис. 3. Индикаторные схемы комбинационной части («ИЛ») и регистра («ИР») ступени конвейера формируют отдельные индикаторные сигналы, объединяемые трехвходовым Г-триггером (G3) вместе с индикатором регистра соседней ступени. Явной зависимости сигнала управления фазой регистра от индикатор-

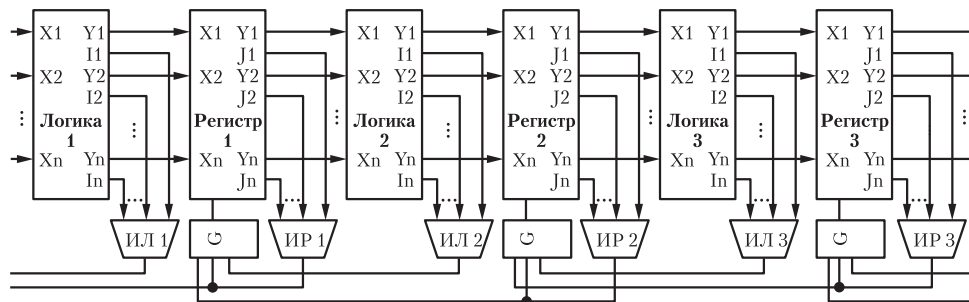


Рис. 3 Схема оптимизированного СС-конвейера

ного выхода предыдущей ступени конвейера нет, но схема конвейера остается СС, так как регистр текущей ступени конвейера не переключится в очередную фазу работы, пока в ту же фазу не переключатся информационные выходы предыдущей ступени конвейера. Однако сигнал управления фазой работы регистра по-прежнему остается общим для всех разрядов регистра и формируется по окончании переключения всех разрядов комбинационной части и регистра данной ступени.

Длительность цикла работы ступени оптимизированного СС-конвейера равна

$$T_{РЦО} = T_{РФР} + T_{РФИР} + T_{СФР} + T_{СФИР} + 2T_{G3} + 2 \max \{T_{СФР} + T_{СФЛ} + T_{СФИР}, T_{РФЛ} + T_{РФР} + T_{РФИР}\} \quad (2)$$

в предположении, что формирование индикаторного выхода комбинационной части в любой фазе заканчивается не позже, чем формирование индикаторного выхода последующего регистра. Здесь  $T_{РФИР}$  и  $T_{СФИР}$  — задержки переключения индикаторной схемы регистра ступени конвейера в рабочую фазу и в спейсер.

Схема СС-конвейера на рис. 3 оказывается более быстродействующей по сравнению со схемой рис. 1, если выполняется неравенство  $T_{РЦК} - T_{РЦО} > 0$ , которое после подстановки формул (1) и (2) преобразуется к виду:

$$(T_{РФИ} - T_{РФИР}) + (T_{СФИ} - T_{СФИР}) + 2(T_{G2} - T_{G3}) + 2 \max \{T_{СФР} + T_{СФЛ} + T_{СФИ}, T_{РФЛ} + T_{РФР} + T_{РФИ}\} - 2 \max \{T_{СФР} + T_{СФЛ} + T_{СФИР}, T_{РФЛ} + T_{РФР} + T_{РФИР}\} > 0. \quad (3)$$

В неравенстве (3) задержка  $T_{РФИР}$  заведомо меньше задержки  $T_{РФИ}$ , а задержка  $T_{СФИР}$  гарантированно меньше задержки  $T_{СФИ}$ . Пусть  $T_{РФИР} = K_{РИ}T_{РФИ}$ ,  $T_{СФИР} = K_{СИ}T_{СФИ}$ , где  $K_{РИ}$  и  $K_{СИ}$  — коэффициенты пропорциональности, меньшие единицы, а отношение задержек переключения в рабочую фазу и в спейсер примерно одинаково для комбинационной части и регистра:  $T_{РФР} = K_{РС}T_{СФР}$ ,  $T_{РФЛ} = K_{РС}T_{СФЛ}$ . Тогда неравенство (3) придет к виду:

$$(1 - K_{РИ}) T_{РФИ} + (1 - K_{СИ}) T_{СФИ} + 2(T_{G2} - T_{G3}) + \\ + 2 \max \{T_{СФР} + T_{СФЛ} + T_{СФИ}, K_{РС} (T_{СФР} + T_{СФЛ}) + T_{РФИ}\} - \\ - 2 \max \{T_{СФР} + T_{СФЛ} + K_{СИ} T_{СФИ}, K_{РС} (T_{СФР} + T_{СФЛ}) + K_{РИ} T_{РФИ}\} > \\ > 0. \quad (4)$$

Статистически задержка переключения СС-схемы в спейсер меньше задержки ее переключения в рабочую фазу, а задержки переключения индикаторной схемы в рабочую фазу и в спейсер примерно одинаковы. Тогда  $K_{РС} > 1$ ,  $T_{РФИ} = T_{СФИ} = T_{И}$ ,  $K_{РИ} = K_{СИ} = K_{И}$  и неравенство (4) запишется в виде:

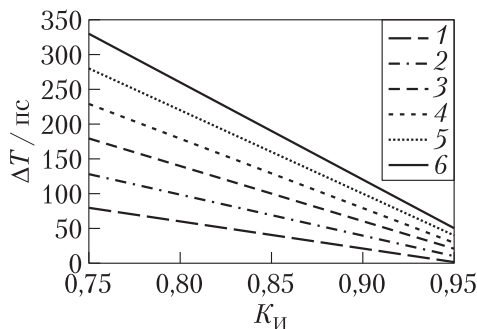
$$\Delta T = 4(1 - K_{И}) T_{И} + 2(T_{G2} - T_{G3}) > 0. \quad (5)$$

В многоразрядной СС-схеме число сигналов, объединяемых в один общий индикаторный сигнал, пропорционально разрядности схемы. Из-за ограничения на допустимое число последовательно соединенных КМОП-транзисторов в элементе (не более трех транзисторов р-типа и не более четырех транзисторов п-типа) индикаторная схема строится на основе двухвходовых и трехвходовых Г-триггеров. Индикаторная схема, свертывающая  $M$  индицирующих сигналов в один, реализуется на  $\lceil M(1 - 1/\log M) \rceil$  G2 или  $\lceil (M/2)(1 - 1/\log_3 M) \rceil$  G3, где скобки  $\lceil \rceil$  означают «ближайшее целое, не меньшее значения выражения в скобках». Число ярусов такой подсхемы равно  $\lceil \log M \rceil$  [или]  $\lceil \log_3 M \rceil$  соответственно, а задержка переключения индикаторной подсхемы в первом приближении равна

$$T_{И} = T_{G3} \lceil \log_3 M \rceil.$$

Например, число индикаторных сигналов первого слоя «дерева» Уоллеса умножителя двойной точности, соответствующего стандарту IEEE754 [6], реализующего модифицированный алгоритм Бута и использующего парафазное кодирование с единичным спейсером, равно 1431. Свертывающая их индикаторная схема будет иметь 716 G13 и G12, расположенных на 7 ярусах пирамидальной структуры. При задержке одного G3 в 65-нанометровой КМОП-технологии, равной 50 пс при типовых условиях, суммарная задержка переключения в одну сторону такой индикаторной схемы составляет около 350 пс, а в обе стороны — 700 пс, что приводит к серьезному замедлению конвейера умножителя.

На рис. 4 приведены графики оценки левой части ( $\Delta T$ ) неравенства (5) в зависимости от коэффициента  $K_{И}$  для разных значений разрядности  $M$  регистра



**Рис. 4** Выигрыш в быстродействии оптимизированного СС-конвейера: 1 —  $M = 3^2$ ; 2 —  $3^3$ ; 3 —  $3^4$ ; 4 —  $3^5$ ; 5 —  $3^6$ ; 6 —  $M = 3^7$

ступени СС-конвейера при  $T_{Г2} \approx 40$  пс и  $T_{Г3} \approx 50$  пс. В качестве значений  $M$  выбраны степени тройки: индикаторные «деревья» для такого числа сигналов, реализованные на ГЗ, оптимальны как по быстродействию, так и по аппаратным затратам. Графики  $\Delta T(K_{И})$  для остальных значений  $M_i$ , таких что  $M_1 < M_i < M_2$ , где  $M_1$  и  $M_2$  — значения из ряда степеней тройки, занимают промежуточные положения между соответствующими графиками для  $M_1$  и  $M_2$ .

Использование поразрядной индикации обеспечивает сокращение вклада индикаторной схемы в задержку переключения многоразрядной СС-схемы.

### 3 Поразрядная индикация ступени конвейера

Структурная схема СС-конвейера с поразрядной индикацией показана на рис. 5. Здесь  $n$  — разрядность тракта обработки информации;  $J_1, \dots, J_n$  — поразрядные индикаторные выходы регистра ступени конвейера;  $K_1, \dots, K_n$  — поразрядные фазовые входы регистра. Поразрядная индикация позволяет инициировать переключение СС-схемы в противоположную фазу работы, не дожидаясь окончания переключения в текущую фазу всех разрядов регистра. Максимальный эффект от использования поразрядной индикации достигается в СС-схемах с абсолютной независимостью соседних разрядов по входам-выходам. При наличии взаимной зависимости разрядов растут аппаратные затраты на поразрядную индикацию и уменьшается выигрыш по быстродействию.

Следует отметить, что использование независимой поразрядной индикации в качестве запрос-ответного взаимодействия соседних ступеней СС-конвейера ограничено только внутренней областью неразветвленного конвейера. Входной

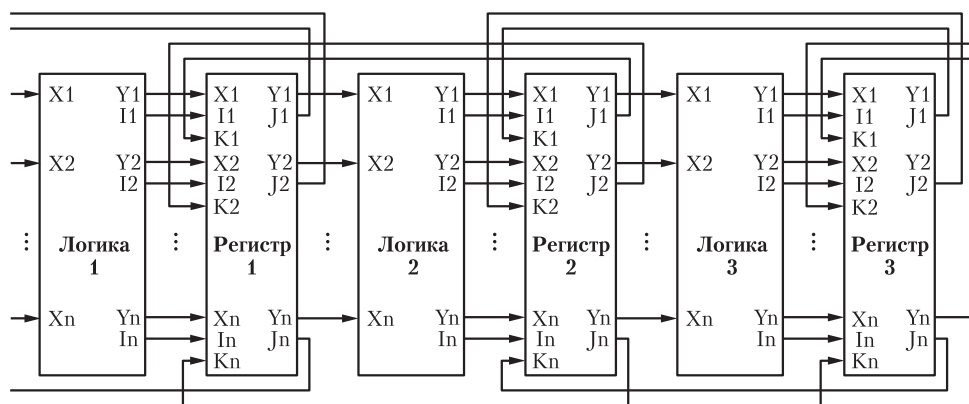


Рис. 5 Схема СС-конвейера с поразрядной индикацией

и выходной блоки такого конвейера обязаны формировать общий индикаторный сигнал для всех разрядов регистра.

Пусть СС-схема имеет  $N$  разрядов и каждые  $L$  соседних разрядов комбинационной части ступени конвейера связаны друг с другом: выходы элементов в младших разрядах подмножеств из  $L$  разрядов используются как входы в старших разрядах этих же подмножеств. Тогда индикация окончания переключения  $L$  соседних разрядов комбинационной части и выходного регистра ступени конвейера в очередную фазу работы гарантирует готовность результата в этих  $L$  разрядах и разрешает переключение самого младшего из них в противоположную фазу. Для полной поразрядной индикации всех  $N$  разрядов ступени конвейера потребуется  $(N - L + 1)$  таких индикаторных подсхем.

Длительность цикла работы ступени СС-конвейера с поразрядной индикацией определяется формулой:

$$T_{РЦП} = T_{РФР} + T_{РФПИР} + T_{СФР} + T_{СФПИР} + 2T_{G2} + \\ + 2 \max \{ T_{СФР} + T_{СФЛ} + T_{СФПИР}, T_{РФЛ} + T_{РФР} + T_{РФПИР} \},$$

где  $T_{РФПИР}$  и  $T_{СФПИР}$  — задержки переключения поразрядных индикаторных выходов регистра ступени конвейера в рабочую фазу и в спейсер соответственно относительно информационных выходов. С учетом вышеприведенных допущений в отношении задержек разного типа в предельном случае (полная независимость соседних разрядов ступени,  $L = 1$ )  $T_{РФПИР} = T_{СФПИР} = T_{G2}$ . В общем случае

$$T_{РФПИР} = T_{СФПИР} = T_{G2} + \lceil \log_3 O \rceil T_{G3}, \quad (6)$$

в то время как аналогичные величины в формуле (2) имеют значение

$$T_{РФИР} = T_{СФИР} = \lceil \log_3 N \rceil T_{G3}. \quad (7)$$

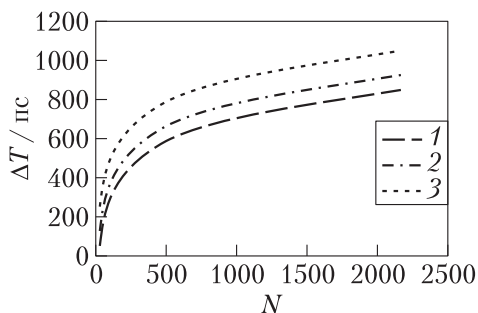
Тогда выигрыш в быстродействии варианта СС-конвейера с поразрядной индикацией можно оценить формулой:

$$T_{РЦО} - T_{РЦП} = T_{СФИР} - T_{СФПИР} + 2(T_{G3} - T_{G2}) + 3(T_{РФИР} - T_{РФПИР}).$$

С учетом формул (6) и (7) получается:

$$\Delta T_{П} = T_{РЦО} - T_{РЦП} = 2(T_{G3} - T_{G2}) + \\ + 4(\lceil \log_3 N \rceil T_{G3} - \lceil \log_3 O \rceil T_{G3}). \quad (8)$$

Рисунок 6 демонстрирует зависимость  $\Delta T_{П}$ , задаваемую формулой (8), от суммарной разрядности тракта обработки данных  $N$  для разных вариантов межразрядной связности  $O$  при фиксированных значениях  $T_{G2} = 40$  пс и  $T_{G3} = 50$  пс.



**Рис. 6** Выигрыш в быстродействии ступени СС-конвейера с поразрядной индикацией: 1 —  $O = 9$ ; 2 — 6; 3 —  $O = 3$

Если  $L < O = (\lceil \log_3 L \rceil)^3$ , то с точки зрения минимизации аппаратных затрат целесообразно индцировать одним общим сигналом  $O$  соседних разрядов, а не  $L$ . Тогда завершение переключения элементов комбинационной части и регистра ступени конвейера в  $O$  соседних разрядах разрешает переключение в противоположную фазу  $(O - L + 1)$  разрядов и аппаратные затраты снижаются до  $C_0 = 16((N - O + 1)/(O - L + 1)) \log_3 O$ . Например, при  $N = 64$ ,  $L = 5$  и  $O = 9$  сложность минимизированной поразрядной индикаторной схемы снижается в 3,9 раза. В этом случае целесообразно говорить о *групповом* индикаторе, так как он управляет фазовыми переключениями группы соседних разрядов.

На практике вариант поразрядной индикации, показанный на рис. 5, встречается редко и может рассматриваться как совокупность параллельных СС-конвейеров. Гораздо чаще встречаются случаи конвейерной обработки многоразрядных цифровых данных, например устройство умножения со сложением-вычитанием (Fused-Multiply-Add, FMA) [7]. В нем умножитель построен на основе многоразрядных сумматоров с сохранением переноса. В результате разряды сумматоров оказываются слабо связанными и допускают применение поразрядной индикации в пределах одной-двух ступеней конвейера.

В ряде случаев целесообразно использовать вариант поразрядной индикации, ориентированный на ускорение переключения ступеней конвейера, длительность рабочего цикла которых намного превышает длительность цикла соседних ступеней.

В формуле (2) длительность цикла работы ступени оптимизированного СС-конвейера в значительной степени определяется задержкой переключения комбинационной части ступени в спейсер ( $T_{СФЛ}$ ) и рабочую фазу ( $T_{РФЛ}$ ). Во многих схемах задержка срабатывания комбинационной части ступени конвейера превалирует над задержками регистра этой же ступени. Это создает базу для применения поразрядной индикации в виде, представленном на рис. 7.

Аппаратные затраты поразрядной индикации избыточны, поскольку один и тот же разряд схемы индцируется одновременно в  $L$  индикаторных подсхемах, и равны примерно  $C_L = 16(N - L + 1) \log_3 L$  КМОП-транзисторов. Но такой способ индикации позволяет ускорить переключение комбинационной части ступени конвейера в рабочую фазу и спейсер на время, соответствующее ожиданию формирования общего индикаторного выхода регистра и комбинационной части ступени для всех разрядов.

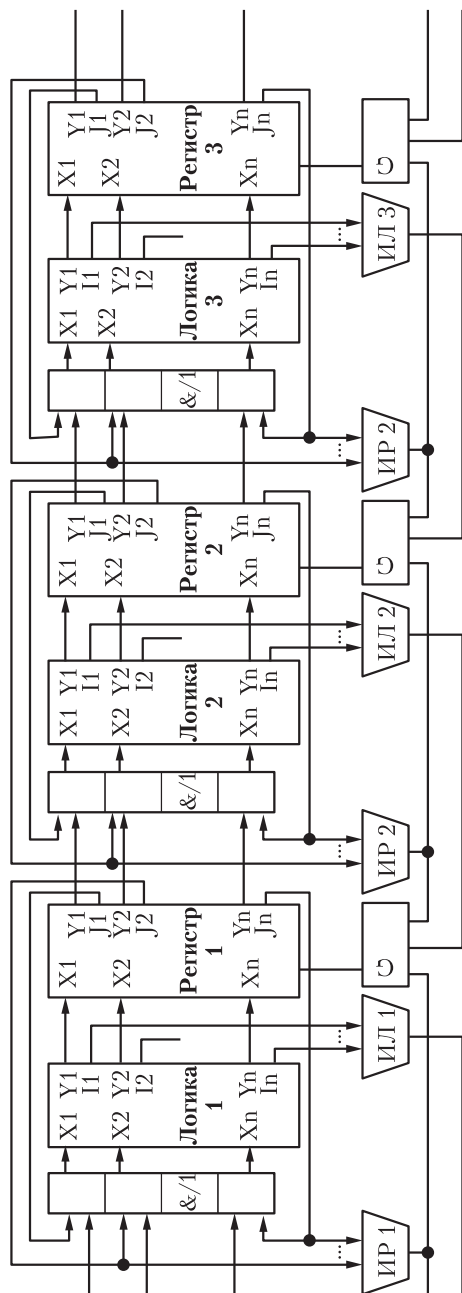


Рис. 7 Схема конвейера с поразрядной индикацией комбинированной части

Цель ее заключается в том, чтобы исключить из критического пути ЗП комбинационной части ступени конвейера из рабочей фазы в спейсер. По сравнению со схемой СС-конвейера в схему конвейера на рис. 5 добавлены блоки принудительного перевода комбинационной части каждой ступени в спейсер (БПС), обозначенные символами «&/1». Они реализуются на  $2N$  элементах 2И, 2И-НЕ, 2ИЛИ или 2ИЛИ-НЕ в зависимости от типа спейсера формируемых ПФС — входов комбинационной части ступени конвейера. По завершении переключения в рабочую фазу разрядов регистра ступени, подтвержденном индикатором совокупности  $O$  соседних разрядов, соответствующие разряды предшествующей комбинационной части принудительно переводятся в спейсер. В этом случае переключение комбинационной части ступени конвейера в спейсер происходит в «фоновом режиме», одновременно с переключением следующей ступени конвейера в рабочую фазу. В результате длительность рабочего цикла ступени конвейера описывается формулой:

$$T_{РЦПЛ} = T_{РФР} + T_{РФИР} + T_{СФР} + T_{СФИР} + 2T_{Гз} + T_{БПС} + 2 \max \{ T_{СФР} + T_{СФИР}, T_{РФЛ} + T_{РФР} + T_{РФИР} \}, \quad (9)$$

где  $T_{БПС}$  — усредненная задержка срабатывания элементов БПС в рабочую фазу (время переключения БПС в спейсер маскируется временем переключения соседних ступеней конвейера).

Сравнение формул (9) и (2) показывает, что применение поразрядной индикации целесообразно только тогда, когда

$$T_{СФЛ} > T_{РФЛ} + T_{РФР} + T_{РФИР} - T_{СФР} - T_{СФИР} - T_{БПС}.$$

В противном случае вместо ускорения работы получится замедление на величину  $T_{БПС}$ . Таким образом, оптимизация СС-конвейера и использование поразрядной индикации позволяют повысить быстродействие СС-конвейера тем больше, чем выше разрядность тракта цифровой обработки данных.

## 4 Заключение

«Узким местом» многоразрядных цифровых СС-схем является индикаторная подсхема. Необходимость индицирования окончания переходных процессов во всех элементах СС-схемы и организации запрос-ответного взаимодействия между СС-блоками вынуждает разработчика тратить дополнительные аппаратурные и временные ресурсы, формируя индикаторную подсхему. Она собирает все внутренние индикаторные сигналы схемы в один общий сигнал, используемый в качестве сигнала управления предшествующими и последующими функциональными блоками (ступенями конвейера) в тракте цифровой обработки данных.

Применение оптимизированной структуры СС-конвейера и поразрядной индикации в многоразрядных схемах позволяет существенно ускорить их работу



за счет относительно небольшого увеличения аппаратных затрат. Поразрядную индикацию целесообразно использовать в схемах, в которых соседние разряды схемы слабо связаны друг с другом общими сигналами.

## Литература

1. *Kishinevsky M., Kondratyev A., Taubin A., Varshavsky V.* Concurrent hardware: The theory and practice of self-timed design. — New York, NY, USA: J. Wiley & Sons, 1994. 368 p.
2. *Соколов И. А., Степченко Ю. А., Петрухин В. С., Дьяченко Ю. Г., Захаров В. Н.* Самосинхронная схемотехника — перспективный путь реализации аппаратуры // Системы высокой доступности, 2007. Т. 3. № 1-2. С. 61–72.
3. *Степченко Ю. А., Дьяченко Ю. Г., Горелкин Г. А.* Самосинхронные схемы — будущее микроэлектроники // Вопросы радиоэлектроники, 2011. № 2. С. 153–184.
4. *Степченко Ю. А., Дьяченко Ю. Г., Денисов А. Н., Фомин Ю. П.* Г-триггер: Патент РФ № 2371842 с приоритетом от 15.11.2007. Опубликовано 27.10.2009.
5. *Степченко Ю. А., Дьяченко Ю. Г., Рождественский Ю. В., Морозов Н. В., Степченко Д. Ю., Рождественские А. В., Сурков А. В.* Самосинхронный умножитель с накоплением: варианты реализации // Системы и средства информатики, 2014. Т. 24. № 3. С. 63–77.
6. IEEE Standard for Floating-Point Arithmetic IEEE Std 754-2008. — IEEE Computer Society, 2008. doi: 10.1109/IEEESTD.2008.4610935.
7. *Stepchenkov Y. A., Zakharov V. N., Rogdestvenski Y. V., Diachenko Y. G., Morozov N. V., Stepchenkov D. Y.* Speed-independent floating point coprocessor // IEEE East-West Design & Test Symposium Proceedings. — IEEE, 2015. P. 111–114.

Поступила в редакцию 11.09.19

---

---

## INDICATION OPTIMIZATION IN MULTIBIT SELF-TIMED CIRCUITS

*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:** Indication subcircuit in self-timed (ST) circuits provides both control of the completion of switching all their cells to the current phase and control of their functional blocks interaction. An increase of the ST-circuit capacity leads to rising contribution of the indication subcircuit to the circuit’s transient delay. The paper discusses the optimization of the indication subcircuit and

the organization of the ST-pipeline intended for improving performance of the whole circuit. Register bit for storing intermediate data in the pipeline stages is implemented on the basis of hysteretic trigger instead of traditional RS-trigger. Such register bit has less complexity and provides storing both the work and the spacer states of a dual-rail data. Discipline of the pipeline stage phase control by means of total indication outputs of the adjacent pipeline stages is replaced with bit-wise indication and bit-wise control principle, which utilizes the parallelism of the calculations in the multibit ST-circuits. Proposed solutions essentially improve ST-circuits performance due to a slight complication of the indication subcircuit.

**Keywords:** self-timed circuit; pipeline; hysteretic trigger

**DOI:** 10.14357/08696527190402

## Acknowledgments

The study was supported by 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*. New York, NY: John Wiley & Sons. 368 p.
2. Sokolov, I. A., Yu. A. Stepchenkov, V. S. Petrukhin, Yu. G. Diachenko, and V. N. Zakharov. 2007. Samosinkhronnaya skhemotekhnika — perspektivnyy put' realizatsii apparatury [Self-timed circuitry is the perspective way for hardware realization]. *Sistemy vysokoy dostupnosti* [Highly Available Systems] 2(1-2):61–72.
3. Stepchenkov, Yu. A., Yu. G. Diachenko, and G. A. Gorelkin. 2011. Samosinkhronnye skhemy — budushchee mikroelektroniki [Self-timed circuits are the future of microelectronics]. *Voprosy radioelektroniki* [Issues of Radio Electronics] 2:153–184.
4. Stepchenkov, Yu. A., Yu. G. Diachenko, A. N. Denisov, and Yu. P. Fomin. 2009. G-trigger [G-trigger]. Patent RF No. 2371842.
5. Stepchenkov, Y. A., Y. G. Diachenko, Y. V. Rogdestvenski, N. V. Morozov, D. Y. Stepchenkov, A. V. Rogdestvenskene, and A. V. Surkov. 2014. Samosinkhronnyy umnozhitel' s nakopleniem: varianty realizatsii [Self-timed fused multiply-add unit: Implementation variants]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 24(3):63–77.
6. IEEE Computer Society. 2008. IEEE Standard for Floating-Point Arithmetic IEEE Std 754-2008. doi: 10.1109/IEEESTD.2008.4610935.
7. Stepchenkov, Y. A., V. N. Zakharov, Y. V. Rogdestvenski, Y. G. Diachenko, N. V. Morozov, and D. Y. Stepchenkov. 2015. Speed-independent floating point coprocessor. *East-West Design and Test Symposium Proceedings*. IEEE. 111–114.

Received September 11, 2019

## Contributors

**Stepchenkov Yuri A.** (b. 1951) — 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; 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, 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, 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

## ПОИСК ЭМПИРИЧЕСКИХ ПРИЧИН СБОЕВ И ОШИБОК В КОМПЬЮТЕРНЫХ СИСТЕМАХ И СЕТЯХ С ИСПОЛЬЗОВАНИЕМ МЕТАДАННЫХ\*

*А. А. Грушо<sup>1</sup>, М. И. Забейло<sup>2</sup>, Н. А. Грушо<sup>3</sup>, Е. Е. Тимонина<sup>4</sup>*

**Аннотация:** Управление распределенными информационно-вычислительными системами (РИВС) предполагает решение двух взаимосвязанных задач: бесперебойной работы системы и ее безопасности. Обе задачи используют схожие методы для выявления причин сбоев и инцидентов безопасности. Целью исследования является адекватная формализация, которую в дальнейшем можно использовать для автоматизации хотя бы части функций поиска причин сбоев и инцидентов безопасности. Основываясь на взаимосвязи функций системного администратора (СА) и методах построения метаданных для управления соединениями в РИВС, предложен способ иерархического построения метаданных и нового способа их применения, позволяющий осуществлять поиск причин сбоев и управлять решением задач в таких системах. Построена модель и алгоритм поиска эмпирической причины сбоя, которые основаны на метаданных и резервировании задач и данных. Показано, что возможно построение дерева решений выявления эмпирической причины неявного сбоя с точностью до имеющейся детализации сбойной информационной технологии. При этом выделяются типовые задачи анализа причинно-следственных связей, из которых можно конструировать весь поиск.

**Ключевые слова:** информационная безопасность; метаданные; причинно-следственные связи; системное администрирование; неявные сбои и ошибки

**DOI:** 10.14357/08696527190403

### 1 Введение

Управление РИВС предполагает решение двух взаимосвязанных задач: бесперебойной работы системы и ее безопасности. Обе задачи требуют исчерпывающих данных для выявления причин сбоев и инцидентов безопасности.

---

\*Работа частично поддержана РФФИ (проект 18-07-00274-а).

<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, m.zabehailo@yandex.ru

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

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

Системный администратор в своей работе по устранению и предотвращению сбоев в компьютерных системах сталкивается со сложным взаимодействием множества программно-аппаратных элементов. Если аппаратное или программное обеспечение не предоставляет достаточно данных для выявления сбоя или нарушения безопасности, то СА и офицеру службы безопасности приходится проверять каждый элемент компьютерной системы, связанный с решением конкретной задачи.

Некоторые элементы системы способны предоставлять подробную диагностическую информацию (журналы работы, диагностическая печать на принтерах), другим же требуется специальное программно-аппаратное обеспечение для отображения такой информации, например специализированные тестеры, адаптеры, отладчики, программы и т. д. Но чем больше элементов в компьютерной системе, тем больше диагностических средств требуется для получения нужной информации.

Работа посвящена построению формализованной модели поиска причин неявных сбоев и инцидентов информационной безопасности. Близость понятий сбоя и инцидента информационной безопасности позволяет провести исследования только неявных сбоев. В качестве исходных данных используются данные о запросах к СА на выявление причин сбоев в компьютерных системах, подключенном к ним периферийном оборудовании и программном обеспечении. Материал собирался в течение 6 мес., на протяжении которых поступило более 1000 таких запросов. В разд. 2 приведен только один пример неявного сбоя, информация о котором позволяет объяснить параметры предлагаемой модели и возможности ее использования. Целью исследования является адекватная формализация, которую в дальнейшем можно использовать для автоматизации хотя бы части функций поиска причин сбоев и инцидентов безопасности.

Средства автоматизации для анализа эмпирических причинно-следственных связей традиционно относятся к интеллектуальному анализу данных [1]. Метаданные и информационные пространства были определены в работах авторов [2, 3]. Иерархическое представление информационных технологий основано на понятии составной задачи [4], для которой из математической модели бизнес-процесса формируются метаданные и редукция задач [5]. Близкие по идеологии исследования причинно-следственных связей можно найти в работах, посвященных FCA (Formal Content Analysis) [6], аргументационному анализу [7] и др.

Основываясь на взаимосвязи функций СА и методах построения метаданных для управления соединениями в РИВС [8], предложен способ иерархического построения метаданных и нового метода их применения, позволяющий осуществлять поиск причин сбоев и управлять решением задач в РИВС.

## **2 Пример поиска неявной причины сбоя в информационной технологии**

На одном из выделенных серверов клиента установлено программное обеспечение (ПО) для резервного копирования. Данное ПО по заданному расписанию

должно осуществлять резервное копирование ценных данных и состояния операционной системы (ОС) на дисковое хранилище. Системный администратор производит наблюдение за системой резервного копирования с помощью специальной консоли, в которой отображается ход выполнения заданий и возникающие ошибки. Системный администратор обнаружил, что одно из выполняемых заданий завершилось с ошибкой «Диспетчер DPM обнаружил несовпадение контрольных сумм при передаче данных по сети (Идентификатор 30087 Сведения: Код внутренней ошибки: 0x809909B2)», т. е. при передаче возник какой-то сбой, помешавший нормальному завершению задачи резервного копирования. Но причина этого сбоя находится вне границ наблюдения системы, а коды ошибок не расшифрованы. После перезагрузки серверов, агентов и пр. ошибка не пропала. В процессе наблюдения за сервером, который подлежал резервированию, выяснилось, что на нем был запущен процесс, который очень сильно загружал процессор, диск и сетевой адаптер. После остановки этого процесса СА перезапустил задание резервного копирования, и оно выполнилось успешно. Даже после возобновления процесса, использовавшего большую часть ресурсов сервера, резервное копирование продолжалось без ошибок. Так как сервер клиента не может быть остановлен для более детального выяснения причины такого поведения и проведения экспериментов, СА пришлось прекратить работы по дальнейшему выявлению причин этого сбоя. Каким образом СА определил взаимосвязь задания резервного копирования с повышенной нагрузкой на сервер? Рассуждения СА начинаются с определения элементов, участвующих в процессе выполнения резервного копирования. Данными элементами являются:

- ПО для резервного копирования;
- ОС этого сервера;
- дисковое хранилище резервных копий;
- локальная компьютерная сеть;
- сервер, подлежащий резервированию, его ОС;
- прикладные программы и данные.

Ошибка указывала на проблему с передачей по локальной компьютерной сети, но сама по себе сеть функционировала нормально при выполнении простых задач, например открытие веб-сайта, доступ к файлам на удаленном сервере. Оперативного физического доступа к локальной сети не было. Таким образом, из взаимодействующих элементов остался только удаленный сервер с выполняемым процессом, сильно использовавший ресурсы, включая локальную сеть. Временное исключение нагрузки на сеть с данного сервера стало ключевым для устранения сбоя.

Для формализации действий СА необходимо описать все задачи, которые в ходе поиска причин сбоя решил СА, и исходные данные, которые СА использовал. Для рассмотренного примера это следующий перечень задач.

1. *Восстановление перечня и порядка решения задач в рассматриваемой информационной технологии резервного копирования.* Отметим, что описание порядка решения задач с однозначно определенной степенью детализации не всегда возможно [9]. В данном примере рассматривается простейшая конфигурация задач, когда все данные решения предыдущей задачи являются входными данными следующей задачи. Информация о порядке решения задач содержится в метаданных, что позволяет их использовать в анализе причин сбоев. В рассматриваемом случае СА знает последовательность решаемых задач, которая в агрегированном виде может быть записана следующим образом:
  - задача формирования данных для резервирования с помощью агента на сервере 1;
  - задача передачи данных резервной копии с сервера 1 на сервер с хранилищем резервных данных;
  - задача сохранения резервной копии в хранилище резервных данных на сервере 2.
2. *Проверка правильности передачи данных по сети.* Проверка показала, что простые задачи с помощью сети решаются правильно, т. е. сама сеть и сетевое оборудование не могут быть причиной сбоя.
3. *Проверка работы сервера, хранящего резервные копии.* Задача, помеченная в метаданных как запись резервной копии в хранилище, детализировалась с использованием иерархической декомпозиции структуры сервера:
  - ПО резервного копирования в хранилище;
  - ОС сервера 2;
  - дисковое хранилище резервных копий;
  - аппаратная часть сервера 2 (центральный процессор, адаптер сети и т. д.).Для решения этих задач используются тесты, построенные на ранее решенных задачах, и вычисляются показатели, традиционные для штатной работы системы.
4. *Проверка работы сервера 1* также осуществлялась с помощью детализации метаданных в агрегированной схеме, приведенной выше. Иерархическая декомпозиция задач на сервере 1 описывается следующим образом:
  - прикладные программы и данные, подлежащие копированию;
  - ОС и драйверы устройств;
  - аппаратная платформа, в которую входят адаптеры локальной сети.

В множество параметров проверки аппаратной платформы входят показатели загрузки центрального процессора, диска и сети. Сравнение значений

этих параметров со средними значениями позволяет выявить аномальную загруженность этих устройств.

5. *Выявление причины аномальной загруженности центрального процессора, диска и адаптера локальной сети.* Эту задачу можно решать перебором запущенных процессов на данном сервере. Решением этой задачи является определение процесса, вызывающего аномальную загрузку. Для определения того, что данный процесс стал причиной сбоя задачи резервного копирования, можно временно приостановить данный процесс и повторно запустить задачу резервного копирования. В этом случае задача резервного копирования успешно завершена. Возобновление приостановленного процесса не помешало повторному резервированию. Отсюда можно сделать вывод о сбое какого-то параметра конфигурации ПО, который был восстановлен при повторном запуске резервного копирования в условиях нормальной загрузки сервера 1.

С точки зрения проведенного анализа причин сбоя задачи резервного копирования нет основания считать, что поиск причин сбоя нельзя автоматизировать. Модель анализа причинно-следственных связей при поиске причин сбоя в РИВС рассмотрена в следующем разделе.

### 3 Модель поиска причин сбоев в распределенной информационно-вычислительной системе

Каждая задача  $C$  описывается следующим набором параметров [4]:

- $\varphi(C)$  — объект, определяющий взаимодействие с другими задачами и обеспечивающий задачу  $C$  входными данными и распределением выходных данных;
- $\pi(C)$  — ПО, реализующее преобразование информации;
- $\psi(C)$  — область определения задачи  $C$ ;
- $\gamma(C)$  — область значений преобразования  $\pi(C)$ .

Если  $x \in \psi(C)$ , то  $x = (y, z)$ , где  $z$  — значения конфигурационных параметров в задаче  $C$ , а  $y$  — данные, которые преобразует задача  $C$  с помощью преобразования  $C((\bullet, z))$ .

Составная задача  $A = \{B_1, \dots, B_s\}$  может быть представлена клиент-серверной архитектурой, в которой взаимодействие задач  $B_1, \dots, B_s$  описывается взаимодействием  $\varphi(A)$  и  $\varphi(B_i)$ ,  $i = 1, \dots, s$ . Простая составная задача  $A$  — это задача, подзадачи которой образуют связь  $A$  с аппаратной платформой и возвратом результата задаче  $A$ . Каждая задача представима корневым деревом редукции подзадач [5]. Глубина этого дерева определяется последовательным разбиением каждой задачи на подзадачи, расположенные на следующем уровне декомпозиции. Возможны повторы задач в дереве редукции, но с разными исходными данными. Последовательное построение решения осуществляется



соединением результатов решения подзадач и подъемом результата к корню дерева.

В дереве редукций все листья порождаются простыми составными задачами. Если дана ветка дерева редукций глубины  $h$ , то его поддереву, получаемое удалением вершин этой ветки на глубине  $h$ , называется агрегацией дерева редукции. Наоборот, построение ветки дерева редукций глубины  $h$  из ветки дерева глубины  $h - 1$  путем операции редукции задачи предыдущего слоя называется операцией детализации дерева редукции.

Операции детализации и агрегации задают иерархическую декомпозицию исходной задачи (частичный порядок). Сбой (отказ в решении) задачи на любой глубине дерева порождает существование сбоя во всех агрегациях данного дерева. Напомним, что метаданные управляют соединениями в сети [10] с помощью задач  $\mathcal{M}$  и  $\mathcal{N}$ , где  $\mathcal{M}$  управляет распределением задач по хостам, а  $\mathcal{N}$  выполняет функции разрешения или запрещения соединений между хостами.  $H(A)$  — бинарное отношение нахождения задачи  $A$  на хосте  $H(A)$ . Используя метаданные, можно автоматически определять непосредственные связи задач  $A_1, \dots, A_k$ , передающие исходные данные для решения задачи  $B$ .

Из введенных понятий очевидно следует связь бинарного отношения  $H(A)$  и объекта  $\varphi(A)$ , а именно: бинарное отношение  $H(A)$  через своего агента обеспечивает взаимодействие задачи  $A$  с задачами на других хостах, т. е. обеспечивает запуск задачи, завершение задачи, обеспечение исходными данными от других задач и распределение результатов вычислений другим задачам. Таким образом, можно считать, что объект  $\varphi(A)$  является обобщенным хостом, содержащим задачу  $A$ , тогда как в случае разрешительной системы с помощью метаданных можно обобщить функционал задач  $\mathcal{M}$  и  $\mathcal{N}$  на распределение и организацию взаимодействий всех задач информационной технологии на данном уровне иерархической декомпозиции этой технологии.

Для простоты будем далее считать, что  $k = 1$  и что в цепочке  $A \rightarrow B$  произошел сбой, причину которого надо определить и исправить. Без ограничения общности можно считать, что все рассматриваемые далее задачи находятся на разных обобщенных хостах. В этом состоит обобщение метаданных с проблемы разрешения соединений между хостами на проблему взаимодействий всех задач между собой. Такое обобщение возможно, если пользоваться иерархической декомпозицией решения задач.

В автоматическом режиме через агента в хосте  $H(\mathcal{N})$  задача  $\mathcal{N}$  получает от задачи  $H(B)$  сигнал сбоя задачи  $B$ . Предположим, что в системе предусмотрено резервирование задач  $A'$  для  $A$  и  $B'$  — для  $B$ . Технологию поиска причины сбоя и его исправления обозначим через  $\mathcal{G}$ . Системный администратор, получив информацию о наличии сбоя, запускает информационную технологию  $\mathcal{G}$  через задачу  $\mathcal{N}$ . Из метаданных информационная технология  $\mathcal{G}$  имеет следующие исходные данные:  $A \rightarrow B, A', B', H(A), H(B), H(A'), H(B')$ .

Сначала технология  $\mathcal{G}$  решает задачу проверки возможности передачи данных от задачи  $H(A)$  задаче  $H(B)$ . С этой целью задача  $\mathcal{G}$  через задачу  $\mathcal{N}$  посылает

агенту в хосте  $H(A)$  случайное число  $r$  и задание агенту в  $H(A)$  переслать это число в  $H(B)$ . Если это удастся, то канал связи с  $H(A)$  работает. Если агент в  $H(B)$  получает от  $H(A)$  число  $r'$  и отправляет его в  $\mathcal{N}$ , то возможны следующие варианты. Если  $r' = r$ , то каналы от  $H(A)$  к  $H(B)$  и от  $H(B)$  к  $\mathcal{N}$  работают правильно. Если  $r' \neq r$ , то сбой произошел либо в стеке сетевых протоколов  $H(A)$ , либо в стеке сетевых протоколов  $H(B)$ . Если эта ошибка повторяется, то необходимо перейти на резервный путь от  $H(A')$  к  $H(B')$ . Исходные данные задачи  $A$  надо передать задаче  $A'$ .

Если эта связь не обеспечила правильную передачу данных, то эмпирической причиной исходного сбоя стала либо ошибка в приложении, реализующем задачу  $A$ , либо ошибка в стеке протоколов связи хоста  $H(A)$ .

Если исходные данные задачи  $A$  переданы правильно задаче  $A'$ , то надо попытаться решить задачу  $B$  на хостах  $H(B')$  и  $H(B)$  с данными от  $A'$ . Если результаты решения задачи  $B$  на хостах  $H(B')$  и  $H(B)$  совпадают (сравнение осуществляется через задачу  $\mathcal{N}$ ), то можно считать, что сбой ликвидирован и исходную информационную технологию можно продолжать через задачу  $A'$ .

Если исходные данные задачи  $A$  неправильно переданы задаче  $A'$ , то, используя метаданные, можно найти задачу  $C$ , которая передавала исходные данные задаче  $A$ , и заменить исходную цепочку на  $C \rightarrow A' \rightarrow B'$ .

В исходном сервере с задачей  $A$  необходимо искать сбой в предположении, что  $A$  является простой составной задачей. Одним из параметров стандартного стека протоколов в этой проверке служит загруженность микропроцессорной техники. Аномалии выявляются сравнением загруженности микропроцессоров в прецедентах, взятых из памяти задачи  $\mathcal{G}$ . Выявив аномалии, необходимо методом перебора определить процессы, порождающие повышенную загрузку. Приостановка этих процессов и повторный запуск цепочки  $A \rightarrow B$  позволяет либо выявить эмпирическую причину сбоя, либо продолжить поиск сбойных конфигурационных параметров.

Необходимо отметить, что реализация построенного алгоритма поиска причин сбоя может не дать ожидаемого результата. В этом слабость поиска эмпирической причины и приведенных правдоподобных рассуждений, поэтому в данной работе используется термин «эмпирическая причина» вместо термина «причина». В случае негативного результата поиска необходимо строить более глубокое приближение исходной технологии в иерархической декомпозиции.

Таким образом, построена модель и алгоритм поиска эмпирической причины сбоя, которые основаны на метаданных и резервировании задач и данных. В частности, предложенный метод решает задачу поиска неявного сбоя в приведенном выше примере и вполне доступен для программной реализации. Фактически построено дерево решений на основе сравнения данных, наблюдаемых при сбое, и данных, полученных за счет резервирования.

## 4 Заключение

В работе исследована возможность автоматизации поиска эмпирических причин неявных сбоев и инцидентов информационной безопасности. В качестве исходных данных могут использоваться метаданные, порожденные бизнес-процессом, и архитектура их применения. Кроме этого необходима возможность дублирования задач и получения правильного их решения.

Показано, что возможно построение дерева решений выявления эмпирической причины неявного сбоя с точностью до имеющейся детализации сбойной информационной технологии. При этом выделяются типовые задачи анализа причинно-следственных связей, из которых можно конструировать весь поиск. В качестве компонентов алгоритмов поиска можно брать конкретные элементы программно-аппаратных решений.

## Литература

1. Финн В. К. Искусственный интеллект: Методология, применения, философия. — М.: Красанд, 2011. 448 с.
2. Grusho A., Grusho N., Zabezhailo M., Zatsarinny A., Timonina E. Information security of SDN on the basis of metadata // Computer network security / Eds. J. Rak, J. Bay, I. V. Kotenko, *et al.* — Lecture notes in computer science ser. — Springer, 2017. Vol. 10446. P. 339–347. [https://link.springer.com/chapter/10.1007/978-3-319-65127-9\\_27](https://link.springer.com/chapter/10.1007/978-3-319-65127-9_27).
3. Грушо А. А., Забежайло М. И., Смирнов Д. В., Тимонина Е. Е. Модель множества информационных пространств в задаче поиска инсайдера // Информатика и её применения, 2017. Т. 11. Вып. 4. С. 65–69.
4. Грушо А. А., Тимонина Е. Е., Шоргин С. Я. Иерархический метод порождения метаданных для управления сетевыми соединениями // Информатика и её применения, 2018. Т. 12. Вып. 2. С. 44–49.
5. Нильсон Н. Искусственный интеллект. Методы поиска решений / Пер. с англ. — М.: Мир, 1973. 272 с. (*Nilsson N. J. Problem-solving methods in artificial intelligence. New York, NY, USA: McGraw-Hill Pub. Co., 1971. 255 p.*)
6. Poelmans J., Elzinga P., Viaene S., Dedene G. Formal concept analysis in knowledge discovery: A survey // Conceptual structures: From information to intelligence / Eds. M. Croitoru, S. Ferré, D. Lukose. — Lecture notes in computer science ser. — Berlin–Heidelberg: Springer, 2010. Vol. 6208. P. 139–153.
7. Rahwan I., Simari G. R. Argumentation in artificial intelligence. — 1st ed. — Boston, MA, USA: Springer, 2009. 494 p.
8. Грушо Н. А. Поиск путей динамической реконфигурации распределенной информационно-вычислительной системы в случае захвата хоста противником // Системы и средства информатики, 2019. Т. 29. № 3. С. 194–201.
9. Grusho A., Timonina E., Shorgin S. Security models based on stochastic metadata // Analytical and computational methods in theory probability / Eds. V. Rykov, N. Singpurwalla, A. Zubkov. — Lecture notes in computer science ser. — Springer, 2017. Vol. 10684. P. 388–400. [https://link.springer.com/chapter/10.1007/978-3-319-71504-9\\_32](https://link.springer.com/chapter/10.1007/978-3-319-71504-9_32).

10. Grusho A. A., Timonina E. E., Shorgin S. Ya. Modelling for ensuring information security of the distributed information systems // 31th European Conference on Modelling and Simulation Proceedings. — Dudweiler, Germany: Digitaldruck Pirrot GmbH, 2017. P. 656–660. [http://www.scs-europe.net/dlib/2017/ecms2017acceptedpapers/0656-probstat\\_ECMS2017\\_0026.pdf](http://www.scs-europe.net/dlib/2017/ecms2017acceptedpapers/0656-probstat_ECMS2017_0026.pdf).

Поступила в редакцию 11.09.19

---

---

## SEARCH OF EMPIRICAL CAUSES OF FAILURES AND ERRORS IN COMPUTER SYSTEMS AND NETWORKS USING METADATA

A. A. Grusho, M. I. Zabezhailo, N. A. Grusho, 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 management of distributed information and computing systems involves two interrelated tasks: business continuity and security. Both tasks use similar methods to identify the causes of failures and security incidents. The aim of the study is adequate formalization, which in the future can be used to automate at least part of the functions of search for causes of failures and security incidents. Based on the relationship between the functions of the system administrator and methods of building metadata for managing connections in distributed information and computing systems, the method of hierarchical building of metadata and the new method of their application are proposed, which allows to search for causes of failures and control the solution of problems in such systems. The model and algorithm of search for an empirical cause of the failure, which are based on metadata and backup of tasks and data, have been built. It is shown that it is possible to build a tree of solutions to identify an empirical cause of an implicit failure with accuracy to an available destabilization of failed information technology. This highlights typical cause-effect analysis tasks from which you can design the entire search.

**Keywords:** information security; metadata; cause-and-effect relationships; system administration; implicit failures and errors

**DOI:** 10.14357/08696527190403

### Acknowledgments

The paper was partially supported by the Russian Foundation for Basic Research (project 18-07-00274-a).

### References

1. Finn, V. K. 2011. *Iskusstvennyy intellekt: metodologiya, primeneniya, filosofiya* [Artificial intelligence: Methodology, applications, and philosophy]. Moscow: Krasand. 448 p.

2. Grusho, A., N. Grusho, M. Zabezhailo, A. Zatsarinny, and E. Timonina. 2017. Information security of SDN on the basis of metadata. *Computer network security*. Eds. J. Rak, J. Bay, I. V. Kotenko, *et al.* Lecture notes in computer science ser. Springer. 10446:339–347. Available at: [https://link.springer.com/chapter/10.1007/978-3-319-65127-9\\_27](https://link.springer.com/chapter/10.1007/978-3-319-65127-9_27) (accessed August 16, 2019).
3. Grusho, A. A., M. I. Zabezhailo, D. V. Smirnov, and E. E. Timonina. 2017. Model' mnozhestva informatsionnykh prostranstv v zadache poiska insaydera [The model of the set of information spaces in the problem of insider detection]. *Informatika i ee Primeneniya — Inform. Appl.* 11(4):65–69.
4. Grusho, A. A., E. E. Timonina, and S. Ya. Shorgin. 2018. Ierarkhicheskiy metod porozhdeniya metadannykh dlya upravleniya setevymi soedineniyami [Hierarchical method of meta data generation for control of network connections]. *Informatika i ee Primeneniya — Inform. Appl.* 12(2):44–49.
5. Nilsson, N. J. 1971. *Problem-solving methods in artificial intelligence*. New York, NY: McGraw-Hill Pub. Co. 255 p.
6. Poelmans, J., P. Elzinga, S. Viaene, and G. Dedene. 2010. Formal concept analysis in knowledge discovery: A survey. *Conceptual structures: From information to intelligence*. Eds. M. Croitoru, S. Ferré, and D. Lukose. Lecture notes in computer science ser. Berlin–Heidelberg: Springer. 6208:139–153.
7. Rahwan, I., and G. R. Simari. 2009. *Argumentation in artificial intelligence*. 1st ed. Boston, MA: Springer. 494 p.
8. Grusho, N. A. 2019. Poisk putey dinamicheskoy rekonfiguratsii raspredelennoy informatsionno-vychislitel'noy sistemy v sluchae zakhvata khosta protivnikom [Methods of identification of “weak” signs of violations of information security]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(3):194–201.
9. Grusho, A., E. Timonina, and S. Shorgin. 2017. Security models based on stochastic metadata. *Analytical and computational methods in theory probability*. Eds. V. Rykov, N. Singpurwalla, and A. Zubkov. Lecture notes in computer science ser. Springer. 10684:388–400. Available at: [https://link.springer.com/chapter/10.1007/978-3-319-71504-9\\_32](https://link.springer.com/chapter/10.1007/978-3-319-71504-9_32) (accessed August 16, 2019).
10. Grusho, A. A., E. E. Timonina, and S. Ya. Shorgin. 2017. Modelling for ensuring information security of the distributed information systems. *31th European Conference on Modelling and Simulation Proceedings*. Dudweiler, Germany: Digitaldruck Pirrot GmbH. 656–660. Available at: [http://www.scs-europe.net/dlib/2017/ecms2017acceptedpapers/0656-probstat\\_ECMS2017\\_0026.pdf](http://www.scs-europe.net/dlib/2017/ecms2017acceptedpapers/0656-probstat_ECMS2017_0026.pdf) (accessed August 16, 2019).

*Received September 11, 2019*

## Contributors

**Grusho Alexander A.** (b. 1946) — Doctor of Science in physics and mathematics, professor, principal 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; [grusho@yandex.ru](mailto:grusho@yandex.ru)

**Zabezhailo Michael I.** (b. 1956) — Doctor of Science in physics and mathematics, principal 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; m.zabezhailo@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 Science and Control” of the Russian Academy of Sciences; 44-2 Vavilov Str., Moscow 119133, Russian Federation; info@itake.ru

**Timonina Elena E.** (b. 1952) — 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 119133, Russian Federation; eltimon@yandex.ru

## ВИЗУАЛЬНЫЙ МЕТАЯЗЫК ОПИСАНИЯ РЕШЕНИЯ ДИАГНОСТИЧЕСКИХ ПРОБЛЕМ\*

*С. Б. Румовская*<sup>1</sup>, *А. В. Колесников*<sup>2</sup>, *А. А. Литвин*<sup>3</sup>

**Аннотация:** Рассматривается идея визуального метаязыка, описывающего решение диагностической проблемы (ДП) комбинацией нескольких взаимосвязанных процессов рассуждений на разных языках, а также подход к редукции ДП с когнитивной визуализацией. Диагностическая проблема в результате представляется как декомпозиция иерархически выстроенных с горизонтальными связями элементов разного типа: предметное диагностическое решение, диагностические задачи и диагностические подзадачи, диагностические решения-связки и скомпенсированные затруднения. На визуальном метаязыке основывается аксиоматическая теория ролевых визуальных диагностических моделей функциональных гибридных интеллектуальных диагностических систем (ФГИДС) как систем поддержки принятия диагностических решений с когнитивной визуализацией проблем и получением интегрированного образа «состояние органов и систем человека».

**Ключевые слова:** диагностическая проблема; когнитивная визуализация; редукция диагностической проблемы; визуальный метаязык

**DOI:** 10.14357/08696527190404

### 1 Введение

От оперативности реакции врача на возникающую проблемную диагностическую ситуацию зависит безопасность принимаемых решений. В среднем около половины граждан РФ сталкивались с ошибками врачей, и примерно в 35% случаев имело место расхождение поликлинического и клинического диагнозов, а в 21,6% прижизненный диагноз оказывался неверен.

Эксперты по психологии мышления [1] считают, что два способа представления информации — символный и картинно-образный, а также работа с ними и соотнесение их относительно друг друга обеспечивают феномен человеческого мышления. Образная система характеризуется:

- (1) высокой скоростью обработки образов, которая особенно поразительна, если учесть низкие скорости передачи сигналов в мозге;
- (2) целостностью образных представлений («гештальтов»).

---

\*Работа выполнена при поддержке РФФИ (проект 19-07-00250А).

<sup>1</sup>Калининградский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, [sophiyabr@gmail.com](mailto:sophiyabr@gmail.com)

<sup>2</sup>Балтийский федеральный университет им. И. Канта, [avkolesnikov@yandex.ru](mailto:avkolesnikov@yandex.ru)

<sup>3</sup>Областная клиническая больница Калининградской области, [aalitvin@gmail.com](mailto:aalitvin@gmail.com)

Оба эти свойства не имеют аналогов в рамках компьютерной парадигмы. Визуализация позволяет перевести во внешне замечаемый план те содержания: образы, представления, и, соответственно, процессы с ними, — которые происходят у специалиста (врача) «в голове». Сделать их видимыми, фиксируемыми и легко управляемыми как для него самого, так и для тех, с кем он взаимодействует по этому поводу.

Исследование когнитивных структур и механизмов человека, специалиста в медицине, которое позволит сделать процессы преодоления ДП видимыми, контрастными, а разработка на их основе ФГИДС с когнитивным моделированием ДП позволит подойти к сглаживанию противоречивых показаний аппаратуры, снизит число врачебных ошибок, повысит качество и, соответственно, безопасность медицинских интеллектуальных систем поддержки принятия диагностических решений.

В работе рассмотрено понятие диагностической проблемы, психологические механизмы решения и ее редукция, идея визуального метаязыка, который позволит описать решение ДП комбинацией нескольких взаимоувязанных процессов рассуждений на разных языках.

## 2 Диагностическая проблема и механизмы ее решения

Диагностическая проблема — это сложная задача распознавания болезни и оценки индивидуальных биологических особенностей и социального статуса объекта диагностики (пациента) путем истолкования результатов целенаправленного медицинского обследования и их обобщения в виде установленного диагноза [2]. Это аморфное образование, без предзадания, формулируемое динамически, по ходу решения.

Одновременно д. п. н., профессор В. Ф. Спиридонов в рамках своих исследований по психологии мышления при решении задач и проблем отметил два механизма решения проблем [3].

*Первый механизм* — это самоопределение решателя (термин ввел Г. П. Щедровицкий — российский философ и методолог, общественный и культурный деятель, создатель системомыследеятельностной методологии). Базируется этот механизм на том, что человек (решатель) — часть проблемы, с которой он столкнулся. В рамках этого уточняются: содержание проблемы; нахождение решателем своего места в проблеме; проясняется и более четко формулируется проблема и, соответственно, итоговая цель; решателем определяются зоны ответственности за достижение решения или его реализацию (в случае практической проблемы), а также описываются критерии успешного результата.

*Второй механизм*, на котором и будет базироваться данное исследование, — изменение представления проблемы в ходе решения. Проблема здесь — образ тесно связанной и иерархически выстроенной системы задач и обслуживающих их подзадач, предполагающих наличие знания о способе их решении. При этом эле-



менты проблемы, согласно исследованию Куколевой [4], опытные специалисты при ее решении дифференцируют на несколько типов составляющих:

- (1) предметное диагностическое решение — конкретные действия, результат которых — это достижение поставленной цели — диагноз (план обследования пациента);
- (2) «диагностические задачи» — более крупные — и «диагностические подзадачи» — более мелкие, связанные с решением крупной (например, перед тем как диагностировать ишемическую болезнь сердца, необходимо выделить информативный набор показателей состояния здоровья пациента относительно данного заболевания);
- (3) скомпенсированные затруднения — часто встречающиеся небольшие задачи, для которых есть типовое решение — диагностические задачи, для решения которых есть четкий и однозначный протокол обследования и постановки диагноза (например, выявление дислипидемии);
- (4) диагностические решения-связки — действия, учитывающие конкретные условия по типу «если... то...».

В результате исследования предполагается на основе данного механизма разработать гибридную интеллектуальную систему с когнитивным моделированием ДП. При этом в рамках проблемно-структурной методологии гибридных интеллектуальных систем [5] диагностические задачи и скомпенсированные затруднения по сути представляют собой функциональные подзадачи, составляющие функциональную структуру ДП, а диагностические подзадачи — это технологические задачи, решаемые для организации эффективной обработки данных и знаний в ходе диагностического процесса. Диагностические же решения-связки определяют часть горизонтальных связей подзадач, а предметное решение представляет собой описание концептуальной модели диагностического процесса.

Рассмотрим далее идею языка, на котором будет базироваться аксиоматическая теория ролевых визуальных диагностических моделей ФГИДС как систем поддержки принятия диагностических решений с когнитивной визуализацией проблем и получением интегрированного образа «состояние органов и систем человека».

### 3 Аксиоматическая теория ролевых визуальных диагностических моделей. Визуальный метаязык

Аксиоматическая теория ролевых визуальных диагностических моделей основывается на понимании визуального языка как семиотической системы:

$$vl_d = \langle KKVY_d, VS, VSe, VP, V\delta \rangle, \quad (1)$$

где  $KKVY_d$  — концептуально-визуальное категориальное ядро;  $VS$ ,  $VSe$  и  $VP$  — множества синтаксических, семантических (знания о предметной области ме-

дицинской диагностики) и прагматических правил (множество правил вывода диагностических решений) соответственно;  $V\delta$  — правила изменения множеств  $KKVY_d$ ,  $VS$ ,  $VSe$  и  $VP$ .

Множества  $KKVY_d$ ,  $VS$ ,  $VSe$  и  $VP$  из (1) определяются выражениями:

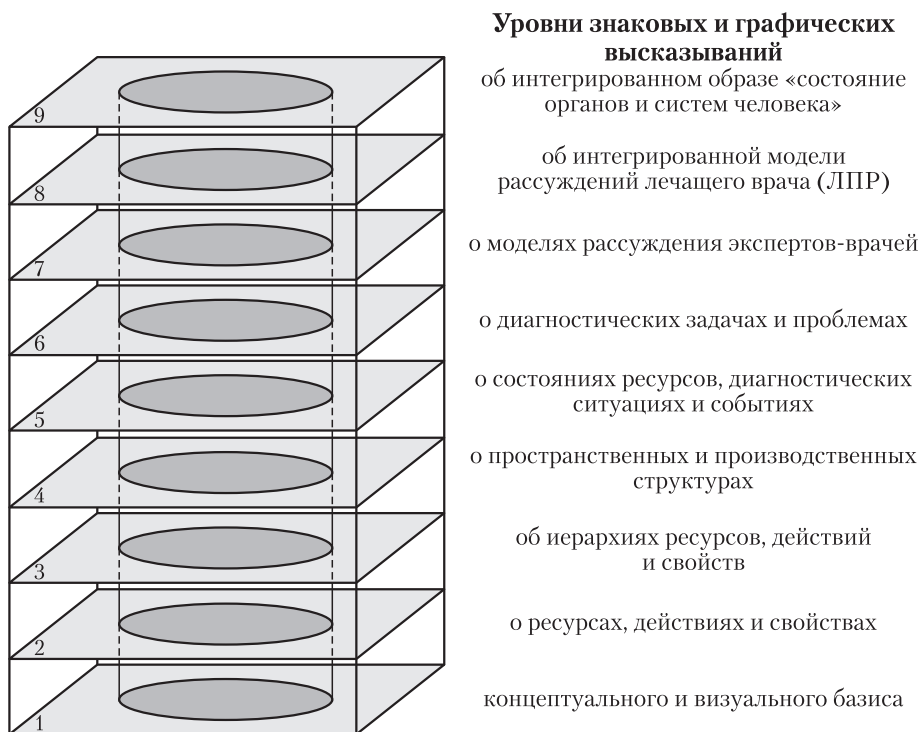
$$\begin{aligned}
 KKVY_d &= \langle KKY_d, KVY_d, G \rangle, \\
 KKY_d &= \langle RES_d, PR_d, ACT_d, R \rangle, \quad KVY_d = \langle P^{RES_d}, P^{PR_d}, P^{ACT_d}, VR \rangle, \\
 G &= \langle G^{RES}, G^{PR}, G^{ACT}, G^R \rangle, \\
 G^{RES} : RES_d &\leftrightarrow P^{RES_d}, \quad G^{PR} : PR_d \leftrightarrow P^{PR_d}, \\
 G^{ACT} : ACT_d &\leftrightarrow P^{ACT_d}, \quad G^R : R_d \leftrightarrow P^{R_d}; \\
 VS &= \langle KVY_d, VN, PRU \rangle; \\
 VSe &= \langle KKY_d, G \rangle; \\
 VP &= \{ \langle Ex_m, act_d, M, Cond \rangle \},
 \end{aligned}$$

где  $KKY_d$  — концептуальное категориальное ядро;  $KVY_d$  — визуальное категориальное ядро;  $RES_d$ ,  $PR_d$  и  $ACT_d$  — множества ресурсов, свойств и действий, раскрывающие структуру диагностики в медицине;  $R$  — множество отношений, заданных на  $RES_d$ ,  $PR_d$  и  $ACT_d$  — «ресурс–ресурс», «свойство–свойство», «действие–действие», «ресурс–свойство», «свойство–ресурс», «ресурс–действие», «действие–ресурс», «действие–свойство», «свойство–действие»;  $P^{RES_d}$ ,  $P^{PR_d}$  и  $P^{ACT_d}$  — множества визуальных символов ресурсов, свойств и действий;  $VR$  — множество ролевых визуальных отношений на одном и более  $P^{RES_d}$ ,  $P^{PR_d}$ ,  $P^{ACT_d}$ , образные (визуально-)символьные графические высказывания;  $VN$  — словарь нетерминальных символов;  $PRU$  — множество продукционных правил;  $Ex_m$  — множество носителей языка (экспертов-врачей), действия которых ограничиваются внешней средой;  $act_d$  — действие, определенное на множестве действий  $ACT$  и являющееся объектом нормативной регуляции;  $M$  — множество систем модальностей;  $Cond$  — множество ситуаций, в которых (не) должно применяться действие.

Метаязык визуального управления предложен в [6] в виде «слоеного пирога», и, следуя данной идее, метаязык аксиоматической теории ролевых визуальных диагностических моделей представлен на рис. 1.

На рис. 1 образно-визуальное ядро на каждом уровне представлено темно-серым овалом, а концептуально-языковая оболочка — светло-серым прямоугольником. В основании лежит концептуально-визуальный базис — словари понятий и отношений. Визуальное ядро и концептуальная оболочка более высокого уровня включают знаки и концепты ядер и оболочек предыдущих уровней.

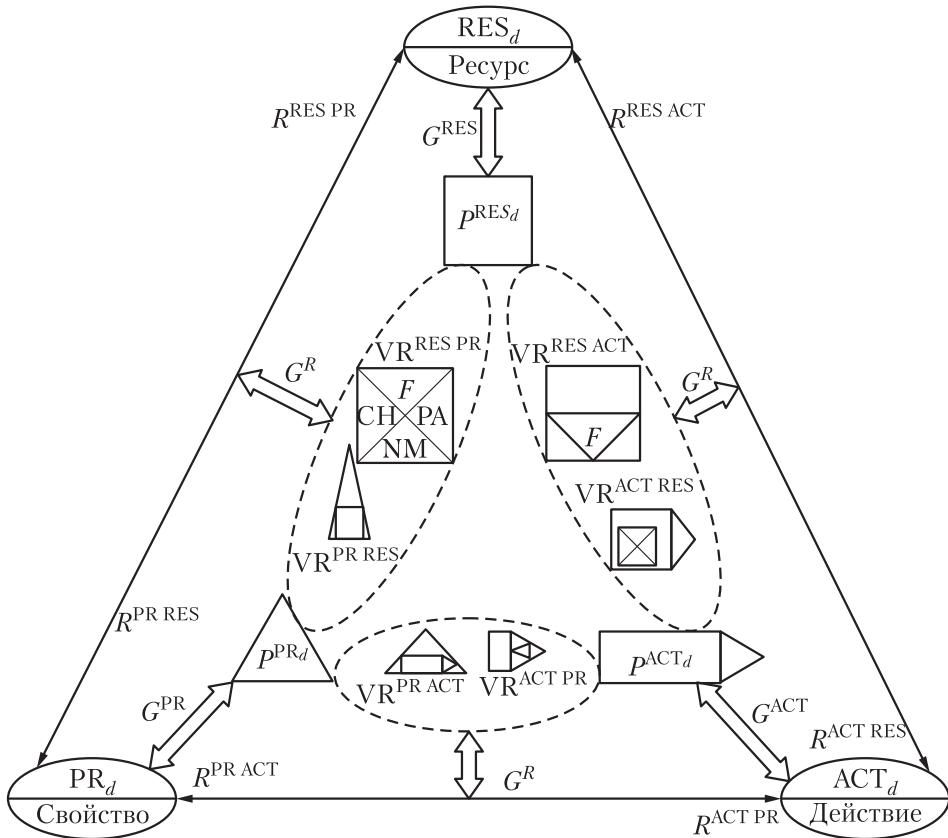
Концептуально-визуальный категориальный базис метаязыка визуальной диагностики «ресурс–свойство–действие–отношение» представлен на рис. 2.



**Рис. 1** Метаязык визуальной диагностики

Ресурс на рис. 2 — форма или понятие, обозначающее или символизирующее вещь (одушевленную и неодушевленную), имеющуюся у субъекта диагностики для решения ДП. Это все то, что есть у врача для принятия решений. Для отображения состояния органов и систем человека выделим среди ресурсов системные функциональные единицы (СФЕ) [7] разной степени детализации, под которыми понимаются различные клеточные и тканевые структуры, например нефрон почек, ацинус поджелудочной железы. Системная функциональная единица реализует лишь одну функцию по принципу максимального или нулевого действия. Тогда, например, в качестве СФЕ поджелудочной железы (ПЖ) можно определить дольки или ацинусы и островки Лангерганса. Есть человеческие и технические ресурсы.

Свойство на рис. 2 — все то, что не является границами данного ресурса. Это то, что, характеризуя ресурсы, не образует новых ресурсов. Выделяют физические свойства (параметры), характеристические свойства (характеристики), именные свойства (имена) и функции (предназначение ресурса). Это, например, признаки, описывающие состояние здоровья пациента, совокупности которых



**Рис. 2** Концептуально-визуальное ядро «ресурс–свойство–действие–отношение»:  $P^{RES_d}$ ,  $P^{PR_d}$  и  $P^{ACT_d}$  — множества визуальных символов ресурсов, свойств и действий;  $P^{RES PR}$ ,  $R^{PR RES}$ ,  $R^{RES ACT}$ ,  $R^{ACT RES}$ ,  $R^{PR ACT}$ ,  $R^{ACT PR}$ ,  $VR^{RES PR}$ ,  $VR^{PR RES}$ ,  $VR^{RES ACT}$ ,  $VR^{ACT RES}$ ,  $VR^{PR ACT}$  и  $VR^{ACT PR}$  — множества отношений «ресурс–свойство», «свойство–ресурс», «ресурс–действие», «действие–ресурс», «свойство–действие», «действие–свойство» и их ролевые визуальные модели соответственно;  $G^{RES}$ ,  $G^{PR}$ ,  $G^{ACT}$  и  $G^R$  — биекция множеств ресурсов, свойств, действий и отношений на множества их визуальных символов соответственно;  $F$ ,  $CH$ ,  $PA$  и  $NM$  — множества функций, характеристик, параметров и имен

могут соответствовать тому или иному заболеванию, параметры аппаратов компьютерной томографии (КТ), ультразвукового обследования и др.

Параметр — свойство, значение которого может быть измерено по некоторой шкале. Например, *скорость* клубочковой фильтрации, *температура*, *концентрация* глюкозагона в сыворотке крови.

Характеристика — свойство, выражающее строение, устройство, особенности. Например, *многослойная КТ, спиральная КТ, альвеолярная железа, экзокринная часть поджелудочной железы, нейроэндокринная опухоль*.

Функция — свойство, выражающее предназначение ресурса. Например, *орган пищеварения, перенос кислорода, секреция гормонов*.

Имя — свойство, идентифицирующее ресурсы, действия, задачи, методы и выделяющее их среди других с одинаковыми параметрами, функциями и характеристиками. Например, *класс КТ-аппарата, модель протеза руки*.

Действие — форма или понятие, обозначающее (или символизирующее) отношения на ресурсах как следствие деятельности (исполняемых функций организма как системы) и внешних воздействий (например, радикальная, симптоматическая, восстановительная терапия). Реакции, в том числе патологические, недостаточные или чрезмерные, в результате которых возникают дефекты (опухоли, гипотрофии, инфильтраты, нагноения и т. д.) — это действия.

Ядро на рис. 2 дает достаточно простое и прозрачное структурирование мира визуальной диагностики, утверждая, что это есть мир ресурсов, действий их свойств и девяти классов отношений (петли «ресурс–ресурс», «действие–действие», «свойство–свойство» на рис. 2 опущены).

Расширим ядро  $KKVY_d$ . На рис. 3 представлены дополняющие схему на рис. 2 конструкты расширенного ядра. Для моделирования предметных знаний об объекте диагностики (пациенте) необходимо расширить концептуально-визуальное ядро  $KKVY_d$  концептами (и соответствующими им визуальными символами) единицами измерения («мерами», MES), «значениями» (VAL) и «состояниями» (ST). Для моделирования предметных знаний о субъекте диагностики (лечащий врач и узкоспециализированные врачи) — «оценки» (EST).

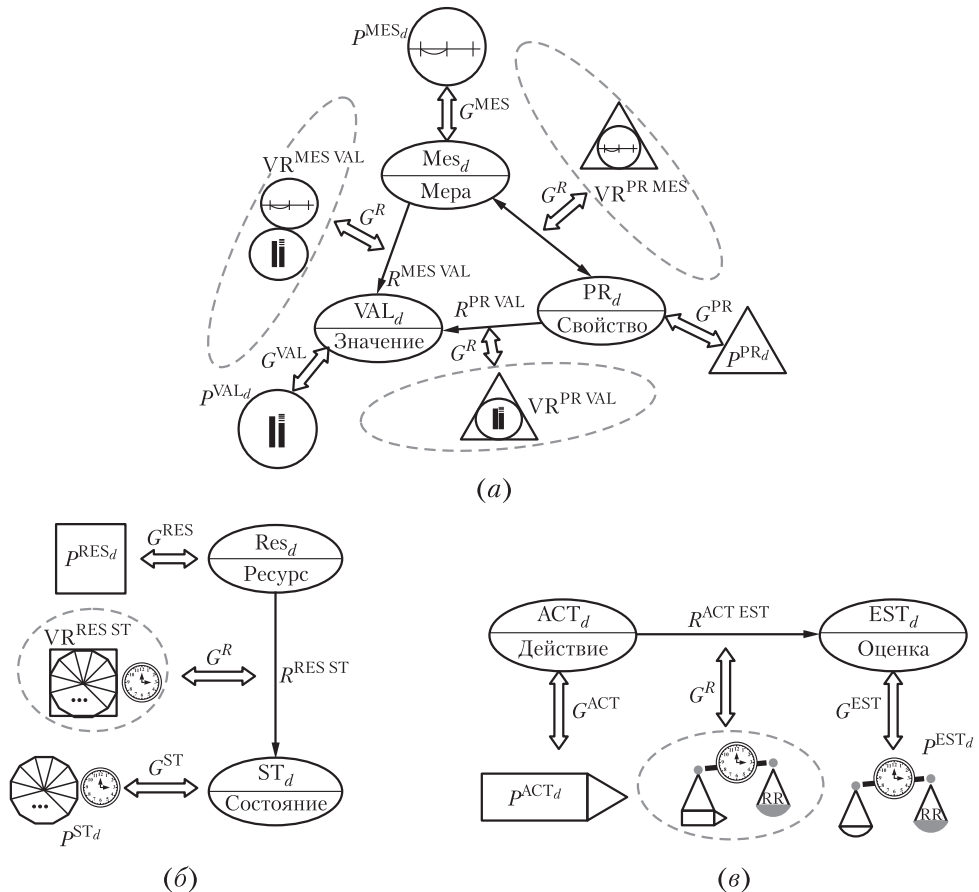
Рассмотрим подробнее концепты расширения концептуально-визуального ядра «мера–значение–состояние–оценка» (см. рис. 3).

Мера — понятие, обозначающее принятую единицу параметра свойства и используемую в качестве меры последнего. Например, *мкмоль/л, частота сердечных сокращений в минуту, единиц/л* и др.

Значение — понятие или число, показывающее количество единиц меры. Например, *активность амилазы в моче 700 единиц/л, уровень сывороточного билирубина 340 мкмоль/л, повышенная концентрация мочевины в сыворотке*.

Состояние — понятие, обозначающее внешнее проявление процессов, происходящих в ресурсе в некоторый момент времени. Например, *нарушена структура слизистой оболочки тонкой кишки, ключевые ферменты гликолиза активны*.

Состояние объекта диагностики  $ST_{od}(t)$  в момент времени  $t$  — интегрированный образ совокупности существенных для решаемой в момент  $t$  диагностической задачи параметров, качественных нечетких характеристик ресурсов и исполняемых функций, а также *d-ситуаций* одновременного исполнения некоторого множества действий над объектом диагностики — диагностических операций и их локализации — и *r-ситуаций* — одновременного выполнения действий ресурсами, т. е. реакции одних систем организма, его подсистем и их элементов на



**Рис. 3** Расширение концептуально-визуального ядра «мера–значение–состояние–оценка»:  $P^{RES_d}$ ,  $P^{PR_d}$  и  $P^{ACT_d}$  — множества визуальных символов ресурсов, свойств и действий;  $R^{MES PR}$ ,  $R^{PR MES}$ ,  $R^{PR VAL}$ ,  $R^{MES VAL}$ ,  $R^{RES ST}$ ,  $R^{ACT EST}$ ,  $VR^{MES PR}$ ,  $VR^{PR MES}$ ,  $VR^{PR VAL}$ ,  $VR^{MES VAL}$ ,  $VR^{RES ST}$  и  $VR^{ACT EST}$  — множество отношений «мера–свойство», «свойство–мера», «свойство–значение», «мера–значение», «ресурс–состояние», «действие–оценка» и их ролевые визуальные модели соответственно;  $G^{RES}$ ,  $G^{PR}$ ,  $G^{ACT}$ ,  $G^R$ ,  $G^{VAL}$ ,  $G^{MES}$ ,  $G^{ST}$  и  $G^{EST}$  — биекции множеств ресурсов, свойств, действий, отношений, значений, мер, свойств и оценок на множества их визуальных символов соответственно; RR (reference range) — референсные значения

деятельность других систем организма, его подсистем и их элементов, а также реакции на внешние воздействия на организм.

Оценка — понятие, отображающее результат выполнения действий на оценочные шкалы. Установление меры близости текущего состояния ресурсов к же-

лаемому. Например, получение *информативных* результатов КТ, *сдвиг влево* лейкоцитарной формулы, *сниженная и неоднородная* эхогенность паренхимы.

На следующем этапе исследования будет необходимо:

- (1) ввести концепт «задача» (PRB) и концепты для моделирования предметных знаний о субъекте моделирования — «метод» (MET), «модель» (MOD);
- (2) адаптировать метод смешанной редукции относительно диагностической проблемы с когнитивной визуализацией ее элементов и проблемы в целом, а также определить множество отношений и способов их репрезентации в иерархической структуре декомпозиции.

#### 4 Заключение

Создание аксиоматической теории ролевых визуальных отношений проблемной диагностической ситуации и ее элементов на категориальном ядре «ресурс–свойство–действие–отношение», обусловливающих релевантное отображение вербально-символьного и визуально-пространственного аспектов интеллектуальной деятельности в архитектуру и алгоритмы искусственного интеллекта, позволит получить новые теоретические, методологические и технологические знания по инфокоммуникационным технологиям визуального мышления в процессах медицинской диагностики.

#### Литература

1. Майоров А. А. Пространственное когнитивное моделирование // Перспективы науки и образования, 2014. № 1. С. 33–37.
2. Малая медицинская энциклопедия: в 6 т. Т. 2: Грудь — Куммеля болезнь / Под ред. В. И. Покровского. — М.: Медицина, 1991. С. 89.
3. Спиридонов В. Ф. Психология мышления: решение задач и проблем. — М.: Генезис, 2006. 319 с.
4. Куколева О. М. Представление проблемной ситуации как маркер компетентности решателя // Исследования обучения и развития в контексте культурно-исторического подхода: Мат-лы вторых чтений, посвященных памяти Л. С. Выготского, 2002. С. 231–242.
5. Румовская С. Б. Исследование методов поддержки принятия коллективных диагностических решений и разработка инструментальных средств «Виртуальный консилиум» (на примере диагностики артериальной гипертензии): Дис. ... канд. техн. наук. — М., 2017. 138 с.
6. Колесников А. В., Листопад С. В. Концептуально-визуальные основы виртуальных гетерогенных коллективов, поддерживающих принятие решений // Гибридные и синергетические интеллектуальные системы: Мат-лы III Всеросс. Пospelовской конференции с международным участием. — Калининград: БФУ им. И. Канта, 2016. С. 8–56.
7. Гайдес М. А. Общая теория систем (системы и системный анализ). — 2-е изд. — М.: Глобус-Пресс, 2005. 201 с.

Поступила в редакцию 21.08.19

## VISUAL METALANGUAGE FOR DESCRIPTION OF SOLUTION OF DIAGNOSTIC PROBLEMS

S. B. Rumovskaya<sup>1</sup>, A. V. Kolesnikov<sup>2</sup>, and A. A. Litvin<sup>3</sup>

<sup>1</sup>Kaliningrad Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 5 Gostinaya Str., Kaliningrad 236022, Russian Federation

<sup>2</sup>Immanuel Kant Baltic Federal University, 14 A. Nevskogo Str., Kaliningrad 236041, Russian Federation

<sup>3</sup>Regional Clinical Hospital of the Kaliningrad Region, 74 Klinicheskaya Str., Kaliningrad 236016, Russian Federation

**Abstract:** The paper considers the idea of the visual metalanguage describing the solution of a diagnostic problem by a combination of several interconnected reasoning processes in different languages. The approach to reducing a diagnostic problem with cognitive visualization is also observed. The diagnostic problem is presented as a decomposition of elements of different types hierarchically arranged with horizontal connections. It includes domain diagnostic solution, diagnostic tasks and subtasks, diagnostic solutions-bundles and compensated difficulties. The axiomatic theory of role visual diagnostic models of functional hybrid intelligent diagnostic systems is based on the considered visual metalanguage. Such systems are the support systems for making diagnostic decisions with cognitive visualization of problems and obtaining an integrated image — the “state of human organs and systems.”

**Keywords:** diagnostic problem; cognitive visualization; reduction of a diagnostic problem; visual metalanguage

**DOI:** 10.14357/08696527190404

### Acknowledgments

The study was funded by the Russian Foundation for Basic Research according to the research project No. 19-07-00250A.

### References

1. Mayorov, A. A. Prostranstvennoe kognitivnoe modelirovanie [Extensional cognitive modeling]. *Perspektivy nauki i obrazovaniya* [Perspectives of science and education] 1:33–37.
2. Pokrovskii, V. I., ed. 1991. *Malaya meditsinskaya entsiklopediya. T. 2: Grud' — Kummelya bolezni'* [Small medical encyclopaedia. Vol. 2: Chest — the disease of Kummel]. Moscow: Medicine. P. 89.
3. Spiridonov, V. F. 2006. *Psikhologiya myshleniya: reshenie zadach i problem* [Psychology of thinking: Problem and task solving]. Moscow: Genezis. 319 p.



4. Kukoleva, O. M. 2002. Predstavlenie problemnoy situatsii kak marker kompetentnosti reshatel'ya [Representation of problem situation as highlighter of competence of the solver]. *Issledovaniya obucheniya i razvitiya v kontekste kul'turno-istoricheskogo podkhoda: Mat-ly vtorykh chteniy, posvyashchennykh pamyati L. S. Vygotskogo* [Investigation of Education and Development Within the Context of Culture-Historical Approach: 2nd Readings Dedicated to the Memory of L. S. Vygotskogii Proceedings]. 231–242.
5. Rumovskaya, S. B. 2017. Issledovanie metodov podderzhki prinyatiya kolektivnykh diagnosticheskikh resheniy i razrabotka instrumental'nykh sredstv «Virtual'nyy konsilium» (na primere diagnostiki arterial'noy gipertenzii) [The research of methods of collective diagnostic decision support and development of instruments of the “Virtual council” (illustrated with the diagnostic of arterial hypertension)]. Moscow. PhD Thesis. 138 p.
6. Kolesnikov, A. V., and S. V. Listopad. 2016. Kontseptual'no-vizual'nye osnovy virtual'nykh geterogennykh kolektivov, podderzhivayushchikh prinyatie resheniy [Conceptual and visual foundations of virtual heterogeneous teams supporting decision-making]. *Gibridnye i sinergeticheskie intellektual'nye sistemy: mat-ly III Vseross. Pospelovskoy konferentsii s mezhdunarodnym uchastiem* [Hybrid and Synergetic Intelligent Systems: 3rd All-Russian Pospelovsky Conference with International Participation Proceedings]. Kaliningrad: Publishing House of Immanuel Kant Baltic Federal University. 8–56.
7. Gaydes, M. A. 2005. *Obshchaya teoriya sistem (sistemy i sistemnyy analiz)* [General theory of systems (systems and systems analysis)]. Moscow: Globus-Press. 201 p.

Received August 21, 2019

## Contributors

**Rumovskaya Sophiya B.** (b. 1985) — Candidate of Science (PhD) in technology, scientist, Kaliningrad Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 5 Gostinaya Str., Kaliningrad 236022, Russian Federation; sophiyabr@gmail.com

**Kolesnikov Alexander V.** (b. 1948) — Doctor of Science in technology, professor, Institute of Physical and Mathematical Sciences and Information Technology, Immanuel Kant Baltic Federal University, 14 A. Nevskogo Str., Kaliningrad 236041, Russian Federation; avkolesnikov@yandex.ru

**Litvin Andrei A.** (b. 1966) — Doctor of Science in medicine, Deputy Head Physician, Regional Clinical Hospital of the Kaliningrad Region, 74 Klinicheskaya Str., Kaliningrad 236016, Russian Federation; aalitvin@gmail.com

## СТАЦИОНАРНЫЕ ХАРАКТЕРИСТИКИ ОБСЛУЖИВАНИЯ В СИСТЕМЕ GI/MSP/ $n/\infty$ С ОБОБЩЕННЫМ ОБНОВЛЕНИЕМ\*

*И. С. Зарядов<sup>1</sup>, Л. А. Мейханаджян<sup>2</sup>, Т. А. Милованова<sup>3</sup>*

**Аннотация:** Рассматривается система обслуживания GI/MSP/ $n/\infty$  с рекуррентным входящим потоком,  $n$  идентичными приборами, обслуживанием марковского типа, очередью неограниченной емкости и обобщенным обновлением. Обобщенное обновление, являющееся разновидностью механизма активного управления очередью, предполагает, что в момент окончания обслуживания покидающая систему заявка может удалить из очереди некоторое случайное число ожидающих заявок с заданным вероятностным распределением. С помощью метода вложенной цепи Маркова найдены стационарные распределения основных показателей функционирования системы. Полученные соотношения дают возможность написания программ расчета, позволяющих вычислить как стационарные вероятности числа заявок в системе по моментам поступления заявок и по времени, так и стационарное распределение времени ожидания начала обслуживания (при прямом порядке обслуживания и обновления).

**Ключевые слова:** система массового обслуживания; обобщенное обновление; марковский процесс обслуживания; управление очередью; вложенная цепь Маркова

**DOI:** 10.14357/08696527190405

### 1 Введение

В работе рассматривается  $n$ -линейная система массового обслуживания (СМО) с одной очередью неограниченной емкости, рекуррентным входящим потоком заявок, специальным обслуживанием марковского типа и механизмом обобщенного обновления согласно [1]. С помощью метода вложенной цепи Маркова найдены стационарные распределения основных характеристик обслуживания. Приведенные здесь результаты обобщают аналогичные результаты статей [1–4], полученные для однолинейной системы  $G/MSP/1/r$  и многолинейных систем  $G/M/n/r$  с обновлением. Кроме того, предложенная модель

---

\*Исследование выполнено при финансовой поддержке РФФИ (проект 19-07-00739).

<sup>1</sup>Российский университет дружбы народов; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, zaryadov-is@rudn.ru

<sup>2</sup>Финансовый университет при Правительстве РФ, lamejkanadzhyan@fa.ru

<sup>3</sup>Российский университет дружбы народов, milovanova-ta@rudn.ru

процесса обслуживания может рассматриваться как обобщение (в случае идентичных приборов) марковского обслуживания (Markovian Service Process), обычно используемого при описании работы приборов в СМО, функционирующих в случайной среде [5–8].

Отметим также, что интерес к моделям, подобным той, что рассматривается в статье, связан с перспективами применения механизма обобщенного обновления в качестве алгоритма активного управления очередями (см. подробнее в [9–13]).

## 2 Описание системы

Рассматривается  $n$ -линейная СМО  $G/MSP/n/\infty$  с очередью бесконечной емкости и обобщенным обновлением. Входящий в систему поток заявок является рекуррентным, причем время между соседними поступлениями заявок имеет произвольную функцию распределения  $A(x)$ . Будем предполагать, что среднее время  $a = \int_0^\infty x dA(x)$  между поступлениями заявок конечно и, кроме того, там, где речь пойдет о стационарном распределении по времени, будем считать, что распределение  $A(x)$  является нерешетчатым.

Марковский процесс обслуживания заявок определяется следующим образом. Если в системе находится  $k$  заявок, то процесс обслуживания может находиться в одном из  $l_k$ ,  $1 \leq l_k < \infty$ , состояний (фаз обслуживания). Тогда если фаза обслуживания равна  $i$ ,  $1 \leq i \leq l_k$ , то за «малое» время  $\Delta$  с вероятностью  $\lambda_{ij}^{(k)} \Delta + o(\Delta)$  фаза изменится на  $j$ -ю,  $1 \leq j \leq l_k$ , и все заявки будут продолжать обслуживаться, а с вероятностью  $n_{ij}^{(k)} \Delta + o(\Delta)$  фаза изменится на  $j$ -ю,  $1 \leq j \leq \leq l_{k-1}$ , но обслуживание одной из заявок закончится и она покинет систему. Матрицы из элементов  $\lambda_{ij}^{(k)}$  и  $n_{ij}^{(k)}$  будем обозначать через  $\Lambda_k$  и  $N_k$ ,  $k \geq 1$ . Кроме того, будем предполагать, что  $l_k = l$  при  $k \geq n$ ; тогда матрицы  $\Lambda_k = \Lambda$  совпадают при  $k \geq n$ , а матрицы  $N_k = N$  совпадают при  $k \geq n + 1$ . Наконец, будем считать, что на свободном периоде фаза обслуживания не изменяется. Если в момент поступления очередной заявки в системе имеется  $k$ ,  $0 \leq k \leq n - 1$ , других заявок и фаза обслуживания равна  $i$ ,  $1 \leq i \leq l_k$ , то после поступления новой заявки она с вероятностью  $\omega_{ij}^{(k)}$  изменится на  $j$ ,  $1 \leq j \leq l_{k+1}$ . Матрицу из элементов  $\omega_{ij}^{(k)}$  будем обозначать через  $\Omega_k$ .

Обобщенное обновление определяется следующим образом (см. также [1]). Находящаяся на приборе заявка в момент окончания обслуживания одновременно с уходом из системы либо с вероятностью  $q(l)$  удаляет («убивает») из очереди ровно  $l$  заявок, если в ней находится более  $l$  заявок, либо с вероятностью  $Q(l) = \sum_{k=l}^\infty q(k)$  полностью опустошает очередь, если в ней было ровно  $l$  заявок. Вероятности  $q(l)$  называются вероятностями обновления. Очевидно, что  $Q(0) = \sum_{l=0}^\infty q(l) = 1$ , а вероятность того, что закончившая обслуживание на приборе заявка покинет систему, не оказывая на нее никакого воздействия, — это  $q(0)$ .

Предполагается, что матрица  $\Lambda + N$  является неразложимой, а матрица  $N$  — ненулевой. Более того, будем предполагать, что при исходных параметрах рассматриваемой СМО введенная далее вложенная цепь Маркова будет неприводимой и существует стационарный режим, необходимым и достаточным условием которого является выполнение неравенства<sup>1</sup>:

$$a^{-1} < \pi_s N \bar{1} \sum_{i=0}^{\infty} (i+1)q(i),$$

где в правой части стоит среднее число покидающих систему за единицу времени заявок при условии, что в системе находится бесконечное число заявок (а  $\bar{\pi}_s$  — вектор стационарных вероятностей марковского процесса с инфинитезимальной матрицей  $\Lambda + N$ ). Наконец, будем считать, что заявки обслуживаются и «убиваются» в порядке поступления.

### 3 Стационарные распределения числа заявок в системе

Рассмотрим последовательные моменты  $\tau_n, n \geq 0$ , поступления заявок в систему. Пусть  $\xi(t)$  — фаза обслуживания заявок в момент времени  $t$ , а  $\nu(t)$  — число заявок в системе в этот момент. Определим случайные величины  $\xi_n = \xi(\tau_n + 0)$  и  $\nu_n = \nu(\tau_n + 0)$ , которые задают соответственно фазу обслуживания и число заявок в системе непосредственно после момента поступления  $n$ -й заявки. Положим  $\zeta_n = (\xi_n, \nu_n)$ . Последовательность  $\{\zeta_n, n \geq 0\}$  образует однородную цепь Маркова, которую будем называть вложенной. Множество состояний  $\mathcal{X}$  вложенной цепи Маркова имеет вид:

$$\mathcal{X} = \{(1, 1), \dots, (l_1, 1)\} \cup \{(1, 2), \dots, (l_2, 2)\} \cup \dots,$$

где первый индекс указывает на фазу обслуживания, а второй — на число заявок в системе непосредственно после момента поступления заявки.

Выпишем матрицу переходных вероятностей вложенной цепи Маркова  $\{\zeta_n, n \geq 0\}$ . Для этого сначала определим следующие матрицы:

- $F_k(x), k \geq 0$ , — матрица, элемент  $(F_k(x))_{ij}, 1 \leq i, j \leq l$ , которой представляет собой условную вероятность того, что за время  $x$  систему покинет ровно  $k$  заявок и процесс обслуживания перейдет на фазу  $j$  при условии, что в начальный момент в системе (вместе с заявками на приборах) было не менее  $k + n$  заявок, процесс обслуживания находился на фазе  $i$  и за время  $x$  в систему не поступила новая заявка;
- $A_k, k \geq 0$ , — матрица, элемент  $(A_k)_{ij}$  которой представляет собой условную вероятность того, что за время между последовательными моментами поступления заявок систему покинут  $k$  заявок и процесс обслуживания перейдет на

<sup>1</sup>Здесь и далее  $\bar{1}$  — вектор из единиц соответствующей размерности.

фазу  $j$ , при условии, что в начальный момент в системе было не менее  $k + n$  заявок и процесс обслуживания находился на фазе  $i$ .

Матричные функции  $F_k(x)$  удовлетворяют рекуррентным соотношениям:

$$F_0(x) = e^{\Lambda x}; \tag{1}$$

$$F_k(x) = \int_0^x \sum_{m=1}^k F_{k-m}(y) Nq(m-1) F_0(x-y) dy, \quad k \geq 1, \tag{2}$$

а матрицы  $A_k$  рассчитываются по формуле:

$$A_k = \int_0^{\infty} F_k(x) dA(x), \quad k \geq 0. \tag{3}$$

Вывод этих соотношений ничем не отличается от вывода аналогичных соотношений для таких СМО, как, например, MAP/G/1/r или SM/MSP/n/r [6, 14]. Для получения расчетных формул для матриц  $F_k(x)$  и  $A_k$  можно воспользоваться методом, изложенным в [14, п. 7.3]. Пусть  $a = \max_{1 \leq i \leq l} |\Lambda_{ii}|$  — максимальный по модулю элемент главной диагонали матрицы  $\Lambda$ . Тогда из (1) имеем (см., например, [14, соотношение (28), с. 350]):

$$F_0(x) = e^{-ax} \sum_{i=0}^{\infty} \frac{(ax)^i}{i!} (I + a^{-1}\Lambda)^i.$$

Здесь и далее  $I$  — единичная матрица, порядок которой определяется из контекста. Матричные функции  $F_k(x)$ ,  $k \geq 1$ , будем искать в виде

$$F_k(x) = \sum_{i=1}^k \tilde{F}_i(x) a_{ki}, \tag{4}$$

где функции  $\tilde{F}_i(x)$  вычисляются по формулам:

$$\tilde{F}_k(x) = e^{-ax} \sum_{i=k}^{\infty} \frac{(ax)^i}{i!} F_{ki},$$

$$F_{ki} = \sum_{j=0}^{i-k} (I + a^{-1}\Lambda)^j (a^{-1}N) F_{k-1, i-j-1}, \quad i \geq k \geq 1.$$

Подставляя (4) в левую и правую части (2), получаем рекуррентную процедуру для расчета коэффициентов  $a_{ki}$ ,  $k \geq 1$ :

$$a_{k1} = q(k-1), \quad a_{ki} = \sum_{m=1}^{k+1-i} q(m-1)a_{k-m,i-1}, \quad 2 \leq j \leq k.$$

Заметим, что расчет матричных коэффициентов  $F_{ki}$  удобно осуществлять рекуррентным образом (см. [14, соотношение (31), с. 350]). Вернемся к матрицам  $A_k$ . Подставляя выражение (4) для  $F_k(x)$  в (3), получаем, что матрицы  $A_k$  можно представить в виде:

$$A_0 = \sum_{j=0}^{\infty} F_{0j} \alpha_j; \quad A_k = \sum_{i=1}^k a_{ki} \sum_{j=i}^{\infty} F_{ij} \alpha_j, \quad k \geq 1,$$

где  $\alpha_i = \int_0^{\infty} ((ax)^i / i!) e^{-ax} dA(x)$  — экспоненциальный момент порядка  $i$  функции распределения  $A(x)$ .

Вернемся к вложенной цепи Маркова. В дальнейшем потребуются следующие матрицы:

- $F_{kw}(x)$ ,  $0 \leq w \leq n-1$ ,  $k \geq w$ , — матрица, элемент  $(F_{kw}(x))_{ij}$ ,  $1 \leq i \leq l_k$ ,  $1 \leq j \leq l_w$ , которой представляет собой условную вероятность того, что за время  $x$  систему покинет  $k-w$  заявок и процесс обслуживания перейдет на фазу  $j$  при условии, что в начальный момент в системе (вместе с заявками на приборах) было  $k$  заявок, процесс обслуживания находился на фазе  $i$  и за время  $x$  в систему не поступила новая заявка;
- $A_{kw}$ ,  $0 \leq w \leq n-1$ ,  $k \geq w$ , — матрица, элемент  $(A_{kw})_{ij}$  которой представляет собой условную вероятность того, что за время между поступлениями заявок систему покинет  $k-w$  заявок и процесс обслуживания после поступления новой заявки перейдет на фазу  $j$ , при условии, что в начальный момент в системе было  $k$  заявок и процесс обслуживания находился на фазе  $i$ .

Матрицы  $F_{kw}(x)$  можно найти из соотношений:

$$F_{00}(x) = I; \quad F_{kk}(x) = e^{\Lambda_k x}, \quad 1 \leq k \leq n-1; \quad (5)$$

$$F_{kw}(x) = \int_0^x F_{k,w+1}(y) N_{w+1} F_{ww}(x-y) dy, \quad 0 \leq w \leq n-2, \quad w+1 \leq k \leq n-1; \quad (6)$$

$$F_{k,n-1}(x) = \int_0^x \sum_{m=0}^{k-n} F_m(y) N_n Q(k-n-m) F_{n-1,n-1}(x-y) dy, \quad k \geq n; \quad (7)$$

$$F_{kw}(x) = \int_0^x F_{k,w+1}(y) N_{w+1} F_{ww}(x-y) dy, \quad 0 \leq w \leq n-2, \quad k \geq n, \quad (8)$$

а матрицы  $A_{kw}$  — из соотношения:

$$A_{kw} = \int_0^\infty F_{kw}(x) \Omega_w dA(x), \quad 0 \leq w \leq n-1, \quad k \geq w. \quad (9)$$

Отметим, что для численного расчета матриц  $F_{kw}(x)$  и  $A_{kw}$  описанный выше метод не всегда пригоден. Это связано с тем, что максимальные (по модулю) диагональные элементы матриц  $\Lambda_k$ ,  $1 \leq k \leq n-1$ , могут быть не равны между собой. В этом случае (некоторым) выходом из положения может стать векторизация соотношений (5)—(9). Например, применяя оператор векторизации  $\text{vec}$  к левой и правой части (9) и воспользовавшись соотношением<sup>1</sup>  $\text{vec}(ABC) = (C^T B^T \otimes I) \text{vec}(A)$  (см. [15]), получим:

$$\text{vec}(A_{kw}) = \left( \Omega_w^T \otimes I \right) \int_0^\infty \text{vec}(F_{kw}(x)) dA(x).$$

Вычисление входящих в последнюю формулу векторов  $\text{vec}(F_{kw}(x))$  и интегралов может быть эффективно реализовано в программах, специально предназначенных для векторных вычислений.

Таким образом, матрица  $P$  переходных вероятностей вложенной цепи Маркова в указанных обозначениях имеет следующий вид:

$$P = \begin{pmatrix} A_{10} & A_{11} & 0 & \cdots & 0 & 0 & 0 & \cdots \\ A_{20} & A_{21} & A_{22} & \cdots & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots \\ A_{n-1,0} & A_{n-1,1} & A_{n-1,2} & \cdots & A_{n-1,n-1} & 0 & 0 & \cdots \\ A_{n0} & A_{n1} & A_{n2} & \cdots & A_{n,n-1} & A_0 & 0 & \cdots \\ A_{n+1,0} & A_{n+1,1} & A_{n+1,2} & \cdots & A_{n+1,n-1} & A_1 & A_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Обозначим через  $\pi_{ik}$  стационарную по вложенной цепи Маркова вероятность того, что в системе имеется  $k$  заявок и фаза обслуживания равна  $i$ , и положим

<sup>1</sup>Напомним, что здесь  $\otimes$  — символ Кронекера, а  $I$  — единичная квадратная матрица, размер которой равен числу строк матрицы  $A$ .

$\vec{\pi} = (\vec{\pi}_1, \vec{\pi}_2, \dots)$ ,  $\vec{\pi}_k = (\pi_{1k}, \dots, \pi_{l_k, k})$ . Для  $\vec{\pi}$  справедлива система уравнений равновесия (СУР)  $\vec{\pi} = \vec{\pi}P$ . Структура матрицы  $P$  позволяет использовать для решения СУР один из множества известных методов расчета стационарного распределения очереди в СМО типа  $G/M/1$  (см., например, [16, 17]) и, в частности, методы, предложенные в [6, разд. 3] и [2, разд. 2]. Вкратце остановимся на общем алгоритме расчета векторов  $\vec{\pi}_k$ . На первом шаге находится решение уравнения  $G = \sum_{k=0}^{\infty} G^k A_k$  в классе матриц, все собственные значения которых по модулю меньше единицы. Согласно [16] при выполнении условия существования стационарного режима такое решение существует и единственно. На следующем шаге решается система линейных уравнений<sup>1</sup>

$$\vec{\pi}_1 = \sum_{w=1}^{n-1} \vec{\pi}_w A_{w0} + \vec{\pi}_n A_0^*; \tag{10}$$

$$\vec{\pi}_k = \sum_{w=k-1}^{n-1} \vec{\pi}_w A_{w, k-1} + \vec{\pi}_n A_{k-1}^*, \quad 2 \leq k \leq n, \tag{11}$$

где

$$A_k^* = \sum_{w=n}^{\infty} G^{w-n} A_{wk}, \quad 0 \leq k \leq n-1,$$

с точностью до константы, которая находится из условия нормировки

$$\sum_{k=1}^{n-1} \vec{\pi}_k \vec{1} + \vec{\pi}_n (I - G)^{-1} \vec{1} = 1.$$

Для решения (10)–(11) удобно, например, воспользоваться методом исключения состояний (см., например, [2, разд. 2] или [18]). Наконец, на третьем шаге рассчитываются вероятности  $\vec{\pi}_k = \vec{\pi}_n G^{k-n}$ ,  $k \geq n+1$ .

Зная стационарное распределение вложенной цепи Маркова, уже можно определить ряд других стационарных характеристик очереди. Например, вводя обозначение  $\vec{\pi}_k^- = (\pi_{1k}^-, \dots, \pi_{l_{k+1}, k}^-)$ ,  $k \geq 0$ , где  $\pi_{ik}^-$  — стационарная вероятность того, что поступающая в систему заявка застанет в ней  $k$  других заявок и марковский процесс обслуживания после ее поступления окажется на фазе  $i$ , немедленно можем записать  $\vec{\pi}_k^- = \vec{\pi}_{k+1}$ .

<sup>1</sup> Известно (см., например, [6, с. 92]), что для того, чтобы эта система уравнений имела решение, достаточно, чтобы  $\sum_{k=0}^{n-1} A_k^* \vec{1} = \vec{1}$ . В справедливости последнего равенства можно убедиться непосредственной проверкой. Действительно, интегрируя (9) по частям, получаем  $\sum_{k=0}^{n-1} A_k^* \vec{1} = \sum_{j=0}^{\infty} q(j) \sum_{i=0}^j G^i U N_n \vec{1}$ , где  $U = \sum_{k=0}^{\infty} G^k T_k$ . Учитывая, что  $G \vec{1} = \sum_{k=0}^{\infty} G^k A_k \vec{1} = \vec{1} - U N \vec{1} + \sum_{i=0}^{\infty} G^{i+1} U N q(i) \vec{1}$ , матрица  $G$  невырождена и  $N_n \vec{1} = N \vec{1}$ , простым переупорядочением слагаемых приходим к искомому равенству.



Для нахождения стационарных вероятностей состояний по времени введем матрицы  $T_k$ ,  $k \geq 0$ , и  $T_{kw}$ ,  $0 \leq w \leq n - 1$ ,  $k \geq w$ . Элемент  $(T_k)_{ij}$ ,  $1 \leq i, j \leq l$ , матрицы  $T_k$  представляет собой среднее время, проведенное на интервале между соседними моментами поступления заявок системой в состоянии  $(j, M - k)$  при условии, что после поступления первой заявки в системе оказалось  $M$ ,  $M > k$ , заявок и фаза обслуживания была  $i$ . Элемент  $(T_{kw})_{ij}$ ,  $1 \leq i \leq l_w$ ,  $0 \leq j \leq l_w$ , матрицы  $T_{kw}$  есть среднее время, проведенное на интервале между соседними моментами поступления заявок системой в состоянии  $(j, w)$  при условии, что после поступления первой заявки в системе оказалось  $k$ ,  $k \geq w$ , заявок и фаза обслуживания была  $i$ . Матрицы  $T_k$  и  $T_{kw}$  определяются соотношениями:

$$T_k = \int_0^{\infty} (1 - A(x)) F_k(x) dx; \quad T_{kw} = \int_0^{\infty} (1 - A(x)) F_{kw}(x) dx.$$

Убедиться в этом можно, применяя рассуждения, аналогичные тем, что приведены в [19, с. 285]. Используя результаты теории полумарковских процессов, получаем для векторов  $\vec{p}_k$ ,  $k \geq 0$ , стационарных вероятностей состояний по времени формулы:

$$\vec{p}_0 = \frac{1}{a} \sum_{w=1}^{\infty} \vec{\pi}_w T_{w0}; \tag{12}$$

$$\vec{p}_k = \begin{cases} \frac{1}{a} \sum_{w=k}^{\infty} \vec{\pi}_w T_{wk}, & 1 \leq k \leq n - 1; \\ \frac{1}{a} \sum_{w=k}^{\infty} \vec{\pi}_w T_{w-k}, & k \geq n. \end{cases} \tag{13}$$

Заметим, что формулы (12) и (13) можно получить и из результатов теории процессов восстановления (см., например, [14, с. 328]).

#### 4 Стационарные распределения времен ожидания и пребывания обслуженной заявки в системе

Выпишем в терминах преобразования Лапласа–Стилтьеса (ПЛС) некоторые стационарные характеристики, связанные с временем ожидания начала обслуживания и временем пребывания заявки в системе.

Обозначим через  $U_k(x)$ ,  $k \geq 1$ , матрицу, элемент  $(U_k(x))_{i,j}$ ,  $1 \leq i, j \leq l$ , которой есть вероятность того, что поступившая в систему заявка перейдет из очереди на обслуживание за время  $x$  и процесс обслуживания перейдет на фазу  $j$ , при условии, что она в момент поступления застала в системе  $n + k - 1$  других

заявок и фаза обслуживания была  $i$ , а через  $\tilde{U}_k(s) = \int_0^\infty e^{-sx} dU_k(x)$  — ПЛС матрицы  $U_k(x)$ . Тогда

$$\tilde{U}_k(s) = \sum_{i=1}^k (sI - \Lambda)^{-1} Nq(i-1) \tilde{U}_{k-i}(s), \quad k \geq 1,$$

где для сокращения записи используется соглашение  $\tilde{U}_0(s) \equiv I$ . Тогда ПЛС  $\tilde{W}_{\text{serv}}(s)$  стационарного распределения  $W_{\text{serv}}(x)$  пребывания в очереди попавшей на обслуживание заявки примет вид:

$$\tilde{W}_{\text{serv}}(s) = \frac{1}{1 - p_{\text{loss}}} \left( \sum_{k=0}^{n-1} \vec{\pi}_k^- \vec{1} + \sum_{k=n}^{\infty} \vec{\pi}_k^- \tilde{U}_{k-n+1}(s) \vec{1} \right), \quad (14)$$

где  $p_{\text{loss}}$  — стационарная вероятность того, что поступившая в систему заявка не попадет на обслуживание (т. е. будет потеряна), которая равна

$$p_{\text{loss}} = \sum_{k=n}^{\infty} \gamma(k-n+1) \vec{\pi}_k^- \vec{1},$$

а  $\gamma(k)$ ,  $k \geq 1$ , — стационарная вероятность того, что заявка, находящаяся в очереди на  $k$ -м месте, будет «убита» (какой-либо) обслуженной заявкой. Для рассматриваемой системы она вычисляется по формуле

$$\gamma(k) = \sum_{m=1}^k \gamma_m(k),$$

где

$$\gamma_1(k) = Q(k), \quad k \geq 1; \quad \gamma_m(k) = \sum_{i=0}^{k-m} q(i) \gamma_{m-1}(k-1-i), \quad 2 \leq m \leq k, \quad k \geq 2,$$

и совпадает с вероятностью  $\gamma(k)$  в [1, с. 136].

Перейдем к нахождению стационарного распределения времени пребывания в системе обслуженной заявки. Введем матрицы  $\tilde{V}_k(s, y)$ ,  $k \geq 1$ ,  $y \geq 0$ . Элемент  $(\tilde{V}_k(s, y))_{ij}$  матрицы  $\tilde{V}_k(s, y)$  есть ПЛС времени пребывания заявки на обслуживании и вероятность того, что в момент ее ухода из системы фаза обслуживания равна  $j$  при условии, что сразу после ее поступления на обслуживание занятыми оказались  $\min(k, n)$  приборов, в очереди было  $\max(0, k-n)$  заявок, время до прихода следующей заявки было равно  $y$  и фаза обслуживания была

равна  $i$ . Учитывая, что все приборы предполагаются одинаковыми и, таким образом, при окончании обслуживания равновероятно может освободиться любой из (занятых) приборов, по свойствам ПЛС имеем:

$$\tilde{V}_k(s, y) = \begin{cases} e^{-(sI-\Lambda_k)y} \tilde{V}_{k+1}(s) + \\ \int_0^y e^{-(sI-\Lambda_k)u} N_k \left( \frac{1}{k} + \frac{k-1}{k} \tilde{V}_{k-1}(s, y-u) \right) du, & 1 \leq k \leq n-1; \\ \int_0^y e^{-(sI-\Lambda)u} N_n \left( \frac{1}{n} + \frac{n-1}{n} \tilde{V}_{n-1}(s, y-u) \right) du + e^{-(sI-\Lambda)y} \tilde{V}_{n+1}(s), & k = n; \\ e^{-(sI-\Lambda)y} \tilde{V}_{k+1}(s) + \int_0^y e^{-(sI-\Lambda)u} N \left( \frac{1}{n} + \right. \\ \left. + \frac{n-1}{n} \left( \sum_{j=0}^{k-n-1} q(j) \tilde{V}_{k-1-j}(s, y-u) + Q(k-n) \tilde{V}_{n-1}(s, y-u) \right) \right) du, & k \geq n+1, \end{cases}$$

где

$$\tilde{V}_k(s) = \int_0^\infty \tilde{V}_k(s, y) dA(y).$$

Тогда заявка, поступающая в систему, когда занято  $k$ ,  $0 \leq k \leq n-1$ , приборов немедленно становится на обслуживания и ПЛС времени ее пребывания в системе (с учетом смен фаз обслуживания) равно  $\tilde{V}_{k+1}(s)$ . Заявка, застающая при поступлении  $k$ ,  $k \geq n$ , заявок в системе, ожидает некоторое время в очереди, которое (с учетом смен фаз обслуживания) с вероятностью  $dU_{k-n+1}(x)$  равно  $x$ . При этом в момент поступления заявки на обслуживание в системе окажется  $m$  заявок в очереди и время до прихода следующей заявки будет равно  $y$  с плотностью вероятности  $a_m(x, y)$ :

$$a_0(x, y) = A^{*1}(x+y); \quad a_m(x, y) = \int_0^x A^{*m}(u) A^{*1}(x-u+y) du, \quad m \geq 1,$$

где  $A^{*1}(x)$  определяются рекуррентно по формулам

$$A^{*1}(x) = A'(x); \quad A^{*m}(x) = \int_0^x A^{*(m-1)}(x-u) A^{*1}(u) du.$$

Таким образом, ПЛС  $\tilde{V}_{\text{serv}}(s)$  стационарного распределения  $V_{\text{serv}}(x)$  времени пребывания в системе попавшей на обслуживание заявки имеет вид:

$$\tilde{V}_{\text{serv}}(s) = \frac{1}{1 - p_{\text{loss}}} \left( \sum_{k=0}^{n-1} \tilde{\pi}_k^- \tilde{V}_{k+1}(s) \bar{\mathbf{1}} + \sum_{k=n}^{\infty} \tilde{\pi}_k^- \int_0^{\infty} e^{-sx} dU_{k-n+1}(x) \int_0^{\infty} \sum_{m=0}^{\infty} a_m(x, y) \tilde{V}_{n+m}(s, y) \bar{\mathbf{1}} dy \right). \quad (15)$$

## 5 Заключение

При некоторых дополнительных ограничениях<sup>1</sup> на  $A(x)$  большую часть выражений для матриц можно упростить и получить в явном виде.

Дифференцируя выражения (14) и (15) необходимое число раз, можно получить моменты любых порядков времени ожидания начала обслуживания и времени пребывания в системе (обслуженной) заявки. Для нахождения средних значений наиболее подходит формула Литтла. Для других моментов (14) еще годится для практических расчетов, а (15) уже становится едва ли пригодной. В связи с этим представляется интересным исследовать специальные случаи входного потока (и процесса обслуживания), когда (15) приобретает более простой вид.

В заключение также отметим, что полученные в работе результаты достаточны для нахождения специальных показателей качества управления очередью (например, распределение числа последовательных потерь) и, таким образом, обеспечивают возможность сравнения механизма обновления с другими известными механизмами активного управления очередями.

## Литература

1. Зарядов И. С. Система массового обслуживания GI/M/n/∞ с обобщенным обновлением // Автоматика и телемеханика, 2010. Вып. 4. С. 130–139.
2. Бочаров П. П., Д'Анчиче Ч., Печинкин А. В., Салерно С. Стационарные характеристики системы массового обслуживания G/MSP/1/r // Автоматика и телемеханика, 2003. № 2. С. 127–143.
3. Бочаров П. П., Зарядов И. С. Стационарное распределение вероятностей в системах массового обслуживания с обновлением // Вестник РУДН. Сер. Математика. Информатика. Физика, 2007. № 1-2. С. 15–25.
4. Зарядов И. С. Стационарные характеристики обслуживания в системе G/M/n/r с обобщенным обновлением // Вестник РУДН. Сер. Математика. Информатика. Физика, 2008. № 2. С. 3–9.

<sup>1</sup>Это, например, случай, когда  $A(x)$  имеет дробно-рациональное ПЛС.

5. Бочаров П. П. Стационарное распределение конечной очереди с рекуррентным потоком и марковским обслуживанием // Автоматика и телемеханика, 1996. № 9. С. 66–78.
6. Печинкин А. В., Чаплыгин В. В. Стационарные характеристики системы массового обслуживания  $G/MSP/n/r$  // Вестник РУДН. Сер. Прикладная математика и информатика, 2003. № 1. С. 119–143.
7. Печинкин А. В., Чаплыгин В. В. Стационарные характеристики системы массового обслуживания  $SM/MSP/n/r$  // Автоматика и телемеханика, 2004. № 9. С. 85–100.
8. Чаплыгин В. В. Система массового обслуживания  $G/MSP/n/r$  с потоком отрицательных заявок // Информационные процессы, 2005. Т. 5. № 1. С. 1–19.
9. Chydzinski A., Chrost L. Analysis of AQM queues with queue size based packet dropping // Int. J. Appl. Math. Comp., 2011. Vol. 21. Iss. 3. P. 567–577. doi: 10.2478/v10006-011-0045-7.
10. Zaryadov I. S., Razumchik R. V., Milovanova T. A. Stationary waiting time distribution in  $G/M/n/r$  with random renovation policy // Comm. Comp. Inf. Sc., 2016. Vol. 678. P. 418–429. doi: 10.1007/978-3-319-51917-331.
11. Chydzinski A., Mrozowski P. Queues with dropping functions and general arrival processes // PLoS ONE, 2016. Vol. 11. Iss. 3. Art. ID: e0150702. doi: 10.1371/journal.pone.0150702.
12. Kononov M., Razumchik R. Comparison of two active queue management schemes through the  $M/D/1/N$  queue // Информатика и её применения, 2018. Т. 12. Вып. 4. С. 9–15. doi: 10.14357/19922264180402.
13. Zaryadov I., Bogdanova E., Milovanova T., Matushenko S., Pyatkina D. Stationary characteristics of the  $GI/M/1$  queue with general renovation and feedback // 10th Congress (International) on Ultra Modern Telecommunications and Control Systems and Workshops, 2018. P. 1–6. doi: 10.1109/icumt.2018.8631244.
14. Бочаров П. П., Печинкин А. В. Теория массового обслуживания. — М.: РУДН, 1995. 529 с.
15. Steeb W. H., Hardy Y. Matrix calculus and Kronecker product: A practical approach to linear and multilinear algebra. — River Edge, NJ, USA: World Scientific, 2011. 324 p.
16. Neuts M. F. Matrix-geometric solutions in stochastic models. An algorithmic approach. — Baltimore–London: The Johns Hopkins University Press, 1981. 332 p.
17. Latouche G., Ramaswami V. Introduction to matrix analytic methods in stochastic modeling. — ASA-SIAM ser. on statistics and applied probability. — Philadelphia, PA, USA: SIAM, 2000. 348 p.
18. Печинкин А. В., Разумчик Р. В. Системы массового обслуживания в дискретном времени. — М.: Физматлит, 2018. 432 с.
19. Bocharov P. P., D'Apice C., Pechinkin A. V., Salerno S. Queueing theory. — Utrecht, Boston: VSP, 2004. 446 p.

Поступила в редакцию 01.09.19

## STATIONARY CHARACTERISTICS OF THE GI/MSP/ $n$ / $\infty$ QUEUE WITH GENERAL RENOVATION

I. S. Zaryadov<sup>1,2</sup>, L. A. Meykhanadzhyan<sup>3</sup>, and T. A. Milovanova<sup>1</sup>

<sup>1</sup>Peoples' Friendship University of Russia (RUDN University), Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation

<sup>2</sup>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

<sup>3</sup>Financial University under the Government of the Russian Federation, 49 Leningradsky Prosp., Moscow 125993, Russian Federation

**Abstract:** Consideration is given to the GI/MSP/ $n$ / $\infty$  queue with general input flow of customers,  $n$  identical servers, service process of markovian type, queue of infinite capacity, and general renovation. General renovation being the variant of an active queue management mechanism, implies that upon a service completion, a customer may remove a random number of customers from the queue (if any is available), with a given probability distribution. Using embedded Markov chain technique, one derives stationary distributions of the main system's performance characteristics. The obtained results are ready for numerical implementation and allow one to compute stationary distributions of the system size, stationary loss probability, and waiting time distribution (under FIFO (first in, first out) service and head-of-the-queue renovation).

**Keywords:** queueing system; general renovation; markovian service process; queue management; embedded Markov chain

**DOI:** 10.14357/08696527190405

### Acknowledgments

The study was funded by the Russian Foundation for Basic Research, project No. 19-07-00739.

### References

1. Zaryadov, I. S. 2010. The GI/M/ $n$ / $\infty$  queuing system with generalized renovation. *Automat. Rem. Contr.* 71(4):663–671.
2. Bocharov, P. P., C. D'Apice, A. V. Pechinkin, and S. Salerno. 2003. The stationary characteristics of the G/MSP/1/ $r$  queueing system. *Automat. Rem. Contr.* 64(2):288–301.
3. Bocharov, P. P., and I. S. Zaryadov. 2007. Statsionarnoe raspredelenie veroyatnostey v sistemakh massovogo obsluzhivaniya s obnovleniem [Stationary probability distribution of the queueing system with renovation]. *Vestnik RUDN. Ser. Matematika. Informatika. Fizika* [Bulletin of Peoples' Friendship University of Russia. Ser. Mathematics. Information Sciences. Physics] 1-2:15–25.

4. Zaryadov, I. S. 2008. Stacionarnye kharakteristiki obsluzhivaniya v sisteme  $G/M/n/r$  s obobshchennym obnovleniem [Queueing system  $G/M/n/r$  with general renovation. Stationary characteristics]. *Vestnik RUDN. Ser. Matematika. Informatika. Fizika* [RUDN J. Mathematics Information Sciences Physics] 2:3–9.
5. Bocharov, P. P. 1996. Stationary distribution of a finite queue with recursive input flow and Markovian service discipline. *Automat. Rem. Contr.* 57(9):1274–1283.
6. Pechinkin, A. V., and V. V. Chaplygin. 2003. Stacionarnye kharakteristiki sistemy massovogo obsluzhivaniya  $G/MSP/n/r$  [The stationary characteristics of the  $G/MSP/n/r$  queueing system]. *Vestnik RUDN. Ser. Prikladnaya matematika i informatika* [Bulletin of Peoples' Friendship University of Russia. Ser. Applied Mathematics and Information Sciences] 1:119–143.
7. Pechinkin, A. V., and V. V. Chaplygin. 2004. Stationary characteristics of the SM/MSP/ $n/r$  queueing system. *Automat. Rem. Contr.* 65(9):1429–1443.
8. Chaplygin, V. V. 2005. Sistema massovogo obsluzhivaniya  $G/MSP/n/r$  s potokom otritsatel'nykh zayavok [A queueing system  $G/MSP/n/r$  with flow of negative claims]. *Informatsionnye protsessy* [Information Processes] 5(1):1–19.
9. Chydzinski, A., and L. Chrost. 2011. Analysis of AQM queues with queue size based packet dropping. *Int. J. Appl. Math. Comp.* 21(3):567–577. doi: 10.2478/v10006-011-0045-7.
10. Zaryadov, I. S., R. V. Razumchik, and T. A. Milovanova. 2016. Stationary waiting time distribution in  $G/M/n/r$  with random renovation policy. *Comm. Comp. Inf. Sc.* 678:418–429. doi: 10.1007/978-3-319-51917-3\_31.
11. Chydzinski, A., and P. Mrozowski. 2016. Queues with dropping functions and general arrival processes. *PLoS ONE* 11(3):e0150702. doi: 10.1371/journal.pone.0150702.
12. Kononov, M., and R. Razumchik. 2018. Comparison of two active queue management schemes through the  $M/D/1/N$  queue. *Informatika i ee Primeneniya — Inform. Appl.* 12(4):9–15. doi: 10.14357/19922264180402.
13. Zaryadov, I., E. Bogdanova, T. Milovanova, S. Matushenko, and D. Pyatkina. 2018. Stationary characteristics of the GI/M/1 queue with general renovation and feedback. *10th Congress (International) on Ultra Modern Telecommunications and Control Systems and Workshops*. 1–6. doi: 10.1109/icumt.2018.8631244.
14. Bocharov, P. P., and A. V. Pechinkin. 1995. *Teoriya massovogo obsluzhivaniya* [Queueing theory]. Moscow: RUDN. 529 p.
15. Steeb, W. H., and Y. Hardy 2011. *Matrix calculus and Kronecker product: A practical approach to linear and multilinear algebra*. River Edge, NJ: World Scientific. 324 p.
16. Neuts, M. F. 1981. *Matrix-geometric solutions in stochastic models. An algorithmic approach*. Baltimore–London: The Johns Hopkins University Press. 332 p.
17. Latouche, G., and V. Ramaswami. 2000. *Introduction to matrix analytic methods in stochastic modeling*. ASA-SIAM ser. on statistics and applied probability. Philadelphia, PA: SIAM. 348 p.
18. Pechinkin, A. V., and R. V. Razumchik. 2018. *Sistemy massovogo obsluzhivaniya v diskretnom vremeni* [Discrete time queueing systems]. Moscow: Fizmatlit. 432 p.
19. Bocharov, P. P., C. D'Apice, A. V. Pechinkin, and S. Salerno. 2004. *Queueing theory*. Utrecht, Boston: VSP. 446 p.

*Received September 1, 2019*

## **Contributors**

**Zaryadov Ivan S.** (b. 1981) — Candidate of Science (PhD) in physics and mathematics, associate professor, Department of Applied Informatics and Probability Theory, Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; senior scientist, Institute of Informatics Problems, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences; izaryadov@sci.pfu.edu.ru

**Meykhanadzhyan Lusine A.** (b. 1990) — Candidate of Science (PhD) in physics and mathematics, associate professor, Department of Data Analysis, Decision-Making and Financial Technology, Financial University under the Government of the Russian Federation, 49 Leningradsky Prosp., Moscow 125993, Russian Federation; lamejkhanadzhyan@fa.ru

**Milovanova Tatiana A.** (b. 1977) — Candidate of Science (PhD) in physics and mathematics, senior lecturer, Department of Applied Informatics and Probability Theory, Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; tmilovanova77@mail.ru



## МЕТОД СЛУЧАЙНОГО ОТБОРА ПРИ ПРОГНОЗИРОВАНИИ ВРЕМЕННЫХ РЯДОВ РЫНКА КРИПТОВАЛЮТ

*О. Е. Сороколетова<sup>1</sup>, Т. В. Захарова<sup>2</sup>*

**Аннотация:** Работа посвящена рассмотрению применения метода случайного отбора, или RSM-метода, в задаче классификации (прогнозирования динамики) нестационарных временных рядов рынка криптовалют. Название метода RSM — аббревиатура его полного названия Random Sampling Method. RSM представляет собой метод глубокого обучения. Одними из основных методов глубокого обучения среди использовавшихся ранее для решения данной задачи и показавших свою эффективность остаются рекуррентные LSTM (long short term memory) нейросети. В настоящей работе представлена более гибкая архитектура, построенная на базе LSTM-ячеек и имеющая таким образом все преимущества традиционного алгоритма, однако более устойчивая к проблеме дисбаланса классов. Основным отличительным признаком RSM является использование метрического обучения.

**Ключевые слова:** криптовалюта; временные ряды; задача классификации; прогнозирование динамики; метрическое обучение; LSTM; нейросети; глубокое обучение

**DOI:** 10.14357/08696527190406

### 1 Введение

Взгляды на вопрос применимости методов машинного (в том числе глубокого) обучения в качестве инструмента для алгоритмической торговли полярны, и в этом кроется причина большого числа дискуссий и появления множества научных, а также экономических и финансовых публикаций, посвященных доказательству эффективности того или иного инструмента. Криптовалютные временные ряды — это феномен недавнего времени, поэтому он мало изучен и представляет интерес для исследований. В данной работе рассматривается алгоритм машинного обучения, эффективный при анализе временных рядов рынка криптовалют.

Методы глубокого обучения требуют большого объема данных для тренировки. В случае криптовалютных временных рядов это означает, что данные должны быть высокочастотными, т. е., как показывает практика, интервал временной метки должен составлять минуту и не более. Однако на большом объеме

---

<sup>1</sup>Московский государственный университет имени М. В. Ломоносова, факультет вычислительной математики и кибернетики, olgas020697@gmail.com

<sup>2</sup>Московский государственный университет имени М. В. Ломоносова, факультет вычислительной математики и кибернетики; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, tvzaharova@mail.ru

высокочастотных данных, особенно таких нестабильных, как криптовалюта, возникает новая проблема — проблема нестационарности, с которой и призван бороться метод случайного отбора.

Метод случайного отбора представляет собой метод метрического обучения, который учится оценивать схожесть между целевым и эталонным экземпляром перед выставлением метки, где количество эталонных экземпляров должно быть равно числу классов в задаче классификации. Важной особенностью метода является алгоритм выбора эталонных последовательностей, который будет определен ниже как схема отбора.

Метрическое обучение хорошо изучено на задаче классификации изображений. В случае значительно ограниченного числа входных данных и, главное, недостаточных знаний об их природе, что, в частности, не позволяет обнаружить подходящее признаковое пространство (как качественно, так и количественно), установить метрику бывает затруднительно. Этот случай полностью соответствует характеристикам рынка криптовалют. При решении задачи классификации изображений в работе [1] схожие затруднения были разрешены при помощи вовлечения нейросетевой архитектуры. В данной статье также при помощи нейронной сети рассмотрен алгоритм решения задачи прогнозирования временных рядов.

Идеи реализации метрического обучения в задаче классификации описаны в источниках [2, 3].

## 2 Постановка задачи классификации

Формальная постановка задачи предсказания движения цены акции на рынке может быть представлена следующим образом.

Имеются три возможных варианта движения цены после временной отметки  $t$ : вверх (up), вниз (down) или без изменений (static). Будем рассматривать распределение логарифмической доходности  $R_t$ :

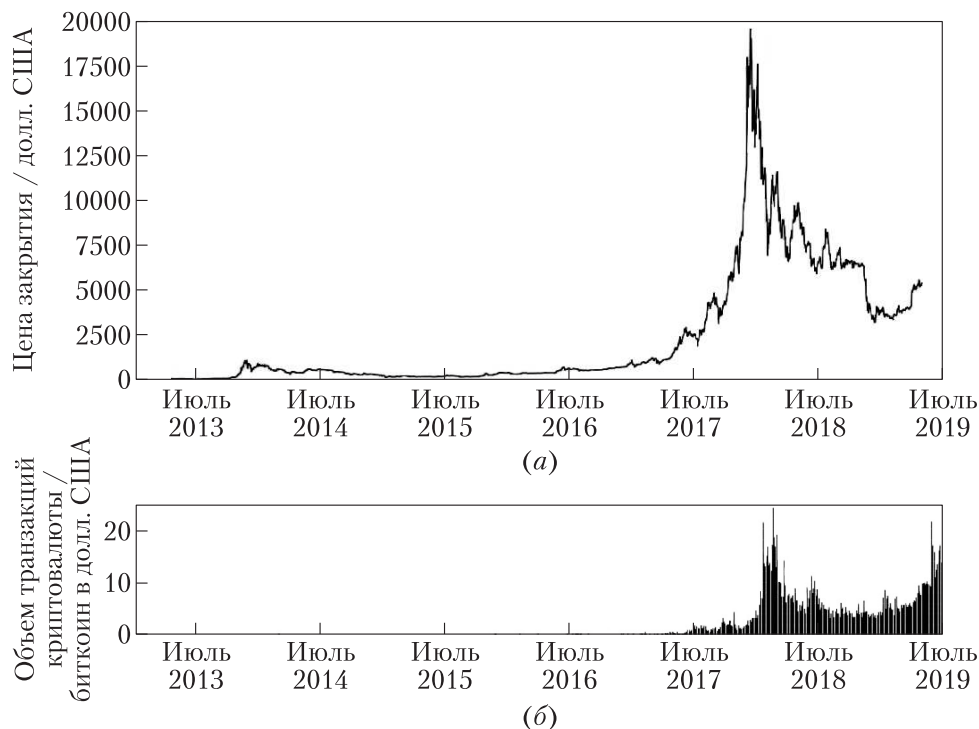
$$R_t = \log \left( \frac{P_t}{P_{t-1}} \right),$$

где  $P_t$  — значение цены в момент времени  $t$ , взятое из временного ряда. Обозначим через  $p_t^{(\text{up})}$ ,  $p_t^{(\text{down})}$  и  $p_t^{(\text{static})}$  вероятности наступления соответствующих верхним индексам событий. Эти вероятности будут получены в результате обучения модели. Таким образом, возникает задача трехклассовой классификации:

$$X_t \in \{\text{up}, \text{down}, \text{static}\}.$$

## 3 Предобработка данных и схема отбора

На вход модели подаются OHLC-временные ряды (Open, High, Low, Close — цена открытия, наиболее высокая цена, наиболее низкая цена и цена



**Рис. 1** Цены закрытия и динамика объема транзакций криптовалюты биткоин в долларах США

закрытия соответственно). Как известно, дисбаланс между значениями признаков может вызвать неустойчивость работы модели, ухудшить результаты обучения и замедлить его процесс, поэтому необходимо производить нормализацию данных. Так как изучаемые ряды нестационарны, будем сначала извлекать последовательность из набора, а затем применять к ней нормализацию. Для того чтобы определиться со стратегией нормализации, обратимся к рис. 1.

Большинство криптовалютных временных рядов, как и тот, что представлен на рис. 1, имеют пик в конце 2017 г., а следовательно, известен максимум ряда. Из этого можно заключить, что оптимальным выбором будет применение минимаксной нормализации:

$$x_{ti}^{\text{norm}} = \frac{x_{ti} - \min(x_t)}{\max(x_t) - \min(x_t)},$$

где индекс  $i$  соответствует номеру признака (Open, High, Low или Close), индекс  $t$  — временной отметке во входной последовательности, а операции взятия

минимума и максимума применяются к компонентам  $x_{ti}$  для всех  $i = 1, \dots, 4$  на отрезке  $[t - j + 1, t]$ , где  $j$  — длина входной последовательности. В результате получим значения от 0 до 1.

Следующим шагом избавимся от очевидных выбросов, например, по порогу  $\alpha$ : удалим имеющие метку «up» или «down» последовательности, для которых  $|R_t| > \alpha$ . На этом вся необходимая предобработка данных будет выполнена, и можно приступать к применению схемы отбора.

Схема отбора — это алгоритм, по которому осуществляется выбор тройки эталонных последовательностей, подающейся затем на вход модели. Она построена, основываясь на предположении, что финансовые данные нестационарны, что обуславливает необходимость выбора последовательностей исключительно среди недавнего прошлого. Более формально, для заданной входной последовательности  $x_t$  схема подразумевает выбор последовательностей из интервала  $[t - k - l, t - k]$ , где  $k$  — размер окна для простого скользящего среднего, а  $l$  — гиперпараметр, определяющий размер этого интервала.

## 4 Основная идея метода случайного отбора

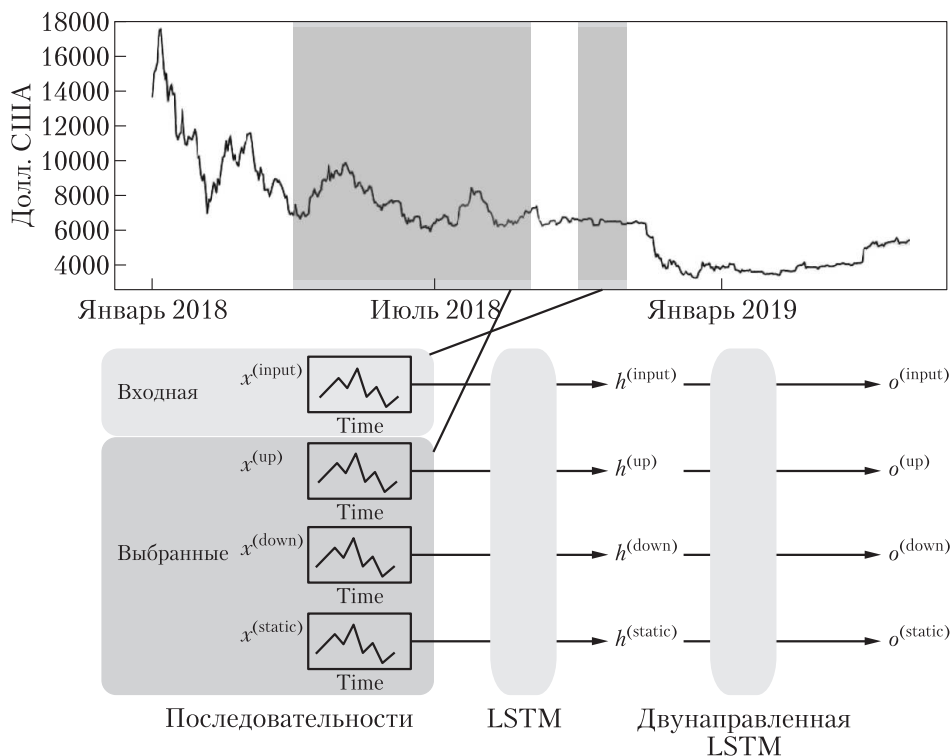
В основе метода лежит предположение о том, что схожесть пары последовательностей может быть охарактеризована при помощи классов, к которым они принадлежат: последовательность с меткой «up» более близка последовательности с меткой «up», чем последовательности с меткой «down» или «static».

На вход модели подаются пары элементов, один из которых — это последовательность, подлежащая классификации, а другой — набор последовательностей (три последовательности, помеченные как «up», «down» и «static» соответственно), выбранных из недавнего прошлого. Затем осуществляется получение внутренних представлений путем независимого кодирования каждой последовательности и сравнение внутреннего представления входной последовательности и внутренних представлений выбранных последовательностей. Выходным значением является класс, к которому предположительно принадлежит целевая последовательность. Работа метода представлена на рис. 2.

Оптимизация параметров модели производится таким образом, чтобы достигались внутренние представления, для которых верно: среди внутренних представлений последовательностей из тройки более близко внутреннему представлению целевой последовательности то, которое имеет метку класса, совпадающую с истинной в данном случае меткой. Близость внутренних представлений оценивается при помощи косинусной меры сходства.

## 5 Работа кодировщика

Рассмотрим процесс получения внутренних представлений подробнее. Кодировщик, осуществляющий эту работу, состоит из двух частей: LSTM и дву-



**Рис. 2** Полная схема метода. Дана входная последовательность  $x_t^{(input)}$ , последовательности  $x_t^{(up)}$ ,  $x_t^{(down)}$  и  $x_t^{(static)}$  случайно выбираются из темного окна  $[t - k - l, t - k]$ . Затем входная и выбранные последовательности независимо проходят через кодировщик с целью получения и обработки внутренних представлений: LSTM и двунаправленную LSTM

направленной LSTM. Долгая краткосрочная память (LSTM) — это особая разновидность архитектуры рекуррентных нейронных сетей, способная к обучению долговременным зависимостям. Не будем останавливаться на ней подробно, но заметим, что в ее блоке используются четыре взаимодействующих слоя, функционирование которых описывает следующая система уравнений:

$$\begin{aligned}
 i_t &= \sigma(W_{xi}x_t + W_{mi}m_{t-1} + W_{ci}c_{t-1} + b_i); \\
 f_t &= \sigma(W_{xf}x_t + W_{mf}m_{t-1} + W_{cf}c_{t-1} + b_f); \\
 c_t &= f_t c_{t-1} + i_t \tanh(W_{xc}x_t + W_{mc}m_{t-1} + b_c); \\
 m_t &= \text{out}_t \tanh(c_t), \quad \text{out}_t = \sigma(W_{xo}x_t + W_{mo}m_{t-1} + W_{co}c_t + b_o),
 \end{aligned}$$

где  $\sigma$  — сигмоидальная функция;  $W_{ij}$  и  $b_j$  — параметры модели;  $x_t$  несет смысл вновь приобретенной информации, а  $m_t$  — накопленной памяти.

Для того чтобы описать первую часть кодировщика, зафиксируем входную последовательность в момент времени  $t$ :  $x_t^{(\text{input})}$ , а также соответствующие ей выбранные  $x_t^{(\text{up})}$ ,  $x_t^{(\text{down})}$  и  $x_t^{(\text{static})}$ . Затем преобразуем их независимо во внутренние представления  $h_t^{(\text{input})}$ ,  $h_t^{(\text{up})}$ ,  $h_t^{(\text{down})}$  и  $h_t^{(\text{static})}$  с помощью LSTM:

$$h_t = \text{LSTM}(x_t).$$

Здесь верхние индексы не указаны для краткости.

Следующим шагом необходимо установить отношения между  $h_t^{(\text{input})}$ ,  $h_t^{(\text{up})}$ ,  $h_t^{(\text{down})}$  и  $h_t^{(\text{static})}$ , за что отвечает вторая часть кодировщика — двунаправленная LSTM-нейросеть. Она пропускает последовательности в прямом ( $h_t^{(\text{input})}$ ,  $h_t^{(\text{up})}$ ,  $h_t^{(\text{down})}$ ,  $h_t^{(\text{static})}$ ) и обратном ( $h_t^{(\text{static})}$ ,  $h_t^{(\text{down})}$ ,  $h_t^{(\text{up})}$ ,  $h_t^{(\text{input})}$ ) порядке, генерируя выходное значение как результат сложения:

$$o_{ti} = h_{ti}^{\rightarrow} + h_{ti}^{\leftarrow} + h_{ti},$$

где  $h_{ti}^{\rightarrow}$  ( $h_{ti}^{\leftarrow}$ ) —  $i$ -й выход из двунаправленной LSTM-нейросети при прямом (обратном) прохождении;  $o_{ti}$  —  $i$ -й выход кодировщика.

Стандартно при решении задачи классификации во время обучения модели производится минимизация кросс-энтропийной функции потерь, не составляет исключения и данный случай. Выход представлен one-hot-векторами, в которых на месте верного класса установлена 1, а все остальные разряды обращены в 0.

Обозначим через  $e_t$  среднее цен скользящего окна размера  $T$ . Тогда целевые метки представимы следующим образом:

$$y_t = \begin{cases} -1, & \text{если } e_t > e_{t+T} + \epsilon; \\ 1, & \text{если } e_t < e_{t+T} - \epsilon; \\ 0 & \text{иначе.} \end{cases}$$

Здесь  $\epsilon$  — пороговый параметр, введенный с целью контроля баланса классов.

## 6 Заключение

Криптовалютный рынок — получившая стремительное развитие и нестабильная область финансовой системы, требующая комплексного изучения. В частности, актуальна задача прогнозирования временных рядов рынка криптовалют, решение которой — трудоемкий процесс, имеющий множество нюансов. Применение глубоких технологий открывает новые возможности во многих прикладных областях, так как нейросети, благодаря своему алгоритмическому устройству, способны выучивать сложные нелинейные зависимости и выявлять скрытые

структурные признаки. Это делает поиски универсального инструмента для алгоритмической торговли на рынке криптовалют среди методов глубокого обучения осмысленными.

Основной целью данной работы ставилось исследование нового метода обучения — метода случайного отбора (RSM) — на задаче классификации (предсказания динамики цен) временных рядов, специфика которого заключается в том, что он имеет высокую эффективность в областях применения, в которых верно предположение о нестационарном характере ряда. Заметим, что данное условие может быть выполнено и в случае финансовых временных рядов общего вида. На практике метод случайного отбора имеет превосходство над ранее использовавшимися для решения рассматриваемой задачи алгоритмами глубокого обучения, в частности над теми из них, которые считаются наиболее производительными: рекуррентными LSTM-нейросетями.

## Литература

1. Hilliard N., Phillips L., Howland S., Yankov A., Corley C. D., Hodas N. O. Few-shot learning with metric-agnostic conditional embeddings. arXiv:1802.04376v1, 2018.
2. Fei-Fei L., Fergus R., Perona P. One-shot learning of object categories // IEEE T. Pattern Anal., 2006. Vol. 28. Iss. 4. P. 594–611. doi: 10.1109/TPAMI.2006.79.
3. Lake B. M., Salakhutdinov R., Tenenbaum J. B. Human-level concept learning through probabilistic program induction // Science, 2015. Vol. 350. P. 1332–1338. doi: 10.1126/science.aab3050.

Поступила в редакцию 08.05.19

---

---

## RANDOM SAMPLING METHOD FOR CRYPTOCURRENCY MARKET TIME SERIES FORECASTING

*O. E. Sorokoletova<sup>1</sup> and T. V. Zakharova<sup>1,2</sup>*

<sup>1</sup>Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, 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:** This paper applies Random Sampling Method (RSM) to classification task for cryptocurrencies time series, which are not-stationary Long Short Term Memory (LSTM) networks have been demonstrated to be particularly useful for learning sequences containing longer term patterns of unknown length, such as at this task. But RSM represents another deep learning algorithm with more

flexible architecture, built on the basis of LSTM cells and thus having all the advantages of the traditional algorithm, but more resistant to the class imbalance problem. The main distinguishing feature of RSM is the use of metric learning and special sampling scheme.

**Keywords:** cryptocurrency; time series; forecasting; classification; metric learning; LSTM; random sampling; neural networks; deep learning

**DOI:** 10.14357/08696527190406

## References

1. Hilliard, N., L. Phillips, S. Howland, A. Yankov, C. D. Corley, and N. O. Hodas. 2018. Few-shot learning with metric-agnostic conditional embeddings. *arXiv*. Available at: <https://arxiv.org/abs/1802.04376v1> (accessed February 12, 2018).
2. Fei-Fei, L., R. Fergus, and P. Perona. 2006. One-shot learning of object categories. *IEEE T. Pattern Anal.* 28(4):594–611. doi: 10.1109/TPAMI.2006.79.
3. Lake, B. M., R. Salakhutdinov, and J. B. Tenenbaum. 2015. Human-level concept learning through probabilistic program induction. *Science* 350:1332–1338. doi: 10.1126/science.aab3050.

*Received May 8, 2019*

## Contributors

**Sorokoletova Olga E.** (b. 1996) — student, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; [olgas020697@gmail.com](mailto:olgas020697@gmail.com)

**Zakharova Tatiana V.** (b. 1962) — Candidate of Science (PhD) in physics and mathematics, associate professor, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; 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; [tvzaharova@mail.ru](mailto:tvzaharova@mail.ru)



# АНАЛИЗ ПРОИЗВОДИТЕЛЬНОСТИ СИСТЕМ «НОВОЕ РАДИО» СЕТИ 5G С ПОМОЩЬЮ СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ С ПЕРЕМЕННЫМИ ТРЕБОВАНИЯМИ К РЕСУРСУ\*

*В. А. Бесчастный<sup>1</sup>, Д. Ю. Острикова<sup>2</sup>, Ю. В. Гайдамака<sup>3</sup>*

**Аннотация:** Развертывание перспективных телекоммуникационных систем пятого поколения «новое радио» (*англ.* 5G New Radio, 5G NR), использующих частоты в миллиметровом диапазоне, позволит достичь беспрецедентного увеличения пропускной способности беспроводных сетей. С помощью таких систем станет возможным внедрение специальных ресурсоемких приложений с крайне высокими требованиями к инфраструктуре сети, которые являются ключевым фактором развития систем 5G. В работе предложен метод анализа 5G NR систем на основе мультисервисной системы массового обслуживания (СМО) с ограниченным ресурсом и заявками, требования к ресурсу которых могут меняться в процессе их обслуживания. Моделирование переменных требований к ресурсу позволяет получать более точную оценку показателей качества функционирования системы по сравнению с традиционными мультисервисными моделями с постоянными требованиями к ресурсу. Для рассмотренной модели предложен итерационный алгоритм расчета стационарного распределения вероятностей с заданной точностью.

**Ключевые слова:** новое радио; миллиметровый диапазон; прямая видимость; вероятность блокировки; СМО; ограниченный ресурс; переменные требования

**DOI:** 10.14357/08696527190407

## 1 Введение

С появлением семейства технологий «новое радио» [1] перед разработчиками телекоммуникационных систем встало множество новых проблем, с которыми не приходилось сталкиваться в прошлом [2–4]. Среди таких проблем можно выделить значительно большие потери при распространении сигнала по сравнению с микроволновой связью [5], блокировки сигнала небольшими динамическими объектами [6], необходимость эффективных интеллектуальных механизмов

---

\* Публикация подготовлена при поддержке программы РУДН «5-100» и при финансовой поддержке РФФИ (проекты 18-07-00576 и 18-37-00380).

<sup>1</sup> Российский университет дружбы народов, [beschastnyy-va@rudn.ru](mailto:beschastnyy-va@rudn.ru)

<sup>2</sup> Российский университет дружбы народов, [ostrikova-dyu@rudn.ru](mailto:ostrikova-dyu@rudn.ru)

<sup>3</sup> Российский университет дружбы народов; Федеральный исследовательский центр «Информатика и управление» Российской академии наук, [gaydamaka-yuv@rudn.ru](mailto:gaydamaka-yuv@rudn.ru)

управления лучом и др. [2, 3, 7]. В реальных условиях эксплуатации подобных систем основной сложностью для обеспечения непрерывного обслуживания абонентов являются подвижные препятствия, такие как люди и транспортные средства, временно блокирующие распространение радиосигнала и называемые «блокираторами».

В зависимости от среды распространения сигнала, а также расстояния от базовой станции (Base Station, BS) пользовательское оборудование в состоянии блокировки прямой видимости до базовой станции может либо выпасть из зоны обслуживания, либо понизить схему модуляции и кодирования для удовлетворения требуемых показателей обслуживания. В качестве решения для случая с выходом оборудования из зоны обслуживания консорциумом 3GPP недавно был предложен механизм так называемого «множественного соединения» [8], при котором одновременно устанавливаются несколько активных каналов между пользовательским устройством и расположенными поблизости базовыми станциями, и в случае появления блокиратора сессия меняет опорную точку доступа, т. е. соединение переключается на один из активных каналов без препятствий прямой видимости (Line of Sight, LoS). В случае понижения схемы модуляции и кодирования активная сессия требует большего объема ресурса [9], и при его нехватке может либо снизиться скорость передачи данных для соответствующего соединения, либо сессия может быть сброшена.

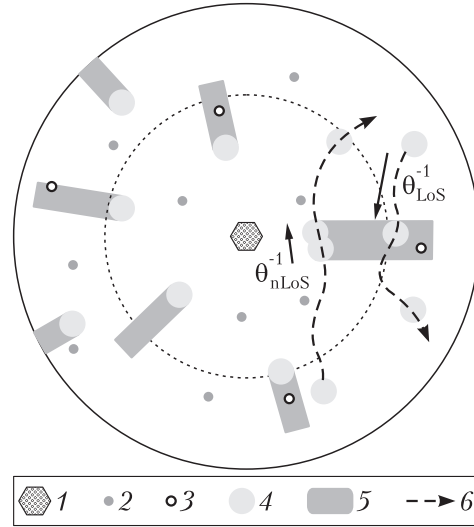
## 2 Описание системной модели

В работе рассматривается модель функционирования базовой станции 5G NR, предложенная в [10]. Базовая станция обеспечивает покрытие внутри окружности радиуса  $d$ , максимальное значение которого  $d_{\text{LoS}}^E$  рассчитывается, согласно [3], с помощью модели распространения сигнала в миллиметровом диапазоне при отсутствии блокираторов и набора схем модуляции и кодирования [1]. Предполагается, что пользователи равномерно распределены по зоне покрытия в соответствии с пуассоновским точечным процессом с параметром  $\rho$ . Таким образом, процесс поступления запросов пользователей на установление соединений является пуассоновским с интенсивностью  $\lambda = \Lambda \rho \pi d^2$ , где  $\Lambda$  — параметр экспоненциально распределенных интервалов между двумя последовательными запросами от одного пользователя, а  $\rho$  — плотность пользователей в зоне покрытия (рис. 1).

Для обслуживания запросов пользователей базовая станция выделяет радиоресурс, при этом объем ресурса зависит от местоположения пользователя и является случайной величиной [11]. Предположим, что с точки зрения требований к объему ресурса область покрытия разбита на внутреннюю и внешнюю зоны — ближний к точке доступа круг, в центре которого расположена базовая станция, и оставшуюся зону вплоть до границ зоны покрытия, которая имеет форму кольца. Радиус круга определяется моделью распространения сигнала как максимальное расстояние  $d_{\text{nLoS}}^E$ , на котором можно установить соединение

в условиях блокировки линии прямой видимости [3]. Для кольца внутренний радиус равен радиусу круга, а внешний радиус определяется как  $d = \min(d_{\text{ISD}} - d_{\text{nLoS}}^E, d_{\text{LoS}}^E)$ , где  $d_{\text{ISD}}$  — расстояние от целевой до ближайшей соседней NR BS. Также предположим, что требования к объему ресурса для обслуживания запросов от пользователей, находящихся в круге, ниже соответствующих требований от пользователей в кольце. Это предположение соответствует простейшей модели распространения сигнала [7], согласно которой объем ресурса обратно пропорционален расстоянию между оборудованием пользователя и точкой доступа.

Модель учитывает возможность блокировки прямой видимости подвижными препятствиями для пользователей в круге и в кольце. При появлении блокиратора на линии прямой видимости между точкой доступа и пользовательским устройством в круге соединение не прерывается, но увеличивается объем ресурса, выделенный точкой доступа для его поддержания. Когда ситуация блокирования прямой видимости (non Line of Sight, nLoS) возникает для устройства в кольце, качество соединения в радиоканале падает настолько, что для его поддержания необходимо сменить опорную точку доступа. В следующем разделе построена СМО фазового типа для моделирования обслуживания пользователей базовой станцией 5G NR с учетом появления и исчезновения блокираторов прямой видимости.



**Рис. 1** Модель движения блокираторов: 1 — NR BS; 2 и 3 — пользовательское оборудование с прямой и без прямой видимости соответственно; 4 — блокиратор; 5 — зона без прямой видимости; 6 — траектория движения

### 3 Анализ показателей эффективности с помощью системы массового обслуживания с переменным требованием к ресурсу

Рассмотрим ресурсную СМО с  $N$  приборами, ресурсом конечного объема  $R$  и пуассоновским поступающим потоком интенсивности  $\lambda$ . Предположим, что длительности обслуживания заявок не зависят от поступающего потока, независимы в совокупности и имеют функции распределения фазового типа

$$B(x) = 1 - \mathbf{a} \exp(x\mathbf{M})\mathbf{u}.$$

Здесь  $\mathbf{a}$  — вектор-строка длины  $n$  распределения вероятностей начальной фазы;  $\mathbf{u}$  — вектор-столбец длины  $n$  из единиц;  $\mathbf{M} = [\mu_{ij}]$  — невырожденная квадратная матрица порядка  $n$  интенсивностей переходов между фазами обслуживания;  $n$  — число фаз обслуживания.

Изменение требований заявок к ресурсам в течение времени обслуживания моделируется с помощью смены фазы обслуживания. На  $j$ -й фазе обслуживания заявке требуется  $r_j$  единиц ресурса. Если при поступлении или в момент смены фазы обслуживания выяснится, что объем требуемого заявке ресурса превышает доступный объем незанятого ресурса, заявка теряется.

Состояние системы в момент  $t$  можно описать однородным марковским процессом  $\mathbf{X}(t) = (X_1(t), \dots, X_n(t))$ , где  $X_j(t)$  — число заявок, обслуживаемых на фазе  $j$ . Процесс  $\mathbf{X}(T)$  имеет конечное пространство состояний

$$\mathbf{X}(R) = \left\{ (k_1, \dots, k_n) \mid k_j \geq 0, j = 1, \dots, n, \sum_{j=1}^n k_j \leq N, \sum_{j=1}^n k_j r_j \leq R \right\}.$$

Обозначим через  $\mu_{i0} = -\sum_{j=1}^n \mu_{ij}$  интенсивность завершения обслуживания заявки после фазы  $i$ . Система уравнений равновесия для процесса  $\mathbf{X}(t)$  имеет следующий вид:

$$\begin{aligned} & \left( \lambda \mathbf{I}(\mathbf{k}\mathbf{u} < N) \sum_{i=1}^n \mathbf{I}(\mathbf{k}\mathbf{r} + r_i \leq R) a_i + \mathbf{k}\boldsymbol{\mu} \right) p(\mathbf{k}) = \\ & = \lambda \sum_{i=1}^n \mathbf{I}(k_i > 0) p(\mathbf{k} - \mathbf{e}_i) a_i + \mathbf{I}(\mathbf{k}\mathbf{u} < N) \sum_{i=1}^n p(\mathbf{k} + \mathbf{e}_i) (k_i + 1) \mu_{i0} + \\ & + \mathbf{I}(\mathbf{k}\mathbf{u} < N) \sum_{i=1}^n \sum_{\substack{j=1, \\ j \neq i}}^n \mathbf{I}(\mathbf{k}\mathbf{r} + r_j > R) p(\mathbf{k} + \mathbf{e}_i) (k_i + 1) \mu_{ij} + \\ & + \mathbf{I}(\mathbf{k}\mathbf{u} < N) \sum_{i=1}^n p(\mathbf{k} + \mathbf{e}_i) (k_i + 1) \mu_{i0} + \\ & + \mathbf{I}(\mathbf{k}\mathbf{u} < N) \sum_{i=1}^n \sum_{\substack{j=1, \\ j \neq i}}^n \mathbf{I}(\mathbf{k}\mathbf{r} + r_j > R) p(\mathbf{k} + \mathbf{e}_i) (k_i + 1) \mu_{ij}, \quad \mathbf{k} \in \mathbf{X}(R), \quad (1) \end{aligned}$$

где  $\mathbf{I}(S)$  — индикатор утверждения  $S$ ;  $\mathbf{r} = (r_1, \dots, r_n)$  и  $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)$  — векторы-столбцы;  $\mathbf{e}_i$  — вектор-строка, все компоненты которого равны 0, кроме стоящей на  $i$ -м месте единицы. Эту систему уравнений можно решить численно

методом последовательных приближений. В качестве начального приближения возьмем равномерное распределение

$$p(\mathbf{k}) = \frac{1}{|\mathcal{X}(R)|}, \quad \mathbf{k} \in \mathcal{X}(R).$$

Затем, до тех пор пока не достигнута требуемая точность, вычисляем очередные приближения, при этом каждая итерация состоит из трех шагов.

1.  $G := 0$ .
2.  $p(\mathbf{k}) :=$

$$\begin{aligned} := & \frac{1}{\lambda \mathbf{I}(\mathbf{k}\mathbf{u} < N) \sum_{i=1}^n \mathbf{I}(\mathbf{k}\mathbf{r} + r_i \leq R) a_i + \mathbf{k}\boldsymbol{\mu}} \left[ \lambda \sum_{i=1}^n \mathbf{I}(k_i > 0) p(\mathbf{k} - \mathbf{e}_i) a_i + \right. \\ & + \mathbf{I}(\mathbf{k}\mathbf{u} < N) \sum_{i=1}^n p(\mathbf{k} + \mathbf{e}_i) (k_i + 1) \mu_{i0} + \\ & + \mathbf{I}(\mathbf{k}\mathbf{u} < N) \sum_{i=1}^n \sum_{\substack{j=1, \\ j \neq i}}^n \mathbf{I}(\mathbf{k}\mathbf{r} + r_j > R) p(\mathbf{k} + \mathbf{e}_i) (k_i + 1) \mu_{ij} + \\ & \left. + \sum_{i=1}^n \sum_{\substack{j=1, \\ j \neq i}}^n \mathbf{I}(k_j > 0) \mathbf{I}(\mathbf{k}\mathbf{r} + r_j \leq R) p(\mathbf{k} + \mathbf{e}_i - \mathbf{e}_j) (k_i + 1) \mu_{ij} \right], \quad (2) \end{aligned}$$

$$G := G + p(\mathbf{k}), \quad \mathbf{k} \in \mathcal{X}(R).$$

3.  $p(\mathbf{k} := (1/G)p(\mathbf{k}), \mathbf{k} \in \mathcal{X}(R)$ .

Система уравнений равновесия (1) и алгоритм (2) дают метод расчета стационарного распределения, который применен в следующем разделе при анализе показателей качества функционирования системы 5G NR.

## 4 Численный анализ

В данном разделе проводится численный анализ функционирования базовой станции системы 5G NR с точки зрения одной из ключевых характеристик качества обслуживания — вероятности сброса сессий, которая для рассматриваемой модели вычисляется как  $B = \mathbf{a}\mathbf{b}$ , где  $\mathbf{b} = (b_1, \dots, b_n)$ , а вероятность блокировки в  $i$ -й фазе вычисляется следующим образом:

$$b_i = \sum_{\substack{\mathbf{k} \in \mathcal{X}(R), \\ \mathbf{k}\mathbf{r} + r_i > R}} p(\mathbf{k} + \sum_{\substack{j \in \{1, \dots, n\} \setminus i, \\ \mathbf{k}\mathbf{r} - r_j + r_i > R}} p(\mathbf{k})), \quad 1 \leq i \leq n.$$

Входные параметры численного анализа

Обозначение	Описание	Значения
$d$	Радиус внешней зоны	400 м
$d_{\text{ISD}}$	Расстояние между соседними BS	600 м
$R$	Число ресурсных блоков в выделенной полосе частот	667
$r_1$	Требования сессий в кольце при наличии прямой видимости	31 RB*
$r_2$	Требования сессий в круге при наличии прямой видимости	16 RB
$r_3$	Требования сессий в круге без прямой видимости	158 RB
$\lambda$	Интенсивность поступления запросов от одного устройства	$10^{-3} \text{ с}^{-1}$
$\mu$	Интенсивность обслуживания сессий	$1/30 \text{ с}^{-1}$
$\alpha = \gamma$	Параметр распределения периодов между последовательными блокировками прямой видимости	$0,033 \text{ с}^{-1}$
$\beta$	Параметр распределения длительностей блокировок	$0,34 \text{ с}^{-1}$

\*RB — resource block (ресурсный блок).

Помимо вероятности сброса также исследуется доля занятого ресурса базовой станции NR:

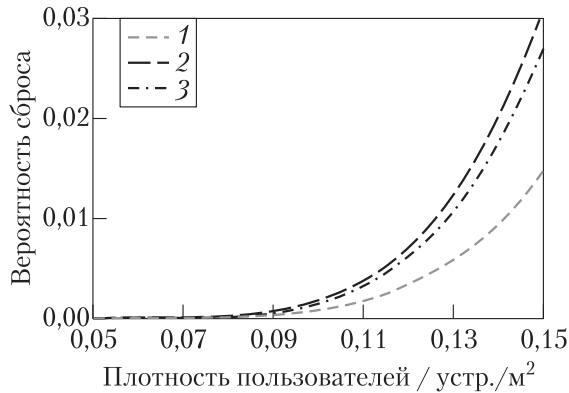
$$\text{UTIL} = \sum_{k \in X(R)} \mathbf{k} r p(\mathbf{k}).$$

Рассматриваемая система функционирует на рабочей частоте 28 ГГц с доступной полосой частот в 1 ГГц и коэффициентом затухания, равным 2,1. Мощность передающей антенны составляет 0,2 Вт, а ее — усиление 2,58 дБ. Предполагаем, что все пользователи получают один тип услуг с фиксированной требуемой скоростью передачи данных 50 Мбит/с (соответствует передаче видео в высоком качестве). Эти параметры определяют требования для установления пользовательских сессий, а также величину доступного ресурса. В роли блокираторов выступают цилиндры радиуса 0,4 м с высотой 1,7 м, имитирующие человеческое тело. Частота наступления и длительность периодов отсутствия прямой видимости зависят от модели движения блокираторов. Для численного анализа выбрано случайное блуждание со средней скоростью 1,1 м/с.

Метод получения исходных данных таблицы для заданных значений параметров системы 5G NR изложен в [3].

Построенная для рассмотренной системы 5G NR модель СМО имеет три фазы обслуживания ( $n = 3$ ): первая фаза соответствует незаблокированным сессиям из кольца, вторая фаза — незаблокированным сессиям из круга, а третья фаза — заблокированным сессиям из круга. Матрица переходов  $\mathbf{M} = [\mu_{ij}]$  между фазами обслуживания имеет следующий вид:

$$\mathbf{M} = \begin{bmatrix} -(\gamma + \mu) & 0 & 0 \\ 0 & -(\alpha + \mu) & \alpha \\ 0 & \beta & -(\beta + \mu) \end{bmatrix},$$



**Рис. 2** Вероятность сброса сессий в зависимости от положения устройства: 1 —  $B_I$ ; 2 —  $B_O$ ; 3 —  $B$

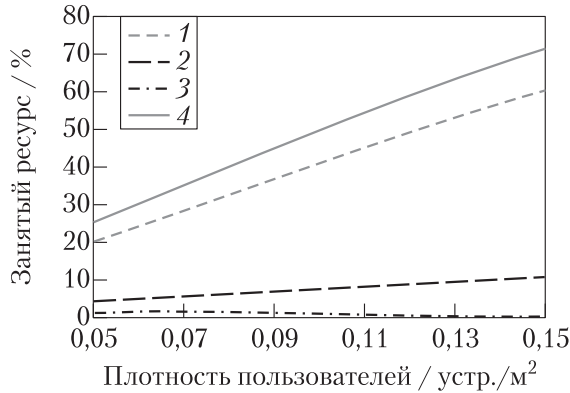
где  $\alpha^{-1}$  и  $\gamma^{-1}$  — средние интервалы между двумя последовательными блокировками прямой видимости в круге и кольце соответственно;  $\beta^{-1}$  — средняя длительность блокировки. В этом случае функция распределения

$$B(x) = 1 - \mathbf{a} \exp(x\mathbf{M})\mathbf{u} = 1 - a_1 e^{-x(\mu+\gamma)} - (1 - a_1) e^{-\mu x}$$

представляет собой гиперэкспоненциальное распределение.

На рис. 2 изображен график зависимости вероятности сброса сессии от положения устройства, где  $B_I$  соответствует вероятности сброса в круге,  $B_O$  — в кольце, а  $B$  — вероятности сброса вне зависимости от положения устройства. Вероятность сброса сессий из внешней зоны (кольцо) значительно выше, чем у сессий из внутренней зоны (круг), так как их общее число и требования к ресурсу выше. Это свидетельствует о необходимости использования принципа множественных соединений [8], который позволил бы снизить вероятность сброса в кольце за счет резервирования на соседних точках доступа дополнительного ресурса для заблокированных сессий, являющихся наиболее ресурсоемкими. Вследствие более жестких требований к ресурсу именно эти сессии в большей степени подвержены отказам в обслуживании.

На рис. 3 изображена зависимость загрузки базовой станции от типа и положения устройств, где  $UTIL$  — средняя доля ресурса, используемая всеми типами пользователей,  $UTIL_{LoS}^I$  — пользователями из круга при наличии прямой видимости до NR BS,  $UTIL_{nLoS}^I$  — пользователями из круга без прямой видимости до NR BS,  $UTIL^O$  — пользователями из кольца. Следует отметить, что на сессии из внешней зоны приходится более 80% от общего используемого ресурса, что подтверждает необходимость распределения нагрузки в соответствующей зоне обслуживания между соседними базовыми станциями. В то же время видно, что



**Рис. 3** Средняя доля занятого ресурса в зависимости от положения устройства и условий прямой видимости: 1 —  $UTIL^O$ ; 2 —  $UTIL_{LoS}^I$ ; 3 —  $UTIL_{nLoS}^I$ ; 4 —  $UTIL$

с ростом плотности устройств доля ресурса, занимаемая незаблокированными сессиями, уменьшается, что объясняется снижением их доли в структуре трафика. Снижение происходит из-за роста вероятности блокировки прямой видимости вследствие увеличения числа потенциальных блокираторов.

Из экспоненциального характера роста вероятности сброса сессий даже при не полностью занятом ресурсе можно сделать вывод о том, что для подобных систем необходимо исследовать механизмы резервирования ресурсов.

## 5 Заключение

В работе рассмотрен процесс выделения радиоресурсов базовой станцией 5G NR с учетом временных блокировок прямой видимости подвижными препятствиями. Предложена математическая модель в виде многоканальной СМО фазового типа с ресурсными ограничениями и переменными требованиями к ресурсу. Также разработан метод расчета стационарных вероятностей для оценки ключевых показателей системы, которые в дальнейшем могут быть использованы для нахождения оптимальных сценариев развертывания мобильных сетей на базе технологии 5G NR.

В качестве задачи дальнейших исследований планируется расширить модель добавлением потока заявок с перекрывающимися требованиями к ресурсу [12], что позволит моделировать процесс обслуживания многоадресных соединений [10, 12, 13], широко применяемых в современных проводных и беспроводных сетях и позволяющих значительно улучшить эффективность использования ресурсов.

Авторы благодарят профессора Исследовательского института инноваций г. Хельсинки (Финляндия) В. А. Наумова за советы по разработке алгоритма для случая СМО фазового типа.



## Литература

1. 5G; NR; Physical channels and modulation (3GPP TS 38.211 version 15.2.0 Release 15). ETSI TS 138 211 V15.2.0 (2018-07).
2. *Petrov V., Solomitckii D., Samuylov A., Lema M.A., Gapeyenko M., Moltchanov D., Andreev S., Naumov V., Samouylov K., Dohler M.* Dynamic multi-connectivity performance in ultra-dense urban mmWave deployments // *IEEE J. Sel. Area. Comm.*, 2017. Vol. 35. Iss. 9. P. 2038–2055. doi: 10.1109/jsac.2017.2720482.
3. *Kovalchukov R., Moltchanov D., Begishev V., Samuylov A., Andreev S., Koucheryavy Ye., Samouylov K.* Improved session continuity in 5G NR with joint use of multi-connectivity and guard bandwidth // *IEEE Global Communications Conference*, 2018. P. 1–7. doi: 10.1109/glocom.2018.8647608.
4. *Чухно О. В., Чухно Н. В., Гайдамака Ю. В., Шоргин С. Я.* Аппроксимация коэффициента усиления направленности антенны для анализа «направленной глухоты» в трехмерном пространстве // *Системы и средства информатики*, 2019. Т. 29. № 3. С. 39–51.
5. *Moltchanov D., Samuylov A., Petrov V., Gapeyenko M., Himayat N., Andreev S., Koucheryavy Ye.* Improving session continuity with bandwidth reservation in mmWave communications // *IEEE Wirel. Commun. Lett.*, 2019. Vol. 8. Iss. 1. P. 105–108. doi: 10.1109/lwc.2018.2859988.
6. *Gapeyenko M., Samuylov A., Gerasimenko M., Moltchanov D., Singh S.A., Riza M., Aryafar E., Himayat N., Andreev S., Koucheryavy Ye.* On the temporal effects of mobile blockers in urban millimeter-wave cellular scenarios // *IEEE T. Veh. Technol.*, 2017. Vol. 66. Iss. 11. P. 10124–10138. doi: 10.1109/tvt.2017.2754543.
7. Qualcomm: mmWave 5G NR prototype demo video. <https://www.qualcomm.com/videos/mmwave-5g-nr-prototype-demo-video>.
8. 3GPP Technical Specification 37.340; NR; Multi-connectivity; Overall description (Release 15), Jan 2018.
9. 3GPP 5G. Study on channel model for frequencies from 0.5 to 100 GHz (Release 14). Technical Report 38.901, May 2017.
10. *Samuylov A., Moltchanov D., Krupko A., Kovalchukov R., Moskaleva F., Gaidamaka Yu.* Performance analysis of mixture of unicast and multicast sessions in 5G NR systems // *Emerging Technologies for Connected Society: 10th Congress (International) on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT) Proceedings*. — IEEE, 2018. P. 287–293. doi: 10.1109/ICUMT.2018.8631230.
11. *Наумов В. А., Самуйлов А. К.* Модель выделения ресурсов беспроводной сети объемами случайной величины // *Вестник РУДН. Сер. Математика. Информатика. Физика*, 2015. Т. 2. С. 38–45.
12. *Samouylov K., Gaidamaka Y.* Analysis of loss systems with overlapping resource requirements // *Stat. Pap.*, 2018. Vol. 59. Iss. 4. P. 1463–1470. doi: 10.1007/s00362-018-1041-x.
13. *Biason A., Zorzi M.* Multicast via point to multipoint transmissions in directional 5G mmWave communications // *IEEE Commun. Mag.*, 2019. Vol. 57. Iss. 2. P. 88–94. doi: 10.1109/mcom.2019.1700679.

Поступила в редакцию 02.09.19

## 5G NEW RADIO BASE STATION PERFORMANCE ANALYSIS USING LIMITED RESOURCE LOSS SYSTEM WITH VARIABLE DEMANDS

V. A. Beschastnyi<sup>1</sup>, D. Yu. Ostrikova<sup>1</sup>, and Yu. V. Gaidamaka<sup>1,2</sup>

<sup>1</sup>Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation

<sup>2</sup>Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** The deployment of prospective 5G New Radio telecommunication systems (5G NR) that utilizes millimeter-wave frequencies will allow achieving an unprecedented increase in capacity of mobile wireless networks. This type of communications will provide the possibility to introduce new special resource-consuming and mission critical applications with extremely high requirements for network infrastructure, which are the key driver in the development of 5G systems. The present authors propose an analytical method for the analysis of 5G NR systems based on a multiservice loss system with limited resources and customers that demand varying amount of resource within their service time. For the considered model, the authors propose an approximate method that allows to calculate the stationary probability distribution and other network performance metrics to the specified accuracy. The findings are illustrated with a numerical example.

**Keywords:** New Radio; mmWave; Line of Sight, LoS; loss probability; loss system; limited resource; variable demand

**DOI:** 10.14357/08696527190407

### Acknowledgments

The publication has been prepared with the support of the "RUDN University Program 5-100" and funded by the Russian Foundation for Basic Research according to the research projects Nos. 18-07-00576 and 18-37-00380.

### References

1. 3GPP TR 38.211; NR; Physical channels and modulation (Release 15), Dec 2017.
2. Petrov, V., D. Solomitckii., A. Samuylov, M. A. Lema, M. Gapeyenko, D. Moltchanov, S. Andreev., V. Naumov, K. Samouylov, and M. Dohler. 2017. Dynamic multi-connectivity performance in ultra-dense urban mmWave deployments. *IEEE J. Sel. Area. Comm.* 35(9):2038–2055.
3. Kovalchukov, R., D. Moltchanov, V. Begishev, A. Samuylov, S. Andreev, Y. Koucheryavy, and K. Samouylov. 2018. Improved session continuity in 5G NR with joint use of multi-connectivity and guard bandwidth. *2018 IEEE Global Communications Conference* 1–7.
4. Chukhno, O. V., N. V. Chukhno, Yu. V. Gaidamaka, and S. Ya. Shorgin. 2019. Ap-proksimatsiya koeffitsienta usileniya napravlennosti anteny dlya analiza "napravlennoy glukhoty" v trekhmernom prostranstve [Approximation of antenna directivity gain for

- directional deafness analysis in three-dimensional space]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(3):39–51.
5. Moltchanov, D., A. Samuylov, V. Petrov, M. Gapeyenko, N. Himayat, S. Andreev, and Ye. Koucheryavy. 2019. Improving session continuity with bandwidth reservation in mmWave communications. *IEEE Wirel. Commun. Lett.* 8(1):105–108.
  6. Gapeyenko, M., A. Samuylov, M. Gerasimenko, D. Moltchanov, S. A. Singh, M. Riza, E. Aryafar, N. Himayat, S. Andreev, and Ye. Koucheryavy. 2017. On the temporal effects of mobile blockers in urban millimeter-wave cellular scenarios. *IEEE T. Veh. Technol.* 66(11):10124–10138.
  7. Qualcomm, mmWave 5G NR prototype demo video. Available at <https://www.qualcomm.com/videos/mmwave-5g-nr-prototype-demo-video> (accessed July 17, 2019).
  8. 3GPP. NR: Multi-connectivity; Overall description (Release 15). 3GPP TS 37.34 V15.2.0. 2018.
  9. 3GPP. 5G. May 2017. Study on channel model for frequencies from 0.5 to 100 GHz (Release 14). Technical Report 38.901.
  10. Samuylov, A., D. Moltchanov, A. Krupko, R. Kovalchukov, F. Moskaleva, and Yu. Gaidamaka. 2018. Performance analysis of mixture of unicast and multicast sessions in 5G NR systems. *10th Congress (International) on Ultra Modern Telecommunications and Control Systems and Workshops*. IEEE. 1–7. doi: 10.1109/ICUMT.2018.8631230.
  11. Naumov, V. A., and A. K. Samuylov. 2015. Model' vydeleniya resursov besprovodnoy seti ob'emami sluchaynoy velichiny [Queuing system with resource allocation of the random volume]. *RUDN J. Mathematics Information Sciences Physics* 2:38–45.
  12. Samouylov, K., and Y. Gaidamaka. Y. 2018. Analysis of loss systems with overlapping resource requirements. *Stat. Pap.* 59(4):1463–1470. doi: 10.1007/s00362-018-1041-x.
  13. Biason, A., and M. Zorzi. 2019. Multicast via point to multipoint transmissions in directional 5G mmWave communications. *IEEE Commun. Mag.* 57(2):88–94. doi: 10.1109/mcom.2019.1700679.

*Received September 2, 2019*

## Contributors

**Beschastnyi Vitalii A.** (b. 1992) — PhD student, Department of Applied Probability and Informatics, Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; [beschastnyy-va@rudn.ru](mailto:beschastnyy-va@rudn.ru)

**Ostrikova Daria Yu.** (b. 1988) — Candidate of Science (PhD) in physics and mathematics, senior lecturer, Department of Applied Probability and Informatics, Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; [ostrikova-dyu@rudn.ru](mailto:ostrikova-dyu@rudn.ru)

**Gaidamaka Yuliya V.** (b. 1971) — Doctor of Science in physics and mathematics, professor, Department of Applied Probability and Informatics, Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; senior scientist, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; [gaydamaka-yuv@rudn.ru](mailto:gaydamaka-yuv@rudn.ru)

## МЕТОД ИЗВЛЕЧЕНИЯ ЭКСПЕРТНЫХ ОЦЕНОК ИЗ СОЦИАЛЬНЫХ СЕТЕЙ ПРИ ПРОВЕДЕНИИ ПРОЦЕССА ГРУППОВОГО ПРИНЯТИЯ РЕШЕНИЙ\*

*О. В. Чухно<sup>1</sup>, Н. В. Чухно<sup>2</sup>, Ю. В. Гайдамака<sup>3</sup>, К. Е. Самуйлов<sup>4</sup>*

**Аннотация:** Социальные сети представляются наиболее предпочтительной средой с точки зрения взаимодействия пользователей, обмена информацией и общения, поэтому их можно использовать в процессе группового принятия решений. Однако в существующих методах группового принятия решений эксперты должны следовать конкретным предложенным структурам при выставлении своих оценок, что неестественно для онлайн-сервисов, в которых пользователи делятся мнениями в свободной форме. Для решения данной проблемы предлагается модель процесса группового принятия решений, в которой эксперты могут чувствовать себя свободнее при оценивании пар альтернатив или высказывании своих мнений. Данный метод открывает возможность для автоматической работы в социальных сетях. Также предложен обзор и на его основании формальное описание алгоритма группового принятия решений в социальной сети. Представлен метод оценивания альтернатив, в котором эксперты выставляют свои оценки в ходе обсуждения проблемы с использованием текстов, сообщений или публикаций на выбранных ими платформах. Кроме того, предложены способы определения достигнутого уровня консенсуса среди экспертов данного процесса.

**Ключевые слова:** групповое принятие решений; социальные сети; консенсусные меры; анализ тональности текстов

**DOI:** 10.14357/08696527190408

### 1 Введение

Благодаря технологии интернета у людей появилась возможность общаться, делиться своими мнениями и чувствами, обсуждать какие-либо проблемы и темы, находясь в любой точке мира. Социальные сети сейчас довольно популярны и представляют важную среду для проведения процессов группового принятия решений (*англ.* Group Decision Making, GDM) [1, 2]. GDM — это процесс,

---

\* Публикация подготовлена при поддержке Программы РУДН «5-100» и при финансовой поддержке РФФИ (проекты 17-07-00845 и 18-07-00576).

<sup>1</sup> Российский университет дружбы народов, olgachukhno95@gmail.com

<sup>2</sup> Российский университет дружбы народов, nvchukhno@gmail.com

<sup>3</sup> Российский университет дружбы народов; Федеральный исследовательский центр «Информатика и управление» Российской академии наук, gaydamaka-yuv@rudn.ru

<sup>4</sup> Российский университет дружбы народов; Федеральный исследовательский центр «Информатика и управление» Российской академии наук, samouylov-ke@rudn.ru

состоящий в выборе наилучшей альтернативы или множества альтернатив из всех возможных с учетом мнений, выраженных группой лиц [3–7].

Основная цель статьи — представить новый метод, способный работать с уже существующими социальными сетями. С его помощью эксперты смогут использовать любую онлайн-платформу для обсуждения какой-либо проблемы, а система будет выполнять необходимые вычисления без взаимодействия с пользователем. Предложенная модель использует анализ тональности текстов (*англ.* Sentiment analysis) для извлечения значений предпочтений из опубликованных в сети сообщений [8, 9]. Благодаря свободе, которую получают эксперты, они смогут проявить себя лучше, а оценки станут более объективными.

В новом методе также представлены два способа определения степени согласия между экспертами в процессе группового принятия решений. Первый использует матрицы предпочтений экспертов, полученные при анализе тональности текстов. Второй опирается на отношения участников процесса друг к другу.

Остальная часть статьи организована следующим образом. В разд. 2 предложен обзор, необходимый для понимания разработанного метода. Третий раздел описывает метод группового принятия решений, в том числе алгоритм извлечения оценок экспертов из социальных сетей, а также способы измерения достигнутого консенсуса среди экспертов. В заключении подведены итоги и поставлены задачи дальнейших исследований.

## 2 Предварительные замечания

Социальные сети — это платформы, которые позволяют пользователям обмениваться информацией, используя сообщения, посты, публикации и т. д. Важная особенность таких сервисов — использование хештегов (*англ.* Hashtags) — наборов ключевых слов, служащих для описания определенной темы. Кроме того, в сети всегда можно связать пользователя с опубликованным им сообщением или постом, а также в процессе GDM определить экспертов, которые согласны или не согласны с другими участниками процесса. Для обращения к конкретному пользователю онлайн-платформы используют упоминания, которые создаются как @username, где username — имя пользователя. Отличительной особенностью проведения GDM в социальных сетях является то, что эксперты могут участвовать в данном процессе независимо от местоположения, времени и друг от друга.

Для определения настроения человека, будем применять анализ тональности текстов, исследуя сообщения экспертов.

Традиционно процесс группового принятия решений определяется следующим образом: множество экспертов  $E = \{e_1, \dots, e_K\}$  должны ранжировать множество альтернатив  $X = \{x_1, \dots, x_M\}$  по возрастанию степени предпочтения, при этом  $|E| = K < \infty$  и  $|X| = M < \infty$ . Для этой цели эксперты предоставляют свои предпочтения  $p_{ij}(k)$  относительно каждой пары альтерна-

тив  $x_i$  и  $x_j$ , которые выражаются матрицами индивидуальных предпочтений  $\mathbf{P} = \{\mathbf{P}(1), \mathbf{P}(2), \dots, \mathbf{P}(K)\}$ .

Метод группового принятия решений включает следующие шаги.

1. Предоставление предпочтений: эксперты решают, какие альтернативы являются наиболее подходящими для них, и делятся с системой своим мнением.
2. Вычисление матрицы коллективных предпочтений: после того как все эксперты предоставили оценки по всем парам альтернатив, предпочтения объединяются в единую матрицу коллективных оценок. Здесь могут применяться различные операторы агрегирования, например OWA (ordered weighted averaging) или IOWA (induced OWA) [10]. Полученный результат содержит общее мнение всех экспертов.
3. Создание рейтинга альтернатив: итоговые результаты ранжирования по всем альтернативам рассчитываются с использованием коллективной матрицы предпочтений. Результатом является выбор лучшей альтернативы или список альтернатив, упорядоченных по возрастанию значимости.
4. Измерение консенсуса. Чем больше экспертов согласно с окончательным решением, тем более надежными будут результаты.

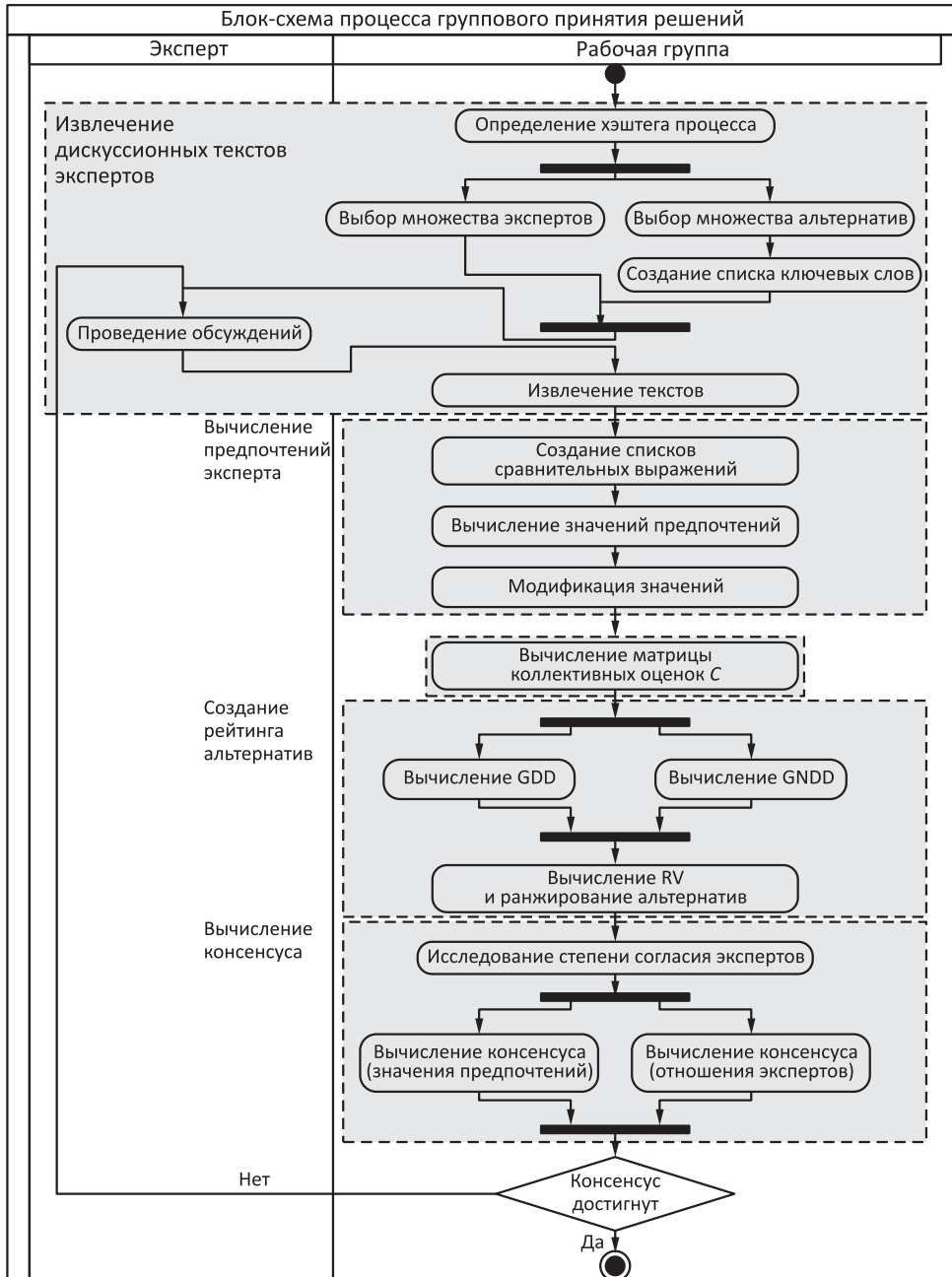
### **3 Метод процесса группового принятия решений с использованием анализа сообщений в социальной сети**

Новый разработанный метод включает следующие этапы (см. рисунок).

#### **I. Извлечение дискуссионных текстов экспертов**

При выполнении данного процесса извлекаются тексты (сообщения), написанные пользователями социальной сети, в том числе выполняются следующие действия.

1. Создание хештега, который будет идентифицировать процесс группового принятия решений, например #GDM. Благодаря этому упрощается задача получения всей необходимой информации по процессу из сети.
2. Выбор учетных записей экспертов, находящихся на целевой онлайн-платформе. Если не ясно, сколько экспертов участвует в процессе, можно определить их число по хештегу.
3. Создание списка ключевых слов, относящихся к альтернативам: до проведения обсуждений необходимо связать множество ключевых слов  $K = \{K_1, \dots, K_M\}$  с конкретной альтернативой. Благодаря этому можно определить, какие альтернативы обсуждают эксперты. Каждое значение  $K_i$  формирует набор слов, которые обычно используются, когда обсуждается альтернатива  $x_i$ .
4. Извлечение дискуссионных текстов в процессе группового принятия решений: для этой цели публикации экспертов, содержащие хештег, выбираются из каждой учетной записи. Выполнение данного шага определяет



Модель группового принятия решений



множество текстов  $A = \{A_1, \dots, A_K\}$ , связанных с множеством экспертов  $E = \{e_1, \dots, e_K\}$ .

## II. Вычисление предпочтений экспертов

После того как тексты экспертов были извлечены, для вычисления значений предпочтений, которые будут использоваться в процессе группового принятия решений, применяется анализ тональности текстов. С этой целью свободный текст, используемый экспертами в их дискуссиях, должен преобразоваться в набор значений предпочтений, которые применяются в процессе GDM для создания окончательного рейтинга альтернатив. Поскольку для решаемой проблемы необходимо определить, есть ли у экспертов положительные, нейтральные или отрицательные чувства при сравнении двух конкретных альтернатив, используются три списка слов: PL, NPL и SL. В PL содержатся слова, определяющие положительные чувства. Если несколько из этих слов используются экспертом, это означает, что он предпочитает первую альтернативу второй. NPL содержит отрицательные слова. Наконец, SL включает слова, которыми часто оперируют при утверждении, что альтернативы имеют одинаковую значимость. Строки PL и NPL имеют сравнительное выражение и его значение — степень предпочтения одной альтернативы над другой. Важно отметить, что SL не нуждается в определении данного значения, так как процесс присваивает среднюю оценку автоматически в случаях, когда две альтернативы считаются одинаковыми.

Предложенный шаг вычисления предпочтений включает:

- (1) извлечение сравнительных предложений: анализируются тексты всех экспертов, участвующих в процессе. Все предложения, содержащие хотя бы одно из выражений из списков слов PL, NPL и SL, извлекаются из сообщений экспертов. В результате этого процесса каждый эксперт  $e_k$  оказывается связан с тремя наборами сравнительных предложений  $\text{spl}_k$ ,  $\text{snpl}_k$  и  $\text{ssl}_k$ ;
- (2) вычисление степени предпочтения  $p_{ij}$  для каждой пары альтернатив. Этот процесс выполняется путем сопоставления значений сравнительного выражения, используемого для сравнения альтернатив  $x_i$  и  $x_j$ . Предполагается, что для заполнения матрицы предпочтений информации всегда достаточно (нет пустых строк). При этом если сравнительное выражение относится к PL и имеет значение 3, то  $p_{ij} = 3$ , в то время как для такого же значения из NPL  $p_{ij} = -3$ . В случае, когда сравнительное выражение принадлежит SL,  $p_{ij} = 0$ ;
- (3) фиксирование интервала значений предпочтений: чтобы избавиться от отрицательных чисел, можно преобразовать значения. Предварительно полученное представление генерирует значения, которые принадлежат интервалу  $[-g, g]$ . Если отрицательные числа нежелательны, можно изменить область и выразить всю информацию с использованием интервала  $[0, 2g]$ :

$$p_{ij[0,2g]}(k) = p_{ij[-g,g]}(k) + g;$$



**Алгоритм вычисления значений предпочтений**

**Data:** Тексты экспертов в социальных сетях  
 $E = \{e_1, \dots, e_K\}$ ,  $X = \{x_1, \dots, x_M\}$ , hashtag,  $K = \{K_1, \dots, K_M\}$ ,  
 $A = \{A_1, \dots, A_K\}$ .  
**procedure**  $A_k \rightarrow \mathbf{P}(k)$   
**for** all  $e_k$ ,  $k = 1$  to  $K$   
**for** all  $A_k$ ,  $k = 1$  to  $K$   
**assign**  $p_{ij}(k)_{i,j \in 1, \dots, M}$   
**if**  $p_{ij}(k) \in \text{PL}$  **do**  
 $p_{ij}(k) = p_{ij}(k)$   
**else if**  $p_{ij}(k) \in \text{NPL}$  **do**  
 $p_{ij}(k) = -p_{ij}(k)$   
**else if**  $p_{ij}(k) \in \text{SL}$  **do**  
 $p_{ij}(k) = 0$   
**and if**  
**and for**  
**for** all  $p_{ij}(k)_{i,j \in 1, \dots, M}$  **do**  
 $p_{ij[0,2g]}(k) = p_{ij[-g,g]}(k) + g$   
 $\tilde{p}_{ij}(k) = \frac{p_{ij[0,2g]}(k)}{2g}$   
**end for**  
**end for**  
Вывести матрицы оценок экспертов  $\mathbf{P}(k) = (\tilde{p}_{ij}(k))_{i,j \in 1, \dots, M}$   
**end procedure**  $A_k \rightarrow \mathbf{P}(k)$

- (4) нормирование оценок: в статье предлагается использовать оценки, выраженные на интервале  $[0, 1]$ , для этого необходимо воспользоваться формулой

$$\tilde{p}_{ij}(k) = \frac{p_{ij[0,2g]}(k)}{2g}.$$

После того как будут рассчитаны матрицы предпочтений для всех экспертов, можно определить рейтинг альтернатив и вычислить значения консенсуса между всеми экспертами, участвующими в процессе группового принятия решений.

**III. Расчет матрицы коллективных предпочтений**

Значения предпочтений всех экспертов объединяются в единую коллективную матрицу, которая представляет информацию об общем мнении участников процесса GDM. Для этой цели, как и в традиционной модели, может применяться любой оператор агрегирования. Будем использовать оператор среднее. Тогда элементы матрицы коллективных предпочтений будут вычисляться следующим образом:

$$c_{ij} = \frac{\sum_{k=1}^K \tilde{p}_{ij}(k)}{K}.$$

С этого момента предполагается, что все эксперты имеют одинаковую значимость в процессе группового принятия решений.

#### IV. Создание рейтинга альтернатив

Для определения ранжированного списка альтернатив можно использовать операторы GDD (*англ.* Guided Dominance Degree) и GNDD (*англ.* Guided Non Dominance Degree). Эти операторы способны вычислять рейтинг альтернатив (создавать ранжированные списки), используя матрицу коллективных предпочтений  $\mathbf{C}$ , полученную на предыдущем шаге:

$$\text{GDD}_i = \sum_{j=1}^M c_{ij}, \quad i = 1, \dots, M; \quad (1)$$

$$\text{GNDD}_i = \sum_{j=1}^M (1 - \max\{c_{ji} - c_{ij}, 0\}), \quad i = 1, \dots, M. \quad (2)$$

Конечные значения ранжирования RV вычисляются с использованием среднего между (1) и (2):

$$\text{RV}_i = \frac{\text{GDD}_i + \text{GNDD}_i}{2}, \quad i = 1, \dots, M.$$

Теперь альтернативы могут быть отсортированы на основании полученных  $\text{RV}_i$ . Стоит отметить, что предпочтительнее будет та альтернатива, у которой это значение выше.

#### V. Вычисление консенсуса

Чтобы определить, согласны ли эксперты с результатами или, напротив, их мнения разошлись, проводится анализ достигнутого консенсуса.

В работе предлагается два способа вычисления степени согласия экспертов.

1. Первый подход опирается на предпочтения, извлеченные из текстов обзуждений, т. е. на матрицу  $\mathbf{P}(k)$ . Полученные результаты определяются из сходств значений предпочтений, предоставляемых каждым из двух экспертов. Для вычисления степени сходства значений предпочтений между экспертами  $e_k$  и  $e_l$  для двух конкретных альтернатив  $x_i$  и  $x_j$  необходимо применить выражение

$$\text{sm}_{ij}(e_k, e_l) = s(\tilde{p}_{ij}(k), \tilde{p}_{ij}(l)) = 1 - |\tilde{p}_{ij}(k) - \tilde{p}_{ij}(l)|.$$

Для того чтобы измерить консенсус между двумя парами альтернатив  $x_i$  и  $x_j$  по всем участникам процесса GDM, используется следующая формула:

$$\text{cp}_{ij} = \frac{\sum_{k=1, k < l}^{K-1} \text{sm}_{ij}(e_k, e_l)}{\binom{K}{2}}.$$

Значение консенсуса на уровне альтернативы определяется выражением

$$ca_i = \frac{\sum_{j=1, i \neq j}^M (cp_{ij} + cp_{ji})}{2(M-1)}.$$

Наконец, можно вычислить глобальный достигнутый консенсус в процессе группового принятия решений:

$$GCP = \frac{\sum_{i=1}^M ca_i}{M}.$$

Таким образом, чем ближе значение глобального консенсуса к единице, тем более единодушны эксперты в выражении своих оценок и, следовательно, надежнее результаты.

2. Вторым способом определения достигнутого консенсуса в процессе группового принятия решений является исследование отношения экспертов к другим участникам обсуждения. Благодаря процедурам анализа тональности текстов можно провести тщательную оценку мнений экспертов и рассчитать консенсусное значение. Вычисление консенсуса включает следующие действия:

- (1) для каждой учетной записи извлекаются тексты, ссылающиеся непосредственно на других экспертов, участвующих в процессе группового принятия решений. Например, в Twitter или Instagram могут быть выбраны тексты обсуждений, которые включают синтаксис @username;
- (2) определение сравнительной парной оценки между экспертами: анализируются предложения, касающиеся экспертов  $e_k$  и  $e_l$ . Рассчитывается число слов, принадлежащих списку положительных и отрицательных слов. Таким образом,  $pwcount(e_k, e_l)$  измеряет согласие между экспертами  $e_k$  и  $e_l$ , а  $nwcount(e_k, e_l)$  дает меру разногласий.

Для вычисления консенсуса между экспертами  $e_k$  и  $e_l$  применяется формула:

$$(e_k, e_l) = \frac{pwcount(e_k, e_l) - nwcount(e_k, e_l)}{pwcount(e_k, e_l) + nwcount(e_k, e_l)}.$$

Полученное значение находится в интервале  $[-1, 1]$ . Для нормирования данных можно применить линейное преобразование:

$$cs_{[0,1]}(e_k, e_l) = \frac{cs_{[-1,1]}(e_k, e_l) + 1}{2}.$$

Отметим, чем меньше значение  $cs_{[0,1]}(e_k, e_l)$ , тем больше разногласий между экспертами  $e_k$  и  $e_l$ ;

- (3) вычисление глобального достигнутого консенсуса в процессе группового принятия решений на основании информации экспертов по отношению друг к другу:

$$\text{GCS} = \frac{\sum_{k=1, k < l}^{K-1} \text{cs}_{[0,1]}(e_k, e_l)}{\binom{K}{2}}.$$

Если полученное значение равно единице (степень согласия 100%), это означает, что достигнут максимальный консенсус.

## 4 Заключение

В статье представлен метод группового принятия решений, новизна которого заключается в том, что использует анализ тональности текстов пользователей для заполнения матриц отношений предпочтений. Благодаря этому эксперты не должны участвовать в утомляющем и неудобном процессе выражения мнений системе. Они могут свободно говорить и обсуждать проблемы в социальных сетях, не беспокоясь о том, как предоставить свои предпочтения.

Поскольку социальные сети представляются наиболее привлекательным средством общения через интернет, эта среда выбрана для осуществления процессов группового принятия решений. Предложенный метод также реализует измерения консенсусных мер, основанных на значениях предпочтений экспертов и мнениях об остальных участниках, принимающих решения.

В дальнейшем планируется разработать программное средство, способное работать с информацией в социальных сетях в реальных условиях.

## Литература

1. Wasserman S., Faust K. Social network analysis: Methods and applications. — Cambridge: Cambridge University Press, 1994. Vol. 8. 857 p.
2. Scott J. Social network analysis. — 4th ed. — London: Sage, 2017. 227 p.
3. Chiclana F., Tapia Garcia J., del Moral M. J., Herrera-Viedma E. A statistical comparative study of different similarity measures of consensus in group decision making // Inform. Sciences, 2013. Vol. 221. P. 110–123.
4. Morente-Molinera J. A., Pérez I. J., Ureña M. R., Herrera-Viedma E. On multi-granular fuzzy linguistic modeling in group decision making problems: A systematic review and future trends // Knowl.-Based Syst., 2015. Vol. 74. P. 49–60.
5. Urea R., Cabrerizo F. J., Morente-Molinera J. A., Herrera-Viedma E. GDM-R: A new framework in R to support fuzzy group decision making processes // Inform. Sciences, 2016. Vol. 357. P. 161–181.
6. Wu J., Xiong R., Chiclana F. Uninorm trust propagation and aggregation methods for group decision making in social network with four tuple information // Knowl.-Based Syst., 2016. Vol. 96. P. 29–39.

7. *Herrera-Viedma E., Cabrerizo F., Chiclana F., Wu J., Cobo M., Samuylov K.* Consensus in group decision making and social networks // *Stud. Inform. Control*, 2017. Vol. 26. Iss. 3. P. 259–268.
8. *Pang B., Lee L.* Opinion mining and sentiment analysis // *Found. Trends Inf. Ret.*, 2008. Vol. 3. P. 1–135.
9. *Baccianella S.* Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining // *Conference (International) on Language Resources and Evaluation Proceedings*, 2010. P. 2200–2204.
10. *Yager R. R.* On ordered weighted averaging aggregation operators in multi-criteria decision making // *IEEE T. Syst. Man Cyber.*, 1988. Vol. 18. P. 183–190.

*Поступила в редакцию 13.02.19*

---

---

## METHOD OF EXTRACTING EXPERTS' ASSESSMENTS FROM SOCIAL NETWORKS IN THE GROUP DECISION MAKING PROCESS

*O. V. Chukhno<sup>1</sup>, N. V. Chukhno<sup>1</sup>, Yu. V. Gaidamaka<sup>1,2</sup>, and K. E. Samouylov<sup>1,2</sup>*

<sup>1</sup>Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation

<sup>2</sup>Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** Social networks are the most preferred mean in terms of user interaction, information sharing, and communication. Therefore, they can be used in a group decision making process. However, in existing methods of group decision making, experts should follow the specific proposed structures when making their decisions. This is unnatural for online services in which users can express their opinions in a free form. To solve this problem, a model of a group decision making process is proposed in which experts can feel more free when evaluating pairs of alternatives or expressing their opinions. This method makes it possible to work in social networks. The article also offers a review and on its basis a formal description of the group decision making algorithm in a social network. A method for evaluating alternatives is presented, in which experts provide their assessments during the discussion of a problem with the use of texts, messages, or publications in chosen platforms. Methods are proposed for determining the level of consensus reached among the experts of this process.

**Keywords:** group decision making; social networks; consensus measures; sentiment analysis

**DOI:** 10.14357/08696527190408

## Acknowledgments

The publication has been prepared with the support of the “RUDN University Program 5-100” and funded by the Russian Foundation for Basic Research according to the research projects Nos. 17-07-00845 and 18-07-00576.

## References

1. Wasserman, S., and K. Faust. 1994. *Social network analysis: Methods and applications*. Cambridge: Cambridge University Press. Vol. 8. 857 p.
2. Scott, J. 2017. *Social network analysis*. 4th ed. London: Sage. 227 p.
3. Chiclana, F., J. Tapia Garcia, M.J. del Moral, and E. Herrera-Viedma. 2013. A statistical comparative study of different similarity measures of consensus in group decision making. *Inform. Sciences* 221:110–123.
4. Morente-Molinera, J. A., I. J. Pérez, M. R. Ureña, and E. Herrera-Viedma. 2015. On multi-granular fuzzy linguistic modeling in group decision making problems: A systematic review and future trends. *Knowl.-Based Syst.* 74:49–60.
5. Urea, R., F. J. Cabrerizo, J. A. Morente-Molinera, and E. Herrera-Viedma. 2016. GDM-R: A new framework in R to support fuzzy group decision making processes. *Inform. Sciences* 357:161–181.
6. Wu, J., R. Xiong, and F. Chiclana. 2016. Uninorm trust propagation and aggregation methods for group decision making in social network with four tuple information. *Knowl.-Based Syst.* 96:29–39.
7. Herrera-Viedma, E., F. Cabrerizo, F. Chiclana, J. Wu, M. Cobo, and K. Samuylov. 2017. Consensus in group decision making and social networks. *Stud. Inform. Control* 26(3):259–268.
8. Pang, B., and L. Lee. 2008. Opinion mining and sentiment analysis. *Found. Trends Inf. Ret.* 3:1–135.
9. Baccianella, S. 2010. Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. *Conference (International) on Language Resources and Evaluation Proceedings*. 2200–2204.
10. Yager, R. R. 1988. On ordered weighted averaging aggregation operators in multi-criteria decision making. *IEEE T. Syst. Man Cyb.* 18:183–190.

Received February 13, 2019

## Contributors

**Chukhno Olga V.** (b. 1995) — student, Department of Applied Probability and Informatics, Peoples’ Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; olgachukhno95@gmail.com

**Chukhno Nadezhda V.** (b. 1995) — student, Department of Applied Probability and Informatics, Peoples’ Friendship University of Russia (RUDN

University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; nvchukhno@gmail.com

**Gaidamaka Yuliya V.** (b. 1971) — Doctor of Science in physics and mathematics, professor, Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; senior scientist, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; gaydamaka-yuv@rudn.university

**Samouylov Konstantin E.** (b. 1955) — Doctor of Science in technology, professor, Head of Department, Director of Institute of Applied Mathematics and Telecommunications, Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; senior scientist, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; samuylov\_ke@rudn.university

## ПРИНЦИПЫ СОЗДАНИЯ И ФУНКЦИИ ИНТЕЛЛЕКТУАЛЬНОЙ СЛОВАРНОЙ СИСТЕМЫ

*В. В. Вакуленко<sup>1</sup>, И. М. Зацман<sup>2</sup>*

**Аннотация:** Дано описание принципов создания и основных функций проектируемого макета интеллектуальной словарной системы (ИСС), которая включает три основных структурных компонента: (1) хранилище параллельных текстов и мультимедийных данных; (2) лексикографическую базу знаний (ЛБЗ) и надкорпусные базы данных (НБД); (3) двуязычные электронные словари, генерируемые ИСС. Отличительные черты ИСС состоят в следующем: во-первых, она дает возможность систематизировать знания о словах, включая личностные и коллективные знания лексикографов, создающих словарные статьи, используя лексические единицы двух естественных языков (ЕЯ); во-вторых, ИСС обеспечивает обновление в словарях ранее сгенерированных с ее помощью словарных статей; в-третьих, в ИСС реализуется персонализация не только интерфейса пользователей двуязычных электронных словарей, но также полноты отображения словарных статей. Цель статьи состоит в описании принципов создания, основных функций и компонентов ИСС.

**Ключевые слова:** интеллектуальная словарная система; лексикографическая база знаний; электронный словарь; параллельные тексты; мультимедийные данные

**DOI:** 10.14357/08696527190409

### 1 Введение

В работе [1] дано описание концепции создания макета ИСС, включающей ЛБЗ и генерируемые ею двуязычные электронные словари<sup>3</sup>. Интеллектуальная словарная система является видом информационной системы, обеспечивающей формирование и регулярное обновление электронных словарей территориально распределенным коллективом лексикографов в режиме онлайн. Лексикографическая база знаний как компонент ИСС относится к категории баз знаний, что отражено и в ее названии. Этот компонент предназначен для систематизации знаний о словах [2]. Именно ЛБЗ обеспечивают работу лексикографов по созданию словарных статей и их обновление в случае необходимости.

---

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

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

<sup>3</sup>В разработке и обсуждении концепции принимали участие А. А. Гончаров, Д. О. Добровольский, А. А. Дурново, Анна А. Зализняк, В. И. Карпов, М. Г. Кружков и А. В. Шарандин.



Концепция создания макета включает перечень из пяти следующих задач, решение которых предполагает применение современных информационных технологий и средств информатики.

1. Систематизировать в ЛБЗ знания о словах, включая личностные и коллективные знания лексикографов, создающих словари, используя лексические единицы двух ЕЯ.
2. Разработать персонализированный интерфейс пользователя двуязычных электронных словарей, сгенерированных с помощью ЛБЗ.
3. Установить в электронных словарях тезаурусные связи между значениями лексических единиц (отношения «часть–целое», родовидовые отношения, синонимы и др.) с возможностью навигации по этим связям.
4. Реализовать в двуязычных словарях поиск словарных статей по лексическим, грамматическим и семантическим критериям поиска.
5. Включить в словарные статьи необходимые мультимедийные данные.

Разработка и применение макета ИСС даст возможность найти решение ряда назревших проблем, связанных с подготовкой электронных словарей, включая внесение оперативных изменений в словарные статьи, сокращение интервалов времени между выходом версий электронных словарей, отражение на временной шкале ретроспективы изменений, внесенных в словарные статьи. Цель статьи состоит в описании принципов реализации перечисленных задач, основных функций и компонентов макета ИСС, ориентированного на решение этих проблем.

## 2 Принципы создания интеллектуальной словарной системы

Рассмотрим предлагаемые принципы создания макета ИСС, реализация которых и должна обеспечить решение перечисленных задач. Отличительная черта первой задачи состоит в систематизации конвенциональных, личностных и коллективных знаний о словах в процессе описания как их значений, так и переводных эквивалентов значений. Новая функция ЛБЗ будет состоять в том, что одновременно будет предлагаться фиксировать степень социализации (термин Вежбицкого–Накамори [3, 4]) новых значений слов в коллективе, формирующем словарные статьи, т. е. какая часть коллектива лексикографов и лингвистов на данный момент будет согласна с предлагаемой авторской дефиницией того или иного концепта (= значения слова или словосочетания).

Планируется также фиксировать стадии эволюции на временной шкале дефиниций концептов и этапы преобразования личностных и коллективных концептов в конвенциональные. Предлагаемый подход к реализации этой задачи предоставит лингвистам и лексикографам, работающим над словарными статьями, принципиально новые функции и информационные ресурсы.

Во-первых, дефиниции концептов могут меняться лексикографами в процессе формирования словарной статьи. При этом нередко возникает необходимость увидеть в ретроспективе весь процесс изменения дефиниций. С этой целью и планируется фиксировать в ЛБЗ историю изменений с привязкой к временной шкале.

Во-вторых, появится возможность различать личностные и коллективные дефиниции концептов и отражать их изменения в ЛБЗ. Это даст возможность увидеть эволюцию во времени значений слов или словосочетаний через их дефиниции.

Предлагаемые нововведения будут полезны не только лексикографам, работающим над словарными статьями, но и специалистам, исследующим эволюцию языка в ретроспективе появления новых значений у ранее существовавших слов. Если на данный момент существующие подходы дают возможность проводить такие исследования только для конвенциональных концептов [5–7], то предлагаемый подход обеспечит исходные данные для исследования эволюции личностных и коллективных концептов.

Согласно второй задаче, в макете ИСС планируется сделать персонализированными и интерфейс, и наполнение электронных словарных статей. Один из недостатков традиционных электронных словарей — ограниченные возможности персонализации интерфейса доступа к словарным статьям. Например, если пользователю нужно посмотреть только список переводов значений слова без контекстных примеров, то он, как правило, должен просмотреть всю словарную статью в поисках значений этого слова и их переводов.

Таким образом, помимо традиционных настроек интерфейса (изменения вида шрифта, размера текста и других элементов оформления страницы) будут присутствовать настройки *полноты* отображения как самих словарных статей, так и их тематических подборок, сгенерированных при помощи ЛБЗ. Пользователь будет также иметь возможность выбрать именно те словарные статьи, которые удовлетворяют критериям частотности слов. Например, пользователь сможет видеть в результатах поиска только те статьи, которые соответствуют заданному критерию частотности слова по тому или иному корпусу, служащему источником текстов для ИСС. Возможна также группировка статей по тематике (например, все дни недели) — это пример еще одной функции, которую планируется обеспечить с помощью макета ИСС.

Согласно третьей задаче, в ЛБЗ и генерируемых электронных словарях будут устанавливаться несколько видов тезаурусных отношений. Реализация этой задачи в процессе проектирования ИСС предоставит пользователям возможность навигации по тому или иному виду тезаурусных связей между значениями слов.

Согласно четвертой задаче, предлагается реализовать традиционный поиск словарных статей. Это даст возможность найти интересующее пользователя слово или словосочетание на одном ЕЯ и/или переводные эквиваленты на другом ЕЯ, что обеспечит поиск примеров релевантных контекстов в оригинале и переводе

примеров предложений в словарных статьях. Поиск подобных примеров контекстов может быть использован переводчиками при выборе варианта перевода.

Согласно пятой задаче, планируется включить в отдельные словарные статьи иллюстрирующий мультимедийный материал. Такая функция не нова и уже присутствует в визуальных словарях [8]. Например, в электронном визуальном словаре (Visual Dictionary Online Website) слово *face* сопровождается аудио-файлом с его произношением, а его значения иллюстрируются 22 рисунками, показывающими также отношения «часть–целое» между словами. Так, на одном из рисунков кроме лица обезьяны перечисляются другие части ее тела и обозначающие их слова (но при этом отсутствуют гиперссылки на соответствующие им словарные статьи) [9].

Визуальные словари доступны в двуязычной и многоязычной версиях, а также в версиях, адаптированных для определенных возрастных групп [10]. Планируется, что ИСС сможет обеспечивать в случае необходимости включение в словарные статьи аналогичных мультимедийных данных.

### **3 Основные функции и компоненты интеллектуальной словарной системы**

Предполагается, что макет ИСС будет включать три основных структурных компонента: (1) хранилище параллельных текстов и мультимедийных данных; (2) ЛБЗ и НБД; (3) двуязычные электронные словари, генерируемые ЛБЗ.

В соответствии с тремя перечисленными структурными компонентами в ИСС различаются три вида интерфейсов: пользовательский интерфейс электронных словарей; интерфейс ЛБЗ; интерфейс хранилища, обеспечивающий поиск искомых слов в параллельных текстах.

Интерфейсы первого вида будут обладать возможностью адаптации к личностным потребностям пользователей электронных словарей. Интерфейсы второго вида обеспечат работу лексикографов с ЛБЗ, которые формируют словарные статьи. Интерфейсы третьего вида нужны для взаимодействия с корпусами параллельных текстов, которые хранятся в ИСС.

Двуязычные электронные словари, располагающиеся в первом компоненте, представляют собой регулярно пополняемые в режиме онлайн информационные продукты ИСС. Предоставление доступа к электронным словарям осуществляется через сеть Интернет. Пользователям словарей нет необходимости понимать, как функционирует вся ИСС в целом, чтобы работать со словарями. Их интерфейс поможет представить словарную информацию в персонифицированной форме. В то же время лексикографам необходимо понимать, как функционирует вся ИСС в целом, чтобы вносить изменения и дополнения в ЛБЗ, генерирующую словарные статьи, которые по мере их готовности автоматически будут загружаться в электронные словари по сети Интернет. Такой подход даст возможность пользователям электронных словарей не ждать годами выхода новых изданий

словарей, так как их обновление будет осуществляться постатейно в режиме онлайн.

Как следует из структуры ИСС, второй ее компонент включает ЛБЗ, доступную территориально распределенному коллективу лингвистов и лексикографов. Лексикографическая база знаний позволяет им совместно работать над определенной языковой единицей или конструкцией, например вводить и согласовывать новое значение для слова, обнаруженное в параллельных текстах, задав его дефиницию. Подобный функционал обеспечит каждому члену территориально распределенного коллектива возможность видеть результаты работы всех его участников. Таким образом, ЛБЗ позволит каждому автоматически проинформировать коллег о внесенных изменениях. При этом каждое личностное нововведение может быть согласовано в рамках коллектива в режиме онлайн.

Кроме ЛБЗ второй компонент включает ряд НБД, с помощью которых лингвисты проводят исследования, необходимые для создания словаря, например классификацию логико-семантических отношений [11].

Актуализация словарных статей будет обеспечиваться в двух основных режимах работы: индивидуальном и коллективном. В индивидуальном режиме решается задача описания в ЛБЗ личностного знания лексикографа о новых значениях слов. В коллективном режиме решается задача согласования личностных дефиниций значений слов между лексикографами и формирование ими коллективных дефиниций.

Обеспечение доступа к корпусам параллельных текстов даст возможность извлекать из них новые значения слов и описывать их в НБД и ЛБЗ [12–16]. Лексикографическая база знаний является ключевой составляющей ИСС и, подобно ядру, находится в ее центре. Лексикографическая база знаний связана как с первым, так и с третьим компонентом ИСС системой информационных связей.

Как было отмечено выше, ИСС будет поддерживать извлечение и описание новых значений исследуемых слов из параллельных текстов, что является отличительной чертой ИСС как интеллектуальной системы. Согласованные и включенные в ЛБЗ коллективные дефиниции новых значений слов автоматически копируются в электронные словари вместе с примерами их употребления, если проставлен признак готовности соответствующих словарных статей к публикации. Поэтому пользователи электронных словарей ИСС смогут начать работу с новой версией словарной статьи, содержащей последние изменения и дополнения, и увидеть результат проделанной специалистами работы сразу же после ее электронной публикации.

Третий компонент ИСС содержит корпуса параллельных текстов и хранилище мультимедийной информации. В данном случае корпуса разделены на два типа: корпуса, хранящиеся непосредственно в ИСС, и внешние корпуса, к которым лексикографы и лингвисты имеют доступ по сети Интернет.

Изначально слова, описываемые в словарных статьях, находятся в текстах, с которыми работают лексикографы и лингвисты. Описание новых значений слов

подразумевает сначала их лингвистическое аннотирование в НБД в контексте тех предложений, в которых они встретились, до размещения в словарной статье. Под лингвистической аннотацией подразумевается структурированное описание слов вместе с их значениями, дополненное метаданными. Один из примеров метаданных — признаки (теги) частей речи.

Наличие хранилища корпусов параллельных текстов непосредственно в ИСС предоставляет более широкие возможности для работы, нежели внешние корпусы. Это обусловлено возможностью работать с целым текстом, а не только с его фрагментами, а также наличием гиперссылок к текстам от словарных статей ЛБЗ и аннотаций НБД. Тем не менее внешние корпусы также играют важную роль, так как дают возможность использовать фрагменты из более широкого спектра текстов в качестве примеров переводов в словарных статьях.

Помимо перечисленных функций ИСС предоставит расширенные поисковые возможности как лексикографам, так и пользователям словарей. На данном этапе разработки принципов создания ИСС предполагается реализовать следующие фильтры и виды поиска [1]:

- синтаксический фильтр, когда при поиске могут использоваться признаки конструкций или предложений, в которые входит искомое слово (например, вопросительное, восклицательное или отрицательное предложение, диалоговая реплика и т. д.);
- статистический фильтр по частотности слов, что позволит отображать или, наоборот, исключать при отображении результатов поиска низкочастотные слова (частотности слов являются корпусно-ориентированными), т. е. выше или ниже заданного порога;
- хронологический фильтр, который позволит использовать датировку текстов, т. е. в словарной статье могут отображаться примеры переводов только в тех текстах, которые относятся к заданному периоду времени;
- поиск по лексикографическому типу: семантические группы слов (времена года, страны света, части речи, дни недели и др.), функциональные группы слов (например, коннекторы [17]), типы устойчивых словосочетаний и т. д.

#### 4 Заключение

На современном этапе разработки словарных систем в них используются, как правило, методы систематизации и перевода конвенциональных значений слов. Это отвечает нуждам пользователей электронных словарей, но лексикографам в процессе формирования словарных статей иногда требуется описывать новые значения слов, извлеченные из предложений параллельных текстов, т. е. приводить для них свое личностное или коллективное описание в виде дефиниции, иллюстрируемое найденными предложениями.

При реализации этой задачи в макете ИСС предполагается отражать степень социализации извлеченных новых значений слов в территориально распреде-

ленном коллективе лексикографов и этапы формирования их коллективных дефиниций, что является существенным элементом новизны проектируемого макета. При этом планируется использовать методы кодирования личностного, коллективного и конвенционального знания, основанные на интерфейсах 3-го порядка, описанных в работе [18]. Эти методы предназначены для *одновременного* кодирования в цифровой среде ИСС и новых значений слов (словосочетаний), и форм их представления в разных ЕЯ. В заключение отметим, что современные средства информатики позволяют создать ИСС, обладающую описанными в статье функциями, с реализацией этих методов кодирования.

## Литература

1. Зацман И. М. Кодирование концептов в цифровой среде // Информатика и её применения, 2019 (в печати). Т. 13. Вып. 4.
2. Дубичинский В. В. Теоретическая и практическая лексикография. — Вена: Verlag Otto Sagner, 1998. 160 с.
3. Wierzbicki A. P., Nakamori Y. Basic dimensions of creative space // Creative space: Models of creative processes for knowledge civilization age / Eds. A. P. Wierzbicki, Y. Nakamori. — Berlin–Heidelberg: Springer Verlag, 2006. P. 59–90.
4. Wierzbicki A. P., Nakamori Y. Knowledge sciences: Some new developments // Z. Betriebswirt., 2007. Vol. 77. Iss. 3. P. 271–295.
5. Kay C., Roberts J., Samuels M., Wotherspoon I., Alexander M. The historical thesaurus of English. Version 4.2. — Glasgow, U.K.: University of Glasgow, 2015. <https://historicalthesaurus.arts.gla.ac.uk>.
6. Ramiro C., Srinivasan M., Malt B. C., Xu Y. Algorithms in the historical emergence of word senses // P. Natl. Acad. Sci. USA, 2018. Vol. 115. Iss. 10. P. 2323–2328.
7. Kay C., Roberts J., Samuels M., Wotherspoon I., Alexander M. The historical thesaurus of English. Face. Version 4.2. — Glasgow, U.K.: University of Glasgow. <http://historicalthesaurus.arts.gla.ac.uk/category-selection/?qsearch=face>.
8. Корбей Ж.-К., Аршамбо А. Новый визуальный энциклопедический словарь / Пер. с англ. — М.: Рипол-Классик, 2014. 1112 с. (Corbeil J.-C., Archambault A. The visual dictionary with definitions. — Montreal: QA International, 2011. 960 p.)
9. Visual Dictionary Online Website. [http://www.visualdictionaryonline.com/search-results.php?s\\_search\\_box=face&s\\_search\\_result=face](http://www.visualdictionaryonline.com/search-results.php?s_search_box=face&s_search_result=face).
10. Англо-русский визуальный словарь для детей. — М.: АСТ, 2017. 192 с.
11. Инькова О., Манзотти Э. Связность текста: мерологические логико-семантические отношения. — М.: Языки славянских культур, 2019 (в печати).
12. Зацман И. М. Процессы целенаправленной генерации и развития кросс-языковых экспертных знаний: семиотические основания моделирования // Информатика и её применения, 2015. Т. 9. Вып. 3. С. 106–123.
13. Зацман И. М. Стадии целенаправленного извлечения знаний, имплицированных в параллельных текстах // Системы и средства информатики, 2018. Т. 28. № 3. С. 175–188.
14. Зацман И. М. Имплицированные знания: основания и технологии извлечения // Информатика и её применения, 2018. Т. 12. Вып. 3. С. 74–82.

15. Зацман И. М. Целенаправленное развитие систем лингвистических знаний: выявление и заполнение лагун // Информатика и её применения, 2019. Т. 13. Вып. 1. С. 91–98.
16. 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.
17. Семантика коннекторов: контрастивное исследование / Под ред. О. Ю. Иньковой. — М.: ТОРУС ПРЕСС, 2018. 368 с.
18. Зацман И. М. Интерфейсы третьего порядка в информатике // Информатика и её применения, 2019. Т. 13. Вып. 3. С. 85–92.

Поступила в редакцию 01.09.19

---

---

## FRAMEWORK FOR THE DESIGN OF INTELLECTUAL DICTIONARY SYSTEMS AND THEIR PRIMARY FUNCTIONS

*V. V. Vakulenko 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 paper proposes the framework for the design of the Intellectual Dictionary System (IDS) prototype, as well as its primary functions, including three main structural components: (i) parallel text and multimedia data storage; (ii) lexicographical knowledge base and supracorpora databases; and (iii) bilingual online dictionaries generated by IDS. Distinguishing features of IDS are as follows: first, the possibility to store the descriptions of personal and collective concepts of the lexicographers’ knowledge, as well as their dictionary representation in two natural languages in the form of dictionary entries; second, IDS allows updating online dictionaries by entries; and third, IDS gives the possibility to customize not only the user interface of the bilingual online dictionaries, but also the amount of the content displayed in an entry. The goal of the article is to describe the framework for the design of IDS, its primary functions and components.

**Keywords:** intellectual dictionary system; lexicographical knowledge base; online dictionary; parallel texts; multimedia data

**DOI:** 10.14357/08696527190409

### References

1. Zatsman, I. M. 2019 (in press). Kodirovanie kontseptov v tsifrovoy srede [Encoding of concepts in the digital medium]. *Informatika i ee Primeneniya — Inform. Appl.* 13(4).



2. Dubichinsky, V. V. 1998. *Teoreticheskaya i prakticheskaya leksikografiya* [Theoretical and practical lexicography]. Vienna: Verlag Otto Sagner. 160 p.
3. Wierzbicki, A. P., and Y. Nakamori. 2006. Basic dimensions of creative space. *Creative space: Models of creative processes for knowledge civilization age*. Eds. A. P. Wierzbicki and Y. Nakamori. Berlin: Springer Verlag. 59–90.
4. Wierzbicki, A. P., and Y. Nakamori. 2007. Knowledge sciences: Some new developments. *Z. Betriebswirt.* 77(3):271–295.
5. Kay, C., J. Roberts, M. Samuels, I. Wotherspoon, and M. Alexander. 2015. *The historical thesaurus of English*. Version 4.2. Glasgow, U.K.: University of Glasgow. Available at: <https://historicalthesaurus.arts.gla.ac.uk> (accessed August 27, 2019).
6. Ramiro, C., M. Srinivasan, B. C. Malt, and Y. Xu. 2018. Algorithms in the historical emergence of word senses. *P. Natl. Acad. Sci. USA* 115(10):2323–2328.
7. Kay, C., J. Roberts, M. Samuels, I. Wotherspoon, and M. Alexander. 2015. The historical thesaurus of English. Face. Version 4.2. Glasgow, U.K.: University of Glasgow. Available at: <http://historicalthesaurus.arts.gla.ac.uk/category-selection/?qsearch=face> (accessed August 27, 2019).
8. Corbeil J.-C., and A. Archambault. 2011. *The visual dictionary with definitions*. Montreal: QA International. 960 p.
9. Visual Dictionary Online Website. Available at: [http://www.visualdictionaryonline.com/search-results.php?s\\_search\\_box=face&s\\_search\\_result=face](http://www.visualdictionaryonline.com/search-results.php?s_search_box=face&s_search_result=face) (accessed August 27, 2019).
10. Vertyagina, A. A., and D. V. Morozova. 2017. *Anglo-russkiy vizual'nyy slovar' dlya detey* [English-Russian visual dictionary for children]. Moscow: AST. 192 p.
11. Inkova, O., and E. Manzotti. 2019 (in press). *Svyaznost' teksta: mereologicheskie logiko-semanticheskie otnosheniya* [Text coherence: Mereological logical semantic relations]. Moscow: LRC Publishing House.
12. Zatsman, I. M. 2015. Protsessy tselenapravlennoy generatsii i razvitiya kross-yazykovykh ekspertnykh znaniy: semioticheskie osnovaniya modelirovaniya [Goal-oriented processes of cross-lingual expert knowledge creation: Semiotic foundations for modeling]. *Informatika i ee Primeneniya — Inform. Appl.* 9(3):106–123.
13. Zatsman, I. M. 2018. Stadii tselenapravlennoy izvlecheniya znaniy, implitsirovannykh v parallel'nykh tekstakh [Stages of goal-oriented discovery of knowledge implied in parallel texts]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(3):175–188.
14. Zatsman, I. M. 2018. Implitsirovanne znaniya: osnovaniya i tekhnologii izvlecheniya [Implied knowledge: Foundations and technologies of explication]. *Informatika i ee Primeneniya — Inform. Appl.* 12(3):74–82.
15. Zatsman, I. M. 2019. Tselenapravlennoe razvitie sistem lingvisticheskikh znaniy: vyyavlenie i zapolnenie lakun [Goal-oriented development of linguistic knowledge systems: Identifying and filling of lacunae]. *Informatika i ee Primeneniya — Inform. Appl.* 13(1):91–98.
16. Goncharov, A. A., and I. M. Zatsman. 2019. Informatsionnye transformatsii parallel'nykh tekstov v zadachakh izvlecheniya znaniy [Information transformations of parallel texts in knowledge extraction]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(1):180–193.
17. Inkova, O. Yu., ed. 2018. *Semantika konnektorov: kontrastivnoe issledovanie* [Semantics of connectives: A contrastive study]. Moscow: TORUS PRESS. 368 p.



18. Zatsman, I. M. 2019. Interfeysy tret'ego poryadka v informatike [Third-order interfaces in informatics]. *Informatika i ee Primeneniya — Inform. Appl.* 13(3):85–92.

*Received September 1, 2019*

## **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; fortydays@protonmail.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

## О МЕТОДАХ ОЦЕНКИ КАЧЕСТВА МАШИННОГО ПЕРЕВОДА\*

*А. К. Рычихин*<sup>1</sup>

**Аннотация:** Рассматриваются подходы к определению качества машинного перевода (МП) и несколько методов его оценки. Целью статьи является обзор ряда методов и подходов к экспертной и автоматической оценке качества МП. В первой части описываются методы относительной экспертной оценки (ранжирование переводов) и абсолютной оценки, основанной на штрафах за ошибки в переводе, а также программное обеспечение и алгоритмы, упрощающие экспертную оценку и обрабатывающие экспертные суждения. Наибольшее внимание уделяется типологии ошибок DQF/MQM (Dynamic Quality Framework — Динамическая модель оценки качества, Multidimensional Quality Metrics — Многомерные метрики качества) как наиболее гибкой и не нацеленной на ограниченную предметную область. Вторая часть статьи посвящена обзору метрик автоматической оценки качества МП, не использующих знания об исходном и переводящем языке, а также коэффициентам корреляции экспертной и автоматической оценки.

**Ключевые слова:** качество машинного перевода; машинный перевод; метрики автоматической оценки качества машинного перевода; ранжирование переводов; типологии ошибок машинного перевода; экспертная оценка качества перевода

**DOI:** 10.14357/08696527190410

### 1 Введение

В теории перевода не существует общепринятого определения качества перевода, однако задача оценки качества МП с каждым годом становится актуальнее, в первую очередь благодаря внедрению МП в производственную и социальную сферы, где задействованы два и более естественных языков. По оценке TAUS<sup>2</sup>, бюджет индустрии МП в 2014 г. составил 250 млн долл. США и продолжает ежегодно расти [1]. В 2012 г. объем текста, который ежедневно обрабатывает онлайн-переводчик Google.Translate, превзошел совокупный объем переводных текстов, созданных профессиональными переводчиками за год [2]. С 2012 по 2016 гг. объем текста, переводимого Google.Translate, увеличился вдвое и составил 143 млрд слов в день в 100 языковых парах [3]. Очевидно, что МП

---

\* Автор выражает благодарность за помощь в подготовке статьи В. А. Нуриеву и И. М. Зацману.

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

<sup>2</sup> Translation Automation User Society — Пользовательское сообщество по автоматизации перевода.

уже широко используется, отсюда и необходимость совершенствования методов оценки его качества [4].

Целью статьи является описание ряда современных методов и подходов к экспертной и автоматической оценке качества МП. Для экспертной оценки качества характерен функциональный подход (качество перевода тем выше, чем успешнее он справляется со своими задачами) [5], который находит существенно меньшее применение в автоматической оценке МП, где качество оценивается через степень схожести с референтным (профессиональным) переводом (РП).

## 2 Экспертная оценка качества машинного перевода

При экспертной оценке основными критериями считаются точность (передача содержания оригинального текста) и естественность (лексическая и грамматическая правильность выражения на языке перевода). Другие критерии зависят от цели перевода и включают: легкость чтения и понимания, приемлемость и др. Детальный обзор критериев представлен в [6].

Для оценки качества двух и более переводов одного и того же источника применяется ранжирование переводов [6], результаты которого затем используются для сравнения систем МП. Эксперты получают задание сравнить качество переводов и упорядочить их иерархически, устанавливая соотношения между ними: «лучше», «хуже», «одинаковое качество»<sup>1</sup>. Основной единицей сравнения служит предложение. Для упрощения процедуры ранжирования разрабатывается программное обеспечение, например Appraise [8], которое создает пользовательский интерфейс и производит подсчет оценок.

Отдельное направление исследований в этой области — создание и тестирование алгоритмов, позволяющих получить результат, наиболее близкий к тому, который получили бы эксперты при сравнении полных текстов, переведенных разными системами МП, используя минимальный объем сравниваемых текстов. В перспективе это позволит привлекать только высококвалифицированных экспертов, не прибегая к краудсорсингу<sup>2</sup>, т. е. увеличит скорость и надежность оценки. Для этих целей используется модель ожидаемого выигрыша (Expected Wins Model) [9], модель Хопкинса и Мэя [10], алгоритм TrueSkill, адаптированный для ранжирования систем МП [11].

Следует заметить, что экспертная оценка предполагает привлечение двух и более экспертов. Однако их действия могут характеризоваться разобщенностью, т. е. быть несогласованными. В качестве меры согласия между экспертами используют коэффициенты каппа Коэна (для двух экспертов), каппа Флейса (для

---

<sup>1</sup>О процедуре ранжирования см. подробнее в работе [7].

<sup>2</sup>Так, компания Amazon предлагает краудсорсинговую (от англ. crowd — толпа, source — источник) платформу Mechanical Turk, на которой пользователи сети Интернет могут выполнять несложные задания онлайн за определенную плату (см. <https://www.mturk.com>).

трех или более экспертов) [12] или тау Кендалла [13]. При низком уровне согласия дальнейшее использование<sup>1</sup> проведенных оценок нецелесообразно.

Экспертная оценка не всегда предполагает сравнение двух и более переводов. Согласно установленным заранее критериям может оцениваться и один имеющийся перевод. В таком случае используется шкала Ликерта, которая представляет собой оценку по пятибалльной шкале, где при оценке точности «1» перевод не передает содержание оригинального текста, «5» — перевод полностью передает содержание оригинального текста [14]. Таким же образом оценивается степень естественности переводного текста, терминология<sup>2</sup> (его терминологическая системность), стиль (стилистическое соответствие оригиналу) и т. д.

Оценка машинного перевода может сопровождаться аннотацией ошибок и составлением их типологий. Такие типологии активно разрабатываются с 1990-х гг.<sup>3</sup>. Они предполагают инвентаризацию ошибок с присвоением каждому типу ошибок определенного веса (в некоторых типологиях они всегда равны 1). Типологии могут иметь разную степень спецификации, т. е. разную глубину проработки. При их составлении устанавливается степень критичности каждой ошибки и штраф за нее.

Типологии ошибок перевода имеют разную целевую направленность. Они разрабатываются в соответствии со следующими целями (см. подробнее в [16]):

- (1) для определения стандартов качества перевода (ISO CD 14080, LISA QA Model, SAE J2450);
- (2) для улучшения инструментов автоматической проверки текста (ApSIC XBench, Okapi CheckMate, Yamagata QA Distiller);
- (3) для оценки качества перевода на экзаменах при обучении переводчиков (ATA Certification Exam);
- (4) для ручной разметки текстов в системах автоматизированного перевода и памяти перевода (SDL TMS Classic).

Вместе с тем использование этих типологий иногда бывает затруднено: как правило, они создаются под конкретную задачу и потому часто являются узконаправленными. Так, в типологию ошибок LISA QA (Localisation Industry Standards Association Quality Assessment — методика оценки качества, разработанная Ассоциацией по формированию стандартов для индустрии локализации) включена рубрика «оценка функционирования ПО»<sup>4</sup> как отдельная категория

---

<sup>1</sup> Например, для проверки корреляций с автоматической оценкой качества МП (см. ниже).

<sup>2</sup> Следует отметить, что англоязычные наименования рубрик в типологии не всегда обозначают конкретные ошибки. Они также могут указывать, например, на какое-либо свойство текста (точность, естественность, стиль) или уровень языка, к которому ошибка относится (грамматика) и др. Подробнее см. таблицу.

<sup>3</sup> Одна из первых попыток провести такую оценку отражена в докладе ALPAC (Automatic Language Processing Advisory Committee — Консультативный комитет по автоматической обработке языка) [15].

<sup>4</sup> Предполагает тестирование конечным пользователем, говорящим на языке перевода.

## Ошибки, выделяемые в типологии DQF/MQM (два первых уровня)

1-й уровень	2-й уровень
Точность (Accuracy)	Непереведенный сегмент; опущение значимой информации; добавление лишней информации; неточный перевод (буквальный, неоднозначный и т. д.); неподходящая вставка из «памяти перевода»
Естественность (Fluency)	Кодировка символов; грамматика; уровень вежливости, требующий обязательного грамматического выражения ( <i>англ.</i> grammatical register); пунктуация; орфография; непоследовательность (разные сокращения одних и тех же слов и т. д.)
Стиль (Style)	Сложность прочтения (излишне сложные грамматические конструкции или лексика); уровень вежливости ( <i>англ.</i> register); корпоративный стиль; непоследовательность при передаче стиля; использование несоответствующего стиля ( <i>third-party style</i> )
Дизайн (Design)	Графика и таблицы; переносы слов; длина текста (значительная разница длины оригинала и перевода); форматирование (поля, сноски, колонтитулы и т. д.); локальное форматирование (шрифт, начертание, отступ, выравнивание) и др.
Соблюдение традиций написания текстов на языке перевода (Local convention)	Формат адреса, дат, валют, мер и весов, имен, телефонных номеров, времени, летоисчисления; специфическая для языка пунктуация; специфические горячие клавиши (важно при переводе инструкций)
Терминология (Terminology)	Ошибки в выборе термина; термины из другой предметной области; непоследовательное употребление разных терминов в тексте для отсылки к одному денотату
Истинность (Verity)	Полнота описания на языке перевода, достаточная для понимания целевой аудиторией (при условии, что оригинал не содержит эту информацию в эксплицированном виде); отсылки к специфичным культурным реалиям; пригодность к использованию конечными пользователями; соответствие правовым нормам и др.

ошибок, тогда как «точность», «естественность», «стиль» включены в одну общую рубрику. Подробный обзор типологий ошибок МП и операций редактирования, необходимых для их исправления, представлен в работе [17].

Отдельного внимания заслуживает типология ошибок DQF/MQM, разработанная в TAUS и DFKI<sup>1</sup> в 2014 г. При ее разработке учитывались 9 ранее созданных классификаций [16]. Типология имеет 4 уровня (наиболее специфицированные типы ошибок относятся к четвертому уровню; при оценке перевода

<sup>1</sup>Deutsches Forschungszentrum für Künstliche Intelligenz — Немецкий центр исследований искусственного интеллекта.

можно выбирать степень спецификации — от одного до четырех уровней в зависимости от пользовательских целей); учитываются четыре степени критичности ошибок:

- (1) критическая ошибка (даже одна критическая ошибка делает перевод непригодным к использованию — штраф 100);
- (2) значительная ошибка (может значительно затруднить понимание переводного текста — штраф 10);
- (3) незначительная (не затрудняет понимание — штраф 1);
- (4) нулевая (текст не содержит ошибок, но при этом его необходимо модифицировать — штраф 0).

В таблице представлены типы ошибок первого и второго уровня<sup>1</sup> типологии DQF/MQM.

Присваивая разные веса категориям первого уровня, эксперт способен настраивать систему оценивания. Веса могут колебаться от 0 (в случае, если некоторые категории ошибок совершенно не важны) до 1 для менее релевантных типов ошибок и выше 1 — для более релевантных.

Используя формулу

$$Q = 1 - \frac{P}{C}, \quad P = \sum_{i=1}^n w_i s_i,$$

где  $Q$  — индикатор качества перевода;  $P$  — суммарный штраф за все ошибки (суммирование ведется по всем категориям ошибок);  $C$  — число слов в документе;  $w_i$  — вес  $i$ -й категории ошибок;  $s_i$  — серьезность ошибки  $i$ -й категории, можно вычислить количественный показатель качества перевода на основе данной типологии ошибок.

В случае если дробь  $P/C > 1$ , значение  $Q$  может быть отрицательным; в случае безошибочного перевода  $Q = 1$ . Например, в переводе из 500 слов при 3 незначительных (штраф 1) и 2 значительных (штраф 10) ошибках результат составит 95,4% (вес всех категорий ошибок по умолчанию равен 1):

$$Q = 1 - \frac{3 \cdot 1 + 2 \cdot 10}{500} = 0,954,$$

что сигнализирует о качестве перевода в 95,4%.

Существенным преимуществом типологии ошибок DQF/MQM является ее универсальность. Она не нацелена на какую-либо предметную область и подходит для текстов любой тематики. Как и в случае любых других методов аннотирования ошибок, использование DQF/MQM остается времязатратным и дорогостоящим. О попытках автоматизации разметки ошибок в переводе (методах Hjerson и Addicter), позволяющих ускорить и удешевить этот процесс, см. в работах [17, 19].

<sup>1</sup>Объем статьи не позволяет рассмотреть всю типологию полностью. См. [18] с примерами и комментариями ошибок на всех уровнях.

### 3 Автоматическая оценка качества машинного перевода

Автоматическая оценка определяется по уровню соответствия машинного перевода референтному (профессиональному). Поэтому оправдано использование нескольких РП. Для установления уровня соответствия МП и РП применяются методы измерения точности (какой процент перевода выполнен без ошибок) и полноты (какая доля того, что было необходимо перевести, действительно переведена) перевода. Ниже представлены некоторые метрики автоматической оценки качества МП.

WER (Word Error Rate — пословный коэффициент ошибок) [2] и PER (Position-Independent WER — позиционно-независимый пословный коэффициент ошибок) [21] были заимствованы для оценки качества МП из автоматической обработки звучащей речи. WER, или расстояние Левенштейна, определенное на словах (WER использует в качестве единиц слова, а не символы, в отличие от расстояния Левенштейна), понимается как количество преобразований (трансформаций), *необходимое для исправления всех ошибок* в МП или фрагменте распознаваемой звучащей речи. Однако понятие «правильности», применимое к звучащей речи, не может быть адаптировано к качеству МП из-за отсутствия единственного верного перевода<sup>1</sup>. По определению количественный показатель качества перевода WER вычисляется по формуле:

$$\text{WER} = \frac{S + I + D}{R},$$

где  $S$  — число замен;  $I$  — число вставок;  $D$  — число удалений;  $R$  — длина РП [22].

Количественный показатель качества перевода PER не учитывает несовпадения МП и РП в порядке слов и предполагает штраф за слишком длинные переводы<sup>2</sup>. PER вычисляется по формуле:

$$\text{PER} = 1 - \frac{C - \max(0, O - R)}{R},$$

где  $C$  — правильно переведенные слова;  $O$  — длина МП;  $R$  — длина РП [22].

Самым распространенным показателем, используемым для автоматической оценки качества МП, является BLEU (*англ.* Bilingual Evaluation Understudy) [23]. BLEU предполагает подсчет  $n$ -граммной точности (чаще всего  $n \in [1; 4]$ ) [22] при определении соответствия МП и РП, а также штраф за слишком короткий перевод;  $n$ -граммная точность определяется как отношение последовательностей из  $n$  слов, совпадающих в МП и РП, к общему числу

<sup>1</sup>Так, Э. Уэй отмечает: «Переводчики сходятся во мнении, что не существует единственно *правильного* перевода» [4]. Одновременно могут существовать два и более правильных перевода одного исходного текста.

<sup>2</sup>Перевод может быть описан как содержащий небольшое число ошибок, но только потому, что по каким-то причинам правильно переведенные слова появились там неоднократно, а какие-то части оригинала упущены совсем.

последовательностей из  $n$  слов в МП. Штраф за слишком короткий перевод BP (brevity penalty) вычисляется как

$$BP = \min \left( 1, \frac{O}{R} \right),$$

Значения штрафа за краткость и  $n$ -граммной точности используются для определения BLEU:

$$BLEU(n) = BP \cdot \prod_{i=1}^n \text{precision},$$

где precision —  $n$ -граммная точность [22].

Подсчет  $n$ -граммных совпадений также лежит в основе статистического фразового<sup>1</sup> МП (СФМП) [22]. Некоторые исследователи выдвигают гипотезу о предвзятом характере метода BLEU, который необъективно высоко оценивает результаты, полученные с помощью систем СФМП: корреляция с экспертным суждением для них ниже<sup>2</sup>, чем для МП, основанного на правилах, и нейронного МП. Высокие результаты BLEU для СФМП могут объясняться использованием одних и тех же принципов для перевода и для оценки качества перевода, которыми понятие качества перевода не исчерпывается [4, 22].

Как уже было отмечено, BLEU использует для оценки качества МП  $n$ -граммную точность, но не апеллирует к полноте (за исключением введенного штрафа за краткость). Другие метрики, как, например, ChrF (Character F-measure — посимвольная F-мера), используют также  $n$ -граммную полноту [24, 25]. В ChrF  $n \in [1; 6]$ . За единицу анализа принимается один символ за исключением пробелов

$$\text{ChrF}\beta = (1 + \beta^2) \frac{\text{ChrP} \cdot \text{ChrR}}{\beta^2 \text{ChrP} + \text{ChrR}},$$

где ChrP — посимвольная  $n$ -граммная точность; ChrR — посимвольная  $n$ -граммная полнота. В данном случае  $n$ -граммная точность определяется как отношение последовательностей из  $n$  символов, совпадающих в МП и РП, к общему числу последовательностей из  $n$  символов в МП, а  $n$ -граммная полнота — как отношение последовательностей из  $n$  символов, совпадающих в МП и РП, к общему числу последовательностей из  $n$  символов в РП. Чем больше значение коэффициента  $\beta$ , тем значения ChrF ближе к значениям полноты, чем точности.

Для проверки корреляции автоматических и экспертных методов оценки качества МП используются коэффициенты корреляции Пирсона (Pearson's  $r$ ),

<sup>1</sup>В СФМП фраза понимается как некоторая последовательность слов на исходном и переводящем языках, которые в обоих языках следуют непосредственно друг за другом. Так, при переводном немецко-английском соответствии предложений *Natürlich hat John Spass am Spiel — Of course John has fun with the game* самыми крупными фразами являются «Spass am» — «fun with the» и «natürlich» — «of course». Остальные фразы являются однословными.

<sup>2</sup>Результаты СФМП могут одновременно высоко оцениваться BLEU и низко — экспертами.



Спирмена (Spearman's  $\rho$ ) и Кендалла (Kendall's  $\tau$ ). Самым распространенным из них является коэффициент Пирсона [26]. Он применяется для исследования взаимосвязи двух переменных, измеренных на одной и той же выборке. Рассмотренные выше методы автоматической оценки качества МП опираются исключительно на статистические инструменты и не используют знания об исходном и переводящем языках. Однако существуют метрики, которые привлекают и интегрируют лингвистические знания. Так, метрика METEOR [27] учитывает не только полные  $n$ -граммные совпадения в МП и РП (обычно чем их больше, тем система МП оценивается выше), но и использует корни слов вместо точных словоформ, найденных в текстах, а также данные о семантической близости зафиксированных в тексте словоформ с помощью лексической базы данных WordNet (см. <https://wordnet.princeton.edu>).

Главным недостатком METEOR считается высокая вычислительная сложность при поиске неполных совпадений и большее число параметров, требующих настройки, чем в представленных выше метриках [22].

В целом, способы интеграции лингвистических знаний в системы оценки качества МП и их преимущества и недостатки заслуживают отдельного обзора.

## 4 Заключение

В данной статье рассматривались подходы к экспертной и автоматической оценке качества МП. Сложность задачи оценки обусловлена тем, что единственно верного или идеального перевода не существует.

С помощью экспертных методов оценивают качество переводов относительно друг друга в терминах «лучше», «хуже», «одинакового качества» или производят количественные оценки МП на основе некоторых критериев и типологий ошибок. Меньше разногласий между экспертами обнаруживается при первом способе — ранжировании, так как даже при самом подробном описании типов ошибок остается возможность межэкспертной вариации оценок по шкале Ликерта.

Автоматическая оценка качества МП лишена таких недостатков, как дороговизна или низкая скорость. Ее релевантность проверяется через соответствие экспертной оценке. Автоматическая оценка предполагает в большинстве случаев сравнение машинного и профессионального (одного или нескольких) переводов. Метрики WER и PER анализируют количество изменений, которые необходимо внести, чтобы МП совпал с РП. Метрика BLEU основана на пословной  $n$ -граммной точности. Метрика ChrF использует в качестве единиц символы и придает большее значение  $n$ -граммной полноте<sup>1</sup>. Лингвистические знания при автоматической оценке качества, как правило, привлекаются ограниченно, так как их использование усложняет методику оценки.

---

<sup>1</sup>Однако значение коэффициента  $\beta$  может варьировать по мере того, как данная метрика будет проверена на парах языков с разной графикой.

Продолжением данной статьи может стать обзор других метрик МП (NIST, ROUGE, BEER), методов приблизительной оценки качества МП (Quality Estimation), обзор попыток использования МП для перевода художественной литературы и оценки качества их результатов, а также способов интеграции лингвистических знаний в системы оценки МП.

## Литература

1. Size machine translation market is \$250 million — TAUS publishes new market report, 2014. <https://www.taus.net/think-tank/news/press-release/size-machine-translation-market-is-250-million-taus-publishes-new-market-report>.
2. Och F. Breaking down the language barrier — six years in, 2012. <https://googleblog.blogspot.ie/2012/04/breaking-down-language-barriersix-years.html>.
3. Pichai S. Google I/O. Keynote speech, 2016. <https://events.google.com/io2016>.
4. Way A. Quality expectations of machine translation // Translation quality assessment: From principles to practice / Eds. J. Moorkens, Sh. Castilho, F. Gaspari, S. Doherty. — Cham, Switzerland: Springer, 2018. P. 159–178.
5. Reiss K., Vermeer H. J. Grundlegung einer allgemeinen Translationstheorie. — Tübingen, Deutschland: Walter de Gruyter, 1984. 245 p.
6. Castilho Sh., Doherty S, Gaspari F., Moorkens J. Approaches to human and machine translation quality assessment // Translation quality assessment: From principles to practice / Eds. J. Moorkens, Sh. Castilho, F. Gaspari, S. Doherty. — Cham, Switzerland: Springer, 2018. P. 9–38.
7. Княжева Е. А. О некоторых возможностях использования методов системного анализа в целях оценки качества перевода // Вестник ВГУ. Сер.: Лингвистика и межкультурная коммуникация, 2015. № 3. С. 113–119.
8. Federmann Ch. Appraise: An open-source toolkit for manual evaluation of machine translation output // Prague Bull. Math. Linguistics, 2012. Vol. 98. P. 25–35.
9. Koehn P. Simulating human judgment in machine translation evaluation campaigns // 9th Workshop (International) on Spoken Language Translation Proceedings. — Hong Kong, China: International Speech Communication Association, 2012. P. 179–184.
10. Hopkins M., May J. Models of translation competitions // 51st Annual Meeting of the Association for Computational Linguistics Proceedings. — Sofia, Bulgaria: Association for Computational Linguistics, 2013. Vol. 1. P. 1416–1424.
11. Sakaguchi K., Post M., Van Durme B. Efficient elicitation of annotations for human evaluation of machine translation // 9th Workshop on Statistical Machine Translation Proceedings. — Baltimore, MA, USA, 2014. P. 1–11.
12. Fleiss J. L., Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability // Educ. Psychol. Meas., 1973. Vol. 33. P. 613–619.
13. Kendall M. A new measure of rank correlation // Biometrika, 1938. Vol. 30. No. 1-2. P. 81–83.
14. Likert R. A technique for the measurement of attitudes // Arch. Psychol., 1932. Vol. 140. P. 1–55.

15. *Pierce J., Carroll J., Hamp E., Hays D., Hockett C., Oettinger A., Perlis A.* Language and machines — computers in translation and linguistics. — Washington, DC, USA: National Academy of Sciences, 1966. ALPAC report. 124 p.
16. *Lommel A.* Metrics for translation quality assessment: A case for standardising error typologies // Translation quality assessment: From principles to practice / Eds. J. Moorkens, Sh. Castilho, F. Gaspari, S. Doherty. — Cham, Switzerland: Springer, 2018. P. 109–127.
17. *Popović M.* Error classification and analysis for machine translation quality assessment // Translation quality assessment: From principles to practice / Eds. J. Moorkens, Sh. Castilho, F. Gaspari, S. Doherty. — Cham, Switzerland: Springer, 2018. P. 129–158.
18. Multidimensional Quality Metrics (MQM) issue types / Eds. A. Lommel, A. Burchardt, A. Görög, H. Uszkoreit, A. Melby. — 2015. <http://www.qt21.eu/mqm-definition/issues-list-2015-12-30.html>.
19. *Popović M.* Herson: An open source tool for automatic error classification of machine translation output // Prague Bull. Math. Linguistics, 2011. Vol. 96. P. 59–68.
20. *Левенштейн В. И.* Двоичные коды с исправлением выпадений, вставок и замещений символов // Докл. Акад. наук СССР, 1965. Т. 163. № 4. С. 845–848.
21. *Tillmann C., Vogel S., Ney H., Sawaf H., Zubiaga A.* Accelerated DP-based search for statistical translation // 5th European Conference on Speech Communication and Technology Proceedings. — Rhodes, Greece, 1997. P. 2667–2670.
22. *Koehn P.* Statistical machine translation. — Cambridge: Cambridge University Press, 2010. 462 p.
23. *Papineni K., Roukos S., Ward T., Zhu W. J.* BLEU: A method for automatic evaluation of machine translation // 40th Annual Meeting on Association for Computational Linguistics Proceedings. — Philadelphia, PA, USA: Association for Computational Linguistics, 2002. P. 311–318.
24. *Popović M.* CHRF: Character n-gram F-score for automatic MT evaluation // 10th Workshop on Statistical Machine Translation Proceedings. — Lisboa, Portugal: Association for Computational Linguistics, 2015. P. 392–395.
25. *Popović M.* CHRF deconstructed:  $\beta$  parameters and n-gram weights // 1st Conference on Machine Translation Proceedings. — Berlin, Germany: Association for Computational Linguistics, 2016. Vol. 2. P. 499–504.
26. *Pearson K.* Notes on regression and inheritance in the case of two parents // P. R. Soc. London, 1895. Vol. 58. Iss. 347–352. P. 240–242.
27. *Banerjee S., Lavie A.* METEOR: An automatic metric for MT evaluation with improved correlation with human judgments // Workshop on Intrinsic and Extrinsic Evaluation Measures for MT and/or Summarization at the 43rd Annual Meeting of the Association of Computational Linguistics Proceedings. — Ann Arbor, MI, USA: Association of Computational Linguistics, 2005. P. 65–72.

Поступила в редакцию 30.08.19

## ON METHODS OF MACHINE TRANSLATION QUALITY ASSESSMENT

A. K. Rychikhin

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

**Abstract:** The article discusses approaches to determining the quality of machine translation (MT) and several methods of translation quality assessment. The aim of the article is to review a number of methods and approaches to human and automatic assessment of MT quality. The first part of the article describes the methods of relative human evaluation (ranking of translations) and absolute evaluation based on penalties for errors in translation, as well as software and algorithms that simplify human assessment. Most attention is paid to the DQF/MQM (Dynamic Quality Framework/Multidimensional Quality Metrics) error typology which is not aimed at a limited subject area as the most flexible one. The second part of the article is devoted to a review of metrics for automatic quality assessment of MT that do not use linguistic data as well as the correlation coefficients of human and automatic evaluation.

**Keywords:** machine translation (MT); machine translation quality; rankings of translations and translation systems; MT quality metrics; error typologies; translation quality assessment; human evaluation of MT quality

**DOI:** 10.14357/08696527190410

### Acknowledgments

The author acknowledges the assistance of V. A. Nuriev and I. M. Zatsman in the preparation of the article.

### References

1. Size machine translation market is \$250 Million — TAUS publishes new market report. 2014. Available at: <https://www.taus.net/think-tank/news/press-release/size-machine-translation-market-is-250-million-taus-publishes-new-market-report> (accessed August 1, 2019).
2. Och, F. 2012. Breaking down the language barrier — six years in. Available at: <https://googleblog.blogspot.ie/2012/04/breaking-down-language-barriersix-years.html> (accessed August 1, 2019).
3. Pichai, S. 2016. Google I/O. Keynote speech. Available at: <https://events.google.com/io2016/> (accessed August 1, 2019).
4. Way, A. 2018. Quality expectations of machine translation. *Translation quality assessment: From principles to practice*. Eds. J. Moorkens, Sh. Castilho, F. Gaspari, and S. Doherty. Cham, Switzerland: Springer. 159–178.

5. Reiss, K., and H. J. Vermeer. 1984. *Grundlegung einer allgemeinen Translationstheorie*. Tübingen, Deutschland: Walter de Gruyter. 245 p.
6. Castilho, Sh., S. Doherty, F. Gaspari, and J. Moorkens. 2018. Approaches to human and machine translation quality assessment. *Translation quality assessment: From principles to practice*. Eds. J. Moorkens, Sh. Castilho, F. Gaspari, and S. Doherty. Cham, Switzerland: Springer. 9–38.
7. Knyazheva, E. A. 2015. O nekotorykh vozmozhnostyakh ispol'zovaniya metodov sistemnogo analiza v tselyakh otsenki kachestva perevoda [Application of system analysis methods to translation quality evaluation]. *Proceedings of Voronezh State University. Ser. Linguistics and Intercultural Communication* 3:113–119.
8. Federmann, Ch. 2012. Appraise: An open-source toolkit for manual evaluation of machine translation output. *Prague Bull. Math. Linguistics* 98:25–35.
9. Koehn, P. 2012. Simulating human judgment in machine translation evaluation campaigns. *9th Workshop (International) on Spoken Language Translation Proceedings*. Hong Kong, China: International Speech Communication Association. 179–184.
10. Hopkins, M., and J. May. 2013. Models of translation competitions. *51st Annual Meeting of the Association for Computational Linguistics Proceedings*. Sofia, Bulgaria: Association for Computational Linguistics. 1:1416–1424.
11. Sakaguchi, K., M. Post, and B. Van Durme. 2014. Efficient elicitation of annotations for human evaluation of machine translation. *9th Workshop on Statistical Machine Translation Proceedings*. Baltimore, MA. 1–11.
12. Fleiss, J. L., and J. Cohen. 1973 The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educ. Psychol. Meas.* 33:613–619.
13. Kendall, M. 1938. A new measure of rank correlation. *Biometrika* 30(1-2):81–83.
14. Likert, R. 1932. A technique for the measurement of attitudes. *Arch. Psychol.* 14:1–55.
15. Pierce, J., J. Carroll, E. Hamp, D. Hays, C. Hockett, A. Oettinger, and A. Perlis. 1966. *Language and machines — computers in translation and linguistics*. Washington, DC: National Academy of Sciences. ALPAC report. 124 p.
16. Lommel, A. 2018. Metrics for translation quality assessment: A case for standardising error typologies. *Translation quality assessment: From principles to practice*. Eds. J. Moorkens, Sh. Castilho, F. Gaspari, and S. Doherty. Cham, Switzerland: Springer. 109–127.
17. Popović, M. 2018. Error classification and analysis for machine translation quality assessment. *Translation quality assessment: From principles to practice*. Eds. J. Moorkens, Sh. Castilho, F. Gaspari, and S. Doherty. Cham, Switzerland: Springer. 129–158.
18. Lommel, A., A. Burchardt, A. Görög, H. Uszkoreit, and A. Melby, eds. 2015. Multidimensional Quality Metrics (MQM) issue types. Available at: <http://www.qt21.eu/mqm-definition/issues-list-2015-12-30.html> (accessed August 5, 2019).
19. Popović, M. 2011. Hjerson: An open source tool for automatic error classification of machine translation output. *Prague Bull. Math. Linguistics* 96:59–68.
20. Levenshtein, V. I. 1965. Dvoichnye kody s ispravleniem vypadeniy, vstavok i zameshchenii simvolov [Binary codes capable of correcting deletions, insertions, and reversals]. *Dokl. Akad. Nauk SSSR [Sov. Math. Dokl.]* 163(4):845–848.

21. Tillmann, C., S. Vogel, H. Ney, H. Sawaf, and A. Zubiaga. 1997. Accelerated DP-based search for statistical translation. *5th European Conference on Speech Communication and Technology Proceedings*. Rhodes, Greece. 2667–2670.
22. Koehn, P. 2010. *Statistical machine translation*. Cambridge: Cambridge University Press. 462 p.
23. Papineni, K., S. Roukos, T. Ward, and W.J. Zhu. 2002. BLEU: A method for automatic evaluation of machine translation. *40th Annual Meeting on Association for Computational Linguistics Proceedings*. Philadelphia, PA: Association for Computational Linguistics. 311–318.
24. Popović, M. 2015. CHRf: Character n-gram F-score for automatic MT evaluation. *10th Workshop on Statistical Machine Translation Proceedings*. Lisboa, Portugal: Association for Computational Linguistics. 392–395.
25. Popović, M. CHRf deconstructed:  $\beta$  parameters and n-gram weights. *1st Conference on Machine Translation Proceedings*. Berlin, Germany: Association for Computational Linguistics, 2016. 2:499–504.
26. Pearson, K. 1895. Notes on regression and inheritance in the case of two parents. *P. R. Soc. London* 58(347-352):240–242.
27. Banerjee, S., and A. Lavie. 2005. METEOR: An automatic metric for MT evaluation with improved correlation with human judgments. *Workshop on Intrinsic and Extrinsic Evaluation Measures for MT and/or Summarization at the 43rd Annual Meeting of the Association of Computational Linguistics Proceedings*. Ann Arbor, MI: Association of Computational Linguistics. 65–72.

*Received August 30, 2019*

## **Contributor**

**Rychikhin Aleksei K.** (b. 1995) — engineer, 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; ak.rychikhin@gmail.com.

## ИССЛЕДОВАНИЕ ПРОЦЕССА ФОРМИРОВАНИЯ ВИРТУАЛЬНОГО СООБЩЕСТВА КАК НЕЛИНЕЙНОЙ ДИНАМИЧЕСКОЙ СИСТЕМЫ

*И. М. Адамович<sup>1</sup>, О. И. Волков<sup>2</sup>*

**Аннотация:** Статья продолжает серию работ, посвященных описанию и анализу распределенной технологии поддержки конкретно-исторических исследований, основанной на принципах краудсорсинга. Данная статья посвящена описанию и обоснованию подхода к моделированию стадии формирования сообщества исследователей и оценке перспектив развертывания проекта в зависимости от объема начального наполнения семантической сети, лежащей в основе технологии. Предложенный подход заключается в анализе процессов изменения числа участников сообщества исследователей и объема семантической сети технологии как нелинейной динамической системы. Анализ производился для различных начальных условий. С помощью данного подхода была получена оценка уровня начального наполнения семантической сети, при котором обеспечивается устойчивый рост сообщества. Проверка адекватности модели была осуществлена за счет сравнения характера роста сообщества, определенного с помощью моделирования, с эмпирическим графиком изменения числа участников сетевых сообществ.

**Ключевые слова:** семантическая сеть; модель; технология; динамическая система; конкретно-историческое исследование

**DOI:** 10.14357/08696527190411

### 1 Введение

Поддержка конкретно-исторических исследований становится одной из актуальных задач современности в связи с вовлечением в исследовательский процесс не только членов профессионального исторического сообщества, но и самых широких слоев непрофессионалов в связи со все увеличивающимся интересом к частной, семейной истории [1].

В [2, 3] описана разработанная в ФИЦ ИУ РАН распределенная технология поддержки конкретно-исторических исследований, основанная на принципах краудсорсинга (мобилизации ресурсов широкого круга добровольцев посредством информационных технологий). Данные технологии организованы в форме семантической сети. Целесообразность данной формы организации данных бы-

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

<sup>3</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, Volkov@amsd.com



ла обоснована в [4]. Также были проведены проверки эффективности данной технологии с точки зрения качества данных и их информационной доступности [5–8]. Было показано, что неизбежные ошибки данных, возникающие в процессе деятельности множества не связанных между собой участников, организованных в виртуальное сообщество исследователей, не приводят к значимому снижению показателей точности и полноты информационного поиска. Но моделирование функционирования технологии относилось исключительно к стационарной фазе развития виртуального сообщества исследователей [9]. Фаза ее формирования не рассматривалась, хотя именно от нее зависит [10], будет ли технология оценена пользователями. Несмотря на то что круг потенциальных пользователей технологии очень широк, они все фактически уже как-то решают свои исследовательские и коммуникационные задачи, и для их перехода на рассматриваемую технологию нужен стимул. Поскольку проект поддержки конкретно-исторических исследований на базе технологии является некоммерческим, единственным стимулом для присоединения к проекту могут послужить дополнительные исследовательские возможности, которые пользователь получит от новой технологии. Данные возможности помимо развитого инструментария [11, 12] определяются возможностями поиска и сопоставления исторических фактов, которые, в свою очередь, зависят от объема семантической сети. Поскольку семантическая сеть технологии создается самими пользователями на принципах краудсорсинга, ее объем на начальном этапе проекта, когда число пользователей мало, также неизбежно невелик. Таким образом, существует опасение, что при малом начальном наполнении семантической сети проекта привлечь в него пользователей не удастся вообще.

Цель настоящей статьи — исследование роста виртуального сообщества пользователей технологии поддержки конкретно-исторических исследований в зависимости от начального объема семантической сети.

## 2 Описание подхода

Опишем динамику виртуального сообщества как динамическую систему.

Будем полагать, что исходно виртуальное сообщество представляет собой группу энтузиастов, сформировавших к моменту представления технологии широкому кругу потенциальных пользователей некоторый начальный объем семантической сети. Число членов исходной группы энтузиастов обозначим как  $n_0$ , начальный объем семантической сети как  $v_0$ , общее число потенциальных пользователей технологии как  $N$ , объем семантической сети в момент времени  $t$  как  $v(t)$ , а текущее число пользователей как  $n(t)$ . Задача состоит в исследовании роста функций  $v$  и  $n$  в зависимости от начальных значений  $v_0$  и  $n_0$ .

На динамику  $v$  и  $n$  влияют следующие факторы:

- (1) прирост семантической сети пропорционален числу участников  $n(t)$  и средней производительности одного участника, которую будем считать константой и обозначим как  $\varepsilon$ ;



- (2) скорость притока новых участников пропорциональна вероятности перехода на использование технологии для одного потенциального участника и числу еще не перешедших потенциальных участников. Будем считать, что потенциальный пользователь тем скорее перейдет на использование новой технологии, чем больше пользователей из его окружения уже осуществили такой переход. Если объем окружения считать константой, то вероятность перехода пропорциональна доле уже перешедших пользователей  $n(t)/N$ . Коэффициент пропорциональности обозначим как  $\alpha$ ;
- (3) помимо притока пользователей неизбежно будет наблюдаться и отток разочаровавшихся в возможностях технологии, которые определяются прежде всего объемом семантической сети. Будем считать, что участники, принявшие решение об уходе из сообщества пользователей технологии, не изменяют свое решение длительное время и перестают быть потенциальными пользователями. Таким образом, если число участников, покинувших сообщество, обозначить как  $m(t)$ , текущее число еще не перешедших потенциальных участников будет выражаться формулой  $N - n(t) - m(t)$ .

Скорость оттока участников пропорциональна числу участников и вероятности ухода участника. Коэффициент пропорциональности обозначим как  $\beta$ . Вероятность ухода высока при малом объеме семантической сети и стремится к нулю по мере ее роста. Такая зависимость по аналогии с формулой Моно [13], отражающей зависимость скорости потребления микроорганизмом субстрата от его концентрации, может быть описана функцией  $e^{-\gamma v(t)}$ , где  $\gamma$  — постоянный коэффициент, отвечающий за скорость убывания вероятности ухода по мере роста семантической сети;

- (4) число пользователей ограничено значением  $N$ . Предельное значение для числа участников может быть описано уравнением Ферхюльста (логистическим уравнением) [14]

$$\dot{n} = \delta n \left( 1 - \frac{n}{N} \right),$$

обеспечивающим ограничение роста популяции  $n(t)$  предельным значением  $N$  (потенциальная емкость системы). Параметр  $\delta$  характеризует скорость роста популяции.

Результирующая система дифференциальных уравнений, описывающих поведение системы, выглядит следующим образом:

$$\begin{cases} \frac{dn}{dt} = \alpha \frac{n(t)}{N} (N - n(t) - m(t)) - \beta e^{-\gamma v(t)} n(t) + \delta n(t) \left( 1 - \frac{n(t)}{N} \right); \\ \frac{dv}{dt} = \varepsilon n(t); \\ \frac{dm}{dt} = \beta e^{-\gamma v(t)} n(t). \end{cases}$$

### 3 Анализ системы

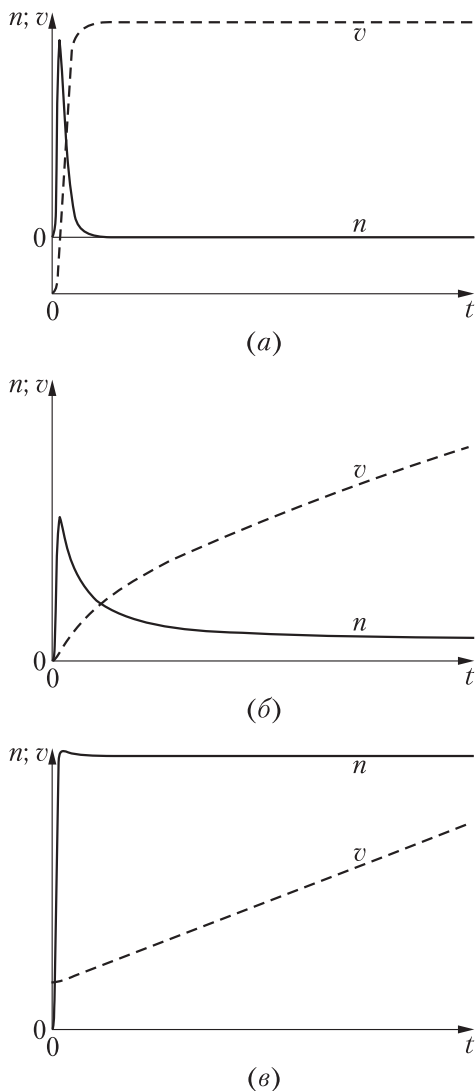
Полученная система дифференциальных уравнений анализировалась численно с помощью пакета MatLab. Анализ производился для трех различных начальных условий  $(n_0, v_0, m_0)$ , отличающихся значениями  $v_0$ : 300, 500 и 1000 узлов семантической сети соответственно. Значение  $n_0$  во всех трех случаях было принято равным 50 участников, а  $m_0$  равным нулю, поскольку отток участников изначально отсутствует.

На рис. 1 приведен вид решений системы для функций  $n(t)$  и  $v(t)$ . На рис. 2 приведен фазовый портрет системы в пространстве  $(n, v)$  для тех же начальных условий.

Рисунок 1, *а* подтверждает опасение, что при малом начальном наполнении семантической сети проекта привлечь в него пользователей не удастся вообще. После недолгого подъема число участников начинает падать до нуля и более уже не меняется. Рисунок 1, *б* демонстрирует, что при достаточном начальном наполнении семантической сети практически все потенциальные участники примыкают к сообществу пользователей технологии и объем сообщества быстро стабилизируется на значении, близком к  $N$ . Рисунок 1, *в*, соответствующий среднему начальному наполнению семантической сети, демонстрирует стабилизацию объема сообщества на некотором промежуточном значении.

### 4 Адекватность модели

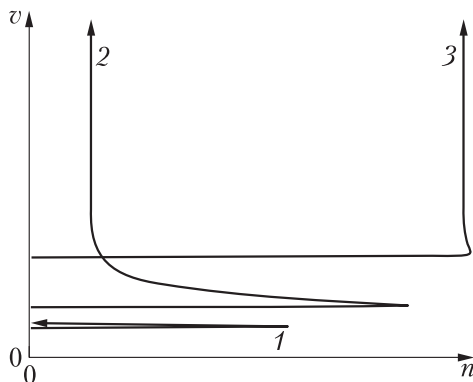
На рис. 3 представлен эмпирический график изменения числа участников сетевых сообществ, приведенный в статье [15], посвященной анализу стадий



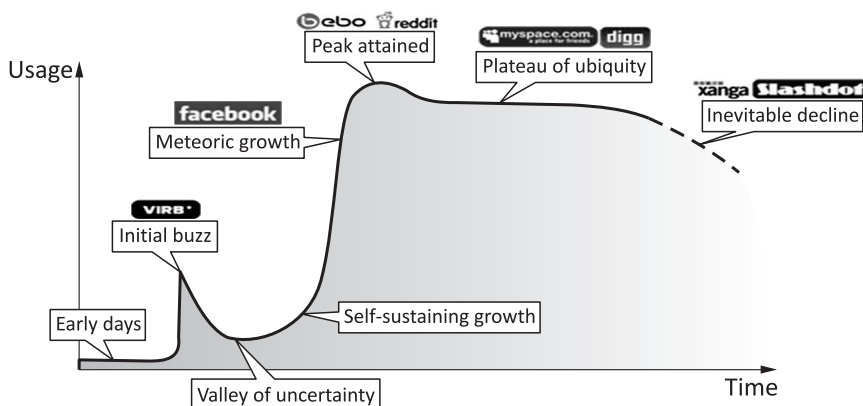
**Рис. 1** Вид решений системы при  $v_0 = 300$  (а); 500 (б) и 1000 (в) узлов

развития социальных сетей и моделированию роста сетевых сообществ. На графике выделены следующие типичные стадии развития сообществ:

- (1) Early days (ранние времена) — стадия, поддерживаемая исключительно создателями проекта;
- (2) Initial Buzz (первоначальный слух) — стадия, когда за счет первичного распространения информации о проекте в сообщество приходят энтузиасты, движимые прежде всего любопытством;
- (3) Valley of Uncertainty (долина неопределенности) — стадия, когда поток энтузиастов схлынул, но сообщество успело получить от их деятельности толчок к развитию;
- (4) Self-Sustaining Growth (самоподдерживающийся рост) — первая фаза стадии развития, определяемой притоком пользователей-прагматиков;
- (5) Meteoric Growth (стремительный рост) — вторая фаза стадии развития, определяемой притоком пользователей-прагматиков;
- (6) Peak Attained (достигнут пик) — приток пользователей-прагматиков достиг максимума;



**Рис. 2** Фазовый портрет системы: 1 —  $v_0 = 300$  узлов; 2 — 500; 3 —  $v_0 = 1000$  узлов



**Рис. 3** График изменения числа участников и стадии развития сетевых сообществ [15]

- (7) Plateau of Ubiquity (плато распространенности) — стационарная фаза развития сообщества;
- (8) Inevitable Decline (неизбежное падение) — стадия упадка проекта.

Поскольку технология поддержки конкретно-исторических исследований включает в себя онлайн-платформу для коммуникации и совместной деятельности исследователей, а сообщество пользователей технологии обладает всеми признаками научно-профессионального сетевого сообщества [16], следует ожидать, что динамика числа участников будет подчиняться тому же закону.

Предметом настоящего исследования было развитие, определяемое исключительно пользователями-прагматиками, пришедшими в сообщество исключительно с целью решения своих исследовательских задач и научной коммуникации. Поэтому исследуемый период соответствует стадиям 4–7 (Self-Sustaining Growth, Meteoric Growth, Peak Attained и Plateau of Ubiquity) развития сетевых сообществ.

Сравнение формы графиков функции  $n(t)$ , приведенных на рис. 1, и формы эмпирического графика, приведенного на рис. 3, ограниченного стадиями 4–7, показывает их полную идентичность, что подтверждает адекватность выбранной модели.

## **5 Выводы**

Проведенное исследование показало, что предположение о наличии явной зависимости перспектив проекта от объема начального наполнения семантической сети оказалось верным и при малом начальном наполнении семантической сети проекта привлечь в него пользователей не удастся вообще. Но менее чем двукратное увеличение начального объема приводит к устойчивому развитию проекта, а трехкратное увеличение — к вовлечению в проект почти всех (более 90%) потенциальных пользователей технологии.

Данный вывод подтверждает жизнеспособность и перспективность предложенной технологии поддержки конкретно-исторических исследований.

## **Литература**

1. *Грибач С. В.* Исследование семейных кризисов посредством психолингвистического эксперимента // Сборники конференций НИЦ Социосфера, 2010. № 6. С. 45–54.
2. *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.
3. *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.

4. *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.
5. *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.
6. *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.
7. *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.
8. *Адамович И. М., Волков О. И.* Процесс коррекции ошибок семантической сети как нелинейная динамическая система // Системы и средства информатики, 2019. Т. 29. № 3. С. 160–168. doi: 10.14357/08696527190314.
9. *Сазанов В. М.* Сравнительный анализ социально-сетевых проектов // Социально-ориентированные сети и технологии: Исследовательский ресурс социального интернета, 2009. <http://v-school.narod.ru/PAPERS/analiz.doc>.
10. *Пономарев С. В.* Краудсорсинг — технология создания виртуальных сообществ // Вестник Пермского государственного университета. Сер.: Политология, 2011. № 3(15). С. 107–121.
11. *Адамович И. М., Волков О. И.* Система извлечения биографических фактов из текстов исторической направленности // Системы и средства информатики, 2015. Т. 25. № 3. С. 235–250. doi: 10.14357/08696527150315.
12. *Маркова Н. А.* Технология поддержки конкретно-исторических исследований на основе модели фактоподобных высказываний // Программная инженерия, 2015. № 5. С. 43–48.
13. *Monod J.* Recherches sur la croissance des cultures bacteriennes. — Paris: Hermann et Cie, Editeurs, 1942. 210 p.
14. *Братусь А. С., Новожилов А. С., Платонов А. П.* Динамические системы и модели биологии. — М.: Физматлит, 2010. 400 с.
15. *Brown S.* The Ebb and flow of social networking, 2007. <https://www.wisdump.com/web/the-ebb-and-flow-of-social-networking>.
16. *Каменский Е. Г., Гримов О. А.* Сетевые сообщества в социальных сетях как фактор развития личностной субъектности // Вестник Нижегородского университета им. Н. И. Лобачевского. Сер. Социальные науки, 2014. № 2(34). С. 62–67.

Поступила в редакцию 22.07.19

## THE FORMING OF VIRTUAL COMMUNITIES PROCESS AS A NONLINEAR DYNAMIC SYSTEM

*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 119133, Russian Federation

**Abstract:** The article continues a series of works devoted to the description and analysis of distributed technology of concrete historical investigation support based on the principles of crowdsourcing. This article is devoted to the description and substantiation of the approach to modeling the stage of formation of the research community and evaluation of the prospects of project deployment depending on the volume of the initial content of the semantic network underlying the technology. The proposed approach is to analyze the processes of changing the number of participants in the research community and the volume of the semantic network of technology as a nonlinear dynamic system. The analysis was performed for different initial conditions. With the help of this approach, the estimation of the level of the initial filling of the semantic network was obtained at which the steady growth of the community is provided. The adequacy of the model was checked by comparing the character of community growth determined by the simulation with the empirical graph of changes in the number of participants in network communities.

**Keywords:** semantic net; model; technology; dynamic systems; specific historical investigation

**DOI:** 10.14357/08696527190411

### References

1. Gribach, S. V. 2010. Issledovanie semeynykh krizisov posredstvom psikholingvisticheskogo eksperimenta [The study of family crises through a psycholinguistic experiment]. *Sborniki konferentsiy NITs Sotsiosfera* [Conference Proceedings NIC Sociosfera] 6:45–54.
2. Adamovich, I. M., and O. I. Volkov. 2016. Tekhnologiya raspredelenного avtomatizirovannogo analiza istoricheskikh tekstov [The distributed automated technology of historical texts analysis]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 26(3):148–161. doi: 10.14357/08696527160311.
3. 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* 29(1):194–205. doi: 10.14357/08696527190116.
4. 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* 29(2):161–171. doi: 10.14357/08696527190214.
5. Adamovich, I. M., and O. I. Volkov. 2017. Ob odnom podkhode k modelirovaniyu protsessa razvitiya semanticheskoy seti [On some approach to modelling of semantic net evolution]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 27(2):143–154. doi: 10.14357/08696527170212.
  6. 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* 27(4):150–163. doi: 10.14357/08696527170412.
  7. 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* 28(1):65–76. doi: 10.14357/08696527180105.
  8. Adamovich, I. M., and O. I. Volkov. 2019. Issledovanie protsessa korrektsii oshibok v semanticheskoy seti kak nelineynoy dinamicheskoy sistemy [The error correction process in the semantic network as a nonlinear dynamic system]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(3):160–168. doi: 10.14357/08696527190314.
  9. Sazanov, V. M. 2009. Sravnitel'nyy analiz sotsial'no-setevykh proektov [Comparative analysis of social-network projects]. *Sotsial'no-orientirovannyye seti i tekhnologii. Issledovatel'skiy resurs sotsial'nogo interneta* [Socially-oriented networks and technologies. Research resource of Social Internet]. Available at: <http://v-school.narod.ru/PAPERS/analiz.doc> (accessed July 12, 2019).
  10. Ponomarev, S. V. 2011. Kraudsorsing — tekhnologiya sozdaniya virtual'nykh soobshchestv [Crowdsourcing as a technology of network communities formation]. *Vestnik Permskogo universiteta. Ser. Politologiya* [Perm University Herald. Political science ser.] 15(3):107–121.
  11. Adamovich, I. M., and O. I. Volkov. 2015. Sistema izvlecheniya biograficheskikh faktov iz tekstov istoricheskoy napravlenosti [The system of facts extraction from historical texts]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 25(2):235–250. doi: 10.14357/08696527150315.
  12. Markova, N. A. 2015. Tekhnologiya podderzhki konkretno-istoricheskikh issledovaniy na osnove modeli faktopodobnykh vyskazyvaniy [Support technology for specific historical studies on the base of fact-like propositions model]. *Programmnyaya inzheneriya* [Software Engineering] 5:43–48.
  13. Monod, J. 1942. *Recherches sur la croissance des cultures bacteriennes*. Paris: Hermann et Cie, Editeurs. 210 p.
  14. Bratus, A. S., A. S. Novozhilov, and A. P. Platonov. 2010. *Dinamicheskie sistemy i modeli biologii* [Dynamical systems and models in biology]. Moscow: Fizmatlit. 400 p.
  15. Stuart Brown. 2007. The Ebb and flow of social networking. Available at: <https://www.wisdump.com/web/the-ebb-and-flow-of-social-networking/> (accessed July 12, 2019).
  16. Kamensky, E. G., and O. A. Grimov. 2014. Setevye soobshchestva v social'nykh setyakh kak faktor razvitiya lichnostnoy sub"ektnosti [Online communities at social

networking services as a factor of personal subjectness development]. *Vestnik Nizhegorodskogo universiteta im. N. I. Lobachevskogo. Ser. social'nye nauki* [Bulletin of Lobachevsky State University of Nizhny Novgorod. Ser. Social Sciences] 34(2):62–67.

*Received July 22, 2019*

## **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



## ФУНКЦИОНАЛЬНАЯ МОДЕЛЬ СИСТЕМЫ УПРАВЛЕНИЯ ПРОЦЕССАМИ СТРАТЕГИЧЕСКОГО ПЛАНИРОВАНИЯ

*А. П. Сучков*<sup>1</sup>

**Аннотация:** Рассматриваются подходы к созданию процессной модели системы управления в области стратегического планирования. С точки зрения научного подхода методы стратегического планирования являются разделом теории управления сложными системами. Ключевым является понимание стратегического планирования как циклической последовательности взаимосвязанных процессов, в числе которых можно выделить два полных и взаимосвязанных цикла управления, которые будем условно называть «планирование» и «реализация». При этом цикл управления «реализация» явно носит ситуационный характер. В его рамках осуществляется постоянный мониторинг хода выполнения стратегического плана, ситуационный анализ потоков событий контролируемого пространства, текущее реагирование на изменение обстановки вплоть до корректировки документов стратегического планирования.

**Ключевые слова:** стратегическое планирование; процессный подход; процессная модель; система управления

**DOI:** 10.14357/08696527190412

### 1 Введение

В настоящее время происходит активное использование методов стратегического планирования на всех уровнях управления государством. Это прежде всего обусловлено политической волей высшего руководства страны, выраженной в последние годы в виде определения конкретных среднесрочных и долгосрочных целей по всем основным областям народного хозяйства, задаваемых в ежегодных посланиях Президента РФ Федеральному собранию и указами Президента РФ.

В соответствии с Федеральным законом от 28 июня 2014 г. № 172-ФЗ «О стратегическом планировании в Российской Федерации» стратегическое планирование — это деятельность участников стратегического планирования по целеполаганию, прогнозированию, планированию и программированию социально-экономического развития (СЭР) РФ, субъектов РФ и муниципальных образований, отраслей экономики и сфер государственного и муниципального управления, обеспечения национальной безопасности РФ. С точки зрения научного подхода методы стратегического планирования являются разделом теории управления сложными системами. Обычно, с разными вариациями назва-

---

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

ний, выделяют следующие этапы: сверхзадача (миссия) – целеполагание – мониторинг – анализ – выбор стратегии – реализация стратегии – контроль (в частности, с образованием обратной связи на этапах анализа и контроля). Ключевым здесь является понимание стратегического планирования как циклической последовательности взаимосвязанных процессов, в числе которых можно выделить два полных и взаимосвязанных цикла управления, которые будем условно называть «планирование» и «реализация» [1]. При этом цикл управления «реализация» явно носит ситуационный характер. В его рамках осуществляется постоянный мониторинг хода выполнения стратегического плана, ситуационный анализ потоков событий контролируемого пространства, текущее реагирование на изменение обстановки вплоть до корректировки документов стратегического планирования.

Законченных теорий и методов стратегического планирования пока не разработано, что подтверждается активными дискуссионными обсуждениями данной проблематики в научном сообществе [2]. Одной из основных проблем представляется различный масштаб времени стратегического целеполагания и процессов текущего управления процессами реализации целей стратегического планирования, что вызывает необходимость сложного согласования процессов управления [1, 3]. Решение этой проблемы возможно путем создания и обоснования информационной процессной модели системы стратегического планирования, содержащей два полных и взаимосвязанных цикла управления с выделением процессов ситуационного управления [2]. Конкретизации такой процессной модели и посвящена данная статья.

## 2 Основные составляющие информационной модели

Основными составляющими обсуждаемой функциональной модели являются субъекты, объекты, документы и процессы стратегического планирования, определяемые законом № 172-ФЗ.

Хотелось бы подчеркнуть, что нормативно-правовые акты РФ одним из основных субъектов стратегического планирования определяют Совет безопасности РФ. С этой целью Совбез реализует функцию осуществления стратегического планирования в области обеспечения национальной безопасности, в том числе:

- оценку эффективности работы по реализации документов стратегического планирования;
- корректировку документов стратегического планирования и стратегических национальных приоритетов;
- координацию деятельности по разработке в федеральных округах документов стратегического планирования.

Другим важным субъектом стратегического планирования является Российская академия наук. В соответствии с Федеральным законом от 27.09.2013 № 253-ФЗ «О Российской академии наук...» она осуществляет свою деятель-

ность в целях *экспертного научного обеспечения деятельности органов государственной власти*, в частности в процессах анализа, прогнозирования и оценки угроз национальной безопасности.

Объектами стратегического планирования являются параметры СЭР РФ, субъектов РФ и муниципальных образований, отраслей экономики и сфер государственного и муниципального управления, обеспечения национальной безопасности РФ. Одними из основных объектов стратегического планирования являются система целевых показателей и угрозы национальной безопасности.

Документы стратегического планирования представляют собой иерархическую систему, отражающую вертикальное разделение органов власти и его стадийность (см. таблицу). Хотелось бы отметить большую наукоемкость содержания этих документов.

### **3 Процессы стратегического планирования**

Итак, представляется целесообразным выделить два полных и взаимосвязанных цикла управления, которые можно условно назвать «планирование» и «реализация» (см. рис. 1, где пунктиром обозначены обратные связи).

Представленная информационная процессная функциональная модель системы управления стратегическим планированием (СУСП) исходит из того, что последовательность процессов как в рамках цикла «планирование», так и цикла «реализация» включает стандартные стадии управления: целеполагание, мониторинг, анализ, принятие решений и их реализацию [4].

Главным основанием целесообразности функционирования СУСП представляется факт наличия *процессов стратегического целеполагания и планирования*, т. е. определения направлений, целей и приоритетов обеспечения СЭР РФ, субъектов РФ и муниципальных образований, отраслей экономики и сфер государственного и муниципального управления, обеспечения национальной безопасности РФ и их реализации в виде документов стратегического планирования. В блок «планирование» входят следующие группы процессов.

*Сверхзадача* — определение высшими должностными лицами государства основных среднесрочных и долгосрочных целей по СЭР РФ. Цели являются интегральными, измеримыми и привязанными ко времени реализации, они формулируются в основополагающих документах стратегического планирования — ежегодных посланиях Президента РФ Федеральному собранию и указах Президента РФ.

*Стратегическое целеполагание* включает формирование единой иерархической системы целевых показателей в двух измерениях — отраслевом и территориальном, а также обоснование основных документов стратегического планирования:

- Стратегия СЭР РФ;
- Стратегия национальной безопасности РФ;

Система документов стратегического планирования в соответствии с законом № 172-ФЗ

Этапы	Уровни		
	Федеральный	Региональный	Муниципальный
Целеполагание	<p><i>Базовые:</i>                      (а) ежегодное послание Президента РФ Федеральному собранию РФ;                      (б) стратегия СЭР РФ;                      (в) стратегия национальной безопасности РФ;                      (г) стратегия научно-технологического развития РФ</p> <p><i>Конкретизирующие:</i>                      (а) отраслевые документы стратегического планирования РФ;                      (б) стратегия пространственного развития РФ;                      (в) стратегии СЭР макро-регионов</p>	Стратегия СЭР субъекта РФ	Стратегия СЭР муниципального образования
Прогнозирование	(а) прогноз научно-технологического развития РФ; (б) стратегический прогноз РФ; (в) прогноз СЭР РФ на долгосрочный период; (г) бюджетный прогноз РФ на долгосрочный период; (д) прогноз СЭР РФ на среднесрочный период	(а) прогноз СЭР субъекта РФ на долгосрочный период; (б) бюджетный прогноз субъекта РФ на долгосрочный период; (в) прогноз СЭР субъекта РФ на среднесрочный период	(а) прогноз СЭР муниципального образования на среднесрочный или долгосрочный период; (б) бюджетный прогноз муниципального образования на долгосрочный период
Планирование и программирование	(а) основные направления деятельности Правительства РФ; (б) государственные программы РФ; (в) государственная программа вооружения; (г) схемы территориального планирования РФ; (д) планы деятельности ФОИВ	(а) план мероприятий по реализации стратегии СЭР субъекта РФ; (б) государственные программы субъекта РФ; (в) схема территориального планирования двух и более субъектов РФ, схема территориального планирования субъекта РФ	(а) план мероприятий по реализации стратегии СЭР муниципального образования; (б) муниципальные программы

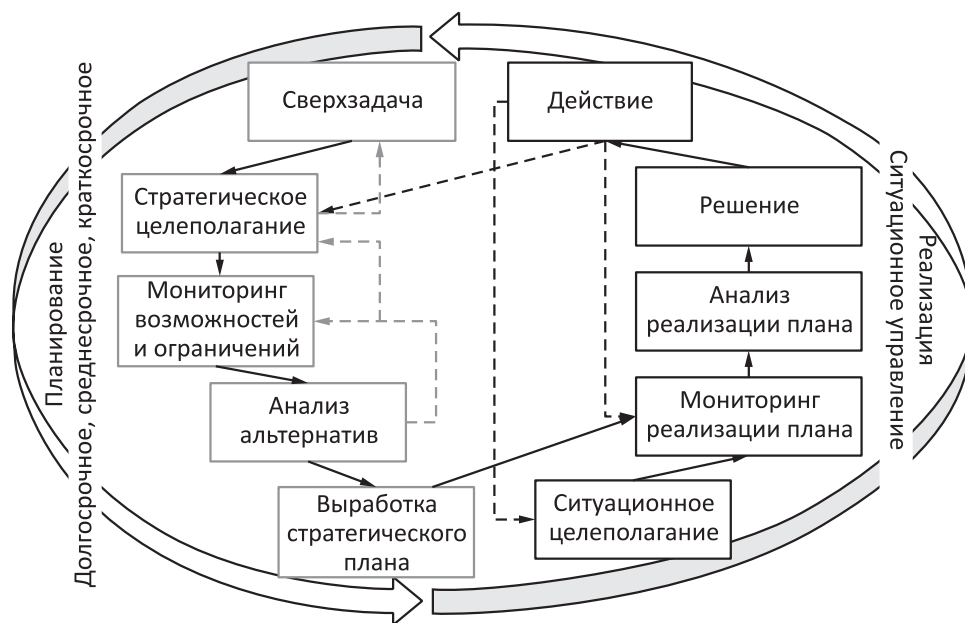


Рис. 1 Структура процессов стратегического планирования

- Стратегия научно-технологического развития РФ;
- Стратегия пространственного и инфраструктурного развития.

Эта группа процессов взаимодействует с первой группой с целью уточнения положений сверхзадачи.

*Мониторинг возможностей и ограничений* позволяет актуализировать текущие и перспективные сильные и слабые стороны СЭР РФ, внешние и внутренние возможности, угрозы и риски с целью выработки решений по реализации стратегических целей. Источниками данных мониторинга служат данные государственной статистической отчетности, данные социологических опросов, экспертного сообщества и независимых источников.

*Анализ альтернатив* — эта группа процессов позволяет на основе мониторинга возможностей и ограничений определить альтернативные варианты решений для формирования стратегического плана. Осуществляется создание моделей управления (оценка достижимости целей, указанных в стратегических документах, на основе математического моделирования и прогнозирования).

*Выработка стратегического плана* — формирование системы конкретных, определенных и обоснованно достижимых целей в виде количественных значений целевых показателей, привязанных к точке/интервалу времени и обеспеченных ресурсами, с определением ответственных исполнителей по отраслевому и меж-

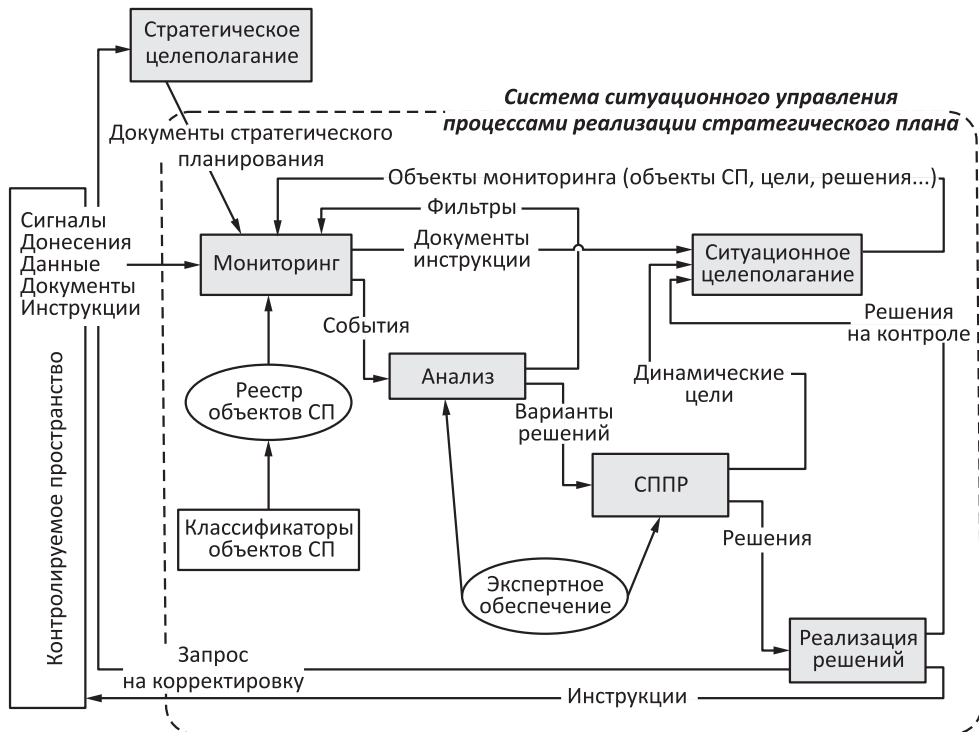


Рис. 2 Функциональная модель системы ситуационного управления

отраслевому принципу [5]. Сущностью стратегического планирования является *системное согласование* балансов доходов и расходов производителей и конечных потребителей — государства, предприятий, домашних хозяйств (государственный баланс), экспортеров и импортеров (внешнеэкономический баланс), а также интересов производителей и потребителей. Стратегический план задается основными направлениями деятельности правительства, планами деятельности федеральных органов исполнительной власти (ФОИВ), госпрограммами, планами мероприятий по реализации СЭР РФ. При этом увязываются планы по уровням планирования: долгосрочный (стратегия), среднесрочный (госпрограмма) и краткосрочный (бюджет).

Первая группа процессов СУСП взаимодействует со второй группой, отвечающей за реализацию задач стратегических планирования, которая по сути своей объединяет текущие процессы ситуационного управления (рис. 2).

Группа процессов *ситуационного целеполагания* обеспечивает контроль исполнения динамических целей, формируемых в СУСП в ходе текущего управления ходом реализации стратегического плана.

В рамках *мониторинга реализации плана* осуществляется сбор данных о контролируемых объектах, событиях и складывающихся ситуациях. Мониторинг опирается на реестр объектов стратегического планирования и соответствующие классификаторы, на источники данных мониторинга, включая документы стратегического планирования, на критерии и правила отбора данных из различных информационных ресурсов. В рамках процессов мониторинга осуществляется оценка параметров ненаблюдаемых (скрытых) элементов обстановки на основе выборочных или косвенных данных, анализ неструктурированных данных, выявление и извлечение фактов.

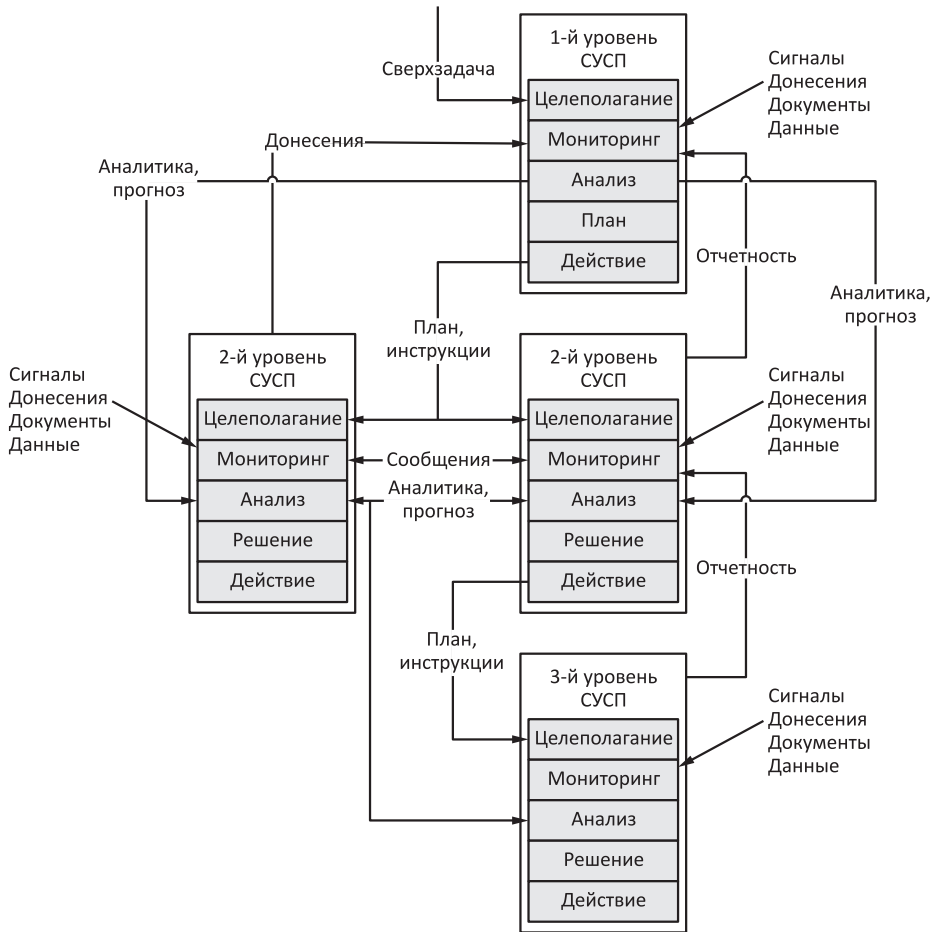
Процессы мониторинга фиксируют события, представляющие собой изменения состояния объектов мониторинга, которые поступают на вход функционального блока *анализа хода реализации документов стратегического планирования*. Здесь осуществляется оценка сложившейся ситуации (штатная, нештатная) и формирование гипотез о состоянии объектов стратегического планирования, а также анализ и учет существующего опыта. Процессы анализа опираются на методы декомпозиции складывающихся ситуаций на известные или достаточно простые ситуации для нормализации с помощью ресурсов системы. Наконец, в зависимости от того, какая ситуация идентифицирована, с использованием экспертного обеспечения вырабатываются варианты реакции СУСП. Важнейшей составляющей анализа являются процессы краткосрочного и стратегического прогнозирования.

Функциональный блок *системы поддержки принятия решений* (СППР) предназначен для формирования решения на основе возможных альтернатив с использованием методов многокритериального выбора. Оценка качества выбранных альтернатив решений может также осуществляться с использованием методов сценарного прогнозирования, что позволяет учесть влияние управляющих воздействий СУСП на развитие обстановки. Блок СППР должен обеспечить также формирование динамических целей для обеспечения процессов мониторинга хода реализации решений.

Процессы блока *реализации решений* обеспечиваются инфраструктурой формализованного электронного документооборота для их доведения до исполнителей и контроля исполнения. Данный функциональный блок при необходимости осуществляет обратную связь с системой стратегического целеполагания с целью корректировки документов стратегического планирования.

#### **4 Информационное взаимодействие в многоуровневой системе управления стратегическим планированием**

Информационные взаимодействия, происходящие в иерархической многоуровневой системе ситуационного управления стратегическим планированием, подразделяются на внешние и внутренние взаимодействия (рис. 3). Внешние взаимодействия с окружающей средой происходят с помощью сенсорных подсистем



**Рис. 3** Процессы взаимодействия в иерархической СУСП

тем, которые могут осуществлять прием сигналов, донесений, различных видов структурированных и неструктурированных данных, а также инструкций от вышестоящих органов, осуществляющих целеполагание и другое регулирование деятельности системы управления на основе нормативно-правовых документов.

В структуре информационных взаимодействий внутри иерархической системы управления можно выделить три вида данных: инструкции, отчеты и сведения.

Инструкции представляют собой управляющие воздействия в виде, как правило, формализованных сообщений (сигналов, целеуказаний, распоряжений, команд), передаваемых по иерархии подчиненности «сверху вниз» — от выше-



стоящих к нижестоящим подсистемам. С помощью инструкций выполняются действия по исполнению принятых решений, включая установление динамических ситуационных целей, определение сроков, ответственных исполнителей и выделение требуемых ресурсов. Такой функционал обычно реализуется средствами формализованного электронного документооборота.

Отчеты передаются от нижестоящих к вышестоящим подсистемам в виде формализованных донесений о реализации поставленных целей, которые находятся на контроле вышестоящего органа управления. Сюда могут включаться доклады о готовности и работоспособности, а также результаты деятельности по реализации решений.

Сведения представляют собой формализованные и неформализованные сообщения, передаваемые с целью информирования и координации деятельности взаимодействующих подсистем, а также обмена данными о результатах аналитики (прогнозирование, идентификация ситуаций, выявленные тренды, аномалии, артефакты). В [6] рассмотрены вопросы организации обмена аналитическими данными различных видов и обоснованы способы их формализации.

Приведенная процессная модель позволяет сформулировать достаточные функциональные требования и требования по назначению при подготовке технического задания на СУСП и ее составные части.

## Литература

1. Проблемы стратегического государственного планирования и управления в современной России: Материалы постоянно действующего научного семинара, 2006–2018 гг.
2. Сучков А. П., Босов А. В., Макоско А. А. Ситуационный анализ в процессах стратегического планирования в области обеспечения национальной безопасности // Инновации, 2018. № 8(238). С. 33–39.
3. Проблемы стратегического государственного планирования и управления в современной России: Материалы постоянно действующего научного семинара. — М.: Научный эксперт, 2011. Вып. 5(43). 96 с.
4. Сучков А. П. Полнофункциональный процессный подход к реализации систем ситуационного управления // Системы и средства информатики, 2017. Т. 27. № 1. С. 86–100.
5. Сучков А. П. Формирование системы целей для ситуационного управления // Системы и средства информатики, 2013. Т. 23. № 2. С. 171–182.
6. Сучков А. П. Некоторые подходы к интеграции аналитических данных существующих и перспективных систем поддержки принятия решений // Системы и средства информатики, 2015. Т. 25. № 3. С. 201–211.

*Поступила в редакцию 02.04.19*

## FUNCTIONAL MODEL OF STRATEGIC PLANNING PROCESS 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 article deals with the approaches to the development of process models of management systems in the field of strategic planning. From the point of view of the scientific approach, the methods of strategic planning can be attributed to the theory of complex systems management. The key is to understand strategic planning as a cyclical sequence of interrelated processes, among which there are two complete and interrelated management cycles which will be conventionally called “planning” and “implementation.” The management cycle “implementation” is clearly situational. It provides continuous monitoring of the implementation of the strategic plan, situational analysis of the flows of events in the controlled space, and the current response to changes in the situation up to the adjustment of strategic planning documents.

**Keywords:** strategic planning; process approach; process model; management system

**DOI:** 10.14357/08696527190412

### References

1. Problemy strategicheskogo gosudarstvennogo planirovaniya i upravleniya v sovremennoy Rossii: Materialy postoyanno deystvuyushchego nauchnogo seminar 2006–2018 [Problems of strategic state planning and management in modern Russia. Materials of the permanent scientific seminar 2006–2018]. Available at: <http://rusrand.ru/library/reports/seminar/psguvr> (accessed September 20, 2019).
2. Suchkov, A. P., A. V. Bosov, and A. A. Makosko. 2018. Situatsionnyy analiz v protsessakh strategicheskogo planirovaniya v oblasti obespecheniya natsional'noy bezopasnosti [Situational analysis in the processes of strategic planning in the field of national security]. *Innovatsii* [Innovations] 8(238):33–39.
3. Problemy strategicheskogo gosudarstvennogo planirovaniya i upravleniya v sovremennoy Rossii: Materialy postoyanno deystvuyushchego nauchnogo seminar [Problems of strategic state planning and management in modern Russia. Materials of the permanent scientific seminar]. 2011. Moscow: Nauchnyy expert. 5(43). 96 p. Available at: <https://rucont.ru/efd/136938> (accessed September 20, 2019).
4. Suchkov, A. P. 2017. Polnofunktsional'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.

5. 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.
6. Suchkov, A. P. 2015. Nekotorye podkhody k integratsii analiticheskikh dannykh sushchestvuyushchikh i perspektivnykh sistem podderzhki prinyatiya resheniy [Some approaches to the analytical data integration of the existing and future decision support systems]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 25(3):201–211.

*Received April 2, 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@frcsc.ru

## СИМВОЛЫ, КОДЫ, СИГНАЛЫ

*В. Д. Ильин*<sup>1</sup>

**Аннотация:** Совершенствование символьного, кодового и сигнального арсенала современной среды построения, интерпретации, сохранения, манипулирования и обмена сообщениями — постоянно актуальная научно-техническая задача, от решения которой зависит продвижение по многим направлениям развития интерфейсов «человек–машина» и «машина–машина». Символы, коды и сигналы рассматриваются в статье как конструктивные объекты для построения, сохранения и передачи сообщений в цифровой среде. Символ определен как заменитель некоторого объекта, принадлежащий определенному набору для формирования сообщений по заданным правилам. Каждый элемент такого набора наделен совокупностью свойств (одинаковой для всех элементов набора), обеспечивающей применимость в заданной среде формирования, передачи, интерпретации, сохранения сообщений и манипулирования ими (копирования, поиска и др.). Определены виды (визуальный, аудио и др.) символов и для каждого вида — типы символов. Код определен как заменитель символа или сообщения, удовлетворяющий требованиям решения базовых задач представления, преобразования, распознавания, конструирования, интерпретации, обмена, сохранения, накопления, поиска и защиты сообщений. Сигнал изучается как физически реализованное представление кода, рассчитанное на распознавание и интерпретацию человеком или аппаратным средством (микропроцессором, видеоконтроллером или др.).

**Ключевые слова:** интерфейсы; сообщения; символьное моделирование; символы; виды символов; типы символов; коды; сигналы

**DOI:** 10.14357/08696527190413

### 1 Введение

В современной *среде решения задач с помощью компьютеров и компьютерных устройств (s-среде)* с символами и построенными из них сообщениями имеют дело люди, с кодами (символов и сообщений) — программные средства, а с сигналами (физическими реализациями кодов) — аппаратные средства<sup>2</sup>.

В наши дни инфраструктурное ядро *применения цифровых информационных технологий (цифровизации)* включает технологии цифровых двойников (*англ. digital twins*) [1], облачных вычислений (*англ. cloud computing*) [2–4], электронных сервисов [5–7], Интернета вещей (*англ. Internet of Things, IoT*) [8, 9] и M2M-технологии (*англ. Machine-to-Machine, M2M*) [10, 11].

<sup>1</sup>Вычислительный центр им. А. А. Дородницына Федерального исследовательского центра «Информатика и управление» Российской академии наук, vdilyin@yandex.ru

<sup>2</sup>*Ильин В. Д.* Символьное моделирование // Большая российская энциклопедия, 2019. [Электронный ресурс]. <http://dev.bigenc.ru/technology and technique/text/4010980> (дата обращения: 22.08.2019).

**Выделение фрагментов текста.** Для выделения определений, замечаний и примеров далее используются средства языка TSM-комплекса (TSM: Textual Symbolic Modeling), разработанного для формализованного описания текстовых моделей [12].

В статье применены следующие средства выделения фрагментов текста: □ ⟨фрагмент описания⟩ □ ≈ утверждение (определение, аксиома и др.) (здесь и далее символ ≈ заменяет слово «означает»);

◇ ⟨фрагмент описания⟩ ◇ ≈ замечание;

○ ⟨фрагмент описания⟩ ○ ≈ пример.

Курсивом выделены первые вхождения названий понятий и фрагменты описания, к которым автор хочет привлечь внимание.

Префикс *s*- означает принадлежность к системе понятий *символьного моделирования произвольных объектов в человеко-машинной среде решения задач (s-среде)*, названного *s-моделированием* [12].

**Рассматриваемые результаты.** Рассматриваемые в статье 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.

## 2 S-символы

В *s*-среде с помощью *компьютеров и компьютерных устройств (в теории s-моделирования [12] названных s-машинами)* формируются сообщения, представляемые на *языках программирования*<sup>1</sup>, запросов и др. Выполняются различные преобразования [○ из аналоговой формы в цифровую и обратно; из несжатой в сжатую и обратно; из одной формы представления документа в другую (○ \*.doc в \*.pdf ○) ○] и распознавания сообщений. Сохраняемые сообщения используются для конструирования новых сообщений (программ, документов и др.). Сообщения интерпретируются на *моделях систем понятий* [12], которые хранятся в памяти интерпретатора также в форме сообщений. Сообщениями обмениваются [используя при этом программно-аппаратно реализованные системы правил (○ *s-сетевые*<sup>2</sup> протоколы ○)]; их сохраняют и накапливают (создавая электронные библиотеки, энциклопедии и другие информационные ресурсы). Сообщения служат объектами поиска и защиты [12].

□ *S-сообщение* рассматриваем как конечную упорядоченную совокупность *s-символов*, рассчитанную на распознавание и интерпретацию получателем-человеком, или ее *s-код*, удовлетворяющий требованиям решения базовых задач

<sup>1</sup>Ильин А. В. Языки программирования // Большая российская энциклопедия, 2019. [Электронный ресурс]. [https://bigenc.ru/technology\\_and\\_technique/text/4924750](https://bigenc.ru/technology_and_technique/text/4924750) (дата обращения: 22.08.2019).

<sup>2</sup>Ильин В. Д. Компьютерная сеть // Большая российская энциклопедия, 2019. [Электронный ресурс]. [https://bigenc.ru/technology\\_and\\_technique/text/2087809](https://bigenc.ru/technology_and_technique/text/2087809) (дата обращения: 22.08.2019).

*s*-(представления, преобразования, распознавания, конструирования, интерпретации, обмена, сохранения, накопления, поиска и защиты) в *s*-среде. □

○ *S*-модели систем понятий и систем знаний, в которых представлены результаты изучения произвольных сущностей (объектов исследований) [12]; программы, определяющие поведение *s*-машин; веб-страницы и файлы документов — все это *s*-сообщения ○.

□ *S*-символ — заместитель некоторого объекта, принадлежащий определенному набору, предназначенному для формирования *s*-сообщений по заданным правилам. Каждый элемент такого набора наделен совокупностью свойств (одинаковой для всех элементов набора), обеспечивающей применимость в *s*-среде для формирования, передачи, распознавания, интерпретации, сохранения сообщений и манипулирования ими (копирования, поиска и др.). Представлен в двух формах, одна из которых рассчитана на распознавание и интерпретацию человеком, другая — программой *s*-машины.

*S*-символ является реализуемым в *s*-среде средством представления произвольного объекта, наделенным набором свойств, обеспечивающих возможности: распознавания и интерпретации человеком; применения без участия изобретателя; неограниченного числа воспроизведений по заданным правилам, копирования, удаления и перемещения в *s*-среде; применения в задачах *s*-(представления, распознавания, преобразования, конструирования, интерпретации, обмена, сохранения, накопления, поиска, защиты) сообщений. Обозначение в *s*-моделировании — *s-symbol*. □

Роль набора символов может выполнять что угодно [если это «что угодно» наделено совокупностью свойств, обеспечивающих применимость в заданной среде формирования, передачи, интерпретации, сохранения сообщений и манипулирования ими (копирования, поиска и др.)].

○ Набор жестовых символов языка глухонемых; система Брайля для слепых (используются фактурные символы, рассчитанные на осязание посредством пальцев рук); набор звуковых символов музыкальной композиции и соответствующий ему набор графических символов нотного письма; набор графических символов шахматной нотации. ○

Не предполагается никаких ограничений на виды и типы заменяемых символами объектов: они могут иметь любую физическую сущность, размещение, происхождение и назначение. Символы одного вида могут заменять символы другого вида (то же справедливо и для типов символов).

*Воспроизводимость.* Оригинальный символ, хранящийся только в памяти человека, который его изобрел, приобретает шанс стать воспроизводимым, когда он представлен в *s*-среде. Это означает, что новый символ описан его изобретателем с использованием других символов, считающихся известными. Другими словами, воспроизводимость предполагает обязательную опору на уже известные символы и средства их изготовления и применения.

*Дискретность.* Символы, используемые в *s*-среде, дискретны. Дискретность символов — необходимое условие их различимости.

*Пороговая различимость.* Определяет допустимое уменьшение размеров визуальных  $s$ -символов, продолжительности звучания аудио- $s$ -символов и т. д. [12].

**Виды и типы символов.** Видам символов (*визуальному, аудио* и др. [12]) соответствуют средства производства и восприятия сообщений, которыми наделен человек и  $s$ -машинные устройства ввода-вывода. Виды символов делятся на типы.

□ Типу  $ts[A[at, op]]$  символов соответствует множество  $A$  символов, для которых определены набор атрибутов  $at$  и семейство  $op$  допустимых операций. □

*Специализацией*  $ts[B < A[at, op]]$  типа  $ts[A[at, op]]$  будем называть подмножество  $B < A[at, op]$  символов с набором атрибутов  $at[B]$  ( $at[B] > at[A]$ ) и семейством допустимых операций  $op[B]$  ( $op[B] > op[A]$ ). Тип  $ts[A[at, op]]$  будем называть *обобщением* типа  $ts[B < A[at, op]]$ .

○ *Текстовый, гипертекстовый, анимационный* — специализации типа *графический*. Элементы множества символов типа *текстовый* — графические изображения букв, знаков препинания и др., для которых задан набор ограничений (на размеры, взаимное размещение и др.). Тип *числовой* — специализация типа *текстовый* (◇ в семействе операций типа *числовой* есть арифметические операции, которых нет в семействе операций типа *текстовый* ◇). ○

**Отношения.** □ *Эквивалентность* определяет взаимозаменяемость  $s$ -символов. □ Может быть задана для любого числа  $s$ -символов (не менее двух). ○ Вибровывод мобильного телефона ( $s$ -сообщение, составленное из тактильных символов) может быть эквивалентен любому из аудиовыводов ○.

□ *Порядок* определяет последовательность  $s$ -символов. □ ○ Алфавит, натуральный ряд чисел ○.

□ *Принадлежность* определяет вхождение  $s$ -символа в некоторый набор символов. □ ○ Символ «□», используемый для выделения определений, аксиом и других утверждений, принадлежит набору символов TSM [12]. ○

□ ***Визуальный вид  $s$ -символов*** — совокупность символов, применимых для построения сообщений, содержащих неподвижные и подвижные изображения, рассчитанные на восприятие человеком посредством зрения, а программами  $s$ -машин — посредством устройств графического и видеоввода; специализация  $s$ -symbol по параметру *means* (средства):  $s\text{-symbol}[\text{visual}] \approx \approx s\text{-symbol}[:, \text{means}=\text{sight}]$ . □

Значение *sight* выделяет из множества символов те, которые рассчитаны на зрительный аппарат человека и устройства графического и видеоввода в составе  $s$ -машин (веб-камеры, сканеры и др.). Вывод сообщений, содержащих визуальные символы и рассчитанных на прием человеком, осуществляется на экраны мобильных устройств, мониторов и др.

Виду *визуальный* соответствуют типы: *графический* [для построения сообщений, содержащих неподвижные изображения (фотографии, схемы и др.)]; *видео* (для построения сообщений, содержащих подвижные изображения).

◇ *Текстовый тип s-символов* — специализация графического типа символов.

○ Текстовый набор в стандарте Unicode включает буквы национальных алфавитов и различные символы (математические, музыкальные и др.), используемые при создании текстовых сообщений с помощью s-машин. ○

Текстовые символы могут быть графическими заменителями речевых символов (○ буквы русского алфавита — заменители звуков русской речи ○). Текстовыми наборами символов (включающими специальные символы) представлены алфавиты языков программирования, формальных языков математики и др. ◇

□ **Звуковой (аудио) вид s-символов** — совокупность символов, применимых для построения сообщений, содержащих звуки, рассчитанные на звукопроизводящие и звукоприемные средства человека и s-машин. Специализация s-symbol по параметру means: s-symbol[audio] ≈ s-symbol[:means=audio]. □ Значение audio выделяет из множества символов те, которые рассчитаны на голосовой и слуховой аппараты человека, музыкальные инструменты, аудиоустройства ввода и вывода в составе s-машин.

Виду аудио соответствуют типы: *речевой, музыкальный* и др. [12].

□ **Тактильный вид s-символов** — совокупность символов, применимых для построения сообщений путем прикосновений к элементам *интерфейсных<sup>1</sup> устройств s-машин*, изменения фактуры и температуры их поверхности, которые рассчитаны на восприятие человеком посредством осязания, а программами s-машин — посредством клавиатур, сенсорных экранов и других тактильных устройств s-машинного ввода. Специализация s-symbol по параметру means: s-symbol[touch] ≈ s-symbol[:means=touch]. □

Значение touch выделяет из множества символов подмножество, рассчитанное на осязательный аппарат человека и тактильные устройства ввода (клавиатуру, сенсорный экран и др.) и вывода (○ механизм вибровызова мобильного телефона ○) в составе s-машин.

Виду тактильный соответствуют типы: *жестовый* (для передачи сообщений тактильными жестами); *кинетический* (для передачи сообщений путем перемещения элементов интерфейсных устройств s-машины (○ ввод с клавиатуры, манипулирование мышью, вибровызов мобильного телефона ○)); *фактурный* (для передачи сообщений изменением фактуры поверхности (○ плоская, волнистая и т. д. ○) интерфейсных устройств); *термический* (для передачи сообщений путем изменения температуры поверхности интерфейсных устройств) [12].

### 3 S-коды

□ *S-код (s-code)* — заменитель s-символа или s-сообщения, удовлетворяющий требованиям решения базовых задач s-(представления, преобразования,

---

<sup>1</sup>Ильин А. В. Интерфейс // Большая российская энциклопедия, 2019. [Электронный ресурс]. [https://bigenc.ru/technology\\_and\\_technique/text/4426494](https://bigenc.ru/technology_and_technique/text/4426494) (дата обращения: 22.08.2019).



распознавания, конструирования, интерпретации, обмена, сохранения, накопления, поиска и защиты) в  $s$ -среде. □

Исследователям и инженерам, имеющим дело с различными символьными моделями, свойственно стремление представлять их числовыми кодами. Это позволяет применять методы решения задач, которые можно представить в виде программ, рассчитанных на выполнение  $s$ -машинами.  $S$ -символам любого вида ставят во взаимно однозначное соответствие числа, которые можно представить в памяти  $s$ -машины.

○ ASCII (American Standard Code for Information Interchange) — набор, состоящий из 128 текстовых символов (букв, цифр, знаков пунктуации и др.). Расширенный набор ASCII включает 256 текстовых символов. Каждому символу назначен уникальный номер, называемый кодом ASCII. ○

Изобретая цифровые  $s$ -машины, выбирают систему счисления и число разрядов для представления чисел в памяти  $s$ -машин. При этом выбор направляется стремлением обеспечить наиболее эффективное манипулирование числовыми кодами. Выбор ограничен рядом условий, среди которых физико-техническая реализуемость элементов  $s$ -машин. Обычно основание системы счисления выбирают равным количеству устойчивых состояний, в которых могут находиться элементарные составляющие, из которых построена  $s$ -машина.

○ Поскольку транзисторы современных электронных  $s$ -машин имеют два устойчивых состояния, в них используется двоичное представление числовых кодов программ и *данных*<sup>1</sup> [12]. ○

## 4 $S$ -сигналы

□  $S$ -сигнал ( $s$ -signal) — физически реализованное представление  $s$ -кода, рассчитанное на распознавание и интерпретацию человеком или аппаратным средством  $s$ -машины (микропроцессором, видеоконтроллером или др.). □

○ В цифровой  $s$ -машине с двоичным представлением  $s$ -кодов машинных команд и данных  $s$ -сигналы представлены композициями двух типов импульсов напряжения на выходах транзисторов, один из которых соответствует двоичному нулю, а другой — двоичной единице [12]. ○

## 5 Заключение

1. Одной из постоянно актуальных задач  $s$ -моделирования является реализация в  $s$ -среде знаний и умений, накопленных человечеством. Для ее решения природный сенсорный комплекс (зрение, слух, осязание, обоняние, вкус) человека дополняется в  $s$ -среде изобретенными сенсорами, рассчитанными на восприятие визуальных, аудио-, тактильных и запаховых символов.

<sup>1</sup>Ильин В. Д. Данные // Большая российская энциклопедия, 2019. [Электронный ресурс]. [https://bigenc.ru/technology\\_and\\_technique/text/3954919](https://bigenc.ru/technology_and_technique/text/3954919) (дата обращения: 22.08.2019).

2. Совершенствование интерфейсов «человек–машина» и «машина–машина» предполагает развитие символического, кодового и сигнального арсенала цифровой среды построения, интерпретации, сохранения, манипулирования и обмена сообщениями.
3. В процессе развития теории *s*-моделирования [12] уточняются определения и истолкование ключевых понятий (*s*-символ, *s*-код, *s*-сигнал) и понятий, связанных с ними (вид *s*-символа, тип *s*-символа и др.).

## Литература

1. 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).
2. *Armbrust M., Fox A., Griffith R., et al.* A view of cloud computing // Commun. ACM, 2010. Vol. 53. Iss. 4. P. 50–58. doi: 10.1145/1721654.1721672.
3. *Wang L., Laszewski G., Younge A., et al.* Cloud computing: A perspective study // New Gener. Comput., 2010. Vol. 28. Iss. 2. P. 137–146. doi: 10.1007/s00354-008-0081-5.
4. *Jamsa K. A.* Cloud computing. — Burlington: Jones & Bartlett Learning, 2013. 322 p.
5. *Wei Y., Blake M. B.* Service-oriented computing and cloud computing: Challenges and opportunities // IEEE Internet Comput., 2010. Vol. 14. Iss. 6. P. 72–75. doi: 10.1109/MIC.2010.147.
6. *Ильин А. В.* Интернет-сервис планирования расходов // Системы и средства информатики, 2015. Т. 25. № 2. С. 111–122. doi: 10.14357/08696527150207.
7. *Jede A., Teuteberg F.* Understanding socio-technical impacts arising from software-as-a-service usage in companies // Business Information Syst. Eng., 2016. Vol. 58. Iss. 3. P. 161–176. doi: 10.1007/s12599-016-0429-1.
8. *Perera C., Liu C. H., Jayawardena S.* The emerging Internet of Things marketplace from an industrial perspective: A survey // IEEE T. Emerging Topics Computing, 2015. Vol. 3. Iss. 4. P. 585–598. doi: 10.1109/TETC.2015.2390034.
9. «Интернет вещей» (IoT) в России. Технология будущего, доступная уже сейчас. — PWC, 2017. [https://www.pwc.ru/ru/publications/iot/loT-inRussia-research\\_rus.pdf](https://www.pwc.ru/ru/publications/iot/loT-inRussia-research_rus.pdf).
10. *Kim R. Y.* Efficient wireless communications schemes for machine to machine communications // Comm. Com. Inf. Sc., 2011. Vol. 181. Iss. 3. P. 313–323. doi: 10.1007/978-3-642-22203-0\_28.
11. *Lien S. Y., Liao T. H., Kao C. Y., et al.* Cooperative access class barring for machine-to-machine communications // IEEE T. Wirel. Commun., 2012. Vol. 11. Iss. 1. P. 27–32. doi: 10.1109/TWC.2011.111611.110350.
12. *Ильин А. В., Ильин В. Д.* Основы теории *s*-моделирования. — М.: ИПИ РАН, 2009. 143 с.

Поступила в редакцию 27.08.19

## SYMBOLS, CODES, SIGNALS

*V. D. Ilyin*

Dorodnicyn Computing Center of the Russian Academy of Sciences, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 40 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** Improving the symbolic, code, and signal arsenal of the modern environment for constructing, interpreting, saving, manipulating, and exchanging messages is a constantly relevant scientific and technical problem. The advancement in many areas of human–machine and machine–machine interfaces development depends on solution of this problem. In this article, symbols, codes, and signals are considered as constructive objects for building, storing, and transmitting messages in a digital environment. A symbol is defined as a substitute for an object, belonging to a specific set for generating messages according to the specified rules. Each element of such set is endowed with a set of properties (the same for all elements of the set) which ensures applicability in a given environment of the forming, transmitting, interpreting, storing, and manipulating the messages. The symbols kinds (visual, audio, etc.) and the symbol types inside each kind have been defined. The code is defined as a substitute for symbol or message that meets the requirements for solving the basic problems of representing, transforming, recognizing, constructing, interpreting, exchanging, saving, accumulating, searching, and protecting messages. The signal is studied as a physically realized code representation designed for recognition and interpretation by human or hardware (microprocessor, video controller, etc.).

**Keywords:** interfaces; messages; symbolic modeling; symbols; symbol kinds; symbol types; codes; signals

**DOI:** 10.14357/08696527190413

### References

1. General Electric. 2018. 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 August 22, 2019).
2. Armbrust, M., A. Fox, R. Griffith, *et al.* 2010. A view of cloud computing. *Commun. ACM* 53(4):50–58. doi: 10.1145/1721654.1721672.
3. Wang, L., G. Laszewski, A. Younge, *et al.* 2010. Cloud computing: A perspective study. *New Gener. Comput.* (28):137–146. doi: 10.1007/s00354-008-0081-5.
4. Jamsa, K. A. 2013. *Cloud computing*. Burlington: Jones & Bartlett Learning. 322 p.
5. Wei, Y., and M. B. Blake. 2010. Service-oriented computing and cloud computing: Challenges and opportunities. *IEEE Internet Comput.* 14(6):72–75. doi: 10.1109/MIC.2010.147.

6. Ilyin, A. V. 2015. Internet-servis planirovaniya raskhodov [The online service for cost planning]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 25(2):111–122. doi: 10.14357/08696527150207.
7. Jede, A., and F. Teuteberg. 2016. Understanding socio-technical impacts arising from software-as-a-service usage in companies. *Business Information Syst. Eng.* 58(3):161–176. doi: 10.1007/s12599-016-0429-1.
8. Perera, C., C. H. Liu, and S. Jayawardena. 2015. The emerging Internet of Things marketplace from an industrial perspective: A survey. *IEEE T. Emerging Topics Computing* 3(4):585–598. doi: 10.1109/TETC.2015.2390034.
9. PWC. 2017. “Internet veshchey” (IoT) v Rossii. Tekhnologiya budushchego, dostupnaya uzhe seychas [Internet of things (IoT) in Russia. Technology of the future, available now]. Available at: [https://www.pwc.ru/ru/publications/iot/loT-inRussia-research\\_rus.pdf](https://www.pwc.ru/ru/publications/iot/loT-inRussia-research_rus.pdf) (accessed August 22, 2019).
10. Kim, R. Y. 2011. Efficient wireless communications schemes for machine to machine communications. *Comm. Com. Inf. Sc.* 181(3):313–323. doi: 10.1007/978-3-642-22203-0\_28.
11. Lien, S. Y., T. H. Liau, C. Y. Kao, *et al.* 2012. Cooperative access class barring for machine-to-machine communications. *IEEE T. Wirel. Commun.* 11(1):27–32. doi: 10.1109/TWC.2011.111611.110350.
12. Ilyin, A. V., and V. D. Ilyin. 2009. *Osnovy teorii s-modelirovaniya* [Basics of the theory of s-modeling]. Moscow: IPI RAN.143 p.

*Received August 27, 2019*

## Contributor

**Ilyin Vladimir D.** (b. 1937) — Doctor of Science in technology, professor, leading scientist, Dorodnicyn Computing Center of the Russian Academy of Sciences, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 40 Vavilov Str., Moscow 119333, Russian Federation; vdilyin@yandex.ru

**Адамович Игорь Михайлович** (р. 1934) — кандидат технических наук, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Бесчастный Виталий Александрович** (р. 1992) — аспирант кафедры прикладной информатики и теории вероятностей Российского университета дружбы народов

**Вакуленко Василий Васильевич** (р. 1995) — инженер-исследователь Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Волков Олег Игоревич** (р. 1964) — ведущий программист Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Гайдамака Юлия Васильевна** (р. 1971) — доктор физико-математических наук, профессор кафедры прикладной информатики и теории вероятностей Российского университета дружбы народов; старший научный сотрудник 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

**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

**Забейайло Михаил Иванович** (р. 1956) — доктор физико-математических наук, доцент, главный научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Зарядов Иван Сергеевич** (р. 1981) — кандидат физико-математических наук, доцент кафедры прикладной информатики и теории вероятностей Российского университета дружбы народов; старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Захарова Татьяна Валерьевна** (р. 1962) — кандидат физико-математических наук, доцент кафедры математической статистики факультета вычислительной математики и кибернетики МГУ имени М. В. Ломоносова; старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Зацман Игорь Моисеевич** (р. 1952) — доктор технических наук, заведующий отделом Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Ильин Владимир Дмитриевич** (р. 1937) — доктор технических наук, профессор, ведущий научный сотрудник Вычислительного центра им. А. А. Дородницына Российской академии наук Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Колесников Александр Васильевич** (р. 1948) — доктор технических наук, профессор Института физико-математических наук и информационных технологий Балтийского федерального университета имени Иммануила Канта

**Кривенко Михаил Петрович** (р. 1946) — доктор технических наук, профессор, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Литвин Андрей Антонович** (р. 1966) — доктор медицинских наук, заместитель главного врача по медицинской части Областной клинической больницы Калининградской области

**Мейханаджян Лусине Акововна** (р. 1990) — кандидат физико-математических наук, доцент Департамента анализа данных, принятия решений и финансовых технологий Финансового университета при Правительстве РФ

**Милованова Татьяна Александровна** (р. 1977) — кандидат физико-математических наук, старший преподаватель кафедры прикладной информатики и теории вероятностей Российского университета дружбы народов

**Морозов Николай Викторович** (р. 1956) — старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Острикова Дарья Юрьевна** (р. 1988) — кандидат физико-математических наук, старший преподаватель кафедры прикладной информатики и теории вероятностей Российского университета дружбы народов

**Рождественский Юрий Владимирович** (р. 1952) — кандидат технических наук, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Румовская София Борисовна** (р. 1985) — кандидат технических наук, научный сотрудник Калининградского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Рычихин Алексей Константинович** (р. 1995) — инженер Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Самуйлов Константин Евгеньевич** (р. 1955) — доктор технических наук, профессор, заведующий кафедрой прикладной информатики и теории вероятностей Российского университета дружбы народов; старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Сороколетова Ольга Евгеньевна** (р. 1996) — студентка кафедры математической статистики факультета вычислительной математики и кибернетики МГУ имени М. В. Ломоносова

**Степченко Дмитрий Юрьевич** (р. 1973) — старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Степченко Юрий Афанасьевич** (р. 1951) — кандидат технических наук, 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

**Тимонина Елена Евгеньевна** (р. 1952) — доктор технических наук, профессор, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Чухно Надежда Викторовна** (р. 1995) — студентка кафедры прикладной информатики и теории вероятностей Российского университета дружбы народов

**Чухно Ольга Викторовна** (р. 1995) — студентка кафедры прикладной информатики и теории вероятностей Российского университета дружбы народов



---

## АВТОРСКИЙ УКАЗАТЕЛЬ ЗА 2019 г.

---

	№	Стр.
<b>Абгарян К. К., Гаврилов Е. С.</b> Информационная поддержка интеграционной платформы многомасштабного моделирования	1	53–62
<b>Адамович И. М., Волков О. И.</b> Единая технология поддержки конкретно-исторических исследований	1	194–205
<b>Адамович И. М., Волков О. И.</b> Исследование процесса формирования виртуального сообщества как нелинейной динамической системы	4	119–128
<b>Адамович И. М., Волков О. И.</b> Принципы организации данных для технологии поддержки конкретно-исторических исследований	2	161–171
<b>Адамович И. М., Волков О. И.</b> Процесс коррекции ошибок семантической сети как нелинейная динамическая система	3	160–168
<b>Арутюнов Е. Н., Кудрявцев А. А., Титова А. И.</b> Априорное вейбулловское распределение в байесовских моделях баланса	1	174–179
<b>Архипов А. Е.</b> см. Бобырь М. В.		
<b>Бекешева И. С.</b> см. Бобылева О. В.		
<b>Бесчастный В. А., Острикова Д. Ю., Гайдамака Ю. В.</b> Анализ производительности систем «новое радио» сети 5G с помощью системы массового обслуживания с переменными требованиями к ресурсу	4	73–83
<b>Бобылев В. А.</b> см. Бобылева О. В.		
<b>Бобылева О. В., Бекешева И. С., Бобылев В. А., Чаркова В. В.</b> Алгоритм нечеткого сравнения при обработке персональных данных	3	149–159
<b>Бобырь М. В., Архипов А. Е., Милостная Н. А.</b> Метод расчета карты глубин на основе мягких операторов	2	71–84
<b>Бунтман Н. В.</b> см. Гончаров А. А.		
<b>Вакуленко В. В., Зацман И. М.</b> Принципы создания и функции интеллектуальной словарной системы	4	96–105
<b>Власова С. А., Каленов Н. Е.</b> Интернет-каталог Библиотеки по естественным наукам Российской академии наук как специальная информационно-поисковая система, ориентированная на квалифицированного пользователя	1	86–95
<b>Волков О. И.</b> см. Адамович И. М.		
<b>Волков О. И.</b> см. Адамович И. М.		
<b>Волков О. И.</b> см. Адамович И. М.		
<b>Волков О. И.</b> см. Адамович И. М.		
<b>Волович К. И.</b> см. Зацаринный А. А.		

	№	Стр.
<b>Гаврилов Е. С.</b> см. Абгарян К. К.		
<b>Гайдамака Ю. В.</b> см. Бесчастный В. А.		
<b>Гайдамака Ю. В.</b> см. Чухно О. В.		
<b>Гайдамака Ю. В.</b> см. Чухно О. В.		
<b>Галина И. В., Шарнин М. М.</b> Индекс контекстного научного цитирования	3	66–76
<b>Галина И. В.</b> см. Шарнин М. М.		
<b>Гаранин А. И.</b> см. Зацаринный А. А.		
<b>Гончаров А. А., Бунтман Н. В., Нуриев В. А.</b> Ошибки в машинном переводе: проблемы классификации	3	92–103
<b>Гончаров А. А., Зацман И. М.</b> Информационные трансформации параллельных текстов в задачах извлечения знаний	1	180–193
<b>Гончаров А. А., Инькова О. Ю., Кружков М. Г.</b> Методология аннотирования в надкорпусных базах данных	2	148–160
<b>Грушо А. А., Грушо Н. А., Тимонина Е. Е.</b> Методы защиты от массового вывода из строя слабозащищенных компьютеров цифровой экономики	1	4–11
<b>Грушо А. А., Забежайло М. И., Грушо Н. А., Тимонина Е. Е.</b> Поиск эмпирических причин сбоев и ошибок в компьютерных системах и сетях с использованием метаданных	4	28–38
<b>Грушо А. А., Зацаринный А. А., Тимонина Е. Е.</b> Электронная бухгалтерская книга на базе ситуационных центров для цифровой экономики	2	4–11
<b>Грушо Н. А.</b> Поиск путей динамической реконфигурации распределенной информационно-вычислительной системы в случае захвата хоста противником	3	194–201
<b>Грушо Н. А.</b> см. Грушо А. А.		
<b>Грушо Н. А.</b> см. Грушо А. А.		
<b>Дулин С. К., Розенберг И. Н., Уманский В. И.</b> О проблеме интеграции информационных ресурсов	3	127–138
<b>Дьяченко Д. Ю.</b> см. Степченков Ю. А.		
<b>Дьяченко Д. Ю.</b> см. Степченков Ю. А.		
<b>Дьяченко Ю. Г.</b> см. Степченков Ю. А.		
<b>Дьяченко Ю. Г.</b> см. Степченков Ю. А.		
<b>Егоров В. Б.</b> Практическое определение «программно определяемого»	2	85–94
<b>Егоров В. Б.</b> Способ повышения пропускной способности пакетных коммутаторов на основе интегрированных сетевых процессоров	1	63–73
<b>Егорова А. Ю., Зацман И. М., Мамонова О. С.</b> Надкорпусные базы данных в лингвистических проектах	3	77–91
<b>Забежайло М. И.</b> см. Грушо А. А.		

	№	Стр.
<b>Зарядов И. С., Мейханаджян Л. А., Милованова Т. А.</b> Стационарные характеристики обслуживания в системе GI/MSP/n/∞ собобщенным обновлением	4	50–64
<b>Захаров В. Н., Мусабаяев Р. Р., Красовицкий А. М., Козлов- ская Я. Д., Хорошилов Ал-др А., Хорошилов Ал-ей А.</b> Метод кластеризации новостных сообщений средств массо- вой информации на основе их концептуального анализа	3	52–65
<b>Захарова Т. В.</b> см. Сороколетова О. Е.		
<b>Зацаринный А. А., Гаранин А. И., Кондрашев В. А., Во- лович К. И., Мальковский С. И.</b> Оценка надежности гибридного высокопроизводительного вычислительного комплекса при решении научных задач	2	135–147
<b>Зацаринный А. А., Ионенков Ю. С.</b> Метод выбора вариан- та построения информационно-телекоммуникационной сис- темы	3	114–126
<b>Зацаринный А. А., Кондрашев В. А., Сучков А. П.</b> Систе- ма научных сервисов как актуальный компонент научных исследований	1	25–40
<b>Зацаринный А. А., Королёв В. И.</b> Технологические аспек- ты обеспечения информационной безопасности финансово- го рынка в условиях цифровой трансформации экономики России	1	12–24
<b>Зацаринный А. А.</b> см. Грушо А. А.		
<b>Зацман И. М.</b> см. Вакуленко В. В.		
<b>Зацман И. М.</b> см. Гончаров А. А.		
<b>Зацман И. М.</b> см. Егорова А. Ю.		
<b>Иванов Д. В.</b> Условные границы средних максимумов случай- ных величин и их достижимость	1	140–163
<b>Ильин А. В., Ильин В. Д.</b> Формирование ситуационно зави- симых систем требований к решениям задач планирования расходов	3	169–179
<b>Ильин В. Д.</b> Модель кооперативного решателя задач на основе цифровых двойников	2	172–179
<b>Ильин В. Д.</b> Символы, коды, сигналы	4	140–148
<b>Ильин В. Д.</b> см. Ильин А. В.		
<b>Инькова О. Ю.</b> см. Гончаров А. А.		
<b>Ионенков Ю. С.</b> см. Зацаринный А. А.		
<b>Каленов Н. Е.</b> см. Власова С. А.		
<b>Кириков И. А.</b> см. Листопад С. В.		
<b>Киселёва Н. Н.</b> см. Ожерельев И. С.		
<b>Коднянко В. А.</b> О вычислительной избыточности метода ди- хотомии и условной минимизации унимодальных функций методом экономной дихотомии	1	164–173

	№	Стр.
<b>Козловская Я. Д.</b> см. Захаров В. Н.		
<b>Колесников А. В., Листопад С. В.</b> Функциональная структура гибридной интеллектуальной многоагентной системы гетерогенного мышления для решения проблемы восстановления распределительной электросети	1	41–52
<b>Колесников А. В., Листопад С. В., Майтаков Ф. Г.</b> Интеллектуализация оперативно-диспетчерского управления сложной региональной динамической электроэнергетической системой	2	46–59
<b>Колесников А. В.</b> см. Румовская С. Б.		
<b>Кондрашев В. А.</b> см. Зацаринный А. А.		
<b>Кондрашев В. А.</b> см. Зацаринный А. А.		
<b>Королёв В. И.</b> см. Зацаринный А. А.		
<b>Косолапов Ю. В.</b> Способ вкрапления данных на основе одной схемы разделения секрета	3	180–193
<b>Красовицкий А. М.</b> см. Захаров В. Н.		
<b>Кривенко М. П.</b> Выбор размерностей для смеси вероятностных анализаторов главных компонент	3	4–15
<b>Кривенко М. П.</b> Снижение размерности для смеси вероятностных анализаторов главных компонент применительно к задачам медицинской диагностики	4	4–13
<b>Кружков М. Г.</b> см. Гончаров А. А.		
<b>Кудрявцев А. А., Палионная С. И., Шестаков О. В.</b> Индекс преимущества в Байесовских моделях надежности и баланса с бета-полиномиальными априорными плотностями	3	29–38
<b>Кудрявцев А. А., Палионная С. И., Шоргин В. С.</b> Априорное обобщенное распределение Фреше в байесовских моделях баланса	2	39–45
<b>Кудрявцев А. А.</b> см. Арутюнов Е. Н.		
<b>Листопад С. В., Кириков И. А.</b> Моделирование конфликтов агентов в гибридных интеллектуальных многоагентных системах	3	139–148
<b>Листопад С. В.</b> см. Колесников А. В.		
<b>Листопад С. В.</b> см. Колесников А. В.		
<b>Литвин А. А.</b> см. Румовская С. Б.		
<b>Лукьянов Л. А.</b> см. Спивак С. И.		
<b>Майоров Б. Г.</b> Восстановление сигнала в системах реального времени по равномерным выборкам с уменьшением интервала Найквиста	2	95–112
<b>Майтаков Ф. Г.</b> см. Колесников А. В.		
<b>Мальковский С. И.</b> см. Зацаринный А. А.		
<b>Мамонова О. С.</b> см. Егорова А. Ю.		
<b>Мейханаджян Л. А.</b> см. Зарядов И. С.		

	№	Стр.
<b>Милованова Т. А.</b> см. Зарядов И. С.		
<b>Милостная Н. А.</b> см. Бобырь М. В.		
<b>Морозкин Н. Д.</b> см. Спивак С. И.		
<b>Морозов Н. В.</b> см. Степченков Ю. А.		
<b>Морозов Н. В.</b> см. Степченков Ю. А.		
<b>Мусабаев Р. Р.</b> см. Захаров В. Н.		
<b>Никишин Д. А.</b> Виды неоднородностей в структуре генерализации геоданных	1	74–85
<b>Никишин Д. А.</b> Информационные трансформации в процессах обработки геоданных	2	60–70
<b>Нуриев В. А.</b> см. Гончаров А. А.		
<b>Одинаев Р. Н.</b> Математическая модель задачи защиты растений в биосистеме типа «вредные насекомые – полезные насекомые» с произвольными трофическими функциями	1	96–108
<b>Ожерельев И. С., Сенько О. В., Киселёва Н. Н.</b> Метод поиска выпадающих объектов с использованием параметров неустойчивости обучения	2	122–134
<b>Острикова Д. Ю.</b> см. Бесчастный В. А.		
<b>Палионная С. И.</b> см. Кудрявцев А. А.		
<b>Палионная С. И.</b> см. Кудрявцев А. А.		
<b>Рождественский Ю. В.</b> см. Степченков Ю. А.		
<b>Розенберг И. Н.</b> см. Дулин С. К.		
<b>Рудак А. О.</b> см. Шнурков П. В.		
<b>Румовская С. Б., Колесников А. В., Литвин А. А.</b> Визуальный метаязык описания решения диагностических проблем	4	39–49
<b>Рычихин А. К.</b> О методах оценки качества машинного перевода	4	106–118
<b>Самуйлов К. Е.</b> см. Чухно О. В.		
<b>Сенько О. В.</b> см. Ожерельев И. С.		
<b>Синицын В. И.</b> см. Синицын И. Н.		
<b>Синицын В. И.</b> см. Синицын И. Н.		
<b>Синицын И. Н., Синицын В. И.</b> Аналитическое моделирование процессов в вольтерровских стохастических системах методом канонических разложений	1	109–127
<b>Синицын И. Н., Синицын В. И.</b> Условно-оптимальное линейное оценивание нормальных процессов в вольтерровских стохастических системах	3	16–28
<b>Сороколетова О. Е., Захарова Т. В.</b> Метод случайного отбора при прогнозировании временных рядов рынка криптовалют	4	65–72
<b>Спивак С. И., Морозкин Н. Д., Лукьянов Л. А.</b> Метаобъектный подход к моделированию бизнес-процессов предприятия в рамках единой ERP-системы	2	113–121
<b>Степченков Д. Ю.</b> см. Степченков Ю. А.		
<b>Степченков Д. Ю.</b> см. Степченков Ю. А.		

	№	Стр.
<b>Степченков Ю. А., Дьяченко Ю. Г., Морозов Н. В., Степченков Д. Ю., Дьяченко Д. Ю.</b> Характеризация последовательностных самосинхронных элементов	3	104–113
<b>Степченков Ю. А., Дьяченко Ю. Г., Рождественский Ю. В., Морозов Н. В., Степченков Д. Ю., Дьяченко Д. Ю.</b> Оптимизация индикации многозарядных самосинхронных схем	4	14–27
<b>Стрижов В. В.</b> см. Усманова К. Р.		
<b>Сучков А. П.</b> Функциональная модель системы управления процессами стратегического планирования	4	129–139
<b>Сучков А. П.</b> см. Зацаринный А. А.		
<b>Тимонина Е. Е.</b> см. Грушо А. А.		
<b>Тимонина Е. Е.</b> см. Грушо А. А.		
<b>Тимонина Е. Е.</b> см. Грушо А. А.		
<b>Титова А. И.</b> см. Арутюнов Е. Н.		
<b>Уманский В. И.</b> см. Дулин С. К.		
<b>Усманова К. Р., Стрижов В. В.</b> Модели обнаружения зависимостей во временных рядах в задачах построения прогностических моделей	2	12–30
<b>Хорошилов Ал-др А.</b> см. Захаров В. Н.		
<b>Хорошилов Ал-ей А.</b> см. Захаров В. Н.		
<b>Чаркова В. В.</b> см. Бобылева О. В.		
<b>Чухно Н. В.</b> см. Чухно О. В.		
<b>Чухно Н. В.</b> см. Чухно О. В.		
<b>Чухно О. В., Чухно Н. В., Гайдамака Ю. В., Самуйлов К. Е.</b> Метод извлечения экспертных оценок из социальных сетей при проведении процесса группового принятия решений	4	84–95
<b>Чухно О. В., Чухно Н. В., Гайдамака Ю. В., Шоргин С. Я.</b> Аппроксимация коэффициента усиления направленности антенны для анализа «направленной глухоты» в трехмерном пространстве	3	39–51
<b>Шарнин М. М., Галина И. В.</b> Некоторые аспекты автоматизации поиска и анализа веб-текстов экстремистской семантики	1	206–217
<b>Шарнин М. М.</b> см. Галина И. В.		
<b>Шестаков О. В.</b> Сходимость распределения оценки риска пороговой обработки к смеси нормальных законов при случайном объеме выборки	2	31–38
<b>Шестаков О. В.</b> см. Кудрявцев А. А.		
<b>Шнурков П. В., Рудак А. О.</b> Алгоритмическое решение проблемы оптимального управления в динамической односекторной экономической модели с дискретным временем на основе метода динамического программирования	1	128–139
<b>Шоргин В. С.</b> см. Кудрявцев А. А.		
<b>Шоргин С. Я.</b> см. Чухно О. В.		

---

## 2019 AUTHOR INDEX

---

	No.	Page
<b>Abgaryan K.K. and Gavrilov E.S.</b> Informational support of the multiscale modeling integration platform	1	53–62
<b>Adamovich I.M. and Volkov O.I.</b> The error correction process in the semantic network as a nonlinear dynamic system	3	160–168
<b>Adamovich I.M. and Volkov O.I.</b> The forming of virtual communities process as a nonlinear dynamic system	4	119–128
<b>Adamovich I.M. and Volkov O.I.</b> The principles of data organization for the technology of concrete historical research support	2	161–171
<b>Adamovich I.M. and Volkov O.I.</b> Unified technology of concrete historical research support	1	194–205
<b>Arkipov A.E.</b> see Bobilev M.V.		
<b>Arutyunov E.N., Kudryavtsev A.A., and Titova A.I.</b> <i>A priori</i> Weibull distribution in Bayesian balance models	1	174–179
<b>Bekesheva I.S.</b> see Bobileva O.V.		
<b>Beschastnyi V.A., Ostrikova D.Yu., and Gaidamaka Yu.V.</b> 5G new radio base station performance analysis using limited resource loss system with variable demands	4	73–83
<b>Bobilev V.A.</b> see Bobileva O.V.		
<b>Bobileva O.V., Bekesheva I.S., Bobilev V.A., and Charkova V.V.</b> Fuzzy string comparison when processing personal data	3	149–159
<b>Bobyry M.V., Arkipov A.E., and Milostnaya N.A.</b> The method of depth map calculation based on soft operators	2	71–84
<b>Buntman N.V.</b> see Goncharov A.A.		
<b>Charkova V.V.</b> see Bobileva O.V.		
<b>Charnine M.M. and Galina I.V.</b> Some aspects of automation of search and analysis of web-texts of extremist semantics	1	206–217
<b>Charnine M.M.</b> see Galina I.V.		
<b>Chukhno N.V.</b> see Chukhno O.V.		
<b>Chukhno N.V.</b> see Chukhno O.V.		
<b>Chukhno O.V., Chukhno N.V., Gaidamaka Yu.V., and Samouylov K.E.</b> Method of extracting experts' assessments from social networks in the group decision making process	4	84–95
<b>Chukhno O.V., Chukhno N.V., Gaidamaka Yu.V., and Shorgin S.Ya.</b> Approximation of antenna directivity gain for directional deafness analysis in three-dimensional space	3	39–51

	<b>No.</b>	<b>Page</b>
<b>Diachenko D. Yu.</b> see Stepchenkov Yu. A.		
<b>Diachenko D. Yu.</b> see Stepchenkov Yu. A.		
<b>Diachenko Yu. G.</b> see Stepchenkov Yu. A.		
<b>Diachenko Yu. G.</b> see Stepchenkov Yu. A.		
<b>Dulin S. K., Rozenberg I. N., and Umanskiy V. I.</b> About the problem of information resources integration	3	127–138
<b>Egorov V. B.</b> A practical definition of “software-defined”	2	85–94
<b>Egorov V. B.</b> A way to enhance throughput of packet switches built on the basis of integrated network processors	1	63–73
<b>Egorova A. Yu., Zatsman I. M., and Mamonova O. S.</b> Supracorpora databases in linguistic projects	3	77–91
<b>Gaidamaka Yu. V.</b> see Beschastnyi V. A.		
<b>Gaidamaka Yu. V.</b> see Chukhno O. V.		
<b>Gaidamaka Yu. V.</b> see Chukhno O. V.		
<b>Galina I. V. and Charnine M. M.</b> The science contextual citation index	3	66–76
<b>Galina I. V.</b> see Charnine M. M.		
<b>Garanin A. I.</b> see Zatsarinny A. A.		
<b>Gavrilov E. S.</b> see Abgaryan K. K.		
<b>Goncharov A. A., Buntman N. V., and Nuriev V. A.</b> Machine translation errors: Problems of classification	3	92–103
<b>Goncharov A. A., Inkova O. Yu., and Kruzhkov M. G.</b> Annotation methodology of supracorpora databases	2	148–160
<b>Goncharov A. A. and Zatsman I. M.</b> Information transformations of parallel texts in knowledge extraction	1	180–193
<b>Grusho A. A., Grusho N. A., and Timonina E. E.</b> Methods of protection against mass inactivation of low protected computers of digital economy	1	4–11
<b>Grusho A. A., Zabezhailo M. I., Grusho N. A., and Timonina E. E.</b> Search of empirical causes of failures and errors in computer systems and networks using metadata	4	28–38
<b>Grusho A. A., Zatsarinny A. A., and Timonina E. E.</b> The electronic ledger on the basis of situational centers for digital economy	2	4–11
<b>Grusho N. A.</b> Methods of identification of “weak” signs of violations of information security	3	194–201
<b>Grusho N. A.</b> see Grusho A. A.		
<b>Grusho N. A.</b> see Grusho A. A.		
<b>Ilyin A. V. and Ilyin V. D.</b> Formation of situationally dependent systems of requirements for solving the problems of cost planning	3	169–179
<b>Ilyin V. D.</b> Symbols, codes, signals	4	140–148



	<b>No.</b>	<b>Page</b>
<b>Ilyin V. D.</b> The model of the cooperative problem solver based on digital twins	2	172–179
<b>Ilyin V. D.</b> see Ilyin A. V.		
<b>Inkova O. Yu.</b> see Goncharov A. A.		
<b>Ionenkov Yu. S.</b> see Zatsarinny A. A.		
<b>Ivanov D. V.</b> Conditional bounds of expected maxima of random variables and their reachability	1	140–163
<b>Kalenov N. E.</b> see Vlasova S. A.		
<b>Khoroshilov Al-dr A.</b> see Zakharov V. I.		
<b>Khoroshilov Al-ey A.</b> see Zakharov V. I.		
<b>Kirikov I. A.</b> see Listopad S. V.		
<b>Kiseleva N. N.</b> see Ozhereliev I. S.		
<b>Kodnyanko V. A.</b> On computational redundancy of the dichotomous search and conditional minimization of unimodal functions by the economical dichotomous search	1	164–173
<b>Kolesnikov A. V. and Listopad S. V.</b> Functional structure of the hybrid intelligent multiagent system of heterogeneous thinking for solving the problem of restoring the distribution power grid	1	41–52
<b>Kolesnikov A. V., Listopad S. V., and Maitakov F. G.</b> Intellectualization of operational dispatch control of complex regional dynamic power system	2	46–59
<b>Kolesnikov A. V.</b> see Rumovskaya S. B.		
<b>Kondrashev V. A.</b> see Zatsarinny A. A.		
<b>Kondrashev V. A.</b> see Zatsarinny A. A.		
<b>Korolev V. I.</b> see Zatsarinny A. A.		
<b>Kosolapov Yu. V.</b> The data embedding method based on the secret sharing scheme	3	180–193
<b>Kozlovskaya Y. D.</b> see Zakharov V. I.		
<b>Krasovitskiy A. M.</b> see Zakharov V. I.		
<b>Krivenko M. P.</b> Dimensionality reduction for mixture of probabilistic principal component analyzers in relation to the tasks of medical diagnostics	4	4–13
<b>Krivenko M. P.</b> Selecting the dimensionality for mixture of probabilistic principal component analyzers	3	4–15
<b>Kruzhkov M. G.</b> see Goncharov A. A.		
<b>Kudryavtsev A. A., Palionnaia S. I., and Shestakov O. V.</b> Advantage index in Bayesian reliability and balance models with beta-polynomial <i>a priori</i> densities	3	29–38
<b>Kudryavtsev A. A., Palionnaia S. I., and Shorgin V. S.</b> <i>A priori</i> generalized Frechet distribution in Bayesian balance models	2	39–45
<b>Kudryavtsev A. A.</b> see Arutyunov E. N.		

	<b>No.</b>	<b>Page</b>
<b>Listopad S. V. and Kirikov I. A.</b> Modeling of agent conflicts in hybrid intelligent multiagent systems	3	139–148
<b>Listopad S. V.</b> see Kolesnikov A. V.		
<b>Listopad S. V.</b> see Kolesnikov A. V.		
<b>Litvin A. A.</b> see Rumovskaya S. B.		
<b>Lukyanov L. A.</b> see Spivak S. I.		
<b>Maitakov F. G.</b> see Kolesnikov A. V.		
<b>Malkovsky S. I.</b> see Zatsarinny A. A.		
<b>Mamonova O. S.</b> see Egorova A. Yu.		
<b>Mayorov B. G.</b> Recovery of signal in systems of real time on uniform selections with reduction of Nyquist interval	2	95–112
<b>Meykhanadzhyan L. A.</b> see Zaryadov I. S.		
<b>Milostnaya N. A.</b> see Bobyr M. V.		
<b>Milovanova T. A.</b> see Zaryadov I. S.		
<b>Morozkin N. D.</b> see Spivak S. I.		
<b>Morozov N. V.</b> see Stepchenkov Yu. A.		
<b>Morozov N. V.</b> see Stepchenkov Yu. A.		
<b>Musabaev R. R.</b> see Zakharov V. I.		
<b>Nikishin D. A.</b> Information transformation in the processing of geodata	2	60–70
<b>Nikishin D. A.</b> Types of inhomogeneities in the structure of geodata generalization	1	74–85
<b>Nuriev V. A.</b> see Goncharov A. A.		
<b>Odinaev R. N.</b> A mathematical model of the problem of plant protection in a biosystem of the type “harmful insects – beneficial insects” with arbitrary trophic functions	1	96–108
<b>Ostrikova D. Yu.</b> see Beschastnyi V. A.		
<b>Ozhereliev I. S., Senko O. V., and Kiseleva N. N.</b> Method for searching outlier objects using parameters of learning instability	2	122–134
<b>Palionnaia S. I.</b> see Kudryavtsev A. A.		
<b>Palionnaia S. I.</b> see Kudryavtsev A. A.		
<b>Rogdestvenski Yu. V.</b> see Stepchenkov Yu. A.		
<b>Rozenberg I. N.</b> see Dulin S. K.		
<b>Rudak A. O.</b> see Shnurkov P. V.		
<b>Rumovskaya S. B., Kolesnikov A. V., and Litvin A. A.</b> Visual metalanguage for description of solution of diagnostic problems	4	39–49
<b>Rychikhin A. K.</b> On methods of machine translation quality assessment	4	106–118
<b>Samouylov K. E.</b> see Chukhno O. V.,		
<b>Senko O. V.</b> see Ozhereliev I. S.		

	<b>No.</b>	<b>Page</b>
<b>Shestakov O. V.</b> Convergence of the distribution of the threshold processing risk estimate to a mixture of normal laws at a random sample size	2	31–38
<b>Shestakov O. V.</b> see Kudryavtsev A. A.		
<b>Shnurkov P. V. and Rudak A. O.</b> Algorithmic solution of the problem of optimal control in a dynamic one-sector economic model with discrete time based on the dynamic programming method	1	128–137
<b>Shorgin S. Ya.</b> see Chukhno O. V.		
<b>Shorgin V. S.</b> see Kudryavtsev A. A.		
<b>Sinitsyn I. N. and Sinitsyn V. I.</b> Analytical modeling of processes in Volterra stochastic systems by the canonical expansions method	1	109–127
<b>Sinitsyn I. N. and Sinitsyn V. I.</b> Conditionally optimal linear estimation of normal processes in Volterra stochastic systems	3	16–28
<b>Sinitsyn V. I.</b> see Sinitsyn I. N.		
<b>Sinitsyn V. I.</b> see Sinitsyn I. N.		
<b>Sorokoletova O. E. and Zakharova T. V.</b> Random sampling method for cryptocurrency market time series forecasting	4	65–72
<b>Spivak S. I., Lukyanov L. A., and Morozkin N. D.</b> Metaobject approach for modeling enterprise business processes inside single ERP system	2	113–121
<b>Stepchenkov D. Yu.</b> see Stepchenkov Yu. A.		
<b>Stepchenkov D. Yu.</b> see Stepchenkov Yu. A.		
<b>Stepchenkov Yu. A., Diachenko Yu. G., Morozov N. V., Stepchenkov D. Yu., and Diachenko D. Yu.</b> Sequential self-timed cell characterization	3	104–113
<b>Stepchenkov Yu. A., Diachenko Yu. G., Rogdestvenski Yu. V., Morozov N. V., Stepchenkov D. Yu., and Diachenko D. Yu.</b> Indication optimization in multibit self-timed circuits	4	14–27
<b>Strijov V. V.</b> see Usmanova K. R.		
<b>Suchkov A. P.</b> Functional model of strategic planning process management system	4	129–139
<b>Suchkov A. P.</b> see Zatsarinny A. A.		
<b>Timonina E. E.</b> see Grusho A. A.		
<b>Timonina E. E.</b> see Grusho A. A.		
<b>Timonina E. E.</b> see Grusho A. A.		
<b>Titova A. I.</b> see Arutyunov E. N.		
<b>Umanskiy V. I.</b> see Dulin S. K.		
<b>Usmanova K. R. and Strijov V. V.</b> Models of detection relationship between time series in forecasting problems	2	12–30

	<b>No.</b>	<b>Page</b>
<b>Vakulenko V. V. and Zatsman I. M.</b> Framework for the design of intellectual dictionary systems and their primary functions	4	96–105
<b>Vlasova S. A. and Kalenov N. E.</b> Electronic catalog of the library for natural sciences of the Russian Academy of Sciences as a specialized information retrieval system focused on a qualified user	1	86–95
<b>Volkov O. I.</b> see Adamovich I. M.		
<b>Volkov O. I.</b> see Adamovich I. M.		
<b>Volkov O. I.</b> see Adamovich I. M.		
<b>Volkov O. I.</b> see Adamovich I. M.		
<b>Volovich K. I.</b> see Zatsarinny A. A.		
<b>Zabezhaïlo M. I.</b> see Grusho A. A.		
<b>Zakharov V. I., Musabaev R. R., Krasovitskiy A. M., Kozlovskaya Y. D., Khoroshilov Al-dr A., and Khoroshilov Al-ey A.</b> Clustering method of news media reports based on conceptual analysis	3	52–65
<b>Zakharova T. V.</b> see Sorokoletova O. E.		
<b>Zaryadov I. S., Meykhanadzhyan L. A., and Milovanova T. A.</b> Stationary characteristics of the GI/MSP/ $n/\infty$ queue with general renovation	4	50–64
<b>Zatsarinny A. A., Garanin A. I., Kondrashev V. A., Volovich K. I., and Malkovsky S. I.</b> Evaluation of reliability of the hybrid high-performance computing complex in solution of scientific problems	2	135–147
<b>Zatsarinny A. A. and Ionenkov Yu. S.</b> The method of selecting a variant of the construction of information and telecommunication systems	3	114–126
<b>Zatsarinny A. A., Kondrashev V. A., and Suchkov A. P.</b> The system of scientific services as a relevant component of scientific research	1	25–40
<b>Zatsarinny A. A. and Korolev V. I.</b> Technological aspects of information security in the financial market in terms of digital transformation of the Russian economy	1	12–24
<b>Zatsarinny A. A.</b> see Grusho A. A.		
<b>Zatsman I. M.</b> see Egorova A. Yu.		
<b>Zatsman I. M.</b> see Goncharov A. A.		
<b>Zatsman I. M.</b> see Vakulenko V. V.		

---

## 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 редакции 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/2012.22.02\\_rus/authors.asp](http://www.ipiran.ru/journal/collected/2012.22.02_rus/authors.asp) и  
[http://www.ipiran.ru/journal/collected/2012.22.02\\_eng/authors.asp](http://www.ipiran.ru/journal/collected/2012.22.02_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 (ссылка 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. 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 resursoberegayushchie tekhnologii nedropol’zovaniya i povyshe-niya 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. Mamikonians, eds. 1972. *Ekspluatatsiya turbogeneratorov s neposredstvennym okhlazhdeniem* [Operation of turbine generators with direct cooling]. Moscow: Energy Publ. 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 informatsionnogo 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 rashkoda 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 Pubs. 10 p.

**Описание патента:**

Bolshakov, M. V., A. V. Kulakov, A. N. Lavrenov, and M. V. Palkin. 2006. Spособ orientirovaniya po krenu letatel'nogo apparata s opticheskoy golovkoy 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) (Сейфуль-Мулюков Рустем Бадриевич)

<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 15 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 resursoberegayushchie tekhnologii nedropol'zovaniya i povyshe-niya 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. Mamikonians, eds. 1972. *Ekspluatatsiya turbogenera-torov s neposredstvennym okhlazhdeniem* [Operation of turbine generators with direct cooling]. Moscow: Energy Pubs. 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 kolichstva 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 Pubs. 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 golovkoy 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 Prof. Rustem Seyful-Mulyukov)

[http://www.ipiran.ru/english/journal\\_systems.asp](http://www.ipiran.ru/english/journal_systems.asp)

# SYSTEMS AND MEANS OF INFORMATICS (SISTEMY I SREDSTVA INFORMATIKI)

SCIENTIFIC JOURNAL

Volume 29 No.4 Year 2019

Editor-in-Chief and Chair of Editorial Council  
Academician I. A. Sokolov

## IN THIS ISSUE:

DIMENSIONALITY REDUCTION FOR MIXTURE OF PROBABILISTIC PRINCIPAL COMPONENT ANALYZERS IN RELATION TO THE TASKS OF MEDICAL DIAGNOSTICS <i>M. P. Krivenko</i>	4
INDICATION OPTIMIZATION IN MULTIBIT SELF-TIMED CIRCUITS <i>Yu. A. Stepchenkov, Yu. G. Diachenko, Yu. V. Rogdestvenski, N. V. Morozov, D. Yu. Stepchenkov, and D. Yu. Diachenko</i>	14
SEARCH OF EMPIRICAL CAUSES OF FAILURES AND ERRORS IN COMPUTER SYSTEMS AND NETWORKS USING METADATA <i>A. A. Grusho, M. I. Zabezhailo, N. A. Grusho, and E. E. Timonina</i>	28
VISUAL METALANGUAGE FOR DESCRIPTION OF SOLUTION OF DIAGNOSTIC PROBLEMS <i>S. B. Rumovskaya, A. V. Kolesnikov, and A. A. Litvin</i>	39
STATIONARY CHARACTERISTICS OF THE GI/MSP/ $n/\infty$ QUEUE WITH GENERAL RENOVATION <i>I. S. Zaryadov, L. A. Meykhanadzhyan, and T. A. Milovanova</i>	50
RANDOM SAMPLING METHOD FOR CRYPTOCURRENCY MARKET TIME SERIES FORECASTING <i>O. E. Sorokoletova and T. V. Zakharova</i>	65
5G NEW RADIO BASE STATION PERFORMANCE ANALYSIS USING LIMITED RESOURCE LOSS SYSTEM WITH VARIABLE DEMANDS <i>V. A. Beschastnyi, D. Yu. Ostrikova, and Yu. V. Gaidamaka</i>	73
METHOD OF EXTRACTING EXPERTS' ASSESSMENTS FROM SOCIAL NETWORKS IN THE GROUP DECISION MAKING PROCESS <i>O. V. Chukhno, N. V. Chukhno, Yu. V. Gaidamaka, and K. E. Samouylov</i>	84