

СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ

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

Издается с 1989 года

Журнал выходит ежеквартально

Учредитель:

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

РЕДАКЦИОННЫЙ СОВЕТ

академик РАН И. А. Соколов — председатель Редакционного совета

академик РАН Г. И. Савин

академик РАН А. Л. Стемпковский

член-корреспондент РАН Ю. Б. Зубарев

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

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

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

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

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

академик РАН И. А. Соколов — главный редактор

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

д.т.н. В. Н. Захаров

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

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

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, 2021

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» ВИНИТИ
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 31 № 3 Год 2021

СОДЕРЖАНИЕ

Совместное стационарное распределение в системе $GI/M/n/\infty$ с обобщенным обновлением	
<i>Т. А. Милованова, И. С. Зарядов, Л. А. Мейханаджян</i>	4
Некоторые вероятностно-статистические свойства гамма-экспоненциального распределения	
<i>М. О. Воронцов, А. А. Кудрявцев, О. В. Шестаков</i>	18
Исследование проблемы управления запасом непрерывного продукта в стохастической модели регенерации при наличии двух параметров оптимизации	
<i>П. В. Шнурков, К. А. Адамова</i>	36
Интероперабельность как ключевое условие реализации цифровой трансформации	
<i>И. Н. Розенберг, С. К. Дулин, Н. Г. Дулина</i>	48
Безопасное масштабирование электронных бухгалтерских книг на основе тангла	
<i>А. А. Грушо, А. А. Зацаринный, Е. Е. Тимонина</i>	60
Вычисления на основе вероятностной модели анализа главных компонент	
<i>М. П. Кривенко</i>	70
Использование надежной грамматики при информационном поиске в базе знаний экспертной системы о конструкциях летательных аппаратов	
<i>Н. И. Сидняев, Ю. И. Бутенко, Е. Е. Синева</i>	80
Методика оценки производственных рисков разработки средств вооружения и военной техники	
<i>А. В. Босов, А. А. Крюков</i>	88
Концепция построения надкорпусных баз данных	
<i>М. Г. Кружков</i>	101

СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ

Том 31 № 3 Год 2021

СОДЕРЖАНИЕ

Аппаратная реализация рекуррентного обработчика сигналов Ю. А. Степченко, Н. В. Морозов, Ю. Г. Дьяченко, Д. В. Хилько	113
Модель для анализа приоритетного доступа трафика URLLC при прерывании обслуживания и снижении скорости передачи сессий eMBB в сети 5G И. А. Кочеткова, А. И. Кущазли, П. А. Харин, С. Я. Шоргин	123
Пример применения аппарата нейронных сетей при назначении модуляционно-кодовой схемы планировщиком базовой станции сети 5G Е. В. Бобрикова, А. А. Платонова, Ю. В. Гайдамака, С. Я. Шоргин	135
Экспертная оценка машинного перевода: классификация ошибок А. Ю. Егорова, И. М. Зацман, В. А. Нуриев	144
Использование геоинформационных систем в технологии поддержки конкретно-исторических исследований И. М. Адамович, О. И. Волков	158
Символьное моделирование задач и конструирование программ В. Д. Ильин	170
Поправка к статье М. О. Воронцова, А. А. Кудрявцева, С. Я. Шоргина «Аналитические свойства и аспекты вычисления гамма-экспоненциальной функции» (Системы и средства информатики, 2021. Т. 31. № 2. С. 108–118)	178
Об авторах	179
Правила подготовки рукописей статей	183
Requirements for manuscripts	187

СОВМЕСТНОЕ СТАЦИОНАРНОЕ РАСПРЕДЕЛЕНИЕ В СИСТЕМЕ $GI/M/n/\infty$ С ОБОБЩЕННЫМ ОБНОВЛЕНИЕМ*

Т. А. Милованова¹, И. С. Зарядов², Л. А. Мейханаджян³

Аннотация: Рассматривается многолинейная система массового обслуживания (СМО) с конечным числом идентичных приборов и одной очередью неограниченной емкости. Заявки поступают в систему по одной в соответствии с рекуррентным потоком. Времена обслуживания заявок на приборах имеют экспоненциальное распределение с одним и тем же параметром. В системе реализован механизм обобщенного обновления: заявка, обслуживание которой было закончено, в момент ухода из системы удаляет из очереди случайное число заявок в соответствии с заранее заданным вероятностным распределением. Предложен метод нахождения совместного стационарного распределения общего числа заявок в системе и времени, прошедшего с момента последнего поступления заявки в систему. Представлены формулы в терминах преобразований для вычисления совместного распределения в нестационарном режиме.

Ключевые слова: система массового обслуживания; обобщенное обновление; управление очередью

DOI: 10.14357/08696527210301

1 Описание системы

Рассмотрим систему, состоящую из n идентичных приборов и одной очереди неограниченной емкости. Из очереди заявки обслуживаются по одной по правилу FIFO (first in, first out), LIFO (last in, first out) или RANDOM. Входящий поток заявок — рекуррентный, причем время между соседними поступлениями заявок имеет произвольную функцию распределения $A(x)$. От функции $A(x)$ потребуем, чтобы существовала (непрерывная ограниченная) плотность $a(x) = A'(x)$. Положим

$$\lambda(x) = \frac{a(x)}{1 - A(x)}, \quad x \geq 0.$$

Времена обслуживания заявок на приборах являются независимыми случайными величинами, распределенными по экспоненциальному закону с параметром μ .

* Исследование выполнено при финансовой поддержке РФФИ (проекты 19-07-00739 и 20-07-00804).

¹ Российский университет дружбы народов, milovanova-ta@rudn.ru

² Российский университет дружбы народов; Федеральный исследовательский центр «Информатика и управление» Российской академии наук, zaryadov-is@rudn.ru

³ Финансовый университет при Правительстве РФ, lamejkhanaadzhyan@fa.ru

В системе реализован механизм обобщенного обновления [1–4], работающий следующим образом. Находящаяся на приборе заявка в момент окончания обслуживания одновременно со своим уходом из системы либо с вероятностью $q_l \geq 0$, $l = 0, 1, \dots$, удаляет из очереди 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. Отметим, что

$$Q_0 = \sum_{l=0}^{\infty} q_l = 1.$$

Следуя [2], вероятности q_l будем называть вероятностями обновления.

Интерес к механизму обновления связан, в частности, с потенциальными возможностями его применения в качестве алгоритма активного управления очередями (см., например, [5]). Несмотря на то что он и позволяет вывести систему на (по крайней мере тот же) уровень производительности, что достигается классическими RED-подобными (RED — random early detection, произвольное раннее обнаружение) алгоритмами [6], реализация в узлах сетей связи этого механизма в том виде, как это описано выше, трудноосуществима. Связано это с предписанием осуществлять управление после каждого момента окончания обслуживания. Как показывают вычислительные эксперименты, есть способы отказаться от этой установки без существенной потери в качестве управления. К примеру, можно «обновлять» очередь периодически (например, через заранее фиксированное число окончаний обслуживания¹). Варьируя значения нового параметра — периода обновления — и вероятностей обновления $\{q_l, l \geq 0\}$, можно добиться (в некоторых случаях) практически того же уровня производительности, что достигается с помощью классического варианта обновления. Другая, вполне естественная, но неизученная его модификация подразумевает, что вероятность обновления q_l при каждом l является не константой, а функцией какой-то (численной) характеристики системы в момент окончания обслуживания; при этом управление по-прежнему осуществляется в каждый момент окончания обслуживания. Для получения каких-либо аналитических результатов для СМО с подобной модификацией механизма обновления необходимо уметь вычислять совместные распределения ее основных характеристик.

В [2] для описанной в начале этого раздела СМО были получены математические соотношения для вычисления стационарных распределений основных характеристик — числа заявок в системе по вложенной цепи Маркова и по времени. В явном виде были получены распределения времен пребывания в очереди (при прямом порядке обслуживания) «убитой», обслуженной и произвольной заявок. Среди всевозможных вариантов зависимости вероятности обновления q_l от состояния системы в момент окончания обслуживания наиболее простой и со-

¹Отметим, что счетчик числа завершений можно организовать по-разному: считать все окончания обслуживания либо начиная с момента начала функционирования системы, либо с момента последнего опустошения системы.

держательной представляется зависимость q_l от прошедшего времени обслуживания¹. Однако получение совместного стационарного распределения числа заявок в системе и прошедшего времени обслуживания предложенными в [2] методами невозможно. В этой статье показано, что задача определения совместного стационарного распределения сводится к решению системы рекуррентных дифференциальных уравнений (для плотностей) с обычными граничными условиями, и предложен реализуемый в вычислительном плане алгоритм ее расчета.

2 Основная система уравнений

Пусть $\nu(t)$ — число заявок в системе в момент времени t , а $\xi(t)$ — время, прошедшее с момента поступления последней заявки в момент времени t . Аналогично тому, как это делается для классической системы $M/G/1/\infty$ [7, с. 289], показывается, что $\{(\nu(t), \xi(t)), t \geq 0\}$ — линейчатый марковский процесс.

Введем обозначения:

$$P_j(t, x) = P\{\nu(t) = j, \xi(t) < x\};$$

$$p_j(t, x) = \frac{\partial}{\partial x} P_j(t, x), \quad x \geq 0, \quad j = 0, 1, \dots$$

Относительно плотностей $p_j(t, x)$ будем предполагать, что они существуют и являются непрерывными [7, гл. 5]. Рассматривая моменты t и $t + \Delta$ и рассуждая стандартным образом, получаем систему линейных дифференциальных уравнений:

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x}\right) p_j(t, x) = -(\lambda(x) + j\mu)p_j(t, x) + (j+1)\mu p_{j+1}(t, x),$$

$$0 \leq j \leq n-2; \quad (1)$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x}\right) p_{n-1}(t, x) =$$

$$= -(\lambda(x) + (n-1)\mu)p_{n-1}(t, x) + n\mu \sum_{m=n}^{\infty} p_m(t, x)Q_{m-n}; \quad (2)$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x}\right) p_j(t, x) = -(\lambda(x) + n\mu)p_j(t, x) + n\mu \sum_{m=j+1}^{\infty} p_m(t, x)q_{m-j-1},$$

$$j \geq n. \quad (3)$$

Поскольку эргодичность линейчатого процесса эквивалентна эргодичности его вложенной цепи, то при выполнении условия $(\mu n \sum_{k=0}^{\infty} (k+1)q_k)^{-1} < EA$ (полученном в [2]) существуют стационарные вероятности $P_j(x) = \lim_{t \rightarrow \infty} P_j(t, x)$ (и плотности $p_j(x) = P'_j(x)$).

¹Отметим, что значение этой характеристики, в отличие от остаточного времени обслуживания, является наблюдаемым.

3 Совместное стационарное распределение

Введем функции $r_j(x) = p_j(x)/(1 - A(x))$. Система уравнений для r_j следует из (1)–(3) и имеет вид:

$$r'_j(x) = -j\mu r_j(x) + (j+1)\mu r_{j+1}(x), \quad 0 \leq j \leq n-2; \quad (4)$$

$$r'_{n-1}(x) = -(n-1)\mu r_{n-1}(x) + n\mu \sum_{m=n}^{\infty} r_m(x) Q_{m-n}; \quad (5)$$

$$r'_j(x) = -n\mu r_j(x) + n\mu \sum_{m=j+1}^{\infty} q_{m-j-1} r_m(x), \quad j \geq n. \quad (6)$$

Для ее решения необходимы дополнительные условия на $r_j(0)$, которые получаются из граничных условий для $p_j(0)$. Стандартные рассуждения (которые могут быть восстановлены, например, по [7, с. 291]) приводят к равенствам:

$$r_j(0) = \int_0^{\infty} r_{j-1}(u) dA(u), \quad j \geq 1. \quad (7)$$

Введем обозначения:

$$\begin{aligned} \varrho_j(s) &= \int_0^{\infty} e^{-sx} r_j(x) dx; \\ \vec{\varrho} &= (\varrho_n(s), \varrho_{n+1}(s), \dots); \\ \vec{r} &= \left(\frac{r_n(0)}{s + n\mu}, \frac{r_{n+1}(0)}{s + n\mu}, \dots \right) \end{aligned}$$

и перейдем в (4)–(6) к преобразованию Лапласа (ПЛ). Перепишывая (6) с учетом только что введенных обозначений, получаем матричное уравнение $\vec{\varrho} \mathbb{T} = \vec{r}$, в котором $\mathbb{T}$ — полуциркулянтная матрица [8, с. 14] вида

$$\mathbb{T} = \begin{pmatrix} 1 & 0 & 0 & 0 \dots \\ -\frac{n\mu}{s+n\mu} q_0 & 1 & 0 & 0 \dots \\ -\frac{n\mu}{s+n\mu} q_1 - \frac{n\mu}{s+n\mu} q_0 & 1 & 0 & 0 \dots \\ -\frac{n\mu}{s+n\mu} q_2 - \frac{n\mu}{s+n\mu} q_1 - \frac{n\mu}{s+n\mu} q_0 & 1 & 0 & 0 \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Поскольку все диагональные элементы матрицы $\mathbb{T}$ отличны от нуля, она обратима (см., например, [8, теорема 1.3]). Обратная матрица $\mathbb{T}^{-1}$, так же как и $\mathbb{T}$, является полуструктурной и имеет следующую структуру:

$$\mathbb{T}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots \\ \tau_1(s) & 1 & 0 & 0 & \cdots \\ \tau_2(s) & \tau_1(s) & 1 & 0 & \cdots \\ \tau_3(s) & \tau_2(s) & \tau_1(s) & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

где

$$\tau_i(s) = \frac{n\mu}{s + n\mu} q_{i-1} + \frac{n\mu}{s + n\mu} \sum_{j=1}^{i-1} \tau_j(s) q_{i-1-j}, \quad i \geq 1, \quad s \geq 0. \quad (8)$$

Функция τ_i есть ПЛ некоторой непрерывной и ограниченной функции, скажем $t_i(x)$, $x \geq 0$, которую нетрудно найти, обращая (8). В дальнейшем, однако, вместо явного вида $t_i(x)$ понадобятся лишь выражения для функций $\int_0^x t_i(u) e^{n\mu u} du$, $x \geq 0$, которые могут быть найдены из (8) следующим образом. Положим $\tau(z, s) = \sum_{i=1}^{\infty} z^i \tau_i(s)$, $z \in [0, 1]$. Домножая обе части (8) на z^i и суммируя по всем i , получаем:

$$\tau(z, s) = \frac{n\mu z Q(z)}{s + n\mu - n\mu z Q(z)},$$

где $Q(z) = \sum_{i=0}^{\infty} z^i q_i$ — производящая функция (ПФ) числа заявок, удаляемых из системы при обновлении. Обратное по s преобразование $\tau(z, s)$ приводит к равенству

$$\sum_{i=1}^{\infty} z^i e^{n\mu u} t_i(u) = n\mu z Q(z) e^{n\mu z Q(z)u}, \quad z \in [0, 1], \quad u \geq 0,$$

из которого, проинтегрировав по u в пределах от 0 до произвольного $x > 0$, имеем:

$$\sum_{i=1}^{\infty} z^i \int_0^x e^{n\mu u} t_i(u) du = e^{n\mu z Q(z)x} - 1, \quad z \in [0, 1], \quad x \geq 0.$$

Вводя для сокращения записи обозначение $f_n(z) = n\mu z Q(z)$ и дифференцируя левую и правую части предыдущего соотношения k раз по z в точке $z = 0$, находим

$$\int_0^x e^{n\mu u} t_k(u) du = \frac{1}{k!} \sum_{j=1}^k B_{k,j} \left(f'_n(0), f''_n(0), \dots, f_n^{(k-j+1)}(0) \right) x^j, \quad k \geq 1, \quad (9)$$

где через $B_{k,j}$ обозначены частичные полиномы Белла¹ [9, с. 46] и [10]. Таким образом, $\int_0^x e^{n\mu u} t_k(u) du$ есть не что иное, как многочлен степени k от переменной x без свободного члена. Отметим, что коэффициенты $B_{k,j}$ в (9) зависят только от производных функции f_n в нуле (причем $f_n^{(k)}(0) = k!n\mu q_{k-1}$) и для их расчета можно действовать стандартным образом [9, с. 61] или использовать специальные методы (см., например, [11, 12]).

Вернемся к уравнению $\vec{q}\mathbb{T} = \vec{r}$. Записывая его решение в скалярной форме и проводя обратное преобразование (с учетом (9)), получаем явный вид функций r_j при $j \geq n$:

$$r_j(x) = e^{-n\mu x} \sum_{m=j}^{\infty} \frac{r_m(0)}{(m-j)!} \sum_{k=0}^{m-j} x^k B_{m-j,k}, \quad (10)$$

в котором, однако, коэффициенты $r_j(0)$ остаются неизвестными. Для их нахождения воспользуемся граничными условиями (7). Обозначим через $\alpha(s)$ ПЛ $a(x)$, а через $\alpha^{(k)}(s)$ — k -ю производную $\alpha(s)$ по s . Подставляя $r_j(x)$ в приведенном выше виде в (7), приводя подобные слагаемые, с учетом обозначений $\beta_{n,j}(x) = (1/j!) \sum_{k=1}^j (-1)^k \alpha^{(k)}(nx) B_{jk}$ и $\beta_{n,0}(x) = \alpha(nx)$, $x > 0$, получаем

$$r_j(0) = \sum_{k=0}^{\infty} \beta_{n,k}(\mu) r_{j+k-1}(0) \quad j \geq n+1. \quad (11)$$

Отметим, что эта система уравнений (с точностью до обозначений) совпадает с системой уравнений равновесия для стационарных вероятностей по вложенной цепи Маркова для классической системы G/M/1/∞. Аналогичным образом, но из (5) и (7) сначала находится уравнение для $r_{n-1}(x)$, а затем и для $r_n(0)$:

$$r_{n-1}(x) = e^{-(n-1)\mu x} r_{n-1}(0) + \sum_{k=n}^{\infty} g_{n,k}(x; \mu) r_k(0); \quad (12)$$

$$r_n(0) = \alpha((n-1)\mu) r_{n-1}(0) + \sum_{k=n}^{\infty} \gamma_{n,k}(\mu) r_k(0), \quad (13)$$

где $\gamma_{n,m}(y) = \int_0^{\infty} g_{n,m}(u; y) a(u) du$, а значения функции $g_{n,m}$ в точке $(x; y)$ вычисляются по формуле

$$g_{n,m}(x; y) = n e^{xy} \sum_{j=n}^m Q_{j-n} \sum_{k=0}^{m-j} \frac{k! B_{m-j,k}}{(m-j)! \mu^k} \left(e^{-n\mu x} - \sum_{i=0}^k \frac{(xy)^i}{i!} e^{-(n+1)\mu x} \right),$$

$$x \geq 0, \quad y > 0.$$

¹ Отметим, что $B_{0,0} = 1$, $B_{0,m} = 0$ и $B_{m,0} = 0$, $m \geq 1$,

Для решения системы (11), (13) недостает только уравнения для $r_{n-1}(0)$ и (уравнений) для входящих в него слагаемых, содержащих неизвестные значения $r_{n-2}(0), \dots, r_1(0)$. Их вывод начнем с уравнения (4), решение которого (методом вариации постоянной) имеет вид:

$$r_j(x) = e^{-j\mu x} r_j(0) + (j+1)\mu e^{-j\mu x} \int_0^x e^{j\mu u} r_{j+1}(u) du, \quad 0 \leq j \leq n-2. \quad (14)$$

Использование в граничных условиях (7) этого представления для функций r_j затруднительно ввиду наличия интеграла во втором слагаемом в правой части (14). Преодолеть эту трудность можно, например, путем следующих построений. Пусть последовательность положительных при $x \geq 0$ функций $\{d_k(x), k \geq 1\}$ такова, что

$$d_1(x) = \frac{1}{\mu} (1 - e^{-\mu x}); \quad d_k(x) = \int_0^x \mu e^{-\mu u} d_{k-1}(u) du, \quad k \geq 2.$$

Непосредственной проверкой нетрудно убедиться, что общий член этой последовательности равен¹

$$d_k(x) = \sum_{i=1}^{k-1} \frac{(-1)^{k-i-1}}{(k-i)!} d_i(x) + \frac{(-1)^{k-1}}{k!} \frac{1 - e^{-k\mu x}}{\mu}, \quad k \geq 1.$$

Положим $d_0(x) \equiv 1$. Тогда при $1 \leq j \leq n-2$ для интеграла $\int_0^x e^{j\mu u} r_{j+1}(u) du$, с учетом найденного в (10) и (14) явного вида функций r_j , имеет место следующее представление:

$$\begin{aligned} \int_0^x e^{j\mu u} r_{j+1}(u) du &= \\ &= e^{-(n-j)\mu x} \sum_{t=n}^{\infty} r_t(0) \sum_{i=n-j}^{\infty} x^i \Theta_{t,i}^{(j+1)} + \sum_{i=j+1}^{n-1} \frac{i!}{(j+1)!} d_{i-j}(x) r_i(0), \end{aligned} \quad (15)$$

где коэффициенты $\Theta_{t,i}^{(j+1)}$ вычисляются рекуррентным образом (начиная с $j = n-2$) по формуле

¹Напомним, что всюду для сокращения записи используется соглашение $\sum_{i=0}^{-1} = 0$.

$$\Theta_{t,i}^{(j+1)} = \begin{cases} \sum_{j=n}^t \sum_{k=1}^{i-1} \sum_{s=0}^{\min(i-2, t-j)} Q_{j-n} \frac{s! n 2^{i-k} \mu^{i-s-1}}{i!(t-j)!} B_{t-j,s}, & j = n-2; \\ \sum_{k=n-j-1}^{i-1} \Theta_{t,k}^{(j+2)} \frac{k!(j+2)\mu((n-j)\mu)^{i-k-1}}{i!}, & 0 \leq j \leq n-3. \end{cases}$$

Воспользуемся теперь граничными условиями (7) при $1 \leq j \leq n-1$. Подставляя в них выражение для $r_{j-1}(x)$ с учетом (15) и вводя обозначение¹

$$\delta_{ij} = \int_0^\infty e^{-j\mu x} \mu d_{i-j}(x) dA(x),$$

получаем дополнительные $n-1$ соотношений между $\{r_j(0), j \geq 0\}$:

$$r_j(0) = j\mu \sum_{t=n}^\infty r_t(0) \sum_{i=n-j+1}^\infty (-1)^i \alpha^{(i)}(n\mu) \Theta_{t,i}^{(j)} + \sum_{i=j-1}^{n-1} \frac{i! \delta_{i,j-1}}{(j-1)!} r_i(0),$$

$$1 \leq j \leq n-1. \quad (16)$$

Перейдем к решению системы уравнений (11), (13), (16). Для начала заметим, что поскольку $A(+0) = 0$, то $r_j(0)/\sum_{j=1}^\infty r_j(0)$, $j \geq 1$, — суть стационарные вероятности² вложенной цепи Маркова по моментам (сразу после) поступления заявки в систему. Отсюда следует, что ряд $\sum_{j=1}^\infty r_j(0)$ сходится.

Пусть ρ — решение уравнения $\rho = \sum_{k=0}^\infty \beta_{n,k}(\mu) \rho^k$; такое ρ существует, единственно и $\rho \in (0, 1)$. Положим $r_j(0) = r_n(0) \rho^{j-n}$. Тогда вместо бесконечной системы (11), (13), (16) для $\{r_j(0), j \geq 0\}$ имеем следующую систему линейных алгебраических уравнений для $\{r_j(0), 0 \leq j \leq n\}$:

$$r_j(0) = \alpha((j-1)\mu) r_{j-1}(0) + \sum_{i=j}^{n-1} \frac{i! \delta_{i,j-1}}{(j-1)!} r_i(0) + \gamma_j r_n(0), \quad 1 \leq j \leq n, \quad (17)$$

¹Отметим, что константы δ_{ij} могут быть рассчитаны рекуррентно (при фиксированном j последовательно по i) по формуле:

$$\delta_{ij} = \sum_{m=1}^{i-j-1} \frac{(-1)^{i-j-m-1}}{(i-j-m)!} \delta_{m+j,j} + \frac{(-1)^{i-j-1}}{(i-j)!} (\alpha(j\mu) - \alpha(i\mu)).$$

²Которые существуют при выполнении условия, приведенного в конце разд. 2.

где

$$\gamma_j = \begin{cases} j\mu \sum_{t=n}^{\infty} \rho^{t-n} \sum_{i=n-j+1}^{\infty} (-1)^i \alpha^{(i)}(n\mu) \Theta_{t,i}^{(j)}, & 1 \leq j \leq n-1; \\ \sum_{k=n}^{\infty} \gamma_{n,k}(\mu) \rho^{k-n}, & j = n. \end{cases}$$

Как обычно, к такой системе уравнений необходимо добавить условие нормировки, которое имеет вид:

$$\sum_{j=0}^{\infty} \int_0^{\infty} r_j(u)(1 - A(u)) du = 1.$$

После некоторых простых, но утомительных преобразований его можно переписать следующим образом:

$$\begin{aligned} & r_0(0)EA + \sum_{i=1}^{n-1} r_i(0) \sum_{j=0}^i \frac{i!\mu}{j!} \int_0^{\infty} e^{-j\mu x} (1 - A(x)) d_{i-j}(x) dx + \\ & + r_n(0) \sum_{l=n}^{\infty} \left[\sum_{j=0}^{n-2} \sum_{k=n-j}^{\infty} \frac{k!(j+1)\mu \Theta_{l,k}^{(j+1)} \rho^{l-n}}{(n\mu)^{k+1}} \left(1 - \sum_{i=0}^k \frac{(-n\mu)^i}{i!} \alpha^{(i)}(n\mu) \right) + \right. \\ & + \sum_{j=n}^l \sum_{k=0}^{l-j} \frac{k! B_{l-j,k} \rho^{l-n}}{(n\mu)^{k+1} (l-j)!} \left\{ 1 - \frac{n^{k+2} \alpha((n-1)\mu) - n}{n-1} Q_{j-n} + \right. \\ & \left. \left. + \sum_{i=0}^k \frac{(-n\mu)^i}{i!} \alpha^{(i)}(n\mu) \left(\frac{n^{k-i+2} - n}{n-1} Q_{j-n} - 1 \right) \right\} \right] = 1. \quad (18) \end{aligned}$$

Решение системы (17), (18) уже не связано с принципиальными трудностями и может быть получено многими способами: например, хороший результат дает метод исключения состояний (см. [13, разд. 2] и [7, 14]). После расчета $\{r_j(0), 0 \leq j \leq n\}$ и определения остальных значений $\{r_j(0), j \geq n+1\}$ по формуле $r_j(0) = r_n(0)\rho^{j-n}$, вычисляются функции $r_j(x)$ по формулам (10), (12), (14), совпадающие со стационарными плотностями $p_j(x)$ с точностью до множителя $(1 - A(x))$.

4 Заключение

В том случае, когда $A(x)$ является функцией распределения марковского типа, целесообразнее перейти (с заменой непрерывного аргумента на дискретный)

к известным матричным методам анализа марковских систем обслуживания [15]. Если же емкость очереди ограничена, то вычисления несколько упрощаются: отпадает необходимость суммировать ряды с бесконечным числом слагаемых и плотности $p_j(x)$ могут быть вычислены рекуррентным образом.

Скажем несколько слов и о расчете нестационарного распределения $p_j(t, x)$. Даже в простейшем случае, когда $n = 1$, изложенные в статье результаты не позволяют продвинуться далеко. Будем считать, что в момент начала функционирования ($t = 0$) система пуста. С учетом этого начальные условия для системы (1)—(3) выберем следующим образом:

$$p_0(0, x) = \frac{1 - A(x)}{EA}; \quad p_j(0, x) = 0, \quad j \geq 1, \quad x \geq 0.$$

Тогда, применяя к системе линейных дифференциальных уравнений в частных производных первого порядка (1)—(3) двойное ПЛ (по t и по x), с учетом обозначения

$$\varrho_j(v, x) = \int_0^\infty e^{-vx} r_j(t, x) dt$$

приходим к системе уравнений, которую можно решить, следуя по изложенному в предыдущем разделе пути. Прделав это и произведя обратное по s преобразование, находим

$$\varrho_0(v, x) = \frac{1}{vEA} (1 - e^{-vx}) + e^{-vx} \varrho_0(v, 0) + \mu e^{-vx} \sum_{m=1}^{\infty} Q_{m-1} \int_0^x e^{vu} \varrho_m(v, u) du;$$

$$\varrho_j(v, x) = e^{-(v+\mu)x} \varrho_j(v, 0) + e^{-(v+\mu)x} \sum_{m=j+1}^{\infty} \varrho_j(v, 0) \sum_{i=1}^{m-j} \frac{B_{m-j,i}}{(m-j)!} x^i, \quad j \geq 1.$$

Однако в этих соотношениях неизвестны значения функций ϱ_j в точке $(v, 0)$; они (как выше значения $r_j(0)$) могут быть найдены путем решения системы уравнений:

$$\varrho_1(v, 0) = \frac{1}{vEA} (1 - \alpha(v)) + \alpha(v) \varrho_0(v, 0) + \sum_{m=1}^{\infty} \gamma_{1,m}(v + \mu) \varrho_m(v, 0);$$

$$\varrho_j(v, 0) = \sum_{k=0}^{\infty} \beta_{1,k}(v + \mu) \varrho_{j+k-1}(v, 0), \quad j \geq 2,$$

с условием

$$\sum_{j=0}^{\infty} \int_0^\infty \varrho_j(v, u) (1 - A(u)) du = v^{-1}.$$

Однако такой результат трудно использовать на практике. Действительно, например, для расчета среднего числа заявок в системе в некоторый один момент времени необходимо обратиться (воспользовавшись известными методами, см., например, [16, 17]) выражение

$$\sum_{j=1}^{\infty} j \int_0^{\infty} \varrho_j(v, u)(1 - A(u)) du,$$

что потребует¹ многократного решения приведенной выше системы уравнений.

Несомненный интерес в плане дальнейших исследований представляют два вопроса. Первый — это существование функциональной зависимости вероятностей q_i от прошедшего времени обслуживания, которая позволила бы решить систему (1)–(3) (хотя бы в стационарном случае), следуя либо по намеченному выше пути, либо по какому-то другому. Второй вопрос — это разработка эффективных методов расчета нестационарного распределения $p_j(t, x)$.

Литература

1. 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. 14–23.
2. 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.
3. Зарядов И. С. Стационарные временные характеристики системы $G/M/n/r$ с некоторыми вариантами дисциплины обобщенного обновления // Информационные процессы, 2008. Т. 8. № 2. С. 108–124.
4. Зарядов И. С. Система массового обслуживания $GI/M/n/\infty$ с обобщенным обновлением // Автоматика и телемеханика, 2010. № 4. С. 130–139.
5. 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.
6. 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.
7. Бочаров П. П., Печинкин А. В. Теория массового обслуживания. — М.: РУДН, 1995. 529 с.
8. Henrici P. Applied and computational complex analysis. — New York, NY, USA: John Wiley & Sons, 1974. Vol. 1. 682 p.
9. Риордан Дж. Введение в комбинаторный анализ / Пер. с англ. — М.: ИЛ, 1963. 288 с. (Riordan J. An introduction to combinatorial analysis. — 2nd ed. — Wiley publication in mathematical statistics ser. — Wiley, 1958. 244 p.)

¹Помимо преодоления возможных технических сложностей с выбором весов и узлов.

10. Маричев О. И., Славянов С. Ю., Брычков Ю. А. Полиномы Белла в системе Mathematica и асимптотические решения интегральных уравнений // Теоретическая и математическая физика, 2019. Т. 201. № 3. С. 446–456.
11. Cvijovic D. New identities for the partial Bell polynomials // Appl. Math. Lett., 2011. Vol. 24. P. 1544–1547. doi: 10.1016/j.aml.2011.03.043.
12. Qi F., Niu D.-W., Lim D., Yao Y.-H. Special values of the Bell polynomials of the second kind for some sequences and functions // J. Math. Anal. Appl., 2020. Vol. 491. Iss. 2. Art. 124382. 31 p. doi: 10.1016/j.jmaa.2020.124382.
13. 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. 127–142.
14. Печинкин А. В., Разумчик Р. В. Системы массового обслуживания в дискретном времени. — М.: Физматлит, 2018. 432 с.
15. Вишнеvский В. М., Дудин А. Н., Клименок В. И. Стохастические системы с коррелированными потоками. Теория и применение в телекоммуникационных сетях. — М.: Техносфера, 2018. 564 с.
16. Abate J., Whitt W. A unified framework for numerically inverting Laplace transforms // INFORMS J. Comput., 2006. Vol. 18. Iss. 4. P. 408–421. doi: 10.1287/ijoc.1050.0137.
17. Horvath I., Horvath G., Almousa S. A.-D., Telek M. Numerical inverse Laplace transformation using concentrated matrix exponential distributions // Performance Evaluation, 2019. Vol. 137. Art. 102067. 30 p. doi: 10.1016/j.peva.2019.102067.

Поступила в редакцию 15.08.21

JOINT STATIONARY DISTRIBUTION IN THE GI/M/n/∞ QUEUE WITH GENERAL RENOVATION

T. A. Milovanova¹, I. S. Zaryadov^{1,2}, and L. A. Meykhanadzhyan³

¹Peoples’ Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation

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

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

Abstract: Multiserver queuing system with a finite number of identical servers and one queue of unlimited capacity is being considered. Customers enter the system one by one in accordance with a recurrent flow. Service times are exponentially distributed with the same parameter. General renovation mechanism is implemented in the system: a customer, whose service has been completed, upon leaving the system removes a random number of other customers from the queue according to a given probability distribution. The method is

proposed for finding the joint stationary distribution of the total number of customers in the system and the time elapsed since the last arrival. Expressions (in terms of transforms) for the calculation of the transient joint distribution are presented.

Keywords: queueing system; general renovation; queue management

DOI: 10.14357/08696527210301

Acknowledgments

The reported study was funded by RFBR, projects Nos. 19-07-00739 and 20-07-00804.

References

1. Bocharov, P. P., and I. S. Zaryadov. 2007. Statsionarnoe raspredelenie veroyatnostey v sistemakh massovogo obsluzhivaniya s obnovleniem [Queueing systems with renovation. Stationary probability distribution]. *Vestnik RUDN. Ser. Matematika. Informatika. Fizika* [Bulletin of Peoples' Friendship University of Russia. Ser. Mathematics. Information Sciences. Physics] 1-2:14–23.
2. Zaryadov, I. S. 2008. Statsionarnye 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* [Bulletin of Peoples' Friendship University of Russia. Ser. Mathematics. Information Sciences. Physics] 2:3–9.
3. Zaryadov, I. S. 2008. Statsionarnye vremennye kharakteristiki sistemy $G/M/n/r$ s nekotorymi variantami distsipliny obobshchonnogo obnovleniya [Stationary temporal characteristics of the $G/M/n/r$ system with some variations of the generalized renovation discipline]. *Informatsionnye protsessy* [Information Processes] 8(2):108–124.
4. Zaryadov, I. S. 2010. The $GI/M/n/\infty$ queueing system with generalized renovation. *Automat. Rem. Contr.* 71(4): 663–671. doi: 10.1134/S0005117910040077.
5. Konovalov, 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.
6. 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.
7. Bocharov, P. P., and A. V. Pechinkin. 1995. *Teoriya massovogo obsluzhivaniya* [Queueing theory]. Moscow: RUDN. 529 p.
8. Henrici, P. 1974. *Applied and computational complex analysis*. New York, NY: John Wiley & Sons. Vol. 1. 682 p.
9. Riordan, J. 1958. *An introduction to combinatorial analysis*. 2nd ed. Wiley publication in mathematical statistics ser. Wiley. 244 p.
10. Marichev, O. I., S. Yu. Slavyanov, and Yu. A. Brychkov. 2019. Bell Polynomials in the Mathematica system and asymptotic solutions of integral equations. *Theor. Math. Phys.* 201(3):1798–1807.

11. Cvijovic, D. 2011. New identities for the partial Bell polynomials. *Appl. Math. Lett.* 24:1544–1547. doi: 10.1016/j.aml.2011.03.043.
12. Qi, F., D.-W. Niu, Lim D., and Y.-H. Yao. 2020. Special values of the Bell polynomials of the second kind for some sequences and functions. *J. Math. Anal. Appl.* 491(2):124382. 31 p. doi: 10.1016/j.jmaa.2020.124382.
13. 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. doi: 10.1023/A:1022219232282.
14. Pechinkin, A. V., and R. V. Razumchik. 2018. *Sistemy massovogo obsluzhivaniya v diskretnom vremeni* [Discrete time queueing systems]. Moscow: Fizmatlit. 432 p.
15. Vishnevskii, V. M., A. N. Dudin, and V. I. Klimenok. 2018. *Stokhasticheskie sistemy s korrelirovannymi potokami. Teoriya i primeneniye v telekommunikatsionnykh setyakh* [Stochastic systems with correlated streams. Theory and applications in telecommunication networks]. Moscow: Tekhnosfera. 564 p.
16. Abate, J., and W. Whitt. 2006. A unified framework for numerically inverting Laplace transforms. *INFORMS J. Comput.* 18(4):408–421. doi: 10.1287/ijoc.1050.0137.
17. Horvath, I., G. Horvath, S. A.-D. Almousa, and M. Telek. 2019. Numerical inverse Laplace transformation using concentrated matrix exponential distributions. *Performance Evaluation* 137:102067. 30 p. doi: 10.1016/j.peva.2019.102067.

Received August 15, 2021

Contributors

Milovanova Tatiana A. (b. 1977) — 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; tmilovanova77@mail.ru

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, 44-2 Vavilov Str., Moscow 119333, Russian Federation; izaryadov@sci.pfu.edu.ru

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

НЕКОТОРЫЕ ВЕРОЯТНОСТНО-СТАТИСТИЧЕСКИЕ СВОЙСТВА ГАММА-ЭКСПОНЕНЦИАЛЬНОГО РАСПРЕДЕЛЕНИЯ*

М. О. Воронцов¹, А. А. Кудрявцев², О. В. Шестаков³

Аннотация: В настоящее время большое внимание исследователей уделяется обобщениям известных математических объектов с целью получения адекватных моделей, описывающих реальные явления. Большую роль в прикладной теории вероятностей и математической статистике играет гамма-класс распределений, зарекомендовавший себя удобным и эффективным инструментом при моделировании многих реальных процессов. Гамма-класс довольно широк и включает распределения, обладающие такими полезными свойствами, как, например, безграничная делимость и устойчивость, что позволяет использовать распределения из этого класса в качестве асимптотических аппроксимаций в различных предельных теоремах. Одной из важнейших задач прикладной статистики является получение оценок параметров модельного распределения из имеющихся реальных данных. В работе рассматривается гамма-экспоненциальное распределение, представляющее собой обобщение распределений из гамма-класса. Приводятся оценки и асимптотические доверительные интервалы для некоторых параметров этого распределения. Обсуждается вопрос компьютерного моделирования реализаций выборки из гамма-экспоненциального распределения и численного оценивания параметров по выборке. Результаты работы могут найти широкое применение при изучении вероятностных моделей, основанных на непрерывных распределениях с неограниченным неотрицательным носителем.

Ключевые слова: компьютерное моделирование; оценивание параметров; гамма-экспоненциальное распределение; смешанные распределения; обобщенное гамма-распределение

DOI: 10.14357/08696527210302

* Работа выполнена при частичной финансовой поддержке РФФИ (проект 20-07-00655); исследования проводились в рамках программы Московского центра фундаментальной и прикладной математики.

¹Московский государственный университет имени М. В. Ломоносова, факультет вычислительной математики и кибернетики, m.vtsov@mail.ru

²Московский государственный университет имени М. В. Ломоносова, факультет вычислительной математики и кибернетики, nubigena@mail.ru

³Московский государственный университет имени М. В. Ломоносова, факультет вычислительной математики и кибернетики; Федеральный исследовательский центр «Информатика и управление» Российской академии наук, oshestakov@cs.msu.ru

1 Введение

Проблема оценивания неизвестных параметров играет традиционно важную роль в задачах прикладной математической статистики. При этом с целью повышения согласованности математических моделей и анализируемых реальных процессов исследователи рассматривают все более сложные математические абстракции. Применительно к моделям, описывающим реальные явления при помощи непрерывных распределений, имеющих неограниченные неотрицательные носители, традиционно используются частные случаи обобщенного гамма-распределения и обобщенного бета-распределения второго рода. В работе рассматривается задача оценивания параметров предложенного в [1] распределения, тесно связанного с перечисленными популярными распределениями.

Определение 1. Будем говорить, что случайная величина ζ имеет гамма-экспоненциальное распределение $\text{GE}(r, \nu, s, t, \delta)$ с параметрами изгиба $0 \leq r < 1$, формы $\nu \neq 0$, концентрации $s, t > 0$ и масштаба $\delta > 0$, если ее плотность при $x > 0$ задается соотношением

$$g_E(x) = \frac{|\nu| x^{t\nu-1}}{\delta^{t\nu} \Gamma(s) \Gamma(t)} \text{Ge}_{r, tr+s} \left(- \left(\frac{x}{\delta} \right)^\nu \right), \quad (1)$$

где $E = (r, \nu, s, t, \delta)$; $\text{Ge}_{\alpha, \beta}(x)$ — гамма-экспоненциальная функция [2]:

$$\text{Ge}_{\alpha, \beta}(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} \Gamma(\alpha k + \beta), \quad x \in \mathbb{R}, \quad 0 \leq \alpha < 1, \quad \beta > 0. \quad (2)$$

Функция (2) обобщает на случай $\beta \neq 1$ преобразование, введенное Э. Леруа [3] для исследования производящих функций специального вида. Кроме того, функцию (2) можно рассматривать (при некоторых допущениях) как частный случай функции Сривастава–Томовски [4], обобщающей функцию Миттаг-Леффлера [5].

В работе было показано, что распределение (1) адекватно описывает байесовские модели баланса [6]. Это прежде всего вызвано тем, что распределение с плотностью (1) может быть представлено [1] как масштабная смесь двух случайных величин, имеющих обобщенное гамма-распределение [1].

В свою очередь, обобщенное гамма-распределение $\text{GG}(v, q, \theta)$ с плотностью

$$f(x) = \frac{|v| x^{vq-1} e^{-(x/\theta)^v}}{\theta^{vq} \Gamma(q)}, \quad v \neq 0, \quad q > 0, \quad \theta > 0, \quad x > 0, \quad (3)$$

предложенное в 1925 г. итальянским экономистом Л. Аморозо для изучения экономической теории динамического равновесия [7], активно применяемое в гидрологии начиная с 1940-х гг. С. Н. Крицким и М. Ф. Менкелем [8, 9], но зачастую

ассоциирующееся с именем Е. Стэйси, рассмотревшего в 1962 г. частный случай распределения Аморозо [10], доказало свою состоятельность во многих прикладных задачах, используя для моделирования непрерывные распределения с неограниченным неотрицательным носителем. Еще во второй половине XIX в. распределение Рэля использовалось для описания результирующей амплитуды в задаче суммирования колебаний со случайными фазами, а распределение Максвелла–Больцмана — для статистического описания поведения параметров частиц идеального газа. Класс распределений (3) достаточно широк и включает кроме перечисленных экспоненциальное распределение; χ^2 -распределение; распределение Эрланга; гамма-распределение; полунормальное распределение, или распределение максимума процесса броуновского движения; χ -распределение; t -распределение Накагами; распределение Вильсона–Хильфerti; распределение Вейбулла–Гнеденко и многие другие, включая масштабированные и обратные аналоги перечисленных.

Использование обобщенного гамма-распределения (3) и его частных видов и смесей для описания различных моделей наряду с решением проблемы оценивания его параметров имеет богатую историю и до сих пор актуально. Например, в [11–14] предлагается использовать распределения из гамма-класса в задачах обработки радиолокационных сигналов и изображений, для исследования прочности материалов и надежности оборудования, для оценки концентрации вредных газов в промышленных зонах, а также изучения периода ремиссии онкологических больных. Можно привести ряд других примеров. Гамма-экспоненциальное распределение, рассматриваемое в статье, обобщает распределение Аморозо. Поэтому можно утверждать, что результаты статьи будут востребованы при изучении различных моделей, описывающих реальные процессы с помощью распределений с неотрицательным неограниченным носителем.

В работе [1] было показано, что гамма-экспоненциальное распределение обладает следующими свойствами.

Лемма 1. 1. Пусть независимые случайные величины λ и μ имеют соответственно распределения $GG(v, q, \theta)$ и $GG(u, p, \alpha)$, $uv > 0$. Тогда распределение λ совпадает с $GE(0, v, \cdot, q, \theta)$; распределение λ/μ при $|u| > |v|$ совпадает с $GE(v/u, v, p, q, \theta/\alpha)$; распределение λ/μ при $|v| > |u|$ совпадает с $GE(u/v, -u, q, p, \theta/\alpha)$.

2. При $0 < r < 1$ плотность $g_E(x)$, $E = (r, v, s, t, \delta)$, совпадает с плотностью отношения независимых случайных величин, имеющих обобщенные гамма-распределения $GG(v, t, \delta)$ и $GG(v/r, s, 1)$.

Возможность представления гамма-экспоненциального распределения как частного случайных величин, имеющих обобщенное гамма-распределение, позволяет применять его в широком круге прикладных задач. Так, в демографии уровень младенческой смертности определяется как отношение числа умерших в возрасте до года к числу родившихся за некоторый период времени [15], а индекс

разводимости — как отношение коэффициента суммарной разводимости к коэффициенту суммарной брачности [16]; в физике коэффициент трансформации — это отношение выходного напряжения к входному, а универсальная функция Кирхгофа — отношение излучательной способности тела к поглощательной [17]; в теории массового обслуживания отношение интенсивности входящего потока к интенсивности обслуживания определяет коэффициент загрузки системы [18]; при моделировании чрезвычайных ситуаций пожароопасность объекта определяется отношением угрозы возникновения пожара к фактору пожарозащиты [19]; в теории

надежности ожидаемое время безотказной работы представимо в виде отношения среднего времени безотказной работы к среднему времени восстановления [20]. Ряд примеров можно продолжить. Каждую из таких характеристик можно рассматривать как индекс баланса системы [6]. Применение рандомизационного байесовского подхода к описанным моделям дает возможность изучать характеристики индекса баланса как масштабной смеси вероятностных законов.

Помимо перечисленных свойств пятипараметрическое гамма-экспоненциальное распределение может применяться для моделирования широкого круга реальных явлений ввиду большого разнообразия его возможных плотностей (рис. 1).

На практике исследователь имеет дело с наблюдаемыми величинами, отражающими проявление анализируемого реального процесса, в отношении которых делаются некоторые модельные предположения о виде их распределения. В случае моделирования реального процесса при помощи гамма-экспоненциального распределения неизбежно возникает вопрос оценивания неизвестных параметров по реальным данным. Ввиду представления плотности $g_E(x)$ в терминах специальной гамма-экспоненциальной функции (2) метод максимального правдоподобия представляется затруднительным. То же можно сказать и о прямом методе моментов. По этой причине в работах [21, 22] было предложено оценивать параметры гамма-экспоненциального распределения при помощи модифицированного метода, основанного на логарифмических моментах.

В статье обсуждаются вычислительные аспекты построения реализаций выборок из гамма-экспоненциального распределения и нюансы численного оценивания неизвестных параметров.

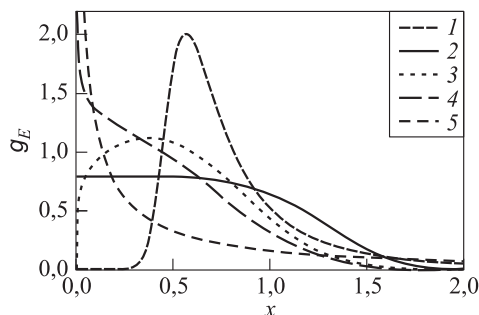


Рис. 1 Плотности гамма-экспоненциального распределения при различных значениях параметров: 1 — $E = (0,7; -6; 1,5; 0,4; 0,5)$; 2 — $E = (0,2; 5; 2; 0,2; 1,4)$; 3 — $E = (0,4; 3; 2; 0,4; 1)$; 4 — $E = (0,4; 2,7; 2; 0,33; 1)$; 5 — $E = (0,7; 0,4; 1,5; 2; 0,2)$

2 Оценки параметров гамма-экспоненциального распределения

Приведем известные оценки параметров гамма-экспоненциального распределения. Для этого введем в рассмотрение дигамма-функцию $\psi(z) = \Gamma'(z)/\Gamma(z)$ и функции

$$R(x) = \sqrt{\frac{\nu^2 x - \psi'(t)}{\psi'(s)}}; \quad (4)$$

$$D_r(x, y) = \exp \left\{ x - \frac{\psi(t)}{\nu} + \frac{\psi(s)}{\nu} \sqrt{\frac{\nu^2 y - \psi'(t)}{\psi'(s)}} \right\};$$

$$V(x) = \sqrt[3]{\frac{\psi''(t) - r^3 \psi''(s)}{x}}; \quad (5)$$

$$D_\nu(x, y) = \exp \left\{ x - (\psi(t) - r\psi(s)) \sqrt[3]{\frac{y}{\psi''(t) - r^3 \psi''(s)}} \right\}.$$

Введем обозначение для выборочных логарифмических моментов случайной величины ζ :

$$L_k(X) = \frac{1}{n} \sum_{i=1}^n \ln^k X_i,$$

где $X = (X_1, \dots, X_n)$ — выборка из распределения ζ .

В работе [21] были доказаны следующие утверждения.

Лемма 2. При фиксированных параметрах ν , t и s распределения GE(r, ν, s, t, δ) оценки

$$\hat{r}(X) = R(L_2(X) - L_1^2(X)); \quad (6)$$

$$\hat{\delta}_r(X) = D_r(L_1(X), L_2(X) - L_1^2(X)) \quad (7)$$

параметров r и δ обладают свойством сильной состоятельности.

Лемма 3. При фиксированных параметрах r , t и s распределения GE(r, ν, s, t, δ) оценки

$$\hat{\nu}(X) = V(L_3(X) - 3L_1(X)L_2(X) + 2L_1^3(X)); \quad (8)$$

$$\hat{\delta}_\nu(X) = D_\nu(L_1(X), L_3(X) - 3L_1(X)L_2(X) + 2L_1^3(X)) \quad (9)$$

параметров ν и δ обладают свойством сильной состоятельности.

Основываясь на приведенных оценках и их асимптотической нормальности [22], можно построить асимптотические доверительные интервалы для оценок (6)–(9) [22]. Для этого понадобятся следующие обозначения производных функций (4) и (5):

$$R_1(x) \equiv \frac{dR}{dx}(x) = \frac{\nu^2}{2\psi'(s)} \sqrt{\frac{\psi'(s)}{\nu^2 x - \psi'(t)}}; \quad (10)$$

$$V_1(x) \equiv \frac{dV}{dx}(x) = -\frac{\psi''(t) - r^3\psi''(s)}{3x^2} \sqrt[3]{\frac{x^2}{(\psi''(t) - r^3\psi''(s))^2}}. \quad (11)$$

В дальнейших рассуждениях понадобятся некоторые моментные характеристики гамма-экспоненциального распределения (1), получаемые путем дифференцирования характеристической функции логарифма случайной величины ζ , имеющей гамма-экспоненциальное распределение (подробнее см. [21, 22]). При фиксированных параметрах t и s получаем:

$$\begin{aligned} \sigma_1^2(r, \nu) &\equiv D \ln \zeta = \frac{\psi'(t) + r^2\psi'(s)}{\nu^2}; \\ \sigma_2^2(r, \nu, \delta) &\equiv D \ln^2 \zeta = \frac{4(\psi'(t) + r^2\psi'(s))[\nu \ln \delta + \psi(t) - r\psi(s)]^2}{\nu^4} + \\ &+ \frac{4(\psi''(t) - r^3\psi''(s))[\nu \ln \delta + \psi(t) - r\psi(s)]}{\nu^4} + \\ &+ \frac{2(\psi'(t) + r^2\psi'(s))^2}{\nu^4} + \frac{\psi'''(t) + r^4\psi'''(s)}{\nu^4}; \quad (12) \\ \sigma_3^2(r, \nu, \delta) &\equiv D \ln^3 \zeta = \frac{9(\psi'(t) + r^2\psi'(s))[\nu \ln \delta + \psi(t) - r\psi(s)]^4}{\nu^6} + \\ &+ \frac{18[\psi''(t) - r^3\psi''(s)][\nu \ln \delta + \psi(t) - r\psi(s)]^3}{\nu^6} + \\ &+ \frac{36[\psi'(t) + r^2\psi'(s)]^2[\nu \ln \delta + \psi(t) - r\psi(s)]^2}{\nu^6} + \\ &+ \frac{15[\psi'''(t) + r^4\psi'''(s)][\nu \ln \delta + \psi(t) - r\psi(s)]^2}{\nu^6} + \\ &+ \frac{34[\psi''(t) - r^3\psi''(s)][\psi'(t) + r^2\psi'(s)][\nu \ln \delta + \psi(t) - r\psi(s)]}{\nu^6} + \end{aligned}$$

$$\begin{aligned}
 & + \frac{6 \left[\psi^{(4)}(t) - r^5 \psi^{(4)}(s) \right] \left[\nu \ln \delta + \psi(t) - r\psi(s) \right]}{\nu^6} + \\
 & + \frac{15 \left[\psi'(t) + r^2 \psi'(s) \right]^3}{\nu^6} + \frac{9 \left[\psi''(t) - r^3 \psi''(s) \right]^2}{\nu^6} + \\
 & + \frac{15 \left[\psi'''(t) + r^4 \psi'''(s) \right] \left[\psi'(t) + r^2 \psi'(s) \right]}{\nu^6} + \frac{\left[\psi^{(5)}(t) + r^6 \psi^{(5)}(s) \right]}{\nu^6}. \quad (13)
 \end{aligned}$$

Введем дополнительные обозначения:

$$\begin{aligned}
 \sigma_{12}(r, \nu) &= \frac{\psi''(t) - r^3 \psi''(s)}{\nu^3}; \\
 s_{12}(r, \nu, \delta) &= \frac{\psi'''(t) + r^4 \psi'''(s)}{\nu^4}; \\
 s_1^2(r, \nu, \delta) &= \delta^2 \sigma_1^2(r, \nu) + \frac{\nu \delta^2 \psi(s) \sigma_{12}(r, \nu)}{r \psi'(s)} + \frac{\nu^2 \delta^2 \psi^2(s) \sigma_2^2(r, \nu, \delta)}{4 r^2 (\psi'(s))^2}; \quad (14) \\
 s_2^2(r, \nu, \delta) &= \delta^2 \sigma_1^2(r, \nu) - \frac{2 \nu^2 \delta^2 (\psi(t) - r \psi(s)) s_{12}(r, \nu, \delta)}{3 (\psi''(t) - r^3 \psi''(s))} + \\
 &+ \frac{\nu^4 \delta^2 (\psi(t) - r \psi(s))^2 \sigma_3^2(r, \nu, \delta)}{9 (\psi''(t) - r^3 \psi''(s))^2}. \quad (15)
 \end{aligned}$$

В работе [22] было показано, что доверительные интервалы для неизвестных параметров r , ν и δ имеют следующий вид. Через u_γ будем обозначать квантиль порядка $(1 + \gamma)/2$ стандартного нормального распределения.

Лемма 4. Асимптотический доверительный интервал с уровнем доверия γ , основанный на оценке (6) неизвестного параметра r , имеет вид:

$$(S_r(X), T_r(X)) = \left(\hat{r}(X) - \frac{u_\gamma}{\sqrt{n}} A_r(X), \hat{r}(X) + \frac{u_\gamma}{\sqrt{n}} A_r(X) \right),$$

где

$$A_r(X) = \sqrt{(R_1(L_2(X) - L_1^2(X)))^2 \sigma_2^2(\hat{r}(X), \nu, \hat{\delta}_r(X))},$$

функции R_1 и σ_2^2 задаются соотношениями (10) и (12), а статистики $\hat{r}(X)$ и $\hat{\delta}_r(X)$ определены в (6) и (7).

Лемма 5. Асимптотический доверительный интервал с уровнем доверия γ , основанный на оценке (7) неизвестного параметра δ , имеет вид:

$$(S_{\delta_r}(X), T_{\delta_r}(X)) = \left(\hat{\delta}_r(X) - \frac{u_\gamma}{\sqrt{n}} A_{\delta_r}(X), \hat{\delta}_r(X) + \frac{u_\gamma}{\sqrt{n}} A_{\delta_r}(X) \right),$$

где

$$A_{\delta_r}(X) = \sqrt{s_1^2(\hat{r}(X), \nu, \hat{\delta}_r(X))},$$

функция s_1^2 задается соотношением (14), а статистики $\hat{r}(X)$ и $\hat{\delta}_r(X)$ определены в (6) и (7).

Лемма 6. Асимптотический доверительный интервал с уровнем доверия γ , основанный на оценке (8) неизвестного параметра ν , имеет вид:

$$(S_\nu(X), T_\nu(X)) = \left(\hat{\nu}(X) - \frac{u_\gamma}{\sqrt{n}} A_\nu(X), \hat{\nu}(X) + \frac{u_\gamma}{\sqrt{n}} A_\nu(X) \right),$$

где

$$A_\nu(X) = \sqrt{(V_1(L_3(X) - 3L_1(X)L_2(X) + 2L_1^3(X)))^2 \sigma_3^2(r, \hat{\nu}(X), \hat{\delta}_\nu(X))},$$

функции V_1 и σ_3^2 задаются соотношениями (11) и (13), а статистики $\hat{\nu}(X)$ и $\hat{\delta}_\nu(X)$ определены в (8) и (9).

Лемма 7. Асимптотический доверительный интервал с уровнем доверия γ , основанный на оценке (9) неизвестного параметра δ , имеет вид:

$$(S_{\delta_\nu}(X), T_{\delta_\nu}(X)) = \left(\hat{\delta}_\nu(X) - \frac{u_\gamma}{\sqrt{n}} A_{\delta_\nu}(X), \hat{\delta}_\nu(X) + \frac{u_\gamma}{\sqrt{n}} A_{\delta_\nu}(X) \right),$$

где

$$A_{\delta_\nu}(X) = \sqrt{s_2^2(r, \hat{\nu}(X), \hat{\delta}_\nu(X))},$$

функция s_2^2 задается соотношением (15), а статистики $\hat{\nu}(X)$ и $\hat{\delta}_\nu(X)$ определены в (8) и (9).

3 Алгоритм построения реализаций выборок из гамма-экспоненциального распределения

Используя лемму 1 и хорошо известное равенство для случайной величины λ , имеющей обобщенное гамма-распределение $\text{GG}(\nu, t, \delta)$,

$$\lambda \stackrel{d}{=} \delta \xi^{1/\nu},$$

где случайная величина ξ имеет гамма-распределение $\text{GG}(1, t, 1)$, легко получить аналогичное представление для гамма-экспоненциального распределения.

Лемма 8. Пусть ξ и η — независимые случайные величины, имеющие соответственно гамма-распределения $GG(1, t, 1)$ и $GG(1, s, 1)$, а случайная величина ζ имеет гамма-экспоненциальное распределение $GE(r, \nu, s, t, \delta)$. Тогда

$$\zeta \stackrel{d}{=} \delta(\xi\eta^{-r})^{1/\nu}.$$

Основываясь на лемме 8, можно предложить следующий алгоритм построения выборки $Z_1, \dots, Z_n$ из гамма-экспоненциального распределения $GE(r, \nu, s, t, \delta)$. На первом этапе (например, с помощью метода выборки с отклонением [23]) строятся независимые выборки $X_1, \dots, X_n$ и $Y_1, \dots, Y_n$ соответственно из гамма-распределений $GG(1, t, 1)$ и $GG(1, s, 1)$. На втором этапе применяется лемма 8 и строятся компоненты

$$Z_k = \delta(X_k Y_k^{-r})^{1/\nu}.$$

4 Вычисление оценок параметров изгиба, формы и масштаба

При помощи описанного в предыдущем разделе алгоритма реализации выборок и лемм 2 и 3 можно построить гистограммы по выборке из модельного гамма-экспоненциального распределения с известными параметрами и оценить по реализованной выборке пары параметров изгиба–масштаба и формы–масштаба в предположении, что они неизвестны. Гистограммы на рис. 2–4 соответствуют выборке $X = (X_1, \dots, X_n)$ объема $n = 10\,000$ из распределения с параметрами $E = (r, \nu, s, t, \delta)$.

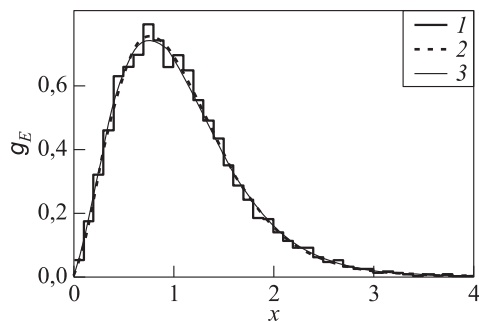


Рис. 2 Иллюстрация вычисления оценок параметров формы и масштаба: 1 — гистограмма; 2 — истинная плотность с $E = (0,4; 1,7; 1,8; 1,4; 1)$; 3 — оцененная плотность с $E = (0,4; 1,663; 1,8; 1,4; 0,998)$

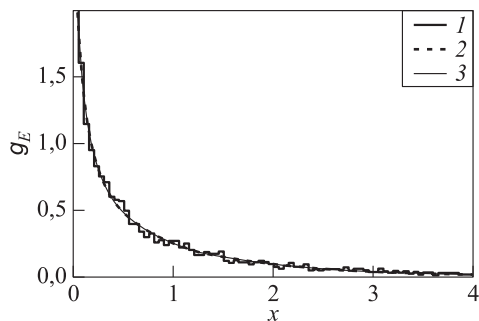


Рис. 3 Иллюстрация вычисления оценок параметров формы и масштаба: 1 — гистограмма; 2 — истинная плотность с $E = (0,3; 0,8; 3; 0,8; 1,5)$; 3 — оцененная плотность с $E = (0,3; 0,790; 3; 0,8; 1,486)$

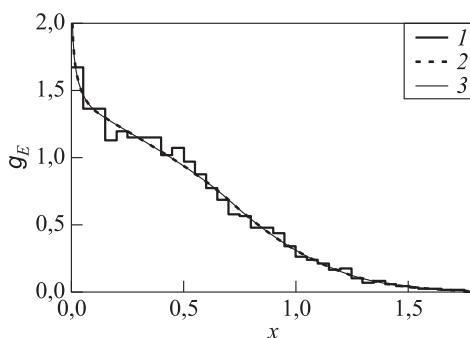


Рис. 4 Иллюстрация вычисления оценок параметров изгиба и масштаба: 1 — гистограмма; 2 — истинная плотность с $E = (0,4; 2,7, 2; 0,33; 1)$; 3 — оцененная плотность с $E = (0,458; 2,7; 2; 0,33; 1,012)$

В табл. 1–6 приведены значения оценок из лемм 2 и 3 и асимптотические доверительные интервалы из лемм 4–7; уровень доверия γ равен 95%, объем выборки равен n .

В табл. 7–9 приведено число выборок (из $N = 1000$ выборок) из модельного распределения с различным набором параметров, при оценке которых асимптотический доверительный интервал с уровнем доверия 95% покрыл истинное значение параметра.

Таблица 1 Оценки параметров r и δ с соответствующими доверительными интервалами, построенные по выборке из модельного распределения с набором параметров $E = (0,7; 0,5; 5,2; 3,7; 1)$

n	$\hat{r}(X)$	$S_r(X)$	$T_r(X)$	$\hat{\delta}_r(X)$	$S_{\delta_r}(X)$	$T_{\delta_r}(X)$
10^2	0,5832	0,1227	1,0437	0,6637	-0,2579	1,5854
10^3	0,6979	0,5673	0,8285	1,0239	0,6178	1,4300
10^4	0,7367	0,6964	0,7769	1,1130	0,9766	1,2495
10^5	0,6904	0,6772	0,7035	0,9704	0,9317	1,0091

Таблица 2 Оценки параметров ν и δ с соответствующими доверительными интервалами, построенные по выборке из модельного распределения с набором параметров $E = (0,7; 0,5; 5,2; 3,7; 1)$

n	$\hat{\nu}(X)$	$S_\nu(X)$	$T_\nu(X)$	$\hat{\delta}_\nu(X)$	$S_{\delta_\nu}(X)$	$T_{\delta_\nu}(X)$
10^2	0,4413	-0,0029	0,8857	0,9323	0,6030	1,2615
10^3	0,5164	0,3530	0,6798	1,0359	0,9382	1,1336
10^4	0,5024	0,4521	0,5528	0,9941	0,9634	1,0247
10^5	0,5029	0,4869	0,5188	1,0006	0,9909	1,0103

Таблица 3 Оценки параметров r и δ с соответствующими доверительными интервалами, построенные по выборке из модельного распределения с набором параметров $E = (0,3; 2,3; 1,2; 1,7; 3,5)$

n	$\hat{r}(X)$	$S_r(X)$	$T_r(X)$	$\hat{\delta}_r(X)$	$S_{\delta_r}(X)$	$T_{\delta_r}(X)$
10^2	0,1776	-2,2194	2,5747	3,7180	2,5285	4,9075
10^3	0,2080	-0,4260	0,8422	3,5452	3,2394	3,8509
10^4	0,2905	0,1406	0,4403	3,5084	3,4337	3,5830
10^5	0,3074	0,2621	0,3526	3,4981	3,4755	3,5208

Таблица 4 Оценки параметров ν и δ с соответствующими доверительными интервалами, построенные по выборке из модельного распределения с набором параметров $E = (0,3; 2,3; 1,2; 1,7; 3,5)$

n	$\hat{\nu}(X)$	$S_\nu(X)$	$T_\nu(X)$	$\hat{\delta}_\nu(X)$	$S_{\delta_\nu}(X)$	$T_{\delta_\nu}(X)$
10^2	2,6075	-8,3730	13,5882	3,4872	2,3907	4,5838
10^3	2,2301	-0,0267	4,4871	3,4345	3,1850	3,6840
10^4	2,2634	1,5055	3,0214	3,4936	3,4075	3,5797
10^5	2,2827	2,0376	2,5278	3,4942	3,4665	3,5219

Таблица 5 Оценки параметров r и δ с соответствующими доверительными интервалами, построенные по выборке из модельного распределения с набором параметров $E = (0,7; -4; 2,6; 0,4; 0,5)$

n	$\hat{r}(X)$	$S_r(X)$	$T_r(X)$	$\hat{\delta}_r(X)$	$S_{\delta_r}(X)$	$T_{\delta_r}(X)$
10^2	0,4900	-8,4680	9,4481	0,5635	-0,3439	1,4710
10^3	0,8747	-0,6446	2,3941	0,4889	0,3604	0,6175
10^4	0,8298	0,3268	1,3328	0,4895	0,4468	0,5322
10^5	0,7949	0,6291	0,9607	0,4923	0,4781	0,5065

Таблица 6 Оценки параметров ν и δ с соответствующими доверительными интервалами, построенные по выборке из модельного распределения с набором параметров $E = (0,7; -4; 2,6; 0,4; 0,5)$

n	$\hat{\nu}(X)$	$S_\nu(X)$	$T_\nu(X)$	$\hat{\delta}_\nu(X)$	$S_{\delta_\nu}(X)$	$T_{\delta_\nu}(X)$
10^2	-4,6608	-7,5225	-1,7991	0,6044	0,3666	0,8421
10^3	-4,0980	-4,8447	-3,3512	0,5147	0,4472	0,5822
10^4	-3,9794	-4,2072	-3,7516	0,4995	0,4784	0,5207
10^5	-3,9611	-4,0327	-3,8894	0,4974	0,4907	0,5041

Таблица 7 Число выборок (из $N = 1000$ выборок) из модельного распределения с набором параметров $E = (0,3; 2,3; 1,2; 1,7; 3,5)$

n	r	δ_r	ν	δ_ν
10^2	1000	991	1000	940
10^3	1000	999	1000	996
10^4	1000	999	1000	999
10^5	1000	1000	1000	1000

Таблица 8 Число выборок (из $N = 1000$ выборок) из модельного распределения с различным набором параметров $E = (0,7; 0,5; 5,2; 3,7; 1)$

n	r	δ_r	ν	δ_ν
10^2	957	920	871	839
10^3	957	944	971	945
10^4	935	940	995	976
10^5	949	946	997	965

Таблица 9 Число выборок (из $N = 1000$ выборок) из модельного распределения с набором параметров $E = (0,7; -4; 2,6; 0,4; 0,5)$

n	r	δ_r	ν	δ_ν
10^2	725	910	987	981
10^3	772	832	985	981
10^4	901	906	978	986
10^5	967	969	982	985

5 Связь гамма-экспоненциального распределения с обобщенным бета-распределением второго рода

Ранее (например, в работе [24]) отмечалась связь отношения двух независимых случайных величин, имеющих обобщенные гамма-распределения с одинаковыми параметрами формы, с обобщенным бета-распределением второго рода $GB2(\nu, \delta, t, s)$ [25] с плотностью

$$f_\beta(x) = \frac{|\nu| (x/\delta)^{t\nu-1} (1 + (x/\delta)^\nu)^{-t-s}}{\delta B(s, t)}, \quad x > 0, \quad (16)$$

частные случаи которого представляют собой распределение Бурра, или распределение Сингх–Маддала, распределение Дагума, распределение Ломакса, F-распределение Фишера–Снедекора и др. Представление гамма-экспоненциального распределения как масштабной смеси обобщенных гамма-распределений, таким образом, приводит к закономерному вопросу о его асимптотике при $r \rightarrow 1$ и связи с обобщенным бета-распределением второго рода.

По аналогии со свойством 19 из [26] можно показать справедливость следующего утверждения.

Лемма 9. Для $x \leq 0$, $t > 0$, $s > 0$ справедливо равенство:

$$\lim_{r \rightarrow 1} \text{Ge}_{r, tr+s}(x) = \frac{\Gamma(t+s)}{(1-x)^{t+s}}.$$

Обозначим функцию гамма-экспоненциального распределения с плотностью $g_E(x)$, определенной в (1), через $G_E(x)$, где $E = (r, \nu, s, t, \delta)$. Обозначим функцию обобщенного бета-распределения второго рода с плотностью (16) через $F_\beta(x)$.

Справедливо следующее утверждение.

Теорема 1. При $r \rightarrow 1$ гамма-экспоненциальное распределение слабо сходится к обобщенному бета-распределению второго рода:

$$\text{GE}(r, \nu, s, t, \delta) \implies \text{GB2}(\nu, \delta, t, s).$$

Доказательство. Покажем, что $g_E(x) \rightarrow f_\beta(x)$ для всех действительных x при $r \rightarrow 1$. Для $x \leq 0$ утверждение очевидно; при $x > 0$, используя лемму 9, имеем:

$$\lim_{r \rightarrow 1} g_E(x) = \frac{|\nu| x^{t\nu-1}}{\delta^{t\nu} \Gamma(s) \Gamma(t)} \lim_{r \rightarrow 1} \text{Ge}_{r, tr+s}(-(x/\delta)^\nu) = \frac{|\nu| (x/\delta)^{t\nu-1}}{\delta(1+(x/\delta)^\nu)^{t+s}} \frac{\Gamma(t+s)}{\Gamma(s) \Gamma(t)}.$$

Используя этот факт и теорему Шеффе [27], получаем (относительно меры Лебега l):

$$\lim_{r \rightarrow 1} \sup_{A \in \mathcal{B}(\mathbb{R})} \left| \int_A g_E dl - \int_A f dl \right| = 0.$$

Следовательно,

$$\lim_{r \rightarrow 1} \sup_x |G_E(x) - F_\beta(x)| = 0,$$

что доказывает утверждение теоремы.

6 Заключение

Традиционно частные случаи обобщенного гамма-распределения и обобщенного бета-распределения второго рода активно применяются при моделировании всевозможных реальных процессов. Тесная связь гамма-экспоненциального распределения с перечисленными популярными распределениями дает возможность утверждать, что приведенные в статье результаты могут найти широкое применение в исследованиях, использующих для моделирования распределения с неограниченными неотрицательными носителями.

Литература

1. Кудрявцев А. А. О представлении гамма-экспоненциального и обобщенного отрицательного биномиального распределений // Информатика и её применения, 2019. Т. 13. Вып. 4. С. 78–82.
2. Кудрявцев А. А., Титова А. И. Гамма-экспоненциальная функция в байесовских моделях массового обслуживания // Информатика и её применения, 2017. Т. 11. Вып. 4. С. 104–108.
3. Le Roy É. Sur les séries divergentes et les fonctions définies par un développement de Taylor // Ann. Fac. Sci. Toulouse 2 ser., 1900. Vol. 2. No. 3. P. 317–384.
4. Srivastava H. M., Tomovski Z. Fractional calculus with an integral operator containing a generalized Mittag-Leffler function in the kernel // Appl. Math. Comput., 2009. Vol. 211. P. 198–210.
5. Gorenflo R., Kilbas A. A., Mainardi F., Rogosin S. V. Mittag-Leffler functions, related topics and applications. — Berlin–Heidelberg: Springer-Verlag, 2014. 443 p.
6. Кудрявцев А. А. Байесовские модели баланса // Информатика и её применения, 2018. Т. 12. Вып. 3. С. 18–27.
7. Amoroso L. Ricerche intorno alla curva dei redditi // Ann. Mat. Pur. Appl., 1925. Vol. 21. P. 123–159.
8. Крицкий С. Н., Менкель М. Ф. О приемах исследования случайных колебаний речного стока // Труды НИУ ГУГМС. Сер. IV, 1946. Вып. 29. С. 3–32.
9. Крицкий С. Н., Менкель М. Ф. Выбор кривых распределения вероятностей для расчетов речного стока // Известия АН СССР. Отд. техн. наук, 1948. № 6. С. 15–21.
10. Stacy E. W. A generalization of the gamma distribution // Ann. Math. Stat., 1962. Vol. 33. P. 1187–1192.
11. Gao G., Ouyang K., Luo Y., Liang S., Zhou S. Scheme of parameter estimation for generalized gamma distribution and its application to ship detection in SAR images // IEEE T. Geosci. Remote, 2017. Vol. 55. No. 3. P. 1812–1832.
12. Zhou Y., Zhu H. Image segmentation using a trimmed likelihood estimator in the asymmetric mixture model based on generalized gamma and Gaussian distributions // Math. Probl. Eng., 2018. Vol. 2018. Art. 3468967. 17 p. doi: 10.1155/2018/3468967.
13. Iriarte Y. A., Varela H., Gómez H. J., Gómez H. W. A gamma-type distribution with applications // Symmetry, 2020. Vol. 12. Iss. 5. Art. 870. 15 p. doi: 10.3390/sym12050870.
14. Rivera P. A., Barranco-Chamorro I., Gallardo D. I., Gómez H. W. Scale mixture of Rayleigh distribution // Mathematics, 2020. Vol. 8. Iss. 10. Art. 1842. 22 p. doi: 10.3390/math8101842.
15. Борисов В. А. Демография. — М.: NOTABENE, 2001. 272 с.
16. Волгин Н. А., Рыбаковский Л. Л., Калмыкова Н. М., Архангельский В. Н., Иванова Е. И., Захарова О. Д., Иванова А. Е., Денисенко М. Б., Тихомиров Н. П., Тихомирова Т. М. Демография / Под ред. Н. А. Волгина, Л. Л. Рыбаковского. — М.: Логос, 2005. 280 с.
17. Кузнецов С. И., Rogozin K. I. Справочник по физике. — Томск: ТПУ, 2012. 224 с.

18. Бочаров П. П., Печинкин А. В. Теория массового обслуживания. — М.: РУДН, 1995. 529 с.
19. Шантала В. Г., Радоуцкий В. Ю., Шантала В. В. Основы моделирования чрезвычайных ситуаций. — Белгород: БГТУ, 2010. 166 с.
20. Здоровцов И. А., Королев В. Ю. Основы теории надежности волоконно-оптических линий передачи железнодорожного транспорта. — М.: МАКС Пресс, 2004. 308 с.
21. Кудрявцев А. А., Шестаков О. В. Метод логарифмических моментов для оценивания параметров гамма-экспоненциального распределения // Информатика и её применения, 2020. Т. 14. Вып. 3. С. 49–54.
22. Kudryavtsev A. A., Shestakov O. V. Asymptotically normal estimators for the parameters of the gamma-exponential distribution // Mathematics, 2021. Vol. 9. Iss. 3. Art. 273. 13 p. doi: 10.3390/math9030273.
23. Devroye L. Non-uniform random variate generation. — New York, NY, USA: Springer-Verlag. 1986. 843 p.
24. Кудрявцев А. А. Априорное обобщенное гамма-распределение в байесовских моделях баланса // Информатика и её применения, 2019. Т. 13. Вып. 3. С. 27–33.
25. McDonald J. B. Some generalized functions for the size distribution of income // Econometrica, 1984. Vol. 52. No. 3. P. 647–665.
26. Воронцов М. О., Кудрявцев А. А., Шоргин С. Я. Аналитические свойства и аспекты вычисления гамма-экспоненциальной функции // Системы и средства информатики, 2021. Т. 31. Вып. 2. С. 106–116.
27. Биллингсли П. Сходимость вероятностных мер / Пер. с англ. — М.: Наука, 1977. 352 с. (Billingsley P. Convergence of probability measures. — New York, NY, USA: John Wiley & Sons, Inc., 1977. 277 p.)

Поступила в редакцию 19.06.21

SOME PROBABILITY-STATISTICAL PROPERTIES OF THE GAMMA-EXPONENTIAL DISTRIBUTION

M. O. Vorontsov¹, A. A. Kudryavtsev¹, and O. V. Shestakov^{1,2}

¹Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskie Gory, GSP-1, Moscow 119991, Russian Federation

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

Abstract: Currently, much attention of researchers is paid to generalizations of well-known mathematical objects in order to obtain adequate models describing real phenomena. An important role in the applied theory of probability and mathematical statistics is played by the gamma class of distributions, which has proven to be a convenient and effective tool for modeling a lot of real processes.

The gamma class is quite wide and includes distributions that have such useful properties as, for example, infinite divisibility and stability, which makes it possible to use distributions from this class as asymptotic approximations in various limit theorems. One of the most important tasks of applied statistics is to obtain estimates of the parameters of the model distribution from the available real data. The paper considers the gamma-exponential distribution which is a generalization of the distributions from the gamma class. Estimates and asymptotic confidence intervals are given for some parameters of this distribution. The problems of computer modeling of sample realizations from the gamma-exponential distribution and the numerical estimation of parameters for the sample are discussed. The results of the work can be widely used in the study of probabilistic models based on continuous distributions with an unbounded nonnegative support.

Keywords: computer modeling; parameter estimation; gamma-exponential distribution; mixed distributions; generalized gamma distribution

DOI: 10.14357/08696527210302

Acknowledgments

The work was partly supported by the Russian Foundation for Basic Research (project 20-07-00655). The research was conducted in accordance with the program of Moscow Center for Fundamental and Applied Mathematics.

References

1. Kudryavtsev, A. A. 2019. O predstavlenii gamma-eksponentsial'nogo i obobshchenno-go otritsatel'nogo binomial'nogo raspredeleniy [On the representation of gamma-exponential and generalized negative binomial distributions]. *Informatika i ee Primeneniya — Inform. Appl.* 13(4):78–82.
2. Kudryavtsev, A. A., and A. I. Titova. 2017. Gamma-eksponentsial'naya funktsiya v bayesovskikh modelyakh massovogo obsluzhivaniya [Gamma-exponential function in Bayesian queuing models]. *Informatika i ee Primeneniya — Inform. Appl.* 11(4):104–108.
3. Le Roy, É. 1900. Sur les séries divergentes et les fonctions définies par un développement de Taylor. *Ann. Fac. Sci. Toulouse 2 ser.* 2(3):317–384.
4. Srivastava, H. M., and Ž. Tomovski. 2009. Fractional calculus with an integral operator containing a generalized Mittag-Leffler function in the kernel. *Appl. Math. Comput.* 211:198–210.
5. Gorenlo, R., A. A. Kilbas, F. Mainardi, and S. V. Rogosin. 2014. *Mittag-Leffler functions, related topics and applications*. Berlin–Heidelberg: Springer-Verlag. 443 p.
6. Kudryavtsev, A. A. 2018. Bayesovskie modeli balansa [Bayesian balance models]. *Informatika i ee Primeneniya — Inform. Appl.* 12(3):18–27.
7. Amoroso, L. 1925. Ricerche intorno alla curva dei redditi. *Ann. Mat. Pur. Appl.* 21:123–159.

8. Kritsky, S.N., and M.F. Menkel. 1946. O priemakh issledovaniya sluchaynykh kolebaniy rechnogo stoka [Methods of investigation of random fluctuations of river flow]. *Trudy NIU GUGMS Ser. IV* [GUGMS Research Institutions Proceedings, Ser. IV] 29:3–32.
9. Kritsky, S.N., and M. F. Menkel. 1948. Vybor krivyykh raspredeleniya veroyatnostey dlya raschetov rechnogo stoka [Selection of probability distribution curves for river flow calculations]. *Izvestiya AN SSSR. Otd. tekhn. nauk* [Herald of the Russian Academy of Sciences. Technical Sciences] 6:15–21.
10. Stacy, E. W. 1962. A generalization of the gamma distribution. *Ann. Math. Stat.* 33:1187–1192.
11. Gao, G., K. Ouyang, Y. Luo, S. Liang, and S. Zhou. 2017. Scheme of parameter estimation for generalized gamma distribution and its application to ship detection in SAR images. *IEEE T. Geosci. Remote* 55(3):1812–1832.
12. Zhou, Y., and H. Zhu. 2018. Image segmentation using a trimmed likelihood estimator in the asymmetric mixture model based on generalized gamma and Gaussian distributions. *Math. Probl. Eng.* 2018:3468967. 17 p. doi: 10.1155/2018/3468967.
13. Iriarte, Y. A., H. Varela, H. J. Gómez, and H. W. Gómez. 2020. A gamma-type distribution with applications. *Symmetry* 12(5):870. 15 p. doi: 10.3390/sym12050870.
14. Rivera, P. A., I. Barranco-Chamorro, D. I. Gallardo, and H. W. Gómez. 2020. Scale mixture of Rayleigh distribution. *Mathematics* 8(10):1842. 22 p. doi: 10.3390/math8101842.
15. Borisov, V. A. 2001. *Demografiya* [Demography]. Moscow: NOTABENE. 272 p.
16. Volgin, N. A., L. L. Rybakovskiy, N. M. Kalmykova, V. N. Arkhangel'skiy, E. I. Ivanova, O. D. Zakharova, A. E. Ivanova, M. B. Denisenko, N. P. Tikhomirov, and T. M. Tikhomirova. 2005. *Demografiya* [Demography]. Eds. N. A. Volgin and L. L. Rybakovskiy. Moscow: Logos. 280 p.
17. Kuznetsov, S. I., and K. I. Rogozin. 2012. *Spravochnik po fizike* [Handbook of physics]. Tomsk: TPU. 224 p.
18. Bocharov, P. P., and A. V. Pechinkin. 1995. *Teoriya massovogo obsluzhivaniya* [Queueing theory]. Moscow: RUDN. 529 p.
19. Shaptala, V. G., V. Yu. Radoutskiy, and V. V. Shaptala. 2010. *Osnovy modelirovaniya chrezvychaynykh situatsiy* [Basics of modeling of emergency situations]. Belgorod: BGTU. 166 p.
20. Zdorovtsov, I. A., and V. Yu. Korolev. 2004. *Osnovy teorii nadezhnosti volokonno-opticheskikh liniy peredachi zheleznodorozhnogo transporta* [Fundamentals of reliability theory of fiber optic transmission lines for railway transport]. Moscow: MAKS Press. 308 p.
21. Kudryavtsev, A. A., and O. V. Shestakov. 2020. Metod logarifmicheskikh momentov dlya otsenivaniya parametrov gamma-eksponentsial'nogo raspredeleniya [Method of logarithmic moments for estimating the gamma-exponential distribution parameters]. *Informatika i ee Primeneniya — Inform. Appl.* 14(3):49–54.
22. Kudryavtsev, A. A., and O. V. Shestakov. 2021. Asymptotically normal estimators for the parameters of the gamma-exponential distribution. *Mathematics* 9(3):273. 13 p. doi: 10.3390/math9030273.
23. Devroye, L. 1986. *Non-uniform random variate generation*. New York, NY: Springer-Verlag. 843 p.

24. Kudryavtsev, A. A. 2019. Apriornoe obobshchennoe gamma-raspredelenie v bayesovskikh modelyakh balansa [*A priori generalized gamma distribution in Bayesian balance models*]. *Informatika i ee Primeneniya — Inform. Appl.* 13(3):27–33.
25. McDonald, J. B. 1984. Some generalized functions for the size distribution of income. *Econometrica* 52(3):647–665.
26. Vorontsov, M. O., A. A. Kudryavtsev, and S. Ya. Shorgin. 2021. Analiticheskie svoystva i aspekty vychisleniya gamma-eksponentsial'noy funktsii [Analytical properties and aspects of computation of the gamma-exponential function]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 31(2):106–116.
27. Billingsley, P. 1977. *Convergence of probability measures*. New York, NY: John Wiley & Sons, Inc. 277 p.

Received June 19, 2021

Contributors

Vorontsov Mikhail O. (b. 1996)— student, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskie Gory, GSP-1, Moscow 119991, Russian Federation; m.vtsov@mail.ru

Kudryavtsev Alexey A. (b. 1978) — 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 Leninskie Gory, GSP-1, Moscow 119991, Russian Federation; nubigena@mail.ru

Shestakov Oleg V. (b. 1976) — Doctor of Science in physics and mathematics, professor, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskie 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; leading scientist, Moscow Center for Fundamental and Applied Mathematics, M. V. Lomonosov Moscow State University, 1 Leninskie Gory, GSP-1, Moscow 119991, Russian Federation; oshestakov@cs.msu.su

ИССЛЕДОВАНИЕ ПРОБЛЕМЫ УПРАВЛЕНИЯ ЗАПАСОМ НЕПРЕРЫВНОГО ПРОДУКТА В СТОХАСТИЧЕСКОЙ МОДЕЛИ РЕГЕНЕРАЦИИ ПРИ НАЛИЧИИ ДВУХ ПАРАМЕТРОВ ОПТИМИЗАЦИИ

П. В. Шнурков¹, К. А. Адамова²

Аннотация: Предлагается и исследуется новая стохастическая модель управления запасом непрерывного продукта в схеме регенерации. Состоянием системы является объем запаса в произвольный момент времени. Параметрами управления служат детерминированный максимально допустимый объем запаса и случайное время от момента очередного пополнения до следующего момента заказа. Получено явное аналитическое представление для стационарного стоимостного показателя эффективности управления — средней удельной прибыли. По своей структуре этот показатель представляет собой дробно-линейный интегральный функционал от распределения случайного параметра управления. При этом подынтегральные функции числителя и знаменателя указанного функционала зависят от второго детерминированного параметра управления. Для решения задачи оптимизации используется специальная форма теоремы об экстремуме дробно-линейного интегрального функционала. Проведено аналитическое исследование основной функции этого функционала на глобальный экстремум. На основании этого исследования установлены аналитические свойства исходных характеристик модели, при выполнении которых существует оптимальное решение задачи управления, а также получено описание этого решения.

Ключевые слова: управление запасом непрерывного продукта; управляемый регенерирующий процесс; дробно-линейный интегральный функционал

DOI: 10.14357/08696527210303

1 Введение. Краткая характеристика проблемы и описание предшествующих исследований

В научной литературе известно немало математических моделей, описывающих реальные сложные системы, предназначенные для временного хранения и поставки потребителю определенного вида продукта. В работах П. В. Шнуркова и его соавторов [1–4] были проанализированы различные стохастические

¹Национальный исследовательский университет «Высшая школа экономики», pshnurkov@hse.ru

²Научно-производственный центр автоматизации и приборостроения им. академика Н. А. Пилюгина, ksenya_an@mail.ru

модели управления запасом непрерывного продукта, основу которых составлял регенерирующий процесс. Общая стохастическая модель в форме регенерации, исследованная в этих работах, представляет собой развитие классической детерминированной модели с повторяющимися циклами, которая неоднократно излагалась в работах монографического характера, посвященных теории управления запасами [5–7]. Другое направление в современной теории систем управления запасами образуют исследования стохастических полумарковских моделей. Отметим некоторые работы из данного направления, опубликованные в последние годы [8, 9].

Стохастическая модель управления запасом непрерывного продукта, рассматриваемая в настоящей работе, принадлежит к классу регенерационных моделей. Таким образом, данная работа продолжает цикл исследований стохастических управляемых моделей, образуемый упомянутыми выше работами [1–4]. Основная отличительная особенность этой модели заключается в наличии двух параметров управления: детерминированного уровня пополнения запаса и случайного времени до момента заказа на пополнение товара.

По своему математическому содержанию проблема оптимального управления запасом, рассматриваемая в данной работе, представляет собой задачу нахождения безусловного экстремума стационарного стоимостного показателя эффективности управления (средней удельной прибыли). Как и в предшествующих исследованиях, этот показатель по своей аналитической форме представляет собой дробно-линейный интегральный функционал, заданный на множестве распределений случайного параметра управления. Однако, в отличие от упомянутых исследований, в рассматриваемой задаче подынтегральные функции числителя и знаменателя этого функционала зависят от второго, детерминированного параметра управления. В связи с этим для решения поставленной задачи нельзя непосредственно использовать известные результаты об экстремуме дробно-линейного интегрального функционала [10, 11].

Теоретическую основу решения рассматриваемой задачи образует специальный вариант утверждения об экстремуме дробно-линейного интегрального функционала, сформулированный и доказанный авторами настоящего исследования в работах [12, 13]. Заметим, что в новом варианте указанного утверждения проблема существования решения задачи оптимального управления, а также и непосредственного нахождения этого оптимального управления, определяется экстремальными свойствами так называемой основной функции дробно-линейного интегрального функционала, которая представляет собой отношение подынтегральных функций числителя и знаменателя, причем зависит от дополнительного детерминированного параметра оптимизации. В связи с этим в данной работе проводится аналитическое исследование основной функции на экстремум по двум параметрам оптимизации. В результате удастся установить аналитические условия на исходные характеристики модели, при выполнении которых оптимальное решение существует, а также охарактеризовать особенности этого решения.

2 Описание функционирования системы и математическая модель

Рассмотрим некоторую торговую систему (склад, магазин), предназначенную для временного хранения и поставки одного определенного вида продукта. Предположим, что множество значений объема этого продукта представляет собой некоторое подмножество множества вещественных чисел. Обозначим через $x(t)$ случайный процесс, значением которого является объем запаса данного продукта в момент времени t , $t \geq 0$. Потребление продукта из системы происходит с постоянной заданной скоростью $\alpha > 0$.

Обозначим через t_n , $n = 0, 1, 2, \dots$, $t_0 = 0$, случайные моменты времени, в которые происходит последовательное пополнение запаса продукта. Назовем интервалы времени $[t_n, t_{n+1}]$, $n = 0, 1, 2, \dots$, периодами функционирования данной системы. Приведем описание функционирования системы на одном периоде.

1. В момент начала очередного периода функционирования складской системы t_n объем продукта на складе принимает максимально возможное значение τ . Заметим, что в данной модели параметр τ принимает одно и то же значение для всех периодов функционирования, однако само это значение определяется в ходе решения задачи оптимизации.
2. Определяется время до следующего заказа на пополнение запаса как реализация некоторой неотрицательной случайной величины $\eta^{(n)}$. Случайные величины $\eta^{(n)}$, $n = 0, 1, 2, \dots$, являются независимыми и одинаково распределенными, имеющими распределение $G(y) = P(\eta^{(n)} < y)$, $n = 0, 1, 2, \dots$. Если случайная величина $\eta^{(n)}$ принимает фиксированное значение $\eta^{(n)} = u$, то следующий заказ на пополнение делается в момент времени $t'_n = t_n + u$. В этот момент начинается период задержки поставки, который продолжается детерминированное время z . Величина z предполагается известной. В течение периода задержки потребление продукта продолжается с постоянной скоростью $\alpha > 0$.

В момент времени $t_{n+1} = t_n + \eta^{(n)} + z = t_n + u + z$ происходит пополнение запаса, которое осуществляется мгновенно. Дальнейшая эволюция процесса $x(t)$ после момента t_{n+1} происходит независимо от прошлого и по описанным выше правилам.

При указанных условиях траектория случайного процесса $x(t)$ может быть задана следующим соотношением (см. рисунок):

$$x(t) = \tau - \alpha(t - t_n), \quad t_n \leq t \leq t_{n+1} = t_n + \eta^{(n)} + z, \quad n = 0, 1, 2, \dots$$

Стохастический объект $x(t)$ принадлежит к управляемым регенерирующим процессам. Управление этим процессом происходит в моменты регенерации t_n , $n = 0, 1, 2, \dots$. Общая теория регенерирующих случайных процессов изложена, например, в работе [14].

Управление процессом, или принимаемое решение, заключается в выборе значений двух независимых параметров $(\tau^{(n)}, \eta^{(n)})$, где $\tau^{(n)}$ — детерминированный параметр, который представляет собой объем продукта после очередного пополнения; $\eta^{(n)}$ — случайный параметр, который представляет собой время до следующего заказа на пополнение.

В рассматриваемой модели предполагается, что значение параметра $\tau^{(n)}$ не зависит от номера периода регенерации n : $\tau^{(n)} = \tau$, $n = 0, 1, 2, \dots$, а случайные параметры $\eta^{(n)}$ имеют одинаковое распределение $G(y) = P(\eta^{(n)} < y)$ при любом значении $n = 0, 1, 2, \dots$.

Задача оптимального управления случайным регенерирующим процессом $x(t)$ будет формально представлять собой задачу поиска экстремума некоторого стоимостного показателя $I(\tau, G(\cdot))$, заданного на множестве пар объектов $(\tau, G(\cdot))$, где $\tau \in S$ — множество возможных значений параметра τ ; $G(\cdot) \in \Gamma$ — множество функций распределения неотрицательных случайных величин.

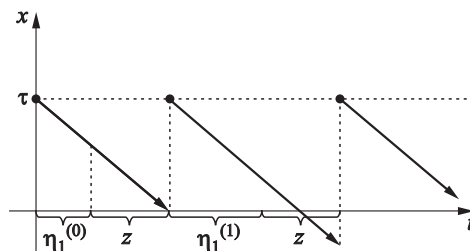
Определим исходные стоимостные характеристики рассматриваемой модели.

Рассмотрим сначала функции, описывающие различные виды затрат.

1. Функция $c_0(y)$ описывает затраты, связанные с пополнением объема запаса на складе на величину y . Функция определена при $y \geq 0$. Считаем, что $c_0(y) = 0$ при $y = 0$.
2. Функция $c_1(x)$ представляет собой затраты, связанные с хранением продукции объема x на складе в единицу времени, $c_1(x) = 0$ при $x \leq 0$, $c_1(x) \geq 0$ при $x > 0$.
3. Функция $c_2(x)$ представляет собой затраты или потери, связанные с дефицитом продукта объемом x в системе в единицу времени. Иначе говоря, данный вид затрат возникает, если реальный запас продукта исчерпан и образуется неудовлетворенный спрос объема x . При этом предполагается, что $c_2(x) = 0$ при $x \geq 0$, $c_2(x) > 0$ при $x < 0$.

Теперь введем две функции, характеризующие доход системы, связанный с оплатой за поставку товара потребителю.

1. Функция $d_1(x)$ представляет собой цену за поставку единицы товара в единицу времени при условии, что в системе имеется реальный запас объема $x > 0$. При этом $d_1(x) \geq 0$, $x > 0$; $d_1(x) = 0$, $x \leq 0$.
2. Функция $d_2(x)$ представляет собой цену за поставку единицы товара в единицу времени при условии, что в системе образовался дефицит объема x . При этом предполагается, что потребление может быть организовано с резервного склада. В этом случае $d_2(x) \geq 0$, $x < 0$; $d_2(x) = 0$, $x \geq 0$.



Траектория случайного процесса $x(t)$

Описанные характеристики отражают экономическую суть реальных процессов.

3 Постановка задачи оптимального управления

Введем аддитивный стоимостный функционал $V(t)$, связанный со случайным процессом $x(t)$. По своему прикладному содержанию значение этого функционала в момент времени t представляет собой случайную прибыль, накопленную за период времени $[0, t]$. Такие функционалы часто используются в моделях управления случайными процессами. Теоретические вопросы, связанные с построением таких функционалов и их важнейшими свойствами, изложены, в частности, в классической работе [15].

Введем следующие обозначения:

$V_n = V(t_n)$ — значение функционала в момент регенерации t_n , $n = 0, 1, 2, \dots$;

$\Delta V_n = V_{n+1} - V_n$ — приращение функционала за период регенерации $[t_n, t_{n+1}]$, $n = 0, 1, 2, \dots$;

$\Delta t_n = t_{n+1} - t_n$ — случайная длительность периода регенерации, $n = 0, 1, 2, \dots$

Заметим, что приращение данного функционала за произвольный период регенерации полностью определяется заданными стоимостными характеристиками модели. В ходе исследования были получены явные аналитические представления для математических ожиданий этих приращений через указанные характеристики.

Известно, что для стоимостных аддитивных функционалов справедливы так называемые эргодические теоремы [15]. В частности, для регенерирующего процесса при достаточно общих условиях, а именно при $E\Delta t_n > 0$, выполняется соотношение:

$$I = \lim_{t \rightarrow \infty} \frac{EV(t)}{t} = \frac{E\Delta V_n}{E\Delta t_n} = \frac{V}{T},$$

где $V = E\Delta V_n$ — математическое ожидание прибыли, полученной за период регенерации; $T = E\Delta t_n$ — математическое ожидание длительности периода регенерации.

По своему прикладному содержанию показатель I представляет собой среднюю удельную прибыль. Следуя [15], будем рассматривать I как показатель эффективности управления в исследуемой стохастической модели.

Поставим формально задачу оптимального управления для регенерирующего процесса $x(t)$:

$$I(\tau, G) \rightarrow \max, \quad \tau \in S, \quad G \in \Gamma,$$

где Γ — множество всех распределений неотрицательных случайных величин.

4 Основное утверждение о представлении стационарного стоимостного показателя

Одним из основных результатов данного исследования является утверждение, в котором определяется явное представление для показателя средней удельной прибыли в форме дробно-линейного интегрального функционала.

Теорема 1. В рассматриваемой модели регенерации стационарный функционал средней удельной прибыли имеет следующий вид:

$$I(\tau, G) = \frac{V}{T} = \frac{\int_0^{\infty} A(\tau, u) dG(u)}{\int_0^{\infty} B(\tau, u) dG(u)},$$

где $A(\tau, u)$ — условное математическое ожидание прибыли, полученной за период регенерации, при условии, что параметры управления принимают фиксированные значения τ, u ; $B(\tau, u)$ — условное математическое ожидание длительности периода регенерации, которое определяется при тех же условиях.

Указанные подынтегральные функции представляются следующими формулами.

1. В случае, когда $z > \tau/\alpha$, $u \geq 0$,

$$A(\tau, u) = \int_0^{\tau/\alpha} [\alpha d_1(\tau - \alpha t) - c_1(\tau - \alpha t)] dt + \int_{\tau/\alpha}^{u+z} [\alpha d_2(\tau - \alpha t) - c_2(\tau - \alpha t)] dt - c_0(\alpha(u + z)). \quad (1)$$

2. В случае, когда $0 < z < \tau/\alpha$,

$$A(\tau, u) = \begin{cases} \int_0^{u+z} [\alpha d_1(\tau - \alpha t) - c_1(\tau - \alpha t)] dt - c_0(\alpha(u + z)), & 0 \leq u < \frac{\tau}{\alpha} - z; \\ \int_0^{\tau/\alpha} [\alpha d_1(\tau - \alpha t) - c_1(\tau - \alpha t)] dt + \int_{\tau/\alpha}^{u+z} [\alpha d_2(\tau - \alpha t) - c_2(\tau - \alpha t)] dt - c_0(\alpha(u + z)), & u > \frac{\tau}{\alpha} - z. \end{cases} \quad (2)$$

Функция $B(\tau, u)$ имеет следующий вид:

$$B(\tau, u) = u + z, \quad u \geq 0. \quad (3)$$

Доказательство теоремы 1 приведено в приложении к данной работе [16].

5 Аналитическое исследование проблемы управления запасом

5.1 Аналитическое представление основной функции дробно-линейного интегрального функционала

Как следует из утверждения теоремы об экстремуме дробно-линейного интегрального функционала, зависящего от параметра [12, 13], решение задачи оптимального управления полностью определяется экстремальными свойствами основной функции

$$C(\tau, u) = \frac{A(\tau, u)}{B(\tau, u)},$$

где функции $A(\tau, u)$ и $B(\tau, u)$ задаются формулами (1)–(3).

Предположим, что $\tau \in [\tau_0, \tau_1]$, где $0 \leq \tau_0 < \tau_1 < \infty$, т. е. τ принадлежит замкнутому ограниченному интервалу.

Выпишем явное представление для основной функции при различных соотношениях между исходными параметрами.

Рассмотрим вариант, когда $\alpha z \in [\tau_0, \tau_1]$. Тогда из соотношений (1)–(3) следует:

(1) в случае, когда $\tau_0 \leq \tau < \alpha z$, $u \geq 0$,

$$C(\tau, u) = \frac{\int_0^{\tau/\alpha} [\alpha d_1(\tau - \alpha t) - c_1(\tau - \alpha t)] dt}{u + z} + \\ + \frac{\int_{\tau/\alpha}^{u+z} [\alpha d_2(\tau - \alpha t) - c_2(\tau - \alpha t)] dt - c_0(\alpha(u + z))}{u + z};$$

(2) в случае, когда $\alpha z < \tau \leq \tau_1$,

$$C(\tau, u) = \begin{cases} \frac{\int_0^{\tau/\alpha} [\alpha d_1(\tau - \alpha t) - c_1(\tau - \alpha t)] dt - c_0(\alpha(u + z))}{u + z}, & 0 \leq u < \frac{\tau}{\alpha} - z; \\ \frac{\int_0^{\tau/\alpha} [\alpha d_1(\tau - \alpha t) - c_1(\tau - \alpha t)] dt}{u + z} + \\ + \frac{\int_{\tau/\alpha}^{u+z} [\alpha d_2(\tau - \alpha t) - c_2(\tau - \alpha t)] dt - c_0(\alpha(u + z))}{u + z}, & u > \frac{\tau}{\alpha} - z. \end{cases}$$

Для других вариантов соотношений исходных параметров, а именно $\alpha z < \tau_0$ и $\alpha z > \tau_1$, соотношение для функции $C(\tau, u)$ выписывается аналогично с использованием формул (1)–(3).

Для доказательства существования и нахождения глобального экстремума можно использовать следующие направления аналитического исследования:

- (1) исследование функции $C(\tau, u)$ на непрерывность по совокупности переменных (τ, u) в рассматриваемой области значений аргументов;
- (2) исследование монотонности функции $C(\tau, u)$ методом анализа частных производных.

5.2 Аналитическое исследование основной функции дробно-линейного интегрального функционала

Имеют место следующие утверждения об аналитических свойствах основной функции.

Теорема 2. *Предположим, что функции $c_0(x)$, $d_1(x)$ и $c_1(x)$ непрерывны при $x \geq 0$, а функции $d_2(x)$ и $c_2(x)$ непрерывны при $x \leq 0$. Тогда основная функция $C(\tau, u)$ непрерывна при любых значениях аргументов: $\tau_0 \leq \tau \leq \tau_1$, $u \geq 0$.*

Введем дополнительные обозначения:

$$r_1(x) = \alpha d_1(x) - c_1(x); \quad r_2(x) = \alpha d_2(x) - c_2(x).$$

Теорема 3. *Предположим, что функция $r_1(x)$ неубывающая по x , $x \geq 0$, а $r_2(x)$ неубывающая по x , $x \leq 0$. Тогда основная функция $C(\tau, u)$ монотонно возрастает по τ , $\tau_0 \leq \tau \leq \tau_1$, для любого фиксированного u , $u \geq 0$.*

Доказательства теорем 2 и 3 приведены в приложении к данной работе [16].

На основании теоремы 2 можно установить существование глобального экстремума основной функции $C(\tau, u)$ в заданной области значений аргументов. Из теоремы 3 следует, что оптимальным значением параметра τ является его максимальное значение $\tau = \tau_1$.

6 Заключение

Утверждение о безусловном экстремуме дробно-линейного интегрального функционала, зависящего от параметра [12, 13], может служить теоретической основой для исследования многих содержательных моделей теории управления запасами и других прикладных стохастических моделей. Оно создает теоретическую возможность для решения соответствующих задач, в которых параметры оптимизации многомерные. При этом часть параметров оптимизации могут быть детерминированными, а остальные могут иметь случайный характер. Аналитические результаты, связанные с оптимальным решением и подобные тем, которые получены в настоящей работе, будут появляться только в специальных вариантах моделей. Однако открываются практически неограниченные возможности численного исследования соответствующих задач оптимального управления, связанные с использованием средств современной вычислительной техники.

Литература

1. Шнурков П. В., Мельников Р. В. Оптимальное управление запасом непрерывного продукта в модели регенерации // Обозрение прикладной и промышленной математики, 2006. Т. 13. № 3. С. 434–452.
2. Шнурков П. В., Мельников Р. В. Исследование проблемы управления запасом непрерывного продукта при детерминированной задержке поставки // Автоматика и телемеханика, 2008. № 10. С. 93–113.
3. Мельников Р. В. Исследование проблем управления запасом непрерывного продукта в стохастической модели регенерации: Дис. . . . канд. физ.-мат. наук. — М.: МГИЭМ, 2010. 133 с.
4. Шнурков П. В., Пименова Е. Ю. Оптимальное управление запасом непрерывного продукта в схеме регенерации с детерминированной задержкой поставки и периодом реального пополнения // Системы и средства информатики, 2017. Т. 27. № 4. С. 80–94.
5. Рыжиков Ю. И. Управление запасами. — М.: Наука, 1969. 344 с.
6. Лотоцкий В. А., Мандель А. С. Модели и методы управления запасами. — М.: Наука, 1991. 188 с.
7. Porteus E. L. Foundations of stochastic inventory theory. — Stanford, CA, USA: Stanford Business Book, 2002. 299 p.
8. Nino-Mora J. Restless bandit marginal productivity indices, diminishing returns, and optimal control of make-to-order/make-to-stock $M/G/1$ queues // Math. Oper. Res., 2006. Vol. 31. No. 1. P. 50–84.

9. Шнурков П. В., Иванов А. В. Анализ дискретной полумарковской модели управления запасом непрерывного продукта при периодическом прекращении потребления // Дискретная математика, 2014. Т. 26. Вып. 1. С. 143–154.
10. Шнурков П. В. О решении задачи безусловного экстремума для дробно-линейного интегрального функционала на множестве вероятностных мер // Докл. Акад. наук, 2016. Т. 470. № 4. С. 387–392.
11. Шнурков П. В., Горшенин А. К., Белоусов В. В. Аналитическое решение задачи оптимального управления полумарковским процессом с конечным множеством состояний // Информатика и её применения, 2016. Т. 10. Вып. 4. С. 72–88.
12. Shnurkov P., Adamova K. Solution of the unconditional extremal problem for a linear-fractional integral functional depending on the parameter // arXiv.org, 2019. arXiv:1906.05824 [math.OC]. 14 p.
13. Шнурков П. В., Адамова К. А. Решение задачи безусловного экстремума для дробно-линейного интегрального функционала, зависящего от параметра // Информатика и её применения, 2020. Т. 14. Вып. 2. С. 98–103.
14. Рыков В. В., Козырев Д. В. Основы теории массового обслуживания. — М.: Инфра-М, 2016. 223 с.
15. Майн Х., Осаки С. Марковские процессы принятия решений / Пер. с англ. — М.: Наука, 1977, 176 с. (Mine H., Osaki S. Markovian decision processes. — New York, NY, USA: Elsevier, 1970. 142 p.)
16. Шнурков П. В., Адамова К. А. Приложение к статье: Проблема оптимизации управления запасом непрерывного продукта в стохастической модели регенерации при наличии двух параметров оптимизации. 2021. 29 с. <http://www.ipiran.ru/publications/Shnurkov.docx>.

Поступила в редакцию 12.03.21

INVESTIGATION OF THE PROBLEM OF CONTINUOUS PRODUCT STOCK CONTROL IN A STOCHASTIC MODEL OF REGENERATION WITH TWO OPTIMIZATION PARAMETERS

P. V. Shnurkov¹ and K. A. Adamova²

¹National Research University Higher School of Economics, 34 Tallinskaya Str., Moscow 123458, Russian Federation

²Academician Pilugin Scientific-Production Center of Automatics and Instrument-Making, 1 Vvedenskogo Str., Moscow 117342, Russian Federation

Abstract: The paper proposes and analyzes a new stochastic model of regeneration of continuous product stock control. The volume of stock at an arbitrary moment of time is the status of the system. The control parameters are a maximum deterministic allowable stock volume and a random time from replenishment until the next order. For a stationary value indicator of the efficiency of

control — medium specific profit, an explicit analytical representation was obtained. By its structure, this indicator is a fractional linear integral functional from the distribution of a random control parameter. The integrands in the numerator and the denominator of the specified functional depend on the second deterministic control parameter. A special form of the theorem on the extremum of a fractional linear integral functional is used to solve the optimization problem. An analytical study of the main function of this functional on the global extremum was conducted. Based on this study, analytical properties of the original characteristics of the model, under which there is an optimal solution of the management task, are established and a description of this solution is obtained.

Keywords: continuous product stock control; controlled regenerating process; linear-fractional integral functional

DOI: 10.14357/08696527210303

References

1. Shnurkov, P. V., and R. V. Mel'nikov. 2006. Optimal'noe upravlenie zapasom nepre-ryvnogo produkta v modeli regeneratsii [Optimal control of a continuous product inventory in the regeneration model]. *Obozrenie prikladnoy i promyshlennoy matematiki* [Review of Applied and Industrial Mathematics] 13(3):434–452.
2. Shnurkov, P. V., and R. V. Mel'nikov. 2008. Analysis of the problem of continuous-product inventory control under deterministic lead time. *Automat. Rem. Contr.* 69(10):1734–1751.
3. Mel'nikov, R. V. 2010. Issledovanie problem upravleniya zapasom v stokhasticheskoy modeli regeneratsii [Research inventory control problems in stochastic recovery model]. Moscow: MGIEM. PhD Diss. 133 p.
4. Shnurkov, P. V., and E. Yu. Pimenova. 2017. Optimal'noe upravlenie zapasom nepre-ryvnogo produkta v skheme regeneratsii s determinirovannoy zaderzhkoy postavki i periodom real'nogo popolneniya [Optimal inventory control of continuous product in regeneration theory with determinate delay of the delivery and the period of real replenishment]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 27(4):80–94.
5. Ryzhikov, Yu. I. 1963. *Upravlenie zapasami* [Inventory management]. Moscow: Nauka. 235 p.
6. Lototskiy, V. A., and A. S. Mandel'. 1991. *Modeli i metody upravleniya zapasami* [Models and inventory control methods]. Moscow: Nauka. 188 p.
7. Porteus, E. L. 2002. *Foundations of stochastic inventory theory*. Stanford, CA: Stanford Business Book. 299 p.
8. Nino-Mora, J. 2006. Restless bandit marginal productivity indices, diminishing returns, and optimal control of make-to-order/make-to-stock $M/G/1$ queues. *Math. Oper. Res.* 31(1):50–84.
9. Snurkov, P. V., and A. V. Ivanov. 2015. Analysis of a discrete semi-Markov model of continuous inventory control with periodic interruptions of consumption. *Discrete Math.* 25(1):59–67.

10. Shnurkov, P. V. 2016. Solution of the unconditional extremum problem for a linear-fractional integral functional on a set of probability measures. *Dokl. Math.* 94(2):550–554.
11. Shnurkov, P. V., A. K. Gorshenin, and V. V. Belousov. 2016. Analiticheskoe reshenie zadachi optimal'nogo upravleniya polumarkovskim protsessom s konechnym mnozhestvom sostoyaniy [An analytic solution of the optimal control problem for a semi-Markov process with a finite set of states]. *Informatika i ee Primeneniya — Inform. Appl.* 10(4):72–88.
12. Shnurkov, P., and K. Adamova. 2019. Solution of the unconditional extremal problem for a linear-fractional integral functional depending on the parameter. *arXiv.org*. 14 p. Available at: <https://arxiv.org/abs/1906.05824> (accessed September 7, 2021).
13. Shnurkov, P. V., and K. A. Adamova. 2020. Reshenie zadachi bezuslovnogo ekstremuma dlya drobno-lineynogo integral'nogo funktsionala, zavisyashchego ot parametra [Solution of the unconditional extremal problem for a linear-fractional integral functional depending on the parameter]. *Informatika i ee Primeneniya — Inform. Appl.* 14(2):98–103.
14. Rykov, V. V., and D. V. Kozyrev. 2016. *Osnovy teorii massovogo obsluzhivaniya* [Queueing theory fundamentals]. Moscow: Infra-M. 223 p.
15. Mine, H., and S. Osaki. 1970. *Markovian decision processes*. New York, NY: Elsevier. 142 p.
16. Shnurkov, P. V., and K. A. Adamova. 2021. Prilozhenie k stat'e: Problema optimizatsii upravleniya zapasom nepreryvnogo produkta v stokhasticheskoy modeli regeneratsii pri nalichii dvukh parametrov optimizatsii [Appendix to article: The problem of optimization of continuous product stock control in a stochastic model of regeneration with two optimization parameters]. 29 p. Available at: <http://www.ipiran.ru/publications/Shnurkov.docx> (accessed September 7, 2021).

Received March 12, 2021

Contributors

Shnurkov Peter V. (b. 1953) — Candidate of Science (PhD) in physics and mathematics, associate professor, National Research University Higher School of Economics, 34 Tallinskaya Str., Moscow 123458, Russian Federation; pshnurkov@hse.ru

Adamova Kseniia A. (b. 1994) — engineer, Academician Pilugin Scientific-Production Center of Automatics and Instrument-Making, 1 Vvedenskogo Str., Moscow 117342, Russian Federation; kseniya_an@mail.ru

ИНТЕРОПЕРАБЕЛЬНОСТЬ КАК КЛЮЧЕВОЕ УСЛОВИЕ РЕАЛИЗАЦИИ ЦИФРОВОЙ ТРАНСФОРМАЦИИ

И. Н. Розенберг¹, С. К. Дулин², Н. Г. Дулина³

Аннотация: Анализируются процессы, характерные для оцифровывания информации, и стандартизация оцифровывания в информационной среде. Рассмотрены три этапа оцифровывания информации и их особенности. Отмечено, что важнейшим условием для проведения цифровой трансформации является обеспечение уровня семантической интероперабельности. Описаны характерные особенности этого уровня.

Ключевые слова: оцифровывание; цифровизация; цифровая трансформация; интероперабельность; семантическая геоинтероперабельность

DOI: 10.14357/08696527210304

1 Введение

Оцифровывание, цифровизация и цифровая трансформация — это различные процессы, которые по-разному используются в различных отраслях промышленности. Тем не менее в сфере информационных технологий (ИТ) они, как правило, означают следующее:

- оцифровывание означает преобразование некоторого информационного ресурса из аналогового представления в цифровое. Например, сканирование рукописей и хранение их в электронной библиотеке;
- цифровизация означает внедрение технологического процесса в области технологий и соответствующего программного обеспечения. Сервис онлайн-покупок в интернете служит примером цифрового бизнес-процесса, объединяющего оформленные электронно покупку и доставку;
- цифровая трансформация означает, что созданы процессы, основанные на программном обеспечении, которые могут развиваться, чтобы повысить гибкость компании и ее конкурентоспособность. Иногда в бизнесе это называется «Конкурировать онлайн или умереть». Можно привести следующее определение: «Цифровая трансформация — это процесс интеграции цифровых

¹ Научно-исследовательский и проектно-конструкторский институт информатизации, автоматизации и связи на железнодорожном транспорте (АО НИИАС), I.Rozenberg@vniias.ru

² Федеральный исследовательский центр «Информатика и управление» Российской академии наук; Научно-исследовательский и проектно-конструкторский институт информатизации, автоматизации и связи на железнодорожном транспорте (АО НИИАС), skdulin@mail.ru

³ Федеральный исследовательский центр «Информатика и управление» Российской академии наук, ngdulina@mail.ru

технологий во все аспекты бизнес-деятельности, требующий внесения коренных изменений в технологии, культуру, операции и принципы создания новых продуктов и услуг» [1].

2 Стандарты цифровизации

На сегодня ведущей инициативой для каждого ИТ-менеджера стала цифровизация. Цифровизация создает беспрецедентный спрос на специалистов, которые знают, как реализовать бизнес-процессы в программном обеспечении (разработчики программного обеспечения и архитекторы) и специалистов, которые знают, как управлять сложными системами приложений в производстве с высокой надежностью и производительностью.

Наиболее важным архитектурным изменением становится переход к архитектуре, которая позволяет самостоятельно работать с каждым микросервисом, что приводит к значительному сокращению сроков разработки приложений, значительному повышению технической и деловой гибкости и резкому повышению онлайн-конкурентоспособности.

Когда дело доходит до цифровой трансформации, то понимание процессов в проблемной области требует значительно большего времени, и это значительно сложнее, чем просто адаптировать их к новым цифровым технологиям. В цифровой трансформации ищутся новые решения, поскольку появляются проблемы, которые решаются с помощью новых технологий. Например, документ не просто оцифровывается, но обрабатывается и анализируется, изучается вопрос о том, нужен ли конкретный процесс или же его можно упростить с помощью новой технологии. Поэтому цифровая трансформация — это не что иное, как решение проблем с помощью наилучших технических средств. В определенной степени это относится и к таким темам, как «гибкое дизайнерское мышление», «мозговой штурм» и другим креативным способам работы, поскольку они необходимы для того, чтобы расширить взгляд на проблемы. Можно решить их техническими средствами с использованием новых технологий, только если есть понимание этих проблем.

Цифровизация привела к тому, что появился цифровой контент. Этот контент может быть обработан цифровыми процессами, а новые технологии приведут к разработке цифровой бизнес-стратегии. Но цифровая трансформация дает значительно больше. Конечно, цифровой контент, цифровые процессы и цифровая бизнес-модель необходимы, но при цифровой трансформации рассматриваются все производственные аспекты. Цифровая трансформация — это работа, которая должна быть тщательно спланирована и занимает много времени (3–5 лет). В отличие от цифровизации цифровая трансформация не просто требует от предприятий применения большого числа передовых технологий в своей деятельности и не может быть завершена в одном проекте. Таким образом, можно сделать вывод, что дорожная карта цифровой трансформации будет включать в себя множество проектов в России и за рубежом (рис. 1) [2].

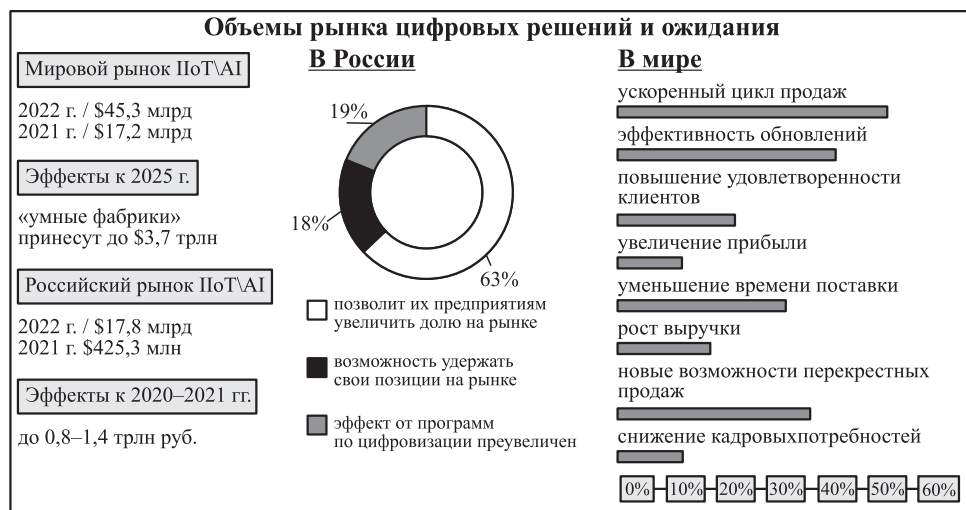


Рис. 1 Оценка эффекта цифровизации (по данным опроса KPMG, MarketsandMarkets, PWC): IoT\AI — Industrial Internet of Things\Artificial Intellect, промышленный интернет вещей/искусственный интеллект

Реализацию развития цифровой трансформации можно представить в виде трех этапов.

- I. Внедрение системы управления информационными технологиями и сервисами.
- II. Разработка и внедрение сервиса по анализу использования, оценки и повышения эффективности цифровых технологий.
- III. Переход от цифрового сервиса к сервису на основе искусственного интеллекта, машинного обучения и больших данных для управления цифровыми технологиями.

Пятого августа 2020 г. были утверждены первые стандарты цифровой промышленности в России. Они разработаны техническим комитетом «Кибер-физические системы» на базе Российской венчурной компании при поддержке Минпромторга РФ. Как отмечается в сообщении ведомства, стандарты направлены на эффективное внедрение цифровых технологий в российскую промышленность, развитие качественных и независимых решений, а также на обеспечение их совместимости. Стандартизация, от которой напрямую зависит успех цифровой трансформации, способна приносить в валовой внутренний продукт страны до 1% в год и около 3% в развитие экспорта.

Перечень утвержденных стандартов:

- предварительный национальный стандарт (ПНСТ) «Умное производство. Двойники цифровые производства. Часть 1. Общие положения»;

- ПНСТ «Умное производство. Двойники цифровые производства. Часть 2. Типовая архитектура»;
- ПНСТ «Умное производство. Двойники цифровые производства. Часть 3. Цифровое представление физических производственных элементов»;
- ПНСТ «Умное производство. Двойники цифровые производства. Часть 4. Обмен информацией»;
- ПНСТ «Умное производство. Двойники цифровые производства. Элементы визуализации цифровых двойников производства»;
- ПНСТ «Умное производство. Унифицированная архитектура ОРС. Часть 1. Общие положения»;
- ПНСТ «Умное производство. Интерфейсы для ухода за автоматизированной машиной. Часть 1. Общие положения»;
- ПНСТ «Умное производство. Интероперабельность единиц возможностей для промышленных прикладных решений. Часть 1. Критерии интероперабельности единиц возможностей согласно требованиям к применению»;
- ПНСТ «Умное производство. Интероперабельность единиц возможностей для промышленных прикладных решений. Часть 2. Шаблоны возможностей и каталогизация программных блоков»;
- ПНСТ «Умное производство. Интероперабельность единиц возможностей для промышленных прикладных решений. Часть 3. Верификация и валидация интероперабельности единиц возможностей».

3 Семантическая интероперабельность и понимание

Следует обратить внимание на последние три стандарта, посвященных термину «интероперабельность», служащему не только базовым понятием платформы Индустрия 4.0, но и тесно связанным с обсуждаемыми выше понятиями «цифровизация» и «цифровая трансформация». Распоряжением Правительства от 28 июля 2017 г. № 1632-р подтверждена важность решения проблемы интероперабельности для госуправления и для обеспечения обороноспособности страны. Согласно ГОСТу (ISO/IEC/IEEE 24765:2010. ГОСТР 55062-2012), «интероперабельность — способность двух или более информационных систем или компонентов к обмену информацией и к использованию информации, полученной в результате обмена».

Исследования интероперабельности привели к необходимости создания моделей интероперабельности, которые обеспечивают установление интероперабельности между системами в соответствии с различными целями и контекстами [3].

Модели интероперабельности позволяют: (1) сформировать общий тезаурус, который обеспечивает согласованное представление семантики; (2) сформулировать предложения по выработке решений; (3) оценить возможные модификации

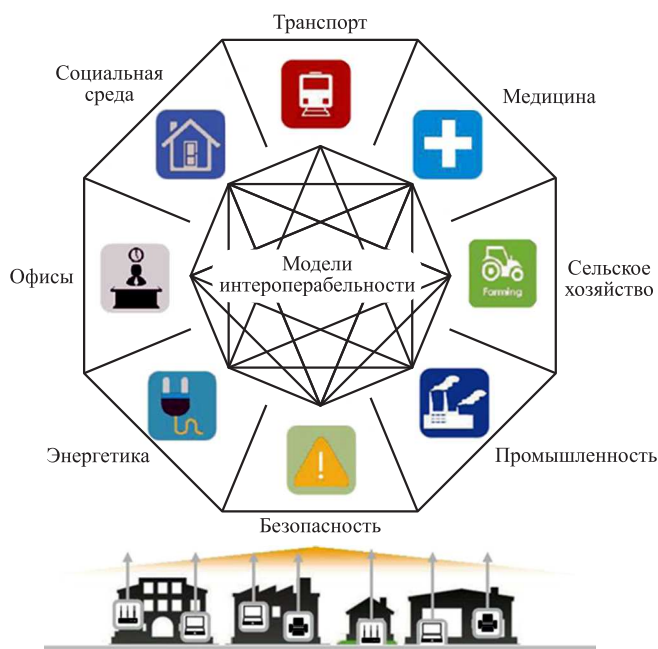


Рис. 2 Использование моделей интероперабельности

существующих и добавление новых опций. Отдельной проблемой оказывается согласование представлений моделей, необходимое для разработки специального метода интеграции моделей. Существующие модели интероперабельности используются на 6 различных уровнях организационной интероперабельности: техническом, синтаксическом, семантическом, прагматическом, динамическом и концептуальном — и применяются практически во всех сферах деятельности человека (рис. 2) [4].

Самым актуальным и важным из уровней интероперабельности на сегодняшний день стал уровень семантической интероперабельности, поскольку лишь при достижении этого уровня можно говорить о проведении цифровой трансформации. Семантическая интероперабельность стала ключевым фактором, позволяющим взаимодействующим системам понимать данные и обрабатывать их на основе этого понимания. Взаимодействующие системы представляют собой компоненты и инструменты в жизненном цикле изготовления продукции и оказания услуг. Стандарты служат одним из средств достижения семантической интероперабельности, поскольку обеспечивают четко определенные и согласованные требования к данным, хотя они зачастую настолько узко определены, что не могут обеспечить полную информационную модель, необходимую для интероперабельности. Тем не менее модели, связанные со стандартами, могут

быть легко интегрированы, обеспечивая достаточно широкую модель предметной области. Стандарты обеспечивают необходимую основу для любой интероперабельности. Для создания интеллектуальных инфраструктур необходим более целостный обмен информацией между различными предметными областями, что требует универсальной интероперабельности, основанной на общих семантических и онтологических основах с безусловным наличием понимания.

Наряду с проблемой понимания вообще существует проблема понимания семантики геоданных [5].

Среди определений понятия «понимание» для задач обеспечения интероперабельности наиболее подходит следующее: «Понимание (усвоение) — соотнесение, сравнение вновь поступающей информации (текста, образа, речи, поведения, явления и т. п.) с уже известной, накопленной и структурированной информацией (т. е. суть знанием) посредством когнитивных структур, схем поведения, понятий и категорий понятий; и ее оценка с определенных позиций, на основе некоторого образца, стандарта, нормы, принципа и т. п., например накопленного опыта или обстоятельств, в которых происходит процесс восприятия и понимания, особенностей когнитивной системы / когнитивного стиля человека» [5].

Непонимание чаще всего обусловлено необоснованностью, неочевидностью или отсутствием такого стандарта, а также может быть вызвано возникающим когнитивным диссонансом (в терминологии информационных систем — нарушением целостности существующего знания), а также, в случае осуществления процесса понимания человеком, субъективными факторами восприятия (неблагоприятными условиями для восприятия и понимания/усвоения информации). В последнем случае понимание также может происходить не сразу, а с задержкой, например:

- по прошествии определенного времени — было время «подумать» или создались благоприятные условия для обдумывания;
- при изменении условий — была получена дополнительная информация, скажем, произошли определенные события, что дополнило, «протокнуло» недопонятую ранее информацию, «поставило все на свои места».

Понимание (а точнее сказать, осмысленное знание) дает возможность осуществлять действия с новыми предметами на основе уже сложившихся умений и навыков, а также рефлексировать о понятой/усвоенной информации (знании), выводить из нее разные следствия.

Относительно процесса понимания (интерпретации) можно сформулировать следующее: автор (генератор) выражает свои субъективные образы или суждения в формализованном виде (текста, речи), а интерпретатор восстанавливает этот образ или суждение в собственном субъективном контексте. Этот субъективный контекст определяется предыдущим опытом субъекта и текущими обстоятельствами, в которых происходит интерпретация. Таким образом, язык выступает «универсальной онтологией», которая обеспечивает интероперабельность субъектов — носителей языка (точнее говоря, пользователей языка).

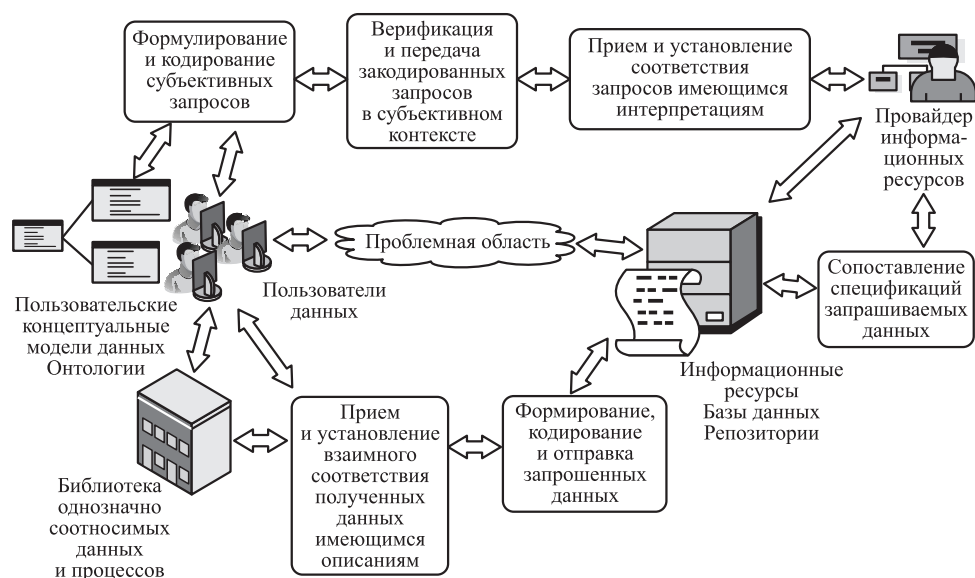


Рис. 3 Структура семантической интероперабельности

Для генератора не существует (скорее, она для него незаметна) проблемы омонимии слов/терминов (неоднозначности используемого слова/термина), поскольку он знает подразумеваемое под ним значение. А вот интерпретатор сталкивается с такой проблемой и вынужден ее решать посредством опоры на контекст сообщения, а если это не помогает, то вынужден использовать вероятностные (частотные) или нечеткие оценки (типа «автор, по всей видимости, имел в виду следующее» или «автор мог иметь в виду либо то, либо это»). Поиск неявно/нечетко выраженного содержания происходит в результате обращения к знаниям, имеющим отношение к данному тексту (рис. 3)

Семиотическая коммуникация предполагает участие отправителя и получателя информационного сообщения, причем в процессе общения каждый из них выступает в роли то генератора, то интерпретатора сообщения.

В практическом плане проблема понимания возникает, когда взаимодействующие агенты сталкиваются с «ситуациями неопределенности ключевых понятий, с нечеткими постановками задач, с лексической полисемией и другими препятствиями к однозначной и эксплицитной трактовке терминов, их смысловых связей и их восприятия специалистами из разных областей» [5].

4 Семантическая геоинтероперабельность

Геоинтероперабельность — это способность геоинформационных систем (1) к обмену геопространственной информацией об объектах и процессах на,

выше и ниже поверхности Земли; (2) к совместному сетевому использованию программной среды для переработки такой информации. Семантическая геоинтероперабельность основывается на согласованном понимании смысла геоданных, например в процессе взаимодействия пользователей геопорталов при решении задач, требующих совместной обработки и использования ими геоданных, или при решении задач концептуального поиска в базах геоданных. Семантическая геоинтероперабельность принципиально отличается от синтаксической [6], так как в последнем случае речь идет только об обеспечении согласованной обработки и обмена геоданными на нескольких уровнях взаимодействия (сетевом, обменных форматах, вызовах удаленных процедур, запросов к базам геоданных).

Проблема обеспечения семантической геоинтероперабельности заключается в обеспечении согласованного взаимодействия специалистов, решающих задачи, требующие совместного использования ими георесурсов, при условии адекватного понимания ими семантики геоданных, полученных в результате обмена. Обеспечение семантической геоинтероперабельности предполагает, что разработка архитектуры средств согласования понимания геоданных является результатом сравнительного анализа существующих метасхем баз геоданных с учетом нескольких уровней взаимодействия и многофакторности взаимодействия пользователей, включая экспертов, на основе геоинформационных описаний и семантики, заложенной в пространственные онтологии и/или геотегаурусы и классификаторы. Ключевой момент обеспечения семантической геоинтероперабельности — создание единой концептуальной модели представления и согласования понимания пользователями геоданных на основе интеграции пространственно распределенной информации (семантической, метрической и топологической), хранимой в геоинформационных базах данных.

Цифровая трансформация геоданных, возможная только при наличии семантической геоинтероперабельности, требует предварительной цифровизации всех геоинформационных ресурсов, которая усложняется присущей геоданным гетерогенностью. При этом необходимы точные и актуальные геоданные о расположении пространственных объектов, их свойствах и характеристиках, их перемещении или изменении во времени, взаимодействии с другими объектами, т. е. речь идет не о расширении атрибутивных данных, а о создании пространственной основы для цифровой трансформации. Полученные на этой основе геоинформационные решения позволяют как упорядочить и организовать различные информационные потоки, так и получать независимые своевременные аналитические отчеты о тех или иных пространственных объектах и происходящих с ними событиях.

Оцифрованные геоданные могут быть получены в результате использования современных технических средств спутникового сканирования зондирования и позиционирования или с помощью авиационных средств аэрофотосъемки земной поверхности. Существенный вклад в формирование геоконтента оказывает фотограмметрия и мобильное лазерное сканирование.

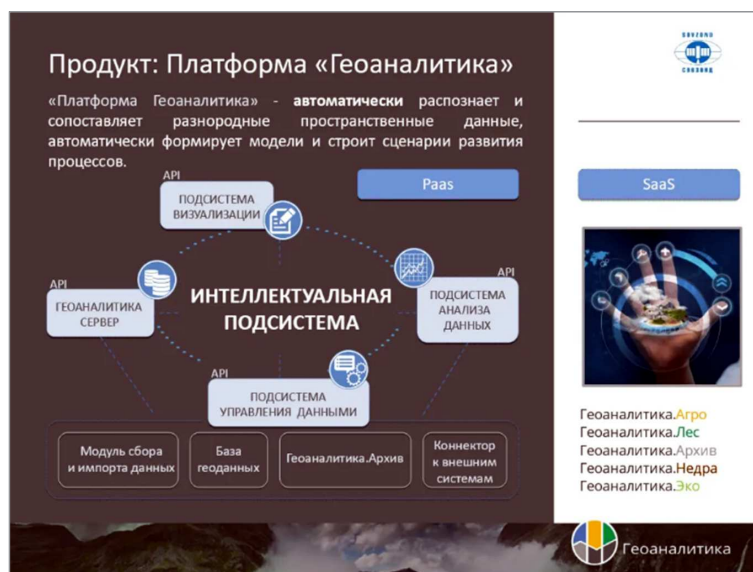


Рис. 4 Реализация цифровой трансформации платформой «Геоаналитика»

Определенный успех в продвижении цифровой трансформации геоданных можно отметить у компании «Совзонд», которая на протяжении 25 лет работает в сфере геопространственных данных, внедряя цифровые информационные геотехнологии. Внедрение современных цифровых технологий наблюдения и пространственного анализа позволило существенно расширить возможности контроля в части использования государственных средств, получать независимые оценки по использованию природных ресурсов и территорий, контролировать собираемость налогов и т. д. Специалистами компании разработана современная геоинформационная облачная платформа «Геоаналитика» [7], решающая широкий спектр задач по получению, хранению и автоматизированной обработке пространственной информации (рис. 4).

Специалисты отмечают определенные достоинства платформы «Геоаналитика», среди которых:

- работа с потоками пространственных данных, алгоритмы BIG DATA;
- централизованное хранение и распределенный доступ к данным;
- автоматизированные процессы обработки и анализа разнородных данных, интерактивная аналитика, моделирование и машинное обучение;
- системы поддержки принятия решений и «облачные» сервисы.

5 Заключение

В работе рассмотрены этапы перехода к цифровому контенту: оцифровывание, цифровизация, цифровая трансформация. Отмечено определяющее значение семантической интероперабельности для реализации цифровой трансформации. Расширение семантической интероперабельности до семантической геоинтероперабельности позволит использовать цифровую трансформацию данных о пространственных объектах и процессах для всей окружающей среды.

Литература

1. Что такое цифровая трансформация? — Hewlett Packard Enterprise, 2020. <https://www.hpe.com/ru/ru/what-is/digital-transformation.html>.
2. Условия и порядок отбора проектов для финансирования по программе «Цифровизация промышленности»: Стандарт Фонда № СФ-И-116. Редакция 2.5. — М.: Фонд развития промышленности, 2021. 31 с. <https://frprf.ru/download/usloviya-programmy-tsifrovizatsiya-promyshlennosti-.pdf>.
3. Макаренко С. И., Олейников А. Я., Черницкая Т. Е. Модели интероперабельности информационных систем // Системы управления, связи и безопасности, 2019. № 4. С. 215–245.
4. Turnitsa C., Tolk A. Battle management language: A triangle with five sides // Spring Simulation Interoperability Workshop Proceedings. — Simulation Interoperability Standards Organization, 2006. Vol. 1. P. 91–100.
5. Тарасов В. Б. Проблема понимания: настоящее и будущее искусственного интеллекта // Открытые семантические технологии проектирования интеллектуальных систем: Мат-лы V Междунар. научн.-технич. конф. — Минск: БГУИР, 2015. С. 25–42.
6. Дулин С. К., Дулина Н. Г. От согласованности геоданных к семантической геоинтероперабельности. — М.: ВЦ РАН, 2014. 28 с.
7. Анонс обновлений ГИС-платформы «Геоаналитика» // Геодезия и картография, 25.09.2018. <https://geocartography.ru/news/industrynews/anons-obnovleniy-gis-platformy-geoanalitika>.

Поступила в редакцию 09.12.20

INTEROPERABILITY AS A KEY CONDITION FOR THE IMPLEMENTATION OF DIGITAL TRANSFORMATION

I. N. Rozenberg¹, S. K. Dulin^{1,2}, and N. G. Dulina²

¹Research & Design Institute for Information Technology, Signalling and Telecommunications on Railway Transport, 27-1 Nizhegorodskaya Str., Moscow 109029, Russian Federation

²Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation

Abstract: The paper analyzes the processes typical for the digitization of information, and the standardization of digitization in the information environment. Three stages of information digitization and their features are considered. It is noted that the most important condition for digital transformation is to ensure the level of semantic interoperability. The characteristic features of this level are described.

Keywords: digitization; digitalization; digital transformation; interoperability; semantic geointeroperability

DOI: 10.14357/08696527210304

References

1. Hewlett Packard Enterprise, 2020. Chto takoe tsifrovaya transformatsiya? [What is digital transformation?]. Available at: <https://www.hpe.com/ru/ru/what-is/digital-transformation.html> (accessed August 23, 2021).
2. Standart Fonda SF-I-116. 2021. Usloviya i poryadok otbora proektov dlya finansirovaniya po programme “Tsifrovizatsiya promyshlennosti” [Fund Standard: Conditions and procedure for selecting projects for financing under the Industry Digitalization program]. Moscow: Industry Development Fund. 31 p. Available at: <https://frprf.ru/download/usloviya-programmy-tsifrovizatsiya-promyshlennosti-.pdf> (accessed August 23, 2021).
3. Makarenko, S. I., A. Ya. Oleynikov, and T. E. Chernitskaya. 2019. Modeli interoperabel'nosti informatsionnykh sistem [Models of interoperability of information systems]. *Sistemy upravleniya, svyazi i bezopasnosti* [Systems of Control, Communication, and Security] 4:215–245.
4. Turnitsa, C., and A. Tolk. 2006. Battle management language: A triangle with five sides. *Spring Simulation Interoperability Workshop Proceedings*. Simulation Interoperability Standards Organization. 1:91–100.
5. Tarasov, V. B. 2015. Problema ponimaniya: nastoyashchee i budushchee iskusstvennogo intellekta [The problem of understanding: The present and the future of artificial intelligence]. *5th Scientific and Technical Conference (International) “Open Semantic Technologies for Intelligent Systems” Proceedings*. Minsk: BSUIR. 25–42

6. Dulin, S. K., and N. G. Dulina. 2014. *Ot soglasovannosti geodannykh k semanticheskoy geointeroperabel'nosti* [From consistency of geodata to semantic geointeroperability]. Moscow: CC RAS. 28 p.
7. Anons obnovleniy GIS-platforny "Geoanalitika" [Announcement of updates to the Geoanalytica GIS platform]. 25.09.2018. Available at: <https://geocartography.ru/news/industrynews/anons-obnovleniy-gis-platforny-geoanalitika> (accessed August 23, 2021).

Received December 9, 2020

Contributors

Rozenberg Igor N. (b. 1965) — Doctor of Science in technology, professor, research advisor, Research & Design Institute for Information Technology, Signalling and Telecommunications on Railway Transport, 27-1 Nizhegorodskaya Str., Moscow 109029, Russian Federation; I.Rozenberg@vniias.ru

Dulin Sergey K. (b. 1950) — Doctor of Science in technology, professor; leading scientist, Institute of Informatics Problems, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; principal scientist, Research & Design Institute for Information Technology, Signalling and Telecommunications on Railway Transport, 27-1 Nizhegorodskaya Str., Moscow 109029, Russian Federation; skdulin@mail.ru

Dulina Natalia G. (b. 1947) — Candidate of Science (PhD) in technology, leading programmer, A. A. Dorodnicyn Computing Center, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences, 40 Vavilov Str., Moscow 119333, Russian Federation; ngdulina@mail.ru

БЕЗОПАСНОЕ МАСШТАБИРОВАНИЕ ЭЛЕКТРОННЫХ БУХГАЛТЕРСКИХ КНИГ НА ОСНОВЕ ТАНГЛА*

А. А. Грушо¹, А. А. Зацаринный², Е. Е. Тимонина³

Аннотация: Рассматривается электронная бухгалтерская книга (ЭБК), основанная на концепции блокчейна. Развитием традиционного блокчейна стал тангл (tangle), который представляется ориентированным ациклическим графом (DAG — Directed Acyclic Graph) и который безопасно хранит информацию о транзакциях. Централизованный консенсус предполагает концентрацию функций контроля в едином органе. Масштабирование ЭБК в ряде случаев порождает угрозы коррупции и мошенничества. Рассматриваются три типа угроз, которые определяются коррупцией и сговором юридических лиц. Обеспечение информационной безопасности при масштабировании ЭБК — это направление, которое начинает развиваться в связи с ростом цифровой экономики.

Ключевые слова: информационная безопасность; тангл; блокчейн; угрозы коррупции и мошенничества

DOI: 10.14357/08696527210305

1 Введение

Электронная бухгалтерская книга основана на концепции блокчейна [1, 2]. Управление записями в блокчейне определяется консенсусом [3]. Все консенсусы можно разделить на централизованные и децентрализованные. С технической точки зрения централизованные и децентрализованные консенсусы схожи, так как каждый узел сети индивидуально ответствен за безопасность и хранение совместно используемых для ЭБК данных. Централизованный консенсус предполагает концентрацию функций контроля в едином органе. Это означает, что имеется единый контактный узел формирования и хранения блокчейнов.

Развитием традиционных блокчейнов стал тангл [4], который представляется ориентированным ациклическим графом (DAG) и который безопасно хранит информацию о транзакциях. Вершины (узлы) в DAG — это транзакции. Сеть тангла состоит из вершин, которые являются сущностями, отражающими транз-

* Работа частично поддержана РФФИ (проект 18-29-03124-мк).

¹Федеральный исследовательский центр «Информатика и управление» Российской академии наук, grusho@yandex.ru

²Федеральный исследовательский центр «Информатика и управление» Российской академии наук, AZatsarinny@ipiran.ru

³Федеральный исследовательский центр «Информатика и управление» Российской академии наук, eltimon@yandex.ru

акции и обеспечивающими валидацию транзакций с помощью ориентированных дуг.

Основная идея тангла состоит в том, что пользователи при регистрации транзакции как узла тангла должны провести определенную работу для «одобрения» других ранее определенных транзакций. Таким образом, каждый пользователь, создающий транзакцию, вносит вклад в безопасность сети. Узлы проверяют, чтобы «одобренные» транзакции были неконфликтными. Если узел выясняет, что транзакция конфликтует с историей в тангле, то этот узел не «одобрит» конфликтную транзакцию.

Для создания транзакций узел выполняет следующие действия:

- (1) выбирает две или более транзакций для «одобрения»;
- (2) проверяет непротиворечивость выбранных для «одобрения» транзакций;
- (3) использует криптографические методы для формирования специальных, сложно вычислимых данных из «одобренных» транзакций и собственной информации.

Первые транзакции в тангле называются генезисом. Все транзакции в тангле прямо или косвенно одобряют генезис.

Основой обеспечения юридической значимости электронного взаимодействия участников экономической деятельности служат договоры (соглашения, контракты), которые содержат три параметра:

- (1) задачу, которую надлежит решить;
- (2) цену решения задачи, которую надлежит решить;
- (3) время, за которое надлежит решить задачу.

Договор, как правило, связывает двух участников: заказчика и исполнителя. Это отношение можно отражать на отдельном графе в виде ориентированной дуги, на которой ставится метка в виде вектора из значений трех указанных параметров.

Результаты выполнения договоров также отражаются в тангле в форме транзакций. При выполнении договоров в транзакциях отражаются акты о решении задачи, документы на оплату выполненных работ или на предоставление авансовых платежей. Эти данные также можно отображать в виде другого ориентированного графа. Таким образом, электронное взаимодействие участников экономической деятельности цифровой экономики должно обеспечивать дистанционное заключение договора и юридическое подтверждение информации о выполнении или невыполнении договора.

Идея ЭБК состоит в том, что каждый субъект экономической деятельности обладает танглом, хранящимся у него и в узле, реализующем централизованный консенсус. В этом тангле фиксируются все договорные обязательства и результаты их выполнения. Основная идея бухгалтерской книги [5] состоит в двойной записи, которая отражена в главной бухгалтерской книге. В ЭБК каждая

транзакция должна найти отражение по крайней мере в двух танглах, которые принадлежат взаимодействующим участникам экономической деятельности. В этом состоит непротиворечивость транзакции, которая проверяется на узле, реализующем централизованный консенсус.

В рамках одной ЭБК вопросы взаимодействий и безопасности рассмотрены в работе [6].

2 Масштабирование электронной бухгалтерской книги

Масштабируемость ЭБК подразумевает взаимодействие нескольких ЭБК. В связи с этим рассмотрим новый вид консенсуса, который назовем сложным централизованным консенсусом (СЦК). Сложный централизованный консенсус предполагает взаимодействие нескольких ЭБК, каждая из которых организована на базе своего централизованного консенсуса.

Основная идея СЦК состоит в построении иерархии ЭБК. Предположим, что должен быть выполнен большой проект, создаваемый на основании нормативного акта (НА) и обеспечивающий выполнение проекта финансирования, которое определяется финансовым документом (ФД). Реализация проекта предполагает начальное участие организаций $O_1, \dots, O_k$, каждая из которых связана со своей ЭБК. Тогда можно создать новый тангл, генезис которого состоит из двух вершин НА и ФД, юридическая значимость которых может быть оформлена соответствующими документами и подписями. Этот тангл можно назвать коренным для проекта. Аналогично каждая ЭБК может создать коренные танглы в своих ЭБК, в которых генезисы будут определяться договорами об участии в проекте между коренным танглом проекта и коренными танглами организаций и соответствующими ФД. Вообще договор между коренным танглом проекта и танглом организации можно сразу создавать на основе уже сделанных ранее проектов организаций, хотя они относятся к разным ЭБК. Наличие общего удостоверяющего центра электронных подписей позволяет это сделать. Основная проблема состоит в обеспечении безопасности (контроля) при взаимодействии различных ЭБК при выполнении проекта.

В работе рассматривается ряд угроз, которые представляются существенными при реализации любого крупного проекта, и исследуется проблема мониторинга с целью выявления признаков реализации этих угроз. Исследование ограничено угрозами, связанными с организацией договорных отношений и финансирования этих отношений. Вопросы выполнения договорных отношений в этой статье не рассматриваются.

3 Угрозы безопасности при масштабировании электронной бухгалтерской книги

Рассматриваются три типа угроз, которые определяются коррупцией и сговором юридических лиц.

Первый класс угроз будем называть циклами. Пусть организация O_1 получила подряд на выполнение одной из задач T проекта с объемом финансирования F . В ходе организации субподрядных работ несколько юридических лиц $O_2, \dots, O_k$ могут договориться о последовательном выполнении задачи T . Каждая из организаций $O_2, \dots, O_k$ получает долю финансирования $\alpha_2 F, \dots, \alpha_k F$, где $\alpha_2 + \dots + \alpha_k = \alpha < 1$. При этом субподрядчики $O_2, \dots, O_k$ не могут и не собирались решать задачу целиком, но договорились (за исключением некоторого процента от F) нелегально передать деньги O_1 , заключив договор между O_k и O_1 на решение исходной задачи T за объем финансирования $(1 - \alpha)F$. Получился цикл, в котором одну и ту же задачу T предлагается решать за различные деньги одной и той же организации O_1 . Этот цикл позволяет перераспределять или обналичивать средства, выделенные на проект. Сюда также входит схема отмывания денег с помощью фирм-однодневок.

Второй класс угроз будем называть коррупцией. Пусть организация O должна организовать решение задачи с трудоемкостью T и с объемом финансирования F , кратко (T, F) . Для решения задачи привлекаются две организации O_1 и O_2 . С первой организацией O_1 есть договоренность об «откате». Тогда распределение работы и финансирования можно организовать следующим образом:

- O_1 получает договор $(T/10, 9F/10)$;
- O_2 получает договор $(9T/10, F/10)$.

Третий класс угроз состоит в существовании первопричины задержки выполнения проекта хотя бы одним из критических узлов реализации проекта, которая позволит увеличить финансирование, требуемое на покрытие останова во время задержки.

Для выявления реализации этих угроз необходимо определить, как и какую информацию надо собирать и где ее обрабатывать. Кроме того, необходимо определить вспомогательную информацию для анализа данных мониторинга.

4 Вспомогательные данные для анализа данных мониторинга выполнения проекта

Пусть организация, обеспечивающая условия функционирования коренного тангла, является центром мониторинга безопасности реализации проекта. Для сбора, хранения и анализа данных мониторинга безопасности должен существовать вычислительный центр проекта (ВЦП), где в базах данных хранятся следующие вспомогательные данные.

1. При подготовке к реализации проекта должна быть разработана номенклатура задач, которые необходимо решить при реализации проекта, а именно: по каждой соответствующей задаче должны быть хотя бы ориентировочно указаны:

- разработка/цена/время;
- готовое изделие/цена/время;
- возможность совместимости/цена/время достижения совместимости.

2. Описание проекта:

- структурная модель;
- разбиение на подсистемы;
- цены/сроки/тестирование результатов.

3. Обновляемая база данных договорных отношений участников проекта (независимо от ЭБК), т. е. совокупность дуг между участниками договорных отношений с метками (задача/цена/время). Сообщения о закрытии договора при его выполнении. Эти данные обновляются и передаются в центр мониторинга организациями, поддерживающими централизованные консенсусы каждой ЭБК. При этом сами танглы экономических субъектов или их фрагменты не передаются, кроме как по отдельным запросам.

4. Средства интеллектуального анализа данных (ИАД).

Далее показано, что этих данных достаточно для выявления признаков реализации указанных выше угроз. Рассмотрим методы выявления признаков угроз.

5 Методы анализа безопасности при масштабировании электронной бухгалтерской книги

Рассмотрим методы выявления циклов. В поступающих данных обновлений договоров можно выявлять факты, когда субъект O заключает договор (исходящая дуга) и субъект O позже (сравнение меток времени) становится исполнителем (входящая дуга). Данный факт может считаться признаком возможного цикла. Тогда выделяются дуги, по которым можно построить ориентированные пути, исходящие из O , и ориентированные пути, входящие в O . Существование хотя бы одного пересечения на множествах этих путей означает существование цикла. Более того, построение всех циклов позволяет выявить все группы сговора с O . Детальный анализ меток на этих циклах позволит восстановить схему и причины образования циклов.

Рассмотрим методы выявления схем коррупции. Для каждой вершины, из которой выходят несколько договорных дуг, необходимо провести сравнение меток этих дуг. С этой целью надо использовать базу данных номенклатуры решаемых задач проекта. Если комплексная задача внесена в эту базу, то должно существовать разбиение этой задачи на подзадачи, иначе зачем нужны несколько субподрядчиков? Тогда подзадачи также должны присутствовать в номенклатуре задач проекта, для этого надо смотреть структурное описание проекта. Для

каждой задачи, по условию, есть данные о ее формулировке и ориентировочной цене. Это дает возможность сравнивать условия договоров субподрядчиков по сложности стоящих задач и цене за решение этих задач. Отсюда сразу можно увидеть следы коррупции. Если описания подзадач отсутствуют, то единственный способ их создать — это квалифицированная экспертиза. К сожалению, схем коррупции, подобной описанной, может быть много, но базовая основа у большинства таких схем одинаковая. Это позволяет априори создавать вспомогательные данные для выявления подобных схем.

Рассмотрим третий класс угроз, связанный с повышением цены проекта за счет задержек в решении отдельных задач. Проблема состоит в том, что существуют несколько способов прятать первопричину задержек и объяснять все «объективными» обстоятельствами. Самый простой способ состоит в формировании условий, когда отсутствует ответственность за порождение задержки. Пусть субъект O должен решить сложную задачу $T = (T_1, T_2)$, в которой T_1 состоит в создании софта для обработки данных в задаче T_2 . Субъект O , заинтересованный в создании задержки, распределяет задачи между субъектами O_1 и O_2 . Созданный софт тестируется в O_1 и показывает хорошие результаты. Однако в O_2 этот софт не работает. Выявление причин отказа работы софта и необходимые работы по модернизации порождают требуемую для O задержку. При этом O не несет ответственности за задержку, так как не участвует в разработке софта и его применении. Субъект O_1 отказывается нести ответственность, так как у него софт хорошо работает. Субъект O_2 отказывается нести ответственность за задержку потому, что софт создавал не он. В данном простейшем случае первопричину легко найти. Она состоит в том, что O некорректно поставил задачу и не организовал взаимодействие между O_1 и O_2 . Некорректность постановки задачи состоит в том, что в условии задачи T_1 должно присутствовать условие о необходимости использовать функционал софта в задаче T_2 , которая будет решаться в O_2 . Для этого тестирование софта должно вестись не только в O_1 , но также одновременно и в O_2 .

Рассмотрим, как можно выявить первопричину задержки на ВЦП. В рассматриваемом простейшем случае задержка проекта определяется невозможностью O решить сложную задачу T , но ответственность за задержку O не берет. Тогда по запросу ВЦП организации, реализующие централизованные консенсусы и имеющие копии танглов O , O_1 и O_2 , делают копии договоров и высылают ВЦП. Если ни в одном договоре не упоминается задача адаптации софта для решения задачи T_2 в O_2 , то ответственность за задержку ложится на O . Если O_1 проигнорировал требование адаптации в O_2 , то ответственность ложится на O_1 . Если O_2 декларирует неработоспособность софта, но O_1 показывает положительные результаты тестирования софта в O_2 , то ответственность за задержку ложится на O_2 . Идентификация ответственности влечет большие потери для ответственного за задержку (штраф, компенсация ущерба и т. д.). Тогда все звенья реализации проекта становятся заинтересованными в недопущении задержек выполнения проекта.

Рассмотренные простейшие примеры реализации угроз не исчерпывают всего многообразия конкретных сценариев реализации рассматриваемых угроз. Однако видно, что потенциальный доступ к танглам и имеющиеся вспомогательные данные в главном тангле позволяют эффективно вести анализ появления признаков реализации угроз. Наиболее часто встречается коррупция. Выявление коррупции не дает достаточной информации для выявления причин коррупции и заинтересованных лиц. Поэтому в работе рассмотрены задачи только мониторинга реализации проекта. Также не рассмотрены задачи поиска первопричин неявных сбоя [7–9], порождающих задержки. Такие сбои могут порождаться не только заинтересованными исполнителями проекта, но также наличием вредоносного кода в импортных программах [10] или даже скрытыми дверями (Backdoors), известными только противнику [11]. Например, при наступлении определенных событий возможен сброс конфигурационных параметров, используемых для согласованной работы различных приложений [12].

6 Заключение

Результаты, полученные в статье, являются развитием серии работ авторов [2, 6], посвященных построению ЭБК на базе централизованных консенсусов и танглов применительно к современным проблемам цифровой трансформации. Так, ЭБК на основе танглов служит, по существу, примером одного из возможных подходов к реализации цифровой экономики.

Показано, что обеспечение информационной безопасности при масштабировании ЭБК становится новым направлением, обусловленным интенсивным развитием цифровой экономики. При этом существенную роль играют блокчейн-технологии и их различные модификации.

В рамках масштабируемости ЭБК, состоящей во взаимодействии нескольких ЭБК, объективно возникает новый вид консенсуса, определенный как СЦК. Основная идея СЦК состоит в построении иерархии ЭБК.

Масштабирование ЭБК порождает в ряде случаев угрозы коррупции и мошенничества. Предложена классификация угроз выполнения крупных проектов, обусловленных коррупцией и сговором юридических лиц. Определены и формализованы три типа угроз.

Предложены методы анализа безопасности при масштабировании ЭБК применительно к трем типам угроз. Приведены простейшие примеры реализации угроз.

Литература

1. Лелу Л. Блокчейн от А до Я. Все о технологии десятилетия / Пер. с фр. — М.: Эксмо, 2018. 190 с. (*Leloup L. Blockchain: La revolution de la confiance. Paris: Groupe Eyrolles, 2017. 160 p.*)

2. Грушо А. А., Зацаринный А. А., Тимонина Е. Е. Электронная бухгалтерская книга на базе ситуационных центров для цифровой экономики // Системы и средства информатики, 2019. Т. 29. № 2. С. 4–11.
3. Wahab A., Mahmood W. Survey of consensus protocols. 2018. https://www.researchgate.net/publication/328160285_Survey_of_Consensus_Protocols.
4. Popov S. The tangle. 2018. https://assets.ctfassets.net/r1dr6vzfxhev/2t4uxvslqk0EUau6g2sw0g/45eae33637ca92f85dd9f4a3a218e1ec/iota1_4_3.pdf.
5. Брыкова Н. В. Теория бухгалтерского учета: Баланс и система счетов. — 2-е изд. — М.: Академия, 2011. 80 с.
6. Грушо А. А., Зацаринный А. А., Тимонина Е. Е. Описание динамики цифровой экономики с помощью электронной бухгалтерской книги // Системы и средства информатики, 2020. Т. 30. № 1. С. 108–114.
7. Grusho A., Grusho N., Zabezhaio 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.
8. Грушо Н. А., Грушо А. А., Тимонина Е. Е. Локализация сбоя с помощью метаданных // Проблемы информационной безопасности. Компьютерные системы, 2020. № 3. С. 9–15.
9. Грушо А. А., Грушо Н. А., Забежайло М. И., Тимонина Е. Е. Локализация исходной причины аномалии // Проблемы информационной безопасности. Компьютерные системы, 2020. № 4. С. 9–16.
10. Грушо А. А., Грушо Н. А., Тимонина Е. Е. Методы защиты информации от атак с помощью скрытых каналов и враждебных программно-аппаратных агентов в распределенных системах // Вестн. РГГУ. Сер.: Документоведение и архивоведение. Информатика. Защита информации и информационная безопасность, 2009. № 10. С. 33–45.
11. Грушо А. А., Грушо Н. А., Забежайло М. И., Тимонина Е. Е. «Закладки» без вредоносного кода // Системы и средства информатики, 2021. Т. 31. № 2. С. 4–15.
12. Грушо А. А., Забежайло М. И., Зацаринный А. А., Николаев А. В., Писковский В. О., Тимонина Е. Е. Классификация ошибочных состояний в распределенных вычислительных системах и источники их возникновения // Системы и средства информатики, 2017. Т. 27. № 2. С. 29–40.

Поступила в редакцию 23.07.21

SECURE SCALING OF ELECTRONIC LEDGERS BASED ON TANGLES

A. A. Grusho, A. A. Zatsarinny, and E. E. Timonina

Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation

Abstract: The paper deals with an electronic ledger based on the concept of blockchain. The development of the traditional blockchain is the tangle (tangle), which is represented by an oriented acyclic graph (DAG — Directed Acyclic Graph) and which securely stores transaction information. A centralized consensus implies a concentration of control functions in a single body. The scaling up of e-books in some cases poses threats of corruption and fraud. Three types of threats are considered that are determined by corruption and collusion of legal entities. Ensuring information security when scaling electronic books is a direction that begins to develop in connection with the growth of the digital economy.

Keywords: information security; tangle; blockchain; threats of corruption and fraud

DOI: 10.14357/08696527210305

Acknowledgments

The paper was partially supported by the Russian Foundation for Basic Research (project 18-29-03124-mk).

References

1. Leloup, L. 2017. *Blockchain: La revolution de la confiance*. Paris: Groupe Eyrolles. 160 p.
2. Grusho, A. A., A. A. Zatsarinny, and E. E. Timonina. 2019. Elektronnaya bukhgalterskaya kniga na baze situatsionnykh tsentrov dlya tsifrovoy ekonomiki [The electronic ledger on the basis of the situational centers for digital economy]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(2):4–11.
3. Wahab, A., and W. Mahmood. 2018. Survey of consensus protocols. Available at: https://www.researchgate.net/publication/328160285_Survey_of_Consensus_Protocols (accessed April 24, 2021).
4. Popov, S. 2018. The tangle. Available at: https://assets.ctfassets.net/r1dr6vzfxhev/2t4uxvslqk0EUau6g2sw0g/45eae33637ca92f85dd9f4a3a218elec/iotal.4_3.pdf (accessed August 24, 2021).
5. Brykova, N. V. 2011. *Teoriya bukhgalterskogo ucheta: Balans i sistema schetov* [Accounting theory: Balance sheet and system of accounts]. 2nd ed. Moscow: Akademiya. 80 p.
6. Grusho, A. A., A. A. Zatsarinny, and E. E. Timonina. 2020. Opisanie dinamiki tsifrovoy ekonomiki s pomoshch'yu elektronnoy bukhgalterskoy knigi [Description

- of digital economy dynamics using an electronic ledger]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(1):108–114.
7. Grusho, A., N. Grusho, M. Zabezhalo, 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.
 8. Grusho, N. A., A. A. Grusho, and E. E. Timonina. 2020. Lokalizatsiya sboev s pomoshch'yu metadannykh [Localizing failures with metadata]. *Problemy informatsionnoy bezopasnosti. Komp'yuternye sistemy* [Information Security Problems. Computer Systems] 3:9–15.
 9. Grusho, A. A., N. A. Grusho, M. I. Zabezhalo, and E. E. Timonina. 2020. Lokalizatsiya iskhodnoy prichiny anomalii [Root cause anomaly localization]. *Problemy informatsionnoy bezopasnosti. Komp'yuternye sistemy* [Information Security Problems. Computer Systems] 4:9–16.
 10. Grusho, A. A., N. A. Grusho, and E. E. Timonina. 2009. Metody zashchity informatsii ot atak s pomoshch'yu skrytykh kanalov i vrazhdebnykh programmno-apparatnykh agentov v raspredelennykh sistemakh [Methods of information protection against covert channels attacks and malicious software/hardware agents in distributed systems]. *Vestnik RGGU. Ser. Dokumentovedenie i arkhivovedenie. Informatika. Zashchita informatsii i informatsionnaya bezopasnost'* [RGGU Bulletin. Document Science and Archive Science. Informatics. Information security and information security ser.] 10:33–45.
 11. Grusho, A. A., N. A. Grusho, M. I. Zabezhalo, and E. E. Timonina. 2021. “Zakladki” bez vredonosnogo koda [Hidden impact without malicious code]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 31(2):4–15.
 12. Grusho, A. A., M. I. Zabezhalo, A. A. Zatsarinny, A. V. Nikolaev, V. O. Piskovski, and E. E. Timonina. 2017. Klassifikatsiya oshibochnykh sostoyaniy v raspredelennykh vychislitel'nykh sistemakh i istochniki ikh vozniknoveniya [Erroneous states classification in distributed computing systems and sources of their occurrence]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 27(2):29–40.

Received July 23, 2021

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

Zatsarinny Alexander A. (b. 1951) — Doctor of Science in technology, 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; AZatsarinny@ipiran.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

ВЫЧИСЛЕНИЯ НА ОСНОВЕ ВЕРОЯТНОСТНОЙ МОДЕЛИ АНАЛИЗА ГЛАВНЫХ КОМПОНЕНТ

*М. П. Кривенко*¹

Аннотация: Эффективное решение задач анализа многомерных данных требует применения сложных вероятностных моделей, среди которых достойное место занимает вероятностная модель анализа главных компонент (PPCA — Probabilistic Principal Component Analysis). Она позволяет адекватно описывать реальные данные, формально ставить и решать задачу выбора ее параметров. Применение PPCA на практике сопряжено с большим объемом трудоемких вычислений. Рассматриваются приемы существенного снижения временных затрат при вычислении плотности многомерного нормального распределения. Для этого обосновывается целесообразность применения разложения Холецкого для ковариационной матрицы и тождества Вудбери для ее PPCA-факторизации. Приводится краткое описание экспериментов, позволяющих оценить реальные временные характеристики алгоритмов и выявить условия их эффективного применения. Попутно даются рекомендации по реализации отдельных операций.

Ключевые слова: вычисления для вероятностных анализаторов главных компонент; факторизация Холецкого; матричное тождество Вудбери; компьютерный эксперимент

DOI: 10.14357/08696527210306

1 Введение

Атрибутами модели обрабатываемых данных являются адекватность, аналитическая и вычислительная доступность, эффективность получаемых на ее основе решений. С теоретической и практической точек зрения плодотворным оказывается предположение о смеси нормальных распределений наблюдений. За счет структурных усложнений оно позволяет лучше описывать многообразие реальных данных, оставляя в силе преимущества относительно простых аналитических решений типовых вероятностных задач. Но в случае анализа многомерных данных использование смеси приводит к проблемам проклятия размерности. Это, в свою очередь, актуализирует вопросы выбора варианта модели и порождает необходимость снижения размерности данных.

К наиболее простым и широко используемым методам снижения размерности путем преобразования относятся линейные, а среди них — анализ главных компонент (PCA — Principal Component Analysis). Существенного прогресса

¹Федеральный исследовательский центр «Информатика и управление» Российской академии наук, mkrivenko@ipiran.ru

в применении PCA для данных, демонстрирующих нелинейность, удалось добиться в рамках вероятностной модели анализаторов главных компонент (PPCA) и ее обобщения до соответствующей смеси. В [1] рассматривались возможности разных подходов к выбору структурных параметров модели. Для реализации выбора размерностей применительно к задачам классификации предлагалось использовать комбинацию известных методов выбора размерностей принятой модели, а для оценивания качества обучаемой классификации данных — методы управления выборкой, применение которых немислимо без применения эффективных вычислительных алгоритмов.

Базовая вероятностная модель PPCA для сниженной размерности основывается на представлении

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

где $\mathbf{y}$ — наблюдаемая $(d \times 1)$ -переменная, $\mathbf{y} \sim N(\boldsymbol{\mu}, \mathbf{C}(k))$; $\mathbf{W}$ — $(d \times k)$ -матрица преобразования; $\mathbf{x}$ — латентная $(k \times 1)$ -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 σ^2 суть параметры модели. В случае модели смеси PPCA каждому элементу смеси соответствует свой набор перечисленных параметров, а также вес элемента смеси.

Для оценивания параметров PPCA при фиксированных значениях структурных параметров существуют, согласно [2], два основных варианта:

- (1) поэтапный с помощью спектрального разложения ковариационных матриц, когда сначала находятся оценки параметров смеси нормальных распределений для исходных значений d , затем осуществляется факторизация ковариационных матриц элементов смеси при заданных значениях сниженной размерности и, наконец, строятся оценки параметров $\mathbf{W}$ и σ^2 ;
- (2) одновременный с помощью ЕМ-алгоритма (Expectation-Maximization) для нахождения оценок всех параметров в совокупности.

Последний 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 подразумевает многократное оценивание ее параметров. Поэтому поэтапный вариант оценивания становится более привлекательным, так как позволяет единойжды получать представления ковариационных матриц элементов смеси, а затем многократно формировать из них нужные оценки.

Ключевым элементом всех вариантов обработки обучающей выборки оказываются вычисления, касающиеся плотности многомерного нормального распределения при заданной факторизации его ковариационной матрицы. Они многократно повторяются при оценивании параметров модели и при исследованиях адекватности выбранной модели. Неэффективная реализация соответствующих процедур приводит к вынужденному сокращению вариантов обработки данных и, как следствие, к искаженным результатам исследований.

2 Методы вычислений

Рассмотрим частную задачу вычисления плотности многомерного нормального распределения с нулевым средним и невырожденной ковариационной матрицей $\mathbf{C}$

$$\varphi(\mathbf{x}) = \frac{1}{(2\pi)^{d/2} |\mathbf{C}|^{1/2}} \exp \left(-\frac{1}{2} \mathbf{x}^T \mathbf{C}^{-1} \mathbf{x} \right).$$

Для ее решения обычно используется спектральное разложение симметричной матрицы $\mathbf{C}$, на основе которого подсчитывается определитель и находится обратная матрица, т. е. используется разложение

$$\mathbf{C} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T,$$

где $\mathbf{V} \mathbf{V}^T = \mathbf{V}^T \mathbf{V} = \mathbf{I}$; $\mathbf{\Lambda}$ — диагональная матрица с $\lambda_{ii} > 0$. Тогда

$$|\mathbf{C}| = \prod_{i=1}^d \lambda_{ii}; \quad \mathbf{C}^{-1} = \mathbf{V} \mathbf{\Lambda}^{-1} \mathbf{V}^T.$$

Далее получаем для показателя экспоненты h в выражении для $\varphi(\mathbf{x})$:

$$h = \mathbf{x}^T \mathbf{C}^{-1} \mathbf{x} = \left\| \mathbf{\Lambda}^{-1/2} \mathbf{V}^T \mathbf{x} \right\|^2.$$

Соответствующий метод вычисления $\varphi(\mathbf{x})$, использующий спектральное разложение ковариационной матрицы, обозначим как SC.

Другим подходом может стать использование разложения Холецкого. Для положительно определенной симметричной матрицы существует единственная нижняя треугольная матрица $\hat{\mathbf{L}}$ размера $d \times d$ с положительными диагональными элементами такая, что

$$\mathbf{C} = \hat{\mathbf{L}} \hat{\mathbf{L}}^T.$$

После того как $\hat{\mathbf{L}}$ получена, имеем:

$$|\mathbf{C}| = \prod_{i=1}^d l_{ii}^2;$$

$$\mathbf{C}^{-1} = \left(\hat{\mathbf{L}}^{-1} \right)^T \hat{\mathbf{L}}^{-1}. \quad (1)$$

Для произвольной нижней треугольной невырожденной матрицы $\mathbf{L}$ получить $\mathbf{T} = \mathbf{L}^{-1}$ достаточно просто. Из определения обратной матрицы, являющейся

также нижней треугольной, для нахождения ее элементов с индексами (i, j) получаем уравнения:

$$\sum_{k=j}^i l_{ik} t_{kj} = \begin{cases} 1, & i = j; \\ 0, & i \neq j. \end{cases}$$

Они при каждом $j = 1, 2, \dots, d$ последовательно для $i = j + 1, \dots, d$ дают решения:

$$t_{ij} = \begin{cases} \frac{1}{l_{ij}}, & i = j; \\ -\frac{\sum_{k=j}^{i-1} l_{ik} t_{kj}}{l_{ii}}, & i > j. \end{cases}$$

Схожие формулы приведены в [3, разд. 11.2.2].

В случае ковариационной матрицы представление (1) позволяет после перемножения матриц получить искомые элементы $\mathbf{C}^{-1}$.

Но в нахождении $\mathbf{C}^{-1}$ нет необходимости, так как фактически нужна величина h показателя экспоненты, принимающая вид:

$$h = \left(\hat{\mathbf{L}}^{-1} \mathbf{x} \right)^T \left(\hat{\mathbf{L}}^{-1} \mathbf{x} \right) = \left\| \hat{\mathbf{L}}^{-1} \mathbf{x} \right\|^2.$$

Для ее получения достаточно относительно неизвестного $\mathbf{z}$ решить уравнение $\hat{\mathbf{L}} \mathbf{z} = \mathbf{x}$ с нижней треугольной матрицей коэффициентов прямой подстановкой, а затем вычислить $h = \|\mathbf{z}\|^2$.

Таким образом, при заданных $\mathbf{C}$ и $\mathbf{x}$ получаем схему вычислений:

- нахождение разложения Холецкого $\mathbf{C} = \hat{\mathbf{L}} \hat{\mathbf{L}}^T$;
- решение относительно $\mathbf{z}$ уравнения $\hat{\mathbf{L}} \mathbf{z} = \mathbf{x}$;
- подсчет $|\mathbf{C}|$ и h .

Описанный метод назовем СС, указывая на использование для ковариационной матрицы разложения Холецкого.

Но данная схема не учитывает в случае РРСА специфики факторизации ковариационной матрицы

$$\mathbf{C} = \sigma^2 \mathbf{I} + \mathbf{W} \mathbf{W}^T.$$

Она создает предпосылки для упрощения сложности вычислений при сниженной размерности признакового пространства, когда $k < d$, а матрица $\mathbf{W}$ имеет размеры $d \times k$. Для этого естественно обратиться к тождеству Вудбери и получить представление (далее пришлось ввести специфическое обозначение для единичной матрицы меньших размеров):

$$\left(\sigma^2 \mathbf{I} + \mathbf{W} \mathbf{W}^T \right)^{-1} = \frac{1}{\sigma^2} \left[\mathbf{I} - \mathbf{W} \left(\sigma^2 \mathbf{I}_{k \times k} + \mathbf{W}^T \mathbf{W} \right)^{-1} \mathbf{W}^T \right].$$

Кроме этого,

$$\left| \sigma^2 \mathbf{I} + \mathbf{W} \mathbf{W}^T \right| = (\sigma^2)^{d-k} \left| \sigma^2 \mathbf{I}_{k \times k} + \mathbf{W}^T \mathbf{W} \right|.$$

Матрица «сниженного размера» $k \times k$ в литературе по РРСА обычно обозначается как $\mathbf{M} = \sigma^2 \mathbf{I}_{k \times k} + \mathbf{W}^T \mathbf{W}$. Понятно, что она является симметричной и неотрицательно определенной. Далее дополнительно к этому потребуем, чтобы матрица $\mathbf{W}^T \mathbf{W}$ была невырожденной. Это вполне естественно в случае РРСА, когда столбцы матрицы $\mathbf{W}$ представляют базис в пространстве сниженной размерности.

С учетом разложения Холецкого $\mathbf{M} = \tilde{\mathbf{L}} \tilde{\mathbf{L}}^T$ получаем:

$$|\mathbf{C}| = \sigma^{2d-2k} \prod_{i=1}^k \tilde{l}_{ii}^2; \quad h = \frac{1}{\sigma^2} \left[\|\mathbf{x}\|^2 - \left\| \tilde{\mathbf{L}}^{-1} (\mathbf{W}^T \mathbf{x}) \right\|^2 \right].$$

Как и ранее, теперь осталось решить относительно неизвестного $\mathbf{z}$ уравнение $\tilde{\mathbf{L}} \mathbf{z} = \mathbf{W}^T \mathbf{x}$ с простой матрицей коэффициентов, после чего вычислить $h = \|\mathbf{z}\|^2$. Напомним, что матрица $\tilde{\mathbf{L}}$ имеет размеры $k \times k$. В результате получили метод СМ, подчеркивая в названии использование матрицы $\mathbf{M}$, а не $\mathbf{C}$.

Для метода СМ возникает дополнительная возможность сокращения времени вычислений. Рассмотрим отдельно нахождение $\|\tilde{\mathbf{L}}^{-1} \mathbf{W}^T \mathbf{x}\|^2$. Дело в том, что для прямоугольной матрицы существует представление, подобное разложению Холецкого, а именно:

$$\mathbf{W}^T = \hat{\mathbf{L}} \hat{\mathbf{U}},$$

где $\hat{\mathbf{L}}$ — нижняя треугольная матрица размера $k \times k$; $\hat{\mathbf{U}}$ — верхняя треугольная матрица размера $k \times d$. Алгоритм его получения именно в этой форме приведен в [4, разд. 3.2.10]. Тогда

$$\left\| \tilde{\mathbf{L}}^{-1} \mathbf{W}^T \mathbf{x} \right\|^2 = \left\| \bar{\bar{\mathbf{L}}}^{-1} (\hat{\mathbf{U}} \mathbf{x}) \right\|^2,$$

т. е. $\bar{\bar{\mathbf{L}}} = \hat{\mathbf{L}}^{-1} \tilde{\mathbf{L}}$ по-прежнему остается нижней треугольной матрицей, а за счет наличия нулей под диагональю в $\hat{\mathbf{U}}$ снижается число операций при вычислении вектора $\hat{\mathbf{U}} \mathbf{x}$.

Прибегнуть к возможностям снижения размерности можно и в случае использования спектрального разложения матриц. Тогда с учетом разложения $\mathbf{M} = \tilde{\mathbf{V}} \tilde{\mathbf{\Lambda}} \tilde{\mathbf{V}}^T$ получаем:

$$|\mathbf{C}| = (\sigma^2)^{d-k} \prod_{i=1}^k \tilde{\lambda}_{ii}^2; \quad h = \frac{1}{\sigma^2} \left[\|\mathbf{x}\|^2 - \left\| \tilde{\mathbf{\Lambda}}^{-1/2} \tilde{\mathbf{V}}^T \mathbf{W}^T \mathbf{x} \right\|^2 \right],$$

что составляет метод SM.

3 Эксперименты

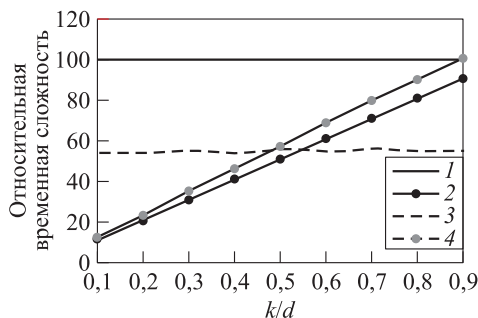
Разнообразие методов приводит к необходимости выбора варианта, обеспечивающего наилучшие временные характеристики. Рассмотрение входных данных большого размера и оценка порядка роста времени работы алгоритма безусловно служат хорошим ориентиром, но могут ввести в заблуждение при оценке эффективности процедур в реальных условиях. Это неприемлемо при многократной обработке данных в ходе подбора модели методами управления обработки результатов наблюдений, при итерационных методах оценивания параметров модели, в ходе оценивания качества полученных решений путем моделирования. Понятно, что точное представление о временной сложности алгоритма зависит от установления элементарных операций, определения способов их реализации, сопряжено с трудоемким процессом программирования и машинных экспериментов. Но только таким путем может быть получено полное представление о временной сложности предлагаемых методов. Далее исследование временной сложности проводилось как аналитически, так и экспериментальным путем. Для оценивания временных затрат использовалась запись TStopwatch модуля System.Diagnostics в среде Embarcadero® Delphi 10.3 для процессора Intel® Core™ i5-9500. Эксперименты заключались в повторном выполнении каждой из процедур N_{exp} раз, замеры суммарного времени корректировались с учетом выполнения вспомогательных операций (например, исключалось время на генерирование случайных данных). Опробовались значения $d = 20, 40, 50, 80, 100$ и $k/d = 0,1; 0,2; \dots; 0,9$ для каждого из d . Исследование сопровождалось контролем точности получаемых результатов, проводимым сравнением с одним из методов, принятых в качестве базового. Точность в случае аппроксимации временных показателей характеризовалась относительной ошибкой вычисления $\ln \varphi(x)$, которая при $N_{\text{exp}} = 10^5$ не превышала 10^{-12} .

Сначала была проведена серия экспериментов, позволяющих упорядочить реальную временную сложность построенных алгоритмов. С этой целью сравнивались суммарные времена вычисления $\ln \varphi(x)$ для случайно выбранных значений $\mathbf{x}$ при $N_{\text{exp}} = 10^5$. Пример получающихся результатов дан на рисунке.

Проведенные исследования позволяют подытожить:

- из методов, не учитывающих снижение размерности, ожидаемым преимуществом обладает СС по сравнению с SC (за счет замены обращения матрицы решением системы линейных уравнений с треугольной матрицей);
- существуют области значений k , для которых рационально использовать факторизацию ковариационной матрицы, учитывающую снижение размерности, и переходить к алгоритмам типа SM или CM.

В ходе компьютерных экспериментов попутно были найдены эффективные решения отдельных задач, которые нашли претворение в анализируемых алгоритмах. В частности, продемонстрировано, что скалярное произведение вектора на себя быстрее вычисляется с помощью возведения в квадрат компонент, а не их



Зависимость относительной временной сложности методов SC (1), SM (2), CC (3) и CM (4) от k/d при фиксированном $d = 100$

перемножения; относительные преимущества достигают 25% при малой размерности вектора и постепенно сходят на нет: при $d = 50$ это 10%, при $d = 100$ — 4%, при $d = 500$ — 2%. Поэтому так ценно за счет разложения обратной ковариационной матрицы сводить вычисление $\mathbf{x}^T \mathbf{C}^{-1} \mathbf{x}$ к нахождению нормы вектора. При этом, если речь идет о разложении Холецкого, выигрыш во времени выполнения достигает ожидаемого значения в 50% уже при $d > 60$.

После того как в качестве объектов исследований остались методы CC и SM, актуальным стал вопрос о критерии выбора одного из них при принятой модели и данных с реальными объемами. Если T_{CC} и T_{SM} суть временные сложности соответствующих алгоритмов при заданных d и k , а $k_0(d)$ есть решение уравнения $T_{CC} - T_{SM} = 0$, то предпочтение при $k < k_0(d)$ надо отдать методу SM, в противном случае — методу CC. Анализ алгоритмов построенных методов позволил выделить набор факторов, определяющих их временную сложность: для SC — это $const, d, d^2$, для SM — $const, k, kd, d$, для CC — $const, d, d^2$, для CM — $const, k, k^2, kd, d$. Тогда для T_{CC} и T_{SM} можно записать:

$$\delta(d, k) = T_{CC}(d, k) - T_{SM}(d, k) = a_0 + a_1 k + a_2 kd + a_3 d + a_4 d^2, \quad (2)$$

где $a_0, \dots, a_4$ — некоторые коэффициенты, зависящие от реализации алгоритма на определенном вычислителе. Далее принимается упрощение, что эти величины — константы, но надо помнить, что это допущение может стать источником ошибок (например, время доступа к элементам матрицы, очевидно, зависит от ее размерности, что может существенно повлиять на значения коэффициентов в (2)).

Из уравнения $\delta(d, k) = 0$ получаем его решение:

$$k_0 = \frac{-a_0 - a_3 d - a_4 d^2}{a_1 + a_2 d} = \frac{-(a_0/a_2)/d - a_3/a_2 - (a_4/a_2)d}{(a_1/a_2)/d + 1}, \quad (3)$$

откуда следует, что

$$\begin{aligned} \frac{k_0}{d} &= \frac{-b_0/d - b_3/d - b_4}{b_1/d + 1} = \left(-\frac{b_0}{d^2} - \frac{b_3}{d} - b_4 \right) \left(1 - \frac{b_1}{d} + O\left(\frac{1}{d^2}\right) \right) = \\ &= -b_4 + \frac{b_1 b_4 - b_3}{d} + O\left(\frac{1}{d^2}\right). \quad (4) \end{aligned}$$

где $b_0 = a_0/a_2$, $b_1 = a_1/a_2$, $b_3 = a_3/a_2$, $b_4 = a_4/a_2$ и асимптотическое разложение правомерно при $d > b_1$.

Смысл проделанных преобразований в следующем:

- переход от коэффициентов $a \dots$ к $b \dots$ избавляет от необходимости обращать внимание на то, в каких единицах измеряется время при экспериментах;
- полученное представление говорит о том, что критерий проще строить на основе величины k_0/d , которая при увеличении d стремится к константе $-a_4/a_2$;
- асимптотическое разложение демонстрирует, когда применим предельный результат, но использовать его для расчета k_0 вряд ли имеет смысл, так как исходная формула сама по себе достаточно проста.

Нахождение коэффициентов в (2) проводилось с помощью регрессионного анализа и метода наименьших квадратов. В результате были получены оценки: $a_0^* = 2,190196$, $a_1^* = -0,666241$, $a_2^* = -0,253967$, $a_3^* = 1,299035$, $a_4^* = 0,127478$, а также, что $k_0^*/d \xrightarrow{d \rightarrow \infty} 0,50$. Кроме этого оказалось, что асимптотическое разложение верно при $d > b_1^* = 2,6$. Для иллюстрации возможностей полученных результатов в таблице приведены отдельные значения $R(d, k_0) = T_{SM}/T_{CC}$ в процентах для некоторых d и соответствующих k_0 , найденных согласно (3) и (4) на основе полученных оценок $a_0^*, \dots, a_4^*$. При этом серой заливкой отмечены ячейки с теми значениями d , по которым строилась регрессионная зависимость. Полученные результаты говорят о точности предложенного варианта нахождения k_0 , а также о практичности замены его простым правилом: если сниженная размерность меньше половины исходной, то следует пользоваться алгоритмом SM, т. е. учитывать особенности факторизации ковариационной матрицы в случае PPCA, если же это не так, то необходимо обращаться к СС, т. е. к обработке разложения Холецкого ковариационной матрицы.

Поведение $R(d, k_0)$ для различных d и двух типов k_0

Тип k_0	d										
	20	30	40	50	60	70	80	90	100	150	200
k_0^*	99	102	102	100	101	101	99	101	101	100	99
k_0^∞	72	82	83	86	90	91	91	93	93	95	95

4 Заключение

Снижение временной сложности базовых алгоритмов обработки данных расширяет возможности их применения, позволяет получать новые результаты. Например, становится реальным опробование большего числа вариантов при выборе структурных параметров модели или увеличение «глубины» при перепроверке полученных результатов.

Для вычисления плотности распределения в случае PPCA исследовались вопросы применения различных разложений матриц и обращения обновления низкого ранга некоторой матрицы. За счет комбинирования этих подходов, а также благодаря «маленьким хитростям» удалось снизить временные вычислительные затраты по крайней мере в 2 раза. Попутно были очерчены области, где для задач с наличием сниженной размерности имеет смысл использовать этот факт, а где нет.

Полученные результаты обладают потенциалом улучшения при уточнении спектра элементарных вычислительных операций, а также применимы в подобных ситуациях (см., например, [5]).

Литература

1. *Кривенко М. П.* Выбор размерностей для смеси вероятностных анализаторов главных компонент // Системы и средства информатики, 2019. Т. 29. № 3. С. 4–15.
2. *Tipping M. E., Bishop C. M.* Mixtures of probabilistic principal component analyzers // Neural Comput., 1999. Vol. 11. Iss. 2. P. 443–482.
3. *Seber G. A. F., Lee A. J.* Linear regression analysis. — Hoboken, NJ, USA: Wiley, 2003. 557 p.
4. *Golub G. H., Van Loan C. F.* Matrix computations. — 4th ed. — Baltimore, MD, USA: The Johns Hopkins University Press, 2013. 780 p.
5. *Yiping Guo Y., Bondell H. D.* On robust probabilistic principal component analysis using multivariate t-distributions // arXiv.org, 2020. arXiv:2010.10786 [stat.ME].

Поступила в редакцию 09.03.21

COMPUTING BASED ON PROBABILISTIC PRINCIPAL COMPONENT ANALYSIS MODEL

M. P. Krivenko

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

Abstract: An effective solution to problems of multivariate data analysis requires the use of complex probabilistic models, among which the probabilistic model of principal component analysis (PPCA) occupies a worthy place. It allows one to adequately describe real data, formally formulate and solve the problem

of choosing its parameters. Practical application of PPCA is associated with a large number of laborious computations. The article discusses techniques for significantly reducing the time spent in calculating the density of the multivariate normal distribution. For this, the expediency of using the Cholesky expansion for the covariance matrix and Woodbury identity for its PPCA factorization is substantiated. A brief description of the experiments is given, which makes it possible to evaluate the real time characteristics of the algorithms and to reveal the conditions for their effective application. Along the way, recommendations are given on the implementation of individual operations.

Keywords: PPCA computing; Cholesky factorization; Woodbury matrix identity; computing experience

DOI: 10.14357/08696527210306

References

1. Krivenko, M. P. 2019. Vybor 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.
2. Tipping, M. E., and C. M. Bishop. 1999. Mixtures of probabilistic principal component analyzers. *Neural Comput.* 11(2):443–482.
3. Seber, G. A. F., and A. J. Lee. 2003. *Linear regression analysis*. Hoboken, NJ: Wiley. 557 p.
4. Golub, G. H., and C. F. Van Loan. 2013. *Matrix computations*. 4th ed. Baltimore, MD: The Johns Hopkins University Press. 780 p.
5. Yiping, G. Y., and H. D. Bondell. 2010. On robust probabilistic principal component analysis using multivariate t-distributions. Available at: <https://arxiv.org/abs/2010.10786> (accessed August 26, 2021).

Received March 9, 2021

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

ИСПОЛЬЗОВАНИЕ ПАДЕЖНОЙ ГРАММАТИКИ ПРИ ИНФОРМАЦИОННОМ ПОИСКЕ В БАЗЕ ЗНАНИЙ ЭКСПЕРТНОЙ СИСТЕМЫ О КОНСТРУКЦИЯХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ

Н. И. Сидняев¹, Ю. И. Бутенко², Е. Е. Синева³

Аннотация: Раскрыты основные принципы построения экспертных систем в области авиационной промышленности. Предложен логико-лингвистический подход к информационному моделированию предметной области при проектировании базы знаний экспертной системы. Разработана модель представления знаний для описания конструкций летательных аппаратов, обеспечивающая информационно-структурную надежность базы знаний экспертной системы. Проведен анализ парадигматических особенностей авиакосмической терминологии с целью улучшения эффективности информационного поиска. Предложен метод разрешения лексической многозначности поискового запроса в базе знаний. Описан механизм вывода информации из базы знаний на основе падежной грамматики Ч. Филлмора, позволяющей определить семантико-синтаксическую структуру выводимого предложения. Показано, что посредством расстановки ограничений на число участников ситуации, которые может присоединять глагол, можно обеспечить эффективный вывод информации из базы знаний.

Ключевые слова: экспертная система; база знаний; летательный аппарат; терминология; знания

DOI: 10.14357/08696527210307

1 Введение

Поддержка принятия решений в сложных ситуациях необходима в различных областях деятельности, где требуются обоснованные, логически доказуемые аргументы, в том числе в оперативном и стратегическом управлении технологиями на авиационных предприятиях. При принятии решений необходимо проведение детального моделирования последствий предполагаемого решения, поиск оптимального пути достижения заданного результата при помощи технологий имитационного моделирования. С подобными задачами успешно справляются

¹Московский государственный технический университет им. Н. Э. Баумана, sidnyaev@yandex.ru

²Московский государственный технический университет им. Н. Э. Баумана, iuliiabutenko2015@yandex.ru

³Московский государственный технический университет им. Н. Э. Баумана, bolotovaee@mail.ru

экспертные системы — комплексы программных средств, способные заменять специалиста при решении сложных задач, возникающих в процессе диагностики, проектирования или эксплуатации аэрокосмической и другой техники [1]. Такие системы основаны на знаниях, полученных в процессе взаимодействия с экспертами в конкретной предметной области.

Параллельно с развитием различных направлений исследований в области искусственного интеллекта идет развитие информационных структур для представления знаний [2, 3]. Появились новые способы описания и представления данных, возникли фреймовые, списочные и иерархические структуры. Процесс построения баз знаний достаточно сложен и носит итеративный характер, заключающийся в циклической модификации баз знаний на основе результатов ее тестирования [4].

2 Абстрактная иерархическая структура как способ описания конструкций летательных аппаратов

Основной задачей проектирования базы знаний в области авиакосмического приборостроения является описание структуры летательного аппарата таким образом, чтобы оно предоставляло наиболее полную и непротиворечивую информацию об описываемом объекте. Если применить лингвистические модели в совокупности с математическим аппаратом, можно достичь более ясного и результативного описания и решения сложной технологической проблемы. Информация воспринимается специалистом на когнитивном уровне посредством основных единиц этого уровня — понятий, идей, концептов, поэтому для представления знаний часто применяется аппарат семантических сетей, который позволяет воссоздать модель, схожую с процессами мышления конструктора, технолога [5, 6]. В инженерии знаний под семантической сетью подразумевается граф с узлами, отображающими объекты предметной области, и дугами, обозначающими отношения между данными понятиями.

Для описания конструкций летательных аппаратов разработана интегрированная модель представления знаний — абстрактная иерархическая структура, объединяющая графовую структуру, фреймы и предикаты. Все знания о предметной области представлены в виде графа, в узлах которого содержатся фреймы — минимально возможные описания сущностей, записанные с помощью предикатов и предикатных символов [1]. Подобная структура гарантирует непротиворечивость описаний объектов и их отношений в реальном производстве.

3 Отражение парадигматических особенностей авиакосмической терминологии в базе знаний

Неотъемлемой частью любого управления является его документационное обеспечение, основа которого — однозначно понимаемая и непротиворечивая терминология. Единое терминологическое обеспечение позволяет избегать недо-

понимания между специалистами как в рамках одной предметной области, так и за ее пределами.

Терминосистема — это некоторая лингвистическая модель, представляющая конкретную предметную область [7]. Аэрокосмическая терминология русского языка представляет собой сложную систему, включающую термины из различных тематических пространств: обозначения субъектов аэрокосмической сферы (*космонавт, летчик, пилот*), наименования организаций и структур области авиации и космонавтики (*НАСА, ИАТА, ИФАТКА*), обозначения действий аэрокосмической деятельности (*взлет, посадка, пилотирование, пристыковка*), наименования деталей и частей летательных аппаратов (*фюзеляж, крыло, подкрылки, двигатель*), обозначение характеристик и состояний летательных аппаратов (*работоспособность, надежность, безопасность*). Стандартизация терминологии данной предметной области — процесс трудоемкий, а отсутствие стандартизации приводит к появлению разных терминов для обозначения одних и тех же объектов и явлений, в связи с чем возникают случаи синонимии и лексической многозначности.

При построении баз знаний интеллектуальных систем синонимия считается крайне неблагоприятным явлением. Различия в номинации объектов приводят к трудностям в сфере профессиональной коммуникации, поэтому базу знаний экспертной системы предлагается пополнять всеми возможными вариантами синонимов тех или иных узкоспециальных и общенаучных терминов.

В качестве иллюстрации приведем несколько примеров синонимичных терминов: *интерцептор–спойлер*; *пусковое устройство–пусковая кнопка–стартер*; *шкала дальности–дистанционная шкала*; *спуск–снижение*; *приборная доска–приборная панель*; *маслоуловитель–маслосборник*.

В аэрокосмической терминологии также наблюдается семантическая оппозиция номинаций «свой–чужой», в результате чего возникают синонимичные пары, например: *космонавт–астронавт* (англ. astronaut); *космонавтика–астронавтика* (англ. astronautica); *высотомер–альтиметр* (франц. altimètre).

Равнозначность подобных синонимов необходимо отразить в проектируемой базе знаний во избежание недопонимания между экспертной системой и ее пользователями. Стоит отметить, что редко используемые аббревиатуры должны быть расшифрованы и понятны пользователям экспертной системы (например, *ШВП* — это *шасси на воздушной подушке* или *шаровой вытяжной парашют*).

Значительные трудности в процессе поиска информации в базе знаний вызывают явления лексической многозначности, т. е. сосуществование различных значений одного слова или фразы, например:

- Свеча* — 1. Приспособление для воспламенения горючей смеси (в двигателях внутреннего сгорания); 2. Крутой взлет, подъем (фигура высшего пилотажа).
Хвост — 1. Хвостовое оперение летательных аппаратов; 2. Вытянутое из пыли и газа кометное вещество, образующееся при приближении кометы к Солнцу.
Атмосфера — 1. Газообразная оболочка, окружающая Землю и некоторые другие планеты; 2. Единица измерения давления.

4 Применение падежной грамматики Ч. Филлмора для вывода знаний из экспертной системы

Один из необходимых компонентов базы знаний экспертной системы — механизм вывода, который содержит правила для решения конкретной задачи. Механизм вывода ссылается на информацию из базы знаний и выбирает факты и правила, которые будут применяться при попытке ответить на запрос пользователя. Данный механизм обеспечивает аргументацию информации в базе знаний и позволяет человеку получать информацию из базы знаний в виде предложений, построенных на естественном языке. Одним из наиболее удачных механизмов, позволяющих моделировать смысл высказываний на естественном языке, считается механизм, основанный на теории падежных грамматик Ч. Филлмора [8]. В падежной грамматике глагол является центром предложения и диктует в силу своего значения набор ролей (глубинных падежей), реализующихся в предложении посредством именных форм. Иными словами, падежная грамматика показывает связь существительного или местоимения с другими словами в предложении. Обрамление падежами подвержено определенным ограничениям, например каждый семантический падеж может встречаться в предложении только один раз [9]. Некоторые падежи являются обязательными, другие — необязательными. Обязательные падежи нельзя удалять, рискуя получить грамматически неправильные предложения.

Для описания события в каждом фрейме иерархической структуры в первую очередь выделяется действие, которое обычно описывается глаголом. Далее определяются: объект, который действует — агенс; объект, над которым это действие выполняется — пациенс [10].

На каждое отношение накладывается множество ограничений, например структура глагола «пилотировать» может включать следующие семантические падежи: (1) агентив — *летчик пилотировал самолет*; (2) пациентив — *летчик пилотировал **самолет***; (3) инструменталис — *летчик пилотировал самолет **боковой ручкой управления***; (4) локатив — *летчик впервые пилотировал самолет **на аэродроме** с плохим покрытием*. Возможно включение в структуру данного глагола и других семантических падежей, а также атрибутивных отношений, например: «*пилотировать каким образом?*» — «*умело*», «*профессионально*».

Такие ограничения необходимо накладывать для того, чтобы система могла строить правильные семантико-синтаксические структуры, что обеспечит связность и внутреннюю интерпретируемость знаний в экспертной системе.

Падежные отношения, а также соответствие между синтагматикой и линейной упорядоченностью слов в предложении можно представить в виде дерева непосредственно составляющих. Глубинная структура пропозиционального компонента любого простого предложения представляет собой конструкцию, состоящую из глагольной и именной группы, которые находятся в специальных помеченных отношениях (семантических падежах) ко всему предложению. Дере-

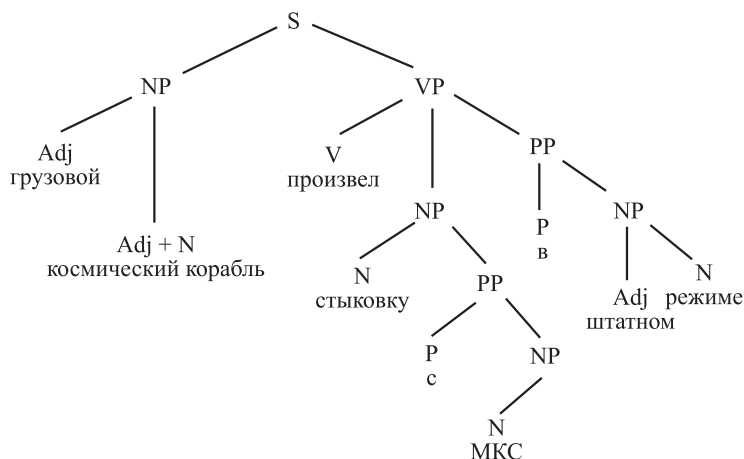


Рис. 1 Представление предложения в виде дерева составляющих

во составляющих позволяет представить синтаксическую структуру предложения в виде иерархии, составленной из связанных друг с другом непересекающихся единиц. Пример описания предложения «Грузовой космический корабль произвел стыковку с МКС в штатном режиме» с помощью дерева составляющих показан на рис. 1.

Исходя из этого, модель семантической целостности предложения (C) можно представить следующим образом:

$$C = \langle V^I, N^I \rangle,$$

где V — глагольная составляющая; N — именная составляющая; I — семантический падеж. Для вывода информации из базы знаний требуются также атрибутивные отношения для описания формы, цвета, размера и других характеристик объектов, а также дополнительные падежи цели, условия, времени и др., т. е. в зависимости от целей экспертной системы перечень, содержащий 6 основных падежей, может быть расширен.

Добавим к структуре рассматриваемого предложения дополнительные семантические падежи и применим теорию семантических падежей для описания семантической структуры знания о событии. Допустим, системе нужно вывести сообщение «Грузовой космический корабль пристыковался к МКС в 15:00 мск в штатном режиме. Груз доставлен на станцию». Семантическая сеть, описывающая событие доставки груза на МКС, представлена на рис. 2.

В процессе поиска информации в базе знаний, основанной на семантической сети, происходит сопоставление общей сети с сетью запроса для поиска нужного фрагмента информации. При выводе в семантических сетях также применяется метод перекрестного поиска, в процессе которого сопоставляются не узлы семантической сети, а дуги — отношения между объектами.



Рис. 2 Пример семантической сети, описывающей событие доставки груза на МКС

5 Заключение

В работе раскрыты основные принципы построения интеллектуальных систем в авиакосмической промышленности. Поднимается проблема стандартизации терминологии как одного из ключевых этапов разработки баз знаний экспертных систем с целью сведения разнообразно представленной информации к общепонятному виду, что позволит организовать ее совместное и многократное использование специалистами. Для выявления парадигматических особенностей терминологии авиакосмической области проведен анализ лексических единиц, входящих в тематическую группу «Авиакосмическое приборостроение». В связи с тем что стандартизация терминологии является трудоемким процессом, предложено пополнять проектируемую базу знаний всеми возможными вариантами синонимичных терминов. Рассмотрен механизм вывода информации из базы знаний экспертной системы на основе падежной грамматики Ч. Филлмора, которая позволяет определить семантико-синтаксическую структуру выводимого предложения. Показано, что посредством расстановки ограничений на число актантов (участников ситуации), которые может присоединять глагол, можно обеспечить эффективный вывод предложений на естественном языке.

Литература

1. Сидняев Н. И., Бутенко Ю. И., Болотова Е. Е. Экспертная система производственного типа для создания базы знаний о конструкциях летательных аппаратов // Авиакосмическое приборостроение, 2019. № 6. С. 38–52. doi: 10.25791/aviakosmos.06.2019.676.
2. Сигов А. С., Нечаев В. В., Кошкарёв М. И. Архитектура предметно-ориентированной базы знаний интеллектуальной системы // Int. J. Open Information Technologies, 2014. № 12. Р. 1–6.
3. Сидняев Н. И., Бутенко Ю. И., Болотова Е. Е. Логическая модель требований информационно-системной надежности для баз знаний интеллектуальных систем // Программная инженерия, 2020. № 4. С. 195–204. doi: 10.17587/prin.11.195-204.

4. *Helbig H.* Knowledge representation and the semantics of natural language. — Berlin/Heidelberg/New York: Springer, 2006. 655 p.
5. *Woods D. D.* Cognitive technologies: The design of joint human-machine cognitive systems // *AI Mag.*, 1985. Vol. 6. Iss. 4. P. 86–92.
6. *Ломов П. А., Шишаев М. Г.* Формирование когнитивных фреймов на основе онтологических паттернов для визуализации онтологий // *Информационные системы и технологии*, 2015. № 6(92). С. 12–22.
7. *Даниленко В. П.* Русская терминология: опыт лингвистического описания. — М.: Наука, 1977. 246 с.
8. *Fillmore Ch.* The case for case // *Universals in linguistic theory* / Eds. E. Bach, R. T. Harms. — New York, NY, USA: Holt, Rinehart, and Winston, 1968. P. 1–88.
9. *Богданов В. В.* Структурно-семантическая организация предложения. — Л.: ЛГУ, 1977. 205 с.
10. *Jackendoff R. S.* The status of thematic relations in linguistic theory // *Linguistic Inq.*, 1987. Vol. 16. P. 369–411.

Поступила в редакцию 18.08.21

USE OF CASE GRAMMAR IN INFORMATION SEARCH IN THE EXPERT SYSTEM KNOWLEDGE BASE ON AIRCRAFT STRUCTURES

N. I. Sidnyaev, Yu. I. Butenko, and E. E. Sineva

Bauman Moscow State Technical University, 5-1 2nd Baumanskaya Str., Moscow 105005, Russian Federation

Abstract: The paper reveals the basic principles of building expert systems in the aviation industry. A model of knowledge representation that provides information and structural reliability of the expert system's knowledge base for describing aircraft structures has been developed. The article analyzes the paradigmatic features of aerospace terminology in order to improve the efficiency of information search. A method for resolving lexical ambiguity of a search query in the knowledge base is proposed. Inference engine and explanation mechanism of the knowledge base premised on the Charles Fillmore's case grammar is described. The grammatical cases make it possible to determine the semantic and syntactic structure of the output sentence. Restrictions show the number of participants in the situation that a verb can attach, it is possible to provide effective output of information from the knowledge base.

Keywords: expert system; knowledge base; aircraft; terminology; knowledge

DOI: 10.14357/08696527210307

References

1. Sidnyaev, N. I., Yu. I. Butenko, and E. E. Bolotova. 2019. Ekspertnaya sistema produktsionnogo tipa dlya sozdaniya bazy znaniy o konstruktivnykh letatel'nykh apparatov [Rule-based expert system for creating a knowledge base on aircraft structures]. *Aviakosmicheskoe priborostroenie* [Aerospace Instrument-Making] 6:38–52. doi: 10.25791/aviakosmos.06.2019.676.
2. Sigov, A. S., V. V. Nechaev, and M. I. Koshkaryov. 2014. Arkhitektura predmetno-oriyentirovannoy bazy znaniy intellektual'noy sistemy [Architecture of the subject-oriented knowledge base of the intellectual system]. *Int. J. Open Information Technologies* 12:1–6.
3. Sidnyaev, N. I., Yu. I. Butenko, and E. E. Bolotova. 2020. Logicheskaya model' trebovaniy informatsionno-sistemnoy nadezhnosti dlya baz znaniy intellektual'nykh sistem [Logical model of information and system reliability requirements for knowledge bases of intellectual systems]. *Programmnaya inzheneriya* [Software Engineering] 4:195–204. doi: 10.17587/prin.11.195-204.
4. Helbig, H. 2006. *Knowledge representation and the semantics of natural language*. Berlin/Heidelberg/New York: Springer. 655 p.
5. Woods, D. D. 1985. Cognitive technologies: The design of joint human-machine cognitive systems. *AI Mag.* 6(4):86–92.
6. Lomov, P. A., and M. G. Shishaev. 2015. Formirovanie kognitivnykh freymov na osnove ontologicheskikh patternov dlya vizualizatsii ontologii [Approach to ontology visualization based cognitive frames]. *Informatsionnye sistemy i tekhnologii* [Information Systems and Technologies] 6(92):12–22.
7. Danilenko, V. P. 1977. *Russkaya terminologiya: opyt lingvisticheskogo opisaniya* [Russian terminology: Experience of linguistic description]. Moscow: Nauka. 246 p.
8. Fillmore, C. J. 1968. The case for case. *Universals in linguistic theory*. Eds. E. Bach and R. T. Harms. New York, NY: Holt, Rinehart, and Winston. 1–88.
9. Bogdanov, V. V. 1977. *Strukturno-semanticheskaya organizatsiya predlozheniya* [Structural and semantic organization of a sentence]. Leningrad: LSU. 205 p.
10. Jackendoff, R. S. 1987. The status of thematic relations in linguistic theory. *Linguistic Inq.* 16:369–411.

Received August 18, 2021

Contributors

Sidnyaev Nikolay I. (b. 1955) — Doctor of Science in technology, professor, Head of Department, Bauman Moscow State Technical University, 5-1 2nd Baumanskaya Str., Moscow 105005, Russian Federation; sid-nyaev@yandex.ru

Butenko Yuliya I. (b. 1987) — Candidate of Science (PhD) in technology, associate professor, Bauman Moscow State Technical University, 5-1 2nd Baumanskaya Str., Moscow 105005, Russian Federation; iuliiabutenko2015@yandex.ru

Sineva Elizaveta E. (b. 1998) — Master student, Faculty of Fundamental Sciences, Bauman Moscow State Technical University, 5-1 2nd Baumanskaya Str., Moscow 105005, Russian Federation; bolotovae@mail.ru

МЕТОДИКА ОЦЕНКИ ПРОИЗВОДСТВЕННЫХ РИСКОВ РАЗРАБОТКИ СРЕДСТВ ВООРУЖЕНИЯ И ВОЕННОЙ ТЕХНИКИ

А. В. Босов¹, А. А. Крюков²

Аннотация: Представлена методика оценки производственных рисков на этапе опытно-конструкторских работ (ОКР) при разработке высокотехнологичной продукции. Для учета сложной классификации производственных рисков, различных источников их возникновения и сфер влияния на результат ОКР предлагается внедрить комбинированный методический подход и оценивать производственные риски частично статистическими методами, частично методом экспертных оценок. Рассмотрено содержание этапов соответствующей методики. Обоснован метод расчета интегральной оценки производственного риска и шкала уровней производственных рисков проекта согласно характеру влияния и приемлемости.

Ключевые слова: методика оценки производственных рисков; опытно-конструкторские работы; метод экспертных оценок; статистические методы

DOI: 10.14357/08696527210308

1 Введение

В существующих условиях геополитической нестабильности и ужесточения глобальной конкуренции особую актуальность приобретает оценка рисков высокотехнологичных производств оборонно-промышленного комплекса (ОПК), формирующих основу национальной безопасности. Система оценки производственных рисков предприятий ОПК при выполнении ОКР обеспечивает своевременное обнаружение, анализ, оценку и управление риском на всех стадиях ОКР и служит одним из основных механизмов обеспечения выполнения договорных обязательств. Особую значимость такая система приобретает при выполнении ОКР в рамках государственного оборонного заказа (ГОЗ). В литературе представлены разнообразные исследования методических подходов к оценке рисков. Так, в [1–3] авторы рассматривают понятие и оценку рисков в целом, не затрагивая проблематику ОПК и специфику производственных рисков. Отдельные аспекты оценки и управления рисками предприятий ОПК, включая производственные риски, раскрыты в работах [4, 5] и др. При этом до настоящего времени

¹Федеральный исследовательский центр «Информатика и управление» Российской академии наук, AVBosov@ipiran.ru

²АО «Концерн воздушно-космической обороны «Алмаз-Антей», alexeykryukov@almaz-antey.ru

остаются открытыми вопросы комплексной системной оценки производственных рисков выполнения ОКР на предприятиях ОПК при выполнении ГОЗ.

2 Постановка задачи

Цель представляемого исследования — разработка методических подходов к оценке производственных рисков предприятий ОПК на этапе проведения ОКР в рамках выполнения ГОЗ при создании средств вооружения.

Для оценки рисков разработки средств вооружения в ходе выполнения ОКР целесообразно применять комплексный подход из нескольких методов оценки рисков. Учитывая сложную классификацию производственных рисков, различные источники их возникновения и сферы влияния на результат ОКР, предлагается комбинированный методический подход, реализующий оценку производственных рисков частично статистическими методами, частично методом экспертных оценок. В значительной степени представленный результат основан на использовании накопленной в Концерне ВКО «Алмаз-Антей», заинтересованном в практическом использовании предлагаемой методики, информационно-статистической базе для оценки индикаторов рисков.

В следующем разделе методика представлена детальным описанием последовательных этапов оценки производственных рисков.

3 Методика

I этап. Систематизация производственных рисков опытно-конструкторских работ

Риски, негативно влияющие на процесс выполнения и успешного достижения целей ОКР, оцениваются с помощью соответствующих индикаторов. Значение каждого из индикаторов устанавливается или в соответствии с накопленной на предприятии информационно-аналитической базой (при ее наличии), или оценивается экспертами при невозможности количественно измерить рассматриваемый риск в силу его сложности и/или отсутствия информационно-аналитической базы [1]. По результатам сбора индикаторов оценки производственных рисков ОКР формируется их перечень, под которым понимается обоснованная и ограниченная совокупность идентифицированных и оцениваемых рисков, возникающих при разработке средств вооружения. Производственные риски, наиболее значимые в рамках конкретной работы и выявленные компетентными экспертами, подкрепляются соответствующими измеримыми индикаторами оценки [3].

II этап. Группировка производственных рисков опытно-конструкторских работ

Результаты систематизации и формирования перечня рисков после группировки на II этапе заносятся в реестр рисков (табл. 1). Методика построения

Таблица 1 Реестр производственных рисков ОКР

Группа рисков	Классификационная группа	Индикаторы оценки риска		Риски проекта	
		Индикатор	Значение	Шифр риска	Риск
Риски, оцениваемые статистическими методами, R_1	Технологические риски R_{11}	...	...	$R_{11}(1)$... $R_{11}(n_1)$	...
	Риски бесперебойности и безотказности технологических процессов R_{12}	...	...	$R_{12}(1)$... $R_{12}(n_2)$	...
Риски, оцениваемые методом экспертных оценок, R_2	Технологические риски R_{21}	...	...	$R_{21}(1)$... $R_{21}(n_3)$	...
	Риски бесперебойности и безотказности технологических процессов R_{22}	...	...	$R_{22}(1)$... $R_{22}(n_4)$	...

реестра производственных рисков для ОКР позволяет выявить важные и значимые риски, создающие угрозу на пути к достижению целей конкретного проекта, а также ведущих к нарушению деятельности предприятия в целом. На этом этапе необходимо обосновать метод оценки риска в зависимости от наличия или отсутствия информационно-аналитической базы оценки риска. Производственные риски ОКР в рамках каждой классификационной группы относятся к группе R_1 рисков, оцениваемых статистическими методами, или к группе R_2 рисков, оцениваемых методами экспертных оценок. Результаты заносятся в табл. 1, каждому риску присваивается шифр, при этом учитывается:

- метод оценки риска (статистический или экспертный, в зависимости от наличия или отсутствия информационно-аналитической базы);
- классификационная группа, к которой отнесен риск;
- взаимосвязь значения оценочного индикатора и риска, который измеряется данным индикатором [6].

Шифр риска определяется принадлежностью к группе рисков в зависимости от метода оценки (R_1 или R_2), принадлежностью к классификационной группе рисков (технологические или бесперебойности и безотказности), порядковым номером (соответствует индикатору).

В общем случае методика может использоваться для произвольного числа K классификационных групп и общего числа рисков во всех группах $N = \sum_{k=1}^K n_k$. В табл. 1 $K = 4$, $N = n_1 + n_2 + n_3 + n_4$.

III этап. Расчет статистических коэффициентов оценки производственных рисков опытно-конструкторских работ

На данном этапе собирается информационно-аналитическая база данных по индикаторам оценки рисков из перечня, для которых она может быть сформирована. Поскольку закон распределения оценочного индикатора неизвестен, то на основе этих данных риск рассчитывается по статистическим коэффициентам: вычисляются выборочные значения среднего, дисперсии, среднеквадратичного отклонения, коэффициента вариации, по которым оценивается риск.

Выберем классификационную группу рисков, оцениваемых статистическими методами, и обозначим i -й индикатор x_i , $i = 1, \dots, n$. В табл. 1 могут выбирать группы R_{11} и R_{12} , соответственно $n = n_1$ или n_2 . Выберем из имеющейся информационно-аналитической базы данных все величины x_i^m отклонений фактического значения оценочного индикатора x_i от нормативного по модулю за

Таблица 2 Значения коэффициента вариации и их интерпретация в статистической оценке производственных рисков ОКР

v_i	$\hat{x}_i$	Интерпретация (характер риска)	Приемлемость риска	Стратегия управления производственным риском проекта
< 10%	0–1	Степень риска несущественна; вложения крайне желательны	Риск приемлем. Вмешательство не требуется	Сохранение уровня производственного риска проекта
10%–20%	1–2	Средняя степень риска	Риск приемлем. Вмешательство не требуется	Сохранение уровня производственного риска проекта
20%–30%	2–3	Высокая степень риска	Риск приемлем. Возможно вмешательство	Разработка мероприятий по снижению уровня возможных потерь от производственного риска проекта
30%–40%	3–4	Модель крайне неустойчива, дальнейшее продолжение ОКР необоснованно	Риск неприемлем. Необходим дополнительный анализ рискообразующих факторов	Разработка мероприятий по снижению уровня возможных потерь от производственного риска проекта
> 40%	4–5	Модель крайне неустойчива, дальнейшее продолжение ОКР необоснованно	Риск неприемлем. Необходим дополнительный анализ рискообразующих факторов	Страхование (передача) производственного риска проекта

период анализа и обозначим $\bar{x}_i = (1/M) \sum_{m=1}^M x_i^m$ среднее значение, где M — общий период анализа, за который в организации накоплены данные о фактических и нормативных значениях оценочных индикаторов производственных рисков ОКР.

Для характеристики фактических отклонений от среднего значения вычислим среднеквадратичное отклонение $\delta_i = \sqrt{(1/M) \sum_{m=1}^M (x_i^m - \bar{x}_i)^2}$ и статистический коэффициент вариации $v_i = \delta_i / \bar{x}_i$. Последний коэффициент помогает определить оптимальное управленческое решение из двух альтернативных вариантов, когда следует выбрать минимизацию производственного риска или анализируемый параметр (ущерб, потери, недостижение целей и задач проекта). На его основе определим статистическую оценку $\hat{x}_i$ индикатора x_i , исходя из данных табл. 2, при этом статистическая оценка производственного риска соответствует условию $0 \leq \hat{x}_i \leq 5$.

Отметим, что нормативного значения для статистической оценки производственного риска ОКР не существует. Однако имеются некоторые опорные критерии, применяемые в инвестиционном анализе, которые помогают при его анализе и интерпретации [6].

IV этап. Ранжирование производственных рисков опытно-конструкторских работ

Данный этап описывает порядок работы с производственными рисками, оцениваемыми экспертными методами (в табл. 1 группы R_{21} и R_{22}). Часть рисков, по которым отсутствует информационно-аналитическая база для оценочных индикаторов и которые нельзя оценить статистическими методами, составляет сложную, субъективную и недетерминированную часть ОКР, и для них нужно применять метод экспертных оценок [7]. Реализация метода предполагает привлечение компетентных экспертов в количестве не менее 5 специалистов для каждой классификационной группы производственных рисков и формирование чек-листа для каждого эксперта, в котором будет содержаться:

- перечень рисков по каждой классификационной группе;
- ранг риска, отражающий важность риска по возрастанию [8] (значения от 1 до n , в табл. 1 $n = n_3$ или $n = n_4$);
- балльная оценка эксперта в соответствии с алгоритмом (табл. 3), который предполагает определение зависимости от уровня влияния риска на достижение задач ОКР и вероятности возникновения или вероятного отклонения от планируемых результатов реализации ОКР [2].

Предполагается, что ранжирование проводится J экспертами, так что для i -го индикатора x_j определяются ранги f_{ij} , $j = 1, \dots, J$, полученные от всех экспертов, которые затем суммируются, и устанавливается результирующий

Таблица 3 Шкала балльной оценки производственных рисков ОКР

Оценка	Качественные характеристики производственных рисков ОКР	
	Уровень влияния	Вероятность наступления
5–4	Высокий	Высокая
3–4	Высокий	Низкая
2–3	Низкий	Высокая
1–2	Низкий	Низкая

ранг для риска $f_i = (1/J) \sum_{j=1}^J f_{ij}$. В классификационных группах R_2 риски упорядочиваются так, чтобы $f_i < f_{i+1}$.

V этап. Определение весовых коэффициентов (коэффициентов значимости) производственных рисков опытно-конструкторских работ

В каждой квалификационной группе рассчитываются весовые коэффициенты ω_i , $i = 1, \dots, n$, которые отвечают условиям $0 \leq \omega_i \leq 1$, $\sum_{i=1}^n \omega_i = 1$, $n = n_k$, для рисков в группах R_2 . Эти коэффициенты отражают вклад каждого риска в общий риск реализации ОКР. Весовые коэффициенты рассчитываются на основании ранга производственного риска, определенного на предыдущем этапе методики следующим образом:

$$\omega_i = \frac{2(n - f_i + 1)}{n(n + 1)}.$$

VI этап. Оценка производственных рисков опытно-конструкторских работ

Производственный риск, оцениваемый экспертами, является неотъемлемой составляющей ОКР и характеризует их реализуемость. Оценка риска компетентными экспертами в соответствии с разработанной методикой на данном этапе будет осуществляться на основе уровня влияния и вероятности возникновения риска [9]. В зависимости от оценки этих параметров формируется оценка риска по балльной шкале, представленной в табл. 3.

Экспертные оценки производственного риска, данные J экспертами для i -го индикатора риска в каждой квалификационной группе, обозначаются $\hat{x}_{ij}$, $i = 1, \dots, n$, $j = 1, \dots, J$.

VII этап. Оценка степени согласованности экспертных оценок в рамках классификационной группы рисков

С целью проверки объективности результатов экспертной оценки производственных рисков ОКР предлагается определять степень согласованности мнений

экспертов в рамках каждой классификационной группы рисков. Эта величина характеризуется коэффициентом конкордации Кендалла $0 \leq W \leq 1$ (коэффициентом согласия) [5]:

$$W = \frac{12 \sum_{i=1}^n \left(\sum_{j=1}^J \hat{x}_{ij} - J(n+1)/2 \right)^2}{J^2(n^3 - n)}.$$

Этот коэффициент при $W = 1$ означает, что эксперты пришли к единогласному мнению и присвоили рискам одинаковые или близкие по значению ранги; чем ближе значение W к нулю, тем менее согласованными выглядят оценки экспертов. Степень согласованности считается удовлетворительной при $W > 0,5$ и хорошей при $W > 0,7$.

VIII этап. Расчет средней взвешенной оценки производственных рисков опытно-конструкторских работ в классификационных группах

Результаты экспертных оценок производственных рисков ОКР в рамках выделенных классификационных групп обобщаются в виде средневзвешенной оценки каждого отдельного риска: $\hat{x}_i = (\omega_j/J) \sum_{j=1}^J \hat{x}_{ij}$, а затем консолидируются в табл. 4.

Таблица 4 Результаты экспертной оценки рисков

Классификационные группы производственных рисков ОКР	Шифр риска	Весовой коэффициент риска	Ранг риска	Эксперт 1	...	Эксперт J	Средняя взвешенная экспертная оценка риска
				Оценка риска		Оценка риска	
R_{21}	$R_{21}(1)$	ω_1	f_1	$\hat{x}_{11}$	...	$\hat{x}_{1J}$	$\hat{x}_1$
	...	...	...	...	...	...	...
	$R_{21}(n_3)$	ω_{n_3}	f_{n_3}	$\hat{x}_{n_3 1}$	...	$\hat{x}_{n_3 J}$	$\hat{x}_{n_3}$
W							$W(R_{21})$
R_{22}	$R_{22}(1)$	ω_1	f_1	$\hat{x}_{11}$	...	$\hat{x}_{1J}$	$\hat{x}_1$
	...	...	...	...	...	...	...
	$R_{22}(n_4)$	ω_{n_4}	f_{n_4}	$\hat{x}_{n_4 1}$	...	$\hat{x}_{n_4 J}$	$\hat{x}_{n_4}$
W							$W(R_{22})$

IX этап. Интегральная оценка производственного риска опытно-конструкторских работ

На данном этапе методики вычисляется интегральная оценка производственных рисков по каждой классификационной группе. Для этого средневзвешенные

Таблица 5 Интегральная оценка классификационной группы рисков

Группа рисков	Шифр риска	Статистическая оценка	Интегральная оценка классификационной группы
R_{11}	$R_{11}(1)$... $R_{11}(n_1)$	$\hat{x}_1$... $\hat{x}_{n_1}$	Y_{11}
R_{12}	$R_{12}(1)$... $R_{12}(n_2)$	$\hat{x}_1$... $\hat{x}_{n_2}$	Y_{12}
R_{21}	$R_{21}(1)$... $R_{21}(n_3)$	$\hat{x}_1$... $\hat{x}_{n_3}$	Y_{21}
R_{22}	$R_{22}(1)$... $R_{22}(n_4)$	$\hat{x}_1$... $\hat{x}_{n_4}$	Y_{22}

балльные оценки i -го риска $\hat{x}_i$ в классификационной группе сводятся в табл. 5. В эту же таблицу по рискам, оцененным статистическими методами, заносятся статистические оценки.

Интегральная оценка производственных рисков ОКР по классификационным группам рисков рассчитывается как средневзвешенная величина с использованием весовых коэффициентов, определенных V этапом методики, т. е. для рисков, оцениваемых статистически, $Y_{R_{1k}} = \sum_{i=1}^n \omega_i \hat{x}_i$, а для рисков, оцененных экспертами, $Y_{R_{2k}} = \sum_{i=1}^n \hat{x}_i$.

X этап. Расчет ранга и весового коэффициента классификационной группы производственных рисков опытно-конструкторских работ

Основанием для определения ранга рисков классификационной группы служит его интегральная оценка, рассчитанная на предыдущем этапе методики. Весовой коэффициент классификационной группы рисков ОКР определяется аналогично весу ω_i каждого отдельного риска внутри классификационной группы. Рассчитанный для группы R коэффициент обозначается φ_R , а соответствующие ранги групп — g_R .

XI этап. Интегральная оценка производственных рисков опытно-конструкторских работ

Расчет интегральной оценки производственных рисков ОКР проводится на основе определенных весовых коэффициентов классификационных групп рисков

Таблица 6 Интегральная оценка производственных рисков ОКР

Классификационная группа	Ранг классификационной группы	Вес классификационной группы	Интегральная оценка классификационной группы	Интегральная оценка производственных рисков проекта
R_{11} R_{12} R_{21} R_{22}	$g_{R_{11}}$ $g_{R_{12}}$ $g_{R_{21}}$ $g_{R_{22}}$	$\varphi_{R_{11}}$ $\varphi_{R_{12}}$ $\varphi_{R_{21}}$ $\varphi_{R_{22}}$	$Y_{R_{11}}$ $Y_{R_{12}}$ $Y_{R_{21}}$ $Y_{R_{22}}$	Y_{int}

и их интегральных оценок, где значения Y_R выбираются из табл. 5. Данные для расчета интегральной оценки сводятся в табл. 6.

Интегральная оценка производственных рисков проекта Y_{int} является компенсирующей, т. е. высокие значения интегральных оценок одних классификационных групп производственных рисков сглаживаются низкими значениями других. На основе полученного значения интегральной оценки $0 \leq Y_{\text{int}} \leq 5$ делаются выводы об общем уровне риска ОКР.

4 Программная реализация и апробация

Для апробации разработанной методики оценки производственных рисков ОКР выполнен проект интегрированного риск-менеджмента на уровне оборонного предприятия. С этой целью была разработана интегрированная система управления рисками (ИСУПР) оборонных предприятий при выполнении ОКР, за основу которой взят подход к управлению рисками стандарта PMBoK (Project Management Body of Knowledge), который включает в себя следующие этапы [10]:

- (1) планирование управления производственными рисками ОКР — процесс формирования системы управления рисками, методологии оценки и управления рисками, разработки плана управления рисками;
- (2) идентификация производственного риска ОКР — процесс определения перечня рисков, которые могут влиять на результаты ОКР, и описание характера их влияния, выбор и обоснование оценочных индикаторов, выбор методов оценки рисков с учетом их специфики;
- (3) качественный анализ производственных рисков ОКР — процесс расстановки приоритетов относительно важности рисков для их дальнейшего анализа (ранжирование рисков) и дальнейших действий по вычислению весовых коэффициентов, отражающих важность того или иного риска для достижения целей и задач ОКР;
- (4) количественный анализ производственных рисков ОКР — процесс численного анализа влияния идентифицированных рисков на цели и достижение

результата ОКР, что предполагает оценку и сопоставление уровня влияния и вероятности возникновения каждого конкретного риска;

- (5) разработка и реализация управленческих мероприятий — процесс обоснования методов и инструментов для минимизации деструктивного влияния производственных рисков на возможность достижения целей и задач ОКР;
- (6) мониторинг и контроль — процесс внедрения принятых управленческих решений в практику управления производственными рисками ОКР оборонного предприятия, мониторинг идентифицированных рисков, идентификация и формализация новых рисков, оценка результативности управленческих мероприятий в ходе выполнения ОКР.

ИСУПР внедрялась в деятельность Концерна ВКО «Алмаз-Антей», обеспечив, в числе прочего, интеграцию инструментов анализа, мониторинга и управления в ключевые корпоративные процессы. В системе линейного и функционального управления были определены уровни ответственности за оценку и управление производственными рисками и составление соответствующей отчетности. Каждому риску присвоен владелец риска, а также система оценочных индикаторов. На уровне управленческих функций, подразделений и ключевых бизнес-процессов назначены ответственные координаторы по рискам среди руководителей, которые внедряют принципы управления рисками в корпоративную культуру, распространяя и поддерживая процессы управления рисками, участвуя в создании системы мотивации для персонала. Определялись сроки и задачи по оценке, анализу и мониторингу рисков с учетом специфики оценочных индикаторов бизнес-процессов, на уровне которых применяется управление рисками.

Подобный интегрированный подход позволил сформировать зоны ответственности в системе управления производственными рисками ОКР и проводить их мониторинг на всех уровнях корпоративного управления, а также обеспечивать разработку целевых планов реагирования на существенные риски как в каждом дочернем обществе Концерна, так и по корпоративной структуре в целом.

5 Заключение

Разработанная методика оценки производственных рисков создания вооружения на этапе ОКР оборонного предприятия является гибкой и учитывает разные источники возникновения, сложность, информационно-аналитическую базу оценки и методы оценки. Основой данной методики служат широко применяемые на практике статистические методы и метод экспертных оценок, однако число применяемых методов оценки рисков не ограничено в целом за счет гибкости методики. Методика позволяет провести оценку рисков проекта на этапе ОКР, когда должно быть принято решение о начале проектирования уникального образца вооружения, возможно, не имеющего аналогов, однако при этом методика не связана с поэтапной структурой процесса проведения ОКР, не зависит от его

сроков, уникальности и сложности. В результате можно сфокусировать ресурсы на управлении производственными рисками по выделенным классификационным группам рисков проекта, установив взаимосвязь между производственными рисками ОКР, портфелем проектов и научно-технической политикой предприятия (т. е. стратегией проведения ОКР) в целом. Методика позволяет комплексно управлять рисками конкретного проекта, а также разрабатывать аналитически обоснованные корректирующие мероприятия по их снижению.

Литература

1. Лапуста М. Г., Шаршукова Л. Г. Риски в предпринимательской деятельности. — М.: ИНФРА-М, 1998. 224 с.
2. Хохлов Н. В. Управление риском. — М.: ЮНИТИ, 2001. 239 с.
3. Тэпман Л. Н. Риски в экономике / Под. ред. В. А. Швандара. — М.: ЮНИТИ, 2002. 380 с.
4. Мухин М. Е., Луговнина С. М. Риски инновационных проектов на предприятиях ОПК // Аллея науки, 2017. Т. 3. № 15. С. 421–424.
5. Мовтян Б. А., Данилаев Д. П. Оценка факторов риска выполнения гособоронзаказа в структуре концерна ОПК // Вестник ИжГТУ имени М. Т. Калашникова, 2018. Т. 21. № 3. С. 38–43.
6. Грачева М. В. Анализ проектных рисков. — М.: Финстатинформ, 1999. 126 с.
7. Прыкин Б. В. Техничко-экономический анализ производства. — М.: ЮНИТИ, 2000. 399 с.
8. Дубров А. М., Лагоша Б. А. Моделирование рискованных ситуаций в экономике и бизнесе. — М.: Финансы и статистика, 2001. 224 с.
9. Методика определения допустимой стоимости разработки и серийного производства образцов ВВТ на начальных стадиях жизненного цикла образцов ВВТ и учета экономических факторов, оказывающих влияние на принятие решений при обосновании и формировании государственной программы вооружения и государственного оборонного заказа / Под ред. С. Н. Остапенко. — М.: ВНИИНС, 2001. 115 с.
10. PMBOK® Guide. A guide to the project management body of knowledge. — Newtown Square, PA, USA: Project Management Institute, 2004. 216 p.

Поступила в редакцию 30.06.21

METHODOLOGY FOR ASSESSING PRODUCTION RISKS FOR DEVELOPING WEAPONS AND MILITARY EQUIPMENT

A. V. Bosov¹ and A. A. Kryukov²

¹Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

²Almaz-Antey Air and Space Defence Corporation, 41 Vereyskaya Str., Moscow 121471, Russian Federation

Abstract: A methodology for assessing production risks at the stage of research and development work (R&D) in the development of high-tech products is presented. To take into account the complex classification of production risks, various sources of their occurrence, and spheres of influence on the R&D result, it is proposed to introduce a combined methodological approach and assess production risks partly by statistical methods and partly by the method of expert assessments. The content of the stages of the corresponding methodology is considered. The method of calculating the integral assessment of production risk and the scale of the levels of production risks of the project in accordance with the nature of the impact and acceptability are substantiated.

Keywords: methodology for assessing industrial risks; experimental and design works; method of expert assessments; statistical methods

DOI: 10.14357/08696527210308

References

1. Lapusta, M. G., and L. G. Sharshukova. 1998. *Riski v predprinimatel'skoy deyatelnosti* [Business risks]. Moscow: INFRA-M. 224 p.
2. Hoxhlov, N. V. 2001. *Upravlenie riskom* [Risk management]. Moscow: YuNITI. 239 p.
3. Tepman, L. N. 2002. *Riski v ekonomike* [Risks in the economy]. Moscow: YuNITI. 380 p.
4. Muhin, M. E., and S. M. Lugovnina. 2017. Riski innovatsionnykh proektov na predpriyatiyakh OPK [Risks of innovative projects at the enterprises of the defense industry]. *Alleya nauki* [Alley of Science] 3(15):421–424.
5. Movtyan, B. A., and D. P. Danilaev. 2018. Otsenka faktorov riska vypolneniya gosoboronzakaza v strukture kontserna OPK [Estimating the risk factors of the state defense orders implementation at the concern structure]. *Vestnik IzhGTU imeni M. T. Kalashnikova* [Bulletin of the Izhevsk State Technical University named after M. T. Kalashnikov] 21(3):38–43.
6. Gracheva, M. V. 1999. *Analiz proektnykh riskov* [Analysis of project risks]. Moscow: Finstatinform. 126 p.
7. Prykin, B. V. 2000. *Tekhniko-ekonomicheskiy analiz proizvodstva* [Technical and economic analysis of production]. Moscow: YuNITI. 399 p.

8. Dubrov, A. M., and B. A. Lagosha. 2001. *Modelirovanie riskovykh situatsiy v ekonomike i biznese* [Modeling of risk situations in economics and business]. Moscow: Finansy i statistika. 224 p.
9. Ostapenko, S. N., ed. 2001. *Metodika opredeleniya dopustimoy stoimosti razrabotki i seriynogo proizvodstva obraztsov VVT na nachal'nykh stadiyakh zhiznennogo tsikla obraztsov VVT i ucheta ekonomicheskikh faktorov, okazyvayushchikh vliyaniye na prinyatie resheniy pri obosnovanii i formirovanii gosudarstvennoy programmy vooruzheniya i gosudarstvennogo oboronnogo zakaza* [Methodology for determining the permissible cost of development and serial production of weapons and military equipment at the initial stages of the life cycle of weapons and military equipment and taking into account the economic factors influencing decision-making in the justification and formation of the state weapons program and the state defense order]. Moscow: VNIINS. 115 p.
10. PMBOK® Guide. 2004. A guide to the project management body of knowledge. Newtown Square, PA: Project Management Institute. 216 p.

Received June 30, 2021

Contributors

Bosov Alexey V. (b. 1969) — Doctor of Science in technology, principal scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; AVBosov@ipiran.ru

Kryukov Alexey A. (b. 1983) — deputy chief of staff of general director, Almaz-Antey Air and Space Defence Corporation, 41 Vereyskaya Str., Moscow 121471, Russian Federation; alexeykryukov@almaz-antey.ru

КОНЦЕПЦИЯ ПОСТРОЕНИЯ НАДКОРПУСНЫХ БАЗ ДАННЫХ*

*М. Г. Кругжков*¹

Аннотация: Представлен обзор концепции, основных структурных составляющих и функций надкорпусных баз данных (НБД). Надкорпусные базы данных представляют собой новый вид структурированных информационных ресурсов. Они существенно расширяют возможности лингвистических текстовых корпусов, в особенности параллельных. Цель статьи — познакомить читателей с основными возможностями параллельных корпусов, а также продемонстрировать, каким образом концепция НБД позволяет расширять эти возможности и преодолевать некоторые их ограничения. В рамках НБД лингвисты могут устанавливать, фиксировать и аннотировать переводные соответствия (ПС) между языковыми единицами (ЯЕ) в языках оригинала и перевода, при этом для их аннотации используются рубрики фасетных классификаций, которые исследователи формируют в соответствии со своими потребностями. Также описана общая архитектура НБД, разработанных в ФИЦ ИУ РАН, которая подразделяется на корпусную и надкорпусную составляющие, взаимодействующие друг с другом в рамках единой базы данных.

Ключевые слова: корпусная лингвистика; надкорпусная база данных; параллельный корпус; лингвистическое аннотирование; информационные технологии; фасетная классификация

DOI: 10.14357/08696527210309

1 Введение

Корпуса параллельных текстов (параллельные корпуса) широко используются при проведении контрастивных (сопоставительных) лингвистических исследований, т. е. исследований, ставящих своей целью сравнение языковых явлений в различных языках. Параллельные корпуса позволяют давать ответы на различные вопросы сопоставительной лингвистики, основываясь на эмпирических данных большого объема, благодаря чему результаты контрастивных исследований во многом освобождаются от субъективности, которая может иметь место, когда исследователи обосновывают свои выводы, основываясь только на своей языковой интуиции.

*Работа выполнена в Институте проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук с использованием ЦКП «Информатика» ФИЦ ИУ РАН.

¹Федеральный исследовательский центр «Информатика и управление» Российской академии наук, magnit75@yandex.ru

Параллельные корпуса представляют собой собрания параллельных текстов, каждый из которых включает в себя некоторый текст на языке оригинала (ЯО) и один или несколько переводов этого текста на другой язык (язык перевода, ЯП) или на несколько разных языков¹. В параллельном тексте оригинал и перевод разбиты на равное число фрагментов, при этом каждому фрагменту в тексте оригинала противопоставляется соответствующий ему по смыслу фрагмент в тексте перевода, благодаря чему в структуре параллельного текста образуются пары «оригинал–перевод». Протяженность этих фрагментов в разных корпусах, как правило, составляет от одного предложения до одного абзаца². Для удобства восприятия очень длинные предложения иногда дробятся на несколько законченных по смыслу фрагментов или, наоборот, несколько кратких предложений могут объединяться в один расширенный фрагмент. Следует также учитывать, что границы предложений в тексте оригинала и тексте перевода могут не совпадать: одно предложение в оригинале может быть разбито на два или больше предложений в переводе или наоборот — одному предложению в переводе может соответствовать два или больше предложений в оригинале.

Кроме того, словоформы обычно снабжаются разметкой, помогающей лингвистам находить в параллельных текстах интересующие их ЯЕ и конструкции по морфологическим и/или синтаксическим признакам, а также посредством задания леммы — базовой формы исследуемой единицы. Подробнее о структуре параллельных корпусов см., например, [2].

Корпуса параллельных текстов используются в контрастивных лингвистических исследованиях, в исследованиях по теории перевода, в переводческой практике, а также в учебных целях (см., например, [3–5]). Однако работа с параллельными корпусами зачастую требует значительных усилий от исследователей и не всегда дает возможность с помощью поисковых запросов напрямую получать ответы на поставленные вопросы, поскольку, составляя запросы, пользователь, как правило, не может достоверно установить, какие именно ЯЕ в ЯП соответствуют искомым ЯЕ в ЯО, и наоборот. В связи с этим исследователям приходится анализировать значительное число примеров вручную и при этом либо удерживать в памяти результаты этого анализа, либо фиксировать их на тех или иных носителях (в текстовых файлах, на бумажных карточках и т. п.), чтобы впоследствии иметь возможность сделать соответствующие обобщения на основе проанализированных данных. Однако зафиксированные подобным образом результаты анализа часто недостаточно хорошо структурированы; кроме того,

¹В данной статье в дальнейшем будет предполагаться, что параллельные тексты включают в себя только один перевод — это позволит избежать постоянных оговорок (*перевод* или *переводы*, *язык* или *языки* и т. д.). Особенности, возникающие при работе с параллельными текстами, включающими в себя более одного варианта перевода, будут оговариваться особо.

²Существуют также пословно выровненные тексты, однако они не представляют собой последовательность выровненных пар фрагментов, а имеют более сложную структуру, что вызвано тем, что разные языки имеют разный порядок слов, разные наборы служебных слов и т. п. (см., например, [1]).

в дальнейшем их технически сложно обрабатывать, особенно когда речь идет о значительных объемах языкового материала. Наконец, исследователи зачастую фиксируют таким образом не все исследуемые единицы, а только наиболее интересные и показательные с их точки зрения примеры, что не позволяет сделать обобщенные выводы о тех или иных явлениях, наблюдаемых во всем представительном массиве параллельных текстов. Следовательно, такая организация работы не всегда позволяет в полной мере использовать главные преимущества корпусного подхода: исследователи не имеют возможности задействовать эффект масштаба, чтобы получить точные и объективные количественные данные по способам и особенностям перевода исследуемых ЯЕ.

Для решения этой проблемы в Институте проблем информатики ФИЦ ИУ РАН была разработана и реализована концепция «надкорпусных баз данных». Надкорпусные базы данных призваны помочь исследователям выполнять масштабные проекты в области сопоставительной лингвистики, опирающиеся на богатый эмпирический материал, заключенный в параллельных корпусах [6–16]. Надкорпусные базы данных позволяют группам исследователей анализировать большое число примеров вместе с их переводами и *сохранять результаты этого анализа в унифицированном формате в рамках специально разработанной базы данных*, что позволяет в дальнейшем легко анализировать полученный массив данных на количественном и качественном уровне.

Плунгян В. А. отмечает, что «корпус... необходим профессиональным лингвистам, тем, кто так или иначе имеет дело с фактами языка, а значит, должен эти факты *сбирать и систематизировать*» [17] (курсив наш. — М. К.). Однако стандартный поисковый инструментарий большинства корпусов в основном позволяет лишь с той или иной степенью точности находить в корпусе определенные факты языка, но очень немногие корпуса предлагают функции, позволяющие именно *сбирать и систематизировать* эти факты, т. е. фиксировать их непосредственно в структуре корпуса и аннотировать их с помощью специальных классификационных признаков, с тем чтобы впоследствии иметь возможность получать обобщенную информацию по накопленному таким образом материалу. Именно эти функции стремятся взять на себя НБД, при этом очевидно, что они могли бы оказаться полезными и при работе с монологическими корпусами. Однако именно при анализе параллельных корпусов концепция НБД оказывается в особенности полезной, поскольку она позволяет не только выделять и аннотировать отдельные факты языка, но и дает возможность устанавливать переводные соответствия между ЯЕ и конструкциями в определенных языковых парах, открывая новые возможности для исследователей в области контрастивной лингвистики.

2 Структура надкорпусной базы данных

В идеале структура НБД должна быть построена таким образом, чтобы пользователи имели возможность работать с параллельным корпусом, находя в нем употребления исследуемых ЯЕ и их переводы, и одновременно ана-

лизировать и размечать выявленные таким образом переводные соответствия в рамках заданной классификационной схемы. Этого можно добиваться разными способами, однако общая архитектура системы должна одновременно обеспечивать возможность работы с параллельным корпусом и возможность сохранения и редактирования формируемых аннотаций переводных соответствий в рамках базы данных. Информацию, которая генерируется на основе корпусных данных в процессе аннотирования выявленных в параллельном корпусе переводных соответствий, будем называть надкорпусной, чтобы отличать ее от собственно корпусной информации, заключенной в исходном параллельном корпусе.

В целях обеспечения более удобного, быстрого и надежного взаимодействия между корпусной и надкорпусной составляющими было принято решение объединить их в рамках единой базы данных. Далее будет описана структура НБД, которая разработана в ФИЦ ИУ РАН и объединяет в себе корпусную и надкорпусную составляющие, а также обеспечивает механизмы взаимодействия между этими составляющими. Следует принимать во внимание, что данная структура использовалась для реализации нескольких НБД, поддерживающих проекты различных кросслингвистических исследований [6–16], вследствие чего некоторые детали этой общей схемы в различных НБД могут незначительно отличаться.

2.1 Корпусная составляющая

Корпусная составляющая предназначена для хранения параллельных текстов, загруженных в НБД. Источником текстов служат параллельные корпуса Национального корпуса русского языка [18]. В этой части НБД содержится информация о текстах (произведениях), содержащихся в корпусе, переводах этих текстов, авторах текстов и их переводов, выровненных парах фрагментов в рамках параллельных текстов, словоупотреблениях, составляющих эти фрагменты, и знаках пунктуации. Также корпусная составляющая содержит информацию о возможных вариантах морфологического разбора всех включенных в корпус словоформ, включая указание их леммы (базовой формы слова) и набора морфологических признаков. На основе этой информации пользователи НБД осуществляют первичный поиск фрагментов, в которых могут находиться интересующие их ЯЕ, после чего они получают возможность формировать аннотации выявленных переводных соответствий. Логическая структурная схема корпусной составляющей НБД представлена на рис. 1.

Примечание. Прямоугольники соответствуют таблицам НБД, линии между прямоугольниками — связям между таблицами. Связи «один к одному» обозначаются простыми линиями, связи «один ко многим» — линиями с одним простым и одним разветвленным концом, связи «многие ко многим», реализуемые с помощью дополнительных промежуточных таблиц, — линиями с двумя разветвленными концами. Например, одной словоформе могут соответствовать несколько морфологических разборов, которые могут относиться к одной или разным леммам (так, словофор-

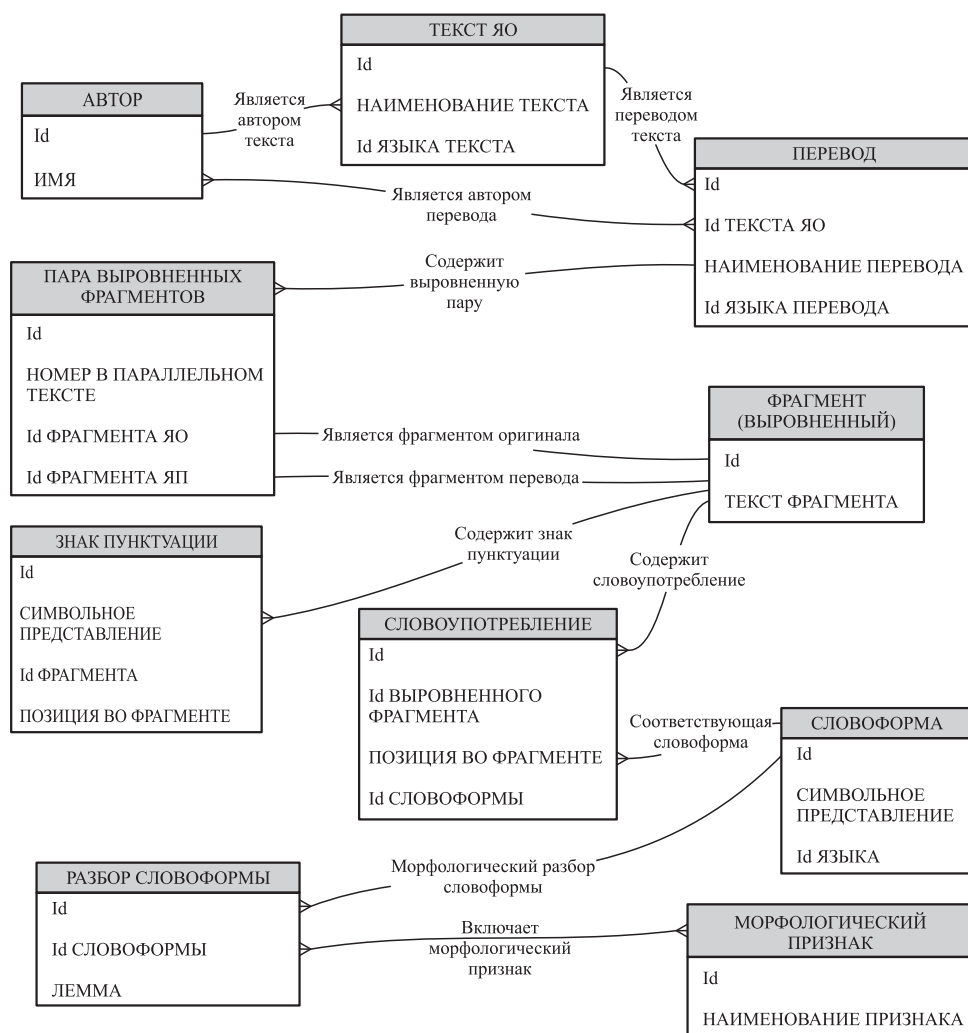


Рис. 1 Логическая структурная схема корпусной составляющей НБД

ме *постели* соответствуют леммы *постель* [сущ., род. падеж] и *постелить* [глагол, повелит. накл.]).

2.2 Надкорпусная составляющая

Структурная схема надкорпусной составляющей НБД предназначена для хранения аннотаций, выявленных в параллельном корпусе ЯЕ, и ПС (рис. 2).

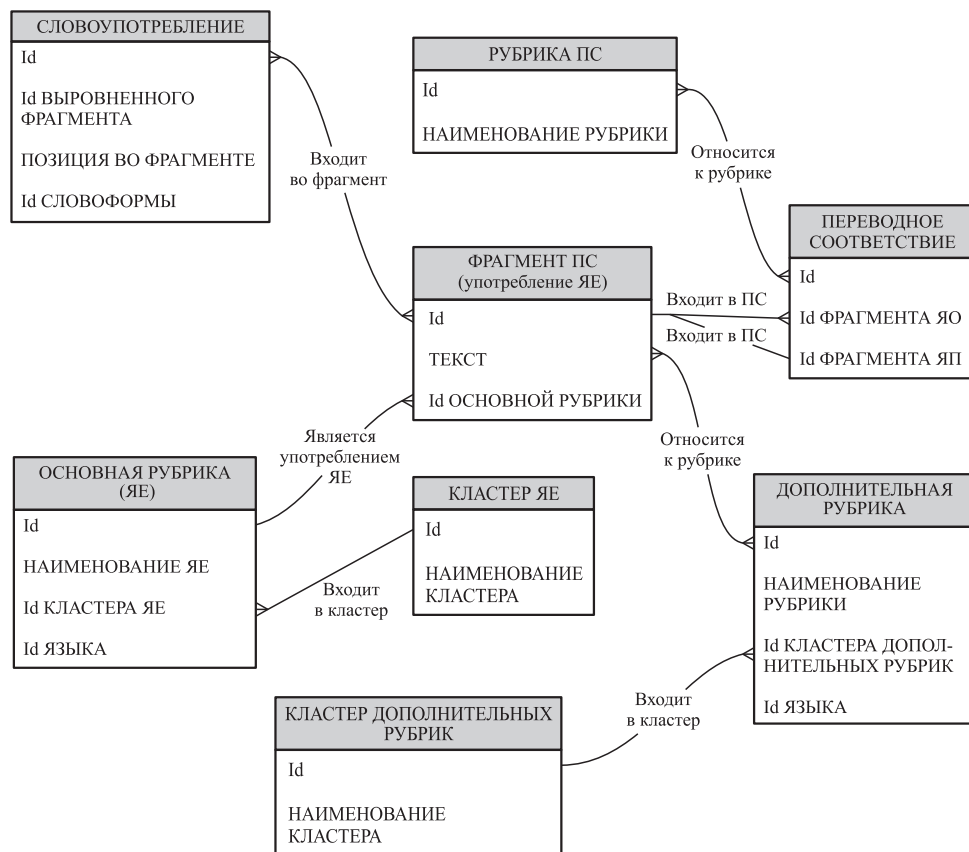


Рис. 2 Логическая структурная схема надкорпусной составляющей НБД

Центральный объект этой составляющей — переводное соответствие¹. Фрагменты ЯО и ЯП, объединенные в рамках ПС, описываются с помощью таблицы «Фрагмент», при этом фрагмент ЯО интерпретируется и фиксируется в НБД как употребление некоторой языковой единицы ЯО, а в соответствующем ему фрагменте ЯП выявляется некоторая ЯЕ или конструкция, которая отвечает за передачу исходной ЯЕ. Примечательно, что один и тот же фраг-

¹Во многих существующих работах вместо термина «переводное соответствие» используется термин моноэквивалентность (МЭ). Термин МЭ несет в себе большую семантическую нагрузку, поскольку само его содержание требует определения того, что следует понимать под эквивалентностью применительно к переводу. При этом в любом случае приходится констатировать, что зачастую никакой «эквивалентности» в зафиксированных ПС может быть не выявлено. Тем не менее термин МЭ традиционно широко используется применительно к проблематике НБД, поэтому оба термина, ПС и МЭ, в настоящее время используются как синонимы.

мент ЯО может быть задействован более чем в одном ПС, если в параллельном корпусе представлено несколько разных переводов одного и того же исходного текста.

Языковые единицы ЯО и ЯП фиксируются в качестве классификационных рубрик фрагментов в таблице «Основные рубрики», при этом для удобства пользователя эти рубрики объединяются в кластеры. Кроме того, при анализе фрагментов исследователи присваивают фрагментам ЯО и ЯП дополнительные классификационные рубрики, специфицирующие различные аспекты употребления рассматриваемых ЯЕ, которые также объединяются в кластеры. Таким образом, фрагменты ЯО и ЯП аннотируются с помощью *фасетной классификации*. При этом следует отметить, что кластеры основных и дополнительных рубрик не всегда непосредственно соответствуют фактическим фасетам, что связано с соображениями удобства пользователей НБД. Иногда несколько функционально или семантически близких фасетов группируются в рамках одного кластера: например, в рамках кластера «Грамматические признаки» объединяются рубрики нескольких фасетов, таких как *время (настоящее, прошедшее, будущее), наклонение (инфинитив, императив, причастие, деепричастие)* и т. д.

Наконец, самим ПС также присваиваются рубрики, специфицирующие те особенности ПС, которые определяют специфику переводной трансформации (например, *перефразирование, смена подлежащего* и т. д.).

3 Заключение

Как было отмечено выше, структура НБД должна обеспечивать возможность параллельной работы и с корпусной, и с надкорпусной составляющими, в связи с чем было принято решение объединить их в рамках одной базы данных. Как видно из рис. 1 и 2, эти составляющие связывает таблица «Словоупотребление». Между таблицами «Фрагмент» и «Словоупотребление» находится связывающая таблица, которая явно специфицирует, какие словоупотребления из параллельного корпуса входят во фрагменты ЯО и ЯП, служащие формообразующими элементами ПС. Данная связывающая таблица также позволяет выделять в рамках этих фрагментов сами рассматриваемые ЯЕ и «функционально значимые слова», т. е. слова, непосредственно влияющие на функционирование этих ЯЕ. Подробнее процесс построения ПС на основе корпусной информации описан в работе [19].

Возможность обращения к корпусной информации также играет важную роль при работе с массивом уже построенных аннотаций ПС. В частности, корпусная информация используется в системе просмотра и поиска ПС, позволяя осуществлять поиск аннотаций посредством задания автора исходного произведения или базовой формы тех или иных лексем (подробнее о поисковых функциях НБД см., например, [20]). Наконец, корпусная информация также может быть задействована при построении запросов к массиву сформированных ПС (см., например, [21]).

Литература

1. Морозова Ю. И., Козеренко Е. Б., Шарнин М. М. Методика извлечения пословных переводных соответствий из параллельных текстов с применением моделей дистрибутивной семантики // Системы и средства информатики, 2014. Т. 24. № 2. С. 131–142.
2. Кружков М. Г. Информационные ресурсы контрастивных лингвистических исследований: электронные корпуса текстов // Системы и средства информатики, 2015. Т. 25. № 2. С. 140–159.
3. Johansson S. Seeing through Multilingual Corpora: On the use of corpora in contrastive studies. — Studies in corpus linguistics ser. — Amsterdam/Philadelphia: John Benjamins Publishing Co., 2007. Vol. 26. 378 p.
4. Добровольский Д. О., Кротова Е. Б., Парина И. С. Корпусная лексикография (материалы мастер-класса) // Русская германистика: ежегодник Российского союза германистов. — М.: Языки славянской культуры, 2014. Т. 11. С. 237–278.
5. Сичинава Д. В. Параллельные корпуса в составе Национального корпуса русского языка: новые языки и новые задачи // Труды Института русского языка им. В. В. Виноградова, 2019. Вып. 21. С. 41–60.
6. Зализняк Анна А., Зацман И. М., Инькова О. Ю., Кружков М. Г. Надкорпусные базы данных как лингвистический ресурс // Корпусная лингвистика: Тр. VII Междунар. конф. — СПб.: СПбГУ, 2015. С. 211–218.
7. Семантика коннекторов: контрастивное исследование / Под. ред. О. Ю. Иньковой. — М.: ТОРУС ПРЕСС, 2018. 368 с.
8. Зацман И. М., Кружков М. Г. Надкорпусная база данных коннекторов: развитие системы терминов проектирования // Системы и средства информатики, 2018. Т. 28. № 4. С. 156–167.
9. Бунтман Н. В., Инькова О. Ю. Французский коннектор *sinon* // Семантика коннекторов: контрастивное исследование / Под ред. О. Ю. Иньковой. — М.: ТОРУС ПРЕСС, 2018. С. 301–330.
10. Добровольский Д. О., Зализняк Анна А. Немецкие конструкции с модальными глаголами и их русские соответствия: проект надкорпусной базы данных // Компьютерная лингвистика и интеллектуальные технологии: По мат-лам Междунар. конф. «Диалог». — М.: РГГУ, 2018. С. 172–184.
11. Структура коннекторов и методы ее описания / Под. ред. О. Ю. Иньковой. — М.: ТОРУС ПРЕСС, 2019. 316 с.
12. Зацман И. М., Кружков М. Г., Лощилова Е. Ю. Методы и средства информатики для описания структуры неоднословных коннекторов // Структура коннекторов и методы ее описания / Под ред. О. Ю. Иньковой. — М.: ТОРУС ПРЕСС, 2019. С. 205–230.
13. Гончаров А. А., Инькова О. Ю., Кружков М. Г. Методология аннотирования в надкорпусных базах данных // Системы и средства информатики, 2019. Т. 29. № 2. С. 148–160.
14. Инькова О. Ю., Кружков М. Г. Сочетаемость логико-семантических отношений: количественные методы анализа // Информатика и её применения, 2019. Т. 13. Вып. 2. С. 83–91.
15. Семантика коннекторов: количественные методы описания / Под ред. О. Иньковой. — Bern/Berlin: Peter Lang AG, 2021. 276 с.

16. Гончаров А. А., Бунтман Н. В., Нуриев В. А. Ошибки в машинном переводе коннекторов: сравнительный анализ работы двух автоматических переводчиков // Семантика коннекторов. Количественные методы описания / Под ред. О. Иньковой. — Bern/Berlin: Peter Lang AG, 2021. С. 225–276.
17. Плунгян В. А. Зачем нужен Национальный корпус русского языка? Неформальное введение // Национальный корпус русского языка: 2003–2005. — М.: Индрик, 2005. С. 6–20.
18. Национальный корпус русского языка. <https://ruscorpora.ru/new/index.html>.
19. Попкова Н. А., Инькова О. Ю., Зацман И. М., Крузиков М. Г. Методика построения моноэквивалентностей в надкорпусной базе данных коннекторов // Задачи современной информатики: Тр. 2-й Молодежной научн. конф. — М.: ФИЦ ИУ РАН, 2015. С. 143–153.
20. Инькова О. Ю., Крузиков М. Г. Надкорпусные русско-французские базы данных глагольных форм и коннекторов // *Langues slaves en contraste*. — Eds. О. Inikova, А. Trovesi. — Bergamo: Bergamo University Press, 2016. P. 365–392.
21. Зацман И. М., Крузиков М. Г., Ложилова Е. Ю. Методы анализа частотности моделей перевода коннекторов и обратимость генерализации статистических моделей // Системы и средства информатики, 2017. Т. 27. № 4. С. 164–176.

Поступила в редакцию 14.08.21

CONCEPTUAL FRAMEWORK FOR SUPRACORPORA DATABASES

M. G. Kruzikov

Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation

Abstract: The paper provides an overview of the concept, main structural constituents, and functions of supracorpora databases (SCDB). Supracorpora databases represent a novel type of structured information resources that significantly expand capabilities of linguistic text corpora, parallel corpora in particular. The paper outlines principle features and limitations of parallel corpora and demonstrates how SCDBs allow extending these features and overcoming the limitations. Supracorpora databases allow linguistic experts to establish, record, and annotate translation correspondences between language units in the source and target texts while relying on faceted classification categories composed by the researchers themselves according to their requirements. The article also describes the general structure of SCDB architecture developed in FRC CSC RAS which incorporates corpus and subcorpus constituents that interact with one another as a part of a common database.

Keywords: corpus linguistics; supracorpora database; parallel corpus; linguistic annotation; information technologies; faceted classification

DOI: 10.14357/08696527210309

Acknowledgments

The research was carried out using the infrastructure of the Shared Research Facilities “High Performance Computing and Big Data” (CKP “Informatics”) of FRC CSC RAS (Moscow).

References

1. Morozova, Yu. I., E. B. Kozerenko, and M. M. Sharnin. 2014. Metodika izvlecheniya poslovnnykh perevodnykh sootvetstviy iz parallel'nykh tekstov s primeneniem modeley distributivnoy semantiki [Method for extracting single-word translation correspondences from parallel texts using distributional semantics models]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 24(2):131–142.
2. Kruzhkov, M. G. 2015. Informatsionnye resursy kontrastivnykh lingvisticheskikh issledovaniy: elektronnye korpusa tekstov [Information resources for contrastive studies: Electronic text corpora]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 25(2):140–159.
3. Johansson, S. 2007. *Seeing through Multilingual Corpora: On the use of corpora in contrastive studies*. Studies in corpus linguistics ser. Amsterdam/Philadelphia: John Benjamins Publishing Co. Vol. 26. 378 p.
4. Dobrovol'skiy, D. O., E. B. Krotova, and I. S. Parina. 2014. Korpusnaya leksikografiya (materialy master-klassa) [Corpus lexicography (workshop materials)]. *Russkaya germanistika: ezhegodnik Rossiyskogo soyuza germanistov* [Russian German studies: Yearbook of the Russian Union of Germanists]. Moscow: Yazyki slavyanskoy kul'tury. 11:237–278.
5. Sichinava, D. V. 2019. Parallel'nye korpusa v sostave Natsional'nogo korpusa russkogo yazyka: novye yazyki i novye zadachi [Parallel corpora within the Russian National Corpus: New languages and new objectives]. *Trudy Instituta russkogo yazyka im. V. V. Vinogradova* [V. V. Vinogradov Russian Language Institute Proceedings]. 21:41–60.
6. Zaliznyak, A. A., I. M. Zatsman, O. Yu. Inkova, and M. G. Kruzhkov. 2015. Nadkorpusnye bazy dannykh kak lingvisticheskiy resurs [Supracorpora databases as linguistic resource]. *Tr. 7-y Mezhdunar. konf. “Korpusnaya lingvistika”* [7th Conference (International) on Corpus Linguistics Proceedings]. St. Petersburg: St. Petersburg State University. 211–218.
7. Inkova, O., ed. 2018. *Semantika konnektorov: kontrastivnoe issledovanie* [Semantics of connectives: A contrastive study]. Moscow: TORUS PRESS. 368 p.
8. Zatsman, I. M., and M. G. Kruzhkov. 2018. Nadkorpusnaya baza dannykh konnektorov: razvitie sistemy terminov proektirovaniya [Supracorpora database of connectives: Design-oriented evolution of the term system]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(4):156–167.
9. Buntman, N. V., and O. Yu. Inkova. 2018. Frantsuzskiy konnektor *sinon* [French connective *sinon*]. *Semantika konnektorov: kontrastivnoe issledovanie* [Semantics of connectives: A contrastive study]. Ed. O. Yu. Inkova. Moscow: TORUS PRESS. 301–330.

10. Dobrovol'skiy, D. O., and Anna A. Zaliznyak. 2018. Nemetskie konstruksii s modal'nymi glagolami i ikh russkie sootvetstviya: proekt nadkorpusnoy bazy dannykh [German constructions with modal verbs and their Russian correlates: A supracorpora database project]. *Komp'yuternaya lingvistika i intellektual'nye tekhnologii. Po mat-lam ezhegodnoy Mezhdunar. konf. "Dialog"* [Computational Linguistics and Intellectual Technologies. Papers from the Annual Conference (International) "Dialogue"]. Moscow: RGGU. 172–184.
11. Inkova, O., ed. 2019. *Struktura konnektorov i metody ee opisaniya* [Structure of connectives and methods for its description]. Moscow: TORUS PRESS. 316 p.
12. Zatsman, I., M. Kruzhkov, and E. Loshchilova. 2019. Metody i sredstva informatiki dlya opisaniya struktury neodnoslovnnykh konnektorov [Methods and means of informatics for multiword connectives structure description]. *Struktura konnektorov i metody ee opisaniya* [Connectives structure and methods of its description]. Ed. O. Yu. Inkova. Moscow: TORUS PRESS. 205–230.
13. Goncharov, A. A., O. Yu. Inkova, and M. G. Kruzhkov. 2019. Metodologiya annotirovaniya v nadkorpusnykh bazakh dannykh [Annotation methodology of supracorpora databases]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(2):148–160.
14. Inkova, O. Yu., and M. G. Kruzhkov. 2019. Sochetaemost' logiko-semanticheskikh otnosheniy: kolichestvennye metody analiza [Compatibility of logical semantic relations: Methods of quantitative analysis]. *Informatika i ee Primeneniya — Inform. Appl.* 13(2):83–91.
15. Inkova, O., ed. 2021. *Semantika konnektorov: kolichestvennye metody opisaniya* [Semantics of connectives: Quantitative methods of analysis]. Bern/Berlin: Peter Lang. 276 p.
16. Goncharov, A. A., N. V. Buntman, and V. A. Nuriev. 2021. Oshibki v mashinnom perevode konnektorov: sravnitel'nyy analiz raboty dvukh avtomaticheskikh perevodchikov [Errors in machine translation of connectives: Comparative analysis of two translation engines]. *Semantika konnektorov: kolichestvennye metody opisaniya* [Semantics of connectives: Quantitative methods of analysis]. Ed. O. Inkova. Bern/Berlin: Peter Lang AG. 225–276.
17. Plungyan, V. A. 2005. Zachem nuzhen Natsional'nyy korpus russkogo yazyka? Neformal'noe vvedenie [What the Russian National Corpus is for? Informal introduction]. *Natsional'nyy korpus russkogo yazyka: 2003–2005* [Russian National Corpus 2003–2005]. Moscow: Indrik. 6–20.
18. Natsional'nyy korpus russkogo yazyka [Russian National Corpus]. Available at: <http://www.ruscorpora.ru> (accessed August 11, 2021).
19. Popkova, N. A., O. Yu. Inkova, I. M. Zatsman, and M. G. Kruzhkov. 2015. Metodika postroeniya monoekvivalentsiy v nadkorpusnoy baze dannykh konnektorov [Methodology of constructing monoequivalences in the supracorpora database of connectors]. *Tr. 2-y Molodezhnoy nauchn. konf. "Zadachi sovremennoy informatiki"* [2nd Scientific Conference "Problems of Modern Informatics" Proceedings]. Moscow: FRC CSC RAS. 143–153.
20. Inkova, O. Yu., and M. G. Kruzhkov. 2016. Nadkorpusnye russko-frantsuzskie bazy dannykh glagol'nykh form konnektorov [Supracorpora databases of Russian and French verbal forms and connectors]. *Lingue slave a confronto*. Eds. O. Inkova and A. Trovesi. Bergamo: Bergamo University Press. 365–392.

21. Zatsman, I. M., M. G. Kruzhkov, and E. Yu. Loshchilova. 2017. Metody analiza chastotnosti modeley perevoda konnektorov i obratimost' generalizatsii statisticheskikh dannykh [Methods of frequency analysis of connectives translations and reversibility of statistical data generalization]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 27(4):164–176.

Received August 14, 2021

Contributor

Kruzhkov Mikhail G. (b. 1975) — 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; magnit75@yandex.ru

АППАРАТНАЯ РЕАЛИЗАЦИЯ РЕКУРРЕНТНОГО ОБРАБОТЧИКА СИГНАЛОВ*

Ю. А. Степченко¹, Н. В. Морозов², Ю. Г. Дьяченко³, Д. В. Хилько⁴

Аннотация: Представлены результаты аппаратной реализации гибридной многоядерной архитектуры рекуррентного сигнального процессора (ГМАРСП) в виде VHDL-модели уровня регистровых передач и ее апробации в виде макетного образца на отладочной плате с программируемой логической интегральной схемой (ПЛИС) Intel Arria10. Гибридная многоядерная архитектура рекуррентного сигнального процессора состоит из ведущего фон-неймановского процессора, реализующего управляющий уровень архитектуры, и потокового процессора с четырьмя вычислительными секциями на операционном уровне архитектуры. Аппаратная модель ГМАРСП представляет собой совокупность программной или аппаратной реализации управляющего процессора (УП) и VHDL-модели операционного уровня ГМАРСП. Программная реализация УП предоставляется системой Quartus автоматизированного проектирования цифровых СБИС на ПЛИС фирмы Intel. Аппаратную реализацию УП в виде двухъядерного процессора Cortex-A9 обеспечивает ПЛИС на отладочной плате.

Ключевые слова: рекуррентный сигнальный процессор; гибридная многоядерная архитектура; VHDL-модель; ПЛИС

DOI: 10.14357/08696527210310

1 Введение

Гибридная многоядерная архитектура рекуррентного сигнального процессора [1] — это альтернатива традиционной вычислительной архитектуре. Она ориентирована на решение задач цифровой обработки сигналов (ЦОС) и относится к классу потоковых архитектур, но свободна от недостатков, присущих потоковым архитектурам массового параллелизма на базе ассоциативной памяти:

- высокой аппаратной сложности и времени сравнения тегированных маркеров;
- неэффективного исполнения последовательных участков кода;

* Исследование выполнено при поддержке Российского научного фонда (проект 19-11-00334).

¹Федеральный исследовательский центр «Информатика и управление» Российской академии наук, YStepchenkov@ipiran.ru

²Федеральный исследовательский центр «Информатика и управление» Российской академии наук, NMorozov@ipiran.ru

³Федеральный исследовательский центр «Информатика и управление» Российской академии наук, diauga@mail.ru

⁴Федеральный исследовательский центр «Информатика и управление» Российской академии наук, dhilko@yandex.ru

- чересчур большого числа команд, требующегося для выполнения программы, по отношению к процессорам традиционной архитектуры.
- использования сложной, медленной и энергоемкой ассоциативной памяти в качестве устройства сопоставления;
- большого размера теговых полей, превышающего в десятки раз размер собственно обрабатываемых данных.

Учет этих и других факторов потребовал особого подхода к внедрению потоковой парадигмы в область ЦОС. ГМАРСП — это компромиссное решение, обеспечивающее совместимость с существующими вычислительными и аппаратными средами.

Основным средством разработки и отладки капсул для ГМАРСП служит имитационная модель ГМАРСП [1]. Она учитывает особенности архитектурной организации ГМАРСП и позволяет разработчику капсулы проследить за продвижением данных по этапам работы алгоритма и их распределением по параллельным вычислительным ресурсам. Имитационная модель ГМАРСП облегчает процесс капсульного программирования. Однако она не учитывает особенности аппаратной реализации архитектуры и не может быть использована в качестве исходного алгоритмического описания ГМАРСП для ее синтеза в виде аппаратуры с помощью промышленных средств автоматизированного проектирования.

Поэтому разработка аппаратной модели ГМАРСП на уровне регистровых передач на языке VHDL, обеспечивающей адекватность аппаратной реализации ГМАРСП ее архитектуре и позволяющей оценить характеристики ГМАРСП на отладочной плате с ПЛИС, становится актуальной задачей. Данная статья описывает особенности реализации аппаратной модели ГМАРСП в виде макетного образца на отладочной плате HAN Pilot Platform с ПЛИС Intel Arria10 SoC 10AS066K3F40E2SG [2], ее технические характеристики и требуемые аппаратные ресурсы в терминах структурных единиц, имеющихся в составе ПЛИС.

2 Особенности аппаратной реализации ГМАРСП

ГМАРСП имеет следующие особенности аппаратной реализации, преодолевающие недостатки известных потоковых архитектур.

1. Работа ассоциативного запоминающего устройства (АЗУ) основана на сравнении теговых полей входного или искомого операнда с теговыми полями всех операндов в АЗУ. Адрес ячейки АЗУ, в которой теговые поля совпали, выбирается для выполнения операции записи/чтения. Это существенно усложняет АЗУ по сравнению с адресуемой памятью, замедляет его и увеличивает энергопотребление. Вместо большого и медленного массива ассоциативной памяти в ГМАРСП используются прямо адресуемые регистровые памяти с намного меньшим объемом, расположенные в тех блоках архитектуры ГМАРСП, где они востребованы:

- 4×16 бит для хранения констант и 16×52 бит для временного хранения операндов в памяти совпадений каждой вычислительной секции;
- 8×16 бит для хранения констант и 16×32 бит для подгружаемых в процессе обработки капсулы данных в вычислителе каждой секции.

Тем самым устраняется узкое место классических потоковых архитектур. Запись и чтение операндов в память на основе сравнения теговых полей выполняется в ГМАРСП в несколько раз быстрее.

2. Последовательные участки кода выполняются в УП без потери времени на сравнение тегированных маркеров и с более высокой скоростью, чем в рекуррентном обработчике сигналов (РОС): частота работы РОС на ПЛИС не превышает 25 МГц, в то время как частота синхронизации УП NIOS-II составляет 160 МГц, а ARM Cortex-A9 — 1,5 ГГц. Кроме того, в ряде случаев есть возможность исполнить последовательный код в УП одновременно с подготовкой к выполнению в РОС параллельного участка кода (загрузка конфигурационных параметров, констант и т. д.).
3. В каждый момент времени выполняется только одна итерация цикла. Это резко сокращает размер тегов, потери времени на их коммуникацию по сети и объем аппаратуры. Если в потоковых системах массового параллелизма разрядность теговых полей в несколько раз превышает полезную разрядность обрабатываемых данных, то в ГМАРСП число разрядов теговых полей (36) сравнимо с полезной разрядностью обрабатываемых данных: 16 и 38 в неупакованных операндах и 64 в упакованных.

Предложенные способы реализации хранилищ данных обеспечивают предпосылки для реанимации потоковой концепции в области ЦОС, стоимостный фактор которой делает нецелесообразным прямолинейное использование в ней технических решений из области потоковых систем массового параллелизма. Они не требуют сложных аппаратных механизмов для поддержки присущей ГМАРСП капсульной модели программирования.

3 Аппаратная модель ГМАРСП

Структура аппаратной модели ГМАРСП представлена на рис. 1. Она включает в себя УП, шину адреса/данных и РОС. К шине адреса/данных могут быть подключены дополнительные устройства памяти и интерфейса, относящиеся к управляющему уровню.



Рис. 1 Структура аппаратной модели ГМАРСП

В рамках методологии построения архитектуры ГМАРСП роль УП может играть традиционный (фон-неймановский) процессор любого типа. Главное

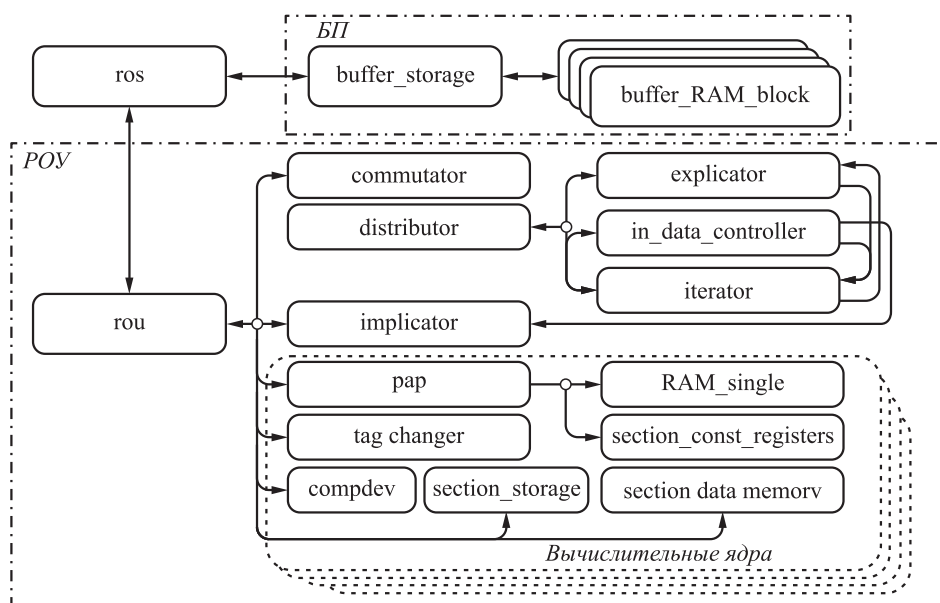


Рис. 2 Структура VHDL-модели РОС

требование к его характеристикам — достаточное быстродействие, обеспечивающее требуемую обработку результатов РОС и своевременную подкачку данных в РОС. В текущей реализации ГМАРСП на ПЛИС Intel Arria10 SoC функции УП выполняет программный процессор NIOS-II или аппаратный процессор ARM Cortex-A9. NIOS-II интегрируется в проект ГМАРСП в виде готового сложного функционального блока, имеющегося в библиотеке системы автоматизированного проектирования (САПР) Quartus Prime SE 18 (Intel) [3]. ARM Cortex-A9 изначально присутствует в ПЛИС на используемой отладочной плате HAN Pilot Platform.

Блок РОС реализован в виде модели уровня регистровых передач на языке VHDL. Целиком аппаратная модель, изображенная на рис. 1, собирается в приложении Qsys, запускаемом из оболочки САПР Quartus.

Структура аппаратной модели РОС на языке VHDL представлена на рис. 2. Она состоит из главного модуля **ros** и модулей, реализующих функционал РОС. Модуль верхнего уровня **ros** организует взаимодействие буферной памяти (БП, **buffer_storage**) и рекуррентного обрабатывающего устройства (РОУ, **rou**) и содержит арбитр обращений к БП со стороны УП (чтение и запись) и РОУ (запись). Запись в БП со стороны РОУ имеет высший приоритет.

Особенности организации работы блока БП:

- (1) БП состоит из 4 банков памяти и контроллера БП, управляющего обменом данными между БП, РОУ и УП. Один операнд капсулы располагается

в четырех банках БП по одному базовому адресу. Каждый операнд сопровождается битом готовности, разрешающим чтение операнда из БП в РОУ;

- (2) банки БП представляют собой двухпортовые ОЗУ. Первый порт служит для записи данных в БП со стороны УП и РОУ и для чтения данных из БП в УП. Он управляется системной частотой шины адреса/данных. Второй порт используется для чтения операндов из БП в РОУ и управляется внутренней частотой РОУ;
- (3) 64-разрядные операнды записываются в БП из УП и считываются из БП в УП по шине адреса/данных за две операции записи, по 32 разряда за одну операцию;
- (4) чтение из БП в РОУ и запись выходных и промежуточных результатов из РОУ в БП выполняется одним 64-битным операндом.

Аппаратная модель РОС реализована в виде кольцевого конвейера, ступени которого используют микроконвейеры для обработки данных в течение одного вычислительного шага (ВШ) длительностью четыре периода тактовой частоты РОУ.

Ступень 1 включает контроллер БП и массив ячеек памяти БП, заполняемый из УП капсулами реализуемых алгоритмов (модуль **buffer_storage**). Контроллер БП считывает по 4 операнда за один ВШ и передает их в РОУ попарно.

Ступень 2 включает контроллер входных данных (КВД, модуль **in_data_controller**). Здесь считанные из БП операнды распаковываются (при необходимости), рассортировываются и передаются в экспликатор (ступень 3) группой из 1–4 операндов. Операнды, предназначенные для итератора (ступень 3) и импликатора (ступень 6), передаются по назначению.

Ступень 3 включает экспликатор (модуль **explicator**) и итератор (модуль **iterator**). Экспликатор распределяет операнды по вычислительным секциям. Итератор организует циклическое использование операндов.

Ступень 4 включает секционные памяти совпадений (ПС, модуль **pap**): ПС подготавливает операнды для обработки в секционном вычислителе и имеет доступ в секционные памяти констант.

Ступень 5 включает секционный вычислитель (модуль **compdev**) и преобразователь тегов (ПТ, модуль **tag_changer**). Они вычисляют содержательную и функциональную части нового операнда.

Ступень 6 включает коммутатор (модуль **commutator**) и импликатор (модуль **implicator**). Коммутатор комплектует операнды результатов обработки капсулы и при необходимости отправляет их в экспликатор для дальнейшего использования внутри РОУ. Импликатор вылавливает в потоке результирующих операндов промежуточные и окончательные выходные данные и отправляет их в БП.

Память операндов есть не только в блоке БП, но и в блоках КВД и экспликатора в виде FIFO (first in, first out). Она обеспечивает безостановочную

Таким образом, латентность конвейера в данном случае равна 6 ВШ или 24 тактам частоты синхронизации РОС (сигнал clk). Как видно из рис. 3, дольше всего операнд обрабатывался в экспликаторе из-за того, что в экспликаторе есть входное FIFO, в котором операнд как минимум два ВШ продвигался к выходу из экспликатора. Эту ситуацию можно охарактеризовать как предвыборку входных операндов из БП и не учитывать указанные два ВШ при определении латентности конвейера РОС. Рисунок 3 подтверждает, что ПС не является узким местом ГМАРСП.

4 Реализация ГМАРСП на ПЛИС

Аппаратные ресурсы, потребовавшиеся для реализации ГМАРСП в ПЛИС в режиме автоматического синтеза с помощью САПР Quartus [3], представлены в таблице для двух вариантов реализации УП: программного процессора NIOS-II и аппаратного процессора ARM Cortex-A9. Рисунок 4 демонстрирует результат синтеза варианта ГМАРСП с NIOS-II.

Сравнение аппаратных ресурсов для двух вариантов УП показывает следующее:

- ресурсы для аппаратной реализации РОС при использовании программного и аппаратного процессора управляющего уровня примерно одинаковы;
- дополнительные ресурсы (около 5% от общего объема использованных ресурсов ПЛИС в случае УП NIOS-II и около 9% в случае ARM Cortex-A9) требуются для аппаратной реализации интерфейса РОС с УП и периферии.

Аппаратные ресурсы при реализации ГМАРСП

Блоки ГМАРСП	Ресурсы ПЛИС с NIOS-II (N) и ARM Cortex-A9 (A)									
	ALMs		Registers		Memory bits		M20K Blocks		DSP Blocks	
	N	A	N	A	N	A	N	A	N	A
Буферная память	5 153	8 609	10 097	10 113	141 900	141 900	16	16	0	0
Контроллер входных данных	1 852	2 025	1 109	1 107	0	0	0	0	0	0
Экспликатор	13 565	12 539	2 687	2 698	0	0	0	0	0	0
Итератор	1 249	1 372	601	640	0	0	0	0	0	0
Память совпадений	4 317	4 192	3 888	3 703	3 200	3 200	12	12	0	0
Вычислитель	18 252	13 953	1 920	2 938	0	0	0	0	4	4
Коммутатор	626	442	589	380	0	0	0	0	0	0
Преобразователь тегов	45	37	92	94	0	0	0	0	0	0
Импликатор	1 416	1 174	653	653	0	0	0	0	0	0
Секционная память	14 212	14 298	20 888	20 888	2 048	2 048	8	8	0	0
Управляющий уровень	3 641	6 675	4 048	9 254	33 788 672	33 555 968	2 080	2 051	4	0
Итого:	64 328	65 316	46 746	51 840	33 935 820	33 703 116	2 106	2 087	8	4

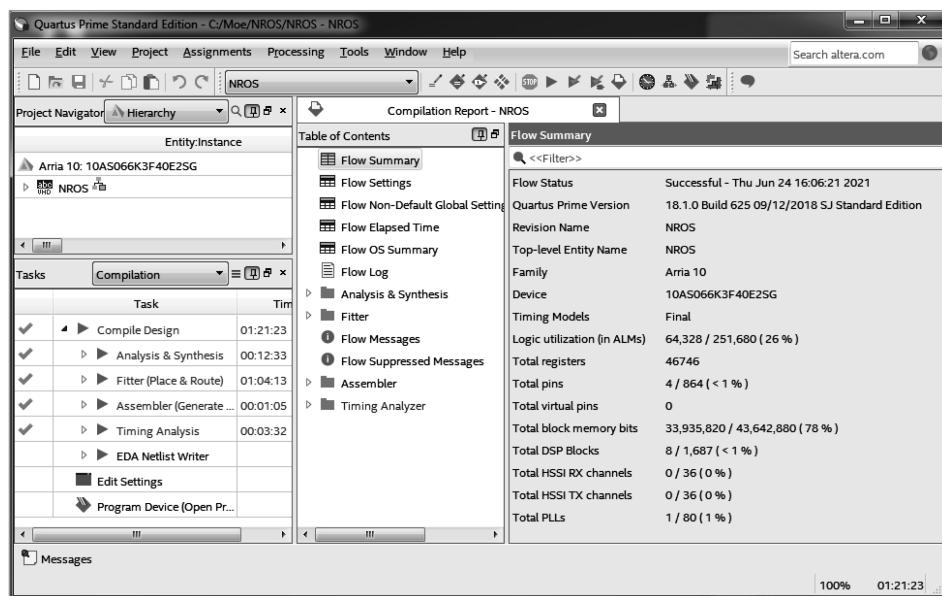


Рис. 4 Скриншот результатов синтеза ГМАРСП с УП NIOS-II

Реализация ГМАРСП на ПЛИС преследовала цель апробации рекуррентной архитектуры в аппаратном виде и получения сравнительных оценок функциональных возможностей рекуррентной архитектуры и ее производительности на синтетических тестах и реальных приложениях ЦОС. ПЛИС обеспечивает быстрое проектирование и корректировку любого проекта за счет программирования готовых аппаратных ресурсов — адаптивных логических блоков, параметризуемых блоков памяти, аппаратных умножителей и т. д., но не претендует на достижение быстродействия, обеспечиваемого в заказной технологии.

Относительно небольшой объем использованных ресурсов ПЛИС (26%) от общего объема ресурсов ПЛИС позволит в дальнейшем расширить функциональность ГМАРСП за счет увеличения числа секций или встраивания аппаратной реализации типовых алгоритмов, например быстрого преобразования Фурье.

5 Заключение

Предложенные и реализованные особенности архитектурной организации ГМАРСП повысили эффективность использования потоковой парадигмы в приложениях ЦОС, сократив требуемый объем аппаратных затрат и повысив ее производительность. ПЛИС ограничивает производительность реализованных на ней приложений. Однако программные средства отладки и верификации при-

ложений на ПЛИС позволяют оценить производительность последних в числе операций на один такт частоты синхронизации.

Аппаратная модель рекуррентной части ГМАРСП на уровне регистровых передач на языке VHDL обеспечила верификацию архитектуры ГМАРСП на макетном образце на базе отладочной платы HAN Pilot Platform с ПЛИС Intel Arria10 SoC 10AS066K3F40E2SG.

Аппаратные ресурсы, требуемые для реализации текущей архитектуры ГМАРСП на ПЛИС, подтверждают возможность использования современных отладочных плат для макетирования следующих поколений ГМАРСП с более широкой функциональностью.

Литература

1. *Shikunov Yu., Stepchenkov Yu., Khilko D.* Recurrent mechanism developments in the data-flow computer architecture // IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering Proceedings. — Piscataway, NJ, USA: IEEE, 2018. P. 1413–1418. doi: 10.1109/EIConRus.2018.8317362.
2. HAN Pilot Platform. Specifications. <https://www.terasic.com.tw/cgi-bin/page/archive.pl?Language=English&CategoryNo=216&No=1133&PartNo=2>.
3. Intel Quartus Prime Software. <https://www.intel.ru/content/www/ru/ru/software/programmable/quartus-prime/download.html>.

Поступила в редакцию 25.06.21

RECURRENT SIGNAL PROCESSOR HARDWARE IMPLEMENTATION

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

Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation

Abstract: The paper presents the results of hybrid architecture of recurrent multicore signal processor (HARMSP) hardware implementation as register transfer level VHDL-model and its prototype approbation on a development board with Intel Arria10 field-programmable gate array (FPGA). HARMSP consists of von-Neumann master processor at control architecture level and data-flow recurrent processor with four computing sections at operational level. Hardware HARMSP model is a complex of software or hardware control processor (CP) implementation and operational level VHDL-model. CAD Quartus (Intel) provides the software CP implementation on FPGA, whereas SoC FPGA on the development board contains the hardware CP implementation as dual-core Cortex-A9 processor.

Keywords: recurrent signal processor; multicore hybrid architecture; VHDL-model; FPGA

DOI: 10.14357/08696527210310

Acknowledgments

The research was supported by the Russian Science Foundation (project No. 19-11-0034).

References

1. Shikunov, Yu., Yu. Stepchenkov, and D. Khilko. 2018. Recurrent mechanism developments in the data-flow computer architecture. *IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering Proceedings*. Piscataway, NJ: IEEE. 1413–1418. doi: 10.1109/EIConRus.2018.8317362.
2. HAN Pilot Platform. Specifications. Available at: <https://www.terasic.com.tw/cgi-bin/page/archive.pl?Language=English&CategoryNo=216&No=1133&PartNo=2> (accessed August 31, 2021). см русский вариант, он другой
3. Intel Quartus Prime Software. Available at: <https://www.intel.ru/content/www/ru/ru/software/programmable/quartus-prime/download.html> (accessed August 31, 2021).

Received June 25, 2021

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 119333, Russian Federation; YStepchenkov@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 119333, Russian Federation; NMorozov@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 119333, Russian Federation; diaura@mail.ru

Khilko Dmitri V. (b. 1987) — 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; dhilko@yandex.ru

МОДЕЛЬ ДЛЯ АНАЛИЗА ПРИОРИТЕТНОГО ДОСТУПА ТРАФИКА URLLC ПРИ ПРЕРЫВАНИИ ОБСЛУЖИВАНИЯ И СНИЖЕНИИ СКОРОСТИ ПЕРЕДАЧИ СЕССИЙ eMBB В СЕТИ 5G*

И. А. Кочеткова¹, А. И. Куцазли², П. А. Харин³, С. Я. Шоргин⁴

Аннотация: Совместная реализация в рамках одной сети пятого поколения (5G) двух сценариев — сверхнадежной передачи данных с малой задержкой URLLC (ultrareliable low-latency communication) и усовершенствованной подвижной широкополосной связи eMBB (enhanced mobile broadband) — является актуальной задачей. В статье для схемы с занятием не целого, а части ресурсного блока предложена модель совместной передачи узкополосного трафика URLLC и широкополосного трафика eMBB в виде системы массового обслуживания с приоритетным доступом — снижением и прерыванием передачи широкополосного трафика при поступлении узкополосного трафика. Сформулирована задача максимизации средней скорости передачи широкополосного трафика по значениям уровней скорости и при ограничениях на вероятность блокировки и вероятность прерывания обслуживания широкополосного трафика. Приведены результаты численного анализа показателей эффективности приоритетного доступа, такие как среднее число сессий eMBB и вероятность прерывания обслуживания.

Ключевые слова: 5G; eMBB; URLLC; система массового обслуживания; приоритетный доступ; прерывание обслуживания; снижение скорости передачи

DOI: 10.14357/08696527210311

1 Введение

Становление беспроводных сетей 5G требует от поставщиков услуг соблюдения строгих требований, связанных с высокой пропускной способностью, высокой скоростью передачи данных, сверхнизкой задержкой и надежностью. К сценариям использования услуг 5G относят крупномасштабные системы межмашинной связи (massive machine-type communication, mMTC), сверхнадежную

* Публикация создана при поддержке Программы стратегического академического лидерства РУДН. Исследование выполнено при финансовой поддержке РФФИ (проект 20-37-70079).

¹ Российский университет дружбы народов; Федеральный исследовательский центр «Информатика и управление» Российской академии наук, gudkova-ia@rudn.ru

² Российский университет дружбы народов, aikushch@yandex.ru

³ Российский университет дружбы народов, pxarin@mail.ru

⁴ Федеральный исследовательский центр «Информатика и управление» Российской академии наук, sshorgin@ipiran.ru

передачу данных с малой задержкой (URLLC) и усовершенствованную подвижную широкополосную связь (eMBB). Система mMTC обеспечивает подключение огромного числа устройств и приложений без требований к задержке; URLLC предназначена для критически важных коммуникаций с минимальной задержкой, высокой безопасностью и надежностью; eMBB требует высокой скорости передачи данных. Ввиду столь различных характеристик возникает задача совместного использования таких сценариев, как URLLC и eMBB.

Среди предлагаемых подходов сосуществования URLLC и eMBB можно выделить подход [1] с резервированием ресурса для трафика URLLC. В работе [2] исследуется применение технологии нарезки радиоресурсов (network slicing) в случае гетерогенного множественного доступа — как ортогонального (heterogeneous orthogonal multiple access, H-OMA), так и неортогонального (heterogeneous non-orthogonal multiple access, H-NOMA), исследованного также в [3,4]. Помимо приоритетного доступа URLLC авторы [5, 6] делают акцент на соблюдении качества обслуживания и eMBB, используя при этом методы стохастической геометрии и теории массового обслуживания. Особенность работы [7] заключается в применении системы массового обслуживания со случайными требованиями.

Данная работа служит продолжением исследований, описанных в [8–10]. В [8,9] авторы рассмотрели реализацию приоритетного доступа URLLC с прерыванием сессии eMBB, а в [10] — с приостановкой обслуживания eMBB. В данной статье, по сравнению с [8,9], предлагается еще и предварительное снижение скорости передачи сессии eMBB до ее прерывания. Статья организована следующим образом. В разд. 2 детально описана процедура приоритетного доступа узкополосного трафика URLLC и функционирования системы в целом, в разд. 3 построена соответствующая система массового обслуживания, а в завершающем разд. 4 приведены результаты численного анализа.

2 Приоритетный доступ узкополосного трафика URLLC

Рассмотрим совместную передачу узкополосного трафика URLLC и широкополосного трафика eMBB. Узкополосный трафик является приоритетным, и приоритет реализуется за счет снижения скорости передачи широкополосного трафика, а если у всех текущих сессий скорость становится минимально возможной, то за счет прерывания обслуживания одной из сессий. В табл. 1 приведено тезисное описание функционирования системы для разных вариантов возникающих событий.

Схема обслуживания построена исходя из соблюдения приоритета для узкополосного трафика, тем не менее настройка уровней скоростей для широкополосного трафика может способствовать максимизации средней скорости его передачи. Такая задача оптимизации сформулирована в следующем разделе после получения формул для расчета показателей эффективности как узкополосного, так и широкополосного трафика.

Таблица 1 События в системе

№ п/п	Условие	Результат
1. Поступление запроса на передачу узкополосного трафика URLLC		
а	Свободного ресурса для обеспечения скорости передачи узкополосного трафика достаточно	Запрос будет принят и передача трафика будет начата
б	Свободного ресурса для обеспечения скорости передачи узкополосного трафика недостаточно, но обслуживается хотя бы одна сессия широкополосного трафика не на минимальной скорости	Запрос будет принят за счет снижения скорости передачи случайно выбранной сессии широкополосного трафика из обслуживаемых на максимальной среди них скорости, и передача трафика будет начата
в	Свободного ресурса для обеспечения скорости передачи узкополосного трафика недостаточно, но все сессии широкополосного трафика обслуживаются на минимальной скорости	Запрос будет принят за счет прерывания обслуживания случайно выбранной сессии широкополосного трафика и передача трафика будет начата
г	Свободного ресурса для обеспечения скорости передачи узкополосного трафика недостаточно, а сессии широкополосного трафика отсутствуют	Запрос будет заблокирован
2. Поступление запроса на передачу широкополосного трафика eMBB		
а	Свободного ресурса для обеспечения максимальной скорости передачи широкополосного трафика достаточно	Запрос будет принят и передача трафика на максимальной скорости будет начата
б	Свободного ресурса для обеспечения максимальной скорости передачи широкополосного трафика недостаточно	Запрос будет заблокирован
3. Завершение передачи узкополосного трафика URLLC		
а	Передается хотя бы одна сессия широкополосного трафика не на максимальной скорости	Сессия будет завершена, а скорость передачи случайно выбранной сессии широкополосного трафика из обслуживаемых на минимальной среди них скорости будет повышена
б	Сессии широкополосного трафика передаются на максимальной скорости или отсутствуют	Сессия будет завершена, и выделенный для нее ресурс будет освобожден
4. Завершение передачи широкополосного трафика eMBB		
а	—	Сессия будет завершена, и выделенный для нее ресурс будет освобожден

3 Система массового обслуживания

Перейдем к построению математической модели для узкополосного и широкополосного трафика в виде системы массового обслуживания с приоритетным об-

служиванием узкополосного трафика — снижением и прерыванием скорости передачи. Предположим, что входящие потоки запросов на передачу узкополосного и широкополосного трафика — пуассоновские с интенсивностями λ_1/λ_2 , а время передачи распределено по экспоненциальному закону с параметрами μ_1/μ_2 . Пусть скорости передачи широкополосного трафика $b_1 > \dots > b_k > \dots > b_K$, тогда с учетом объема ресурса V и максимальной скорости передачи широкополосного трафика b_1 максимальное число обслуживаемых сессий широкополосного трафика составит $M = \lfloor V/b_1 \rfloor$.

Опишем функционирование системы при помощи случайного процесса $\mathbf{X}(t)$ с состояниями вида $\mathbf{x} = (n, m_1, \dots, m_K)$, где n — число обслуживаемых сессий узкополосного трафика; $\mathbf{m} = (m_1, \dots, m_K)$ — число обслуживаемых сессий широкополосного трафика. Тогда пространство состояний $\mathbf{X}(t)$ будет иметь вид:

$$\mathbf{X} = \left\{ (n, m_1, \dots, m_K) : n \geq 0, m_k \geq 0, k = 1, \dots, K, \right. \\ \left. \sum_{k=1}^K m_k \leq M, n + \sum_{k=1}^K b_k m_k \leq V \right\}.$$

В табл. 2 перечислены возможные интенсивности переходов между состоянием $\mathbf{x}$ и другими состояниями системы, нумерация строк соответствует нумерации в табл. 1. Вектор $\mathbf{e}_k$ — единичный вектор с 1 на k -м месте.

Исходя из описанных правил, записывается матрица интенсивностей переходов, и могут быть найдены стационарные вероятности $\pi(\mathbf{x})$, $\mathbf{x} \in \mathbf{X}$.

После нахождения стационарных вероятностей $\pi(\mathbf{x})$, $\mathbf{x} \in \mathbf{X}$, могут быть вычислены следующие показатели эффективности:

- вероятность блокировки (англ. blocking probability) узкополосного трафика URLLC

$$P_1^{\text{bl}} = \pi(V, 0, \dots, 0);$$

- вероятность снижения скорости передачи (англ. bit rate degradation probability) широкополосного трафика eMBB

$$P_2^{\text{dg}} = \sum_{k=1}^{K-1} \sum_{m_k=1}^M \frac{1}{m_k} \frac{\lambda_1}{\lambda_1 + n\mu_1 + \mu_2 \sum_{i=k}^K m_i} \times \\ \times \pi \left(V - \sum_{i=k}^K b_i m_i, 0, \dots, 0, m_k, \dots, m_K \right);$$

Таблица 2 Интенсивности переходов

№ п/п	Интенсивность события	Условие на $\mathbf{x}$	Состояние $\mathbf{x}'$
1а	λ_1	$n + 1 + \sum_{k=1}^K b_k m_k \leq V$	$(n + 1, \mathbf{m})$
1б	λ_1	$n + 1 + \sum_{k=1}^K b_k m_k > V,$ $\sum_{i=1}^{k-1} m_i = 0, m_k > 0,$ $k = 1, \dots, K - 1$	$(n + 1, \mathbf{m} - \mathbf{e}_k + \mathbf{e}_{k+1})$
1в	λ_1	$n + 1 + \sum_{k=1}^K b_k m_k > V,$ $\sum_{k=1}^{K-1} m_k = 0, m_K > 0$	$(n + 1, \mathbf{m} - \mathbf{e}_K)$
2а	λ_2	$n + \sum_{k=1}^K b_k m_k + b_1 \leq V$	$(n, \mathbf{m} + \mathbf{e}_1)$
3а	$n\mu_1$	$n > 0, \sum_{i=k+1}^K m_i = 0, m_k > 0,$ $k = 2, \dots, K$	$(n - 1, \mathbf{m} - \mathbf{e}_k + \mathbf{e}_{k-1})$
3б	$n\mu_1$	$n > 0, \sum_{k=2}^K m_k = 0$	$(n - 1, \mathbf{m})$
4а	$m_k \mu_2,$ $k = 1, \dots, K$	$m_k > 0$	$(n, \mathbf{m} - \mathbf{e}_k)$

– вероятность прерывания (*англ.* interruption probability) обслуживания широкополосного трафика eMBB

$$P_2^{\text{in}} = \sum_{m_K=1}^M \frac{1}{m_K} \frac{\lambda_1}{\lambda_1 + n\mu_1 + m_K\mu_2} \pi(V - b_K m_K, 0, \dots, 0, m_K);$$

– вероятность блокировки широкополосного трафика eMBB

$$P_2^{\text{bl}} = \sum_{\mathbf{x} \in B_2} \pi(n, m_1, \dots, m_L), \quad B_2 = \left\{ \mathbf{x} \in X : n + \sum_{k=1}^K b_k m_k + b_1 > V \right\};$$

- среднее число сессий узкополосного трафика URLLC

$$N_1 = \sum_{x \in X} n \pi(n, m_1, \dots, m_K);$$

- среднее число передаваемых на скорости b_k сессий широкополосного трафика eMBB

$$N_{2k} = \sum_{x \in X} m_k \pi(n, m_1, \dots, m_K), \quad k = 1, \dots, K;$$

- среднее число сессий широкополосного трафика eMBB

$$N_2 = \sum_{k=1}^K N_{2k};$$

- средняя скорость передачи широкополосного трафика eMBB

$$\bar{b} = \left[\sum_{x \in X} \sum_{k=1}^K b_k m_k \left(\sum_{k=1}^K m_k \right)^{-1} \pi(n, m_1, \dots, m_K) \right] \left[1 - \sum_{n=0}^V \pi(n, 0, \dots, 0) \right]^{-1}.$$

Вернемся к задаче максимизации средней скорости передачи широкополосного трафика eMBB по значениям уровней скорости и при ограничениях на вероятность блокировки и вероятность прерывания обслуживания широкополосного трафика eMBB. Она может быть записана в виде:

$$(L^*, b_1^*, \dots, b_K^*) = \arg \max_{L, b_1, \dots, b_K} \bar{b}(L, b_1, \dots, b_K),$$

$$b_l \in D = \{d_1, \dots, d_I\}, \quad k = 1, \dots, K,$$

$$b_1 > \dots > b_K \geq \hat{b},$$

$$P_2^{\text{bl}}(K, b_1, \dots, b_K) \leq \hat{P}_2^{\text{bl}},$$

$$P_2^{\text{in}}(K, b_1, \dots, b_K) \leq \hat{P}_2^{\text{in}}.$$

4 Пример численного анализа

Пример численного анализа проведен для следующих исходных данных: $M = 5$ — максимальное число обслуживаемых сессий eMBB, $K = 3$, $b_1 = 3$, $b_2 = 2$, $b_3 = 1$ — и двух сценариев. В первом случае рассматривается зависимость от интенсивности поступления запросов eMBB и $\lambda_1 = 5, 20, 100$, $\lambda_2 = 0-100$, а во втором — от интенсивности поступления URLLC-запросов и $\lambda_1 = 0-100$, $\lambda_2 = 10, 50, 100$. При этом для обоих вариантов среднее

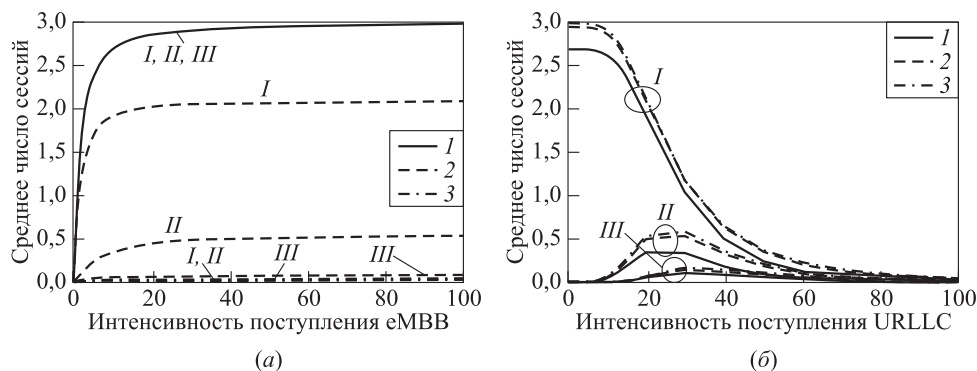


Рис. 1 Среднее число сессий eMBB; *I* — для 1-й скорости; *II* — для 2-й скорости; *III* — для 3-й скорости: (а) 1 — $\lambda_{\text{URLLC}} = 5$; 2 — 20; 3 — $\lambda_{\text{URLLC}} = 100$; (б) 1 — $\lambda_{\text{eMBB}} = 10$; 2 — 50; 3 — $\lambda_{\text{eMBB}} = 100$

время обслуживания фиксированное и равно: $\mu_1^{-1} = 3^{-1}$ и $\mu_1^{-1} = 1^{-1}$. На графиках рассмотрены такие показатели эффективности приоритетного доступа, как среднее число сессий eMBB (рис. 1) и вероятность прерывания обслуживания (рис. 2).

Графики, изображенные на рис. 1, демонстрируют изменения среднего числа сессий eMBB в системе, которые обслуживаются с разной скоростью: первая — с самой высокой, а третья — с самой низкой. При фиксированной интенсивности поступления запросов на передачу трафика URLLC наблюдается увеличение числа сессий eMBB с ростом интенсивности их поступления. При низкой интен-

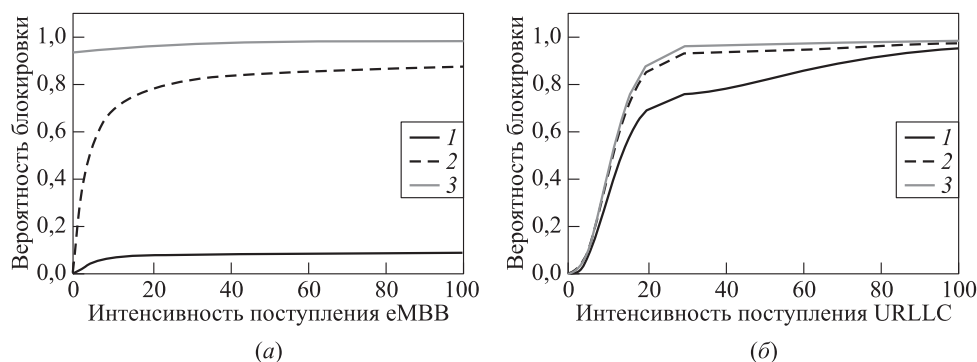


Рис. 2 Вероятность блокировки eMBB: (а) 1 — $\lambda_{\text{URLLC}} = 5$; 2 — 20; 3 — $\lambda_{\text{URLLC}} = 100$; (б) 1 — $\lambda_{\text{eMBB}} = 10$; 2 — 50; 3 — $\lambda_{\text{eMBB}} = 100$

сивности поступления запросов URLLC почти все сессии eMBB обслуживаются на максимальной скорости. Так, на рис. 1, а можно отметить тот факт, что при увеличении интенсивности поступления запросов URLLC сессии eMBB будут обслужены на минимальной скорости. Из рис. 1, б видно, что обслуживаемые на разных скоростях сессии eMBB ведут себя одинаково.

На рис. 2, а и 2, б видно, что с увеличением интенсивности поступления запросов URLLC повышается риск блокировки eMBB-сессий. При невысокой интенсивности поступления запросов URLLC вероятность блокировки eMBB мала, так как система справляется с нагрузкой от потока поступающих на обслуживание eMBB-запросов.

5 Заключение

Разработана математическая модель для узкополосного URLLC- и широкополосного eMBB-трафика в виде системы массового обслуживания с приоритетным обслуживанием узкополосного трафика — снижением и прерыванием скорости передачи. Предложены показатели эффективности приоритетного доступа — вероятность снижения скорости передачи, вероятность прерывания обслуживания, средняя скорость передачи широкополосного трафика.

В дальнейшем предполагается рассмотреть модель с учетом пространственного расположения устройств, генерирующих трафик URLLC и eMBB.

Авторы благодарят студента кафедры прикладной информатики и теории вероятностей РУДН Анастасию Агееву за помощь в проведении численного анализа.

Литература

1. Alsenwi M., Tran N. H., Bennis M., Bairagi A. K., Hong C. S. EMBB-URLLC resource slicing: A risk-sensitive approach // IEEE Commun. Lett., 2019. Vol. 23. Iss. 4. P. 740–743. doi: 10.1109/LCOMM.2019.2900044.
2. Dos Santos E. J., Souza R. D., Rebelatto J. L., Alves H. Network slicing for URLLC and eMBB with max-matching diversity channel allocation // IEEE Commun. Lett., 2020. Vol. 24. Iss. 3. P. 658–661. doi: 10.1109/LCOMM.2019.2959335.
3. De Cola T., Bisio I. QoS optimisation of eMBB services in converged 5G-Satellite networks // IEEE T. Veh. Technol., 2020. Vol. 69. Iss. 10. P. 12098–12110. doi: 10.1109/TVT.2020.3011963.
4. Gerasin I. S., Krasilov A. N., Khorov E. M. Dynamic multiplexing of URLLC traffic and eMBB traffic in an uplink using nonorthogonal multiple access // J. Commun. Technol. El., 2020. Vol. 65. Iss. 6. P. 750–755. doi: 10.1134/S1064226920060108.
5. Markova E., Moltchanov D., Pirmagomedov R., Ivanova D., Koucheryavy Y., Samouylov K. Prioritized service of URLLC traffic in industrial deployments of 5G NR

- systems // Distributed computer communication networks / Eds. V. M. Vishnevskiy, K. E. Samouylov, D. V. Kozyrev. — Lecture notes in computer science ser. — Springer, 2020. Vol. 12563. P. 497–509. doi: 10.1007/978-3-030-66471-8_38.
6. *Markova E., Moltchanov D., Pirmagomedov R., Ivanova D., Koucheryavy Y., Samouylov K.* Priority-based coexistence of eMBB and URLLC traffic in industrial 5G NR deployments // 12th Congress (International) on Ultra Modern Telecommunications and Control Systems and Workshops Proceedings. — Piscataway, NJ, USA: IEEE, 2020. 6 p. doi: 10.1109/ICUMT51630.2020.9222433.
7. *Sopin E., Begishev V., Moltchanov D., Samuylov A.* Resource queuing system with preemptive priority for performance analysis of 5G NR systems // Distributed computer and communication networks / Eds. V. M. Vishnevskiy, K. E. Samouylov, D. V. Kozyrev. — Lecture notes in computer science ser. — Springer, 2020. Vol. 12563. P. 87–99. doi: 10.1007/978-3-030-66471-8_8.
8. *Makeeva E., Polyakov N., Kharin P., Gudkova I.* Probability model for performance analysis of joint URLLC and eMBB transmission in 5G networks // Internet of things, smart spaces, and next generation networks and systems / Eds. O. Galinina, S. Andreev, S. Balandin, Y. Koucheryavy. — Lecture notes in computer science ser. — Springer, 2019. Vol. 11660. P. 635–648. doi: 10.1007/978-3-030-30859-9_55.
9. *Макеева Е. Д., Поляков Н. А., Харин П. А., Гудкова И. А.* Вероятностная модель для анализа характеристик совместной передачи трафика URLLC и eMBB в беспроводных сетях // Вестник Томского гос. ун-та. Управление, вычислительная техника и информатика, 2020. Т. 52. С. 33–42. doi: 10.17223/19988605/52/4.
10. *Харин П. А., Макеева Е. Д., Кочеткова И. А., Ефросинин Д. В., Шоргин С. Я.* Система массового обслуживания с орбитами для анализа совместного обслуживания трафика с малыми задержками URLLC и широкополосного доступа eMBB в беспроводных сетях пятого поколения // Информатика и её применения, 2020. Т. 14. Вып. 4. С. 17–24. doi: 10.14357/19922264200403.

Поступила в редакцию 03.08.21

MODEL FOR ANALYZING PRIORITY URLLC TRANSMISSION WITH eMBB BIT RATE DEGRADATION AND INTERRUPTIONS IN 5G NETWORKS

I. A. Kochetkova^{1,2}, A. I. Kushchazli¹, P. A. Kharin¹, and S. Ya. Shorgin²

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

²Federal Research Center "Computer Sciences and Control" of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation

Abstract: For 5G, the coexistence of two scenarios within the same network — ultrareliable low-latency communication (URLLC) and enhanced mobile broadband (eMBB) — is an actual task. For a scheme with the occupation of not the whole, but a part of the resource block, a model of coexistence of narrow-band URLLC traffic and broadband eMBB traffic in the form of a queuing system with priority access — reducing and interrupting the service of broadband traffic when narrow-band traffic arrives — is proposed. The authors formulate the problem of maximizing the average speed of broadband traffic transmission by the values of speed levels and with restrictions on the probability of blocking and the probability of interrupting broadband traffic service. The results of a numerical analysis of priority access efficiency indicators, such as the average number of eMBB sessions and probability of service interruption, are presented.

Keywords: 5G; eMBB; URLLC; queuing system; priority admission control; interruption; bit rate degradation

DOI: 10.14357/08696527210311

Acknowledgments

This paper has been supported by the RUDN University Strategic Academic Leadership Program. The reported study was funded by RFBR, project number 20-37-70079.

References

1. Alsenwi, M., N. H. Tran, M. Bennis, A. K. Bairagi, and C. S. Hong. 2019. EMBB-URLLC resource slicing: A risk-sensitive approach. *IEEE Commun. Lett.* 23(4):740–743. doi: 10.1109/LCOMM.2019.2900044.
2. Dos Santos, E. J., R. D. Souza, J. L. Rebelatto, and H. Alves. 2020. Network slicing for URLLC and eMBB with max-matching diversity channel allocation. *IEEE Commun. Lett.* 24(3):658–661. doi: 10.1109/LCOMM.2019.2959335.
3. De Cola, T., and I. Bisio. 2020. QoS optimisation of eMBB services in converged 5G-satellite networks. *IEEE T. Veh. Technol.* 69(10):12098–12110. doi: 10.1109/TVT.2020.3011963.

4. Gerasin, I. S., A. N. Krasilov, and E. M. Khorov. 2020. Dynamic multiplexing of URLLC traffic and eMBB traffic in an uplink using nonorthogonal multiple access. *J. Commun. Technol. El.* 65(6):750–755. doi: 10.1134/S1064226920060108.
5. Markova, E., D. Moltchanov, R. Pirmagomedov, D. Ivanova, Y. Koucheryavy, and K. Samouylov. 2020. Prioritized service of URLLC traffic in industrial deployments of 5G NR systems. *Distributed computer communication networks*. Eds. V. M. Vishnevskiy, K. E. Samouylov, and D. V. Kozyrev. Lecture notes in computer science ser. Springer. 12563:497–509. doi: 10.1007/978-3-030-66471-8_38.
6. Markova, E., D. Moltchanov, R. Pirmagomedov, D. Ivanova, Y. Koucheryavy, and K. Samouylov. 2020. Priority-based coexistence of eMBB and URLLC traffic in industrial 5G NR deployments. *12th Congress (International) on Ultra Modern Telecommunications and Control Systems and Workshops Proceedings*. Piscataway, NJ: IEEE. 6 p. doi: 10.1109/ICUMT51630.2020.9222433.
7. Sopin, E., V. Begishev, D. Moltchanov, and A. Samouylov. 2020. Resource queuing system with preemptive priority for performance analysis of 5G NR systems. *Distributed computer and communication networks*. Eds. V. M. Vishnevskiy, K. E. Samouylov, and D. V. Kozyrev. Lecture notes in computer science ser. Springer. 12563:87–99. doi: 10.1007/978-3-030-66471-8_8.
8. Makeeva, E., N. Polyakov, P. Kharin, and I. Gudkova. 2019. Probability model for performance analysis of joint URLLC and eMBB transmission in 5G networks. *Internet of things, smart spaces, and next generation networks and systems*. Eds. O. Galinina, S. Andreev, S. Balandin, and Y. Koucheryavy. Lecture notes in computer science ser. Springer. 11660:635–648. doi: 10.1007/978-3-030-30859-9_55.
9. Makeeva, E. D., N. A. Polyakov, P. A. Kharin, and I. A. Gudkova. 2020. Veroyatnostnaya model' dlya analiza kharakteristik sovmestnoy peredachi trafika URLLC i eMBB v besprovodnykh setyakh [Probability model for performance analysis of joint URLLC and eMBB transmission in 5G networks]. *Vestnik Tomskogo gos. un-ta. Upravlenie, vychislitel'naya tekhnika i informatika* [Tomsk State University J. Control and Computer Science] 52:33–42. doi: 10.17223/19988605/52/4.
10. Kharin, P. A., E. D. Makeeva, I. A. Kochetkova, D. V. Efrosinin, and S. Ya. Shorgin. 2020. Sistema massovogo obsluzhivaniya s orbitami dlya analiza sovmestnogo obsluzhivaniya trafika s malymi zaderzhkami URLLC i shirokopolosnogo dostupa eMBB v besprovodnykh setyakh pyatogo pokoleniya [Retrial queuing model for analyzing joint URLLC and eMBB transmission in 5G networks]. *Informatika i ee Primeneniya — Inform. Appl.* 14(4):17–24. doi: 10.14357/19922264200403.

Received August 3, 2021

Contributors

Kochetkova Irina A. (b. 1985) — Candidate of Science (PhD) in physics and mathematics, associate professor, 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

Sciences and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; gudkova-ia@rudn.ru

Kushchazli Anna I. (b. 1997) — PhD student, Peoples’ Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; aikushch@yandex.ru

Kharin Petr A. (b. 1993) — PhD student, Peoples’ Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; pxarin@mail.ru

Shorgin Sergey Ya. (b. 1952) — 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 119333, Russian Federation; sshorgin@ipiran.ru

ПРИМЕР ПРИМЕНЕНИЯ АППАРАТА НЕЙРОННЫХ СЕТЕЙ ПРИ НАЗНАЧЕНИИ МОДУЛЯЦИОННО-КОВОЙ СХЕМЫ ПЛАНИРОВЩИКОМ БАЗОВОЙ СТАНЦИИ СЕТИ 5G*

Е. В. Бобрикова¹, А. А. Платонова², Ю. В. Гайдамака³, С. Я. Шоргин⁴

Аннотация: Предложен метод назначения модуляционно-кодовой схемы (МКС, *англ.* Modulation Coding Scheme, MCS) планировщиком базовой станции (БС), основанный на предсказании значения отношения сигнал/интерференция на оборудовании подвижного пользователя на следующем временном такте по последовательности известных значений этого отношения в прошлом. Предсказание выполняется с помощью машинного обучения, для этого в работе построена и применена однослойная нейронная сеть для решения многопараметрической задачи оптимизации методом стохастического градиента. Обученная нейронная сеть по прогнозному значению отношения сигнал/интерференция позволяет планировщику корректно подобрать МКС для пользователя, тем самым обеспечив требуемый для предоставления услуги уровень качества передачи данных в радиоканале.

Ключевые слова: отношение сигнал/интерференция; машинное обучение; нейронная сеть

DOI: 10.14357/08696527210312

1 Введение

Одной из задач, которые необходимо решать планировщику БС сетей беспроводной связи LTE (long term evolution) и LTE-A (LTE advanced) пятого поколения (5G), является задача подбора МКС при передаче данных в радиоканале между БС/точкой доступа и пользователем. Корректное назначение МКС позволяет избежать повторной передачи данных, приводящей к неэффективному использованию ресурсов приема-передающего оборудования и снижению показателей качества предоставления услуги пользователю, в частности увеличению задержки предоставления услуги. В случае неподвижных пользователей задача выбора МКС не представляет сложности, поскольку МКС определяется

* Публикация выполнена при поддержке Программы стратегического академического лидерства РУДН и при финансовой поддержке РФФИ (проекты 19-07-00933 и 20-07-01064).

¹ Российский университет дружбы народов, bobrikova-ev@rudn.ru

² Российский университет дружбы народов, platonova-aa@rudn.ru

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

⁴ Федеральный исследовательский центр «Информатика и управление» Российской академии наук, sshorgin@ipiran.ru

значением индикатора качества радиоканала (ИКР, *англ.* Channel Quality Indicator, CQI), которое планировщик БС периодически получает от оборудования пользователя [1]. Индикатор качества радиоканала зависит от значения отношения сигнал/интерференция (*англ.* Signal to Interference and Noise Ratio, SINR) на оборудовании пользователя, на которое влияют технология радиодоступа, расстояние между приемо-передающими устройствами, среда распространения радиосигнала, включая наличие экранов и блокаторов сигнала, мощность интерферирующих передатчиков и пр. При перемещении подвижного пользователя его геопозиция, расстояние до БС и параметры радиоканала меняются, после чего приемник фиксирует и отправляет БС новое значение ИКР, в соответствии с которым планировщик БС должен скорректировать МКС для передачи от БС. Таким образом, появляется задача прогнозирования значения ИКР подвижного пользователя, что должно позволить планировщику уйти от неэффективной схемы подбора МКС по умолчанию, когда при получении ИКР ниже требуемого для предоставления услуги планировщик перебирает последовательно соседние МКС, пока не получит от оборудования пользователя значение ИКР, приемлемое для передачи данных с надлежащим качеством. В статье предлагается подход к решению этой задачи с применением аппарата нейронных сетей [2].

2 Системная модель и математическая постановка задачи

Исследуется системная модель, состоящая из одной соты сети беспроводной связи, по которой в зоне покрытия БС перемещаются пользователи. Считаем модель движения пользователя заданной, тогда перемещение произвольного пользователя в дискретном времени описывается траекторией, которая представляет собой конечную последовательность геопозиций, задаваемых координатами пользователя на соседних временных тактах. Поскольку координаты однозначно определяют расстояние между БС и оборудованием пользователя, то траектория перемещения пользователя соответствует последовательности значений отношения сигнал/интерференция на соседних тактах, а также последовательности зависящих от него значений ИКР и МКС, которые связаны с отношением сигнал/интерференция, как это показано в таблице [1].

Задача планировщика БС состоит в прогнозировании МКС для произвольного пользователя по известной траектории его движения в прошлом. Для решения этой задачи в статье использован один из наиболее популярных методов машинного обучения — нейронные сети. Подход с использованием нейронной сети позволяет по заданной обучающей выборке — траектории движения пользователя на предыдущих тактах — предсказывать ответ, представляющий собой значение исследуемой метрики пользователя на следующем такте. Такой метрикой может быть не только МКС для назначения планировщиком БС, но и геопозиция пользователя, принимаемый базовой станцией ИКР или отношение сигнал/интерференция на оборудовании пользователя. В качестве математиче-

ской модели для предсказания выбрана однослойная нейронная сеть — однослойный перцептрон [3], который в результате решения многопараметрической задачи оптимизации по заданной на входе выборке обучается выдавать на выходе приближенный ответ, удовлетворяющий заданным критериям. Исследования показали, что бóльшая точность предсказаний достигается, если в обучающей выборке в качестве признаков объектов и ответов брать значения отношения сигнал/интерференция на оборудовании пользователя, что иллюстрируется при построении математической модели, а также в численном примере в разд. 3.

Переходя к терминологии машинного обучения, для описанной системной модели определим объект $\mathbf{x}_i$ как набор n последовательных тактов времени, а признаками объекта $\mathbf{x}_i$: $f_1(\mathbf{x}_i), \dots, f_n(\mathbf{x}_i)$ будем считать заданные значения отношения сигнал/интерференция, соответствующие четырем тактам времени объекта $\mathbf{x}_i$, т. е. $n = 4$ — число признаков. Обозначим вектор признаков $\mathbf{f}(\mathbf{x}_i) = (f_1(\mathbf{x}_i), \dots, f_n(\mathbf{x}_i))$. В качестве ответа y_i рассматривается известное значение отношения сигнал/интерференция на следующем пятом такте. Заметим, что число ответов равно числу объектов, пара объект и ответ $(\mathbf{x}_i, y_i)$ называется прецедентом, множество пар $X^l = (\mathbf{x}_i, y_i)_{i=1}^l$ — обучающей выборкой, состоящей из l объектов.

В качестве приближающего алгоритма выберем линейную модель

$$a(\mathbf{x}, \mathbf{w}) = \langle \mathbf{w}, \mathbf{f}(\mathbf{x}) \rangle = \sum_{j=1}^n w_j f_j(\mathbf{x}), \quad (1)$$

где $\mathbf{x}$ — объект; w_j — неизвестные веса признаков; n — число весов; $j = 1, \dots, n$.

Цель исследования — обучение алгоритма a , т. е. подбор оптимального параметра $\mathbf{w}$ модели по обучающей выборке X^l . Оптимальным параметром $\mathbf{w}$ модели считается параметр, доставляющий минимальное значение функционалу качества $Q(a, X^l)$ алгоритма a на выборке X^l , т. е.

$$Q(a, X^l) = \frac{1}{l} \sum_{i=1}^l \mathcal{L}(a, y_i) \rightarrow \min_{\mathbf{w}}. \quad (2)$$

Здесь функция

$$\mathcal{L}(a, y_i) = (a(\mathbf{x}_i, \mathbf{w}) - y_i)^2 \quad (3)$$

носит название функции потерь и отражает точность приближения на объекте $\mathbf{x}_i$.

Соответствие отношения сигнал/интерференция, индикатора качества радиоканала и МКС

SINR, дБ	ИКР	МКС
−6,7	1	QPSK
−4,7	2	QPSK
−2,3	3	QPSK
0,2	4	QPSK
2,4	5	QPSK
4,3	6	QPSK
5,9	7	16QAM
8,1	8	16QAM
10,3	9	16QAM
11,7	10	64QAM
14,1	11	64QAM
16,3	12	64QAM
18,7	13	64QAM
21,0	14	64QAM
22,7	15	64QAM

Для численной оптимизации функционала (2) в работе использован метод стохастического градиента [2, 4]. Описанный процесс подбора оптимального параметра $\mathbf{w}$ по обучающей выборке $X^l = (\mathbf{x}_i, y_i)_{i=1}^l$ представляет собой работу простейшей однослойной нейронной сети. Для объекта $\mathbf{x}_i$ результат применения линейной функции $a(\mathbf{x}_i, \mathbf{w})$ будем называть предсказанным значением отношения сигнал/интерференция на объекте $\mathbf{x}_i$, или просто предсказанием на объекте $\mathbf{x}_i$. Заметим, что $a(\mathbf{x}_i, \mathbf{w})$ — приближение известного значения y_i неизвестной функции $y = y(\mathbf{x})$ на объекте $\mathbf{x}_i$.

3 Пример применения аппарата нейронных сетей для прогнозирования отношения сигнал/интерференция

Как показано в разд. 2, математическая модель в виде нейронной сети для прогнозирования отношения сигнал/интерференция строится на основе информации о модели перемещения пользователя, которая предполагается заданной: выборка (data set) траекторий перемещения может быть либо собрана по результатам наблюдений, либо сгенерирована, если модель перемещения задана в аналитическом виде. Иллюстрацию обучения нейронной сети для прогнозирования отношения сигнал/интерференция в сети LTE проведем для примера выборки траекторий перемещения пользователей, сгенерированной в соответствии с моделью движения «Случайное блуждание по сетке» (Grid Random Walk) [5]. Методом Монте-Карло сгенерированы 100 траекторий разной длины, каждая представляет собой последовательность геопозиций пешехода, движущегося по бульвару через зону покрытия БС, состоящую из $M = 112\,320$ квадратных ячеек. Далее для технологии LTE согласно [1, 5] из траекторий получены соответствующие последовательности значений отношения сигнал/интерференция, а из них — «пятерки», которые составили обучающую выборку длины $l = 53\,638$. Прецеденты, составляющие обучающую выборку, представляют собой пары «объект–ответ», где объектом является набор четырех последовательных тактов времени, для каждого объекта задан вектор признаков длины $n = 4$, причем признаки и ответы — это значения отношения сигнал/интерференция.

Например, последовательность значений отношения сигнал/интерференция $\dots 11,0275 \rightarrow 11,0275 \rightarrow 10,98056 \rightarrow 10,92848 \rightarrow 10,88471 \rightarrow 10,88098 \rightarrow 10,91089 \rightarrow 10,89156 \dots$ дает обучающую выборку со следующими парами «объект–ответ»:

$(11,0275, 11,0275, 10,98056, 10,92848) \rightarrow 10,88471;$
 $(11,0275, 10,98056, 10,92848, 10,88471) \rightarrow 10,88098;$
 $(10,98056, 10,92848, 10,88471, 10,88098) \rightarrow 10,91089;$
 $(10,92848, 10,88471, 10,88098, 10,91089) \rightarrow 10,89156$

и т. д.

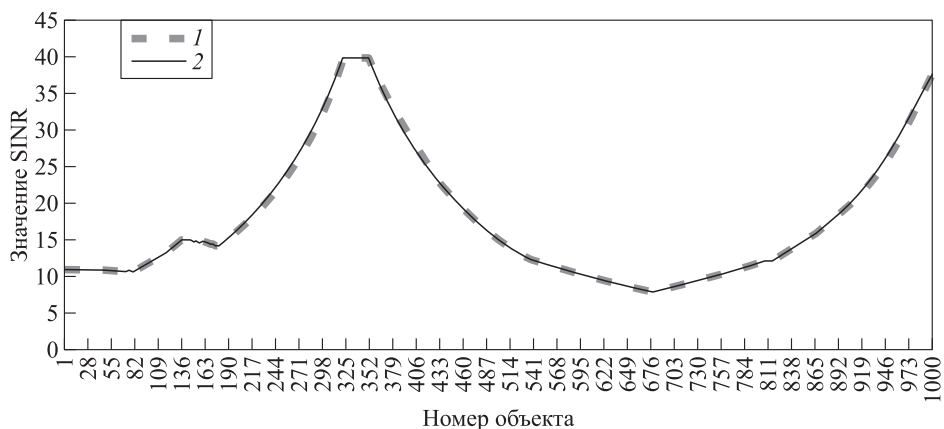


Рис. 1 Предсказанные значения SINR a (1) и ответы y_i (2) на первых 1000 объектах обучающей выборки

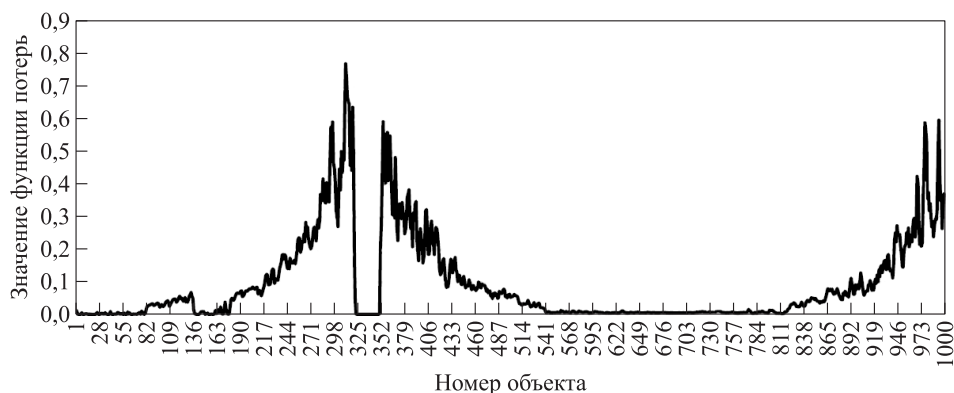


Рис. 2 Функция потерь $\mathcal{L}(a, y_i)$ на первых 1000 объектах обучающей выборки

Начальные значения весов признаков для линейной модели (1) выбраны $\mathbf{w} = (0,25, 0,25, 0,25, 0,25)$, градиентный шаг (темпер обучения) $h = 10^{-4}$ [6].

Далее представлены результаты предсказания отношения сигнал/интерференция на следующем 5-м такте по информации о значениях этого отношения на оборудовании пользователя на четырех последовательных предыдущих тактах с помощью построенной нейронной сети для обучающей выборки из 53 638 прецедентов. На рис. 1 штриховой линией показаны предсказания $a(\mathbf{x}_i, \mathbf{w})$ (1), сплошной линией — ответы y_i для первых 1000 объектов обучающей выборки. Рисунок 2 иллюстрирует изменение функции потерь $\mathcal{L}(a, y_i)$ (3). Критери-

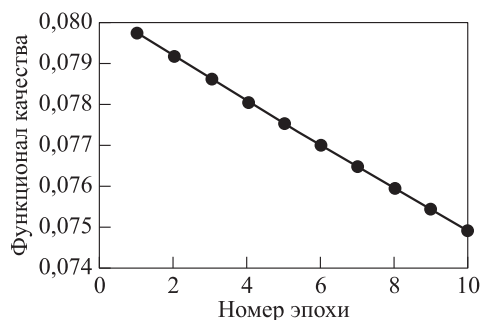


Рис. 3 Изменение функционала качества $Q(a, X^l)$

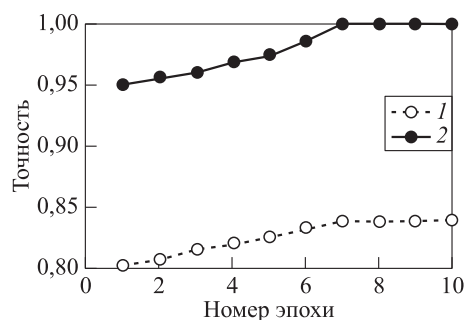


Рис. 4 Изменение точности при $\mathcal{L}(a, y_i) < 10^{-1}$ (1) и при $\mathcal{L}(a, y_i) < 1$ (2)

ем остановка обучения нейронной сети может служить стабилизация одного из следующих параметров модели: весов, функционала качества, точности.

Совпадение предсказанных значений и ответов на рис. 1 подтверждают показатели качества предсказания — функционал качества $Q(a, X^l)$ (2) на рис. 3 и точность, определяемая как отношение числа совпадений предсказанного значения отношения сигнал/интерференция $a(\mathbf{x}_i, \mathbf{w})$ и известного ответа y_i к общему числу l объектов обучающей выборки X^l . На рис. 4 показаны варианты, когда совпадением считаем $\mathcal{L}(a, y_i) < 10^{-1}$ (1) и когда $\mathcal{L}(a, y_i) < 1$ (2).

Для обучения нейронной сети потребовалось 10 эпох, при этом на рис. 4 видно, что максимальная точность для обоих вариантов достигнута уже на 7-й эпохе. Здесь эпоха — это одна итерация метода стохастического градиента по всем объектам обучающей выборки. По завершении обучения нейронной сети значения весов составили $\mathbf{w} = (0,2490; 0,2500; 0,2505; 0,2505)$.

Верификация построенной модели на тестовой выборке из 13 409 прецедентов показала точность предсказания 83,80% для случая $\mathcal{L}(a, y_i) < 10^{-1}$ и точность предсказания 99,96% для случая $\mathcal{L}(a, y_i) < 1$. Согласно таблице, отличие значений отношения сигнал/интерференция для соседних МКС составляет около 2 дБ, поэтому второй случай, достигающий высокой точности, обеспечивает корректное предсказание МКС с помощью построенной нейронной сети. Заметим, что, зная отношение сигнал/интерференция, необходимую МКС планировщик может однозначно определить, например, по таблице.

4 Заключение

Предложенный в статье метод назначения МКС планировщиком БС, основанный на предсказании значения отношения сигнал/интерференция с использованием аппарата нейронных сетей, состоит из двух этапов. На первом этапе для заданной модели перемещения пользователей строится и обучается модель

однослойной нейронной сети, которая позволяет предсказывать значение отношения сигнал/интерференция на оборудовании подвижного пользователя на следующем временном такте по последовательности известных значений этого отношения в прошлом. На втором этапе по предсказанному значению отношения сигнал/интерференция определяется МКС, требуемая для передачи данных пользователю при предоставлении услуги с надлежащим уровнем качества. Отметим, что задача оптимизации, которая решается при обучении сети, является многопараметрической, при этом параметры нейронной сети, а именно: число признаков объекта обучающей выборки, объем выборки, требуемый для достижения заданной точности предсказания, и др. — существенным образом зависят от модели перемещения пользователей. В дальнейшем для решения задачи предсказания отношения сигнал/интерференция планируется как развивать методы обучения с учителем, например применить многослойную нейронную сеть, так и использовать метод обучения с подкреплением для прогнозирования при отсутствии обучающей выборки.

Авторы благодарят магистров кафедры прикладной информатики и теории вероятностей РУДН Э. М. Хайрова и А. А. Мамонова за подготовку обучающей выборки для численного эксперимента.

Литература

1. *Ghosh A., Ratasuk R.* Essentials of LTE and LTE-A. — Cambridge, U.K.: Cambridge University Press, 2011. 264 p. doi: 10.1017/CBO9780511997082.
2. *Воронцов К. В.* Математические методы обучения по прецедентам (теория обучения машин). <http://www.machinelearning.ru/wiki/images/6/6d/Voron-ML-1.pdf>.
3. *Аверкин А. Н., Гаазе-Рапопорт М. Г., Поспелов Д. А.* Толковый словарь по искусственному интеллекту. — М.: Радио и связь, 1992. 256 с.
4. *Траск Э.* Грокаем глубокое обучение / Пер. с англ. — СПб: Питер, 2019. 352 с. (*Trask A. W.* Grokking deep learning. — New York, NY, USA: Manning Publications, 2019. 336 p.)
5. *Гайдамака Ю. В., Самуйлов К. Е., Шоргин С. Я.* Метод моделирования характеристик интерференции при прямом взаимодействии перемещающихся устройств в гетерогенной беспроводной сети пятого поколения // Информатика и её применения, 2017. Т. 11. Вып. 4. С. 2–9.
6. *Бобрикова Е. В., Платонова А. А., Ярцева И. С., Хайров Э. М.* К задаче предсказания SINR в беспроводной сети с подвижными пользователями с помощью аппарата нейронных сетей // Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем. — М.: РУДН, 2021. С. 30–32.

Поступила в редакцию 26.07.21

AN EXAMPLE OF NEURAL NETWORK USAGE FOR ASSIGNING A MODULATION-CODE SCHEME TO A 5G BASE STATION SCHEDULER

E. V. Bobrikova¹, A. A. Platonova¹, Yu. V. Gaidamaka^{1,2}, and S. Ya. Shorgin²

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

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

Abstract: The article proposes a method for assigning a modulation-code scheme by a base station scheduler based on predicting the value of the signal-to-interference ratio on the mobile user's equipment at the next time slot from a sequence of known values of this ratio in the past. For prediction, a model of a single-layer neural network is built in the work, by the example of which a machine learning process is shown for solving a multiparametric optimization problem using the stochastic gradient method. The trained neural network for the predicted value of the signal/interference ratio allows the scheduler to correctly select the modulation-code scheme for the user, thereby ensuring the level of quality of data transmission in the radio channel required for the provision of the service.

Keywords: SINR; machine learning; neural network

DOI: 10.14357/08696527210312

Acknowledgments

This paper has been supported by the RUDN University Strategic Academic Leadership Program and funded by the Russian Foundation for Basic Research according to the research projects Nos. 19-07-00933 and 20-07-01064.

References

1. Ghosh, A., and R. Ratasuk. 2011. *Essentials of LTE and LTE-A*. Cambridge, U.K.: Cambridge University Press. 264 p. doi: 10.1017/CBO9780511997082.
2. Matematicheskie metody obucheniya po pretsedentam (teoriya obucheniya mashin) [Mathematical teaching methods by precedents (machine learning theory)]. Available at: <http://www.machinelearning.ru/wiki/images/6/6d/Voron-ML-1.pdf> (accessed September 3, 2021).
3. Averkin, A. N., M. G. Gaaze-Rapoport, and D. A. Pospelov. 1992. *Tolkovyy slovar' po iskusstvennomu intellektu* [Dictionary of artificial intelligence]. Moscow: Radio i svyaz'. 256 p.
4. Trask, A. W. 2019. *Grokking deep learning*. New York, NY: Manning Publications. 336 p.

5. Gaidamaka, Yu. V., K. E. Samouylov, and S. Ya. Shorgin. 2017. Metod modelirovaniya kharakteristik interferentsii pri pryamom vzaimodeystvii peremeshchayushchikhsya ustroystv v geterogennoy besprovodnoy seti pyatogo pokoleniya [A method for modeling interference characteristics during direct interaction of roaming devices in a heterogeneous fifth generation wireless network]. *Informatika i ee Primeneniya — Inform. Appl.* 11(4):2–9.
6. Bobrikova, E., A. Platonova, I. Yartseva, and E. Khairov. 2021. K zadache predskazaniya SINR v besprovodnoy seti s podvizhnymi pol'zovatelyami s pomoshch'yu apparata neyronnykh setey [To the problem of predicting SINR in a wireless network with mobile users with a neural network apparatus]. *Informatsionno-telekommunikatsionnye tekhnologii i matematicheskoe modelirovanie vysokotekhnologichnykh sistem* [Information and telecommunication technologies and mathematical modeling of high-tech systems]. RUDN University Publ. Moscow. 30–32.

Received July 26, 2021

Contributors

Bobrikova Ekaterina V. (b. 1976) – Candidate of Science (PhD) in physics and mathematics, associate professor, Department of Applied Probability and Informatics, Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; bobrikova-ev@rudn.ru

Platonova Anna A. (b. 1996) — PhD student, Department of Applied Probability and Informatics, Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; platonova-aa@rudn.ru

Gaidamaka Yuliya V. (p. 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, 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; gaydamaka-yuv@rudn.ru

Shorgin Sergey Ya. (b. 1952) — 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; sshorgin@ipiran.ru

ЭКСПЕРТНАЯ ОЦЕНКА МАШИННОГО ПЕРЕВОДА: КЛАССИФИКАЦИЯ ОШИБОК*

А. Ю. Егорова¹, И. М. Зацман², В. А. Нуриев³

Аннотация: Рассматривается оценка качества машинного перевода (МП) с применением классификации его ошибок. Классификация включает рубрики общезыковых ошибок на уровне грамматики, лексики, пунктуации и др., а также категорийно обусловленных ошибок, которые связаны с конкретным типом языковых единиц, выбранным для оценки качества перевода в исследуемой системе МП. Оценка качества выполняется в процессе лингвистического аннотирования МП. Если при аннотировании эксперт обнаруживает ошибки в переводе, то в аннотацию включаются рубрики для типологизации обнаруженных ошибок. Простановка рубрик дает возможность вычислить относительные частотности ошибок МП в рамках всего массива тестовых предложений, отобранных для перевода и оценки его качества. Основная цель статьи состоит в описании предлагаемой классификации общезыковых и категорийно обусловленных ошибок. Ее главное отличие от существующих классификаций (типологий) ошибок заключается в том, что она нацелена на обеспечение интервальной оценки для систем МП, качество работы которых изменяется во времени. Рубрики предлагаемой классификации дают возможность экспертам фиксировать в МП как улучшение, так и снижение качества на заданном интервале времени.

Ключевые слова: машинный перевод; оценка качества; классификация ошибок; общезыковые ошибки; категорийно обусловленные ошибки; лингвистическое аннотирование; интервальная оценка

DOI: 10.14357/08696527210313

1 Введение

За последние несколько лет были опубликованы авторитетные коллективные монографии, посвященные автоматической и экспертной оценке качества МП [1–3], что свидетельствует об актуальности этой тематики. Для экспертной оценки качества с привлечением специалистов (лингвистов, переводчиков)

* Работа выполнена в Федеральном исследовательском центре «Информатика и управление» Российской академии наук с использованием ЦКП «Информатика» ФИЦ ИУ РАН.

¹ Федеральное исследовательское учреждение «Информатика и управление» Российской академии наук, ann.shurova@gmail.com

² Федеральное исследовательское учреждение «Информатика и управление» Российской академии наук, izatsman@yandex.ru

³ Федеральное исследовательское учреждение «Информатика и управление» Российской академии наук, nurieff.v@gmail.com

используются несколько способов [4]: оценивание в соответствии с критериями точности и естественности, ранжирование переводов, прямое оценивание, оценка с учетом коэффициента редактирования перевода человеком, лингвистическое аннотирование перевода с применением классификации, включающей рубрики ошибок перевода. Обзор классификаций/типологий ошибок, которые применяются в процессе экспертной оценки качества, представлен в работе [5].

В данной статье рассматривается классификация, построенная с целью рубрикации ошибок перевода и используемая при лингвистическом аннотировании тестовых предложений и их переводов. Данная классификация включает рубрики общезыковых ошибок на уровне грамматики, лексики, пунктуации и др., а также категорийно обусловленных ошибок, которые связаны с конкретным типом языковых единиц, выбранным для оценки качества в исследуемой системе МП. Если в процессе аннотирования эксперт обнаруживает ошибки в переводе, то в аннотацию тестового предложения и его перевода включаются рубрики для типологизации обнаруженных ошибок.

Простановка рубрик дает возможность вычислять относительные частотности ошибок МП в рамках всего массива тестовых предложений, отобранных для перевода и оценки его качества. Основная цель статьи состоит в описании новой классификации, включающей рубрики ошибок. Ее главное отличие от существующих классификаций/типологий ошибок перевода заключается в том, что она нацелена на обеспечение интервальной оценки для систем МП, качество работы которых изменяется во времени. Рубрики предлагаемой классификации дают возможность экспертам фиксировать в повторных переводах одних и тех же тестовых предложений как повышение, так и снижение качества перевода во времени. Экспертный анализ повторных переводов на заданном интервале времени дает возможность исследовать *темпоральную стабильность* МП.

Отметим, что способ лингвистического аннотирования с использованием классификации ошибок применяется, в числе прочих, группой ученых из Дублина: Sheila Castilho, Pintu Lohar, Joss Moorkens, Maja Popović, Andy Way и их соавторами (см., например, [6–10]). Предметом их анализа являются ошибки МП, выявленные в некоторый момент времени, но не темпоральная стабильность МП на интервале времени.

2 Экспериментальные данные

В статье для описания функциональных возможностей предлагаемой классификации используются предложения на русском языке и их МП на французский язык, полученные с помощью системы нейронного МП (НМП) Google.Translate. Референтные (выполненные человеком-специалистом) переводы приводятся для наглядного представления о качестве МП (см. примеры в табл. 1).

Используемые предложения на русском языке содержат двухкомпонентные коннекторы. Коннектор — это языковая единица, функция которой состоит

Таблица 1 Примеры двух предложений из книги С. Алексиевич «Время секунд хэнд», их МП от 02 августа 2021 г. в 18:00 и референтные переводы, выполненные Софи Бенеш

Исходный текст	Машинный перевод	Референтный перевод
Как живем, так и умираем. . .	Comme nous vivons, nous mourons. . .	On meurt comme on a vécu. . .
Я и в церковь хожу, и крестик ношу, а счастья как не было, так и нет.	Je vais à l'église et porte une croix, mais il n'y avait pas de bonheur, donc il n'y a pas de bonheur.	Je vais à l'église maintenant, et je porte une petite croix, mais du bonheur, j'en ai pas plus qu'avant.

в выражении логико-семантического отношения, существующего между соединенными с ее помощью частями текстового фрагмента [11, с. 17]. Коннектор считается двухкомпонентным, если его части вводят два разных текстовых фрагмента [12, с. 175] (см. два предложения с двухкомпонентным коннектором **как (расстояние), так и** в табл. 1). Выбор этого типа языковых единиц для оценки качества перевода обусловлен тем, что системы МП далеко не всегда опознают двухкомпонентные коннекторы как единую языковую единицу и переводят их как две разные единицы или только ее первый компонент, что приводит к ошибкам при переводе всего предложения (табл. 2). Подробнее о логико-семантических отношениях и коннекторах см. в [13].

Для проверки функциональных возможностей рассматриваемой классификации по рубрицированию общезыковых и категорийно обусловленных ошибок использовался массив из 250 русскоязычных предложений, каждое из которых содержало двухкомпонентный коннектор, в том числе **как (расстояние), так и** [14–17]. Источником предложений служил Национальный корпус русского языка [18]. Каждое такое предложение в течение года было многократно (12 раз) переведено на французский язык с помощью системы НМП Google.Translate и затем проаннотировано вместе с его МП.

Многократность перевода каждого предложения массива дает возможность исследовать темпоральную стабильность МП на заданном интервале времени. Для формирования и хранения аннотаций многократных переводов использовалась надкорпусная база данных (НБД) коннекторов [19–21]. В НБД каждая аннотация (см. табл. 2) состоит из четырех частей:

- (1) исходное предложение (или его фрагмент) с двухкомпонентным коннектором на русском языке;
- (2) рубрики классификации, присвоенные исходному тексту;
- (3) перевод предложения (или его фрагмента) на французский, выполненный с помощью системы НМП Google.Translate;
- (4) рубрики классификации, присвоенные переводу.

Таблица 2 Пример аннотации перевода предложения с двухкомпонентным коннектором

Исходный текстовый фрагмент	Рубрики исходного фрагмента	Перевод	Рубрики перевода
Я и в церковь хожу, и крестик ношу, а счастья как не было, так и нет. (Светлана Алексиевич. Время секунд хэнд, 2013)	как так и <аналогия> <CNT p CNT q> <CNT> <SubCNT> <Дистант>	Je vais à l'église et porte une croix, mais comme il n'y avait pas de bonheur, je ne le fais pas. (Перевод выполнен с помощью системы Google.Translate 03.03.2019 в 17:17)	comme <ErrorSemant> <ErrorCNT>

Примечание. В аннотации используются рубрики двух фасетных классификаций, описание которых дано в следующем разделе:

- (1) рубрика «аналогия» (относится к фасету «Отношения») говорит о том, что коннектор **как (расстояние)**, **так и** выражает логико-семантическое отношение аналогии;
- (2) рубрика «CNT p CNT q» (относится к фасету «Порядок») говорит о том, что элементы двухкомпонентного коннектора находятся в каждом из соединяемых текстовых компонентов **p** (не было) и **q** (нет);
- (3) рубрика «CNT» (относится к фасету «Статус») говорит о том, что аннотация сформирована для всего коннектора, а не для отдельных составляющих его компонентов;
- (4) рубрика «SubCNT» (относится к фасету «Статус») говорит о том, что коннектор **как (расстояние)**, **так и** является встроенным (вложенным) по отношению к встраиваемому коннектору **a**;
- (5) рубрика «Дистант» (относится к фасету «Расположение») говорит о том, что два компонента коннектора разделены текстом «не было»;
- (6) рубрика «ErrorSemant» (относится к фасету «Ошибки») говорит о том, что в переводе есть искажение смысла («je ne le fais pas» переводится как «я этого не делаю»);
- (7) рубрика «ErrorCNT» (относится к фасету «Ошибки») говорит о том, что при переводе коннектора **как (расстояние)**, **так и** допущена семантическая ошибка (французский коннектор **comme** в данном контексте выражает логико-семантическое отношение причины и переводится как «потому как»).

Сравнение результатов МП в табл. 1 и 2 говорит о том, что одно и то же предложение в разные моменты времени может быть переведено системой МП по-разному. Более того, в проведенном эксперименте доля предложений, переводы которых не менялись в течение одного года, была равна всего 2% [17].

По итогам аннотирования многократного МП в НБД был получен массив экспериментальных данных, который использовался для исследования темпоральной стабильности МП и интервальной оценки его качества. В НБД для каждого из 250 тестовых предложений была зафиксирована серия из 12 версий МП и, соответственно, сформировано 12 аннотаций. Таким образом, всего в НБД было сформировано 3000 аннотаций с рубриками рассматриваемой классификации.

3 Фасетная классификация

Рубрики, представленные в столбцах 2 и 4 табл. 2, являются элементами классификаций, основанных на фасетном принципе [22] — одном из наиболее продуктивных принципов при многоаспектном представлении знаний [23]. Формирование (выделение) фасетов классификации и пополнение их рубриками в рамках проведенного эксперимента осуществлялось в процессе лингвистического аннотирования [24] исходных текстовых фрагментов с двухкомпонентным коннектором и их МП.

При аннотировании были задействованы рубрики 8-фасетной классификации для заполнения 2-го столбца и 10-фасетной классификации для 4-го столбца аннотаций (см. табл. 2). При этом в проведенном эксперименте были выделены основные фасеты в обеих классификациях. Таким образом, 8-фасетная классификация включала 1 основной и 7 дополнительных фасетов, а 10-фасетная — 1 основной и 9 дополнительных фасетов.

После завершения эксперимента и определения рубрик верхний (первый) уровень в 8-фасетной классификации включал 7 рубрик, а в 10-фасетной классификации он включал 9 рубрик. Названия рубрик верхних уровней, а также число рубрик второго уровня приведены в табл. 3 и 4 соответственно (числовые значения в таблицах приводятся по состоянию на 21.07.2021).

Наряду с рубриками основных фасетов в процессе аннотирования использовались рубрики и неосновных (дополнительных) фасетов обеих классификаций. В эксперименте были задействованы 7 неосновных фасетов в 8-фасетной классификации и 9 — в 10-фасетной. Названия дополнительных фасетов в обеих классификациях, а также число их рубрик, приведены в табл. 5 и 6 соответственно.

Ключевую роль при интервальной оценке темпоральной стабильности систем МП играет выявление допущенных ими ошибок и определение вида этих ошибок. В табл. 7 представлена третья версия восьмого дополнительного фасета «Ошибки» 10-фасетной классификации, разработанная в ходе эксперимента.

Таблица 3 Основной фасет 8-фасетной классификации

№	Рубрика первого уровня	Число рубрик второго уровня
1	Одноэлементные коннекторы	68
2	Многоэлементные	719
3	Двухкомпонентные	595
4	Многокомпонентные	119
5	Комбинированные	7
6	Пунктуация (в функции коннектора)	2 (двоеточие и тире)
7	Зеро (логико-семантическое отношение в исходном тексте есть, но нет коннектора как показателя отношения)	3

Таблица 4 Основной фасет 10-фасетной классификации

№	Рубрика первого уровня	Число рубрик второго уровня
1	Одноэлементные коннекторы	49
2	Многоэлементные	629
3	Двухкомпонентные	396
4	Многокомпонентные	97
5	Комбинированные	119
6	Пунктуация (в функции коннектора)	2 (двоеточие и тире)
7	Отрицание	223
8	Zero (в переводе отсутствует коннектор для передачи логико-семантического отношения исходного текста)	2
9	Еггог	352

Фасет содержит 19 рубрик ошибок перевода, не считая рубрики NoErgog (отсутствие ошибок в МП). Все 19 рубрик фасета «Ошибки» делятся на две группы. Первая характеризует языковые ошибки, куда входят, например, ошибки на уровне грамматики, лексики, пунктуации и др. (в табл. 7 это строки 1–11). Вторая — категорийно обусловленные ошибки (в табл. 7 это строки 12–19), которые связаны с конкретным типом языковых единиц, выбранным для проведения эксперимента (здесь это коннекторы русского языка).

По завершении аннотирования всех 12 версий перевода одного предложения аннотатор переходит к интервальной оценке темпоральной стабильности его МП. Такая оценка предполагает выявление изменений (повышение/сниже-

Таблица 5 Названия 7 дополнительных фасетов 8-фасетной классификации

№	Фасет	Число рубрик
1	Отношения	57
2	Комбинированные отношения	9
3	Структура	11
4	Позиция	3
5	Порядок	9
6	Статус	5
7	Расположение	2

Таблица 6 Названия 9 дополнительных фасетов 10-фасетной классификации

№	Фасет	Число рубрик
1	Отношения	56
2	Комбинированные отношения	3
3	Структура	9
4	Позиция	3
5	Порядок	9
6	Статус	5
7	Расположение	2
8	Ошибки	22
9	Виды темпоральной стабильности МП	6

Таблица 7 Рубрики фасета «Ошибки» 10-фасетной классификации

Ошибки	№ рубрики	Название рубрики	Код рубрики
Обще-языковые	1	Все предложение аграмматично	AgramTotal
	2	Морфологическая ошибка во фрагменте текста, не вводимом коннектором	ErrorMorph
	3	Морфологическая ошибка во фрагменте текста, вводимом коннектором	ErrorMorphPostCNT
	4	Синтаксическая ошибка во фрагменте текста, не вводимом коннектором	ErrorSyntax
	5	Синтаксическая ошибка во фрагменте текста, вводимом коннектором	ErrorSyntaxPostCNT
	6	Лексическая ошибка, искажение смысла в переводимом фрагменте	ErrorSemant
	7	Пунктуационная ошибка	ErrorPunct
	8	Слова кириллицей в переводе	Cyrillic
	9	Русское слово латинским шрифтом (транслитерация)	Latin
	10	Пропуск фрагмента текста	Lacuna
	11	Избыточный перевод	Pleonasm
Категорийно обусловленные	12	Переведена первая часть неоднословного коннектора	TrPart1CNT
	13	Переведена вторая часть неоднословного коннектора	TrPart2CNT
	14	Первая часть неоднословного коннектора переведена ошибочно	ErrorPart1CNT
	15	Вторая часть неоднословного коннектора переведена ошибочно	ErrorPart2CNT
	16	Коннектор переведен несуществующей языковой единицей	ErrorTotalCNT
	17	Коннектор ошибочно заменен языковой единицей, не являющейся коннектором	ErrorMorphCNT
	18	Орфографическая ошибка в форме коннектора	ErrorOrthCNT
	19	Семантическая ошибка в выборе коннектора	ErrorCNT

ние / колебание качества) в переводе на заданном интервале и определение вида темпоральной стабильности МП. Между тем в разные моменты времени перевод одного и того же исходного текста может содержать разное число и набор ошибок или не содержать их вовсе.

Темпоральной стабильностью МП предлагается считать стабильность содержания перевода (и его качества) во времени (ср. подобную трактовку темпоральной стабильности в [25]). При этом следует различать темпоральную стабильность МП и темпоральную робастность, считая последнюю, как в [26], способностью системы МП проявлять функциональную устойчивость к внешним возмущени-

Таблица 8 Рубрики фасета «Виды темпоральной стабильности МП» 10-фасетной классификации

№	Название рубрики
1	Повышение качества МП
2	Снижение качества МП
3	Колебание качества МП
4	Изменение набора ошибок в МП без динамики его качества
5	Изменение МП без динамики его качества
6	МП без изменений

ям — ошибкам, опечаткам, пропускам текста и т. д. во входных данных. Здесь, однако, сфокусируемся на темпоральной стабильности МП.

В интервальной оценке стабильности МП можно выделить два этапа. Сначала аннотатор проводит анализ ошибок (если таковые имеются) во всех версиях МП одного предложения. Ему/ей требуется проанализировать как число ошибок, так и их виды. Анализ ошибок проводится последовательно, начиная с самой ранней версии НМП, т. е. двигаясь от 1-го к 12-му этапу эксперимента.

На основании проведенного анализа аннотатор оценивает динамику ошибок во времени и проставляет одну из пяти рубрик, характеризующих темпоральную стабильность МП (см. первые пять строк в табл. 8), или шестую рубрику «МП без изменений» — для серий переводов без изменений. Эти рубрики проставляются в аннотации последними по времени, т. е. в 12-й версии перевода, но характеризуют они всю переводную серию.

Если при оценке качества перевода у аннотатора возникают затруднения, то привлекается мнение эксперта. Анализ ошибок в МП и проверка результатов аннотирования осуществлялись одним аннотатором и одним экспертом, поэтому измерение коэффициента согласия между экспертами не выполнялось.

Эксперимент по интервальной оценке позволил получить количественные данные распределения аннотаций по рубрикам девятого фасета 10-фасетной классификации, характеризующим темпоральную стабильность МП (табл. 9).

Наибольшее значение индикатора (30,8% аннотаций) у третьей рубрики, характеризующей колебание качества МП. На втором месте по числу аннотаций — первая рубрика, характеризующая повышение качества перевода (27,6% всех аннотаций). Такое соотношение категорий, возможно, говорит о том, что по мере обучения системы МП качество перевода не всегда изменяется только в сторону повышения или снижения. Оно колеблется, т. е. число допущенных в переводе ошибок растет или уменьшается, при этом их состав может меняться.

Затем следуют пятая рубрика, характеризующая изменение МП без динамики его качества (16,4%), и четвертая рубрика, характеризующая изменение набора ошибок в переводе без динамики его качества (13,6%). Отметим, что число аннотаций с первой рубрикой приблизительно равно числу аннотаций с пятой и четвертой рубриками, а также числу аннотаций с третьей рубрикой. По

Таблица 9 Значения индикаторов интервальной оценки

№	Рубрика фасета «Виды темпоральной стабильности МП»	Число аннотаций	Доля от общего числа аннотаций (3000) как значение индикатора
1	Повышение качества МП	828	27,6%
2	Снижение качества МП	288	9,6%
3	Колебание качества МП	924	30,8%
4	Изменение набора ошибок в МП без динамики его качества	408	13,6%
5	Изменение МП без динамики его качества	492	16,4%
6	МП без изменений	60	2,0%

сравнению с первой рубрикой в 3 раза меньше аннотаций приходится на вторую рубрику, характеризующую снижение качества перевода (9,6%). Стабильными на протяжении всего эксперимента оставались 2% переводов.

4 Заключение

Главное отличие классификации ошибок МП, описанной в статье, от существующих классификаций заключается в том, что она нацелена на обеспечение интервальной оценки для систем НМП, качество работы которых изменяется во времени: повышается, колеблется или снижается.

Проведенный эксперимент позволил определить основные виды изменений в качестве МП и вычислить значения индикаторов его интервальной оценки. Наибольшее значение оказалось у индикатора рубрики, характеризующей колебание качества МП. Это может свидетельствовать о том, что со временем качество работы системы МП не всегда повышается или понижается. Оно может колебаться, а состав и число ошибок — меняться. Результаты эксперимента и изложенные выводы подтверждают актуальность дальнейшего исследования темпоральной стабильности МП на материале других языковых пар и направлений перевода.

Литература

1. Translation quality assessment: From principles to practice / Eds. J. Moorkens, Sh. Castilho, F. Gaspari, S. Doherty. — Machine translation: Technologies and applications ser. — Cham, Switzerland: Springer International Publishing, 2018. Vol. 1. 292 p.
2. *Specia L., Scarton C., Paetzold G. H.* Quality estimation for machine translation. — Synthesis lectures on human language technologies ser. — London: Morgan & Claypool, 2018. 162 p.
3. *Bittner H.* Evaluating the evaluator: A novel perspective on translation quality assessment. — New York, NY, USA: Routledge, 2020. 282 p.

4. Нуриев В. А., Егорова А. Ю. Методы оценки качества машинного перевода: современное состояние // Информатика и её применения, 2021. Т. 15. Вып. 2. С. 104–111.
5. 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. — Machine translation: Technologies and applications ser. — Cham, Switzerland: Springer, 2018. Vol. 1. P. 129–158.
6. Castilho Sh., Moorkens J., Gaspari F., Calixto I., Tinsley J., Way A. Is neural machine translation the new state of the art? // Prague Bull. Math. Linguist., 2017. Vol. 108. No. 1. P. 109–120.
7. 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. — Machine translation: Technologies and applications ser. — Cham, Switzerland: Springer, 2018. Vol. 1. P. 9–38.
8. Popović M. Language-related issues for NMT and PBMT for English–German and English–Serbian // Machine Translation, 2018. Vol. 32. P. 237–253.
9. Castilho Sh., Moorkens J., Gaspari F., et al. Evaluating MT for massive open online courses: A multifaceted comparison between PBSMT and NMT systems // Machine Translation, 2018. Vol. 32. P. 255–278.
10. Castilho Sh., Gaspari F., Moorkens J., Popović M., Toral A. Editors' foreword to the special issue on human factors in neural machine translation // Machine Translation, 2019. Vol. 33. P. 1–7.
11. Инькова-Манзотти О. Ю. Коннекторы противопоставления во французском и русском языках: Сопоставительное исследование. — М.: Информэлектро, 2001. 429 с.
12. Инькова О. Ю., Кружков М. Г. Метод описания структуры неоднословных коннекторов в надкорпусных базах данных // Системы и средства информатики, 2018. Т. 28. № 4. С. 168–181.
13. Инькова О. Ю., Кружков М. Г. Сочетаемость логико-семантических отношений: количественные методы анализа // Информатика и её применения, 2019. Т. 13. Вып. 2. С. 83–91.
14. Егорова А. Ю., Зацман И. М., Косарик В. В., Нуриев В. А. Нестабильность нейронного машинного перевода // Системы и средства информатики, 2020. Т. 30. № 2. С. 124–135.
15. Егорова А. Ю., Зацман И. М., Кружков М. Г., Нуриев В. А. Методика темпоральной оценки неустойчивости машинного перевода // Системы и средства информатики, 2020. Т. 30. № 3. С. 67–80.
16. Егорова А. Ю., Зацман И. М., Кружков М. Г., Нуриев В. А. Машинный перевод: индикаторная оценка результатов обучения искусственной нейронной сети // Системы и средства информатики, 2020. Т. 30. № 4. С. 124–137.
17. Егорова А. Ю., Зацман И. М., Кружков М. Г., Нуриев В. А. Индикаторная оценка неустойчивости нейронного машинного перевода // Системы и средства информатики, 2021. Т. 31. № 2. С. 139–151.
18. Национальный корпус русского языка. <http://www.ruscorpora.ru>.
19. Зализняк А. А., Зацман И. М., Инькова О. Ю. Надкорпусная база данных коннекторов: построение системы терминов // Информатика и её применения, 2017. Т. 11. Вып. 1. С. 100–108.

20. Зацман И. М., Кружков М. Г. Надкорпусная база данных коннекторов: развитие системы терминов проектирования // Системы и средства информатики, 2018. Т. 28. № 4. С. 156–167.
21. Егорова А. Ю., Зацман И. М., Мамонова О. С. Надкорпусные базы данных в лингвистических проектах // Системы и средства информатики, 2019. Т. 29. № 3. С. 77–91.
22. Ranganathan Sh. R. Colon classification. — London: Edward Goldston, 1933. 106 p.
23. Nelson D., Turney L. What's in a word? Rethinking facet headings in a discovery service // Inform. Technol. Libr., 2015. Vol. 34. No. 2. P. 76–91.
24. Handbook of linguistic annotation / Eds. N. Ide, J. Pustejovsky. — Dordrecht, The Netherlands: Springer Science + Business Media, 2017. 1468 p.
25. Huber G. P. Temporal stability and response-order biases in participant descriptions of organizational decisions // Acad. Manage. J., 1985. Vol. 28. No. 4. P. 943–950.
26. Belinkov Y., Bisk Y. Synthetic and natural noise both break neural machine translation // 6th Conference (International) on Learning Representations Proceedings. — La Jolla, CA, USA: ICLR Board, 2018. P. 1–13.

Поступила в редакцию 12.08.21

EXPERT EVALUATION OF MACHINE TRANSLATION: ERROR CLASSIFICATION

A. Yu. Egorova, I. M. Zatsman, and V. A. Nuriev

Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation

Abstract: The paper considers the error classification applied in the expert evaluation of the machine translation quality. The classification includes common error headings (for errors at the level of grammar, vocabulary, punctuation, etc.) as well as headings that are associated with a specific type of linguistic unit selected for evaluating the machine translation quality. The quality evaluation is performed by experts as they linguistically annotate machine translation outcomes. If, while annotating, an expert finds errors, then headings, necessary to characterize these errors, are included in the annotation. The headings allow one to calculate the relative frequency of machine translation errors for the array of test sentences selected for translation and quality evaluation. The main goal of the paper is to describe the proposed classification of common and specific errors. The principal difference of the classification from the existing error classifications is that it is aimed at backing interval evaluation for machine translation systems, whose quality of work may vary over time. The headings of the proposed classification allow experts to record both improvements and decreases in machine translation quality at a given time interval.

Keywords: machine translation; quality evaluation; error classification; common errors; specific errors; linguistic annotation; interval evaluation

DOI: 10.14357/08696527210313

Acknowledgments

The work was performed at the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences using the CKP “Informatics” of FRC CSC RAS.

References

1. Moorkens, J., S. Castilho, F. Gaspari, and S. Doherty, eds. 2018. *Translation quality assessment: From principles to practice*. Machine translation: Technologies and applications ser. Cham, Switzerland: Springer International Publishing. Vol. 1. 299 p.
2. Specia, L., C. Scarton, and G. H. Paetzold. 2018. *Quality estimation for machine translation*. Synthesis lectures on human language technologies ser. London: Morgan & Claypool. 162 p.
3. Bittner, H. 2020. *Evaluating the evaluator: A novel perspective on translation quality assessment*. New York, NY: Routledge. 282 p.
4. Nuriev, V. A., and A. Yu. Egorova. 2021. Metody otsenki kachestva mashinnogo perevoda: sovremennoe sostoyanie [Methods of quality estimation for machine translation: State-of-the-art]. *Informatika i ee Primeneniya — Inform. Appl.* 15(2):104–111.
5. 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. Machine translation: Technologies and applications ser. Cham: Springer International Publishing. 1:129–158.
6. Castilho, Sh., J. Moorkens, F. Gaspari, I. Calixto, J. Tinsley, and A. Way. 2017. Is neural machine translation the new state of the art? *Prague Bull. Math. Linguist.* 108(1):109–120.
7. 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. Machine translation: Technologies and applications ser. Cham, Switzerland: Springer. 1:9–38.
8. Popović, M. 2018. Language-related issues for NMT and PBMT for English–German and English–Serbian. *Machine Translation* 32:237–253.
9. Castilho, S., J. Moorkens, F. Gaspari, et al. 2018. Evaluating MT for massive open online courses. A multifaceted comparison between PBSMT and NMT systems. *Machine Translation* 32:255–278.
10. Castilho, S., F. Gaspari, J. Moorkens, M. Popović, and A. Toral. 2019. Editors’ foreword to the special issue on human factors in neural machine translation. *Machine Translation* 33:1–7.
11. Inkova-Manzotti, O. Yu. 2001. *Konnektory protivopostavleniya vo frantsuzskom i russkom yazykakh. Sopostavitel’noe issledovanie* [Connectors of opposition in French and Russian: A comparative study]. Moscow: Informelektro. 429 p.
12. Inkova, O. Yu., and M. G. Kruzhkov. 2018. Metod opisaniya struktury neodnoslovnnykh konnektorov v nadkorpusnykh bazakh dannykh [Method for description of multiword connectives in supracorpora databases]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(4):168–181.

13. Inkova, O. Yu., and M. G. Kruzhkov. 2019. Sochetaemost' logiko-semanticheskikh otnosheniy: kolichestvennye metody analiza [Compatibility of logical semantic relations: Methods of quantitative analysis]. *Informatika i ee Primeneniya — Inform. Appl.* 13(2):83–91.
14. Egorova, A. Yu., I. M. Zatsman, V. V. Kosarik, and V. A. Nuriev. 2020. Nestabil'nost' neyronnogo mashinnogo perevoda [Instability of neural machine translation]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(2):124–135.
15. Egorova, A. Yu., I. M. Zatsman, M. G. Kruzhkov, and V. A. Nuriev. 2020. Metodika temporal'noy otsenki nestabil'nosti mashinnogo perevoda [The technique allowing for temporal estimation of the machine translation instability]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(3):67–80.
16. Egorova, A. Yu., I. M. Zatsman, M. G. Kruzhkov, and V. A. Nuriev. 2020. Mashinnyy perevod: indikatornaya otsenka rezul'tatov obucheniya iskusstvennoy neyronnoy seti [Machine translation: Indicator-based evaluation of training progress in neural processing]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(4):124–137.
17. Egorova, A. Yu., I. M. Zatsman, M. G. Kruzhkov, and V. A. Nuriev. 2021. Indikatornaya otsenka nestabil'nosti neyronnogo mashinnogo perevoda [Indicator-based evaluation of machine translation instability]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 31(2):139–151.
18. Natsional'nyy korpus russkogo yazyka [Russian National Corpus]. Available at: <http://www.ruscorpora.ru/> (accessed August 11, 2021).
19. Zaliznyak, A. A., I. M. Zatsman, and O. Yu. Inkova. 2017. Nadkorpurnaya baza dannykh konnektorov: postroenie sistemy terminov [Supracorpora database on connectives: Term system development]. *Informatika i ee Primeneniya — Inform. Appl.* 11(1):100–108.
20. Zatsman, I. M., and M. G. Kruzhkov. 2018. Nadkorpurnaya baza dannykh konnektorov: razvitie sistemy terminov proektirovaniya [Supracorpora database of connectives: Design-oriented evolution of the term system]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(4):156–167.
21. Egorova, A. Yu., I. M. Zatsman, and O. S. Mamonova. 2019. Nadkorpurnye bazy dannykh v lingvisticheskikh proektakh [Supracorpora databases in linguistic projects]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(3):77–91.
22. Ranganathan, S. R. 1933. *Colon classification*. London: Edward Goldston. 106 p.
23. Nelson, D., and L. Turney. 2015. What's in a word? Rethinking facet headings in a discovery service. *Inform. Technol. Libr.* 34(2):76–91.
24. Ide, N., and J. Pustejovsky, eds. 2017. *Handbook of linguistic annotation*. Dordrecht, The Netherlands: Springer Science + Business Media. 1468 p.
25. Huber, G. P. 1985. Temporal stability and response-order biases in participant descriptions of organizational decisions. *Acad. Manage. J.* 28(4):943–950.
26. Belinkov, Y., and Y. Bisk. 2018. Synthetic and natural noise both break neural machine translation. *6th Conference (International) on Learning Representations Proceedings*. La Jolla, CA: ICLR Board. 1–13.

Received August 12, 2021

Contributors

Egorova Anna Yu. (b. 1991) — junior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; ann.shurova@gmail.com

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

Nuriev Vitaly A. (b. 1980) — Doctor of Science in philology, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; nurieff.v@gmail.com

ИСПОЛЬЗОВАНИЕ ГЕОИНФОРМАЦИОННЫХ СИСТЕМ В ТЕХНОЛОГИИ ПОДДЕРЖКИ КОНКРЕТНО-ИСТОРИЧЕСКИХ ИССЛЕДОВАНИЙ

И. М. Адамович¹, О. И. Волков²

Аннотация: Статья посвящена дальнейшему развитию распределенной технологии поддержки конкретно-исторических исследований (ПКИИ), основанной на принципах краудсорсинга и ориентированной на широкий круг не относящихся к профессиональным историкам и биографам пользователей. Развитие осуществляется за счет интеграции с геоинформационными системами (ГИС), тенденция применения которых в исторических исследованиях наметилась в последнее время. Проведен анализ проблем интеграции ГИС в рассматриваемую технологию и описанию способов их решения. Проанализированы задачи и особенности ГИС в историческом исследовании, оценена возможность и перспективность использования механизма геохронологического трекинга. Были проанализированы возможные проблемы включения элементов ГИС в технологию, вытекающие из специфики конкретно-исторического исследования. Были предложены методы решения выявленных проблем, описана необходимая модификация объектной модели технологии, предусматривающая введение в нее нового типа данных, служащего для связи элементов семантической сети технологии с картографической информацией.

Ключевые слова: конкретно-историческое исследование; распределенная технология; ГИС; модель данных; геохронологический трекинг

DOI: 10.14357/08696527210314

1 Введение

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

В [2, 3] описана разработанная в ФИЦ ИУ РАН распределенная технология ПКИИ, основанная на принципах краудсорсинга (мобилизации ресурсов широкого круга добровольцев посредством информационных технологий). Данные в этой технологии организованы в форме семантической сети. Узлы сети представляют собой именованные универсальные классы объектов. Факты задаются

¹Федеральный исследовательский центр «Информатика и управление» Российской академии наук, Adam@amsd.com

²Федеральный исследовательский центр «Информатика и управление» Российской академии наук, Volkov@amsd.com

значениями экземпляров классов и связями между ними. Связи наследуются из сети классов [4].

В последнее время наметилась тенденция применения ГИС в исторических исследованиях. Положение о том, что при историческом анализе время невозможно рассматривать отдельно от пространства, становится неотъемлемой частью менталитета историка-исследователя и полноценной методологической посылкой изучения, интерпретации и научного изложения исторических событий. Геоинформационные системы широко применяются в различных направлениях исторических исследований, главным условием которых является пространственная привязка [5]. Отсюда вытекает потребность во включении в технологию ПКИИ средств сбора, хранения, обработки и отображения пространственных данных, а также получение на их основе новой информации и знаний о пространственно-координированных явлениях [6].

Данная статья посвящена анализу проблем интеграции ГИС в технологию ПКИИ и описанию способов их решения.

2 Задачи и особенности геоинформационных систем в историческом исследовании

Источники географической информации для создания исторических ГИС черпаются не только из современных картографических представлений той или иной части пространства. Не меньшую, а подчас и значительно большую роль здесь играют исторические источники.

Историко-биографическая информация, получаемая из архивных источников, при кажущейся естественной упорядоченности характеризуется неполнотой, нечеткостью, а часто и недостаточной точностью. Отмечается и такое качество географической информации, как ее «повсеместность» и множество путей отражения в различных видах исторических источников, которые потенциально могут составить основу исторической ГИС [7].

Выделяют шесть направлений использования ГИС в исторических исследованиях [8]:

- (1) ГИС как инструмент презентации;
- (2) ГИС как инструмент исследования особенностей развития региона;
- (3) ГИС как инструмент исторической реконструкции;
- (4) ГИС как инструмент анализа исторического явления (пространственно-исторический анализ) или процесса (пространственно-временной анализ);
- (5) ГИС как «портал» для доступа к исторической информации;
- (6) ГИС как инструмент, облегчающий историческое исследование за счет повторного использования геоисторической информации.

Технология ПКИИ опирается на конкретно-исторический или идеографический (нарративный) метод исторического познания и предназначена прежде всего для фиксации всех событий и фактов для познания истории объекта в ее конкретных эмпирических формах [9]. Для этого класса исторических задач требуется ориентация на микроанализ пространственных процессов, представленных в виде индивидуальных траекторий — в отличие от ставших уже привычными ГИС-приложений в истории, относящихся к более масштабным пространственно-временным перемещениям (например, при изучении структуры миграционных потоков).

Такие подходы существуют. Геохронологический трекинг представляет собой пример такого специализированного научно-методического и программно-технологического инструментария. Построение геохронологического трека исторической личности (индивидуума) или исторического объекта на основании геопространственной интерпретации его биографической информации есть интеграция хронологических и географических данных в виде графа, соединяющего географические точки нахождения исторической личности (группы и пр.) с цветоградиентной привязкой к тем или иным параметрам этого индивидуума или исторических событий. При этом вершины такого графа имеют строгую историко-географическую привязку, а дуги носят условно-логический характер [10].

Применение алгоритмических и программных механизмов построения, а также корректного отображения геохронологических треков в ГИС для отдельных исторических личностей, членов малых социальных групп и др. позволяет добиться снижения неопределенности, неточности исторического знания при решении таких типов исторических задач, как [10]:

- установление возможности встреч («пересечений» индивидуумов из изучаемой совокупности), зависимости исторических событий и пр.;
- выявление неподтверждаемых гипотез, мифологизированных событий, неточностей;
- выявление «пучков траекторий», т. е. характерных (типологических) пространственных перемещений индивидуумов в определенные периоды времени;
- уточнение компьютерных реконструкций в историко-географическом плане и т. п.

3 Применение геохронологического трекинга в технологии поддержки конкретно-исторических исследований

Анализ задач и особенностей ГИС в исторических исследованиях, а также опыт применения элементов технологии ПКИИ в реальных конкретно-исторических проектах [11] показывает, что применение механизмов построения

и отображения геохронологических треков позволит помимо выполнения функций наглядного отображения и презентации конкретно-исторической информации решать следующие задачи.

1. Установление возможности встреч изучаемых индивидуумов.
2. Оценка правдоподобия некоторых семейных преданий и выявление их реальной подоплеки.

Первые две задачи связаны. В качестве примера можно привести семейное предание о том, что некий изучаемый индивидуум (далее — ИИ) был адъютантом легендарного комдива Г. И. Котовского. Анализ послужного списка ИИ показал, что он никогда под началом Котовского не служил. Предание оказалось опровергнутым, но остался открытым вопрос о причинах его возникновения. Анализ геохронологических треков Котовского и ИИ показал их пересечение осенью 1921 г. на Подолии вблизи Румынской границы. В это время ИИ служил адъютантом 391 стрелкового Таращанского полка 44-й Киевской дивизии. Осенью 1921 г. 44-я дивизия совместно с кавалерийской бригадой Котовского участвовала в крупных войсковых маневрах Вооруженных сил Украины и Крыма, проводимых в этом регионе под руководством Фрунзе. Естественно предположить, что в качестве полкового адъютанта ИИ в ходе маневров мог встречаться с Котовским. Его рассказ об этом эпизоде своей биографии вполне мог быть неверно интерпретирован близкими и привести к возникновению этого предания.

3. Поиск административной (губерния, уезд, волость) или церковной (приход) привязки населенного пункта по ближайшим городам и церквям.
4. Поиск архива, содержащего сведения о жителях населенного пункта, по аналогии с близлежащими населенными пунктами, для которых соответствующие архивы уже найдены.

Третья и четвертая задачи являются сходными. Из-за изменения границ областей и губерний, неоднократных передач материалов из одного архива в другой довольно часто бывает трудно определить архив, содержащий необходимые документы, а также в рамках одного архива определить необходимый фонд и опись. Существенно облегчить решение данной задачи может географический поиск близлежащих крупных населенных пунктов, губернских и уездных центров, а также церквей, в чьих метрических книгах и исповедных ведомостях может содержаться искомая информация. Также данная задача может быть решена по аналогии, если для близлежащих населенных пунктов другими пользователями уже найдены соответствующие архивы и фонды. С большой вероятностью информация по жителям интересующего населенного пункта будет найдена в том же архиве и в том же фонде.

5. Поиск сохранившегося здания по старинному адресу при существенных изменениях в застройке населенного пункта. Примером такой задачи может служить поиск дома в городе Сумы по адресу: «угол ул. Суджанской

и Покрышкина пер., дом Фоменко». Поиск осложняется тем, что в XX в. Сумы были сильно перестроены и Покрышкин пер. был ликвидирован.

6. Определение точного расположения населенного пункта по его описанию. Примером такой задачи может служить поиск деревни Федяево по описанию в «Списках населенных мест по сведениям 1859 года»: «Владимирская губ. уезд Покровский, стан 2. Между рр. Пекшей и Шередарью. Федяево д. вл. при колодце. Расстояние в верстах от уездного города — 24, от становой квартиры — 27». Поиск осложняется тем, что в XX в. три близлежащие деревни Федяево, Акосниково и Анциферово были объединены в одну под именем Кашино и на современных картах название Федяево отсутствует.

Пятая и шестая задача относятся к классу часто возникающих при биографическом поиске задач поиска современного объекта по старинному вербальному описанию. Иногда это можно сделать относительно легко, найдя объект по описанию на исторической карте, а затем определив эту локацию на современной. Но часто задача усложняется следующими факторами:

- отсутствием в распоряжении исследователя точной исторической карты и необходимостью использования вместо нее грубой схемы населенного пункта или местности, что затрудняет поиск современной локации найденного объекта;
- отсутствием в распоряжении исследователя каких-либо исторических карт и схем населенного пункта или местности, что приводит к необходимости поиска по историческому описанию непосредственно на современной карте и может быть затруднительно из-за появления и исчезновения объектов, их переименований и даже изменения таких ориентиров, как леса, русла рек, возвышенности и овраги.

7. Совместное использование карт исследователями.

Поскольку технология ПКИИ основана на принципах краудсорсинга, седьмая задача решается ее средствами автоматически. Это позволяет исследователям использовать найденные другими редкие исторические карты и схемы, что существенно упрощает отдельные этапы конкретно-исторического исследования.

4 Проблемы включения элементов геоинформационных систем в технологию поддержки конкретно-исторических исследований

Историческая геоинформация, как правило, доступна исследователю в виде сканов исторических карт, планов и схем. Эти сканы представлены в растровой форме и не имеют пространственной привязки.

Пространственная привязка осуществляется по опорным точкам, которые связывают известные местоположения в растровых данных с соответствующими местами в геоданных, уже имеющими пространственную привязку [12]. Процесс привязки может быть затруднен за счет следующих факторов:

- (1) использование карт с разными проекциями;
- (2) использование схем и планов с приблизительными и искаженными пропорциями;
- (3) наличие неточностей на картах и схемах;
- (4) сложность определения соответствия точек за счет переименований и различных одновременно используемых названий для населенных пунктов;
- (5) возможные ошибочные привязки.

Проблемы (1)–(3) приводят к невозможности корректного определения расстояний на карте и, следовательно, к проблемам при решении описанных выше задач геохронологического трекинга в технологии ПКИИ, опирающихся на понятие близости точек. Решение проблем (1)–(2) и отчасти проблемы (3) может быть осуществлено трансформацией раstra средствами ГИС. Обычно для этой цели применяется полиномиальная трансформация первого порядка (аффинная), полиномиальная трансформация второго–третьего порядка или трансформация методом сплайна.

Полиномиальная трансформация использует построение полинома на основе опорных точек и алгоритм подбора методом наименьших квадратов. Этот способ оптимизирован для соблюдения общей точности, но не гарантирует локальной точности. Трансформация первого порядка (аффинная) используется для сдвига, изменения масштаба или поворота набора растровых данных. Чем выше порядок трансформации, тем более сложные искажения могут быть скорректированы. Однако трансформация выше третьего порядка используется очень редко. Трансформация методом сплайна — это истинный «метод резинового листа», оптимизированный для обеспечения локальной, но не глобальной точности привязки. Она основана на сплайн-функции — кусочном полиноме, который обеспечивает гладкие непрерывные переходы между соседними полиномами. Это дает точное совмещение исходной и конечной опорных точек, но точность совпадения удаленных пикселей не гарантируется [13].

Проблемы (4) и (5) связаны, поскольку именно сложность определения опорной точки может стать причиной ошибочной привязки. Примером может служить задача поиска г. Лениногорск, из которого в 1941 г. был осуществлен призыв на фронт изучаемого индивидуума. Простой поиск находит единственный г. Лениногорск в Татарстане. Но более глубокое изучение вопроса показывает, что данный населенный пункт стал городом только в 1956 г., а в 1941 г. Лениногорском назывался г. Риддер в Рудном Алтае (Казахстан). Недостаточно тщательная проработка этого и ему подобных вопросов может привести к ошибочным привязкам. Единственно возможным средством минимизации таких ошибок из-за их слабой формализуемости является взаимная верификация привязок пользователями технологии ПКИИ. Поскольку данная технология основана на принципах краудсорсинга, в ней предусмотрены такие меры, основанные на определенных групповых нормах работы с информацией, распределении ролей и использовании

специальных фасилитационных техник (техники управления, отличающиеся от простого управления тем, что их способ не директивный, т.е. не выходит за рамки самоорганизации управляемой системы) [14].

5 Модификация объектной модели технологии поддержки конкретно-исторических исследований

Интеграция ГИС в технологию ПКИИ предполагает установление связей между семантической сетью технологии и географической базой данных (БД) ГИС. Эта задача осложняется тем, что структура и состав семантической сети технологии ПКИИ не определены заранее, а создаются динамически широким кругом не связанных между собой исследователей, т.е. в семантической сети изначально не предусмотрены такие предопределенные типы данных, которые могут быть использованы для поддерживаемых алгоритмом связей с географическими объектами ГИС. Поэтому требуется модификация объектной модели технологии.

Вводится новое понятие ГИС-связки, состоящей из позиции, темпоральных границ связки и ссылки на источник информации.

Позиция представляет собой набор, который включает:

- координатный блок — географические координаты для указания точечного географического объекта или набор географических координат для указания полигона. Ввод координат может осуществляться как посредством указания точного значения в любом из общепринятых форматов, так и интерактивно указанием точки на карте;
- модификатор позиции — указатель, как именно алгоритму интерпретировать указанную позицию. Может принимать значения «точно», «приблизительно», «не далее x км». Значение «точно» определяет положение объекта в строгом соответствии с указанными координатами, а «приблизительно» и «не далее x км» — в некоторой окрестности от указанных координат. Для значения «не далее x км» размер окрестности указывается явно, а для «приблизительно» — в соответствии с настройками умолчаний для алгоритма технологии ПКИИ.

Темпоральные границы необходимы, поскольку пространственные координаты объектов и их границы могут меняться со временем. В качестве примера можно привести д. Матвейково Ступинского района Московской области, расположившуюся по обеим сторонам небольшой реки Киреевки, притока Оки, и имеющую географические координаты $54^{\circ}55'05''$ с. ш., $37^{\circ}56'05''$ в. д. В деревне находится часовня 1878 года постройки. Упоминается в 1577 г. в вотчине за Даниловским монастырем, а «преж была за Семеном Шереметевым в вотчине». Первоначально поселение находилось несколько южнее, чем сейчас, там, где в р. Киреевку впадает Воронка. Но случившаяся моровая язва унесла жизни

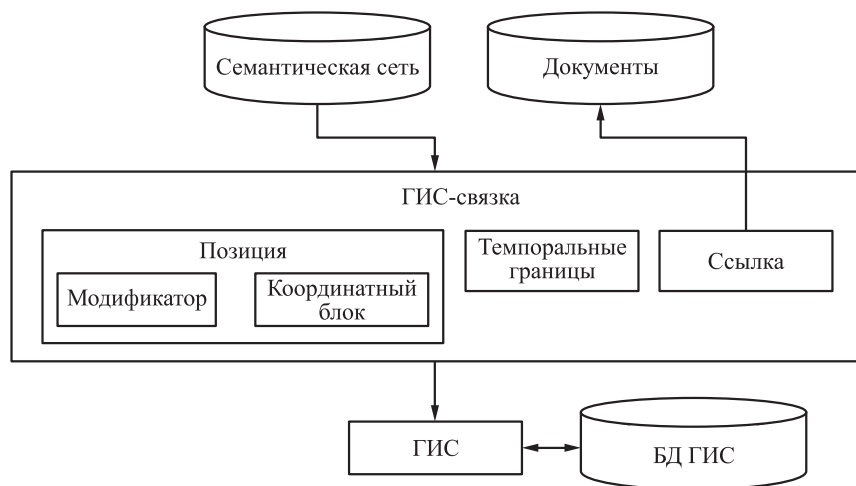


Схема связей данных технологии ПКИИ с ГИС

почти всего населения деревни, поэтому ее забросили и построили на новом месте в верховье реки. Таким образом, современные координаты деревни актуальны для периода от «ранее 1878 г.» до настоящего времени, а предыдущее расположение деревни относится к периоду от «ранее 1577 г.» до «не позднее 1878 г.». Для фиксации таких диапазонов, предполагающих нечеткое и неявное указание дат за счет поддержки типов значений «не задано», «раньше», «позже», «одновременно» и «примерно одновременно», а также отношений дат событий без указания самих значений этих дат, служат специальные темпоральные атрибуты, уже предусмотренные в объектной модели технологии ПКИИ [15].

Ссылка на источник информации представляет собой указатель фрагмента текста [2]. Эта ссылка необходима, поскольку специфика конкретно-исторического исследования состоит в том, что исследователь, как правило, имеет дело не с объективным знанием об объекте исследования, а с его отражением в исторических документах, которые обладают свойством фрагментарности и противоречивости и не существуют в отрыве от своих метаданных, т. е. информации о документе-источнике [2].

На рисунке представлена схема использования ГИС-связки для взаимодействия технологии ПКИИ с ГИС.

6 Выводы

Предложенный механизм интеграции с ГИС существенно дополняет и развивает технологию ПКИИ, ориентированную на широкий круг не относящихся к профессиональным историкам и биографам пользователей, что очень актуаль-

но в связи со все возрастающим общественным интересом к частной, семейной истории.

Механизм в полной мере отражает специфику конкретно-исторического исследования и позволяет исследователю проводить микроанализ пространственных процессов, представленных в виде индивидуальных траекторий, опираясь на инструментарий геохронологического трекинга.

Описанная модификация технологии ПКИИ предполагает усовершенствование ее объектной модели за счет введения в нее нового типа данных, служащего для связи элементов семантической сети с картографической информацией.

Литература

1. Помникова А. Ю. Семейная история в дискурсивном пространстве // Вестник Мининского университета, 2019. Т. 7. № 1. Ст. 9. 22 с.
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. Баталов Р. Н., Радченко Л. К. Обзор основных направлений использования ГИС-технологий в историко-картографических исследованиях // Вестник СГУГиТ, 2020. Т. 25. № 1. С. 119–135.
6. Баранов Ю. Б., Берлянт А. М., Капралов Е. Г., Кошкарев А. В., Серапи-нас Б. Б., Филиппов Ю. А. Геоинформатика: Толковый словарь основных терминов. — М.: ГИС-Ассоциация, 1999. 205 с.
7. Владимиров В. Н. Геоинформационные технологии в исторических исследованиях // Новая и новейшая история, 2006. № 3. С. 133–141.
8. Кончаков Р. Б., Баранова Е. В. Геоинформационные системы в исторических исследованиях: современные тенденции // XV Державинские чтения: Мат-лы Общеросс. научн. конф. — Тамбов: Академия гуманитарного и социального образования, 2010. С. 154–162.
9. Ушаков Е. В. Философия и методология науки. — М.: Юрайт, 2018. 392 с.
10. Ивакин Я. А., Потапычев С. Н. Геохронологический трекинг — специализированный ГИС-инструментарий исторического исследования // Историческая информатика. Информационные технологии и математические методы в исторических исследованиях и образовании, 2016. № 1-2. С. 3–11.
11. Маркова Н. А. Фольклористы-любители в переписке с А. В. Марковым // Традиционная культура, 2017. № 2(66). С. 123–135.
12. Черепанова Е. С., Пьянков С. В., Шихов А. Н. Геоинформатика: основы работы с географическими пространственными данными. — Пермь: ПГНИУ, 2017. 94 с.

13. ArcMap. Обработка и анализ растровых данных. Пространственная привязка. — Environmental Systems Research Institute, 2016. <https://desktop.arcgis.com/ru/arcmap/10.4/manage-data/raster-and-images/fundamentals-for-georeferencing-a-raster-dataset.htm>.
14. Адамович И. М., Волков О. И. Устойчивость технологии поддержки конкретно-исторических исследований к попыткам искажения истории // Системы и средства информатики, 2021. Т. 31. № 2. С. 152–162. doi: 10.14357/08696527210214.
15. Адамович И. М., Волков О. И. Автоматизированный поиск противоречий в конкретно-исторической информации // Системы и средства информатики, 2020. Т. 30. № 3. С. 145–153. doi: 10.14357/08696527200313.

Поступила в редакцию 13.04.21

THE USE OF GEOGRAPHIC INFORMATION SYSTEMS IN TECHNOLOGY OF CONCRETE HISTORICAL INVESTIGATION SUPPORT

I. M. Adamovich and O. I. Volkov

Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation

Abstract: The article is devoted to the further development of the distributed technology of concrete historical investigation support, based on the principles of crowdsourcing and focused on a wide range of users which are nonprofessional historians and biographers. Development is carried out through integration with geographic information systems (GIS), the use of which in historical research is the trend in recent years. The analysis of the problems of GIS integrating into the technology under consideration and the description of the ways to solve them have been made. The article analyzes the tasks and features of GIS in historical research and evaluates the possibility and prospects of using the mechanism of geochronological tracking. Possible problems of including the GIS elements in the technology, which are the consequence of the specificity of concrete historical investigation, are analyzed. The methods for solving the identified problems are proposed. The necessary modification of the object model of the technology in the form of the introduction of a new type of data for the connection of the elements of the semantic network of technology with cartographic information is described.

Keywords: concrete historical investigation; distributed technology; GIS; data model; geochronological tracking

DOI: 10.14357/08696527210314

References

1. Pomnikova, A. Yu. 2019. Semeynaya istoriya v diskursivnom prostranstve [Family stories in different types of discourse]. *Vestnik of Minin University* 7(1):9. 22 p.

2. Adamovich, I. M., and O. I. Volkov. 2016. Tekhnologiya raspredelennogo avtomatizirovannogo analiza istoricheskikh tekstov [The distributed automated technology of historical texts analysis]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 3(26):148–161. doi: 10.14357/08696527160311.
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* 1(29):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* 2(29):161–171. doi: 10.14357/08696527190214.
5. Batalov, R. N., and L. K. Radchenko. 2020. Obzor osnovnykh napravleniy ispol'zovaniya GIS-tekhnologiy v istoriko-kartograficheskikh issledovaniyakh [Overview of main directions of using GIS technologies in historical and cartographical research]. *Vestnik Sibirskogo gosudarstvennogo universiteta geosistem i tekhnologiy* [Vestnik of the Siberian State University of Geosystems and Technologies] 1(25):119–135.
6. Baranov, Ju. B., A. M. Berljant, E. G. Kapralov, A. V. Koshkarev, B. B. Serapinas, and Ju. A. Filippov. 1999. *Geoinformatika: Tolkovyy slovar' osnovnykh terminov* [Geoinformatics: An explanatory dictionary of basic terms]. Moscow: GIS-Assotsiatsiya. 205 p.
7. Vladimirov, V. N. 2006. Geoinformatsionnye tekhnologii v istoricheskikh issledovaniyakh [Geoinformation technologies in historical studies]. *Novaya i noveyshaya istoriya* [Modern and Contemporary History] 3:133–141.
8. Konchakov, R. B., and E. V. Baranova. 2010. Geoinformatsionnye sistemy v istoricheskikh issledovaniyakh: sovremennyye tendentsii [GIS in historical research: The current trend]. *Mat-ly XV Obshcheross. nauchn. konf. "Derzhavinskie chteniya"* [15th All-Russian Scientific Conference "Derzhavin Readings"]. Tambov: Academy of Humanities and Social Education. 154–162.
9. Ushakov, E. V. 2018. *Filosofiya i metodologiya nauki* [Philosophy and methodology of science]. Moscow: Yurayt. 392 p.
10. Ivakin, Ya. A., and S. N. Potapychev. 2016. Geokhronologicheskiy treking — spetsializirovanny GIS-instrumentariy istoricheskogo issledovaniya [Geochronological tracking — a specialized GIS toolkit for historical research]. *Historical Information Science. Information Technology and Quantitative Methods in Historical Research and Education* 1-2:3–11.
11. Markova, N. A. 2017. Fol'kloristy-lyubiteli v perepiske s A. V. Markovym [Folklorist-samateurs in correspondence with A. V. Markov]. *Traditsionnaya kul'tura* [Traditional Culture] 2(66):123–135.
12. Cherepanova, E. S., S. V. P'yankov, and A. N. Shikhov. 2017. *Geoinformatika: osnovy raboty s geograficheskimi prostranstvennymi dannymi* [Geoinformatics: The basics of working with geographic spatial data]. Perm: PGNIU. 94 p.
13. ArcMap. Obrabotka i analiz rastrovyykh dannykh. Prostranstvennaya privyazka [ArcMap. Processing and analyzing of raster data. Georeferencing]. Environmental Systems Research Institute. Available at: <https://desktop.arcgis.com/ru/arcmap/10.4/>

manage-data/raster-and-images/fundamentals-for-georeferencing-a-raster-dataset.htm (accessed August 18, 2021).

14. Adamovich, I. M., and O. I. Volkov. 2021. Ustoychivost' tekhnologii podderzhki konkretno-istoricheskikh issledovaniy k popytkam iskazheniya istorii [The resistance of technology of concrete historical investigation support to attempts of history distortion]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 31(2):152–162. doi: 10.14357/08696527210214.
15. Adamovich, I. M., and O. I. Volkov. 2020. Avtomatizirovannyi poisk protivorechiy v konkretno-istoricheskoy informatsii [Automated search for contradictions in concrete-historical information]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(3):145–153. doi: 10.14357/08696527200313.

Received April 13, 2021

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

СИМВОЛЬНОЕ МОДЕЛИРОВАНИЕ ЗАДАЧ И КОНСТРУИРОВАНИЕ ПРОГРАММ

В. Д. Ильин¹

Аннотация: Представлена обновленная методология символьного моделирования программируемых задач, рассчитанная на построение *систем знаний о задачах*, предназначенных для автоматизированного конструирования программ. Задачные конструкции создаются посредством связей по памяти между задачами, представленными в виде *задачных конструктивных объектов*. Определены три типа функций связи по памяти, каждая из которых позволяет поставить в соответствие паре задач некоторую третью задачу, образованную из этой пары. Представлением задачной области, рассчитанным на формализацию задачных знаний и реализацию процесса конструирования задач, служит *задачный граф*. Искомая конструкция задается спецификацией задачи, содержащей описание ее памяти, ограничений на число задачных вершин (и, если необходимо, ограничений, связанных с точностью результата и др.). Механизм построения *разрешающих структур* ставит в соответствие спецификации исходной задачи подграф на задачном графе путем реализации трех типов поведения в соответствии с тремя типами запросов на конструирование.

Ключевые слова: символьное моделирование задач; задачный конструктивный объект; система знаний о задачах; задачный граф; разрешающая структура; конструирование программ

DOI: 10.14357/08696527210315

1 Введение

В методологии *символьного моделирования задач* (*s-моделирования задач*) программирование рассматривается как создание (на языке программирования) сообщения (называемого программой и адресованного транслятору), которое представляет собой упорядоченный набор инструкций, реализующих алгоритм поведения компьютера, смартфона или другой *s-машины*², определенный в спецификации задачи. Создание программы (в общем случае) включает изучение спецификации задачи, написание исходного текста программы, компиляцию и/или интерпретацию, компоновку [получение кода программы, готового к исполнению, путем сборки из объектных кодов модулей программы (одного или нескольких)] и отладку.

Программирование поведения символьных автоматов (компьютеров, смартфонов и других *s-машин*) занимает особое место среди всех видов деятельности,

¹Федеральный исследовательский центр «Информатика и управление» Российской академии наук, vdilyin@yandex.ru

²Ильин В. Д.. Символьное моделирование // Большая российская энциклопедия — электронная версия. https://bigenc.ru/technology_and_technique/text/4010980.

связанных с управлением. Оно делает потенциально короткой по времени и уникальной по эффективности цепочку: замысел – его символьное воплощение – реализация в виде задуманного поведения автомата. К сожалению, пока только потенциально. Процесс символьного воплощения замысла в виде сегодняшней разработки программного обеспечения никак не назовешь коротким по времени. Из-за этого остается лишь потенциальной и уникальная эффективность всей цепочки. Ее среднее звено продолжает оставаться критическим.

С тех пор как существует программирование, программисты ищут способы и создают средства повышения производительности своего труда [1–9]. В современные инструментальные системы программирования, как правило, встроены механизмы применения шаблонов разрабатываемых продуктов и готовых решений для отдельных составляющих. Такие системы обычно содержат средства редактирования текстов программ, трансляторы, средства отладки, тестирования и оптимизации программ, графические средства проектирования интерфейсов, заготовки текстов для программ различных типов, библиотеки программ и др.

Профессиональным программистам и разработчикам систем автоматизации программирования издавна противостоят любители рассматривать спецификации задач и соответствующие им программы как математические сущности. В первых рядах таких любителей — энтузиасты автоматического синтеза и доказательства правильности программ.

В реальности спецификация нетривиальной задачи и программа соотносятся так же, как проектная документация (в инженерном деле) и изделие, которое предстоит изготовить. Пригодность изделия (его соответствие проектной документации) можно определить только путем тестов, разработанных специалистами, хорошо знающими, что должен уметь испытуемый и какими свойствами (включая надежность и безопасность) должен обладать.

Запись формул и выделение фрагментов текста. Для выделения определений, замечаний и примеров используются средства языка TSM-комплекса (TSM: textual symbolic modeling), разработанного для формализованного описания текстовых s-моделей сообщений.

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.

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

2 S-моделирование задач

Представление связей между понятиями в виде разрешимых задач — необходимое условие построения количественных s-моделей систем понятий.

□ S-задача — это четверка {*Formul*, *Rulsys*, *Alg*, *Prog*}, где *Formul* — постановка задачи; *Rulsys* — множество систем обязательных и ориентирующих

правил решения задачи, поставленных в соответствие *Formul*; *Alg* — объединение множеств алгоритмов, каждое из которых соответствует одному элементу из *Rulsys*; *Prog* — объединение множеств программ, каждое из которых поставлено в соответствие одному из элементов *Alg*. Постановка задачи *Formul* — пара $\{Mem, Rel\}$, где *Mem* — множество *понятий задачи*, на котором задано разбиение $Mem = Inp \cup Out$ ($Inp \cap Out = \emptyset$) и совокупность *Rel* связей между понятиями, определяющая бинарное отношение $Rel \subseteq Inp \times Out$. Множество *Mem* называем *памятью задачи*, а *Inp* и *Out* — ее *входом* и *выходом*, значения которых предполагается соответственно задавать и искать. $\square$

В общем случае множества *Rulsys*, *Alg* и *Prog* могут быть пустыми: числа их элементов зависят от степени изученности задачи.

Для каждого элемента из *Rulsys*, *Alg* и *Prog* задано *описание применения*. Описания применения элементов *Rulsys* включают спецификацию типа *решателя задачи* (автономная s-машина, сетевая кооперация s-машин, кооперация человек-s-машина и др.), требование к информационной безопасности и др. Описания применения элементов из *Alg* включают данные о допустимых режимах работы *решателя задачи* (автоматический локальный, автоматический распределенный, интерактивный локальный и др.), о требованиях к полученному результату и др. Описания применения программ включают данные о языках программирования, операционных системах и др.

◇ Каждая программа сопровождается ссылками на наборы тестовых примеров. ◇

□ *S-алгоритм* — система правил решения задачи (соответствующая одному из элементов *Rulsys*), позволяющая за конечное число шагов поставить в однозначное соответствие заданному набору данных, принадлежащему *Inp*, результирующий набор, принадлежащий *Out*. $\square$

□ *S-программа* — реализованный (на языке программирования высокого уровня, машинно-ориентированном языке и/или в системе машинных команд) s-алгоритм, представленный в форме сообщения, определяющего поведение s-машинного решателя задачи с заданными свойствами. Существует в *символьном, кодовом и сигнальном* воплощениях, связанных отношениями трансляции. $\square$

□ *S-данные* — сообщение, необходимое для решения некоторой задачи или совокупности задач, представленное в форме, рассчитанной на распознавание, преобразование и интерпретацию решателем (программой или человеком). Специализация s-сообщения (*s-message*) по параметру получатель s-сообщения (*s-recipient*), значением которого является решатель s-задачи (*s-solver*): $s-data \approx s-message [::s-recipient = s-solver]$. $\square$

3 S-конструирование задач

Связи по памяти между s-задачами определяются тремя типами функций, каждая из которых является функцией двух аргументов и позволяет поставить

в соответствие паре s -задач некоторую третью s -задачу, образованную из этой пары.

□ S -задача a связана с s -задачей b по памяти, если существует хотя бы одна пара элементов $\{elem\ Met^a, elem\ Met^b\}$, принадлежащих памяти Met^a s -задачи a и памяти Met^b s -задачи b , относительно которой определено общее означивание (элементы имеют одно и то же множество значений). Если S и H — множества s -задач и $D \leq S * S$, а каждой паре (s_i, s_j) элементов из D ставится в соответствие определенный элемент из H , то задана функция связи по памяти $h = conn(s_i, s_j)$. При этом D называем областью определения функции $conn$ и обозначаем D^{conn} . Множество

$$R = \{h : elem\ H; h = conn(s_i, s_j); s_i : elem\ D^{conn}, s_j : elem\ D^{conn}\}$$

называем областью значений функции $conn$. □

Тип связи по памяти зависит от содержимого пересечения по памяти: составлена ли связь из элементов выхода одной и входа другой задачи, из элементов выходов задач или из элементов их входов или же связь получена путем комбинации предыдущих способов.

Элементарная задачная конструкция — задачная пара. Любая задачная конструкция, в свою очередь, может быть использована как составляющая еще более сложной задачной конструкции.

□ Система pS знаний о задачных конструктивных объектах (называемых также p -объектами) — это триада $\langle pA, lng, intr \rangle$, где pA — задачная область; lng — язык спецификации p -объектов; $intr$ — интерпретатор спецификаций искомых p -объектов на pA . Если P — множество всех p -объектов, а $A < P$ — его непустое подмножество, при этом в A (содержащем не менее двух элементов) не существует ни одного элемента, который не был бы связан по памяти хотя бы с одним элементом из A , то s -модель pa задачной области pA — это p -объект, который задается парой $\langle$ память met^A множества задач A задачной области pA $\rangle$, $\langle$ семейство $rul(met^A)$ связей, заданных на met^A $\rangle$. Непустое множество met^A элементов памяти разбито на три подмножества: входов $inpr^A$ задач, выходов out^A задач и подмножество or^A , каждый из элементов которого является и входом, и выходом некоторых задач. Любое одно из этих подмножеств может быть пустым; могут быть одновременно пустыми $inpr^A$ и out^A . □

В отличие от памяти задачи, состоящей из входа и выхода, память задачной области содержит подмножество or элементов памяти, каждый из которых может быть или задан (как входной), или вычислен (как выходной). Такие элементы памяти называем обратимыми, а or — подмножеством обратимых элементов. Подмножество $inpr$ называем подмножеством задаваемых, а подмножество out — подмножеством вычисляемых элементов. S -модель pa задачной области pA служит для интерпретации составленных на языке lng спецификаций искомых задач. Интерпретация заключается в постановке в соответствие некоторому

подмножеству (или паре подмножеств) памяти mem^A некоторой подобласти задачной области pA , названной *разрешающей структурой*. Интерпретация спецификации искомого p -объекта на pA — *конструктивное доказательство существования разрешающей структуры*.

□ *Задачный граф* — представление задачной области, рассчитанное на реализацию процесса *p -конструирования* и формализацию задачных знаний. Множество вершин графа, составленное из задачных объектов, называем *задачным базисом графа* и обозначаем *p -basis*. *Ребро задачного графа* — это пара вершин с непустым пересечением по памяти. *Нагрузка ребра* определяется множеством всех пар элементов памяти, входящих в это пересечение. Каждая вершина графа имеет память. *Память вершины* — это память задачи (или задачной области), которую представляет вершина. □

□ *Составная задача comp* — подобласть задачной области pA , которая содержит не менее двух элементов из множества задач A и на памяти которой задано разбиение: $mem^{comp} = in^{comp} \cup out^{comp}$; $in^{comp} \cap out^{comp} = \emptyset$, определяющее вход in^{comp} и выход out^{comp} составной задачи. Составной задаче поставлен в соответствие *ориентированный граф*, вершинами которого служат задачи. Каждая вершина помечена именем задачи. *Ребра графа* — это пары задач с непустыми пересечениями по памяти. □

В зависимости от состава вершин определены следующие *типы задачных графов*: *U -граф* имеет множество вершин только из *простых задач*; в *C -графе* хотя бы одна вершина представлена составной задачей и нет вершин, представляющих собой задачную область; в *G -графе* — не менее одной вершины представлено задачной областью (остальные могут быть простыми и составными задачами).

G -графы служат средством формализации знаний о p -объектах. Система знаний об s -задачах обеспечивает процессы p -(специализации, конкретизации и конструирования).

◇ Возможность существования в задачном графе одной или нескольких вершин, являющихся задачными областями, имеет принципиальное значение для формализации задачных знаний. ◇

Искомая конструкция задается спецификацией задачи, содержащей описание ее памяти, ограничений на число задачных вершин (и, если необходимо, ограничений, связанных с размером задачи, точностью результата и др.). Заданное описание интерпретируется на задачном графе, который служит представлением интересующей конструктора задачной области. Средством интерпретации спецификаций задач служит механизм конструирования на задачном графе. *Интерпретация на U -графе* в процессе задачного конструирования заключается в постановке в соответствие подмножеству (или паре подмножеств) элементов его памяти такого подграфа, память которого находилась бы в заданном отношении к введенному подмножеству (или паре подмножеств). *Интерпретации на C -графе и G -графе* аналогичны интерпретации на U -графе.

□ *Задача t представима на задачном графе $graph$* , если ее вход inp^t содержится в подмножестве $Giv^{graph} \cup Or^{graph}$, а выход out^t — в подмножестве $Comput^{graph} \cup Or^{graph}$ памяти задачного графа; при этом существует не менее одной задачи из базиса графа, вход которой содержится в inp^t или совпадает с ним. □

□ *Разрешающей структурой $solv^{t-graph}$ на графе $graph$* , поставленной в соответствие некоторой задаче t , называется подграф с минимальным числом задачных вершин, на котором задача t представима. □

Интерпретация задачной вершины U -графа (или C -графа) в процессе поиска разрешающей структуры заключается в соотношении означенности входа и выхода.

Правила интерпретации задачной вершины:

- если полностью означен вход, то полностью означен и выход;
- если означенным полагается хотя бы один элемент выхода, то означенным полагается полностью вход.

Механизм построения разрешающих структур ставит в соответствие спецификации исходной задачи подграф на задачном графе путем реализации трех типов поведения в соответствии с тремя типами запросов на конструирование.

1. Для заданных подмножеств x и y ($x \wedge y = 0$) памяти $mem^{t-graph}$ тогда существует разрешающая структура $solv^{xy}$ (xy — помета) с минимальным числом задачных вершин, вход которой определен посредством x , а выход — посредством y , когда найдется подграф G , множество вершин которого включает хотя бы одну вершину с *разрешимой задачей*, а объединение выходов вершин подграфа G содержит подмножество y (или совпадает с ним).
2. Для подмножества x , заданного на памяти $mem^{t-graph}$ задачного графа, тогда найдется разрешающая структура $solv^x$, вход которой определен подмножеством x , а выход является непустым подмножеством памяти графа, включающим максимальное число элементов, которые могут быть определены при заданном x , когда $x \wedge Comput = 0$ ($Comput < mem^{t-graph}$) и найдется хотя бы одна вершина с *разрешимой задачей*.
3. Для подмножества y , заданного на $mem^{t-graph}$, тогда найдется разрешающая структура $solv^y$ с минимальным числом задачных вершин, выход которой содержит y , а вход составлен из элементов, принадлежащих Giv , когда $y \wedge Giv = 0$.

Для каждого из трех типов запросов получено *конструктивное доказательство существования разрешающей структуры* соответствующего типа.

После того как найдена разрешающая структура, становится осуществимым процесс ее *конкретизации* в соответствии со спецификацией условий применения исходной задачи.

4 Заключение

Обновленная методология *символьного моделирования программируемых задач* рассматривается как теоретическая платформа построения *систем знаний о задачах*, применимых в технологиях *автоматизированного конструирования программ*.

Уточнены определения понятий *задачного конструктивного объекта*, *задачного графа* и др.

Усовершенствован механизм построения *разрешающих структур на задачных графах*.

Литература

1. *Sweger K.* 1130 scientific subroutine package (1130-CM-02X): Programmer's manual. — IBM, 1967. <http://media.ibm1130.org/1130-006-ocr.pdf>.
2. *Dijkstra E. W.* Structured programming // Software engineering techniques / Eds. J. Burton, B. Randell. — Brussels: NATO Science Committee, 1969. P. 89–93.
3. *Wirth N. E.* Program development by stepwise refinement // Commun. ACM, 1971. Vol. 14. No. 4. P. 221–227. doi: 10.1145/362575.362577.
4. *Polster F. J.* Reuse of software through generation of partial systems // IEEE T. Software Eng., 1986. Vol. 12. No. 3. P. 402–416.
5. *Luker P. A., Burns A.* Program generators and generation software // Comput. J., 1986. Vol. 29. No. 4. P. 315–321.
6. *Ильин В. Д.* Система порождения программ. — М.: Наука, 1989. 264 с.
7. *Ilyin V. D.* A methodology for knowledge based engineering of parallel program systems // 8th Conference (International) on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems Proceedings. — Newark, NJ, USA: Gordon and Breach Science Publs., Inc., 1995. P. 805–809.
8. *Арушанян О. Б., Волченкова Н. И.* Обеспечение доступа к библиотеке численного анализа НИВЦ МГУ в системе Интернет // Вычислительные методы и программирование, 2000. Т. 1. № 2. С. 1–8.
9. *Meil D.* The search for unlimited productivity // Commun. ACM, 2021. Vol. 64. No. 6. P. 10–11.

Поступила в редакцию 06.07.21

SYMBOLIC MODELING OF TASKS AND CONSTRUCTING PROGRAMS

V. D. Ilyin

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

Abstract: An updated methodology for symbolic modeling of programmable tasks is presented, designed to build knowledge systems about tasks that are intended for automated design of programs. Task constructions are created by

means of memory links between tasks, represented by task constructive objects. Three types of memory link functions are defined, each of which allows assigning to a pair of tasks some third task formed from this pair. The task graph serves as representation of the problem area, designed to formalize task knowledge and implement the process of constructing tasks. The required construction is described by the problem specification, which contains a description of its memory, restrictions on the number of task vertices (and, if necessary, restrictions related to the accuracy of result, etc.). The mechanism of searching the resolving structures matches the specification of the original problem with a subgraph on the task graph by implementing three types of behavior in accordance with three types of requests for constructing.

Keywords: symbolic modeling of tasks; task constructive object; system of knowledge about tasks; task graph; resolving structure; constructing programs

DOI: 10.14357/08696527210315

References

1. Sweger, K. 1967. 1130 scientific subroutine package (1130-CM-02X): Programmer's manual. Available at: <http://media.ibm1130.org/1130-006-ocr.pdf> (accessed August 20, 2021).
2. Dijkstra, E. 1969. Structured programming. *Software engineering techniques*. Eds. J. Burton and B. Randell. Brussels: NATO Science Committee. 89–93.
3. Wirth, N. E. 1971. Program development by stepwise refinement. *Commun. ACM* 14(4):221–227. doi: 10.1145/362575.362577.
4. Polster, F. J. 1986. Reuse of software through generation of partial systems. *IEEE T. Software Eng.* 12(3):402–416.
5. Luker, P. A., and A. Burns. 1986. Program generators and generation software. *Comput. J.* 29(4):315–321.
6. Ilyin, V. D. 1989. *Sistema porozhdeniya programm* [The system of program generating]. Moscow: Nauka. 264 p.
7. Ilyin, V. D. 1995. A methodology for knowledge based engineering of parallel program systems. *8th Conference (International) on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems Proceedings*. Newark, NJ: Gordon and Breach Science Pubs., Inc. 805–809.
8. Arushanyan, O. B., and N. I. Volchenskova. 2000. Obespechenie dostupa k biblioteke chislennogo analiza NIVTs MGU v sisteme Internet [The access to Numerical Analysis Library of Moscow University Research Computing Center in Internet]. *Vychislitel'nye metody i programmirovaniye* [Numerical Methods and Programming] 1(2):1–8.
9. Meil, D. 2021. The search for unlimited productivity. *Commun. ACM* 64(6):10–11.

Received July 6, 2021

Contributor

Ilyin Vladimir D. (b. 1937) — Doctor of Science in technology, professor, leading scientist, A. A. Dorodnicyn Computing Center, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 40 Vavilov Str., Moscow 119333, Russian Federation; vdilyin@yandex.ru

ПОПРАВКА К СТАТЬЕ

М. О. ВОРОНЦОВА, А. А. КУДРЯВЦЕВА, С. Я. ШОРГИНА
«АНАЛИТИЧЕСКИЕ СВОЙСТВА И АСПЕКТЫ ВЫЧИСЛЕНИЯ
ГАММА-ЭКСПОНЕНЦИАЛЬНОЙ ФУНКЦИИ»
(Системы и средства информатики, 2021. Т. 31. № 2. С. 108–118)
DOI: 10.14357/08696527210210

С. 116, вместо

PROBABILISTIC CHARACTERISTICS OF BALANCE INDEX
OF FACTORS WITH GENERALIZED GAMMA DISTRIBUTION

должно быть

ANALYTICAL PROPERTIES AND ASPECTS OF COMPUTATION
OF THE GAMMA-EXPONENTIAL FUNCTION

ОБ АВТОРАХ

Адамова Ксения Андреевна (р. 1994) — инженер 2-й категории Научно-производственного центра автоматизации и приборостроения им. академика Н. А. Пилюгина

Адамович Игорь Михайлович (р. 1934) — кандидат технических наук, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Бобрикова Екатерина Васильевна (р. 1976) — кандидат физико-математических наук, доцент кафедры прикладной информатики и теории вероятностей Российского университета дружбы народов

Босов Алексей Вячеславович (р. 1969) — доктор технических наук, главный научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Бутенко Юлия Ивановна (р. 1987) — кандидат технических наук, доцент кафедры теоретической информатики и компьютерных технологий Московского государственного технического университета им. Н. Э. Баумана

Волков Олег Игоревич (р. 1964) — ведущий программист Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Воронцов Михаил Олегович (р. 1996) — студент факультета вычислительной математики и кибернетики Московского государственного университета имени М. В. Ломоносова

Гайдамака Юлия Васильевна (р. 1971) — доктор физико-математических наук, профессор 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

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

Дулин Сергей Константинович (р. 1950) — доктор технических наук, профессор, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук; главный научный сотрудник Научно-исследовательского и проектно-конструкторского института информатизации, автоматизации и связи на железнодорожном транспорте

Дулина Наталья Георгиевна (р. 1947) — кандидат технических наук, ведущий программист Вычислительного центра им. А. А. Дородницына 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

Егорова Анна Юрьевна (р. 1991) — младший 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

Зацаринный Александр Алексеевич (р. 1951) — доктор технических наук, профессор, главный 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 Федерального исследовательского центра «Информатика и управление» Российской академии наук

Кочеткова Ирина Андреевна (р. 1985) — 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

Кривенко Михаил Петрович (р. 1946) — доктор технических наук, профессор, ведущий научный сотрудник Института проблем информатики Федерального

исследовательского центра «Информатика и управление» Российской академии наук

Кружков Михаил Григорьевич (р. 1975) — старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Крюков Алексей Анатольевич (р. 1983) — заместитель руководителя аппарата генерального директора АО «Концерн воздушно-космической обороны «Алмаз-Антей»

Кудрявцев Алексей Андреевич (р. 1978) — кандидат физико-математических наук, доцент кафедры математической статистики факультета вычислительной математики и кибернетики Московского государственного университета имени М. В. Ломоносова

Куцазли Анна Ивановна (р. 1997) — аспирант кафедры прикладной информатики и теории вероятностей Российского университета дружбы народов

Мейханаджян Лусине Акобовна (р. 1990) — кандидат физико-математических наук, доцент Департамента анализа данных и машинного обучения Финансового университета при Правительстве РФ

Милованова Татьяна Александровна (р. 1977) — кандидат физико-математических наук, доцент 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

Нуриев Виталий Александрович (р. 1980) — доктор филологических наук, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Платонова Анна Алексеевна (р. 1996) — аспирант кафедры прикладной информатики и теории вероятностей Российского университета дружбы народов

Розенберг Игорь Наумович (р. 1965) — доктор технических наук, профессор, научный руководитель Научно-исследовательского и проектно-конструкторского института информатизации, автоматизации и связи на железнодорожном транспорте

Сидняев Николай Иванович (р. 1955) — доктор технических наук, профессор, заведующий кафедрой высшей математики Московского государственного технического университета им. Н. Э. Баумана

Синева Елизавета Евгеньевна (р. 1998) — студент магистратуры факультета фундаментальных наук Московского государственного технического университета им. Н. Э. Баумана

Степченко Юрий Афанасьевич (р. 1951) — кандидат 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

Харин Петр Александрович (р. 1993) — аспирант кафедры прикладной информатики и теории вероятностей Российского университета дружбы народов

Хилько Дмитрий Владимирович (р. 1987) — старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Шестаков Олег Владимирович (р. 1976) — доктор физико-математических наук, профессор кафедры математической статистики факультета вычислительной математики и кибернетики Московского государственного университета имени М. В. Ломоносова; старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук; ведущий научный сотрудник Московского центра фундаментальной и прикладной математики

Шнурков Петр Викторович (р. 1953) — кандидат физико-математических наук, доцент Национального исследовательского университета «Высшая школа экономики»

Шоргин Сергей Яковлевич (р. 1952) — доктор физико-математических наук, профессор, главный научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Правила подготовки рукописей статей для публикации в журнале «Системы и средства информатики»

Журнал «Системы и средства информатики» публикует теоретические, обзорные и дискуссионные статьи, посвященные научным исследованиям и разработкам в области информационных технологий.

Журнал издается на русском языке. По специальному решению редколлегии отдельные статьи могут печататься на английском языке.

Тематика журнала охватывает следующие направления:

- информационно-телекоммуникационные системы и средства их построения;
- архитектура и программное обеспечение вычислительных машин, комплексов и сетей;
- методы и средства защиты информации.

1. В журнале печатаются статьи, содержащие результаты, ранее не опубликованные и не предназначенные к одновременной публикации в других изданиях.

Публикация предоставленной автором(ами) рукописи не должна нарушать положений глав 69, 70 раздела VII части IV Гражданского кодекса, которые определяют права на результаты интеллектуальной деятельности и средства индивидуализации, в том числе авторские права, в РФ.

Ответственность за нарушение авторских прав, в случае предъявления претензий к редакции журнала, несут авторы статей.

Направляя рукопись в редакцию, авторы сохраняют свои права на данную рукопись и при этом передают учредителям и редколлегии журнала неисключительные права на издание статьи на русском языке (или на языке статьи, если он отличен от русского) и на перевод ее на английский язык, а также на ее распространение в России и за рубежом. Каждый автор должен представить в редакцию подписанный с его стороны «Лицензионный договор о передаче неисключительных прав на использование произведения», текст которого размещен по адресу <http://www.ipiran.ru/publications/licence.doc>. Этот договор может быть представлен в бумажном (в 2-х экз.) или в электронном виде (отсканированная копия заполненного и подписанного документа).

Редколлегия вправе запросить у авторов экспертное заключение о возможности публикации представленной статьи в открытой печати.

2. К статье прилагаются данные автора (авторов) (см. п. 8). При наличии нескольких авторов указывается фамилия автора, ответственного за переписку с редакцией.
3. Редакция журнала осуществляет экспертизу присланных статей в соответствии с принятой в журнале процедурой рецензирования.

Возвращение рукописи на доработку не означает ее принятия к печати.

Доработанный вариант с ответом на замечания рецензента необходимо прислать в редакцию.

4. Решение редколлегии о публикации статьи или ее отклонении сообщается авторам. Редколлегия может также направить авторам текст рецензии на их статью. Дискуссия по поводу отклоненных статей не ведется.

5. Редактура статей высылается авторам для просмотра. Замечания к редакции должны быть присланы авторами в кратчайшие сроки.
6. Рукопись предоставляется в электронном виде в форматах MS WORD (.doc или .docx) или L^AT_EX (.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 — A4; 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 10 страниц указанного формата. При превышении указанного объема редколлегия вправе потребовать от автора сокращения объема рукописи.

Сокращения слов, помимо стандартных, не допускаются. Допускается минимальное количество аббревиатур.

Все страницы рукописи нумеруются.

Шаблоны примеров оформления представлены в Интернете:

<http://www.ipiran.ru/publications/collected/template.doc>

8. Статья должна содержать следующую информацию на **русском и английском языках**:
 - название статьи;
 - Ф.И.О. авторов, на английском можно только имя и фамилию;
 - место работы, с указанием города и страны и электронного адреса каждого автора;
 - сведения об авторах, в соответствии с форматом, образцы которого представлены на страницах:
http://www.ipiran.ru/journal/collected/2019_29_03_rus/authors.asp и
http://www.ipiran.ru/journal/collected/2019_29_03_eng/authors.asp;
 - аннотация (не менее 100 слов на каждом из языков). Аннотация — это краткое резюме работы, которое может публиковаться отдельно. Она является основным источником информации в информационных системах и базах данных. Английская аннотация должна быть оригинальной, может не быть дословным переводом русского текста и должна быть написана хорошим английским языком. В аннотации не должно быть ссылок на литературу и, по возможности, формул;
 - ключевые слова — желательно из принятых в мировой научно-технической литературе тематических тезаурусов. Предложения не могут быть ключевыми словами.
 - источники финансирования работы (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. *Russ. 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 povyseniya 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 Publs. 352 p.

Описание переводной книги (в списке литературы к русскоязычной части необходимо указать: / Пер. с англ. — после названия книги, а в конце ссылки указать оригинал книги в круглых скобках):

1. В русскоязычной части:

Тимошенко С. П., Янг Д. Х., Уивер У. Колебания в инженерном деле / Пер. с англ. — М.: Машиностроение, 1985. 472 с. (*Timoshenko S. P., Young D. H., Weaver W.* Vibration problems in engineering. — 4th ed. — New York, NY, USA: Wiley, 1974. 521 p.)

2. В англоязычной части:

Timoshenko, S. P., D. H. Young, and W. Weaver. 1974. *Vibration problems in engineering*. 4th ed. New York, NY: 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]. Moscow. D.Sc. Diss. 272 p.

Kozhunova, O. S. 2009. Tekhnologiya razrabotki semanticheskogo slovary informatzionnogo monitoringa [Technology of development of semantic dictionary of information monitoring system]. PhD Thesis. Moscow: IPI RAN. 23 p.

Описание ГОСТа:

GOST 8.586.5-2005. 2007. Metodika vypolneniya izmereniy. Izmerenie raskhoda i kolichestva zhidkostey i gazov s pomoshch'yu standartnykh suzhayushchikh ustroystv [Method of measurement. Measurement of flow rate and volume of liquids and gases by means of orifice devices]. Moscow: Standardinform Publs. 10 p.

Описание патента:

Bolshakov, M. V., A. V. Kulakov, A. N. Lavrenov, and M. V. Palkin. 2006. 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: ssi@frccsc.ru (Стригина Светлана Николаевна)

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

Requirements for manuscripts submitted to Journal “Systems and Means of Informatics”

Journal “Systems and Means of Informatics” publishes theoretical, review, and discussion articles on the research and development in the field of information technology.

The journal is published in Russian. By a special decision of the editorial board, some articles can be published in English.

Topics covered include the following areas:

- information and communication systems and tools of their design;
- architecture and software of computational complexes and networks; and
- methods and tools of information protection.

1. The Journal publishes original articles which have not been published before and are not intended for simultaneous publication in other editions. An article submitted to the Journal must not violate the Copyright law. Sending the manuscript to the Editorial Board, the authors retain all rights of the owners of the manuscript and transfer the nonexclusive rights to publish the article in Russian (or the language of the article, if not Russian) and its distribution in Russia and abroad to the Founders and the Editorial Board. Authors should submit a letter to the Editorial Board in the following form:

Agreement on the transfer of rights to publish:

“We, the undersigned authors of the manuscript “. . . ,” pass to the Founder and the Editorial Board of the Journal “Systems and Means of Informatics” the nonexclusive right to publish the manuscript of the article in Russian (or in English) in both print and electronic versions of the Journal. We affirm that this publication does not violate the Copyright of other persons or organizations.

Author(s) signature(s): (name(s), address(es), date).”

This agreement should be submitted in paper form or in the form of a scanned copy (signed by the authors).

The Editorial Board has the right to request from the authors an official expert conclusion that the submitted article has no classified data prohibited for publication.

2. A submitted article should be attached with **the data on the author(s)** (see item 8). If there are several authors, the contact person should be indicated who is responsible for correspondence with the Editorial Board and other authors about revisions and final approval of the proofs.
3. The Editorial Board of the Journal examines the article according to the established reviewing procedure. If authors receive their article for correction after reviewing, it does not mean that the article is approved to be published. The corrected article should be sent to the Editorial Board for the subsequent review and approval.
4. The decision on the article publication or its rejection is communicated to the authors. The Editorial Board may also send the reviews on the submitted articles to the authors. Any discussion upon the rejected articles is not possible.
5. The edited articles will be sent to the authors for proofread. The comments of the authors to the edited text of the article should be sent to the Editorial Board as soon as possible.
6. The manuscript of the article should be presented electronically in the MS WORD (.doc or .docx) or L^AT_EX (.tex) formats, and additionally in the .pdf format. All documents

may be sent by e-mail or provided on a CD or diskette. A hard copy submission is not necessary.

7. The recommended typesetting instructions for manuscript.

Pages parameters: format A4, portrait orientation, document margins (cm): left — 2.5, right — 1.5, above — 2.0, below — 2.0, footer 1.3.

Text: font — Times New Roman, font size — 14, paragraph indent — 0.5, line spacing — 1.5, justified alignment.

The recommended manuscript size: not more than 10 pages of the specified format. If the specified size exceeded, the editorial board is entitled to require the author to reduce the manuscript.

Use only standard abbreviations. Avoid abbreviations in the title and abstract. The full term for which an abbreviation stands should precede its first use in the text unless it is a standard unit of measurement.

All pages of the manuscript should be numbered.

The templates for the manuscript typesetting are presented on site:

<http://www.ipiran.ru/publication/collected/template.doc>

8. Articles should enclose data both in **Russian and English**:

- title;
- author's name and surname;
- affiliation — organization, its address with ZIP code, city, country, and official e-mail address;
- data on authors according to the format (see site):
http://www.ipiran.ru/journal/collected/2019_29_03_rus/authors.asp and
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. *Russ. 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. Mamikoniants, eds. 1972. *Ekspluatatsiya turbogenera-torov s neposredstvennym okhlazhdeniem* [Operation of turbine generators with direct cooling]. Moscow: Energy Publs. 352 p.

Dissertation and Thesis:

Kozhunova, O. S. 2009. Tekhnologiya razrabotki semanticheskogo slovary informat-sionnogo monitoringa [Technology of development of semantic dictionary of information monitoring system]. Moscow: IPI RAN. PhD Thesis. 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 Publs. 10 p.

Bolshakov, M. V., A. V. Kulakov, A. N. Lavrenov, and M. V. Palkin. 2006. Sposob orientirovaniya po krenu letatel'nogo apparata s opticheskoy 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: ssi@frccsc.ru (to Svetlana Strigina)

http://www.ipiran.ru/english/journal_systems.asp

SYSTEMS AND MEANS OF INFORMATICS (SISTEMY I SREDSTVA INFORMATIKI)

SCIENTIFIC JOURNAL

Volume 31 No.3 Year 2021

Editor-in-Chief and Chair of Editorial Council
Academician I. A. Sokolov

IN THIS ISSUE:

JOINT STATIONARY DISTRIBUTION IN THE $GI/M/n/\infty$ QUEUE WITH GENERAL RENOVATION <i>T. A. Milovanova, I. S. Zaryadov, and L. A. Meykhanadzhyan</i>	4
SOME PROBABILITY-STATISTICAL PROPERTIES OF THE GAMMA-EXPONENTIAL DISTRIBUTION <i>M. O. Vorontsov, A. A. Kudryavtsev, and O. V. Shestakov</i>	18
INVESTIGATION OF THE PROBLEM OF CONTINUOUS PRODUCT STOCK CONTROL IN A STOCHASTIC MODEL OF REGENERATION WITH TWO OPTIMIZATION PARAMETERS <i>P. V. Shnurkov and K. A. Adamova</i>	36
INTEROPERABILITY AS A KEY CONDITION FOR THE IMPLEMENTATION OF DIGITAL TRANSFORMATION <i>I. N. Rozenberg, S. K. Dulin, and N. G. Dulina</i>	48
SECURE SCALING OF ELECTRONIC LEDGERS BASED ON TANGLES <i>A. A. Grusho, A. A. Zatsarinny, and E. E. Timonina</i>	60
COMPUTING BASED ON PROBABILISTIC PRINCIPAL COMPONENT ANALYSIS MODEL <i>M. P. Krivenko</i>	70
USE OF CASE GRAMMAR IN INFORMATION SEARCH IN THE EXPERT SYSTEM KNOWLEDGE BASE ON AIRCRAFT STRUCTURES <i>N. I. Sidnyaev, Yu. I. Butenko, and E. E. Sineva</i>	80
METHODOLOGY FOR ASSESSING PRODUCTION RISKS FOR DEVELOPING WEAPONS AND MILITARY EQUIPMENT <i>A. V. Bosov and A. A. Kryukov</i>	88
CONCEPTUAL FRAMEWORK FOR SUPRACORPORA DATABASES <i>M. G. Kruzhkov</i>	101