

Информатика и её применения

Том 16 Выпуск 4 Год 2022

СОДЕРЖАНИЕ

Оценка надежности синхронного и самосинхронного конвейеров И. А. Соколов, Ю. А. Степченков, Ю. Г. Дьяченко, Ю. В. Рождественский	2
Общий порядок аппроксимации оценок фильтрации состояний марковских скачкообразных процессов по дискретизованным наблюдениям А. В. Борисов	8
Несмещенная оценка риска пороговой обработки с двумя пороговыми значениями О. В. Шестаков	14
Обобщение метода выпрямления искаженных из-за мультиколлинеарности коэффициентов для регрессионных моделей с различной степенью корреляции объясняющих переменных М. П. Базилевский	20
Границы экстремального индекса времени ожидания в системе $M/G/1$ с распределением времени обслуживания в виде конечной смеси И. В. Пешкова	26
Оптимальное управление подключением резервного прибора в системе массового обслуживания $G/M/1$ Я. М. Агаларов	34
Об оптимальном расположении антенн для V2X-соединений в субтерагерцевом диапазоне Е. А. Мачнев, В. А. Бесчастный, Д. Ю. Острикова, Ю. В. Гайдамака, С. Я. Шоргин	42
Нечеткие усредняющие операторы в задаче агрегирования нечеткой информации В. Л. Хацкевич	51
О сложности обучения логических процедур классификации Е. В. Дюкова, А. П. Дюкова	57
Технология классификации типов контента электронного учебника А. В. Босов, А. В. Иванов	63
О научной парадигме информатики: верхний уровень классификации объектов ее предметной области И. М. Зацман	73
Модель и технология извлечения новых терминов из медицинских текстов И. М. Зацман, О. В. Золотарев, А. Х. Хакимова, Гу Дунсяо	80
О безопасной архитектуре вычислительной системы на основе микросервисов А. А. Грушо, Н. А. Грушо, М. И. Забежайло, Д. В. Смирнов, Е. Е. Тимонина, С. Я. Шоргин	87
Логическая реляционная модель структур данных для решения задач в предметной области управления землепользованием Д. О. Брюхов, С. А. Ступников	93
Единая модель государственных данных: сценарии развития А. П. Сучков	99
Об авторах	106
Авторский указатель за 2022 г.	108
2022 Author Index	112
Правила подготовки рукописей	116
Requirements for manuscripts	119

ОЦЕНКА НАДЕЖНОСТИ СИНХРОННОГО И САМОСИНХРОННОГО КОНВЕЙЕРОВ*

И. А. Соколов¹, Ю. А. Степченков², Ю. Г. Дьяченко³, Ю. В. Рождественский⁴

Аннотация: Самосинхронная (СС) схемотехника выступает альтернативой синхронным схемам. Самосинхронные схемы обладают рядом преимуществ в сравнении с синхронными аналогами, но аппаратно избыточны. Статья исследует иммунность самосинхронных и синхронных схем к однократным кратковременным логическим сбоям (ЛС) с учетом аппаратурной избыточности СС-схем. Самосинхронные схемы за счет своей неотъемлемой части — индикаторной подсхемы — способны обнаружить ЛС, проявляющийся как инверсия состояния выхода логической ячейки схемы, и приостановить функционирование схемы до его исчезновения. Тем самым СС-схемы маскируют однократный ЛС и предотвращают искажение данных. Использование модифицированного гистерезисного триггера для реализации разряда регистра ступени конвейера маскирует практически все ЛС в комбинационной части (КЧ) ступени конвейера. DICE-подобная реализация этого триггера позволяет в 4 раза снизить чувствительность СС-регистра к ЛС внутри него. Количественные оценки сбоеустойчивости показывают явное (в 2,5–9,4 раза) преимущество СС-конвейера схемы в сравнении с синхронным аналогом.

Ключевые слова: самосинхронные схемы; логический сбой; сбоеустойчивость; конвейер; индикация; вероятностная оценка

DOI: 10.14357/19922264220401

1 Введение

В современных условиях задача обеспечения надежной работы цифровых схем выдвигается на первый план. Повышение тактовой частоты в синхронной схемотехнике, все возрастающая функциональная сложность интегральных микросхем создают предпосылки для повышения их чувствительности к нештатным ситуациям — ЛС и физическим отказам из-за внешних и внутренних причин. Способность схемы к маскированию ЛС и отказов определяет уровень ее надежности. Практика показала, что отказы в цифровых микросхемах встречаются гораздо реже, чем ЛС [1].

Логический сбой проявляется как изменение логического уровня сигнала в цепи, приводящее к искажению результата обработки данных. Методы защиты от ЛС в основном направлены на их маскирование и используют корректирующие коды [2], методы обнаружения и изоляции [3] и некоторые другие подходы [4, 5].

Синхронные безыбыточные схемы не имеют встроенных средств контроля корректности переключений. В отличие от них, асинхронные схемы используют зачатки контроля корректности выполняемых операций [6, 7]. Однако их возможности по маскированию ЛС ограничены.

Альтернативой синхронным и асинхронным схемам выступают СС-схемы [8; 9; 10, р. 61–73]. Они характеризуются двухфазной дисциплиной работы и обязательным подтверждением (индицированием) завершения переключения в каждую фазу. Благодаря этому СС-схемы обладают естественной высокой сбоеустойчивостью [11, 12]. Плата за эти преимущества — увеличение в 1,5–3,3 раза (в зависимости от типа схемы) сложности реализации в сравнении с синхронными аналогами.

При постоянной интенсивности событий, приводящих к сбоям, число сбоев в схеме обычно растет с увеличением сложности схемы [13]. Поэтому решаемая в статье задача сравнительной количественной оценки устойчивости синхронных и самосинхронных схем к однократным ЛС с учетом их аппаратной сложности и особенностей функционирования особенно актуальна.

2 Вероятность появления логического сбоя

В микросхемах, изготовленных по технологии комплементарный металл — диэлектрик — полупроводник (КМДП), ЛС выражается во временном изменении потенциала некоторой цепи из-за

* Исследование выполнено в рамках гранта Российского научного фонда (проект 22-19-00237).

¹ Федеральное исследовательское учреждение «Информатика и управление» Российской академии наук, ISokolov@ipiran.ru

² Федеральное исследовательское учреждение «Информатика и управление» Российской академии наук, YStepchenkov@ipiran.ru

³ Федеральное исследовательское учреждение «Информатика и управление» Российской академии наук, diaura@mail.ru

⁴ Федеральное исследовательское учреждение «Информатика и управление» Российской академии наук, YRogdest@ipiran.ru

индуцирования в ней избыточных неравновесных носителей заряда. В комбинационных схемах логический уровень сигнала восстанавливается спустя какое-то время [14]. В триггерных схемах сбой может запомниться, стать критичным.

При постоянной эксплуатационной плотности потока случайных событий λ_0 (числа событий в единицу времени на единицу площади), инициирующих однократные сбои, интенсивность сбоев λ в схеме оценивается как сумма интенсивностей сбоев отдельных компонентов схемы [13, формула (3.11)]. Обычно в качестве компонента схемы берется КМДП-транзистор [12]:

$$\lambda = N\lambda_0\alpha,$$

где N — число транзисторов в схеме; α — усредненная вероятность появления сбоя при поражении одного транзистора. Тогда отношение интенсивностей сбоев для СС-схемы и синхронного аналога

$$K_\lambda = \frac{\lambda_S}{\lambda_{ST}} = \frac{N_S\lambda_0\alpha_S}{N_{ST}\lambda_0\alpha_{ST}} = \frac{N_S\alpha_S}{N_{ST}\alpha_{ST}} = \frac{\alpha_S}{A_R\alpha_{ST}},$$

где λ_S и λ_{ST} — интенсивности сбоев в синхронной и самосинхронной схемах; N_S и N_{ST} — сложность (число КМДП-транзисторов) синхронной и самосинхронной схем; α_S и α_{ST} — вероятность сбоя при повреждении одного транзистора синхронной и самосинхронной схем; $A_R = N_{ST}/N_S$ — коэффициент аппаратной избыточности СС-схемы в сравнении с синхронным аналогом.

Типовой эффективный диаметр трека ядерной частицы, одного из источников сбоев, достигает 2–2,5 мкм [15]. В КМДП-технологии с проектными нормами 65 нм это, например, соответствует размеру схемы из нескольких транзисторов. Поэтому целесообразно рассматривать вероятность появления сбоя применительно к логическим ячейкам схемы.

Будем рассматривать цифровую схему как совокупность библиотечных ячеек, соединенных сигнальными цепями. Для упрощения будем считать, что появление сбоя в любом месте принципиальной схемы ячейки вызывает инверсию уровня сигнала на ее выходе с вероятностью 0,5. Пусть одно событие, порождающее сбой, может привести к ЛС только в одной ячейке схемы. Тогда интенсивность сбоев λ_i в i -й ячейке схемы равна

$$\lambda_i = \frac{S_i}{2S} \lambda_0 P_\Theta,$$

где S_i и S — площади топологии i -й ячейки и всей схемы; P_Θ — вероятность индуцирования критического избыточного заряда одним сбойным событием.

Однако не все ЛС в ячейках схемы проявляются на ее выходах, поскольку при распространении по схеме они могут быть замаскированы. Интенсивность ЛС на выходах схемы

$$\begin{aligned} \lambda_{\text{вых}} &= \sum_{i=1}^{M_{\text{вых}}} \lambda_i + \sum_{i=M_{\text{вых}}+1}^M (\lambda_i P_{P,i}) = \\ &= \frac{\lambda_0 P_\Theta}{2S} \left(\sum_{i=1}^{M_{\text{вых}}} S_i + \sum_{i=M_{\text{вых}}+1}^M (S_i P_{P,i}) \right), \end{aligned} \quad (1)$$

где M — общее число ячеек в схеме; $M_{\text{вых}}$ — число выходных ячеек в схеме; $P_{P,i}$ — вероятность того, что сбой на выходе i -й внутренней ячейки приведет к сбою на выходе всей схемы.

Оценим вероятность $P_{P,i}$.

3 Маскирование логического сбоя логикой схемы

Введем вероятность $P_{P,ij}$ прохода ЛС с выхода i -й ячейки на выход Y_j схемы. Пусть зависимость Y_j от внутренних и внешних сигналов схемы описывается монотонной функцией $Y_j = F_j(X_1, \dots, X_M)$. С учетом разложения Y_j по X_i ,

$$\begin{aligned} Y_j &= F_{0ij}(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_M) + \\ &+ F_{1ij}(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_M) X_i, \end{aligned}$$

вероятность распространения ЛС от X_i до Y_j :

$$P_{P,ij} = \frac{N_{F_{0ij}} N_{F_{1ij}}}{2^{M_j-1}},$$

где $N_{F_{0ij}}$ и $N_{F_{1ij}}$ — число комбинаций входов схемы, при которых $F_{0ij} = 0$ и $F_{1ij} = 1$ соответственно; M_j — реальное число сигналов, от которых зависит Y_j . Если $F_{0ij} \equiv 0$, то $N_{F_{0ij}} = 1$; если $F_{1ij} \equiv 1$, то $N_{F_{1ij}} = 1$. Вероятность появления на выходе Y_j ЛС, наблюдаемого на выходе одной из ячеек схемы,

$$P_{P,j} = \sum_{i=1}^M \left(\frac{N_{F_{0ij}} N_{F_{1ij}}}{2^{M_j-1}} \prod_{k=1}^{i-1} \left(1 - \frac{N_{F_{0kj}} N_{F_{1kj}}}{2^{M_j-1}} \right) \right),$$

а вероятность появления сбоя, поразившего i -ю ячейку, на выходах схемы

$$P_{P,i} = \sum_{j=1}^{M_{\text{вых}}} \frac{N_{F_{0ij}} N_{F_{1ij}}}{2^{M_j-1}}. \quad (2)$$

Аналогичные вероятности могут быть рассчитаны для любой пары цепей схемы, что позволяет получить оценки чувствительности схемы к ЛС в ее ячейках уже на этапе ее логического синтеза.

Парафазное кодирование информационного сигнала

№	X	{X, XB}	Значение
1	—	00	Нулевой спейсер
2	0	01	Бит «0»
3	1	10	Бит «1»
4	—	11	Единичный спейсер

Сбой выхода СС-схемы часто маскируется СС-дисциплиной за счет избыточного (парафазного [9]) кодирования данных и двухфазной работы. При парафазном кодировании каждый синхронный сигнал X заменяется парафазным сигналом {X, XB}, как показано в таблице.

Парафазный сигнал формируется двумя согласованными логическими ячейками. Следовательно, однократный ЛС изменяет одну компоненту парафазного сигнала, делая его состояние не соответствующим текущей фазе схемы. Индикаторная подсхема обнаруживает это нарушение и останавливает функционирование СС-схемы до исчезновения сбоя.

Практические СС-схемы обычно реализуются в виде конвейера для повышения производительности аналогично синхронным схемам. В типовом СС-конвейере каждая ступень состоит из КЧ и выходного регистра (ВР), как показано на рисунке. Их индикаторные подсхемы ИКЧ и ИВР с помощью гистерезисного триггера (Г-триггера [9], на рисунке обозначен буквой Н) разрешают предыдущей ступени конвейера переключаться в следующую фазу работы.

Анализ возможных ситуаций появления ЛС в ступени конвейера показывает, что сбой в ее КЧ в худшем случае приводит лишь к приостановке работы конвейера, но не портит обрабатываемые данные, если разряды регистра ступени реализованы сбоеустойчивой схемой [16, Fig. 10]. Однако появление ЛС непосредственно в выходном регистре с вероятностью 0,25 вызывает искажение результата обработки данных или «зависание» конвейера. Использование в разряде регистра Г-триггера с DICE-подобной реализацией [16, Fig. 12] в 4 раза улучшает иммунитет регистра.

Индикаторные подсхемы КЧ и регистра ступени СС-конвейера вносят незначительный вклад в чувствительность СС-конвейера к ЛС. Критическая ситуация может возникнуть только в том случае, если ЛС поражает выходной Г-триггер, что в многозарядных СС-схемах крайне маловероятно.

Суммарная вероятность искажения данных в m -й ступени СС-конвейера с n_m -разрядным выходным регистром из-за ЛС зависит от площадей топологии КЧ ($S_{CP,m}$), Г-триггеров (S_H) и индикаторного элемента (S_{IE}) в разрядах регистра и индикаторных подсхем КЧ ($S_{CPI,m}$) и ВР ($S_{ORI,m}$):

$$P_{ST,m} = \frac{0,25n_m(2S_H + S_{IE})}{S_{CP,m} + n_m(2S_H + S_{IE}) + S_{CPI,m} + S_{ORI,m}} \quad (3)$$

Пусть КЧ содержит M ячеек и ее сложность в $K_{CP,m}$ раз превышает сложность регистра. Тогда формула (3) преобразуется к виду:

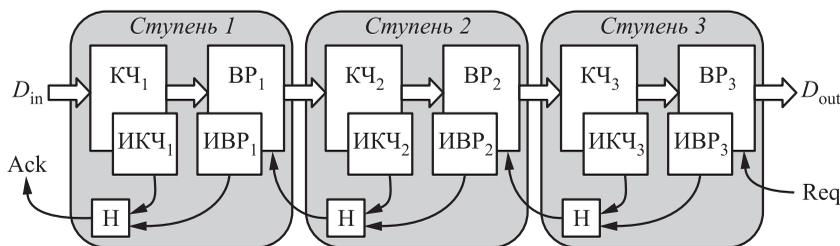
$$P_{ST,m} \approx \frac{0,55}{2,2K_{CP,m} + 2,7 + 0,25M/n_m}$$

При реальных значениях $K_{CP,m} = 4$ и $M = 8n_m$ вероятность критического сбоя $P_{ST,m} \approx 0,041$. При использовании DICE-подобного Г-триггера она уменьшается до величины $P_{DICE,m} \approx 0,011$, т.е. почти в 4 раза.

В синхронном конвейере однократный ЛС, наблюдаемый в любой части m -й ступени, не может замаскироваться дисциплиной работы. С учетом формулы (2) вероятность искажения данных из-за сбоя в КЧ m -й ступени

$$P_{S,m} = \sum_{i=1}^{M_S} \left(\frac{S_i}{S} \sum_{j=1}^{M_{S_{\text{вых}}}} \frac{N_{F_0ij} N_{F_1ij}}{2^{M_j-1}} \right),$$

где M_S — число ячеек в КЧ m -й ступени синхронного конвейера; $M_{S_{\text{вых}}}$ — число ее выходов. В первом приближении эту вероятность можно оценить как $P_{S,m} = 0,25$ [16].



Структура типового СС-конвейера

4 Сравнение сбоеустойчивости синхронных и самосинхронных схем

При заданной эксплуатационной плотности потока сбойных событий λ_0 интенсивность критических сбоев на выходах m -й ступени конвейера равна

$$\lambda_{CF,ST} \approx \lambda_{ST} P_{ST,m} = \frac{0,55\lambda_{ST}}{2,2K_{CP,m} + 2,7 + 0,25M/n_m}$$

для СС-конвейера и

$$\lambda_{CF,S} \approx \lambda_S P_{S,m} = 0,25\lambda_S$$

для синхронного конвейера. Следовательно, отношение интенсивностей критических сбоев для синхронного и самосинхронного конвейера:

$$K_{CF} = \frac{\lambda_{CF,S}}{\lambda_{CF,ST}} = \frac{0,25\lambda_S (2,2K_{CP,m} + 2,7 + 0,25M/n_m)}{0,55\lambda_{ST}}. \quad (4)$$

Поскольку синхронная КЧ в 2 раза проще, чем в СС-конвейере с парафазным кодированием, и индикаторной подсхемы нет, соотношение площадей топологий синхронного ($S_{S,P}$) и самосинхронного конвейеров ($S_{ST,P}$)

$$K_H = \frac{S_{ST,P}}{S_{S,P}} = 2 + \frac{4,25M + 8,5n_m}{20n_m(K_{CP,m} + 1)}.$$

Для реальных значений $K_{CP,m} = 4$, $M = 8n_m$ и $n_m = 32$ получается $K_H = 2,4$. Тогда в соответствии с формулами (1), (2) и (4) сбоеустойчивость СС-конвейера оказывается лучше сбоеустойчивости синхронного конвейера в 2,5–9,4 раза в зависимости от схемы реализации разряда СС-регистра.

5 Заключение

Самосинхронные схемы обладают естественной высокой иммунностью к ЛС благодаря избыточному кодированию данных, двухфазной работе и контролю окончания переключения в каждую фазу. Анализ сбойных ситуаций показывает, что наиболее чувствительной к однократным ЛС частью ступени СС-конвейера оказывается ВР. Однако реализация его разрядов на DICE-подобных Г-триггерах повышает его сбоеустойчивость в 4 раза.

Реализация цифровой схемы в виде СС-конвейера гарантирует повышение ее устойчивости

к однократным ЛС в 2,5–9,4 раза в сравнении с синхронным конвейером, причем при появлении критического сбоя СС-конвейер останавливается и своими индикаторными сигналами локализует место сбоя. Недостатком такой реализации является увеличенная в 2,4 раза сложность и, соответственно, площадь схемы в топологии.

Литература

1. *Викторова В. С., Лубков Н. В., Степанянц А. С.* Анализ надежности отказоустойчивых управляющих вычислительных систем. — М.: ИПУ РАН, 2016. 117 с.
2. *Morelos-Zaragoza R. H.* The art of error correcting coding. — 2nd ed. — Hoboken, NJ, USA: Wiley, 2006. 269 p.
3. *LaFrieda C., Manohar R.* Fault detection and isolation techniques for quasi delay-insensitive circuits // Conference (International) on Dependable Systems and Networks, 2004. P. 41–50. doi: 10.1109/DSN.2004.1311875.
4. *Monnet Y., Renaudin M., Leveugle R.* Hardening techniques against transient faults for asynchronous circuits // 11th On-Line Testing Symposium (International) Proceedings. — IEEE, 2005. P. 129–134.
5. *Dug M., Krstic M., Jokic D.* Implementation and analysis of methods for error detection and correction on FPGA // IFAC-PapersOnLine, 2018. Vol. 51. No. 6. P. 348–353.
6. *Lodhi F. K., Hasan S., Hasan O., Awwad F.* Low power soft error tolerant macro synchronous micro asynchronous (MSMA) pipeline // Computer Society Annual Symposium on VLSI Proceedings. — Piscataway, NJ, USA: IEEE, 2014. P. 601–606. doi: 10.1109/ISVLSI.2014.59.
7. *Gkiokas C., Schoeberl M. A.* Fault-tolerant time-predictable processor // Nordic Circuits and Systems Conference: NORCHIP and Symposium (International) of System-on-Chip Proceedings. — Piscataway, NJ, USA: IEEE, 2019. Art. 8906947. 6 p. doi: 10.1109/NORCHIP.2019.8906947.
8. *Muller D., Bartky W.* A theory of asynchronous circuits // Symposium (International) on the Theory of Switching Proceedings. — Harvard University Press, 1959. Vol. 29. P. 204–243.
9. *Kishinevsky M., Kondratyev A., Taubin A., Varshavsky V.* Concurrent hardware: The theory and practice of self-timed design. — New York, NY, USA: John Wiley & Sons, 1994. 368 p.
10. *Smith S. C., Jia Di.* Designing asynchronous circuits using NULL convention logic (NCL). — Synthesis lectures on digital circuits and systems ser. — Cham: Springer, 2009. 96 p.
11. *Stepchenkov Y. A., Kamenskih A. N., Diachenko Y. G., Rogdestvenski Y. V., Diachenko D. Y.* Improvement of the natural self-timed circuit tolerance to short-term soft errors // Advances Science Technology Engineering Systems J., 2020. Vol. 5. No. 2. P. 44–56.
12. *Соколов И. А., Степченко Ю. А., Рождественский Ю. В., Дьяченко Ю. Г.* Приближенная оценка

эффективности синхронной и самосинхронной методологий в задачах проектирования сбоеустойчивых вычислительно-управляющих систем // Автоматика и телемеханика, 2022. № 2. С. 122–132.

13. Dubrova E. Fault-tolerant design. — New York, NY, USA: Springer, 2013. 185 p. doi: 10.1007/978-1-4614-2113-9.
14. Eaton P., Benedetto J., Mavis D., Avery K., Sibley M., Gadge M., Turflinger T. Single event transient pulse width measurements using a variable temporal latch technique //

IEEE T. Nucl. Sci., 2004. Vol. 51. No. 6. P. 3365–3368. doi: 10.1109/TNS.2004.840020.

15. Emeliyanov V. V., Vataev A. S., Useinov R. G. Impact of heavy ion energy on charge yield in silicon dioxide // IEEE T. Nucl. Sci., 2018. Vol. 65. No. 8. P. 1496–1502.
16. Stepchenkov Y., Diachenko Y., Rogdestvenski Y., Shikunov Y., Diachenko D. Self-timed storage register soft error tolerance improvement // East–West Design & Test Symposium Proceedings. — Piscataway, NJ, USA: IEEE, 2021. P. 145–150.

Поступила в редакцию 20.06.22

SYNCHRONOUS AND SELF-TIMED PIPELINE'S RELIABILITY ESTIMATION

I. A. Sokolov, Yu. A. Stepchenkov, Yu. G. Diachenko, and Yu. V. Rogdestvenski

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

Abstract: Self-timed (ST) circuitry is an alternative to synchronous circuits. Self-timed circuits have a number of advantages over their synchronous counterparts due to their redundant complexity. The article investigates the immunity of self-timed and synchronous circuits to single short-term soft error taking into account the hardware redundancy of ST circuits. Self-timed circuits, due to their indication subcircuit, are able to detect a soft error which occurs as a logical cell's output state inversion and suspend the operation of the circuit until the soft error disappears. Thus, ST circuits mask a single soft error and prevent distortion of the data processing result. The use of a modified hysteretic trigger, which prevents sticking in the antispacer, to implement a pipeline stage register bit masks almost all soft errors in the pipeline stage's combinational part. The DICE-like implementation of this trigger makes it possible to reduce the sensitivity of the ST register to the internal soft errors by a factor of 4. Quantitative estimates of failure tolerance show a clear (by 2.5–9.4 times) advantage of the ST pipeline in comparison with the synchronous counterpart.

Keywords: self-timed circuit; soft error; failure tolerance; pipeline; indication; probabilistic estimate

DOI: 10.14357/19922264220401

Acknowledgments

The research was supported by the Russian Science Foundation (project No. 22-19-00237).

References

1. Viktorova, V. S., N. V. Lubkov, and A. S. Stepanyants. 2016. *Analiz nadezhnosti otkazoustoychivyykh upravlyayushchikh vychislitel'nykh sistem* [Analysis of fault-tolerant computing systems' reliability]. Moscow: IPU RAN. 117 p.
2. Morelos-Zaragoza, R. H. 2006. *The art of error correcting coding*. Hoboken, NJ: Wiley. 269 p.
3. LaFrieda, C., and R. Manohar. 2004. Fault detection and isolation techniques for quasi delay-insensitive circuits. *Conference (International) on Dependable Systems and Networks Proceedings*. 41–50. doi: 10.1109/DSN.2004.1311875.
4. Monnet Y., M. Renaudin, and R. Leveugle. 2005. Hardening techniques against transient faults for asynchronous circuits. *11th On-Line Testing Symposium (International) Proceedings*. IEEE. 129–134.
5. Dug, M., M. Krstic, and D. Jokic. 2018. Implementation and analysis of methods for error detection and correction on FPGA. *IFAC-PapersOnLine* 51(6):348–353.
6. Lodhi, F. K., S. Hasan, O. Hasan, and F. Awwad. 2014. Low power soft error tolerant macro synchronous micro asynchronous pipeline. *Computer Society Annual Symposium on VLSI Proceedings*. Piscataway, NJ: IEEE. 601–606. doi: 10.1109/ISVLSI.2014.59.
7. Gkiokas, C., and M. A. Schoeberl. 2019. Fault-tolerant time-predictable processor. *Nordic Circuits and Systems Conference: NORCHIP Symposium (International) of System-on-Chip Proceedings*. Piscataway, NJ: IEEE. 8906947. 6 p. doi: 10.1109/NORCHIP.2019.8906947.

8. Muller, D. E., and W. C. Bartky. 1959. A theory of asynchronous circuits. *Symposium (International) on the Theory of Switching Proceedings*. Harvard University Press. 29:204–243.
9. Kishinevsky, M., A. Kondratyev, A. Taubin, and V. Varshavsky. 1994. *Concurrent hardware: The theory and practice of self-timed design*. New York, NY: John Wiley & Sons. 368 p.
10. Smith, S. C., and J. Di. 2009. *Designing asynchronous circuits using NULL convention logic (NCL)*. Synthesis lectures on digital circuits systems ser. Cham: Springer. 96 p.
11. Stepchenkov, Y. A., A. N. Kamenskih, Y. G. Diachenko, Y. V. Rogdestvenski, and D. Y. Diachenko. 2020. Improvement of the natural self-timed circuit tolerance to short-term soft errors. *Advances Science Technology Engineering Systems J.* 5(2):44–56.
12. Sokolov, I. A., Yu. A. Stepchenkov, Yu. V. Rozhdestvenskiy, and Yu. G. Diachenko. 2022. An approximate evaluation of the efficiency of synchronous and self-timed methodologies in designing failure-tolerant computing and control systems. *Automat. Rem. Contr.* 83(2):264–272.
13. Dubrova, E. 2013. *Fault-tolerant design*. New York, NY: Springer. 185 p. doi: 10.1007/978-1-4614-2113-9.
14. Eaton, P., J. Benedetto, D. Mavis, K. Avery, M. Sibley, M. Gadlage, and T. Turflinger. 2004. Single event transient pulse width measurements using a variable temporal latch technique. *IEEE T. Nucl. Sci.* 51(6):3365–3368. doi: 10.1109/TNS.2004.840020.
15. Emeliyanov, V. V., A. S. Vátuev, and R. G. Useinov. 2018. Impact of heavy ion energy on charge yield in silicon dioxide. *IEEE T. Nucl. Sci.* 65(8):1496–1502.
16. Stepchenkov, Y., Y. Diachenko, Y. Rogdestvenski, Y. Shikunov, and D. Diachenko. 2021. Self-timed storage register soft error tolerance improvement. *East–West Design & Test Symposium Proceedings*. Piscataway, NJ: IEEE. 145–150.

Received June 20, 2022

Contributors

Sokolov Igor A. (b. 1954) — Doctor of Science in technology, Academician of RAS, director, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; isokolov@ipiran.ru

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

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

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

ОБЩИЙ ПОРЯДОК АППРОКСИМАЦИИ ОЦЕНОК ФИЛЬТРАЦИИ СОСТОЯНИЙ МАРКОВСКИХ СКАЧКООБРАЗНЫХ ПРОЦЕССОВ ПО ДИСКРЕТИЗОВАННЫМ НАБЛЮДЕНИЯМ*

А. В. Борисов¹

Аннотация: Заметка продолжает исследования, посвященные численной аппроксимации оценок фильтрации состояний марковских скачкообразных процессов (МСП) по считающим и диффузионным наблюдениям с мультипликативными шумами. Оценки аппроксимируются с использованием наблюдений, дискретизованных по времени. В отличие от ранних алгоритмов, ограничивающих возможное число скачков состояния на интервале дискретизации, новые оценки предлагается вычислять без подобных ограничений с помощью единой рекуррентной схемы. В заметке получена верхняя граница показателя точности аппроксимации как функция параметров системы наблюдения, используемой схемы численного интегрирования, шага дискретизации и момента оценивания. Численный пример иллюстрирует сублинейное поведение данного показателя в зависимости от последнего аргумента.

Ключевые слова: марковский скачкообразный процесс; оптимальная фильтрация; диффузионные и считающие наблюдения; мультипликативные шумы в наблюдениях; точность численной аппроксимации

DOI: 10.14357/19922264220402

1 Введение

Данная заметка продолжает исследования в области разработки численных методов фильтрации состояний МСП по совокупности диффузионных и считающих наблюдений [1–4]. Интенсивность шумов в диффузионных наблюдениях зависит от оцениваемого состояния.

В [1, 2] была решена задача фильтрации при наличии только диффузионных наблюдений. Помимо аналитического решения был предложен ряд численных алгоритмов, основанных на обработке наблюдений, дискретизованных по времени. Алгоритмы строились, исходя из предположения, что на интервале дискретизации оцениваемый МСП может совершить число скачков, ограниченное некоторой величиной. Был получен показатель точности таких аппроксимаций, как функция параметров системы наблюдения, максимального числа скачков, величины шага временной дискретизации и точности применяемой схемы численного интегрирования.

В работе [3] представлено решение задачи фильтрации по совокупности диффузионных и считающих наблюдений, а в [4] предложен метод приближенного вычисления данной оценки. Следует отметить, что не все схемы численного интегрирования предполагают ограничение числа возможных скачков МСП на интервале дискретизации: напри-

мер, метод квадратур Гаусса может быть использован без данного условия. В [4] предложен *общий показатель* точности в форме средней \mathcal{L}_1 -нормы ошибки аппроксимации оценки фильтрации и доказано утверждение о его локальной зависимости от шага временной дискретизации и схемы численного интегрирования.

Целью данной работы ставится определение глобального поведения показателя точности аппроксимации, т.е. нахождение зависимости данной величины от шага временной дискретизации δ и момента оценивания T .

Статья организована следующим образом. Раздел 2 содержит формальную постановку задачи фильтрации состояний МСП по диффузионным и считающим наблюдениям, дискретизованным по времени с шагом δ . Аналитическое решение данной задачи представлено в разд. 3 в виде формулы Байеса, содержащей интегралы по абстрактной вероятностной мере в числителе и знаменателе. Для ее практической реализации интегралы предлагается заменить некоторыми интегральными суммами. Точность аппроксимации также описывается в форме средней \mathcal{L}_1 -нормы разности оценки фильтрации и ее приближения. Основным результатом работы является утверждение, представляющее различные границы сверху этого показателя точности. Показано, что эти границы представляют собой функции

* Работа выполнялась с использованием инфраструктуры Центра коллективного пользования «Высокопроизводительные вычисления и большие данные» (ЦКП «Информатика») ФИЦ ИУ РАН.

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

шага временной дискретизации δ и момента оценивания T . В зависимости от схемы численного интегрирования, применяемого на одном шаге, верхняя граница может быть как линейной, так и *сублинейной* функцией времени T . Раздел 4 содержит численный пример, иллюстрирующий различие предложенных в работе границ точности. В разд. 5 представлены выводы и заключительные замечания.

2 Постановка задачи фильтрации состояния марковских скачкообразных процессов по дискретизованным наблюдениям

На полном вероятностном пространстве с фильтрацией $(\Omega, \mathcal{F}, \mathcal{P}, \{\mathcal{F}_t\}_{t \geq 0})$ рассматривается непрерывно-дискретная стохастическая динамическая система наблюдения

$$X_t = X_0 + \int_0^t \Lambda^\top X_s ds + \mu_t^X, \quad (1)$$

где $X_t \in \mathbb{S}^N$ — ненаблюдаемое состояние системы — однородный МСП с множеством состояний $\mathbb{S}^N \triangleq \{e_1, \dots, e_N\}$ (\mathbb{S}^N — множество единичных векторов пространства \mathbb{R}^N), матрицей интенсивностей переходов Λ и начальным распределением π ; $\mu_t^X \in \mathbb{R}^N$ — \mathcal{F}_t -согласованный мартингал;

$$\mathcal{Y}_r = \int_{t_{r-1}}^{t_r} f X_s ds + \int_{t_{r-1}}^{t_r} \sum_{n=1}^N X_s^n g_n^{1/2} dW_s, \quad r \in \mathbb{N}, \quad t_r = r\delta, \quad (2)$$

где $\mathcal{Y}_r \in \mathbb{R}^M$ — диффузионные наблюдения $Y_t = \int_0^t f X_s ds + \int_0^t \sum_{n=1}^N X_s^n g_n^{1/2} dW_s$, дискретизованные по времени с шагом δ , зашумленные \mathcal{F}_t -согласованным стандартным винеровским процессом $W_t \in \mathbb{R}^M$; $(M \times N)$ -мерная матрица f характеризует план наблюдений, а набор $(M \times M)$ -мерных симметричных матриц $\{g_n\}_{n=1, \dots, N}$ определяет интенсивности шумов в зависимости от текущего состояния X_t ;

$$\mathcal{Z}_r = \int_{t_{r-1}}^{t_r} h X_s ds + (\mu_{t_r}^Z - \mu_{t_{r-1}}^Z), \quad r \in \mathbb{N}, \quad t_r = r\delta, \quad (3)$$

где $\mathcal{Z}_r \in \mathbb{R}^K$ — считающие наблюдения $Z_t = \int_0^t h X_s ds + \mu_t^Z$, дискретизованные по времени

с шагом δ : элементы $(K \times N)$ -мерной матрицы h определяют интенсивность скачков отдельных компонент в зависимости от текущего состояния X_t ; $\mu_t^Z \in \mathbb{R}^K$ — \mathcal{F}_t -согласованный мартингал.

Обозначим через $\mathfrak{D}_r \triangleq \{\mathcal{Y}_q, \mathcal{Z}_q : q = \overline{1, r}\}$ последовательность σ -алгебр всех дискретизованных наблюдений, полученных на отрезке времени $[0, t_r]$, $r \in \mathbb{N}$, $\mathfrak{D}_0 \triangleq \{\emptyset, \Omega\}$.

Задача оптимальной фильтрации состояния МСП X_t по дискретизованным наблюдениям заключается в построении условного математического ожидания $\hat{X}_r \triangleq \mathbf{E} \{X_{t_r} | \mathfrak{D}_r\}$, $r \in \mathbb{N}$.

Система наблюдения (1)–(3) удовлетворяет следующим ограничениям.

- A. Исследуемый вероятностный базис с фильтрацией является пространством Винера–Пуассона [5].
- B. Шумы в \mathcal{Y} равномерно невырождены [6], т. е. $\min_{1 \leq n \leq N} g_n > 0$.
- B. Компоненты мартингала μ^Z в наблюдениях \mathcal{Z} (3) ортогональны друг другу, т. е. $\langle \mu^Z \rangle_t = \int_0^t \text{diag}(h X_s) ds$. Мартингалы в уравнении состояния μ^X и в наблюдениях μ^Z также ортогональны $\langle \mu^X, \mu^Z \rangle_t \equiv 0$.

3 Аналитическое решение задачи и численная аппроксимация

Ниже в статье будут использоваться следующие обозначения.

1. $\|\alpha\|_Q^2 \triangleq \alpha^\top Q \alpha$.

2. Для $X = \text{col}(x_1, \dots, x_N) \in \mathbb{R}^N$

$$\|X\|_1 \triangleq \sum_{n=1}^N |x_n|;$$

для $A = \|a_{ij}\|_{i,j=\overline{1,N}} \in \mathbb{R}^{N \times N}$

$$\|A\|_1 \triangleq \max_{1 \leq j \leq N} \sum_{i=1}^N |a_{ij}|.$$

3. Единичная матрица подходящей размерности: I ; нулевая матрица подходящей размерности: $\mathbf{0}$; вектор-строка подходящей размерности, составленная из единиц: $\mathbf{1}$.
4. Характеристическая функция множества \mathcal{A} : $\mathbf{1}_{\mathcal{A}}(x)$.
5. Вектор, компоненты которого равны случайному времени пребывания процесса X в каждом из возможных состояний на отрезке $[t_{r-1}, t_r]$:

$$\tau_r \triangleq \int_{t_{r-1}}^{t_r} X_s ds.$$

6. $(N - 1)$ -мерный симплекс в пространстве \mathbb{R}^N , носитель распределения вектора τ_r :

$$D \triangleq \left\{ u = \text{col} (u^1, \dots, u^N) : u^n \geq 0, \sum_{n=1}^N u^n = \delta \right\}.$$

7. «Вероятностный симплекс», множество возможных начальных распределений π :

$$\Pi \triangleq \left\{ \pi = \text{col} (\pi^1, \dots, \pi^N) : \pi^n > 0, \sum_{n=1}^N \pi^n = 1 \right\}.$$

8. Распределение вектора $X_{t_r}^\ell \tau_r$ при условии $X_{t_{r-1}} = e_k$: $\mathbf{P}^{k\ell}(du)$, т. е. для любого $\mathcal{G} \in \mathcal{B}(\mathbb{R}^N)$ верно равенство

$$\mathbf{E} \{ \mathbf{I}_{\mathcal{G}}(\tau_r) X_{t_r}^\ell | X_{t_{r-1}} = e_k \} = \int_{\mathcal{G}} \mathbf{P}^{k\ell}(du).$$

9. M -мерная плотность гауссовского распределения с математическим ожиданием m и невырожденной ковариационной матрицей Q :

$$\mathcal{N}(y, m, Q) \triangleq (2\pi)^{-M/2} \det^{-1/2} Q \exp \left\{ -\frac{1}{2} \|y - m\|_{Q^{-1}}^2 \right\}.$$

10. Распределение K -мерного случайного вектора с независимыми компонентами, имеющими пуассоновское распределение с параметрами $\alpha^k, k = \overline{1, K}$:

$$\mathcal{P}(z, \alpha) \triangleq e^{-\sum_{k=1}^K \alpha^k} \prod_{k=1}^K \frac{(\alpha^k)^{z^k}}{(z^k)!},$$

где

$$z = \text{col} (z^1, \dots, z^K) \in \mathbb{Z}_+^K; \\ \alpha = \text{col} (\alpha^1, \dots, \alpha^K) \in \mathbb{R}_+^K.$$

11. Функции

$$\theta^{kj} = \theta^{kj}(y, z) \triangleq \int_{\mathcal{D}} \mathcal{N} \left(y, fu, \sum_{p=1}^N u^p g_p \right) \mathcal{P}(z, hu) \mathbf{P}^{kj}(du).$$

Матрицы, составленные из этих функций:

$$\theta = \theta(y, z) \triangleq \|\theta^{kj}(y, z)\|_{k,j=\overline{1,N}}.$$

Матричнозначные функции случайных величин — дискретизованных наблюдений, полученных в момент времени t_r :

$$\theta_r \triangleq \theta(\mathcal{Y}_r, \mathcal{Z}_r).$$

Произведения матриц $\{\theta_i\}$ с нарастающим итогом:

$$\Theta_r = \Theta(y_1, \dots, y_r, z_1, \dots, z_r) \triangleq \theta(y_1, z_1) \theta(y_2, z_2) \cdots \theta(y_r, z_r).$$

Матричнозначные функции дискретизованных наблюдений, полученных на промежутке времени $[0, t_r]$:

$$\Theta_r \triangleq \theta_1 \theta_2 \cdots \theta_r.$$

Используя введенные выше определения, оптимальную оценку $\hat{\mathcal{X}}_r$ [4] по дискретизованным наблюдениям можно записать в явном виде:

$$\hat{\mathcal{X}}_r = \frac{1}{\mathbf{1} \Theta_r^\top \pi} \Theta_r^\top \pi.$$

Матричнозначные функции θ и Θ_r обладают следующими очевидными свойствами:

- (1) $\sum_{z \in \mathbb{Z}_+^K} \int_{\mathbb{R}^M} \theta^{kj}(y, z) dy = \mathbf{P} \{ X_h = e_j | X_0 = e_k \} = (\exp\{h\Lambda\})^{kj}$;
- (2) $\sum_{z \in \mathbb{Z}_+^K} \int_{\mathbb{R}^M} \|\theta^\top(y, z)\|_1 dy = 1$;
- (3) $\sum_{(z_1, \dots, z_r) \in \mathbb{Z}_+^{rK}} \int_{\mathbb{R}^{rM}} \|\Theta_r^\top(y_1, \dots, y_r, z_1, \dots, z_r)\|_1 dy_1 \cdots dy_r = 1$.

Интегралы в θ^{kj} не могут быть вычислены аналитически, поэтому их предлагается аппроксимировать интегральными суммами общего вида. Для этого будем использовать следующие функции и матрицы, составленные из них:

– аппроксимация θ^{kj}

$$\psi^{kj} \triangleq \sum_{\ell=1}^L \mathcal{N} \left(y, fu_\ell, \sum_{p=1}^N u_\ell^p g_p \right) \mathcal{P}(z, hu_\ell) \rho_\ell^{kj};$$

– аппроксимация θ

$$\psi = \psi(y, z) \triangleq \|\psi^{kj}(y, z)\|_{k,j=\overline{1,N}};$$

– аппроксимация θ_r

$$\psi_r \triangleq \psi(\mathcal{Y}_r, \mathcal{Z}_r);$$

– аппроксимация Θ_r

$$\Psi_r = \Psi(y_1, \dots, y_r, z_1, \dots, z_r) \triangleq \psi(y_1, z_1) \psi(y_2, z_2) \cdots \psi(y_r, z_r);$$

– аппроксимация Θ_r

$$\Psi_r \triangleq \psi_1 \psi_2 \cdots \psi_r.$$

В определенных выше интегральных суммах $\{u_\ell\}_{\ell=1, L} \subset \mathcal{D}$ ($u_\ell = u_\ell(\delta)$) — набор точек; $\{\rho_\ell^{kj}\}_{\ell, k, j} \subset \mathbb{R}_+$ ($\rho_\ell^{kj} = \rho_\ell^{kj}(\delta)$) — набор весов. Будем использовать только аппроксимации, обладающие следующим свойством:

$$(4) \max_{1 \leq k \leq N} \sum_{j=1}^N \sum_{\ell=1}^L \rho_\ell^{kj}(\delta) \triangleq \mathfrak{W}(\delta) \leq 1.$$

Условие 4 обеспечивает аппроксимациям выполнение свойств, близких по смыслу 1–3:

$$(5) \sum_{z \in \mathbb{Z}_+^K} \int_{\mathbb{R}^M} \psi^{kj}(y, z) dy = \sum_{\ell=1}^L \rho_\ell^{kj} \leq \mathfrak{W}(\delta) \leq 1;$$

$$(6) \sum_{z \in \mathbb{Z}_+^K} \int_{\mathbb{R}^M} \|\psi^\top(y, z)\|_1 dy \leq \mathfrak{W}(\delta) \leq 1;$$

$$(7) \sum_{(z_1, \dots, z_r) \in \mathbb{Z}_+^{rK}} \int_{\mathbb{R}^{rM}} \|\Psi_r^\top(y_1, \dots, y_r, z_1, \dots, z_r)\|_1 dy_1 \cdots dy_r \leq \mathfrak{W}^r(\delta) \leq 1.$$

Таким образом, вместо $\widehat{\mathcal{X}}_r$ можно вычислить ее аппроксимацию

$$\overline{\mathcal{X}}_r = \frac{1}{\mathbf{1} \Psi_r^\top \pi} \Psi_r^\top \pi.$$

Определим глобальную характеристику близости оценки $\{\widehat{\mathcal{X}}\}$ и ее аппроксимации $\{\overline{\mathcal{X}}\}$ следующим образом. Пусть $T > 0$ — фиксированный момент времени, $r \triangleq T/\delta$ (без ограничения общности считаем, что T делится на δ нацело). Показатель точности определяется следующим образом:

$$\Sigma(T, \delta) \triangleq \sup_{\pi \in \Pi} \mathbf{E}_\pi \left\{ \left\| \widehat{\mathcal{X}}_{T/\delta} - \overline{\mathcal{X}}_{T/\delta} \right\|_1 \right\}.$$

Теорема 1. Пусть для аппроксимации ψ выполнено условие

$$\sum_{z \in \mathbb{Z}_+^K} \int_{\mathbb{R}^M} \|(\psi(y, z) - \theta(y, z))^\top\|_1 dy \leq C \delta^{1+\alpha} \quad (4)$$

при некоторых параметрах C , $\alpha > 0$. Тогда

$$\Sigma(T, \delta) \leq 2C \frac{1 - \mathfrak{W}^{T/\delta}(\delta)}{1 - \mathfrak{W}(\delta)} \delta^{1+\alpha}, \text{ если } \mathfrak{W}(\delta) < 1; \quad (5)$$

$$\Sigma(T, \delta) \leq 2TC \delta^\alpha, \quad \text{если } \mathfrak{W}(\delta) = 1. \quad (6)$$

Доказательство теоремы 1 приведено в приложении.

Суммарный вес \mathfrak{W} зависит не только от шага дискретизации δ , но и от матрицы Λ в уравнении динамики (1). В то же время значение C зависит от параметров f , $\{g_n\}$ и h модели наблюдения (2) и (3).

4 Численный пример

Проиллюстрируем различия верхних границ (5) и (6) на конкретном примере. В статье [2] была предложена составная схема «средних» прямоугольников для фильтрации состояний МСП по

наблюдениям с мультипликативными шумами при наличии только диффузионных наблюдений. При наличии дополнительных считающих наблюдений эта схема принимает вид

$$\begin{aligned} \psi^{kj} \triangleq & \delta_{kj} e^{\lambda_{jj} \delta} \mathcal{N}(y, \delta f^j, \delta g_j) \mathcal{P}(y, \delta f^j) + \\ & + (1 - \delta_{kj}) \lambda_{kj} \delta^{1+\alpha} e^{\lambda_{jj} \delta} \times \\ & \times \sum_{\ell=1}^{[\delta^{-\alpha}]} e^{(\lambda_{kk} - \lambda_{jj}) u} \mathcal{N}(y, u f^k + (\delta - u) f^j, u g_k + \\ & + (\delta - u) g_j) \mathcal{P}(y, u h^k + (\delta - u) h^j) \Big|_{u=(\ell-1/2)\delta^{1+\alpha}}, \end{aligned}$$

где f^j и h^j — j -е столбцы матриц f и h .

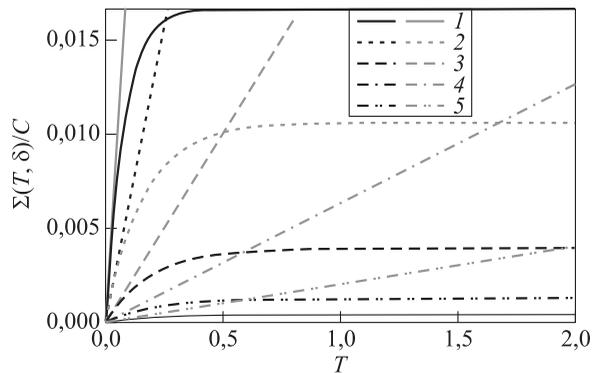
Действуя аналогично [2], можно показать, что для данной схемы вычисления коэффициентов ψ^{kj} выполнено условие (4), в котором величина $C > 0$ зависит от параметров модели наблюдения, а параметр $\alpha > 0$ описывает степень «дополнительного дробления» отрезка дискретизации δ .

Весовые коэффициенты в данной схеме выражаются следующим образом:

$$\begin{aligned} \varrho_\ell^{kj} = & \delta_{kj} e^{\lambda_{jj} \delta} + \\ & + (1 - \delta_{kj}) \lambda_{kj} \delta^{1+\alpha} \exp \left(\delta \left(\lambda_{kk} \left(\ell - \frac{1}{2} \right) \delta^\alpha + \right. \right. \\ & \left. \left. + \lambda_{jj} \left(1 - \left(\ell - \frac{1}{2} \right) \delta^\alpha \right) \right) \right). \end{aligned}$$

Используя второй замечательный предел, можно показать, что $\mathfrak{W}(\delta) < 1$ для достаточно малого $\delta > 0$.

Численный эксперимент был проведен для значений параметров $N = 10$, $\lambda_{ij} \equiv 5$ при $i \neq j$, $T \in [0; 2]$. Показатель точности $\Sigma(T, \delta)$ прямо пропорционален параметру C , поэтому нормируем его на C и исследуем частное $\Sigma(T, \delta)/C$.



Зависимость $\Sigma(T, \delta)/C$ от T , вычисленная по формулам (5) (черные кривые) и (6) (серые кривые) при различных параметрах δ : $1 - \delta = 10^{-2}$; $2 - 10^{-3}$; $3 - 10^{-4}$; $4 - 10^{-5}$; $5 - \delta = 10^{-6}$

На рисунке представлена зависимость от T функции $\Sigma(T, \delta)/C$ при различных значениях шага дискретизации δ .

Анализируя графики, легко прийти к следующему заключению. Если удастся построить схемы численного интегрирования, для которых выполняется условие $\mathfrak{W}(\delta) < 1$, то пессимистичные линейные оценки (6) верхней границы $\Sigma(T, \delta)$ могут быть существенно уточнены с использованием формулы (5).

5 Заключение

Формулы (5) и (6), описывающие оценки сверху для средней \mathcal{L}_1 -нормы ошибки аппроксимации как функции шага временной дискретизации δ и момента построения оценки фильтрации T , являются главными результатами данной работы. Тот факт, что в (6) порядок глобальной точности аппроксимации на 1 меньше порядка локальной точности, вполне ожидаем: потеря порядка точности на единицу наблюдается для численных методов решения как обыкновенных, так и стохастических дифференциальных уравнений [7]. Линейный рост величины ошибки, определяемой формулой (6), с ростом момента оценивания T также представляется достаточно рутинным и следует из дискретного варианта неравенства Гронуолла–Беллмана и свойств 1–3.

Вместе с тем вариант верхней оценки глобального показателя точности, описываемый (5), обладает несколькими необычными качествами. Во-первых, данная оценка демонстрирует не линейный, но *сублинейный* рост по T . Во-вторых, порядок глобальной точности понизился меньше чем на единицу. Эти результаты получены благодаря тому, что матричнозначные функции Θ_r с единичной нормой удалось приблизить функциями Ψ_r , имеющими норму, меньшую 1. Именно это качество позволяет описать эволюцию во времени верхней границы с помощью устойчивой рекуррентной зависимости, доставляющей в итоге менее консервативную оценку показателю точности.

Приложение

Доказательство теоремы 1. Преобразуем ошибку аппроксимации следующим образом:

$$\begin{aligned} \Delta_r &= \bar{\mathcal{X}}_r - \hat{\mathcal{X}}_r = \\ &= \frac{1}{\mathbf{1}\Psi_r^\top \pi \mathbf{1}\Theta_r^\top \pi} \left(\mathbf{1}\Theta_r^\top \pi \Psi_r^\top \pi - \mathbf{1}\Psi_r^\top \pi \Theta_r^\top \pi \right) = \\ &= \frac{1}{\mathbf{1}\Psi_r^\top \pi \mathbf{1}\Theta_r^\top \pi} \left(\Psi_r^\top \pi \mathbf{1}\Theta_r^\top \pi - \mathbf{1}\Psi_r^\top \pi \Theta_r^\top \pi \right) = \end{aligned}$$

$$\begin{aligned} &= \frac{1}{\mathbf{1}\Psi_r^\top \pi \mathbf{1}\Theta_r^\top \pi} \left(\Psi_r^\top \pi \mathbf{1} - \mathbf{1}\Psi_r^\top \pi \mathbf{1} \right) \Theta_r^\top \pi = \\ &= \frac{1}{\mathbf{1}\Psi_r^\top \pi \mathbf{1}\Theta_r^\top \pi} \left(\Psi_r^\top \pi \mathbf{1} - \mathbf{1}\Psi_r^\top \pi \mathbf{1} \right) \left(\Theta_r^\top \pi - \Psi_r^\top \pi \right) = \\ &= \left(\bar{\mathcal{X}}_r \mathbf{1} - \mathbf{1} \right) \frac{1}{\mathbf{1}\Theta_r^\top \pi} \left(\Theta_r - \Psi_r \right)^\top \pi. \end{aligned}$$

Из свойств нормы $\|\cdot\|_1$ следует, что $\|\pi\|_1 \equiv 1$ и $\|\bar{\mathcal{X}}_r \mathbf{1} - \mathbf{1}\|_1 \leq 2$, поэтому

$$\|\Delta_r\|_1 \leq 2 \left\| \frac{1}{\mathbf{1}\Theta_r^\top \pi} \left(\Theta_r - \Psi_r \right)^\top \right\|_1. \quad (7)$$

Используя обозначения $\gamma_r \triangleq \psi_r - \theta_r$ и $\Gamma_r \triangleq \Psi_r - \Theta_r$, можно записать рекурсию

$$\Gamma_r = \Gamma_{r-1} \psi_r + \Theta_{r-1} \gamma_r, \quad r > 1, \quad \Gamma_1 = \gamma_1. \quad (8)$$

Согласно лемме 3 [1] для любой интегрируемой функции $a(y_1, \dots, y_r, z_1, \dots, z_r)$ имеет место равенство:

$$\begin{aligned} \mathbf{E}_\pi \left\{ \frac{a(\mathcal{Y}_1, \dots, \mathcal{Y}_r, \mathcal{Z}_1, \dots, \mathcal{Z}_r)}{\mathbf{1}\Theta_r^\top (\mathcal{Y}_1, \dots, \mathcal{Y}_r, \mathcal{Z}_1, \dots, \mathcal{Z}_r) \pi} \right\} &= \\ = \sum_{(z_1, \dots, z_r) \in \mathbb{Z}_+^{rK} \mathbb{R}^{rM}} \int a(y_1, \dots, y_r, z_1, \dots, z_r) dy_1 \cdots dy_r. \end{aligned}$$

Взяв математическое ожидание левой и правой частей (7), благодаря утверждению данной леммы получаем

$$\mathbf{E}_\pi \{ \|\Delta_r\|_1 \} \leq 2 \sum_{(z_1, \dots, z_r) \in \mathbb{Z}_+^{rK} \mathbb{R}^{rM}} \int \left\| \Gamma_r^\top (y_1, \dots, y_r, z_1, \dots, z_r) \right\|_1 dy_1 \cdots dy_r. \quad (9)$$

Обозначим

$$J_r \triangleq \sum_{(z_1, \dots, z_r) \in \mathbb{Z}_+^{rK} \mathbb{R}^{rM}} \int \left\| \Gamma_r^\top (y_1, \dots, y_r, z_1, \dots, z_r) \right\|_1 dy_1 \cdots dy_r.$$

Тогда, вычисляя интеграл нормы $\|\cdot\|_1$ от левой и правой частей (8) и учитывая свойства 3, 6 и 7, для $r > 1$ можно получить рекуррентное неравенство $J_r \leq \mathfrak{W}(\delta) J_{r-1} + C\delta^{1+\alpha}$ и начальное условие $J_1 = C\delta^{1+\alpha}$, откуда по свойству убывающей геометрической прогрессии следует, что

$$J_r \leq \begin{cases} C \left(\frac{1 - \mathfrak{W}^{T/\delta}(\delta)}{1 - \mathfrak{W}(\delta)} \right) \delta^{1+\alpha}, & \text{если } \mathfrak{W}(\delta) < 1; \\ C \left(\frac{T}{\delta} \right) \delta^{1+\alpha} = CT\delta^\alpha, & \text{если } \mathfrak{W}(\delta) = 1. \end{cases}$$

Неравенства для J_r не зависят от начального распределения π . Утверждение теоремы следует из подстановки неравенств в правую часть (9) и последующего взятия от обеих частей точной верхней грани по $\pi \in \Pi$. Теорема 1 доказана.

Литература

1. Борисов А. \mathcal{L}_1 -оптимальная фильтрация марковских скачкообразных процессов I: точное решение и численные схемы реализации // Автоматика и телемеханика, 2020. Вып. 11. С. 11–31.
2. Борисов А. \mathcal{L}_1 -оптимальная фильтрация марковских скачкообразных процессов II: численный анализ конкретных схем // Автоматика и телемеханика, 2020. Вып. 12. С. 24–49.
3. Борисов А., Казанчян Д. Фильтрация состояний марковских скачкообразных процессов по комплексным наблюдениям I: точное решение задачи // Информатика и её применения, 2021. Т. 15. Вып. 2. С. 11–18.
4. Борисов А., Казанчян Д. Фильтрация состояний марковских скачкообразных процессов по комплексным наблюдениям II: численный алгоритм // Информатика и её применения, 2021. Т. 15. Вып. 3. С. 9–15.
5. Ishikawa Y., Kunita H. Malliavin calculus on the Wiener–Poisson space and its application to canonical SDE with jumps // *Stoch. Proc. Appl.*, 2006. Vol. 116. P. 1743–1769.
6. Liptser R., Shiryaev A. *Statistics of random processes I: General theory*. — Berlin/Heidelberg: Springer, 2001. 427 p.
7. Kloeden P., Platen E. *Numerical solution of stochastic differential equations*. — Berlin: Springer, 1992. 636 p.

Поступила в редакцию 27.12.21

TOTAL APPROXIMATION ORDER FOR MARKOV JUMP PROCESS FILTERING GIVEN DISCRETIZED OBSERVATIONS

A. V. Borisov

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

Abstract: The note proceeds the investigation devoted to the numerical approximation of the Markov jump process filtering given both the counting and diffusion observations with the multiplicative noise. The filtering estimates are approximated using the observations, previously discretized by time. By contrast with the previous algorithms which limit the number of the Markov state transitions that occurred during the time discretization interval, the new estimates are free of these restrictions and constructed via a unified scheme. The note presents an upper bound for the approximation accuracy as a function of the observation system parameters, applied scheme of the numerical integration, the time discretization step, and the filtering moment. A numerical example illustrates a sublinear character of the bound towards the latter argument.

Keywords: Markov jump process; optimal filtering; diffusion and counting observations; multiplicative observation noise; numerical approximation accuracy

DOI: 10.14357/19922264220402

Acknowledgments

The research was carried out using the infrastructure of shared research facilities СКР “Informatics” of FRC CSC RAS.

References

1. Borisov, A. 2020. \mathcal{L}_1 -optimal filtering of Markov jump processes. I. Exact solution and numerical implementation schemes. *Automat. Rem. Contr.* 81(11):1945–1962.
2. Borisov, A. 2020. \mathcal{L}_1 -optimal filtering of Markov jump processes. II. Numerical analysis of particular realizations schemes. *Automat. Rem. Contr.* 81(12):2160–2180.
3. Borisov, A., and D. Kazanchyan. 2021. Fil'tratsiya sostoyaniy markovskikh skachkoobraznykh protsessov po kompleksnym nablyudeniym I: Tochnoe reshenie zadachi [Filtering of Markov jump processes given composite observation I: Exact solution]. *Informatika i ee Primeneniya — Inform. Appl.* 15(2):11–18.
4. Borisov, A., and D. Kazanchyan. 2021. Fil'tratsiya sostoyaniy markovskikh skachkoobraznykh protsessov po kompleksnym nablyudeniym II: Chislennyy algoritm [Filtering of Markov jump processes given composite observations II: Numerical algorithm]. *Informatika i ee Primeneniya — Inform. Appl.* 15(3):9–15.
5. Ishikawa, Y., and H. Kunita. 2006. Malliavin calculus on the Wiener–Poisson space and its application to canonical SDE with jumps. *Stoch. Proc. Appl.* 116:1743–1769.
6. Liptser, R., and A. Shiryaev. 2001. *Statistics of random processes I: General theory*. Berlin/Heidelberg: Springer. 427 p.
7. Kloeden, P., and E. Platen. 1992. *Numerical solution of stochastic differential equations*. Berlin: Springer. 636 p.

Received December 27, 2021

Contributor

Borisov Andrey V. (b. 1965) — Doctor of Science in physics and mathematics, principal scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; aborisov@frcsc.ru

НЕСМЕЩЕННАЯ ОЦЕНКА РИСКА ПОРОГОВОЙ ОБРАБОТКИ С ДВУМЯ ПОРОГОВЫМИ ЗНАЧЕНИЯМИ

О. В. Шестаков¹

Аннотация: Задачи уменьшения уровня шума в сигналах возникают во многих прикладных областях. В случаях когда сигналы не являются стационарными, хорошо зарекомендовали себя методы подавления шума, основанные на вейвлет-преобразовании и процедурах пороговой обработки. Эти методы вычислительно эффективны и хорошо адаптируются к локальным особенностям сигналов. Наиболее распространенными видами пороговой обработки стали жесткая и мягкая пороговая обработка. Однако при использовании жесткой пороговой обработки получаются оценки с большой дисперсией, а мягкая пороговая обработка приводит к появлению дополнительного смещения. В попытке избавиться от этих недостатков в последние годы были предложены различные альтернативные виды пороговой обработки. В данной работе рассматривается процедура пороговой обработки с двумя порогами, которая ведет себя как мягкая пороговая обработка при малых значениях вейвлет-коэффициентов и как жесткая при больших. Для данного вида пороговой обработки строится несмещенная оценка среднеквадратичного риска и анализируются ее статистические свойства. Также описывается алгоритм вычисления пороговых значений, минимизирующих эту оценку.

Ключевые слова: вейвлеты; пороговая обработка; несмещенная оценка риска

DOI: 10.14357/19922264220403

1 Введение

В процессе генерации и передачи сигналы часто загрязняются аддитивным шумом, который в большинстве случаев предполагается гауссовым. Одной из основных задач ставится уменьшение уровня шума с сохранением характерных особенностей сигнала. Кратномасштабный анализ, осуществляемый с помощью вейвлет-разложений, доказал свою эффективность при решении подобных задач.

При использовании вейвлет-разложения шум равномерно распределяется по всем коэффициентам, в то время как основная часть полезного сигнала сосредоточивается в нескольких самых больших коэффициентах, т. е. вейвлет-разложение обеспечивает разреженное (экономное) представление сигнала.

Самым распространенным методом удаления шума из сигнала в пространстве вейвлет-коэффициентов стала пороговая обработка, в результате которой коэффициенты, величина которых меньше определенного порога, обнуляются. Наиболее популярны методы жесткой и мягкой пороговой обработки. Асимптотически оба этих метода оптимальны в смысле среднеквадратичного риска с точностью до множителя, пропорционального логарифму числа отсчетов сигнала [1].

Как жесткая, так и мягкая пороговая обработка имеют свои достоинства и недостатки. Оценки, полученные с помощью мягкой пороговой обработки, имеют большее смещение, поскольку обработке подвергаются все коэффициенты, в том числе и большие. При жесткой пороговой обработке используется разрывная пороговая функция, что приводит к увеличению дисперсии оценок и отсутствию устойчивости, т. е. чувствительности к незначительным изменениям в данных. Для устранения этих недостатков предложено множество альтернативных видов пороговой функции, которые представляют собой компромисс между жесткой и мягкой пороговой обработкой [2–11]. В частности, в работе [2] предлагается использовать пороговую функцию с двумя пороговыми значениями, которая для малых значений коэффициентов ведет себя как мягкая пороговая обработка, а для больших — как жесткая.

В данной работе с помощью метода Стейна [12] строится несмещенная оценка среднеквадратичного риска при такой пороговой обработке. Показывается, что при определенных условиях данная оценка является сильно состоятельной и асимптотически нормальной. Также обсуждается метод вычисления адаптивного порога, минимизирующего оценку среднеквадратичного риска.

¹Московский государственный университет имени М. В. Ломоносова, факультет вычислительной математики и кибернетики; Федеральный исследовательский центр «Информатика и управление» Российской академии наук; Московский центр фундаментальной и прикладной математики, oshestakov@cs.msu.ru

2 Удаление шума с помощью пороговой обработки эмпирических вейвлет-коэффициентов

Предположим, что наблюдается загрязненный шумом дискретный сигнал

$$X_i = f_i + z_i, \quad i = 1, \dots, N,$$

где f_i — чистые значения функции сигнала f ; z_i — независимые случайные величины, имеющие нормальное распределение с нулевым средним и дисперсией σ^2 . Требуется оценить вектор «полезного» сигнала f_1, \dots, f_N .

Донохо и Джонстон [1] разработали методику WaveShrink для оценивания данного вектора с помощью дискретного вейвлет-преобразования и пороговой обработки. Данная методика в определенном смысле «почти» оптимальная. В частности, для широкого класса функций сигналов она с точностью до множителя $\log N$ обеспечивает минимальный порядок среднеквадратичного риска построенной оценки. При этом вейвлет-функция, используемая для дискретного вейвлет-преобразования, должна удовлетворять определенным условиям (подробнее см. [13]). В дальнейшем будем полагать, что эти условия выполнены.

Методика WaveShrink основана на принципе обнуления малых по значению вейвлет-коэффициентов для удаления шума. Ее суть заключается в следующем: к вектору X_1, \dots, X_N применяется дискретное вейвлет-преобразование с помощью умножения на ортогональную матрицу W , определяемую типом выбранных вейвлетов. В результате получают эмпирические вейвлет-коэффициенты

$$Y_{j,k} = \mu_{j,k} + \epsilon_{j,k}, \quad j = 0, \dots, J-1, \quad k = 0, \dots, 2^j - 1,$$

где $\epsilon_{j,k}$ также независимы и нормально распределены с нулевым средним и дисперсией σ^2 . Затем к получившимся эмпирическим вейвлет-коэффициентам применяется пороговая функция, смысл которой заключается в обнулении достаточно малых коэффициентов, которые считаются шумом. После пороговой обработки осуществляется обратное вейвлет-преобразование, которое в силу ортогональности матрицы W заключается просто в умножении вектора обработанных эмпирических вейвлет-коэффициентов на транспонированную матрицу W^* .

Донохо и Джонстон предложили два вида функции пороговой функции. Это функция жесткой пороговой обработки

$$\rho_H(y, T) = \begin{cases} y & \text{при } |y| > T; \\ 0 & \text{при } |y| \leq T \end{cases}$$

и функция мягкой пороговой обработки

$$\rho_S(y, T) = \begin{cases} y - T & \text{при } y > T; \\ y + T & \text{при } y < -T; \\ 0 & \text{при } |y| \leq T. \end{cases}$$

Здесь T — некоторое пороговое значение. Жесткая пороговая обработка приводит к оценкам, имеющим большую дисперсию (из-за разрывности пороговой функции), а мягкая пороговая обработка вызывает дополнительное смещение в оценках (поскольку обработке подвергаются все коэффициенты) [14]. Один из способов справиться с недостатками жесткой и мягкой пороговой обработки — использовать предложенную в работе [2] пороговую функцию с двумя пороговыми значениями T_1 и T_2 :

$$\rho_F(y, T_1, T_2) = \begin{cases} y & \text{при } |y| \geq T_2; \\ \text{sgn}(y) \frac{T_2(|y| - T_1)}{T_2 - T_1} & \text{при } T_1 < |y| < T_2; \\ 0 & \text{при } |y| \leq T_1. \end{cases} \quad (1)$$

Для абсолютных значений y вблизи нижнего порога $\rho_F(y, T_1, T_2)$ ведет себя как $\rho_S(y, T_1)$. Для абсолютных значений y больше верхнего порога $\rho_F(y, T_1, T_2) = y$. Отметим, что жесткая пороговая обработка при $T_1 = T_2$ и мягкая пороговая обработка при $T_2 = \infty$ являются предельными случаями (1).

3 Статистические свойства оценки среднеквадратичного риска

Целью пороговой обработки, как правило, ставится уменьшение среднеквадратичного риска, который для пороговой функции ρ_F определяется как

$$R_J(T_1, T_2) = \sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} E \left(\hat{Y}_{j,k} - \mu_{j,k} \right)^2, \quad (2)$$

где $\hat{Y}_{j,k} = \rho_F(Y_{j,k}, T_1, T_2)$, и методы выбора пороговых значений ориентированы на минимизацию (2). В работе [2] рассматривается минимаксный подход. Обозначим через μ вектор вейвлет-коэффициентов $\{\mu_{j,k}\}$. Вводится в рассмотрение величина

$$\Lambda_J = \inf_{T_1 \leq T_2} \sup_{\mu} \left\{ \frac{R_J(T_1, T_2)}{\sigma^2 + \sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} \min(\mu_{j,k}^2, \sigma^2)} \right\}.$$

Величина $\min(\mu_{j,k}^2, \sigma^2)$ служит идеальным средне-квадратичным риском обработки одного коэффициента, при которой коэффициент обнуляется или остается неизменным [1]. Таким образом, при минимаксных значениях T_1 и T_2

$$R_J(T_1, T_2) \leq \Lambda_J \left(\sigma^2 + \sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} \min(\mu_{j,k}^2, \sigma^2) \right).$$

Для жесткой и мягкой пороговой обработки $\Lambda_J \sim 2 \log 2^J$, а минимаксный порог имеет значение $T \sim \sigma \sqrt{2 \log 2^J}$ [15]. Для пороговой функции ρ_F в [2] численно исследуются минимаксные значения T_1 и T_2 и показывается, что удается построить более точные границы для минимаксного риска, чем в случае жесткой или мягкой пороговой обработки.

Заметим, что в выражении (2) присутствуют неизвестные величины $\mu_{j,k}$, поэтому вычислить значение $R_J(T_1, T_2)$ для конкретного вектора μ на практике нельзя. Однако с помощью метода Стейна для $R_J(T_1, T_2)$ можно вычислить несмещенную оценку. Пусть $g(Y_{j,k}) = Y_{j,k} - \hat{Y}_{j,k}$. Тогда [12]

$$E \left(\hat{Y}_{j,k} - \mu_{j,k} \right)^2 = \sigma^2 + E g^2(Y_{j,k}) - 2\sigma^2 E g'(Y_{j,k}),$$

и, следовательно, величина $\sigma^2 + g^2(Y_{j,k}) - 2\sigma^2 g'(Y_{j,k})$ является несмещенной оценкой для $E \left(\hat{Y}_{j,k} - \mu_{j,k} \right)^2$. Из вида функции ρ_F следует, что

$$\begin{aligned} g(Y_{j,k}) &= Y_{j,k} \mathbf{1}(|Y_{j,k}| \leq T_1) + \\ &+ \frac{T_1(T_2 \operatorname{sgn}(Y_{j,k}) - Y_{j,k})}{T_2 - T_1} \mathbf{1}(T_1 < |Y_{j,k}| < T_2); \\ g'(Y_{j,k}) &= \mathbf{1}(|Y_{j,k}| \leq T_1) - \\ &- \frac{T_1}{T_2 - T_1} \mathbf{1}(T_1 < |Y_{j,k}| < T_2). \end{aligned}$$

Таким образом,

$$\hat{R}_J(T_1, T_2) = \sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} H(Y_{j,k}, T_1, T_2), \quad (3)$$

где

$$H(Y_{j,k}, T_1, T_2) = \begin{cases} \sigma^2 & \text{при } |Y_{j,k}| \geq T_2; \\ \sigma^2 + \frac{T_1^2(T_2 \operatorname{sgn}(Y_{j,k}) - Y_{j,k})^2}{(T_2 - T_1)^2} + \frac{2\sigma^2 T_1}{T_2 - T_1} & \text{при } T_1 < |Y_{j,k}| < T_2; \\ Y_{j,k}^2 - \sigma^2 & \text{при } |Y_{j,k}| \leq T_1 \end{cases}$$

является несмещенной оценкой для $R_J(T_1, T_2)$.

Исследования пороговых значений показывают, что пороги, обеспечивающие величину средне-квадратичного риска, близкую к минимальному, разумно выбирать растущими со скоростью порядка $\sqrt{2 \log 2^J}$ [16, 17]. В этом случае, повторяя рассуждения из работы [18], можно показать, что при условии принадлежности функции сигнала определенному классу Липшица оценка $\hat{R}_J(T_1, T_2)$ является сильно состоятельной и асимптотически нормальной.

Теорема 1. Пусть f задана на отрезке $[a, b]$ и равномерно регулярна по Липшицу с показателем $\gamma > 1/2$, $T_1 = c_1 \sigma \sqrt{2 \log 2^J}$ и $T_2 = c_2 \sigma \sqrt{2 \log 2^J}$ ($c_1 \leq c_2$). Тогда

$$P \left(\frac{\hat{R}_J(T_1, T_2) - R_J(T_1, T_2)}{\sigma^2 \sqrt{2^{J+1}}} < x \right) \rightarrow \Phi(x) \quad \text{при } J \rightarrow \infty,$$

где $\Phi(x)$ — функция распределения стандартного нормального закона.

Теорема 2. Пусть f задана на отрезке $[a, b]$, $f \in L^2([a, b])$, $T_1 = c_1 \sigma \sqrt{2 \log 2^J}$ и $T_2 = c_2 \sigma \sqrt{2 \log 2^J}$ ($c_1 \leq c_2$). Тогда при любом $\lambda > 0$

$$\frac{\hat{R}_J(T_1, T_2) - R_J(T_1, T_2)}{2^{J(1/2+\lambda)}} \rightarrow 0 \text{ п.в. при } J \rightarrow \infty.$$

Эти утверждения служат основанием использования величины (3) для оценивания качества при практической обработке сигналов, поскольку для ее вычисления используются только наблюдаемые данные.

4 Минимизация оценки риска

Еще одним подходом к выбору пороговых значений T_1 и T_2 может служить минимизация оценки (3), т. е. в качестве порогов выбираются такие T_1^* и T_2^* , что

$$\hat{R}_J(T_1^*, T_2^*) = \min_{T_1 > 0, T_2 > 0} \hat{R}_J(T_1, T_2).$$

Для поиска этих значений можно использовать следующий алгоритм.

Упорядочим эмпирические коэффициенты $Y_{j,k}$ по убыванию абсолютных значений и обозначим через $Y_{(i)}$ i -ю координату полученного упорядоченного вектора (без ограничения общности можно полагать, что среди этих значений нет совпадающих). Пусть l и m ($l \leq m$) — такие индексы, что

$$|Y_{(m)}| \leq T_1 < |Y_{(m-1)}|; \quad |Y_{(l)}| < T_2 \leq |Y_{(l-1)}|. \quad (4)$$

Если $l = m$, то функция $\rho_F(y, T_1, T_2)$ фактически сводится к функции жесткой пороговой обработки

с порогом T_2 , поэтому сначала рассмотрим случай $l < m$. Имеем

$$\begin{aligned} \widehat{R}_J(T_1, T_2) &= \sum_{i=1}^{N-1} H(Y_{(i)}, T_1, T_2) = \\ &= \sum_{i=1}^{l-1} H(Y_{(i)}, T_1, T_2) + \sum_{i=l}^{m-1} H(Y_{(i)}, T_1, T_2) + \\ &\quad + \sum_{i=m}^{N-1} H(Y_{(i)}, T_1, T_2) = \sum_{i=1}^{l-1} \sigma^2 + \\ &+ \sum_{i=l}^{m-1} \left(\sigma^2 + \frac{T_1^2 (T_2 \operatorname{sgn}(Y_{(i)}) - Y_{(i)})^2}{(T_2 - T_1)^2} + \frac{2\sigma^2 T_1}{T_2 - T_1} \right) + \\ &\quad + \sum_{i=m}^{N-1} (Y_{(i)}^2 - \sigma^2) = \sum_{i=l}^{m-1} \left(\frac{T_1^2 (T_2 - |Y_{(i)}|)^2}{(T_2 - T_1)^2} + \right. \\ &\quad \left. + \frac{2\sigma^2 T_1}{T_2 - T_1} \right) + \sum_{i=m}^{N-1} Y_{(i)}^2 + (2m - 1 - N)\sigma^2. \quad (5) \end{aligned}$$

При ограничениях (4) для любого значения T_2 слагаемые в первой сумме могут достигать минимального значения только при $T_1 = |Y_{(m)}|$. Далее, для всех слагаемых в первой сумме выполнено $T_1 < |Y_{(i)}| < T_2$. Следовательно, выражение

$$\frac{(T_2 - |Y_{(i)}|)^2}{(T_2 - T_1)^2} = 1 + \frac{2(|Y_{(i)}| - T_1)}{(T_2 - T_1)} + \frac{(T_1 - |Y_{(i)}|)^2}{(T_2 - T_1)^2}$$

убывает по T_2 , т.е. минимальное значение слагаемых в первой сумме достигается при $T_2 = |Y_{(l-1)}|$. Если $l = m$, то выражение (5) не содержит первой суммы и при ограничениях (4) не зависит от T_1 и T_2 . Следовательно, можно также положить $T_1 = |Y_{(m)}|$ и $T_2 = |Y_{(l-1)}|$. Таким образом, пороги T_1^* и T_2^* следует искать, сравнивая значения (5) в точках $T_1 = |Y_{(s)}|$ и $T_2 = |Y_{(t)}|$ при $1 \leq t < s \leq N - 1$. Те значения, на которых достигается минимум, принимаются за T_1^* и T_2^* .

Описанный алгоритм обобщает алгоритм поиска порога, минимизирующего оценку риска в случае мягкой пороговой обработки [13].

Литература

1. Donoho D., Johnstone I. M. Ideal spatial adaptation via wavelet shrinkage // *Biometrika*, 1994. Vol. 81. No. 3. P. 425–455.
2. Bruce A. G., Gao H.-Y. WaveShrink with firm shrinkage // *Stat. Sinica*, 1997. Vol. 7. P. 855–874.
3. Gao H.-Y. Wavelet shrinkage denoising using the non-negative garrote // *J. Comput. Graph. Stat.*, 1998. Vol. 7. No. 4. P. 469–488.

4. Chmelka L., Kozumplik J. Wavelet-based wiener filter for electrocardiogram signal denoising // *Comput. Cardiol.*, 2005. Vol. 32. P. 771–774.
5. Poornachandra S., Kumaravel N., Saravanan T. K., Somaskandan R. WaveShrink using modified hyper-shrinkage function // 27th Annual Conference (International) of the IEEE Engineering in Medicine and Biology Society Proceedings. — Piscataway, NJ, USA: IEEE, 2005. P. 30–32.
6. Lin Y., Cai J. A new threshold function for signal denoising based on wavelet transform // *Conference (International) on Measuring Technology and Mechatronics Automation Proceedings*. — Piscataway, NJ, USA: IEEE, 2010. P. 200–203.
7. Huang H.-C., Lee T. C. M. Stabilized thresholding with generalized sure for image denoising // 17th Conference (International) on Image Processing Proceedings. — Piscataway, NJ, USA: IEEE, 2010. P. 1881–1884.
8. Zhao R.-M., Cui H.-M. Improved threshold denoising method based on wavelet transform // 7th Conference (International) on Modelling, Identification and Control Proceedings. — Piscataway, NJ, USA: IEEE, 2015. Art. 7409352. 4 p.
9. He C., Xing J., Li J., Yang Q., Wang R. A new wavelet thresholding function based on hyperbolic tangent function // *Math. Probl. Eng.*, 2015. Vol. 2015. Art. 528656. 10 p.
10. Priya K. D., Rao G. S., Rao P. S. Comparative analysis of wavelet thresholding techniques with wavelet-wiener filter on ECG signal // *Procedia Comput. Sci.*, 2016. Vol. 87. P. 178–183.
11. He H., Tan Y. A novel adaptive wavelet thresholding with identical correlation shrinkage function for ECG noise removal // *Chinese J. Electron.*, 2018. Vol. 27. No. 3. P. 507–513.
12. Stein C. Estimation of the mean of a multivariate normal distribution // *Ann. Stat.*, 1981. Vol. 9. No. 6. P. 1135–1151.
13. Mallat S. A Wavelet tour of signal processing. — New York, NY, USA: Academic Press, 1999. 857 p.
14. Bruce A. G., Gao H.-Y. Understanding WaveShrink: Variance and bias estimation // *Biometrika*, 1996. Vol. 83. P. 727–745.
15. Donoho D., Johnstone I. M. Minimax estimation via wavelet shrinkage // *Ann. Stat.*, 1998. Vol. 26. No. 3. P. 879–921.
16. Marron J. S., Adak S., Johnstone I. M., Neumann M. H., Patil P. Exact risk analysis of wavelet regression // *J. Comput. Graph. Stat.*, 1998. Vol. 7. P. 278–309.
17. Jansen M. Noise reduction by wavelet thresholding. — Lecture notes in statistics ser. — Springer, 2001. Vol. 161. 217 p.
18. Попенова П. С., Шестаков О. В. Анализ статистических свойств метода гибридной пороговой обработки // *Вестник ТвГУ. Сер.: Прикладная математика*, 2019. № 1. С. 15–22.

Поступила в редакцию 05.08.22

UNBIASED THRESHOLDING RISK ESTIMATE WITH TWO THRESHOLD VALUES

O. V. Shestakov^{1,2,3}

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

³Moscow Center for Fundamental and Applied Mathematics, M. V. Lomonosov Moscow State University, 1 Leninskie Gory, GSP-1, Moscow 119991, Russian Federation

Abstract: The problems of noise reduction in signals arise in many application areas. In cases where the signals are not stationary, noise suppression methods based on wavelet transform and thresholding procedures have proven themselves well. These methods are computationally efficient and adapt well to the local features of the signals. The most common types of thresholding are hard and soft thresholding. However, when using hard thresholding, estimates with large variance are obtained, and soft thresholding leads to additional bias. In an attempt to get rid of these shortcomings, various alternative types of thresholding have been proposed in recent years. This paper considers a thresholding procedure with two thresholds which behaves as soft thresholding for small values of wavelet coefficients and as hard thresholding for the large ones. For this type of thresholding, an unbiased estimate of the mean-square risk is constructed and its statistical properties are analyzed. An algorithm for calculating the threshold values that minimizes this estimate is described.

Keywords: wavelets; thresholding; unbiased risk estimate

DOI: 10.14357/19922264220403

References

1. Donoho, D., and I. M. Johnstone. 1994. Ideal spatial adaptation via wavelet shrinkage. *Biometrika* 81(3):425–455.
2. Bruce, A. G., and H.-Y. Gao. 1997. WaveShrink with firm shrinkage. *Stat. Sinica* 7:855–874.
3. Gao, H.-Y. 1998. Wavelet shrinkage denoising using the non-negative garrote. *J. Comput. Graph. Stat.* 7(4):469–488.
4. Chmelka, L., and J. Kozumplik. 2005. Wavelet-based wiener filter for electrocardiogram signal denoising. *Comput. Cardiol.* 32:771–774.
5. Poornachandra, S., N. Kumaravel, T. K. Saravanan, and R. Somaskandan. 2005. WaveShrink using modified hyper-shrinkage function. *27th Annual Conference (International) of the IEEE Engineering in Medicine and Biology Society Proceedings*. Piscataway, NJ: IEEE. 30–32.
6. Lin, Y., and J. Cai. 2010. A new threshold function for signal denoising based on wavelet transform. *Conference (International) on Measuring Technology and Mechatronics Automation Proceedings*. Piscataway, NJ: IEEE. 200–203.
7. Huang, H.-C., and T. C. M. Lee. 2010. Stabilized thresholding with generalized sure for image denoising. *17th Conference (International) on Image Processing Proceedings*. Piscataway, NJ: IEEE. 1881–1884.
8. Zhao, R.-M., and H.-M. Cui. 2015. Improved threshold denoising method based on wavelet transform. *7th Conference (International) on Modelling, Identification and Control Proceedings*. Piscataway, NJ: IEEE. Art. 7409352. 4 p.
9. He, C., J. Xing, J. Li, Q. Yang, and R. Wang. 2015. A new wavelet thresholding function based on hyperbolic tangent function. *Math. Probl. Eng.* 2015:528656. 10 p.
10. Priya, K. D., G. S. Rao, and P. S. Rao. 2016. Comparative analysis of wavelet thresholding techniques with wavelet-wiener filter on ECG signal. *Procedia Comput. Sci.* 87:178–183.
11. He, H., and Y. Tan. 2018. A novel adaptive wavelet thresholding with identical correlation shrinkage function for ECG noise removal. *Chinese J. Electron.* 27(3):507–513.
12. Stein, C. 1981. Estimation of the mean of a multivariate normal distribution. *Ann. Stat.* 9(6):1135–1151.
13. Mallat, S. 1999. *A wavelet tour of signal processing*. New York, NY: Academic Press. 857 p.
14. Bruce, A. G., and Gao H.-Y. 1996. Understanding WaveShrink: Variance and bias estimation. *Biometrika* 83:727–745.
15. Donoho, D., and I. M. Johnstone. 1998. Minimax estimation via wavelet shrinkage. *Ann. Stat.* 26(3):879–921.
16. Marron, J. S., S. Adak, I. M. Johnstone, M. H. Neumann, and P. Patil. 1998. Exact risk analysis of wavelet regression. *J. Comput. Graph. Stat.* 7:278–309.
17. Jansen, M. 2001. *Noise reduction by wavelet thresholding*. Lecture notes in statistics ser. New York, NY: Springer Verlag. Vol. 161. 217 p.
18. Popenova, P. S., and O. V. Shestakov. 2019. Analiz statisticheskikh svoystv metoda gibridnoy porogovoy obrabotki

[Analysis of statistical properties of the hybrid thresholding technique]. *Vestnik Tverskogo gosudarstvennogo un-ta.*

Ser. Prikladnaya matematika [Herald of Tver State University. Ser. Applied Mathematics] 1:15–22.

Received August 5, 2022

Contributor

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

ОБОБЩЕНИЕ МЕТОДА ВЫПРЯМЛЕНИЯ ИСКАЖЕННЫХ ИЗ-ЗА МУЛЬТИКОЛЛИНЕАРНОСТИ КОЭФФИЦИЕНТОВ ДЛЯ РЕГРЕССИОННЫХ МОДЕЛЕЙ С РАЗЛИЧНОЙ СТЕПЕНЬЮ КОРРЕЛЯЦИИ ОБЪЯСНЯЮЩИХ ПЕРЕМЕННЫХ

М. П. Базилевский¹

Аннотация: При построении регрессионных моделей одну из главных проблем представляет мультиколлинеарность. Это негативное явление приводит к искажению коэффициентов регрессии, в частности их знаков. Ранее для решения проблемы мультиколлинеарности был разработан метод выпрямления искаженных коэффициентов (МВИК), который основан на построении модели полностью линейной регрессии (МПЛР). Одно из условий его применимости — тесная корреляция абсолютно всех пар объясняющих переменных. Но при решении реальных прикладных задач это условие выполняется редко. Чаще всего объясняющие переменные коррелируют друг с другом по-разному. Для этого в данной статье предложен новый итерационный алгоритм МВИК. Особенность алгоритма заключается в том, что он сочетает в себе достоинства как традиционных множественных моделей, так и новых полностью связанных регрессий. Разработанный алгоритм универсален и может применяться для построения регрессионного уравнения с любой структурой корреляционной матрицы. Новый алгоритм успешно применен для моделирования грузоперевозок железнодорожного транспорта в Иркутской области.

Ключевые слова: регрессионный анализ; корреляция; мультиколлинеарность; метод выпрямления искаженных коэффициентов; модель полностью связанной линейной регрессии

DOI: 10.14357/19922264220404

1 Введение

Одной из главных проблем регрессионного анализа [1] является мультиколлинеарность — наличие сильной линейной зависимости между двумя или более объясняющими переменными. Мультиколлинеарность приводит к искажению оценок регрессии, в частности их знаков, что затрудняет или вовсе делает невозможным интерпретацию влияния отдельных факторов на выходную переменную. На сегодняшний день проблема мультиколлинеарности не до конца решена. Для этого применяются различные приемы: исключение сильно коррелирующих объясняющих переменных, метод главных компонент [2, 3], гребневая (ридж) регрессия [4, 5], добавление новых наблюдений, преобразование исходных переменных. Современным направлением стала разработка новых методов оценивания регрессионных моделей в условиях мультиколлинеарности (см., например, [6, 7]).

Все перечисленные приемы имеют свои недостатки. Так, метод исключения приводит к потере информативности и возможности изучения совместного влияния всех переменных на выходной показатель. Преобразовать исходные переменные в набор слабо коррелирующих факторов или найти новые наблюдения удастся крайне редко. При

построении гребневой регрессии возникает проблема с выбором параметра регуляризации. Кроме того, оценивание гребневой регрессии заключается в искажении формулы для метода наименьших квадратов (МНК), поэтому для полученного уравнения будут несправедливы коэффициент детерминации, критерий Фишера и t -критерий Стьюдента. То же самое относится и к новым методам оценивания в условиях мультиколлинеарности. При реализации метода главных компонент возникает проблема с интерпретацией главных компонент. Самый же главный недостаток всех этих методов кроется в том, что ни один из них не гарантирует согласованность знаков коэффициентов регрессии при объясняющих переменных с соответствующими знаками их коэффициентов корреляции с выходной переменной.

В работе [8] автором были предложены МПЛР, которые строятся при условии сильной корреляции между объясняющими переменными. В результате оценивания МПЛР формируются линейно зависимые латентные переменные в виде линейных комбинаций объясняющих переменных. Было установлено, что знаки коэффициентов в этих линейных комбинациях согласуются с соответствующими знаками коэффициентов линейной корреляции.

¹Иркутский государственный университет путей сообщения, кафедра математики, mik2178@yandex.ru

ляии. На этой основе в работе [9] был разработан МВИК, который полностью решает проблему мультиколлинеарности. Однако МВИК работает только тогда, когда абсолютно все объясняющие переменные тесно коррелируют между собой. Если же между всеми ними слабая корреляция, то справедливым будет построение традиционной модели множественной линейной регрессии (ММЛР). Но не до конца понятно, как поступать в ситуации, когда объясняющие переменные коррелируют по-разному: одни — сильно, вторые — слабо, третьи — средне. В этом случае уже нельзя применять ни ММЛР, ни МПЛР. Цель данной работы состоит в обобщении МВИК для регрессионных моделей с различной степенью корреляции объясняющих переменных.

2 Модели множественной и полносвязной линейной регрессии

Рассмотрим ММЛР

$$y_i = \alpha_0 + \alpha_1 x_{i1} + \alpha_2 x_{i2} + \dots + \alpha_m x_{im} + \varepsilon_i, \quad (1)$$

где $y_i, i = \overline{1, n}$, — значения зависимой (объясняемой) переменной y ; $x_{i1}, x_{i2}, \dots, x_{im}, i = \overline{1, n}$, — значения m независимых (объясняющих) переменных x_1, x_2, \dots, x_m (x -переменные); $\varepsilon_i, i = \overline{1, n}$, — ошибки аппроксимации; $\alpha_0, \alpha_1, \dots, \alpha_m$ — неизвестные параметры; n — объем выборки.

При полной коллинеарности между объясняющими переменными определитель $|X^T X|$, где X — матрица наблюдений факторов, равен 0, поэтому МНК-оценки получить невозможно. При мультиколлинеарности, т. е. при $|X^T X| \approx 0$, МНК-оценки становятся неустойчивыми. Таким образом, из-за мультиколлинеарности ММЛР целесообразно строить только тогда, когда абсолютно все объясняющие переменные слабо коррелируют между собой.

Модель полносвязной линейной регрессии [8, 9] имеет следующий вид:

$$\left. \begin{aligned} x_{ij} &= x_{ij}^* + \varepsilon_i^{(x_j)}, \quad i = \overline{1, n}, \quad j = \overline{1, m}; \\ x_j^* &= a_j + b_j x_m^*, \quad j = \overline{1, m-1}, \end{aligned} \right\} \quad (2)$$

где $x_{i1}^*, x_{i2}^*, \dots, x_{im}^*, i = \overline{1, n}$, — истинные значения объясняющих переменных; $a_j, b_j, j = \overline{1, m-1}$, — неизвестные параметры; $\varepsilon_i^{(x_j)}, i = \overline{1, n}$, — ошибки j -й переменной.

Модель полносвязной линейной регрессии (2) целесообразно строить только тогда, когда абсо-

лютно все объясняющие переменные сильно коррелируют между собой, т. е. при сильной мультиколлинеарности.

Модель полносвязной линейной регрессии оценивается с помощью взвешенного метода наименьших полных квадратов (ВМНПК):

$$\begin{aligned} &\lambda_1 \sum_{i=1}^n (x_{i1} - a_1 - b_1 x_{im}^*)^2 + \\ &+ \lambda_2 \sum_{i=1}^n (x_{i2} - a_2 - b_2 x_{im}^*)^2 + \dots \\ &\dots + \lambda_{m-1} \sum_{i=1}^n (x_{i,m-1} - a_{m-1} - b_{m-1} x_{im}^*)^2 + \\ &+ \sum_{i=1}^n (x_{im} - x_{im}^*)^2 \rightarrow \min, \quad (3) \end{aligned}$$

где $\lambda_1, \lambda_2, \dots, \lambda_{m-1}$ — положительные весовые коэффициенты (лямбда-параметры).

Для заданных значений лямбда-параметров задача (3) решается по следующей схеме.

1. Находятся оценки $\tilde{b}_1, \tilde{b}_2, \dots, \tilde{b}_{m-1}$ из нелинейной системы

$$\begin{aligned} &b_p \left(D_{x_m} + \sum_{j=1}^{m-1} \lambda_j^2 b_j^2 D_{x_j} + \right. \\ &+ 2 \sum_{j_1=1}^{m-2} \sum_{j_2=j_1+1}^{m-1} \lambda_{j_1} \lambda_{j_2} b_{j_1} b_{j_2} K_{x_{j_1} x_{j_2}} + \\ &\left. + 2 \sum_{j=1}^{m-1} \lambda_j b_j K_{x_j x_m} \right) = \\ &= \left(1 + \sum_{j=1}^{m-1} \lambda_j b_j^2 \right) \left(\sum_{j=1}^{m-1} \lambda_j b_j K_{x_j x_p} + K_{x_m x_p} \right), \\ & \quad p = \overline{1, m-1}. \end{aligned}$$

В работе [9] предложен численный метод ее решения.

2. Определяются оценки $\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_{m-1}$ по формулам

$$\tilde{a}_j = \overline{x_j} - \tilde{b}_j \overline{x_m}, \quad j = \overline{1, m-1}.$$

3. Вычисляются оценки истинных значений переменной x_m^* по формулам

$$\begin{aligned} \tilde{x}_{im}^* &= \left(1 + \sum_{j=1}^{m-1} \lambda_j \tilde{b}_j^2 \right)^{-1} \left(- \sum_{j=1}^{m-1} \lambda_j \tilde{a}_j \tilde{b}_j + \right. \\ &\left. + \sum_{j=1}^{m-1} \lambda_j \tilde{b}_j x_{ij} + x_{im} \right), \quad i = \overline{1, n}. \quad (4) \end{aligned}$$

В [9] установлено, что при сильной корреляции объясняющих переменных x_1, x_2, \dots, x_m в (4) знаки коэффициентов при объясняющих переменных будут согласованы с соответствующими знаками коэффициентов корреляции $r_{x_j x_m}$, $j = \overline{1, m-1}$. На основе этого факта в [9] был разработан реализующий МВИК алгоритм “Straight B”:

1. При $\lambda_1 = D_{x_m}/D_{x_1}, \lambda_2 = D_{x_m}/D_{x_2}, \dots, \lambda_{m-1} = D_{x_m}/D_{x_{m-1}}$ решается задача (3).
2. С помощью МНК оценивается модель $y_i = c_0 + c_1 \tilde{x}_{im}^* + \varepsilon_i$, $i = \overline{1, n}$, где c_0 и c_1 — неизвестные параметры.
3. Путем подстановки (4) в равенство $\tilde{y} = \tilde{c}_0 + \tilde{c}_1 \tilde{x}_m^*$ определяется искомое уравнение регрессии.

Условия применимости МВИК по алгоритму “Straight B”:

- (1) каждая объясняющая переменная должна тесно коррелировать с выходной переменной y ;
- (2) все объясняющие переменные должны тесно коррелировать друг с другом.

Если второе условие не выполняется, то могут возникнуть проблемы с численным оцениванием МПЛР, а также с согласованностью знаков коэффициентов регрессии с соответствующими знаками коэффициентов корреляции.

Достоинство алгоритма “Straight B” заключается в полном отсутствии эффекта мультиколлинеарности в процессе получения регрессионного уравнения и сохранении в модели всех исходных переменных.

К сожалению, при решении прикладных задач анализа данных крайне редко бывает так, чтобы сразу все объясняющие переменные коррелировали между собой или только слабо, или только сильно. Поэтому возникает проблема с применением как ММЛР (1), так и МПЛР (2). Для ее решения был разработан итерационный алгоритм “Straight C”, основанный на постепенном слиянии сильно коррелирующих переменных.

3 Алгоритм “Straight C”

Предварительно исключаются объясняющие переменные, которые слабо коррелируют с выходной переменной y . Также следует исключить те переменные, знаки коэффициентов корреляции которых с y противоречат содержательному смыслу решаемой задачи.

Задается граница между областями слабой и сильной мультиколлинеарности — $r_{\text{гран}} \in (0, 1)$. Будем считать, что если все абсолютные значения

коэффициентов корреляции объясняющих переменных меньше $r_{\text{гран}}$, то можно применять ММЛР, а если все больше $r_{\text{гран}}$, то — МПЛР.

Шаг 1. Находится корреляционная матрица объясняющих переменных R . Когда число объясняющих переменных

- (1) равно двум, тогда если их коэффициент корреляции $r < r_{\text{гран}}$, то находятся МНК-оценки ММЛР, а если $r \geq r_{\text{гран}}$, то выполняется алгоритм “Straight B”. На этом алгоритм “Straight C” завершается;
- (2) более двух, тогда если все абсолютные значения коэффициентов корреляции матрицы R :
 - попадают в промежуток $[0, r_{\text{гран}})$, то с помощью МНК оценивается ММЛР и алгоритм “Straight C” завершен;
 - попадают в промежуток $[r_{\text{гран}}, 1]$, то действуем по алгоритму “Straight B” и алгоритм “Straight C” завершен;
 - не попадают в эти промежутки, то переходим к шагу 2.

Шаг 2. Генерируются все разбиения множества объясняющих переменных, общее число которых называется числом Белла. Исключаются разбиения, в которых число подмножеств равно числу переменных и 1, поскольку ММЛР и МПЛР уже прошли проверку на шаге 1. Каждое подмножество переменных в разбиении будем называть кластером.

Шаг 3 (цикл по всем разбиениям). Для данного разбиения выбираются все кластеры, состоящие хотя бы из двух переменных. Для каждого такого кластера проверяется условие попадания абсолютных значений всех парных коэффициентов корреляции его переменных в промежуток $[r_{\text{гран}}, 1]$. Если все кластеры удовлетворяют этим условиям, то для каждого из них оценивается МПЛР, по формуле (4) определяются новые z -переменные и все результаты для данного разбиения сохраняются в память № 1.

Шаг 4 (цикл по всем разбиениям из памяти № 1). Для данного разбиения находится корреляционная матрица объясняющих переменных, среди которых могут быть либо z -переменные, либо x - и z -переменные. Проверяется условие попадания абсолютных значений всех парных коэффициентов корреляции в промежуток $[0, r_{\text{гран}})$. Если условие выполняется, то по x - и z -переменным с помощью МНК оценивается ММЛР, проверяется согласованность знаков ее коэффициентов и в случае успеха все результаты для данного разбиения сохраняются в память № 2.

Шаг 5. Если память № 2 пуста, то в памяти № 1 выбираем модель с наибольшим значением коэффициента детерминации R^2 и переходим к шагу 1. В противном случае выбираем в памяти № 2 модель с наибольшим значением коэффициента детерминации R^2 и алгоритм “Straight C” завершается.

Таким образом, алгоритм “Straight C” гарантирует сохранение в окончательной модели всех исходных переменных. Помимо этого итерационно снижается эффект мультиколлинеарности, что приводит к выпрямлению искаженных коэффициентов. К недостатку предложенного алгоритма следует отнести проблему выбора начальной границы $r_{гран}$, от которой зависят итоговые коэффициенты регрессионного уравнения, а значит, и согласованность их знаков со знаками коэффициентов корреляции $r_{yx_j}, j = \overline{1, m}$.

Из-за слияния x -переменных в z -переменные невозможно проверить значимость первых с помощью t -критерия Стьюдента. Для оценки степени влияния x -переменных на y можно воспользоваться следующей процедурой.

Известно, что коэффициент детерминации R^2 оцененной с помощью МНК ММЛР (1) находится по формуле

$$R^2 = \sum_{j=1}^m \tilde{\alpha}_j \frac{\sigma_{x_j}}{\sigma_y} r_{yx_j}. \quad (5)$$

Формула (5) справедлива также и для регрессионного уравнения, полученного с помощью алгоритма “Straight C”. Если в этом уравнении согласованы знаки коэффициентов $\tilde{\alpha}_j, j = \overline{1, m}$, и знаки коэффициентов корреляции $r_{yx_j}, j = \overline{1, m}$, то можно ввести относительные вклады x -переменных в общую детерминацию:

$$C_j^{отн} = 100 \tilde{\alpha}_j \frac{\sigma_{x_j}}{\sigma_y} \frac{r_{yx_j}}{R^2}, \quad j = \overline{1, m}. \quad (6)$$

Заметим, что относительные вклады (6) вводятся не впервые. Так, в работе [10] эти характеристики названы дельта-коэффициентами.

4 Пример

Для демонстрации МВИК на основе алгоритма “Straight C” решалась задача моделирования грузоперевозок железнодорожным транспортом в Иркутской области. Для этого были использованы статистические данные с сайта Федеральной службы государственной статистики (<https://rosstat.gov.ru>) за период с 2000 по 2018 гг. по следующим переменным: y — отправление грузов железнодорожным транспортом общего пользования (млн т); x_2 — население в трудоспособном возрасте (в % от общей численности населения); x_3 — численность рабочей силы (тыс. чел.); x_5 — численность пенсионеров (тыс. чел.); x_8 — число собственных легковых автомобилей на 1000 чел. населения (шт.); x_{18} — число предприятий и организаций; x_{35} — удельный вес автомобильных дорог с твердым покрытием в общей протяженности автомобильных дорог общего пользования (в %); x_{36} — удельный вес автомобильных дорог с усовершенствованным покрытием в протяженности автомобильных дорог с твердым покрытием общего пользования (в %); x_{58} — индексы тарифов на грузовые перевозки (железнодорожный транспорт) (в %).

Матрица парных коэффициентов корреляции для этих переменных представлена в таблице.

Как видно, в этой матрице некоторые переменные коррелируют слабо, например x_2 и x_{18} , а некоторые сильно, например x_5 и x_{58} . Построенная с помощью МНК по этим данным ММЛР имеет вид:

$$\tilde{y} = -385,919 + 4,856x_2 + 0,0286x_3 + 0,2206x_5 - 0,03487x_8 - 0,0001018x_{18} - 0,1653x_{35} - 0,5952x_{36} + 0,003528x_{58}, \quad (7)$$

а ее коэффициент детерминации $R^2 = 0,890647$.

Из-за мультиколлинеарности в уравнении (7) искажены знаки сразу у пяти переменных: $x_5, x_{18}, x_{35}, x_{36}$ и x_{58} .

Корреляционная матрица

	y	x_2	x_3	x_5	x_8	x_{18}	x_{35}	x_{36}	x_{58}
y	1	0,736	0,389	-0,372	-0,354	0,464	0,454	0,693	-0,322
x_2	0,736	1	0,781	-0,879	-0,810	-0,079	0,881	0,851	-0,847
x_3	0,389	0,781	1	-0,859	-0,742	-0,291	0,746	0,611	-0,853
x_5	-0,372	-0,879	-0,859	1	0,935	0,498	-0,940	-0,667	0,988
x_8	-0,354	-0,810	-0,742	0,935	1	0,496	-0,944	-0,544	0,913
x_{18}	0,464	-0,079	-0,291	0,498	0,496	1	-0,357	0,075	0,548
x_{35}	0,454	0,881	0,746	-0,940	-0,944	-0,357	1	0,712	-0,896
x_{36}	0,693	0,851	0,611	-0,667	-0,544	0,075	0,712	1	-0,620
x_{58}	-0,322	-0,847	-0,853	0,988	0,913	0,548	-0,896	-0,620	1

Полученная с использованием МВИК на основе алгоритма “Straight B” регрессия имеет вид:

$$\tilde{y} = 26,807 + 0,1679x_2 + 0,011795x_3 - 0,015763x_5 - 0,00987x_8 - 3,30878 \cdot 10^{-5}x_{18} + 0,09106x_{35} + 0,3636x_{36} - 0,0000438x_{58}, \quad (8)$$

а ее $R^2 = 0,21219$.

В уравнении (8) искажен знак только у переменной x_{18} . Но при этом R^2 для модели (8) оказался существенно ниже, чем для регрессии (7). Эти неудовлетворительные результаты связаны с тем, что нарушено второе условие применимости алгоритма “Straight B”.

Затем к данным был применен МВИК на основе алгоритма “Straight C”. Граница степени мультиколлинеарности $r_{\text{гран}} = 0,72$. Единственной в памяти № 2 оказалась модель

$$\tilde{y} = -40,2784 - 0,0033877z_1 + 0,00069865x_{18} + 1,6017x_{36}, \quad (9)$$

где

$$z_1 = 10069,3873 - 67,570056x_2 - 4,802657x_3 + 6,359486x_5 + 4,02684x_8 - 36,745825x_{35} + 0,176752x_{58}. \quad (10)$$

Коэффициент детерминации регрессии (9) $R^2 = 0,741431$.

Подставляя (10) в (9), получим согласованное по знакам окончательное уравнение регрессии

$$\tilde{y} = -74,3904 + \underset{(9,497\%)}{0,2289}x_2 + \underset{(4,791\%)}{0,0163}x_3 - \underset{(5,177\%)}{0,0215}x_5 - \underset{(4,71\%)}{0,0136}x_8 + \underset{(42,035\%)}{0,0006986}x_{18} + \underset{(6,103\%)}{0,1244}x_{35} + \underset{(23,282\%)}{1,6017}x_{36} - \underset{(4,403\%)}{0,0005987}x_{58}. \quad (11)$$

В (11) в скобках под коэффициентами приведены относительные вклады переменных в общую

детерминацию, найденные по формулам (6). Как видно, наименьший относительный вклад составляет 4,403% для переменной x_{58} , что говорит о значимости всех входящих в модель переменных.

Литература

1. *Montgomery D. C., Peck E. A., Vining G. G.* Introduction to linear regression analysis. — Hoboken, NJ, USA: John Wiley & Sons, 2012. 672 p.
2. *Jolliffe I. T., Cadima J.* Principal component analysis: A review and recent developments // *Philos. T. Roy. Soc. A*, 2016. Vol. 374. No. 2065. Art. 20150202. 16 p.
3. *Naik G. R.* Advances in principal component analysis: Research and development. — Springer, 2018. 252 p.
4. *Kibria B. M. G., Banik S.* Some ridge regression estimators and their performances // *J. Modern Applied Statistical Methods*, 2016. Vol. 15. No. 1. P. 206–238.
5. *Hoerl R. W.* Ridge regression: A historical context // *Technometrics*, 2020. Vol. 62. P. 420–425.
6. *Giacalone M., Panarello D., Mattera R.* Multicollinearity in regression: An efficiency comparison between Lp-norm and least squares estimators // *Qual. Quant.*, 2018. Vol. 52. No. 4. P. 1831–1859.
7. *Dawoud I., Kibria B. M. G.* A new biased estimator to combat the multicollinearity of the Gaussian linear regression model // *Stats*, 2020. Vol. 3. No. 4. P. 526–541.
8. *Базилевский М. П.* Многофакторные модели полностью линейной регрессии без ограничений на соотношения дисперсий ошибок переменных // *Информатика и её применения*, 2020. Т. 14. Вып. 2. С. 92–97.
9. *Базилевский М. П.* Метод выпрямления искаженных из-за мультиколлинеарности коэффициентов в регрессионных моделях // *Информатика и её применения*, 2021. Т. 15. Вып. 2. С. 60–65.
10. *Орлова И. В.* Подход к решению проблемы мультиколлинеарности при анализе влияния факторов на результирующую переменную в моделях регрессии // *Фундаментальные исследования*, 2018. № 3. С. 58–63.

Поступила в редакцию 31.08.21

GENERALIZATION OF A METHOD FOR STRAIGHTENING COEFFICIENTS DISTORTED DUE TO MULTICOLLINEARITY IN REGRESSION MODELS WITH DIFFERENT DEGREES OF EXPLANATORY VARIABLES CORRELATION

M. P. Bazilevskiy

Irkutsk State Transport University, 15 Chernyshevskogo Str., Irkutsk 664074, Russian Federation

Abstract: When constructing regression models, one of the main problems is multicollinearity. This negative phenomenon leads to distortion of the regression coefficients, in particular, their signs. Earlier, to solve the problem of multicollinearity, a method for straightening distorted coefficients was developed which is based on the construction of a fully connected linear regression model. One of the conditions for its applicability is a close correlation of absolutely all pairs of explanatory variables. But when solving real applied problems, this condition is rarely met. Most often, explanatory variables correlate with each other in different ways. The authors propose a new iterative algorithm for the method of straightening distorted coefficients. A feature of the algorithm is that it combines the advantages of both traditional multiple models and new fully connected regressions. The developed algorithm is universal and can be used to construct a regression equation with any structure of the correlation matrix. The new algorithm has been successfully applied to simulate freight transportation by rail in the Irkutsk region.

Keywords: regression analysis; correlation; multicollinearity; method for straightening distorted coefficients; fully connected linear regression model

DOI: 10.14357/19922264220404

References

1. Montgomery, D. C., E. A. Peck, and G. G. Vining. 2012. *Introduction to linear regression analysis*. Hoboken, NJ: John Wiley & Sons. 672 p.
2. Jolliffe, I. T., and J. Cadima. 2016. Principal component analysis: A review and recent developments. *Philos. T. Roy. Soc. A* 374(2065):20150202. 16 p.
3. Naik, G. R. 2018. *Advances in principal component analysis: Research and development*. Springer. 252 p.
4. Kibria, B. M. G., and S. Banik. 2016. Some ridge regression estimators and their performances. *J. Modern Applied Statistical Methods* 15(1):206–238.
5. Hoerl, R. W. 2020. Ridge regression: A historical context. *Technometrics* 62:420–425.
6. Giacalone, M., D. Panarello, and R. Mattera. 2018. Multicollinearity in regression: An efficiency comparison between Lp-norm and least squares estimators. *Qual. Quant.* 52(4):1831–1859.
7. Dawoud, I., and B. M. G. Kibria. 2020. A new biased estimator to combat the multicollinearity of the Gaussian linear regression model. *Stats* 3(4):526–541.
8. Bazilevskiy, M. P. 2020. Mnogofaktornye modeli polnosvyaznoy lineynoy regressii bez ogranicheniy na sootnosheniya dispersiy oshibok peremennykh [Multifactor fully connected linear regression models without constraints to the ratios of variables errors variances]. *Informatika i ee Primeneniya — Inform. Appl.* 14(2):92–97.
9. Bazilevskiy, M. P. 2021. Metod vypryamleniya iskazhenykh iz-za mul'tikollinearnosti koeffitsientov v regressionnykh modelyakh [Method of straightening distorted due to multicollinearity coefficients in regression models]. *Informatika i ee Primeneniya — Inform. Appl.* 15(2):60–65.
10. Orlova, I. V. 2018. Podkhod k resheniyu problemy mul'tikollinearnosti pri analize vliyaniya faktorov na rezul'tiruyushchuyu peremennuyu v modelyakh regressii [Approach to the solution of the multicollinearity problem at the analysis of the influence of the factors on the resulting variable in models of regression]. *Fundamental'nyye issledovaniya* [Fundamental Research] 3:58–63.

Received August 31, 2021

Contributor

Bazilevskiy Mikhail P. (b. 1987) — Candidate of Science (PhD) in technology, associate professor, Department of Mathematics, Irkutsk State Transport University, 15 Chernyshevskogo Str., Irkutsk 664074, Russian Federation; mik2178@yandex.ru

ГРАНИЦЫ ЭКСТРЕМАЛЬНОГО ИНДЕКСА ВРЕМЕНИ ОЖИДАНИЯ В СИСТЕМЕ $M/G/1$ С РАСПРЕДЕЛЕНИЕМ ВРЕМЕНИ ОБСЛУЖИВАНИЯ В ВИДЕ КОНЕЧНОЙ СМЕСИ*

И. В. Пешкова¹

Аннотация: Доказывается следующая теорема: если исходная стационарная последовательность имеет распределение в виде m -компонентной смеси, компоненты которой стохастически упорядочены, и существуют предельные распределения для максимумов всех компонент, а также упорядочены нормализующие последовательности, то экстремальный индекс исходной последовательности находится в границах экстремальных индексов наименьшей и наибольшей компонент. Этот результат применен для оценки экстремального индекса стационарного времени ожидания в системе обслуживания типа $M/G/1$, в которой время обслуживания задано m -компонентной смесью распределений. Рассмотрен пример системы $M/H_m/1$ с гиперэкспоненциальным временем обслуживания. Методом точного моделирования получены результаты оценки экстремального индекса стационарного времени ожидания в системе $M/H_2/1$.

Ключевые слова: распределение экстремальных значений; экстремальный индекс; система обслуживания; стохастическая упорядоченность

DOI: 10.14357/19922264220405

1 Введение

Изучение экстремальных значений находит широкое применение в различных прикладных областях. Так, превышение высоких пороговых значений может означать отказы технических устройств из-за сбоев оборудования, природные катастрофы, потери данных при передаче и т. д. При этом превышения часто происходят не по одному, а группируются во времени, образуя так называемые кластеры [1]. Такая кластеризация в теории экстремальных значений характеризуется с помощью *экстремального индекса* [2], который определяет предельное распределение экстремальных значений стационарных случайных последовательностей [3–5].

Экстремальный индекс θ отражает кластерную структуру базовой последовательности или ее локальную зависимость. Например, для последовательности независимых одинаково распределенных величин $\theta = 1$. Значение экстремального индекса, равное нулю, $\theta = 0$, подразумевает общую зависимость, которая соответствует «очень широким» кластерам превышений пороговых значений. Одна из наиболее практически важных интерпретаций экстремального индекса состоит в том, что $1/\theta$ приблизительно соответствует среднему разме-

ру кластера. Это позволяет оценить экстремальный индекс непараметрическими методами [2, 6, 7].

Экстремальные значения характеристик производительности систем массового обслуживания изучались в работах [8–12]. В статье [13] получены распределения кластерных и межкластерных размеров стационарных времен ожидания в системе $G/G/1$. Предельное распределение и математическое ожидание времени первого превышения порогового значения выведены в [14]. Выбор порогового значения обсуждается в статье [15].

Главная идея данного исследования состоит в построении границ экстремального индекса стационарного времени ожидания в системе обслуживания типа $M/G/1$, в которой время обслуживания задано m -компонентной смесью распределений со *стохастически упорядоченными* компонентами. В качестве примера рассматривается гиперэкспоненциальное распределение, которое представляет собой конечную смесь показательных распределений. Такое распределение играет важную роль в исследовании современных систем, поскольку может быть использовано в качестве распределения, аппроксимирующего распределения с так называемыми «длинными хвостами», широко применяемыми в коммуникационных системах [16].

* Работа выполнена при финансовой поддержке РФФИ (проект 21-71-10135).

¹ Петрозаводский государственный университет; Институт прикладных математических исследований Карельского научного центра РАН, iamnova@petsu.ru

Данное исследование опирается на результаты, полученные автором в предшествующих работах [17–20]. Например, в статье [18] доказана теорема о сравнении экстремальных индексов двух стационарных последовательностей, а в работе [19] получены результаты стохастической упорядоченности и упорядоченности по интенсивности отказов для m -компонентных смесей распределений.

В разд. 2 настоящей статьи доказана теорема 1, которая утверждает, что если стационарная последовательность задана m -компонентной смесью распределений и ее компоненты стохастически упорядочены, т. е. $\overline{F}_{Y_1}(x) \leq \overline{F}_{Y_2}(x) \leq \dots \leq \overline{F}_{Y_m}(x)$ (где $\overline{F}(x)$ — хвост функции распределения), и соответствующие нормализующие последовательности также упорядочены, то экстремальный индекс θ_X исходной последовательности X_n удовлетворяет неравенству $\theta_{Y_1} \geq \theta_X \geq \theta_{Y_m}$. В разд. 3 сформулирована теорема 2, которая позволяет построить границы экстремального индекса стационарного времени ожидания в односерверной системе обслуживания с m -компонентным распределением времени обслуживания. В разд. 4 приведен пример применения теоремы 2 для системы $M/H_m/1$ с гиперэкспоненциальным временем обслуживания.

2 Границы экстремального индекса строго стационарной последовательности, заданной m -компонентной смесью распределений

Рассмотрим стационарную в узком смысле последовательность случайных величин (СВ) $\{X_n, n \geq 1\}$, удовлетворяющую тождеству

$$(X_{k+1}, \dots, X_{k+n}) \stackrel{d}{=} (X_1, \dots, X_n) \text{ для всех } k, n \geq 1,$$

где $\stackrel{d}{=}$ означает равенство по распределению. Обозначим M_n максимум первых n членов последовательности, т. е.

$$M_n = \max(X_1, \dots, X_n).$$

Введем совместную функцию распределения (ф. р.)

$$F_{i_1, \dots, i_k}(x_1, \dots, x_k) = \mathbb{P}(X_1 \leq x_1, \dots, X_k \leq x_k).$$

Для заданной последовательности вещественных чисел $\{u_n = a_n x + b_n, \}$, $a_n > 0$, будем говорить, что выполнено условие перемешивания $D(u_n)$, если для любых целых чисел

$$1 \leq i_1 < \dots < i_k < j_1 < \dots < j_k \leq n,$$

для которых $j_1 - i_k \geq l_n$,

$$F_{i_1, \dots, i_k, j_1, \dots, j_k}(u_n) - F_{i_1, \dots, i_k}(u_n)F_{j_1, \dots, j_k}(u_n) \leq v_{n,l}.$$

Здесь $v_{n,l} \rightarrow 0$ при $n \rightarrow \infty$, $l \rightarrow \infty$; $F_{i_1, \dots, i_k}(u_n) := F_{i_1, \dots, i_k}(u_n, \dots, u_n)$.

Предположим, что для некоторого числа $\tau > 0$ определена такая последовательность $\{u_n(\tau)\}$, что имеет место сходимость

$$n\overline{F}(u_n(\tau)) \rightarrow \tau \text{ при } n \rightarrow \infty \quad (1)$$

и условие $D(u_n(\tau))$ выполнено для каждого такого τ . Тогда [3]

$$\mathbb{P}(M_n \leq u_n(\tau)) \rightarrow e^{-\theta\tau} \text{ при } n \rightarrow \infty, \quad (2)$$

где параметр $\theta \in [0, 1]$. Далее будем рассматривать стационарные последовательности, для которых выполнено условие перемешивания $D(u_n(\tau))$.

Заметим, что если $\mathbb{P}(M_n \leq u_n(\tau))$ сходится хотя бы для одного значения $\tau > 0$, то (2) выполнено при всех $\tau > 0$ для некоторого фиксированного θ .

Будем говорить, что стационарная последовательность $\{X_n, n \geq 1\}$ с общей ф. р. F имеет экстремальный индекс $\theta \in [0, 1]$, если при данном θ для каждого $\tau > 0$ существует такая последовательность вещественных чисел $\{u_n = u_n(\tau)\}$, что соотношения (1) и (2) выполнены (см. [4]).

Из определения экстремального индекса следует, что для больших n можно использовать следующую аппроксимацию распределения максимума M_n стационарной последовательности:

$$\mathbb{P}(M_n \leq u_n) \sim F^{n\theta}(u_n) \text{ для больших } n,$$

где $a \sim b$ означает $\lim a/b = 1$ при $n \rightarrow \infty$. Следовательно, экстремальный индекс θ служит ключевым параметром для изучения асимптотических распределений максимумов стационарных последовательностей.

Следующая теорема позволяет найти границы экстремального индекса строго стационарной последовательности $\{X_n\}$, заданной m -компонентной смесью распределений.

Теорема 1. Пусть строго стационарная последовательность $\{X_n\}$ задана общей ф. р. F_X вида

$$F_X(x) = p_1 F_{Y_1}(x) + \dots + p_m F_{Y_m}(x),$$

$$\sum_{i=1}^m p_i = 1, \quad p_i \geq 0,$$

т. е. каждый член последовательности X_i задан m -компонентной смесью распределений СВ Y^1, \dots, Y^m .

Пусть существуют такие нормализующие последовательности $\{u_n(x) = a_n x + b_n\}$ и $\{u_n^i(x) = a_n^i x + b_n^i\}$, что $a_n, a_n^i > 0$, $n \geq 1$, $u_n(x), u_n^i(x) \rightarrow \infty$ для каждого x при $n \rightarrow \infty, i = 1, \dots, m$, и пусть выполнены соотношения

$$\mathbb{P}(M_n^X \leq u_n(x)) \rightarrow H_X(x),$$

$$\mathbb{P}(M_n^{Y^i} \leq u_n^i(x)) \rightarrow H_{Y^i}(x), \quad n \rightarrow \infty; \quad (3)$$

$$n\overline{F}_X(u_n(x)) \rightarrow \tau(x), \quad n\overline{F}_{Y^i}(u_n^i(x)) \rightarrow \tau^i(x),$$

$$n \rightarrow \infty; \quad (4)$$

$$u_n^1(x) \geq u_n(x) \geq u_n^m(x) \text{ для всех } x, n \geq 1. \quad (5)$$

Пусть также стохастически упорядочены компоненты смеси, т. е.

$$\overline{F}_{Y^1}(x) \leq \overline{F}_{Y^2}(x) \leq \dots \leq \overline{F}_{Y^m}(x) \text{ для всех } x \geq 0.$$

Тогда экстремальные индексы θ_X, θ_{Y^1} и θ_{Y^m} последовательностей $\{X_n\}, \{Y_n^1\}$ и $\{Y_n^m\}$ упорядочены следующим образом:

$$\theta_{Y^1} \geq \theta_X \geq \theta_{Y^m}. \quad (6)$$

Доказательство. Из стохастической упорядоченности компонент смеси

$$Y^1 \underset{\text{st}}{\leq} \dots \underset{\text{st}}{\leq} Y^m$$

вытекает соотношение [19, теорема 1]:

$$Y^1 \underset{\text{st}}{\leq} X \underset{\text{st}}{\leq} Y^m.$$

Поэтому, если выполнены условия (3)–(5), то экстремальные индексы, соответствующие этим последовательностям, связаны неравенством (6) [18, теорема 1]. \square

3 Система $M/G/1$ с временем обслуживания, заданным конечной смесью распределений

Применим теорему 1 для оценки экстремального индекса стационарного времени ожидания в односерверной системе обслуживания с временем обслуживания, заданным конечной смесью распределений.

Рассмотрим односерверную систему Σ с пуассоновским входным потоком. Дисциплина обслуживания — первым пришел, первым обслужен. Обозначим S — типичное время обслуживания заявки и T — типичный интервал между приходами заявок, $\mathbb{E}T = 1/\lambda, \mathbb{E}S = 1/\mu$. Пусть время обслуживания S задано m -компонентной смесью распределений с ф. р. вида

$$B(x) = \sum_{i=1}^m p_i B_i(x),$$

$$\sum_{i=1}^m p_i = 1, \quad p_i \geq 0, \quad i = 1, \dots, m. \quad (7)$$

Предположим, что СВ $S^{(1)}, \dots, S^{(m)}$ независимы и имеют ф. р. $B_i(x), i = 1, \dots, m$. Пусть индикаторы I_i принимают два значения: $I_i = 0$ с вероятностью p_i и $I_i = 1$ с вероятностью $1 - p_i$ — и не зависят от $\{S^{(i)}\}, i = 1, \dots, m$. Тогда СВ

$$S = I_1 S^{(1)} + \dots + I_m S^{(m)}$$

является m -компонентной смесью с ф. р. (7). Математическое ожидание времени обслуживания вычисляется по формуле:

$$\mathbb{E}S = \frac{1}{\mu} = \sum_{i=1}^m \frac{p_i}{\mu_i},$$

где $\mathbb{E}S^{(i)} = 1/\mu_i$. Обозначим

$$\rho = \lambda \mathbb{E}S = \sum_{i=1}^m \frac{\lambda p_i}{\mu_i} = \sum_{i=1}^m p_i \rho_i$$

коэффициент загрузки системы Σ , где $\rho_i = \lambda/\mu_i, i = 1, \dots, m$.

Предположим, что компоненты времени обслуживания S стохастически упорядочены, т. е.

$$S^{(1)} \underset{\text{st}}{\leq} S^{(2)} \underset{\text{st}}{\leq} \dots \underset{\text{st}}{\leq} S^{(m)}.$$

Рассмотрим две системы обслуживания $M/G/1$: $\Sigma^{(1)}$ и $\Sigma^{(m)}$ с таким же входным потоком, как в исходной системе Σ , $\mathbb{E}T^{(i)} = 1/\lambda_i = 1/\lambda, i = 1, m$. (Будем снабжать индексом i величины, относящиеся к i -й системе.) Время обслуживания $S^{(1)}$ в системе $\Sigma^{(1)}$ задано ф. р. B_1 . Время обслуживания $S^{(m)}$ в системе $\Sigma^{(m)}$ задано ф. р. B_m . Будем говорить, что система $\Sigma^{(1)}$ является минорантной, а система $\Sigma^{(m)}$ — мажорантной для исходной системы Σ .

Теперь сравним в системах $\Sigma, \Sigma^{(1)}$ и $\Sigma^{(m)}$ экстремальные индексы стационарных времен ожидания.

Пусть $\nu_n(\nu_n^{(i)})$ — число заявок в системе, $Q_n(Q_n^{(i)})$ — размер очереди и $W_n(W_n^{(i)})$ — время ожидания в очереди в момент прихода заявки с номером n в систему $\Sigma(\Sigma^{(i)})$ соответственно, $i = \overline{1, m}$.

Обозначим (в случае их существования) пределы по распределению

$$Q_n^{(i)} \Rightarrow Q^{(i)}; \nu_n^{(i)} \Rightarrow \nu^{(i)}; W_n^{(i)} \Rightarrow W^{(i)}; \quad n \rightarrow \infty. \quad (8)$$

Эти пределы существуют, в частности, если времена между приходами заявок T являются *нерешетчатыми* и $\rho_i < 1, i = \overline{1, m}$ [12]. Заметим, что если соотношения (8) выполнены и $\rho_i < 1, i = \overline{1, m}$, то можно показать, что и для исходной системы Σ слабые пределы ν, Q и W также существуют.

Рассмотрим стационарные последовательности времен ожидания $\{W_n, n \geq 1\}, \{W_n^{(1)}, n \geq 1\}$ и $\{W_n^{(m)}, n \geq 1\}$ в системах $\Sigma, \Sigma^{(1)}$ и $\Sigma^{(m)}$ соответственно. Обозначим максимумы из n членов этих последовательностей ($n \geq 1$):

$$\begin{aligned} W_n^* &= \max(W_1, \dots, W_n); \\ W_n^{(1)*} &= \max(W_1^{(1)}, \dots, W_n^{(1)}); \\ W_n^{(m)*} &= \max(W_1^{(m)}, \dots, W_n^{(m)}). \end{aligned}$$

Теорема 2. *Предположим, что для систем $\Sigma^{(1)}$ и $\Sigma^{(m)}$ коэффициенты загрузки $\rho_i < 1, i = \overline{1, m}$, и выполнены следующие стохастические соотношения:*

$$\nu_1^{(1)} = \nu_1^{(m)} = \nu_1 = 0; \quad T_{st} = T_{st}^{(1)} = T_{st}^{(m)}. \quad (9)$$

Пусть компоненты смеси времени обслуживания стохастически упорядочены, т. е.

$$\overline{B}_1(x) \leq \dots \leq \overline{B}_m(x). \quad (10)$$

Предположим, что существуют такие нормализующие последовательности $\{u_n(x) = a_n x + b_n\}$ и $\{u_n^{(i)}(x) = a_n^{(i)} x + b_n^{(i)}\}$, что $a_n, a_n^{(i)} > 0, n \geq 1$, для каждого x $u_n(x), u_n^{(i)}(x) \rightarrow \infty$. Пусть

$$u_n^{(1)}(x) \geq u_n(x) \geq u_n^{(m)}(x)$$

и выполнены соотношения (3)–(4) для максимумов стационарных времен ожидания $W_n^, W_n^{(1)*}$ и $W_n^{(m)*}$. Тогда*

$$\theta_{W^{(1)}} \geq \theta_W \geq \theta_{W^{(m)}}, \quad (11)$$

где θ_W — экстремальный индекс W_n^ ; $\theta_{W^{(i)}}$ — экстремальный индекс $W_n^{(i)*}, i = 1, m$.*

Доказательство. Из неравенств (10) следует, что

$$S_{st}^{(1)} \leq S_{st} \leq S_{st}^{(m)},$$

что в совокупности с соотношениями (9) гарантирует (см. [21, теорема 4]), что

$$Q_n^{(1)} \leq Q_n \leq Q_n^{(m)}; \quad W_n^{(1)} \leq W_n \leq W_n^{(m)}, \quad n \geq 1.$$

Таким образом, все условия теоремы 1 выполнены, что влечет справедливость неравенств (11) для экстремальных индексов стационарных времен ожидания рассматриваемых трех систем. \square

4 Экстремальный индекс времени ожидания в системе $M/H_m/1$

В качестве исходной системы Σ рассмотрим систему $M/H_m/1$ с гиперэкспоненциальным (m -компонентным) распределением времени обслуживания с ф. р.

$$\begin{aligned} B(x) &= 1 - \sum_{i=1}^m p_i e^{-\mu_i x}, \\ \mu_i &> 0, \quad p_i \geq 0, \quad i = 1, \dots, m, \quad \sum_{i=1}^m p_i = 1. \end{aligned}$$

В качестве минорантной и мажорантной систем рассмотрим две $M/M/1$ системы $\Sigma^{(1)}$ и $\Sigma^{(m)}$, в которых времена обслуживания $S^{(i)}$ имеют показательный закон распределения с ф. р. $\overline{B}_i(x) = e^{-\mu_i x}, i = \overline{1, m}$. Пусть условия стационарности во всех системах выполнены:

$$\rho_i = \frac{\lambda}{\mu_i} < 1, \quad i = 1, \dots, m.$$

Функция интенсивности отказов для времени обслуживания S в исходной системе Σ равна

$$r(x) := \frac{f_B(x)}{\overline{B}(x)} = \frac{\sum_{i=1}^m p_i \mu_i e^{-\mu_i x}}{\sum_{i=1}^m p_i e^{-\mu_i x}},$$

где $f_B(x)$ — плотность распределения СВ S . Функции интенсивности отказов времен обслуживания $S^{(1)}$ и $S^{(m)}$ в системах $\Sigma^{(1)}$ и $\Sigma^{(m)}$ соответственно равны $r_1(x) = \mu_1$ и $r_m(x) = \mu_m$. Нетрудно проверить, что при условии выполнения неравенства

$$\mu_1 \geq \mu_2 \geq \dots \geq \mu_m \quad (12)$$

функции интенсивности отказов упорядочены следующим образом:

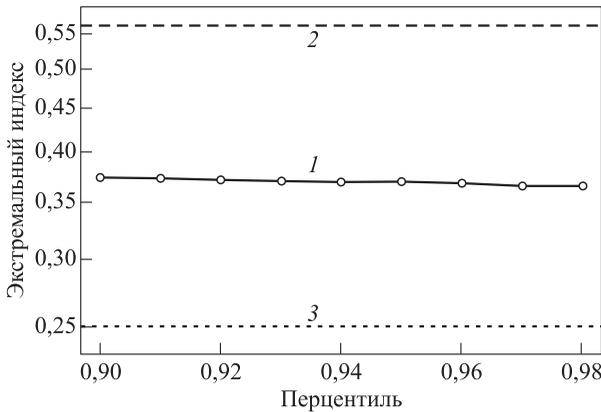
$$r_1(x) \geq r(x) \geq r_m(x), \quad x \geq 0,$$

а следовательно, времена обслуживания в этих системах упорядочены по интенсивности отказов

$$S_{st}^{(1)} \leq S_{st} \leq S_{st}^{(m)},$$

откуда, в свою очередь, вытекает их стохастическая упорядоченность [22]

$$S_{st}^{(1)} \leq S_{st} \leq S_{st}^{(m)}.$$



Экстремальные индексы в системах $M/M(\mu_1 = 4)/1$, $M/H_2(\mu_1 = 4, \mu_2 = 2)/1$ и $M/M(\mu_2 = 2)/1$, $\lambda = \lambda_1 = \lambda_2 = 1$

Поскольку $\Sigma^{(1)}$ и $\Sigma^{(m)}$ относятся к системам типа $M/M/1$, то для них известна ф. р. стационарного времени ожидания $F_{W^{(i)}}$ в явном виде:

$$F_{W^{(i)}}(x) = 1 - \rho_i e^{-(\mu_i - \lambda)x}.$$

Выбирая нормализующие последовательности для максимума стационарного времени ожидания вида

$$u_n^i(x) = \frac{x + \log(\rho_i n)}{\mu_i - \lambda}, \quad i = \overline{1, m},$$

получим его предельное распределение [8]

$$\lim_{n \rightarrow \infty} \mathbb{P}(W_n^{(i)*} \leq u_n^i(x)) = \Lambda(x), \quad i = \overline{1, m},$$

где $\Lambda(x) = \exp(-e^{-x})$ — распределение Гумбеля. Кроме того, для систем $M/M/1$ известны выражения для экстремальных индексов стационарных времен ожидания в явном виде [10]:

$$\theta_{W^{(1)}} = (1 - \rho_1)^2; \quad \theta_{W^{(m)}} = (1 - \rho_m)^2.$$

Вернемся к исходной системе $M/H_m/1$. Для систем $M/G/1$ экстремальный индекс можно вычислить по формуле [10]:

$$\theta_W^* = \frac{\gamma(1 - \rho)}{\gamma + \lambda},$$

где γ — единственное положительное решение уравнения

$$\mathbb{E}e^{\gamma(S-T)} = 1 \tag{13}$$

при условии, что $\mathbb{E}[|S - T|e^{\gamma(S-T)}] < \infty$. (В случае m -компонентного гиперэкспоненциального распределения уравнение (13) имеет несколько корней). Согласно теореме 2, при условии выполнения

соотношений (12) экстремальные индексы в этих системах связаны неравенством

$$\theta_{W^{(1)}} \geq \theta_W \geq \theta_{W^{(2)}}.$$

На рисунке приведены результаты эксперимента по моделированию экстремального индекса стационарного времени ожидания системы $M/H_2/1$ с параметрами $\lambda = 1$, $\mu_1 = 4$ и $\mu_2 = 2$. Для получения стационарного времени ожидания использовался метод точного моделирования [23]. Экстремальный индекс в исходной системе сравнивался с экстремальными индексами двух систем $M/M/1$ с таким же входным потоком и временем обслуживания, распределенным по показательному закону с параметром $\mu_1 = 4$ (в минорантной системе) и $\mu_2 = 2$ (в мажорантной системе) соответственно. Результаты моделирования подтверждают справедливость формулы (11).

5 Заключение

В работе показано, что если в исходной $M/G/1$ системе время обслуживания задано m -компонентной смесью распределений, компоненты которой стохастически упорядочены, то для нее можно построить две новые системы: минорантную и мажорантную системы, экстремальные индексы стационарных времен ожидания которых станут границами для экстремального индекса стационарного времени ожидания исходной системы. В качестве примера рассмотрена система $M/H_m/1$ с гиперэкспоненциальным временем обслуживания. Минорантной (мажорантной) системой послужила система $M/M/1$ с таким же входным потоком, время обслуживания которой имеет показательное распределение с наибольшим (наименьшим) параметром среди всех параметров гиперэкспоненциального распределения $\mu_1 = \max(\mu_1, \dots, \mu_m)$ ($\mu_m = \min(\mu_1, \dots, \mu_m)$ для мажорантной системы). Результаты моделирования $M/H_2/1$ показывают устойчивость оценки экстремального индекса в исходной системе при разных перцентилях. При этом оценка экстремального индекса ограничена значениями экстремальных индексов мажорантной и минорантной систем. Метод точного моделирования для оценки экстремального индекса стационарного времени ожидания может быть применен для других распределений времени обслуживания, для которых не существует решения уравнения (13) в явном виде, что станет темой будущих исследований.

Литература

1. Bertail P., Clemenccon S., Tressou J. Extreme values statistics for Markov chains via the (pseudo-) regenerative

- method // *Extremes*, 2009. Vol. 12. Iss. 4. P. 327–360. doi: 10.1007/s10687-009-0081-y.
2. *Resnick S.* Extreme values, regular variation and point processes. — New York, NY, USA: Springer, 1987. 320 p.
 3. *Leadbetter M. R., Lindgren G., Rootzén H.* Extremes and related properties of random sequences and processes. — New York, NY, USA: Springer, 1983. 336 p.
 4. *Embrechts P., Kluppelberg C., Mikosch T.* Modelling extremal events for insurance and finance. Applications of mathematics. — Berlin, Heidelberg: Springer, 1997. 660 p.
 5. *De Haan L., Ferreira A.* Extreme value theory: An introduction. — New York, NY, USA: Springer Science + Business Media LLC, 2006. 491 p.
 6. *Smith L., Weissman I.* Estimating the extremal index // *J. Roy. Stat. Soc. B Met.*, 1994. Vol. 56. Iss. 3. P. 515–528. doi: 10.1111/J.2517-6161.1994.TB01997.X.
 7. *Weissman I., Novak S.* On blocks and runs estimators of extremal index // *J. Stat. Plan. Infer.*, 1998. Vol. 66. Iss. 2. P. 281–288. doi: 10.1016/S0378-3758(97)00095-5.
 8. *Iglehart D. L.* Extreme values in $GI/G/1$ queue // *Ann. Math. Stat.*, 1972. Vol. 3. Iss. 2. P. 627–635. doi: 10.1214/aoms/1177692642.
 9. *Rootzén H.* Maxima and exceedances of stationary Markov chains // *Adv. Appl. Probab.*, 1988. Vol. 20. Iss. 2. P. 371–390. doi: 10.2307/1427395.
 10. *Hooghiemstra G., Meester L. E.* Computing the extremal index of special Markov chains and queues // *Stoch. Proc. Appl.*, 1996. Vol. 65. Iss. 2. P. 171–185. doi: 10.1016/S0304-4149(96)00111-1.
 11. *Asmussen S.* Extreme value theory for queues via cycle maxima // *Extremes*, 1998. Vol. 1. Iss. 2. P. 137–168. doi: 10.1023/A:1009970005784.
 12. *Asmussen S.* Applied probability and queues. Stochastic modelling and applied probability. — New York, NY, USA: Springer-Verlag, 2003. 438 p.
 13. *Markovich N., Razumchik R.* Cluster modeling of Lindley process with application to queuing // *Distributed computer and communication networks*. — Lecture notes in computer science ser. — Cham, Switzerland: Springer, 2019. Vol. 11965. P. 330–341. doi: 10.1007/978-3-030-36614-8_25.
 14. *Markovich N.* Clustering and hitting times of threshold exceedances and applications // *Int. J. Data Analysis Techniques Strategies*, 2017. Vol. 9. Iss. 4. P. 331–347. doi: 10.1504/IJDATS.2017.10009424.
 15. *Rodionov I.* On threshold selection problem for extremal index estimation // *Recent developments in stochastic methods and applications* / Eds. A. N. Shiryaev, K. E. Samouylov, D. V. Kozyrev. — Springer proceedings in mathematics & statistics ser. — Cham, Switzerland: Springer, 2021. Vol. 371. P. 3–16. doi: 10.1007/978-3-030-83266-7_1.
 16. *Feldman A., Whitt W.* Fitting mixtures of exponentials to long-tail distributions to analyze network performance models // *Perform. Evaluation*, 1998. Vol. 31. Iss. 3-4. P. 245–279. doi: 10.1016/S0166-5316(97)00003-5.
 17. *Peshkova I., Morozov E., Maltseva M.* On regenerative estimation of extremal index in queueing systems // *Distributed computer and communication networks*. — Lecture notes in computer science ser. — Cham, Switzerland: Springer, 2021. Vol. 13144. P. 251–264. doi: 10.1007/978-3-030-92507-9_21.
 18. *Peshkova I.* Сравнение экстремальных индексов времени ожидания в системах обслуживания $M/G/1$ // *Информатика и её применения*, 2022. Том. 16. Вып. 1. С. 61–67. doi: 10.14357/19922264220109.
 19. *Peshkova I., Morozov E.* On comparison of multiserver systems with multicomponent mixture distributions // *J. Math. Sci.*, 2022. Vol. 267. No. 2. P. 260–272. doi: 10.1007/s10958-022-06132-z.
 20. *Peshkova I., Morozov E., Maltseva M.* On comparison of waiting time extremal indexes in queueing systems with Weibull service times // *Information technologies and mathematical modelling. Queueing theory and applications*. — Communications in computer and information science ser. — Cham, Switzerland: Springer, 2022. Vol. 1605. P. 80–92. doi: 10.1007/978-3-031-09331-9_7.
 21. *Whitt W.* Comparing counting processes and queues // *Adv. Appl. Probab.*, 1981. Vol. 13. P. 207–220. doi: 10.2307/1426475.
 22. *Ross S., Shanthikumar J., Zhu Z.* On increasing-failure-rate random variables // *J. Appl. Probab.*, 2005. Vol. 42. P. 797–809. doi: 10.1239/jap/1127322028.
 23. *Sigman K.* Exact simulation of the stationary distribution of the FIFO $M/G/c$ queue: The general case for $\rho < c$ // *Queueing Syst.*, 2012. Vol. 70. P. 37–43. doi: 10.1007/s11134-011-9266-6.

Поступила в редакцию 15.10.22

ON BOUNDS OF THE STATIONARY WAITING TIME EXTREMAL INDEX IN $M/G/1$ SYSTEM WITH MIXTURE SERVICE TIMES

I. V. Peshkova^{1,2}¹Petrozavodsk State University, 33 Lenina Prosp., Petrozavodsk 185910, Russian Federation²Karelian Research Centre of the Russian Academy of Sciences, 11 Pushkinskaya Str., Petrozavodsk 185910, Russian Federation

Abstract: It is proved that if the original stationary sequence has m -component mixture distribution with stochastically ordered components, there are limit distributions for the maxima of all components, and the normalizing sequences are ordered, then the extremal index of the original sequence is within the boundaries of the extremal indexes of the smallest and largest components. This result is used to estimate the extremal index of the stationary waiting time in a queuing system of type $M/G/1$ in which the queuing time is given by an m -component distribution mixture. An example of a system $M/H_m/1$ with hyperexponential service time is considered. Using the exact simulation approach, the results of estimating the extremal index of stationary waiting time in the system $M/H_2/1$ are obtained.

Keywords: extreme value distributions; extremal index; queuing system; stochastic ordering

DOI: 10.14357/19922264220405

Acknowledgments

The research has been prepared with the support of the Russian Science Foundation (project No. 21-71-10135).

References

- Bertail, P., S. Clemencon, and J. Tressou. 2009. Extreme values statistics for Markov chains via the (pseudo-) regenerative method. *Extremes* 12(4):327–360. doi: 10.1007/s10687-009-0081-y.
- Resnick, S. 1987. *Extreme values, regular variation and point processes*. New York, NY: Springer. 320 p.
- Leadbetter, M. R., G. Lindgren, and H. Rootzen. 1983. *Extremes and related properties of random sequences and processes*. New York, NY: Springer. 336 p.
- Embrechts, P., C. Kluppelberg, and T. Mikosch. 1997. *Modelling extremal events for insurance and finance*. Berlin, Heidelberg: Springer. 660 p.
- De Haan, L., and A. Ferreira. 2006. *Extreme value theory: An introduction*. New York, NY: Springer Science + Business Media LLC. 491 p.
- Smith, L., and I. Weissman. 1994. Estimating the extremal index. *J. Roy. Stat. Soc. B Met.* 56(3):515–528. doi: 10.1111/J.2517-6161.1994.TB01997.X.
- Weissman, I., and S. Novak. 1998. On blocks and runs estimators of extremal index. *J. Stat. Plan. Infer.* 66(2):281–288. doi: 10.1016/S0378-3758(97)00095-5.
- Iglehart, D. L. 1972. Extreme values in GI/G/1 queue. *Ann. Math. Stat.* 3(2):627–635. doi: 10.1214/aoms/1177692642.
- Rootzen, H. 1988. Maxima and exceedances of stationary Markov chains. *Adv. Appl. Probab.* 20(2):371–390. doi: 10.2307/1427395.
- Hooghiemstra, G., and L. E. Meester. 1996. Computing the extremal index of special Markov chains and queues. *Stoch. Proc. Appl.* 65(2):171–185. doi: 10.1016/S0304-4149(96)00111-1.
- Asmussen, S. 1998. Extreme value theory for queues via cycle maxima. *Extremes* 1(2):137–168. doi: 10.1023/A:1009970005784.
- Asmussen, S. 2003. *Applied probability and queues*. New York, NY: Springer-Verlag. 438 p.
- Markovich, N., and R. Razumchik. 2019. Cluster modeling of Lindley process with application to queuing. *Distributed computer and communication networks. Lecture notes in computer science ser.* Cham, Switzerland: Springer. 11965:330–341. doi: 10.1007/978-3-030-36614-8_25.
- Markovich, N. 2017. Clustering and hitting times of threshold exceedances and applications. *Int. J. Data Analysis Techniques Strategies* 9(4):331–347. doi: 10.1504/IJDATS.2017.10009424.
- Rodionov, I. 2021. On threshold selection problem for extremal index estimation. *Recent developments in stochastic methods and applications*. Eds. A. N. Shiryaev, K. E. Samouylov, and D. V. Kozyrev. Springer proceedings in mathematics & statistics ser. Cham, Switzerland: Springer. 371:3–16. doi: 10.1007/978-3-030-83266-7_1.
- Feldman, A., and W. Whitt. 1998. Fitting mixtures of exponentials to long-tail distributions to analyze network performance models. *Perform. Evaluation* 31(3–4):245–279. doi: 10.1016/S0166-5316(97)00003-5.
- Peshkova, I., E. Morozov, and M. Maltseva. 2021. On regenerative estimation of extremal index in queueing systems. *Distributed computer and communication networks*. Eds. V. M. Vishnevskiy, K. E. Samouylov, and

- D.V. Kozyrev. Lecture notes in computer science ser. Cham, Switzerland: Springer. 13144:251–264. doi: 10.1007/978-3-030-92507-9_21.
18. Peshkova, I. 2022. Sravneniye ekstremal'nykh indeksov vremen ozhidaniya v sistemakh obsluzhivaniya $M/G/1$ [The comparison of waiting time extremal indexes in $M/G/1$ queueing systems]. *Informatika i ee Primeneniya — Inform. Appl.* 16(1):61–67. doi: 10.14357/19922264220109.
 19. Peshkova, I., and E. Morozov. 2022. On comparison of multiserver systems with multicomponent mixture distributions. *J. Math. Sci.* 267(2):260–272. doi: 10.1007/s10958-022-06132-z.
 20. Peshkova, I., E. Morozov, and M. Maltseva. 2022. On comparison of waiting time extremal indexes in queueing systems with Weibull service times. *Information technologies and mathematical modelling. Queueing theory and applications*. Communications in computer and information science ser. Cham, Switzerland: Springer. 1605:80–92. doi: 10.1007/978-3-031-09331-9_7.
 21. Whitt, W. 1981. Comparing counting processes and queues. *Adv. Appl. Probab.* 13:207–220. doi: 10.2307/1426475.
 22. Ross, S., J. Shanthikumar, and Z. Zhu. 2005. On increasing–failure–rate random variables. *J. Appl. Probab.* 42:797–809. doi: 10.1239/jap/1127322028.
 23. Sigman, K. 2012. Exact simulation of the stationary distribution of the FIFO $M/G/c$ queue: The general case for $\rho < c$. *Queueing Syst.* 70: 37–43. doi: 10.1007/s11134-011-9266-6.

Received October 15, 2022

Contributor

Peshkova Irina V. (b. 1975) — Candidate of Science (PhD) in physics and mathematics, associate professor, Petrozavodsk State University, 33 Lenina Prosp., Petrozavodsk 185910, Russian Federation; senior scientist, Karelian Research Centre of the Russian Academy of Sciences, 11 Pushkinskaya Str., Petrozavodsk 185910, Russian Federation; iaminova@petsu.ru

ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ ПОДКЛЮЧЕНИЕМ РЕЗЕРВНОГО ПРИБОРА В СИСТЕМЕ МАССОВОГО ОБСЛУЖИВАНИЯ $G/M/1^*$

Я. М. Агаларов¹

Аннотация: Рассматривается задача оптимизации режима подключения и отключения резервного прибора в системе массового обслуживания (СМО) $G/M/1$ в зависимости от длины очереди заявок на обслуживание. В качестве целевой функции используется суммарный предельный средний доход системы в единицу времени, рассчитываемый как разность платы за обслуживание заявок и суммарных потерь за задержку заявки в очереди и амортизацию резервного прибора. Рассмотрена также аналогичная задача для СМО $G/M/1/r$ с целевой функцией, учитывающей дополнительные потери дохода из-за отклонения заявок. Доказано свойство унимодальности функции дохода систем $G/M/1/r$ и $G/M/1$ и сформулированы необходимые и достаточные условия существования конечного оптимального порога длины очереди, который служит критерием при подключении и отключении резервного прибора. Предложен простой алгоритм оптимального управления порогом, требующий только наблюдения за величиной дохода системы.

Ключевые слова: система массового обслуживания; резервирование; управление; оптимизация

DOI: 10.14357/19922264220406

1 Введение

Одним из способов повышения надежности и эффективности системы служит включение в ее состав (на стадии проектирования или эксплуатации) резервных элементов. Очень часто при исследовании реальных систем с резервными элементами (как и многих других систем), предназначенных для обслуживания случайных потоков заявок, на практике появляется необходимость применения математических моделей типа СМО, и в ряде публикаций, посвященных этим моделям, рассмотрена задача управления включением и отключением резервного прибора с целью повышения качества обслуживания заявок при одновременном снижении затрат на ресурсы (минимизации потерь) СМО [1–6]. Задача в них сформулирована как задача поиска оптимального порогового управления подключением или отключением резервного прибора в зависимости от достижения длиной очереди или временем пребывания заявки соответствующих пороговых значений. Подобные задачи относятся к классу управляемых СМО, в которых допускается воздействие на поведение системы (например, путем введения управления некоторыми параметрами) в зависимости от текущего состояния параметров системы с целью повышения качества работы системы. В публикациях, посвященных этому классу систем, основное внимание уделялось задачам поиска оптимальных дисциплин обслуживания заявок, и мало встречается работ, поставив-

ших задачу по управлению интенсивностью обслуживания, в том числе по принятию оптимальных решений при подключении резервных приборов. В известных автору публикациях задача по выбору оптимального режима подключения резервного прибора решена в предположении пуассоновского входящего потока, предлагаемые методы решения являются численными и сложны для построения простых процедур оптимального управления.

Ниже приведены результаты исследования задачи оптимизации простой пороговой стратегии подключения и отключения резервного прибора СМО типа $G/M/1$. Данная работа дополняет результаты, полученные в [6], и использует метод решения, приведенный в работах [7, 8].

2 Описание системы и постановка задачи

Рассматривается однолинейная СМО с неограниченной очередью и резервным прибором, на которую поступает рекуррентный поток заявок с функцией распределения вероятностей $U(t)$ со средним значением $0 < \bar{v} < \infty$. Заявки становятся в очередь на обслуживание в порядке поступления и покидают систему только при завершении обслуживания, освободив одновременно прибор и накопитель, а на освободившийся прибор (если он не отключен) поступает первая в очереди заявка из накопителя (если таковая есть). При подключении

* Работа выполнена при поддержке РФФИ (проект 20-07-00804).

¹ Федеральное исследовательское учреждение «Информатика и управление» Российской академии наук, agglar@yandex.ru

резервный прибор сразу начинает обслуживать первую в очереди заявку. Подключение и отключение прибора происходит в зависимости от числа заявок в системе следующим образом: резервный прибор подключается к обслуживанию (если он еще не подключен) в момент поступления заявки в систему, если число заявок в системе (с учетом поступившей) не меньше заранее заданного порогового значения $h_1 \geq 2$, и резервный прибор отключается (если он подключен) от системы, если в момент освобождения любого из приборов в системе остается не более $(h_1 - 1)$ заявок. При этом предполагается, что если уходит заявка с основного прибора, то заявка с резервного прибора переходит на основной.

Время обслуживания заявки на основном и резервном приборах распределено по экспоненциальному закону с параметрами $\mu_1 > 0$ и $\mu_2 > 0$ соответственно, $\bar{v} < (\mu_1 + \mu_2)^{-1}$. В дальнейшем в каждом состоянии заявки в очереди будем считать помеченными условными номерами $1, 2, \dots$ в порядке поступления, не меняющимися до перехода в новое состояние.

Ниже в качестве целевой функции для этой системы используется интегральный показатель (который будем называть функцией дохода), учитывающий следующие стоимостные параметры:

$C_0 \geq 0$ — плата, получаемая системой, если поступившая заявка обслужена системой (принята в накопитель);

$C_1 \geq 0$ — потери в единицу времени за ожидание заявки в системе;

$C_2 \geq 0$ — потери на амортизацию резервного прибора в единицу времени, когда он подключен к системе и занят заявкой;

$C_3 \geq 0$ — потери на амортизацию резервного прибора в единицу времени, когда он простаивает ($C_3 \leq C_2$);

$C_4 \geq 0$ — потери из-за отклонения поступившей заявки при переполнении накопителя.

Обозначим:

D^{h_1} — предельное среднее значение дохода системы в единицу времени при пороге h_1 ;

$g^{h_1}(a)$ — предельное среднее значение суммарного дохода системы, усредненного по числу поступивших заявок при пороге h_1 .

Ставится задача максимизации функции D^{h_1} по $h_1 > 1$, которая эквивалентна задаче (так как $D^{h_1} = g^{h_1}/\bar{v}$) найти оптимальный порог $h_1^* > 1$ такой, что

$$\max_{h_1 > 1} g^{h_1} = g^{h_1^*}. \quad (1)$$

3 Метод решения

Рассмотрим задачу (1) для СМО с накопителем емкостью $h_2 = h_1 + a$, $h_1 \geq 2$, $a = const \geq 0$, которая отличается от исходной только тем, что заявка, заставшая систему в состоянии h_2 , теряется.

Отметим, что процесс обслуживания заявок в данной системе описывается цепью Маркова, где переходы цепи определяются моментами поступления заявок, а состояние цепи — это число заявок, находящихся в системе в момент поступления [9].

Введем обозначения:

$\{\pi_i^{h_1}, 0 \leq i \leq h_2\}$ — стационарное распределение вероятностей цепи при пороге h_1 ($\pi_i^{h_1}$ — стационарная вероятность того, что цепь находится в состоянии i);

$d_i^{h_1}$ — средний доход, получаемый системой в состоянии i при пороге h_1 ;

$g^{h_1}(a)$ — предельное среднее значение суммарного дохода системы.

Из определения вложенной цепи Маркова следует

$$g^{h_1}(a) = \sum_{i=0}^{h_2} \pi_i^{h_1} d_i^{h_1}. \quad (2)$$

Выпишем выражения для стационарных вероятностей $\pi_i^{h_1}$, $i = 0, \dots, h_2$.

Обозначим:

$r_{i,m}^{h_1}$ — вероятность того, что в состоянии i будут обслужены ровно m заявок при условии, что в этом состоянии число подключенных приборов не изменится;

$a_{i,m}^{h_1}$ — вероятность того, что в состоянии i будут обслужены ровно m заявок при условии, что в этом состоянии число подключенных приборов изменится.

Справедливы формулы:

$$r_{i,m}^{h_1} = \int_0^\infty \frac{(\mu_i^{h_1} v)^m}{m!} e^{-\mu_i^{h_1} t} dU(v);$$

$$a_{i,m}^{h_1} = \int_0^v \int_0^v \frac{\mu_i^{h_1} (\mu_i^{h_1} t)^{i-h_1+1}}{(i-h_1+1)!} e^{-\mu_i^{h_1} t} \times \frac{[\mu_1(v-t)]^{m-i+h_1-2}}{(m-i+h_1-2)!} e^{-\mu_1(v-t)} dt dU(v),$$

где

$$\mu_i^{h_1} = \begin{cases} \mu_1 & \text{при } 0 \leq i \leq h_1 - 2; \\ \mu_1 + \mu_2 & \text{при } h_1 - 1 \leq i \leq h_2. \end{cases}$$

Составив систему уравнений равновесия для неизвестных стационарных вероятностей состояний рассматриваемой цепи $\pi_i^{h_1}$, $0 \leq i \leq h_2$, и заменив в каждом уравнении $\pi_j^{h_1}$ на $R_j^{h_1} \pi_{h_2}^{h_1}$ и исключив последовательно в каждом уравнении $R_j^{h_1}$ с наименьшим индексом, получим рекуррентные формулы для вычисления стационарных вероятностей состояний $\pi_j^{h_1}$, $0 \leq j \leq h_2$:

$$\pi_j^{h_1} = \frac{R_j^{h_1}}{\sum_{i=0}^{h_2} R_i^{h_1}}, \quad 0 \leq j \leq h_2, \quad (3)$$

где

$$\begin{aligned} R_{h_2}^{h_1} &= 1; & R_{h_2-1}^{h_1} &= \frac{1 - r_{h_2,0}^{h_1}}{r_{h_2,0}^{h_1}}; \\ R_{j-1}^{h_1} &= \frac{R_j^{h_1}(1 - r_{j,1}^{h_1})}{r_{j-1,0}^{h_1}} - \\ &= \frac{\sum_{i=j+1}^{h_2-1} R_i^{h_1} r_{j,i+1-j}^{h_1} + R_{h_2}^{h_1} r_{h_2,h_2-j}^{h_1}}{r_{j-1,0}^{h_1}}, & h_1 \leq j \leq h_2 - 1; \\ R_{h_1-2}^{h_1} &= \frac{R_{h_1-1}^{h_1}(1 - r_{h_1-1,1}^{h_1})}{r_{h_1-2,0}^{h_1}} - \\ &= \frac{\sum_{i=h_1}^{h_2-1} R_i^{h_1} r_{i,i+2-h_1}^{h_1} + R_{h_2}^{h_1} r_{h_2,h_2-h_1+1}^{h_1}}{r_{h_1-2,0}^{h_1}}; \\ R_{j-1}^{h_1} &= \frac{1}{r_{j-1,0}^{h_1}} \left(R_j^{h_1}(1 - r_{j,1}^{h_1}) - \sum_{i=j+1}^{h_1-2} R_i^{h_1} r_{i,i+1-j}^{h_1} - \right. \\ &\quad \left. - \sum_{i=h_1-1}^{h_2-1} R_i^{h_1} a_{i,i+1-j}^{h_1} \right) - \frac{R_{h_2}^{h_1} a_{h_2,h_2-j}^{h_1}}{r_{j-1,0}^{h_1}}, & 1 \leq j \leq h_1 - 2. \end{aligned}$$

Из (3) следует соотношение:

$$\pi_{j+1}^{h_1+1} = (1 - \pi_0^{h_1+1}) \pi_j^{h_1}, \quad 0 \leq j \leq h_2. \quad (4)$$

В дальнейшем для краткости изложения будем пользоваться обозначениями:

$$\begin{aligned} r_{i,m}^{h_1}(v) &= \frac{(\mu_i^{h_1} v)^m}{m!} e^{-\mu_i^{h_1} v}; \\ z_i^{h_1}(v, t, m) &= \frac{\mu_i^{h_1} (\mu_i^{h_1} t)^{i-h_1+1}}{(i-h_1+1)!} e^{-\mu_i^{h_1} t} \times \\ &\quad \times \frac{[\mu_1(v-t)]^{m-i+h_1-2}}{(m-i+h_1-2)!} e^{-\mu_1(v-t)}. \end{aligned}$$

Получим явные выражения для параметров $d_i^{h_1}$ в формуле (2) для $g^{h_1}(a)$.

Лемма 1. Среднее значение дохода, получаемого системой при пороге h_1 в состоянии i , равно

$$d_i^{h_1} = \begin{cases} C_0 - C_1 \bar{T}_i^{h_1} - C_2 (\bar{v} - \bar{T}_{\text{пп1},i}^{h_1}) - \\ \quad - C_3 \bar{T}_{\text{пп1},i}^{h_1}, & 0 \leq i \leq h_2 - 1; \\ d_{h_2-1}^{h_1} - C_0 - C_4, & i = h_2, \end{cases} \quad (5)$$

где $\bar{T}_i^{h_1}$ — среднее суммарное время задержки заявок в очереди за время нахождения системы в состоянии i :

$$\begin{aligned} \bar{T}_i^{h_1} &= \\ &= \left\{ \begin{aligned} &\frac{1}{\mu_i^{h_1}} \left[i \sum_{m=1}^{i+1} m r_{i,m}^{h_1} - \frac{1}{2} \sum_{m=1}^{i+1} (m-1) m r_{i,m}^{h_1} + \right. \\ &\quad \left. + \frac{1}{2} i(i+1) \sum_{m=i+2}^{\infty} r_{i,m}^{h_1} \right], \quad i \leq h_1 - 2; \\ &\frac{1}{\mu_i^{h_1}} \left[i \sum_{m=1}^{i-h_1+3} m r_{i,m}^{h_1} - \right. \\ &\quad \left. - \frac{1}{2} \sum_{m=1}^{i-h_1+3} (m-1) m r_{i,m}^{h_1} \right] + \\ &\quad + \sum_{m=i-h_1+3}^{\infty} \int_0^v \int_0^v \left[\frac{i-h_1+1}{2} t + \right. \\ &\quad \left. + (m-i+h_1-1)t + \right. \\ &\quad \left. + \frac{(m-i+h_1-2)(v-t)}{2} \right] \times \\ &\quad \times x_i^{h_1}(v, t, m) dt dU(v) + \\ &\quad + \sum_{m=i-h_1+3}^{i-1} \int_0^v \int_0^v (i-m)v \times \\ &\quad \times z_i^{h_1}(v, t, m) dt dU(v) + \\ &\quad + \sum_{m=i+1}^{\infty} \int_0^v \int_0^v \left[\frac{i-h_1+1}{2} t + (h_1-1)t + \right. \\ &\quad \left. + \frac{(h_1-2)(h_1-1)(v-t)}{2(m-i+h_1-1)} \right] \times \\ &\quad \times z_i^{h_1}(v, t, m) dt dU(v), \quad i \geq h_1 - 1; \end{aligned} \right. \quad (6) \end{aligned}$$

$\bar{T}_{\text{пп1},i}^{h_1}$ — среднее время простоя резервного прибора в состоянии i :

$$\begin{aligned} \bar{T}_{\text{пп1},i}^{h_1} &= \\ &= \begin{cases} \bar{v}, & i \leq h_1 - 2; \\ \bar{v} - \int_0^v \int_0^v \frac{\mu_i^{h_1} (\mu_i^{h_1} t)^{i-h_1+1}}{(i-h_1+1)!} e^{-\mu_i^{h_1} t} dt dU(v), & i \geq h_1 - 1. \end{cases} \quad (7) \end{aligned}$$

Среднее значение времени простоя основного прибора $\bar{T}_{\text{пр}2,i}^{h_1}$ при пороге h_1 в состоянии i равно

$$\bar{T}_{\text{пр}2,i}^{h_1} = \begin{cases} \sum_{m=i+1}^{\infty} \int_0^{\infty} \int_0^v \left[v-t - \frac{(h_1-1)(v-t)}{m-i+h_1-1} \right] \times \\ \times z_i^{h_1}(v,t,m) dt dU(v), & i \geq h_1 - 1; \\ \frac{1}{\mu_1} \sum_{m=i+2}^{\infty} (m-i-1) r_{i,m}^{h_1}, & i \leq h_1 - 2. \end{cases} \quad (8)$$

Доказательство. Справедливость формулы (5) следует непосредственно из определений $d_i^{h_1}$, $\bar{T}_{\text{пр}1,i}^{h_1}$ и из того, что при $0 \leq i \leq h_2 - 1$ система получает плату C_0 и нет потерь C_4 , а при $i = h_2$ система из-за отклонения поступившей заявки не получает плату и несет потери C_4 .

Докажем равенство (6). Фиксируем состояние i и v — время нахождения системы в состоянии i . Найдем сначала выражения для суммарного среднего времени ожидания всех заявок в очереди и средние времена простоя резервного и основного приборов за время v .

Пусть T_l , $l \geq 1$, — время ожидания l -й по номеру заявки за время нахождения системы в состоянии i , если ее обслуживание на приборе завершилось или началось; $A_{m,t}$ — событие, состоящее в том, что за время t завершилось обслуживание ровно m заявок и за это время число подключенных приборов не изменилось; $E_{m,t}$ — событие, состоящее в том, что за время v нахождения в состоянии $i \geq h_1 - 1$ завершилось обслуживание ровно $m > i - h_1 + 2$ заявок и обслуживание $(i - h_1 + 2)$ -й заявки завершилось в момент $t \in (0, v]$.

Отметим, что математическое ожидание случайной величины (СВ) T_l при условии выполнения события $A_{m,t}$ равно $lt/(m+1)$, $l \leq m$ (см., например, [6, 9]). Используя эту формулу, получим следующие выражения условных математических ожиданий для СВ T_l :

- среднее СВ T_l при условии выполнения события $A_{m,v}$ и $l \leq m$ равно

$$\bar{T}_{l/m} = \frac{l}{m+1} v \text{ при } i \leq h_1 - 2 \\ \text{или } i \geq h_1 - 1, l \leq i - h_1 + 1;$$

- среднее СВ T_l при $l \geq i - h_1 + 2 \geq 1$ и условии, что произошло событие $E_{m,t}$, равно

$$\bar{T}_{l/m,t} = t + \frac{(l-i+h_1-2)(v-t)}{m-i+h_1-1}.$$

Следовательно, суммарное среднее время ожидания всех заявок в системе, обслуживание которых завершилось или началось в состоянии i , равно:

$$\bar{T}_{\text{обсл}/m} = \begin{cases} \frac{mv}{2}, & m \leq i, i \leq h_1 - 2 \text{ или} \\ & i \geq h_1 - 1, m \leq i - h_1 + 2; \\ \frac{i(i+1)v}{2(m+1)}, & m > i, i \leq h_1 - 2, \end{cases}$$

при условии что произошло $A_{m,v}$;

$$\bar{T}_{\text{обсл}/m,t} = \begin{cases} mt - \frac{i-h_1+1}{2}t + \frac{(m-i+h_1-2)(v-t)}{2}, & i \geq m > i - h_1 + 2, i \geq h_1; \\ \frac{i-h_1+1}{2}t + (h_1-1)t + \\ + \frac{(h_1-2)(h_1-1)(v-t)}{2(m-i+h_1-1)}, & m > i, i \geq h_1, \end{cases}$$

при условии что произошло $E_{m,t}$.

Суммарное среднее время ожидания заявок в системе, обслуживание которых не началось в состоянии i , при условии что число уже обслуженных заявок $m < i$, равно $\bar{T}_{\text{необсл}/m} = (i-m)v$.

Для $\bar{T}_i^{h_1}$, используя полученные выше выражения для $\bar{T}_{\text{обсл}/m}$, $\bar{T}_{\text{обсл}/m,t}$, $\bar{T}_{\text{необсл}/m}$ и формулу полной вероятности на соответствующей группе несовместных событий, получим формулу (6).

Рассуждая аналогично выводу формулы (6), получим формулы (7) и (8).

Приведем еще три леммы.

Лемма 2. Справедливы равенства

$$d_{i+1}^{h_1+1} - d_i^{h_1} = \begin{cases} -\frac{C_1}{\mu_i^{h_1}} \sum_{m=1}^{i+1} m r_{i,m}^{h_1} - \frac{C_1(i+1)}{\mu_i^{h_1}} \sum_{m=i+2}^{\infty} r_{i,m}^{h_1} \\ \text{при } i \leq h_1 - 2; \\ -C_1 \int_0^{\infty} v \left[\sum_{m=0}^{i-h_1+2} r_{i,m}^{h_1}(v) + \right. \\ \left. + \sum_{m=i-h_1+3}^i \int_0^v z_i^{h_1}(v,t,m) dt \right] dU(v) - \\ - C_1 \sum_{m=i+1}^{\infty} \int_0^{\infty} \int_0^v \left[t + \frac{(h_1-1)(v-t)}{m-i+h_1-1} \right] \times \\ \times z_i^{h_1}(v,t,m) dt dU(v) \text{ при } i \geq h_1 - 1. \end{cases} \quad (9)$$

Доказательство. Равенства (9) непосредственно следуют из (5), (6) и (7) после несложных преобразований.

Введем обозначения:

$$\begin{aligned}
 w(h_1) &= \sum_{i=0}^{h_1-2} \pi_i^{h_1} \frac{1}{\mu_i^{h_1}} \sum_{m=i+2}^{\infty} (m-i-1)r_{i,m}^{h_1} + \\
 &+ \sum_{i=h_1-1}^{h_2-1} \pi_i^{h_1} \sum_{m=i+1}^{\infty} \int_0^{\infty} \left[v - \int_0^v \left[t + \frac{(h_1-1)(v-t)}{m-i+h_1-1} \right] \times \right. \\
 &\quad \left. \times z_i^{h_1}(v, t, m) dt \right] dU(v) + \\
 &+ \pi_{h_2}^{h_1} \sum_{m=h_2}^{\infty} \int_0^{\infty} \left[v - \int_0^v \left[t + \frac{(h_1-1)(v-t)}{m-h_2+h_1} \right] \times \right. \\
 &\quad \left. \times z_i^{h_1}(v, t, m) dt \right] dU(v); \quad (10)
 \end{aligned}$$

$$\begin{aligned}
 f(h_1, a) &= C_0 - C_3 \bar{v} - \\
 &- C_1 \frac{1 - \pi_0^{h_1+1}}{\pi_0^{h_1+1}} [\bar{v} - w(h_1, a)]. \quad (11)
 \end{aligned}$$

Лемма 3. *Справедливо соотношение*

$$g^{h_1}(a) - g^{h_1+1}(a) = \pi_0^{h_1+1} [g^{h_1}(a) - f(h_1, a)]. \quad (12)$$

Доказательство. Используя (2) и (4), находим:

$$\begin{aligned}
 g^{h_1}(a) - g^{h_1+1}(a) &= \sum_{i=0}^{h_2-1} \pi_i^{h_1} d_i^{h_1} + \pi_{h_2}^{h_1} d_{h_2}^{h_1} - \\
 - \sum_{i=0}^{h_2} \pi_i^{h_1+1} d_i^{h_1+1} - \pi_{h_2+1}^{h_1+1} d_{h_2+1}^{h_1+1} &= \sum_{i=0}^{h_2-1} \pi_i^{h_1} d_i^{h_1} + \pi_{h_2}^{h_1} d_{h_2}^{h_1} - \\
 - (1 - \pi_0^{h_1+1}) \left(\sum_{i=0}^{h_2-1} \pi_i^{h_1} d_{i+1}^{h_1+1} + \pi_{h_2}^{h_1} d_{h_2+1}^{h_1+1} \right) - \\
 - \pi_0^{h_1+1} d_0^{h_1+1} &= \sum_{i=0}^{h_2-1} \pi_i^{h_1} d_i^{h_1} + \pi_{h_2}^{h_1} d_{h_2}^{h_1} - (1 - \pi_0^{h_1+1}) \times \\
 &\times \left(\sum_{i=0}^{h_2-1} \pi_i^{h_1} (d_i^{h_1} + d_{i+1}^{h_1+1} - d_i^{h_1}) + \right. \\
 &\left. + \pi_{h_2}^{h_1} (d_{h_2}^{h_1} + d_{h_2+1}^{h_1+1} - d_{h_2-1}^{h_1}) \right) - \pi_0^{h_1+1} d_0^{h_1+1} = \\
 = \pi_0^{h_1+1} \left\{ g^{h_1} - \frac{1 - \pi_0^{h_1+1}}{\pi_0^{h_1+1}} \left[\sum_{i=0}^{h_2-1} \pi_i^{h_1} (d_{i+1}^{h_1+1} - d_i^{h_1}) + \right. \right. \\
 &\quad \left. \left. + \pi_{h_2}^{h_1} (d_{h_2+1}^{h_1+1} - d_{h_2-1}^{h_1}) \right] - d_0^{h_1+1} \right\}.
 \end{aligned}$$

Подставив в правую часть последнего равенства вместо $(d_{i+1}^{h_1+1} - d_i^{h_1})$ их выражения из формулы (9), после несложных преобразований получим

$$\begin{aligned}
 g^{h_1}(a) - g^{h_1+1}(a) &= \pi_0^{h_1+1} \left\{ g^{h_1} + C_1 \frac{1 - \pi_0^{h_1+1}}{\pi_0^{h_1+1}} \times \right. \\
 &\times \left[\sum_{i=0}^{h_1-2} \pi_i^{h_1} \left[\bar{v} - \frac{1}{\mu_i^{h_1}} \sum_{m=i+2}^{\infty} (m-i-1)r_{i,m}^{h_1} \right] - \right. \\
 &- \sum_{i=h_1-1}^{h_2-1} \pi_i^{h_1} \left[\sum_{m=i+1}^{\infty} \int_0^{\infty} \left[v - \int_0^v \left[t + \frac{(h_1-1)(v-t)}{m-i+h_1-1} \right] \times \right. \right. \\
 &\quad \left. \left. \times z_i^{h_1}(v, t, m) dt \right] dU(v) \right] - \pi_{h_2}^{h_1} \left[\sum_{m=h_2}^{\infty} \int_0^{\infty} \left[v - \right. \right. \\
 &\quad \left. \left. - \int_0^v \left[t + \frac{(h_1-1)(v-t)}{m-i+h_1-1} \right] z_i^{h_1}(v, t, m) dt \right] dU(v) \right] + \\
 &+ \sum_{i=h_1-1}^{h_2-1} \pi_i^{h_1} \int_0^{\infty} v \left[\sum_{m=0}^{i-h_1+2} r_{i,m}^{h_1}(v) + \right. \\
 &+ \sum_{m=i-h_1+3}^i \int_0^v z_i^{h_1}(v, t, m) dt + \\
 &+ \sum_{m=i+1}^{\infty} \int_0^v z_i^{h_1}(v, t, m) dt \left. \right] dU(v) + \\
 &+ \pi_{h_2}^{h_1} \int_0^{\infty} v \left[\sum_{m=0}^{h_2-h_1+2} r_{i,m}^{h_1}(v) + \right. \\
 &+ \sum_{m=h_2-h_1+3}^{h_2-1} \int_0^v z_i^{h_1}(v, t, m) dt + \\
 &+ \sum_{m=h_2}^{\infty} \int_0^v z_i^{h_1}(v, t, m) dt \left. \right] dU(v) - d_0^{h_1+1} \left. \right\} = \\
 &= \pi_0^{h_1+1} \left\{ g^{h_1} + C_1 \frac{1 - \pi_0^{h_1+1}}{\pi_0^{h_1+1}} \times \right. \\
 &\times \left[\sum_{i=0}^{h_1-2} \pi_i^{h_1} \left[\bar{v} - \frac{1}{\mu_i^{h_1}} \sum_{m=i+2}^{\infty} (m-i-1)r_{i,m}^{h_1} \right] + \right. \\
 &+ \sum_{i=h_1-1}^{h_2-1} \pi_i^{h_1} \left[\bar{v} - \sum_{m=i+1}^{\infty} \int_0^{\infty} \int_0^v \left[v - t - \frac{(h_1-1)(v-t)}{m-i+h_1-1} \right] \times \right. \\
 &\quad \left. \times z_i^{h_1}(v, t, m) dt dU(v) \right] + \\
 &+ \pi_{h_2}^{h_1} \left[\bar{v} - \sum_{m=h_2}^{\infty} \int_0^{\infty} \int_0^v \left[v - t - \frac{(h_1-1)(v-t)}{m-i+h_1-1} \right] \times \right. \\
 &\quad \left. \times z_i^{h_1}(v, t, m) dt dU(v) \right] \left. \right] - d_0^{h_1+1} \left. \right\}.
 \end{aligned}$$

В обозначениях (10) и (11) полученное выше для разности $(g^{h_1}(a) - g^{h_1+1}(a))$ последнее равенство после замены $d_0^{h_1+1} = C_0 - C_3\bar{v}$ примет вид (12).

Лемма 4. Функция $f(h_1, a)$ не возрастает по $h_1 > 0$.

Доказательство. Как следует из (8) и (10), $w(h_1, a)$ — среднее время простоя основного прибора. Функция $w(h_1, a)$ и вероятность $\pi_0^{h_1}$ убывают по h_1 (доказательство аналогично их доказательству в лемме 4 работы [6]). Тогда из (11) следует доказательство леммы 4.

Пусть h_1^* — решение задачи (1), $f(h_1) = \lim_{a \rightarrow \infty} f(h_1, a)$ и $g^{h_1} = \lim_{a \rightarrow \infty} g^{h_1}(a)$ (существование пределов следует из ограниченности функций). Справедлива следующая теорема.

Теорема 1. Справедливы утверждения:

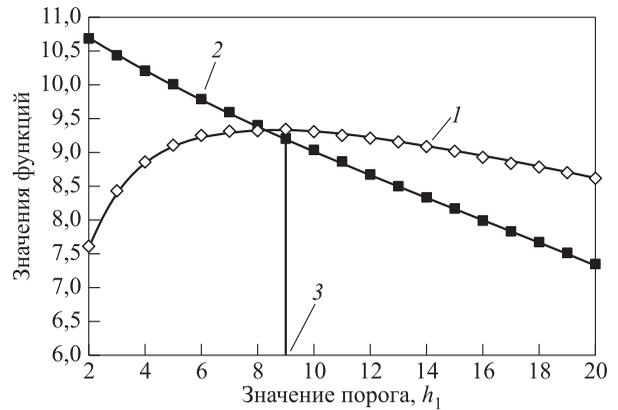
- (1) существует порог $h_1^* < \infty$, если $\inf_{h_1 > 1} f(h_1) < \sup_{h_1 > 1} g^{h_1}$ и $C_1 > 0$;
- (2) $h_1^* = \infty$, если $\inf_{h_1 > 1} f(h_1) \geq \sup_{h_1 > 1} g^{h_1}$ и $C_1 > 0$ или $g^{h_1} > f(h_1)$ при $h_1 = 2$ и $C_1 = 0$;
- (3) $h_1^* = 2$, если $g^{h_1} \leq f(h_1)$ при $h_1 = 2$;
- (4) для существования $1 < h_1^* < \infty$ необходимо и достаточно выполнение условий $g^{h_1^*-1} < g^{h_1^*}$ и $g^{h_1^*+1} \leq g^{h_1^*}$.

Доказательство. Доказательство теоремы следует из справедливости утверждений:

- (1) существует порог $h_1^* < \infty$, если $\inf_{h_1 > 1} f(h_1, a) < \sup_{h_1 > 1} g^{h_1}(a)$ и $C_1 > 0$;
- (2) $h_1^* = \infty$, если $\inf_{h_1 > 1} f(h_1, a) \geq \sup_{h_1 > 1} g^{h_1}(a)$ и $C_1 > 0$ или $g^{h_1}(a) > f(h_1, a)$ при $h_1 = 2$ и $C_1 = 0$;
- (3) $h_1^* = 2$, если $g^{h_1}(a) \leq f(h_1, a)$ при $h_1 = 2$;
- (4) для существования $1 < h_1^* < \infty$ необходимо и достаточно выполнение условий $g^{h_1^*-1}(a) < g^{h_1^*}(a)$ и $g^{h_1^*+1}(a) \leq g^{h_1^*}(a)$.

Доказательство последних утверждений для краткости изложения не приводим, их доказательство полностью следует из лемм 2–4 и теоремы из работы [10] и совпадает с доказательством аналогичных утверждений теоремы 1 в [6].

На рисунке проиллюстрировано поведение функций $g^{h_1}(a)$, $f(h_1, a)$ при изменении значения порога h_1 . Графики вычислены для следующих исходных данных: $U(t) = f_1(1 - e^{-\lambda_1 t}) + f_2(1 - e^{-\lambda_2 t})$, $f_i > 0$, $\lambda_i > 0$, $i = 1, 2$; $f_1 + f_2 = 1$; $f_1 = 0,3$; $f_2 = 0,7$; $\lambda_1 = 2$; $\lambda_2 = 3$; $a = 20$; $C_0 = 5$; $C_1 = 0,1$; $C_2 = 10$; $C_3 = 2$; $C_4 = 0$; $\mu_1 = 2,5$; $\mu_2 = 1,25$.



Зависимости функций $g^{h_1}(a)/\bar{v}$ (1) и $f(h_1, a)/\bar{v}$ (2) от порогового значения h_1 ; 3 — оптимальное значение h_1^*

4 Заключение

Отметим следующие основные выводы, вытекающие из полученных выше результатов:

- целевая функция (2) при $a < \infty$ для любых $0 < \mu_1 + \mu_2 < \infty$ и при $a = \infty$ для $0 < (\mu_1 + \mu_2)/\bar{v} < 1$ унимодальна по h_1 ;
- для существования конечного оптимального порога подключения резервного прибора необходимо и достаточно, чтобы выполнялись условия: либо $\inf_{h_1 > 1} f(h_1) < \sup_{h_1 > 1} g^{h_1}$ и $C_1 > 0$, либо $g^2 \leq f(2)$;
- если $\inf_{h_1 > 1} f(h_1) \geq \sup_{h_1 > 1} g^{h_1}$ и $C_1 > 0$ или $g^2 > f(2)$ и $C_1 = 0$, то подключение резервного прибора уменьшит предельный доход СМО при любом пороге;
- для поиска оптимального конечного порога h_1^* достаточно найти порог h_1 , при котором выполняются условия $g^{h_1-1} < g^{h_1}$ и $g^{h_1+1} \leq g^{h_1}$;
- на практике при устойчивой работе системы для оптимизации режима использования резервного прибора следует применить следующее простое правило: если при увеличении порога среднее значение дохода увеличивается (или при уменьшении порога уменьшается доход), то значение порога необходимо увеличить, а в противном случае значение порога необходимо уменьшить.

Результаты данной работы могут быть использованы на этапах проектирования и эксплуатации с целью повышения экономической эффективности реальных систем, для которых в качестве модели применима СМО типа $G/M/1$ ($G/M/1/r$) с резервированием.

Литература

1. Горцев А. М. Система массового обслуживания с произвольным числом резервных каналов и гистерезисным управлением включением и выключением резервных каналов // Автоматика и телемеханика, 1977. Вып. 10. С. 30–37.
2. Дудин А. Н. О задаче оптимального управления многоскоростной системой массового обслуживания // Автоматика и телемеханика, 1980. Вып. 9. С. 43–51.
3. Рыков В. В. Об условиях монотонности оптимальных политик управления системами массового обслуживания // Автоматика и телемеханика, 1999. Вып. 9. С. 92–106.
4. Самочернова Е. С., Петров Л. И. Оптимизация системы массового обслуживания с однотипным резервным прибором // Известия Томского политехнического университета, 2010. Т. 317. № 5. С. 28–31.
5. Клименок В. И. Многолинейная система массового обслуживания с резервными приборами // Ж. Белорусского государственного университета. Математика. Информатика, 2019. № 3. С. 57–70.
6. Агаларов Я. М. Оптимизация порогового управления переключением скорости обслуживания в системе массового обслуживания $G/M/1$ // Информатика и её применения, 2022. Т. 16. Вып. 1. С. 73–81.
7. Агаларов Я. М., Ушаков В. Г. Об унимодальности функции дохода системы массового обслуживания типа $G/M/s$ с управляемой очередью // Информатика и её применения, 2019. Т. 13. Вып. 1. С. 55–61.
8. Агаларов Я. М., Коновалов М. Г. Доказательство унимодальности целевой функции в задаче порогового управления нагрузкой на сервер // Информатика и её применения, 2019. Т. 13. Вып. 2. С. 2–6.
9. Карлин С. Основы теории случайных процессов / Пер. с англ. — М.: Мир, 1971. 537 с. (*Karlin S. A first course in stochastic processes. — New York; London: Academic Press, 1968. 502 p.*)
10. Агаларов Я. М. Признак унимодальности целочисленной функции одной переменной // Обозрение прикладной и промышленной математики, 2019. Т. 26. Вып. 1. С. 65–66.

Поступила в редакцию 10.02.22

OPTIMAL CONTROL OF A QUEUE-LENGTH DEPENDENT ADDITIONAL SERVER IN $GI/M/1$ QUEUE

Ya. M. Agalarov

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

Abstract: Consideration is given to a $GI/M/1$ queue in which there is an additional server available for serving customers from the queue. The additional server can be turned on and off depending on the current queue length. The long-run total cost per unit time, equal to the difference between the paid amount for service and the losses due to the waiting of customers and additional server depreciation, is being optimized. The case of finite queue capacity is also considered in which the losses also account for lost customers. It is proved that the cost function considered is unimodal. Necessary and sufficient conditions are given for the existence of the decision point (queue length) at which application of the additional server is optimal. A simple algorithm for controlling the decision point, requiring only observations of the cost function value, is provided.

Keywords: queuing system; redundancy; management; optimization

DOI: 10.14357/19922264220406

Acknowledgments

The reported study was partly funded by RFBR, project number 20-07-00804.

References

1. Gortsev, A. M. 1978. A queueing system with an arbitrary number of stand-by channels and hysteresis control of their connection and disconnection. *Automat. Rem. Contr.* 38(10):1451–1457.
2. Dudin, A. N. 1981. On optimal control of a multi-rate service system. *Automat. Rem. Contr.* 41(9):1221–1228.
3. Rykov, V. V. 1999. On monotonicity conditions for optimal policies for the control of queueing systems. *Automat. Rem. Contr.* 60(9):1290–1301.
4. Samochernova, E. S., and L. I. Petrov. 2010. Optimizatsiya sistemy massovogo obsluzhivaniya s odnotipnym rezervnym priborom [Optimization of the queueing system with the same type of backup device]. *Bulletin Tomsk Polytechnic University* 317(5):28–31.

5. Klimenok, V. I. 2019. Mnogolineynaya sistema massovogo obsluzhivaniya s rezervnymi priborami [Multi-server queueing system with reserve servers]. *J. Belarusian State University. Mathematics Informatics* 3:57–70.
6. Agalarov, Ya. M. 2022. Optimizatsiya porogovogo upravleniya pereklyucheniem skorosti obsluzhivaniya v sisteme massovogo obsluzhivaniya $G/M/1$ [Optimization of the threshold service speed control in the $G/M/1$ queue]. *Informatika i ee primeneniya — Inform. Appl.* 16(1):73–81.
7. Agalarov, Ya. M., and V. G. Ushakov. 2019. Ob unimodal'nosti funktsii dokhoda sistemy massovogo obsluzhivaniya tipa $G/M/s$ s upravlyаемoy ochered'yu [On the unimodality of the income function of a type $G/M/s$ queueing system with controlled queue]. *Informatika i ee primeneniya — Inform. Appl.* 13(1):55–61.
8. Agalarov, Ya. M., and M. G. Kononov. 2019. Dokazatel'stvo unimodal'nosti tselevoy funktsii v zadache porogovogo upravleniya nagruzkoy na server [Proof of the unimodality of the objective function in $M/M/N$ queue with threshold-based congestion control]. *Informatika i ee primeneniya — Inform. Appl.* 13(2): 2–6.
9. Karlin, S. 1968. *A first course in stochastic processes*. New York, NY: Academic Press. 502 p.
10. Agalarov, Ya. M. 2019. Priznak unimodal'nosti tselochislennoy funktsii odnoy peremennoy [A sign of unimodality of an integer function of one variable]. *Obozrenie prikladnoy i promyshlennoy matematiki* [Surveys Applied Industrial Mathematics] 26(1):65–66.

Received February 10, 2022

Contributor

Agalarov Yaver M. (b. 1952) — Candidate of Science (PhD) in technology, associate 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; agglar@yandex.ru

ОБ ОПТИМАЛЬНОМ РАСПОЛОЖЕНИИ АНТЕНН ДЛЯ V2X-СОЕДИНЕНИЙ В СУБТЕРАГЕРЦЕВОМ ДИАПАЗОНЕ*

Е. А. Мачнев¹, В. А. Бесчастный², Д. Ю. Острикова³, Ю. В. Гайдамака⁴, С. Я. Шоргин⁵

Аннотация: Субтерагерцевая (суб-ТГц, 100–300 ГГц) связь должна обеспечить огромную скорость передачи данных в системах 6G. Однако зона покрытия базовых станций (БС) ограничена, так как сигнал существенно затухает с увеличением дистанции, а также легко блокируется различными объектами, встречающимися на пути распространения сигнала. Таким образом, БС необходимо располагать достаточно часто, что делает такое решение дорогостоящим. Для снижения плотности развертывания БС можно использовать механизм ретрансляции сигнала с помощью транспортных средств (ТС). Данный способ в большой степени зависит от зоны размещения приемо-передающей антенны на кузове ТС, что ставит вопрос о поиске расположения антенны, при котором механизм ретрансляции будет эффективным с точки зрения скорости передачи данных и расстояния между ТС-источником и БС. В данной работе на основе спецификации IEEE 802.15.3d и экспериментальных данных о распространении сигнала на частоте 300 ГГц предложена математическая модель для анализа многозвеневой системы ретрансляции сигнала для трех зон размещения антенны на кузове ТС. Полученные результаты показывают, что расположение передатчика в зоне лобового стекла характеризуется более низкой скоростью передачи данных, но при этом гораздо большим покрытием, чем расположение в зонах бампера и двигателя.

Ключевые слова: 5G; «новое радио»; V2V; V2X; ретрансляция

DOI: 10.14357/19922264220407

1 Введение

Системы 5G New Radio (NR), работающие в микроволновом (μ Wave) и миллиметровом (mmWave) диапазонах, уже вышли на рынок. В то же время исследователи приступили к изучению субтерагерцевых (100–300 ГГц) диапазонов в контексте систем 6G [1, 2]. Однако чрезвычайно высокое затухание сигнала, эффекты динамической блокировки [3], а также микромобильность [4, 5] ограничивают дальность действия таких систем несколькими сотнями метров.

Одна из серьезных проблем для сотовых операторов — обеспечение расширенного мобильного широкополосного доступа (Enhanced mobile broadband, eMBB) для пользователей в движущихся ТС. Эта услуга требует не только постоянного подключения, но и высокой скорости передачи данных. Одним из решений могут стать высокопроизводительные системы 5G NR, работающие в диапазоне миллиметровых волн, или 6G в субтерагерцевом диапазоне. Однако высокое затухание в субтерагерцевом диапазоне и другие особенности

распространения сигнала, такие как блокировка прямой видимости, ограничивают зону покрытия расположенных по краям дороги БС несколькими сотнями метров, что требует высокой плотности развертывания, а значит, ставит вопрос о рентабельности системы.

Для обеспечения постоянной связи и снижения финансовых затрат сетевых операторов можно использовать механизм ретрансляции сигнала [6]. Чтобы обеспечить поддержку этого механизма, консорциум 3GPP недавно стандартизировал технологию интегрированного доступа и транспортной сети (Integrated Access and Backhaul, IAB) [7]. В рамках этой технологии ТС-ретрансляторы с высокоскоростными передатчиками в диапазоне субтерагерцевых частот могут организовать так называемые «мосты», пересылая по цепочке данные от ТС, которые в настоящее время не имеют прямого подключения к БС. Эффективность механизма многозвеневой ретрансляции сигнала оценивается на основании скорости передачи данных по установленному мосту и длины моста, т. е. расстояния между ТС-источником сигнала и БС. При этом рас-

* Исследование выполнено за счет гранта Российского научного фонда № 22-29-00694.

¹ Российский университет дружбы народов, 1042200071@pfur.ru

² Российский университет дружбы народов, beschastnyy-va@rudn.ru

³ Российский университет дружбы народов, ostrikova-dyu@rudn.ru

⁴ Российский университет дружбы народов; Федеральный исследовательский центр «Информатика и управление» Российской академии наук, gaydamaka-yuv@rudn.ru

⁵ Федеральный исследовательский центр «Информатика и управление» Российской академии наук, sshorgin@ipiran.ru

положение приемо-передающей антенны на кузове ТС влияет на оба указанных показателя.

Конечная цель данного исследования — выработать рекомендации по выбору зоны размещения антенны на ТС, при котором механизм ретрансляции обеспечит минимальную плотность развертывания БС для заданного набора параметров и условий дорожного движения, включая плотность ТС на дороге и скорости их движения. Разработанная математическая модель основана на недавних исследованиях распространения сигнала в субтерагерцевом диапазоне в среде V2V (Vehicle-to-Vehicle) и использует реалистичные параметры связи из стандарта IEEE 802.15.3d [8].

2 Системная модель

Рассматривается участок автомагистрали, например городская улица или скоростное шоссе, покрытие которого беспроводной связью обеспечивается БС 6G, работающими в субтерагерцевом диапазоне. Предполагается, что БС установлены по обеим сторонам дороги в шахматном порядке, например на фонарных столбах, на постоянной высоте h_A . Расстояние между БС по одной стороне дороги равно d , таким образом, БС образуют равнобедренные треугольники с основанием d , т. е. БС на противоположных сторонах сдвинуты друг относительно друга на расстояние $d/2$. Эти БС служат точками доступа в интернет для пользователей, находящихся в ТС, например в автомобилях (рис. 1).

Предполагается, что дорога имеет четное число N_l полос, при этом возможно движение ТС в противоположных направлениях. Ширина полос постоянна и равна w . Скорость ТС предполагается равной v . Расположение ТС на дорожной полосе

определяется пуассоновским процессом интенсивности λ , согласно которому расположены центры ТС. Далее предположим, что автомобили имеют одинаковую постоянную длину l_v , а дорожный просвет равен h_C . Дорожное движение предполагается однородным на каждой полосе, т. е. v , λ и l_v не зависят от рассматриваемой полосы.

Согласно [9], новые технологии V2V обеспечат эффективный и безопасный контроль над мобильностью ТС при минимально допустимом расстоянии между любыми двумя транспортными средствами $d_s = t_s v$, где $t_s = 0,5$ с — минимальное время, необходимое системам автоматического управления для оценки условий движения в режиме реального времени и принятия превентивных мер.

Для снижения капитальных затрат за счет увеличения расстояния между БС предполагается, что доля ТС P_E оснащена устройствами ретрансляции сигнала. Параметр P_E называется «степенью внедрения технологии». Транспортные средства оборудованы двумя приемо-передающими антеннами: одна в передней и одна в задней части ТС. Возможность подключения приемопередатчиков к высокоскоростной внутренней шине со скоростью передачи данных, достаточной для обработки ретранслируемого трафика, подтверждается последними разработками [10].

Транспортные средства, оснащенные средствами связи, нуждаются в услугах eMBB, предоставляемых через БС. Если прямое соединение невозможно из-за того, что ТС находится вне зоны действия ближайшей БС, они используют ретрансляцию для формирования моста, состоящего из одной или нескольких точек ретрансляции, в качестве которых выступают ТС-ретрансляторы. Длина этого моста и скорость передачи данных по нему существенно

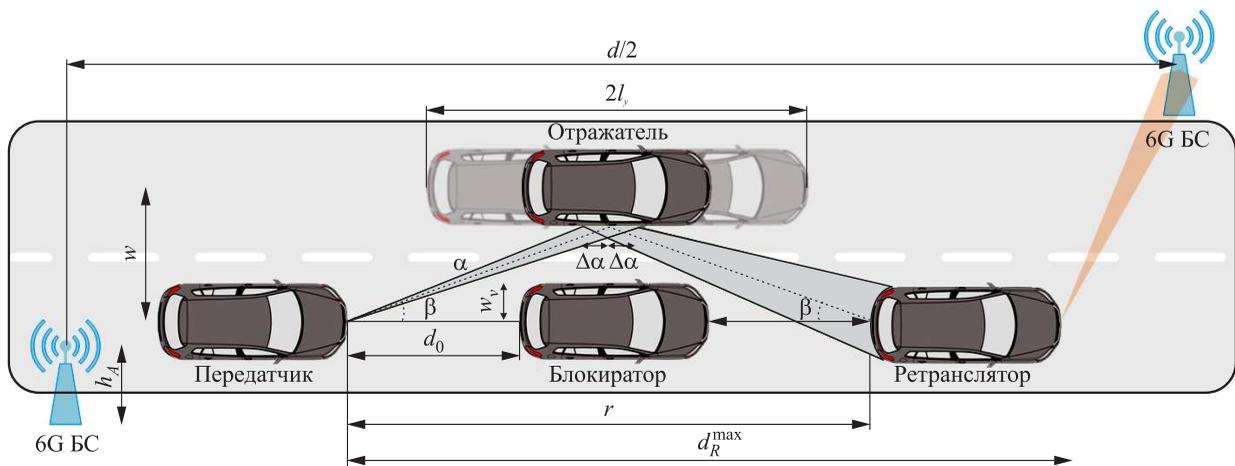


Рис. 1 Схема моста из одного ТС-ретранслятора с участием ТС-отражателя

зависят от плотности развертывания БС, характеристик окружающей среды, влияющих на распространение сигнала, включая плотность ТС на дороге и сценарий дорожного движения, а также зоны размещения антенны.

В работе рассмотрены следующие потенциальные варианты расположения антенны: на уровне бампера (0,3–0,4 м), на уровне двигателя (0,4–1,0 м) и на лобовом стекле (1,0–1,5 м). Определим параметр σ для обозначения зоны размещения антенны, т. е. $\sigma \in \{B, E, W\}$ соответственно.

Значение отношения уровня сигнала к уровню шума (Signal to Interference plus Noise Ratio, SINR) на ТС на расстоянии x от БС записывается в виде

$$S(x) = P_T G_A G_U \frac{x^{-\gamma}}{(N_0 + I) L_A(f, x) L_B},$$

где P_T — излучаемая мощность; G_A и G_U — коэффициенты усиления на стороне приема и передачи; γ — коэффициент потерь на пути сигнала; N_0 — тепловой шум; I — помехи; $L_A(f, x)$ — коэффициент затухания сигнала; L_B — потери, вызванные блокировкой или отражениями от других объектов.

3 Математическая модель установления соединения для механизма многозвеньевой ретрансляции сигнала

Построение модели начнем с анализа прямых подключений между ТС-источником сигнала и БС, а именно: с определения максимальной дистанции прямого подключения и скорости передачи данных. Чтобы вычислить данные параметры, используем набор схем модуляции и кодирования, указанных в стандарте IEEE 802.15.3d [8].

Следуя [11], максимальное расстояние подключения автомобиля-источника к БС с допустимой мощностью принимаемого сигнала S можно записать как

$$d_\xi(S) = \left(\frac{P_T \sqrt[10]{10^{G_A + G_U}}}{\sqrt[10]{10^{N_0 + S}} \sqrt[10]{10^{2 \log_{10} f_c - 14,86 + L_B}}} \right)^{1/\gamma}, \quad (1)$$

где P_T — мощность передачи; G_A и G_U — коэффициенты усиления на приеме и передаче; f_c — несущая частота.

Если следующий автомобиль не поддерживает ретрансляцию, расположение антенны на уровне бампера в дополнение к прямой видимости (L) может использовать пути отражения под ТС (U) и от

соседнего автомобиля (R). Если антенна расположена на уровне двигателя или лобовом стекле, отражение сигнала под автомобилем недоступно ввиду геометрических ограничений, при этом отражение от соседнего автомобиля доступно. Для расположения на стекле существует также возможность прохождения сигнала сквозь стекло следующего автомобиля (W), но затухание сигнала при этом будет в несколько раз выше затухания при прямой видимости. Для обозначения типов распространения сигнала используем параметр $\xi \in \{L, U, R, W\}$.

Следует обратить внимание, что минимальная требуемая мощность сигнала приема $S = S_{\min}$ позволяет получить максимальное расстояние для связи между соседними автомобилями, далее обозначаемое как d_ξ^{\max} , $\xi \in \{L, U, R, W\}$.

Функции плотности вероятности расстояния до i -го соседнего ТС [12] в случае распределения центров ТС согласно процессу Пуассона подчиняются распределению Эрланга [13]:

$$f_i(x) = \frac{2(\pi\lambda)^i}{(i-1)!} x^{2i-1} e^{-\pi\lambda x^2}, \quad x > 0, \quad i = 1, \dots, N.$$

Подключение в зоне прямой видимости. Связь в зоне прямой видимости на расстоянии r возможна для всех рассматриваемых зон размещения антенны, если расстояние между взаимодействующими ТС меньше, чем максимальное расстояние связи в условиях прямой видимости d^{\max} из (1), и следующее ТС оборудовано приемно-передающей антенной. Таким образом, вероятность успешного подключения в условиях прямой видимости на расстоянии r имеет вид:

$$f_{H,LoS}(r) = I(r < d^{\max}) P_E f_0(r),$$

где $f_0(r)$ вычисляется как

$$f_0(r) = 0,5f_1(r) + 0,5f_2(r).$$

Отражение от соседнего автомобиля. Напомним, что боковое отражение доступно для всех рассматриваемых зон размещения антенны. В случае блокировки прямой видимости, т. е. когда следующее ТС не оборудовано приемно-передающей антенной, предполагаем, что расстояние d_0 до блокирующего ТС (см. рис. 1) равновероятно подчиняется распределению расстояния либо до первого соседа, либо до второго. Соответственно, плотность вероятности расстояния с обходом блокирующего автомобиля определяется как свертка расстояний до двух ближайших соседей блокирующего ТС, т. е.

$$f_B(r) = (f_1 * f_2)(r) = \int_0^{\infty} f_1(s) f_2(r-s) ds.$$

В модели предполагается, что возможен обход не более одного блокирующего ТС через одно боковое отражение (см. рис. 1). При этом автомобиль-«отражатель» должен находиться на таком участке соседней полосы, где сигнал от «передатчика» не будет заблокирован автомобилем-«блокиратором», находящимся слишком близко либо к «передатчику», либо к «ретранслятору». Вероятность незаблокированного отражения может быть выражена как

$$\delta_R(r) = \frac{w_v}{\eta(r)} = \frac{w_v r}{2(w - w_v)},$$

где $\eta(r) = 2(w - w_v)/r$ — тангенс угла отклонения луча β ; w_v — половина ширины ТС.

Вероятность $P_R(r)$ наличия на соседней полосе ТС, обеспечивающего незаблокированное отражение сигнала, можно получить, используя свойство независимости процесса Пуассона. Предполагая, что точка центра отражателя на соседней полосе равномерно распределена на отрезке, соответствующем расстоянию между передатчиком и ретранслятором [14], искомую вероятность можно выразить в следующем виде:

$$P_R(r) = \begin{cases} \left(\frac{N_l}{2} - 1 \right) \frac{\delta_R(r)\lambda}{[l_v + 2\Delta_\alpha(r)]^{-1}}, & 2d_s < r < d_R^{\max}; \\ 0 & \text{в других случаях,} \end{cases} \quad (2)$$

где Δ_α — допустимое смещение соседнего ТС.

Отражение под автомобилем. В отличие от зон лобового стекла и двигателя, при размещении антенны в зоне бампера становится возможной передача сигнала под ТС за счет отражения от дорожного полотна. Из-за свойства симметричности отраженного пути (рис. 2) минимальное расстояние между бампером блокиратора и бампером с антенной с обеих сторон одинаково и определяется выражением

$$\delta_U(r) = \frac{r}{2} \left(1 - \frac{h_C}{h_B} \right) - \Delta_\alpha,$$

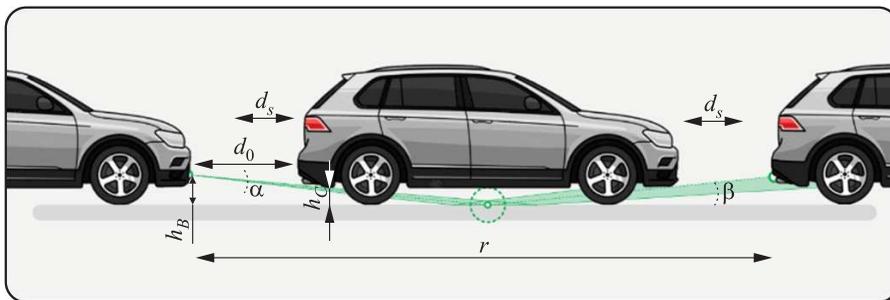


Рис. 2 Схема отражения под автомобилем

где h_C — дорожный просвет; h_B — высота размещения антенны; Δ_α — смещение, допускаемое для блокирующего ТС.

Вероятность передачи сигнала под ТС можно найти аналогично (2):

$$P_{B,B}(r) = \begin{cases} \delta_U(r) \left(\frac{h_C}{h_B} + \frac{\Delta_\alpha}{r} \right), & 2d_s < r < d_U^{\max}; \\ 0 & \text{в других случаях.} \end{cases} \quad (3)$$

Прохождение сигнала сквозь стекло. Наконец, найдем вероятность установления соединения через стекло блокирующего ТС:

$$P_{B,W}(r) = \begin{cases} 1, & 2d_s < r < d_W^{\max}; \\ 0 & \text{в других случаях.} \end{cases}$$

Вероятность успешного многозвеньевое подключение. С учетом вышеизложенных возможностей установления соединения для механизма многозвеньевой ретрансляции сигнала вероятность успешного подключения на расстоянии r , $f_{H,\sigma}(r)$, $\sigma \in \{B, E, W\}$, определяется как сумма вероятностей двух событий: либо следующее ТС оснащено антенной и можно использовать подключение в условиях прямой видимости, либо следующее ТС на полосе не имеет передающего устройства с вероятностью $1 - P_E$ и единственный вариант — обойти его с помощью описанных выше способов. Таким образом, расширяя (3), приходим к следующей формуле для плотности вероятности успешного подключения на расстоянии r :

$$f_{H,\sigma}(r) = I(r < d^{\max}) P_E f_0(r) + (P_E - P_E^2) \times (P_R(r) + P_{B,\sigma}(r) - P_R(r)P_{B,\sigma}(r)) f_B(r). \quad (4)$$

Определив вероятность многозвеньевое подключения в (4), можно перейти к описанию показателей эффективности. В частности, вероятность того, что мост из n ТС-ретрансляторов обеспечит передачу данных на расстояние r , может быть получена с помощью n -кратной свертки (4):

$$p_{C,\sigma}(n, r) = \int_{r-d^{\max}}^{\infty} f_{H,\sigma}^{(n)}(y) dy.$$

Если допустить возможность бесконечного числа ТС-ретрансляторов в соединении, вероятность того, что длина моста окажется не меньше расстояния r от ТС-источника сигнала до БС, может быть получена следующим образом:

$$p_{S,\sigma}(r) = \sum_{n=1}^{\infty} p_{C,\sigma}(n, r) \prod_{j=1}^{n-1} (1 - p_{C,\sigma}(j, r)), \quad \sigma \in \{B, E, W\}.$$

Скорость передачи по установленному мосту [14] определяется звеном с наихудшими условиями канала. Тогда, имея пороговые значения s_i , $i = 1, \dots, N_C$, чувствительности приемника, соответствующие набору показателей качества канала $\{1, \dots, N_C\}$, можно найти вероятность использования схемы кодирования i в звене для каждого типа распространения сигнала $\xi \in \{L, U, R, W\}$:

$$q_{\xi,i} = \int_{d_{\xi}(s_i)}^{d_{\xi}(s_{i+1})} f_{H,\xi}(r) dr, \quad i = 1, \dots, N_C,$$

где $d_{\xi}(S_{N_C+1}) = \infty$.

Теперь определим среднюю скорость передачи данных по установленному мосту, включающему n ТС-ретрансляторов, используя биномиальное распределение

$$\rho_{\xi,n} = \sum_{i=1}^{N_C} \omega_i \left(\sum_{k=i}^{N_C} q_{\xi,k} \right)^n, \quad \xi \in \{L, U, R, W\},$$

где ω_i — спектральная эффективность канала согласно схеме i .

Тем не менее в целях контроля качества соединения имеет смысл ограничивать максимальное число ТС-ретрансляторов некоторым заданным значением N , $N \geq 1$:

$$\begin{aligned} \rho_{S,\sigma}(r) &= \\ &= \sum_{\xi \in \{L, R, U, W\}} P_{\xi} \sum_{n=1}^N \rho_{\xi,n} p_{C,\sigma}(n, r) \prod_{j=1}^n (1 - p_{C,\sigma}(j, r)), \quad \sigma \in \{B, E, W\}. \end{aligned}$$

4 Численный анализ

В качестве исходных данных для численного эксперимента рассматриваются три сценария дорожного движения: пробка, нормальные условия движения в городе и скоростное шоссе. Сценарии различаются скоростью ТС и средним расстоянием между ними, как показано в табл. 1. Остальные параметры системы приведены в табл. 2.

Начнем с исследования основных зависимостей между средней длиной моста и скоростью передачи данных, показанных на рис. 3, для различных зон размещения антенны в зависимости от плотности ТС на полосе дорожного движения, где максимальное число ТС-ретрансляторов $N_{\text{установлено}}$ равно 10, а степень внедрения технологии $P_E = 0,7$.

Таблица 1 Параметры дорожного движения

Сценарий	Скорость автомобиля v , км/ч	Расстояние между автомобилями d_0 , м
Пробка	20	10
Нормальный городской трафик	60	30
Шоссе	120	60

Таблица 2 Входные параметры системы

Обозначение	Значение	Описание
l_v	4,5 м	Длина автомобиля
λ	0,02 авт/м	Средняя плотность автомобилей
h_A	3 м	Высота БС
h_B	0,4/0,7/1,2 м	Высота антенны
h_C	0,2 м	Дорожный просвет автомобиля
v	25 м/с	Скорость автомобиля
f_C	304,2 ГГц	Несущая частота
P_T	$4,2 \cdot 10^{-6}$ Вт	Мощность передатчика БС/автомобиля
N_0	-84 дБ	Мощность шума
S	-56 дБ	Минимальный SINR
G_A, G_U	17,58 дБ	Коэффициенты усиления на приеме и на передаче
γ	2,1	Коэффициент затухания сигнала

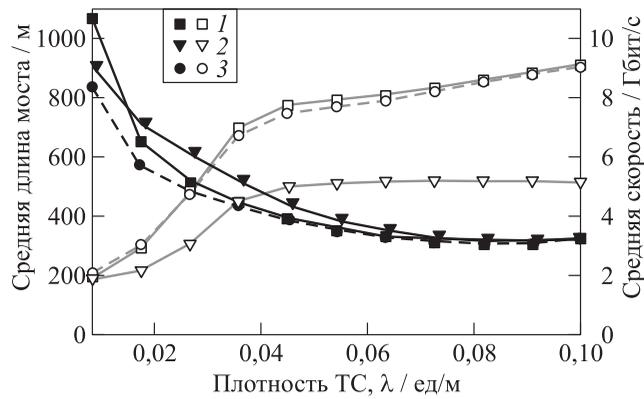


Рис. 3 Средняя длина моста (залитые значки) и скорость передачи (пустые значки) в зависимости от плотности ТС: 1 — бампер; 2 — стекло; 3 — двигатель

По результатам эксперимента можно отметить, что средняя длина моста уменьшается по мере увеличения плотности ТС на дороге для всех рассмотренных вариантов расположения антенны. Это объясняется тем, что с увеличением плотности уменьшается среднее расстояние между ТС. Обратный эффект наблюдается для скорости передачи данных, и это связано с более короткими расстояниями между автомобилями, а следовательно, лучшим качеством канала.

Анализируя влияние зоны размещения антенны, можно сделать вывод, что расположение у лобового стекла обеспечивает большую среднюю длину моста почти для всех рассмотренных плотностей ТС. Однако этот выигрыш достигается за счет гораздо меньшей скорости передачи данных. Обоснование данного наблюдения заключается в том, что экраны в виде заднего и переднего стекол блокирующего автомобиля создают высокие потери при передаче сигнала.

Одной из характеристик, отвечающих за гарантии производительности для пользователей коммерческих систем, служит доступность подключения к БС посредством моста. Рассмотрим доступность подключения как функцию от расстояния между БС (Inter-site distance, ISD), показанную на рис. 4, для различных значений степени внедрения технологии и нормальных условий городского трафика ($\lambda = 1/30$). При достижении 1250 м наблюдается резкий спад, характеризующий расстояние, на котором начинает работать механизм ретрансляции сигнала. Из всех рассмотренных зон размещения антенны лобовое стекло показывает наилучшую доступность 0,95 при степени внедрения технологии 0,9.

Влияние условий дорожного движения на вероятность подключения для различных зон размещения антенны показано на рис. 5. Здесь можно заметить, что наихудший возможный сценарий — пробки, где не только все рассмотренные зоны при-

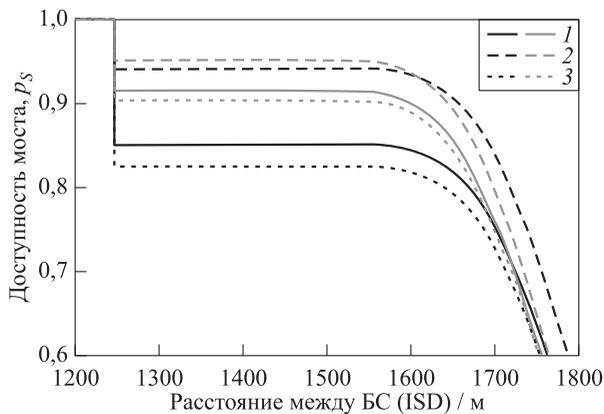


Рис. 4 Вероятность подключения в зависимости от расстояния между БС и степени внедрения технологии (черные кривые — $P_E = 0,7$; серые кривые — $P_E = 0,9$): 1 — бампер; 2 — стекло; 3 — двигатель

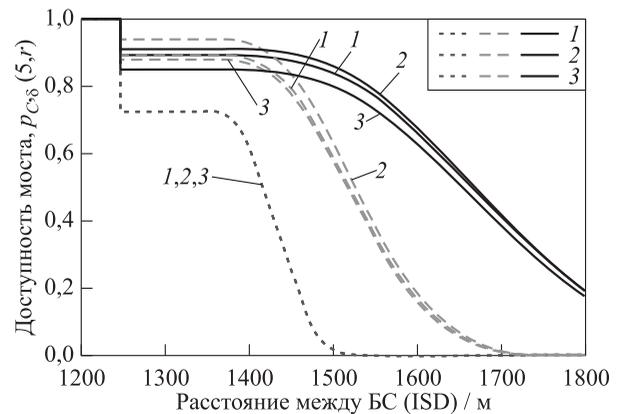


Рис. 5 Вероятность подключения в зависимости от расстояния между БС и сценариев дорожного движения (пунктирные кривые — пробка; штриховые — нормальный городской трафик; сплошные кривые — шоссе): 1 — бампер; 2 — стекло; 3 — двигатель

водят к одному и тому же значению параметра ISD, но и связанное с этим улучшение ISD незначительно. Причина в том, что ТС расположены очень плотно, блокируя сразу несколько вариантов отражений, в том числе под ТС. Тем не менее для нормального и шоссейного сценариев наилучшая зона размещения антенны с точки зрения длины установленного моста — на лобовом стекле.

5 Заключение

В работе предложена математическая модель для оценки производительности механизма многозвеньевой ретрансляции сигнала для V2X-соединений в частотных диапазонах субтерагерцевых волн при различных условиях распространения сигнала. Рассматриваемые метрики включают среднее расстояние подключения к БС и скорость передачи данных с учетом возможной многозвеньевой ретрансляции сигнала, а также критический параметр качества обслуживания — доступность сети.

Представленные численные результаты для типичных сценариев предоставления услуг связи по технологии 6G с использованием механизма многозвеньевой ретрансляции сигнала показывают, что можно рекомендовать размещение антенны в зоне лобового стекла, которое, несмотря на более низкую скорость передачи данных, значительно менее чувствительно к степени внедрения технологии и, как правило, характеризуется гораздо большим покрытием сети.

Литература

1. Moltchanov D., Gaidamaka Y., Ostrikova D., Beschastnyi V., Koucheryavy Y., Samouylov K. Ergodic outage and capacity of terahertz systems under micromobility and blockage impairments // IEEE T. Wirel. Commun., 2021. Vol. 21. Iss. 5. P. 3024–3039. doi: 10.1109/TWC.2021.3117583.
2. Moltchanov D., Beschastnyi V., Ostrikova D., Gaidamaka Y., Koucheryavy Y. Uninterrupted connectivity time in THz systems under user micromobility and blockage // Global Communications Conference Proceedings. — Piscataway, NJ, USA: IEEE, 2021. Art. 9685384. 6 p. doi: 10.1109/GLOBECOM46510.2021.9685384.
3. Gapeyenko M., Samuylov A., Gerasimenko M., Moltchanov D., Singh S., Akdeniz M. R., Aryafar E., Himayat N., Andreev S., Koucheryavy Y. On the temporal effects of mobile blockers in urban millimeter-wave cellular scenarios // IEEE T. Veh. Technol., 2017. Vol. 66. Iss. 11. P. 10124–10138. doi: 10.1109/TVT.2017.2754543.
4. Stepanov N. V., Moltchanov D., Begishev V., Turlikov A., Koucheryavy Y. Statistical analysis and modeling of user micromobility for THz cellular communications // IEEE T. Veh. Technol., 2021. Vol. 71. Iss. 1. P. 725–738. doi: 10.1109/TVT.2021.3124870.
5. Beschastnyi V., Ostrikova D., Moltchanov D., Gaidamaka Y., Koucheryavy Y., Samouylov K. Balancing latency and energy efficiency in mmWave 5G NR systems with multiconnectivity // IEEE Commun. Lett., 2022. Vol. 26. Iss. 8. P. 1952–1956. doi: 10.1109/LCOMM.2022.3175663.
6. Petrov V., Moltchanov D., Andreev S., Heath R. W. Analysis of intelligent vehicular relaying in urban 5G + millimeter-wave cellular deployments // Global Communications Conference Proceedings. — Piscataway, NJ, USA: IEEE, 2019. Art. 05946. 6 p. doi: 10.48550/arXiv.1908.05946.
7. Study on integrated access and backhaul (Release 17): Technical Specification 38.874 V17.0.0. 3GPP, 2020. https://www.3gpp.org/ftp/Specs/archive/38_series/38.874/38874-g00.zip.
8. Petrov V., Kurner T., Hosako I. IEEE 802.15.3d: First standardization efforts for sub-terahertz band communications toward 6G // IEEE Commun. Mag., 2020. Vol. 58. No. 11. P. 28–33. doi: 10.1109/MCOM.001.2000273.
9. Ozpolat M., Bhargava K., Kampert E., Higgins M. D. Multi-lane urban mmWave V2V networks: A path loss behavior dependent coverage analysis // Vehicular Communications, 2021. Vol. 30. Art. 100348. 11 p. doi: 10.1016/j.vehcom.2021.100348.
10. Wang J., Liu J., Kato N. Networking and communications in autonomous driving: A survey // IEEE Commun. Surv. Tut., 2018. Vol. 21. No. 2. P. 1243–1274. doi: 10.1109/COMST.2018.2888904.
11. Eckhardt J. M., Petrov V., Moltchanov D., Koucheryavy Y., Kurner T. Channel measurements and modeling for low-terahertz band vehicular communications // IEEE J. Sel. Area. Comm., 2021. Vol. 39. No. 6. P. 1590–1603. doi: 10.1109/JSAC.2021.3071843.
12. Moltchanov D. Distance distributions in random networks // Ad Hoc Netw., 2012. Vol. 10. P. 1146–1166. doi: 10.48550/arXiv.0804.4204.
13. Basharin G., Gaidamaka Y. V., Samouylov K. E. Mathematical theory of teletraffic and its application to the analysis of multiservice communication of next generation networks // Autom. Control Comp. S., 2013. Vol. 47. No. 2. P. 62–69. doi: 10.3103/S0146411613020028.
14. Kingman J. F. C. Poisson processes. — Oxford studies in probability ser. — Clarendon Press, 1993. 112 p. doi: 10.1002/0470011815.B2A07042.

Поступила в редакцию 15.10.22

ON THE OPTIMAL ANTENNA DEPLOYMENT FOR SUBTERAHERTZ V2X COMMUNICATIONS

E. A. Machnev¹, V. A. Beschastnyi¹, D. Yu. Ostrikova¹, 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: Subterahertz (sub-THz, 100–300 GHz) communication should provide huge data transfer rates in 6G systems. However, the coverage area of base stations (BS) will be very limited, since the signal is quite strongly attenuated from the distance and is also easily blocked by the presence of any objects in the signal path. Thus, the BS will need to be located too often which is a costly process. To reduce the deployment density of the BS, a mechanism was proposed for relaying the signal using vehicles (V2V). This relaying method is characterized by various options for the location of the antenna on vehicles which raises the question of finding the optimal location. In this work, guided by the IEEE 802.15.3d specification and measurements of the signal propagation level at a frequency of 300 GHz, the authors developed a mathematical model for comparing multihop signal relay systems with different antenna locations. The authors consider the following quality of service indicators: coverage, BS availability, and data transfer rate. The results show that the windshield transmitter location has a lower data rate but more coverage while the bumper and engine levels show similar performance. A windshield location is recommended as it is less sensitive to the rate of technology integration and has a larger coverage area.

Keywords: 5G; New Radio; V2V; V2X; multihop communications

DOI: 10.14357/19922264220407

Acknowledgments

The reported study was funded by the Russian Science Foundation, project number 22-29-00694 (<https://rscf.ru/en/project/22-29-00694>).

References

- Moltchanov, D., Y. Gaidamaka, D. Ostrikova, V. Beschastnyi, Y. Koucheryavy, and K. Samouylov. 2021. Ergodic outage and capacity of terahertz systems under micromobility and blockage impairments. *IEEE T. Wirel. Commun.* 21(5):3024–3039. doi: 10.1109/TWC.2021.3117583.
- Moltchanov, D., V. Beschastnyi, D. Ostrikova, Y. Gaidamaka, and Y. Koucheryavy. 2021. Uninterrupted connectivity time in THz systems under user micromobility and blockage. *Global Communications Conference Proceedings*. Piscataway, NJ: IEEE. 9685384. 6 p. doi: 10.1109/GLOBECOM46510.2021.9685384.
- Gapeyenko, M., A. Samuylov, M. Gerasimenko, D. Moltchanov, S. Singh, M. R. Akdeniz, E. Aryafar, N. Himayat, S. Andreev, and Y. Koucheryavy. 2017. On the temporal effects of mobile blockers in urban millimeter-wave cellular scenarios. *IEEE T. Veh. Technol.* 66(11):10124–10138. doi: 10.1109/TVT.2017.2754543.
- Stepanov, N. V., D. Moltchanov, V. Begishev, A. Turlikov, and Y. Koucheryavy. 2021. Statistical analysis and modeling of user micromobility for THz cellular communications. *IEEE T. Veh. Technol.* 71(1):725–738. doi: 10.1109/TVT.2021.3124870.
- Beschastnyi, V., D. Ostrikova, D. Moltchanov, Y. Gaidamaka, Y. Koucheryavy, and K. Samouylov. 2022. Balancing latency and energy efficiency in mmWave 5G NR systems with multiconnectivity. *IEEE Commun. Lett.* 26(8):1952–1956. doi: 10.1109/LCOMM.2022.3175663.
- Petrov, V., D. Moltchanov, S. Andreev, and R. W. Heath. 2019. Analysis of intelligent vehicular relaying in urban 5G + millimeter-wave cellular deployments. *Global Communications Conference Proceedings*. Piscataway, NJ: IEEE. 05946. 6 p. doi: 10.48550/arXiv.1908.05946.
- 3GPP. 2020. NR. Study on integrated access and backhaul (Release 17): Technical Specification 38.874 V17.0.0. Available at: https://www.3gpp.org/ftp/Specs/archive/38_series/38.874/38874-g00.zip (accessed November 28, 2022).
- Petrov, V., T. Kurner, and I. Hosako. 2020. IEEE 802.15.3d: First standardization efforts for sub-terahertz band communications toward 6G. *IEEE Commun. Mag.* 58(11):28–33. doi: 10.1109/MCOM.001.2000273
- Ozpolat, M., K. Bhargava, E. Kampert, and M. D. Higgins. 2021. Multi-lane urban mmwave V2V networks:

- A path loss behavior dependent coverage analysis. *Vehicular Communications* 30:100348. 11 p. doi: 10.1016/j.vehcom.2021.100348.
10. Wang, J., J. Liu, and N. Kato. 2018. Networking and communications in autonomous driving: A survey. *IEEE Commun. Surv. Tut.* 21(2):1243–1274. doi: 10.1109/COMST.2018.2888904.
 11. Eckhardt, J. M., V. Petrov, D. Moltchanov, Y. Koucheryavy, and T. Kurner. 2021. Channel measurements and modeling for low-terahertz band vehicular communications. *IEEE J. Sel. Area. Comm.* 39(6):1590–1603. doi: 10.1109/JSAC.2021.3071843.
 12. Moltchanov, D. 2012. Distance distributions in random networks. *Ad Hoc Netw.* 10(6):1146–1166. doi: 10.48550/arXiv.0804.4204.
 13. Basharin, G. P., Yu. V. Gaidamaka, and K. E. Samouylov. 2013. Mathematical theory of teletraffic and its application to the analysis of multiservice communication of next generation networks. *Autom. Control Comp. S.* 47(2):62–69. doi: 10.3103/S0146411613020028.
 14. Kingman, J. F. C. 1993. *Poisson processes*. Oxford studies in probability ser. Clarendon Press. 112 p. doi: 10.1002/0470011815.B2A07042.

Received October 15, 2022

Contributors

Machnev Egor 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; 1032143100@rudn.ru

Beschastnyi Vitalii A. (b. 1992) — Candidate of Science (PhD) in physics and mathematics, assistant professor, Department of Applied Probability and Informatics, Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; beschastnyy-va@rudn.ru

Ostrikova Daria Yu. (b. 1988) — 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; ostrikova-dyu@rudn.ru

Gaidamaka Yuliya V. (b. 1971) — Doctor of Science in physics and mathematics, professor, Department of Applied Probability and Informatics, Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; senior scientist, 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

НЕЧЕТКИЕ УСРЕДНЯЮЩИЕ ОПЕРАТОРЫ В ЗАДАЧЕ АГРЕГИРОВАНИЯ НЕЧЕТКОЙ ИНФОРМАЦИИ

В. Л. Хацкевич¹

Аннотация: Рассматривается задача агрегирования нечеткой информации посредством построения нечетких усредняющих операторов. Изучены взвешенные нечеткие средние систем нечетких чисел и введен класс нелинейных нечетких средних систем нечетких чисел, являющийся модификацией на нечеткие числа общего класса диссипативных числовых средних. Установлены свойства соответствующих усредняющих операторов, которые представляют собой «нечеткие» аналоги характерных свойств скалярных агрегирующих функций. Это дает обоснование использованию введенных нечетких усредняющих операторов в задаче агрегирования нечеткой информации. При этом под результатом агрегирования нечеткой информации, заданной набором нечетких чисел, понимается нечеткое число, отражающее существенные особенности этой совокупности.

Ключевые слова: усредняющие нечеткие операторы; агрегирование нечеткой информации

DOI: 10.14357/19922264220408

1 Введение

При разработке и анализе информационно-аналитических систем возникает задача построения обобщенной оценки состояния сложного объекта с помощью операции агрегирования, понимаемой как отображение, которое векторной оценке объекта $X = (x_1, \dots, x_n)$ ставит в соответствие скалярную величину $A(X) = A(x_1, \dots, x_n)$, характеризующую этот объект. Компоненты x_i векторной оценки называют частными оценками.

Обозначим через I отрезок расширенной числовой прямой. Под n -мерной агрегирующей функцией обычно понимают функцию $A : I^n \rightarrow I$, которая обладает следующими свойствами (см., например, [1–3]).

Свойство 1. Идемпотентность: $A(x, \dots, x) = x$.

Свойство 2. Монотонность: для любой пары векторных оценок $X = (x_1, \dots, x_n)$ и $Y = (y_1, \dots, y_n)$, таких что $x_i \leq y_i$, $i = 1, \dots, n$, выполняется неравенство $A(X) \leq A(Y)$.

Свойство 3. Непрерывность: функция $A(x_1, \dots, x_n)$ непрерывна.

Свойство 4. Граничные условия:

$$\inf_{X \in I^n} A(X) = \inf I; \quad \sup_{X \in I^n} A(X) = \sup I.$$

Перечисленные свойства отражают интуитивные требования к результату агрегирования. Идемпотентность отражает тот факт, что если все частные оценки равны между собой, то и обобщенная оценка должна быть такой же. Монотонность означает,

что если хотя бы одна частная оценка увеличится (уменьшится), то и обобщенная оценка увеличится (уменьшится). Согласно свойству непрерывности малым изменениям частных оценок соответствуют малые изменения обобщенных оценок.

Иногда в качестве I рассматривают нормированный отрезок $[0, 1]$. Отдельные классы агрегирующих функций обладают рядом дополнительных свойств.

Часто в качестве агрегирующих функций рассматриваются различные средние [4, 5], а именно:

– среднее арифметическое

$$M(x_1, \dots, x_n) = \frac{1}{n} \sum_{i=1}^n x_i;$$

– среднее геометрическое

$$G(x_1, \dots, x_n) = \left(\prod_{i=1}^n x_i \right)^{1/n};$$

– среднее гармоническое

$$H(x_1, \dots, x_n) = n \left(\sum_{i=1}^n \frac{1}{x_i} \right)^{-1}.$$

Отметим, что все приведенные средние обладают свойствами 1–4. Кроме того, среднее арифметическое $M(x_1, \dots, x_n)$ обладает свойствами аддитивности и однородности.

Также в литературе рассматриваются соответствующие взвешенные средние. В частности, порядковые операторы взвешенного агрегирования — OWA (ordered weighted averaging) и WOWA (weighted OWA) [6, 7].

¹ Военно-воздушная академия им. проф. Н. Е. Жуковского и Ю. А. Гагарина, vlkhats@mail.ru

С другой стороны, в последние десятилетия успешно развивается теория нечетких множеств, представляющая собой современный аппарат формализации различных видов неопределенностей (см., например, [8, 9]). Основоположником этого направления стал Л. Заде [10]. Для агрегирования нечеткой информации активно развивается подход, связанный с использованием так называемых треугольных норм и соответствующих им аддитивных генераторов, а также треугольных конорм (см., например, [8, гл. 3]). Кроме того, в теории нечеткого вывода широко распространен метод привлечения взвешенного среднего или, в более общей ситуации, нечеткого интеграла Шоке (см., например, [11–13]). Отметим, что оба этих подхода связаны с агрегированием результирующей функции принадлежности по заданным агрегируемым функциям принадлежности.

В предлагаемой работе рассматривается задача агрегирования нечеткой информации посредством построения нечетких усредняющих операторов. При этом под результатом агрегирования нечеткой информации, заданной набором нечетких чисел, понимается нечеткое число, отражающее существенные особенности этой совокупности.

Цель работы — установление свойств нечетких агрегирующих функций, порождаемых нечеткими усредняющими операторами, аналогичных свойствам скалярных агрегирующих функций. Это служит обоснованием возможности использования рассматриваемых нечетких усредняющих операторов для агрегирования нечеткой информации. Такой подход представляется новым.

Ниже \mathbb{R} — множество действительных чисел. Под нечетким числом \tilde{z} , заданным на универсальном пространстве \mathbb{R} , будем понимать совокупность упорядоченных пар $(\mu_{\tilde{z}}(x), x)$, где функция принадлежности $\mu_{\tilde{z}} : \mathbb{R} \rightarrow [0, 1]$ определяет степень принадлежности $\forall x \in \mathbb{R}$ множеству \tilde{z} . Носителем нечеткого числа \tilde{z} называют множество $x \in \mathbb{R}$, для которых $\mu_{\tilde{z}}(x) > 0$, и обозначают $\text{Supp}(\tilde{z})$.

Будем дополнительно предполагать выполнение следующих условий (ср. [9, гл. 2–4]):

- (а) носитель нечеткого числа — замкнутое и ограниченное (компактное) множество действительных чисел;
- (б) функция принадлежности нечеткого числа $\mu_{\tilde{z}}(x)$ выпукла;
- (в) функция принадлежности нечеткого числа $\mu_{\tilde{z}}(x)$ нормальна, т. е. $\sup_x \mu_{\tilde{z}}(x) = 1$;
- (г) функция принадлежности нечеткого числа $\mu_{\tilde{z}}(x)$ полунепрерывна сверху.

Совокупность таких нечетких чисел будем обозначать J .

Ниже будет использовано интервальное представление нечетких чисел. Как известно, интервал α -уровня нечеткого числа \tilde{z} с функцией принадлежности $\mu_{\tilde{z}}(x)$ определяется соотношением

$$Z_{\alpha} = \{x | \mu_{\tilde{z}}(x) \geq \alpha\}, \quad \alpha \in (0, 1], \quad Z_0 = \text{Supp}(\tilde{z}).$$

Согласно предположениям (а)–(г) на нечеткие числа все α -уровни нечеткого числа — замкнутые и ограниченные интервалы вещественной оси. Обозначим левую границу интервала через $z^{-}(\alpha)$, а правую — через $z^{+}(\alpha)$. Иногда $z^{-}(\alpha)$ и $z^{+}(\alpha)$ называют соответственно левым и правым индексами нечеткого числа.

Суммой нечетких чисел \tilde{z} и \tilde{u} с α -интервалами $[z^{-}(\alpha), z^{+}(\alpha)]$ и $[u^{-}(\alpha), u^{+}(\alpha)]$ называют нечеткое число с α -интервалами $[z^{-}(\alpha) + u^{-}(\alpha), z^{+}(\alpha) + u^{+}(\alpha)]$. Умножение на положительное число означает умножение индексов на это число, а умножение на отрицательное число означает умножение индексов на это число и перемену их местами.

Два нечетких числа равны, если совпадают все их соответствующие α -интервалы.

Рассмотрим совокупность J^n векторов с нечеткими компонентами (нечетких векторов) вида $\tilde{Z} = (\tilde{z}_1, \dots, \tilde{z}_n)$, где нечеткие числа \tilde{z}_i ($i = 1, \dots, n$) удовлетворяют условиям (а)–(г). Для двух нечетких векторов \tilde{Z} и \tilde{W} их сумму и умножение на числа будем понимать покоординатно.

Рассмотрим на множестве J нечетких чисел метрику, определяемую для $\tilde{z}, \tilde{u} \in J$ формулой [14]:

$$\rho(\tilde{z}, \tilde{u}) = \sup_{\alpha \in [0, 1]} \max(|z^{-}(\alpha) - u^{-}(\alpha)|, |z^{+}(\alpha) - u^{+}(\alpha)|),$$

где $[z^{-}(\alpha), z^{+}(\alpha)]$ и $[u^{-}(\alpha), u^{+}(\alpha)]$ — интервалы α -уровней чисел \tilde{z} и \tilde{u} .

На множестве нечетких векторов J^n для нечетких векторов \tilde{Z} и \tilde{W} с компонентами \tilde{z}_i и \tilde{w}_i ($i = 1, \dots, n$) зададим метрику выражением

$$\rho_n(\tilde{Z}, \tilde{W}) = \sum_{i=1}^n \rho(\tilde{z}_i, \tilde{w}_i).$$

Рассмотрим следующий критерий сравнения нечетких чисел, заданных в интервальной форме [15, гл. 4, 5]. Будем писать $\tilde{z} < \tilde{w}$ для нечетких чисел \tilde{z} и \tilde{w} , если одновременно

$$z^{-}(\alpha) \leq w^{-}(\alpha); \quad z^{+}(\alpha) \leq w^{+}(\alpha) \quad (\forall \alpha \in (0, 1]). \quad (1)$$

По-существу, это введение частичной упорядоченности на множестве нечетких чисел, т. е. введение бинарного отношения, обладающего свойствами рефлексивности, транзитивности и антисимметричности. Неравенства нечетких

векторов $\tilde{Z} \prec \tilde{W}$ будем понимать покомпонентно в смысле определения (1).

2 Нечеткие средние систем нечетких чисел и усредняющие операторы как агрегаторы нечеткой информации

Пусть заданы вещественные числа $\beta_i \in R$ ($i = 1, \dots, n$) такие, что $\beta_i \geq 0$, $\sum_{i=1}^n \beta_i = 1$.

Рассмотрим взвешенное нечеткое среднее нечетких чисел $\tilde{z}_1, \dots, \tilde{z}_n$ [16, гл. 7; 17]

$$\tilde{z}_{cp} = \sum_{i=1}^n \beta_i \tilde{z}_i. \quad (2)$$

Обозначим через $z_i^-(\alpha)$ и $z_i^+(\alpha)$ левые и правые индексы нечетких чисел \tilde{z}_i , фигурирующих в формуле (2).

Лемма 1. *Левый индекс нечеткого среднего \tilde{z}_{cp} , определяемого формулой (2), равен*

$$z_{cp}^-(\alpha) = \sum_{i=1}^n \beta_i z_i^-(\alpha),$$

а правый индекс —

$$z_{cp}^+(\alpha) = \sum_{i=1}^n \beta_i z_i^+(\alpha).$$

Она вытекает из определения интервального сложения нечетких чисел и умножения на положительное число.

Назовем оператор $A : J^n \rightarrow J$ нечетким агрегатором, если он обладает следующими свойствами:

- (1) если $\tilde{Z} = (\tilde{z}, \tilde{z}, \dots, \tilde{z})$, то $A(\tilde{Z}) = \tilde{z}$ (идемпотентность);
- (2) если $\tilde{Z} \prec \tilde{W}$, то $A(\tilde{Z}) \prec A(\tilde{W})$ (монотонность);
- (3) оператор $A : J^n \rightarrow J$ непрерывен.

Это модификация на случай нечетких чисел характерных свойств скалярного оператора агрегатора (см. свойства 1–3 из разд. 1).

Рассмотрим усредняющий оператор A_β , определяемый для заданного $\tilde{Z} \in J^n$, $\tilde{Z} = (\tilde{z}_1, \dots, \tilde{z}_n)$ взвешенной суммой (2):

$$A_\beta(\tilde{Z}) = \tilde{z}_{cp} = \sum_{i=1}^n \beta_i \tilde{z}_i. \quad (3)$$

Подчеркнем, что оператор $A_\beta(\tilde{Z})$ агрегирует информацию, заданную нечетким вектором $\tilde{Z} = (\tilde{z}_1, \dots, \tilde{z}_n)$ при весовых коэффициентах β_i ($i = 1, \dots, n$).

Теорема 1. *Усредняющий оператор A_β , определяемый формулой (3), является нечетким агрегатором.*

Доказательство. Проверим свойство 1 — идемпотентность. Пусть $\tilde{Z} = (\tilde{z}, \tilde{z}, \dots, \tilde{z})$, причем нечеткое число \tilde{z} имеет индексы $z^-(\alpha)$ и $z^+(\alpha)$. Тогда согласно лемме 1 \tilde{z}_{cp} имеет левый индекс $\sum_{i=1}^n \beta_i z^-(\alpha) = z^-(\alpha)$, поскольку $\sum_{i=1}^n \beta_i = 1$. И аналогично правый индекс совпадает с $z^+(\alpha)$. Это и доказывает свойство 1.

Покажем свойство 2 — монотонность. Пусть $\tilde{Z}, \tilde{W} \in J^n$ и $\tilde{Z} \prec \tilde{W}$. По определению это означает, что для компонент этих векторов имеем $\tilde{z}_i \prec \tilde{w}_i$ ($i = 1, \dots, n$), т.е. согласно (1) для любого $\alpha \in (0, 1]$ и $i = 1, \dots, n$ для соответствующих индексов имеем $z_i^-(\alpha) \leq w_i^-(\alpha)$ и аналогично $z_i^+(\alpha) \leq w_i^+(\alpha)$. Умножая обе части указанных неравенств на β_i и суммируя по $i = 1, \dots, n$, получаем:

$$\begin{aligned} \sum_{i=1}^n \beta_i z_i^-(\alpha) &\leq \sum_{i=1}^n \beta_i w_i^-(\alpha); \\ \sum_{i=1}^n \beta_i z_i^+(\alpha) &\leq \sum_{i=1}^n \beta_i w_i^+(\alpha). \end{aligned}$$

В этих неравенствах согласно лемме 1 слева стоят α -индексы нечеткого числа $A_\beta(\tilde{Z})$, а справа — $A_\beta(\tilde{W})$, что и влечет соотношение $A_\beta(\tilde{Z}) \prec A_\beta(\tilde{W})$.

Покажем свойство 3 — непрерывность. Пусть задано $\varepsilon > 0$ и для нечетких векторов $\tilde{Z}, \tilde{W} \in J^n$ выполнено неравенство $\rho_n(\tilde{Z}, \tilde{W}) \leq \varepsilon/(2n)$. По определению ρ_n это влечет совокупность неравенств $\rho(z_i, w_i) \leq \varepsilon/(2n)$, $i = 1, \dots, n$.

Далее, в силу (3) и согласно лемме 1 имеем:

$$\begin{aligned} \rho \left(A_\beta(\tilde{Z}), A_\beta(\tilde{W}) \right) &= \\ &= \sup_{\alpha \in [0,1]} \max \left\{ \left| \sum_{i=1}^n \beta_i (z_i^-(\alpha) - w_i^-(\alpha)) \right|, \right. \\ &\quad \left. \left| \sum_{i=1}^n \beta_i (z_i^+(\alpha) - w_i^+(\alpha)) \right| \right\}. \end{aligned}$$

При этом

$$\left| \sum_{i=1}^n \beta_i (z_i^-(\alpha) - w_i^-(\alpha)) \right| \leq \sum_{i=1}^n |z_i^-(\alpha) - w_i^-(\alpha)|.$$

Аналогично для индексов с плюсом. Так что окончательно имеем:

$$\rho \left(A_\beta(\tilde{Z}), A_\beta(\tilde{W}) \right) \leq 2 \sum_{i=1}^n \rho(z_i, w_i) \leq \varepsilon,$$

что и влечет указанную непрерывность.

Замечание 1. Граничные условия для агрегирующих функций (см. условие 4 разд. 1) выполняются для агрегирующей функции $A_\beta(\tilde{Z})$ в смысле их выполнения для левых и правых индексов результирующего нечеткого числа при каждом $\alpha \in [0, 1]$.

Свойство 2 влечет

Следствие 1. Если $\tilde{Z}, \tilde{W} \in J^n$ и $\tilde{Z} < \tilde{W}$, то $A_\beta(\tilde{Z} + \tilde{V}) < A_\beta(\tilde{W} + \tilde{V})$ для любых $\tilde{V} \in J^n$.

Кроме указанных свойств имеет место

Теорема 2. Усредняющий оператор $A_\beta : J^n \rightarrow J$, определяемый формулой (3), аддитивен и однороден.

Действительно, покажем аддитивность A_β . Пусть заданы нечеткие векторы $\tilde{Z}, \tilde{W} \in J^n$. В соответствии с леммой 1 левый и правый индексы $A_\beta(\tilde{Z})$ совпадают с $\sum_{i=1}^n \beta_i z_i^-(\alpha)$ и $\sum_{i=1}^n \beta_i z_i^+(\alpha)$ соответственно. Аналогично для $A_\beta(\tilde{W})$ левый и правый индексы имеют вид $\sum_{i=1}^n \beta_i w_i^-(\alpha)$ и $\sum_{i=1}^n \beta_i w_i^+(\alpha)$. Тогда левый индекс суммы $A_\beta(\tilde{Z}) + A_\beta(\tilde{W})$ имеет вид $\sum_{i=1}^n \beta_i (\tilde{z}^-(\alpha) + w_i^-(\alpha))$, а правый индекс — $\sum_{i=1}^n \beta_i (\tilde{z}^+(\alpha) + w_i^+(\alpha))$, что совпадает с левым и соответственно правым индексом нечеткого числа $A_\beta(\tilde{Z} + \tilde{W})$ и влечет аддитивность.

Однородность есть следствие формулы (3) и определения интервального умножения нечетких чисел на положительное и отрицательное число.

3 Нелинейные нечеткие усредняющие операторы как агрегаторы нечеткой информации

Определим понятие функции от нечеткого числа, используя интервальный подход. Пусть задана непрерывная монотонно возрастающая (монотонно убывающая) вещественная функция $\phi : R \rightarrow R$. Приведем в удобном виде формулировку результата из [18].

Лемма 2. Если \tilde{z} — нечеткое число с левым и правым индексами $z^-(\alpha)$ и $z^+(\alpha)$ и $\phi : R \rightarrow R$ — непрерывная монотонно возрастающая функция, то $\phi(z^-(\alpha))$ и $\phi(z^+(\alpha))$ суть соответственно левый и правый индексы нечеткого числа $\phi(\tilde{z})$. Если $\phi(x)$ — непрерывная монотонно убывающая функция, то $\phi(z^+(\alpha))$ и $\phi(z^-(\alpha))$ — левый и правый индексы $\phi(\tilde{z})$ соответственно.

Пусть заданы нечеткие числа $\tilde{z}_1, \dots, \tilde{z}_n$, а также действительные числа $\beta_i \in \mathbb{R}$ ($i = 1, \dots, n$), причем $\beta_i \geq 0$, $\sum_{i=1}^n \beta_i = 1$. Рассмотрим нелинейное нечеткое среднее общего вида для заданной непрерыв-

ной строго монотонно возрастающей (убывающей) функции $\phi : R \rightarrow R$:

$$\tilde{z}_\phi = \phi^{-1} \left(\sum_{i=1}^n \beta_i \phi(\tilde{z}_i) \right). \quad (4)$$

Определение (4) — аналог вещественного нелинейного ассоциативного среднего [19, гл. I]. Функцию ϕ в этом случае называют определяющей.

В случае определяющей функции $\phi_p(x) = x^p$ ($p > 1$) (или $0 < p < 1$) в (4) получаем аналог взвешенной средней степенной. Если $\phi_G(x) = \log_a x$ ($a > 1$) — это аналог взвешенной средней геометрической, если $\phi_H(x) = 1/x$ — аналог взвешенной средней гармонической.

Лемма 3. Нечеткое число, определяемое формулой (4), имеет левый индекс $\phi^{-1}(\sum_{i=1}^n \beta_i \phi(z_i^-(\alpha)))$ и правый индекс $\phi^{-1}(\sum_{i=1}^n \beta_i \phi(z_i^+(\alpha)))$, где $z_i^-(\alpha)$ и $z_i^+(\alpha)$ — левый и правый индексы нечеткого числа \tilde{z}_i соответственно.

Действительно, проведем рассуждения для левых индексов в предположении, что функция ϕ непрерывна и монотонно возрастает. Заметим, что нечеткое число $\sum_{i=1}^n \beta_i \phi(\tilde{z}_i)$ в силу монотонного возрастания функции ϕ на основании леммы 2 и по правилу интервального сложения нечетких чисел и умножения их на положительные числа имеет левый индекс $\sum_{i=1}^n \beta_i \phi(z_i^-(\alpha))$. Так как ϕ^{-1} также монотонно возрастающая функция (вместе с ϕ), то нечеткое число $\phi^{-1}(\sum_{i=1}^n \beta_i \phi(\tilde{z}_i)) = \tilde{z}_\phi$ имеет левый индекс $\phi^{-1}(\sum_{i=1}^n \beta_i \phi(z_i^-(\alpha)))$.

Аналогично для правых индексов.

Случай монотонного убывания непрерывной функции ϕ в силу леммы 2 приводит к такому же результату.

Определим теперь для фиксированного набора чисел $\beta_i \geq 0$ ($i = 1, \dots, n$) таких, что $\sum_{i=1}^n \beta_i = 1$, и заданной непрерывной строго монотонной функции $\phi : R \rightarrow R$ нечеткий нелинейный усредняющий оператор $F_{\beta, \phi} : J^n \rightarrow J$ равенством

$$F_{\beta, \phi}(\tilde{Z}) = \phi^{-1} \left(\sum_{i=1}^n \beta_i \phi(\tilde{z}_i) \right). \quad (5)$$

Теорема 3. Для нелинейного усредняющего оператора $F_{\beta, \phi}$ выполнены условия 1–3, т. е. он является нечетким агрегатором.

Доказательство. Пусть, например, функция ϕ непрерывна и строго монотонно возрастает. Проверим свойство 1 — идемпотентность. Пусть \tilde{Z} имеет одинаковые компоненты $\tilde{z}_i = \tilde{z}$. Тогда согласно (5)

$$F_{\beta, \phi}(\tilde{Z}) = \phi^{-1} \left(\sum_{i=1}^n \beta_i \phi(\tilde{z}) \right) = \phi^{-1}(\phi(\tilde{z})) = \tilde{z}.$$

Здесь последнее равенство вытекает из леммы 2.

Покажем свойство 2 — монотонность. Пусть $\tilde{Z} \prec \tilde{W}$. Обозначим α -интервалы нечетких чисел \tilde{z}_i и \tilde{w}_i (компонент векторов \tilde{Z}, \tilde{W}) как $[z_i^-(\alpha), z_i^+(\alpha)]$ и $[w_i^-(\alpha), w_i^+(\alpha)]$ соответственно. Так как, по предположению, ϕ — строго монотонно возрастающая функция, то на основании леммы 2 и по правилу сравнения (1) нечетких чисел $\phi(z_i^-(\alpha)) \leq \phi(w_i^-(\alpha))$ и, следовательно,

$$\sum_{i=1}^n \beta_i \phi(\tilde{z}_i^-) \leq \sum_{i=1}^n \beta_i \phi(\tilde{w}_i^-).$$

Аналогично для правых индексов.

Так как ϕ^{-1} возрастает вместе с ϕ , то в соответствии с леммой 3 $F_{\beta, \phi}(\tilde{Z}) \prec F_{\beta, \phi}(\tilde{W})$.

Свойство 3 — непрерывность — проверяется рассуждениями, близкими к доказательству теоремы 1.

В заключение отметим, что взвешенное нечеткое среднее (2) рассматривалось в нечеткой статистике (см., например, [16, гл. 7; 17]). Однако свойства нечетких усредняющих операторов (3), приведенные в настоящей работе, ранее не отмечались. При этом свойства идемпотентности, аддитивности и однородности в случае усредняющего оператора (3) вполне предсказуемы. Непрерывность можно рассматривать в различных метриках на множестве нечетких чисел. Здесь приводится один из вариантов. Что касается свойства монотонности для усредняющих операторов (3), то существенным моментом оказался подбор определения ранжирования нечетких чисел, при котором это свойство выполняется и которое обеспечивает смысловую нагрузку операции агрегирования.

Нелинейное нечеткое среднее (4) и свойства соответствующего нечеткого нелинейного усредняющего оператора (5) ранее, по-видимому, не рассматривались.

Совокупность установленных в настоящей работе свойств для операторов нечеткого усреднения (3) и (5) обеспечивает естественность и адекватность их применения для агрегирования нечеткой информации.

Литература

1. Mesiar R., Kolesarova A., Calvo T., Komornakova M. A. Review of aggregation functions // Fuzzy sets and their extensions: Representation, aggregation and models / Eds. H. Bustince, F. Herrera, J. Montero. — Studies in fuzziness and soft computing ser. — Springer, 2008. Vol. 220. P. 121–144.
2. Леденева Т. М., Подвальный С. Л. Агрегирование информации в оценочных системах // Вестник ВГУ.

- Сер. Системный анализ и информационные технологии, 2016. Vol. 4. P. 155–164.
3. Lopez de Hierro A. F. R., Roldin C., Bustince H., Fernandez J., Rodriguez I., Fardoun H., Lafuente J. Affine construction methodology of aggregation functions // Fuzzy Set. Syst., 2021. Vol. 414. P. 146–164.
4. Dubois D., Prade H. On the use of aggregation operations in information fusion processes // Fuzzy Set. Syst., 2014. Vol. 142. P. 143–161.
5. Beliakov G., Bustince H., Calvo T. A practical guide to averaging functions. — Cham: Springer, 2016. 352 p.
6. Yager R. Quantifier guided aggregation using OWA operators // Int. J. Intell. Syst., 1996. Vol. 11. Iss. 1. P. 49–73.
7. Torra V. Andness directedness for operators of the OWA and WOWA families // Fuzzy Set. Syst., 2021. Vol. 144. P. 28–37.
8. Дюбуа Д., Прад А. Теория возможностей. Приложения к представлению знаний в информатике / Пер. с англ. — М.: Радио и связь, 1990. 288 с. (Dubois D., Prade H. Possibility theory. — New York, NY, USA: Springer, 1988. 280 p.)
9. Пегат А. Нечеткое моделирование и управление / Пер. с англ. — М.: Бином, 2015. 786 с. (Piegat A. Fuzzy modeling and control. — Springer Science & Business Media, 2001. 728 p.)
10. Zadeh L. A. Fuzzy sets // Inform. Control, 1965. Vol. 8. P. 338–353.
11. Tahani H., Keller J. M. Information fusion in computer vision using the fuzzy integral // IEEE T. Syst. Man Cyb., 1990. Vol. 20. Iss. 3. P. 733–741.
12. Kwak K., Pedrycz W. Face recognition: A study in information fusion using fuzzy integral // Pattern Recogn. Lett., 2005. Vol. 26. P. 719–733.
13. Bustince H., Mesiar R., Fernandez J., Galar M., Paterlain D., Altalhi A., Dimuro G. P., Bedregal B., Taka Z. d-Choquet integrals: Choquet integrals based on dissimilarities // Fuzzy Set. Syst., 2021. Vol. 414. P. 1–27.
14. Kaleva O., Seikkala S. On fuzzy metric spaces // Fuzzy Set. Syst., 1984. Vol. 12. P. 215–229.
15. Смоляк С. А. Оценки эффективности инвестиционных проектов в условиях риска и неопределенности. — М.: Наука, 2002. 182 с.
16. Nguyen H. T., Wu B. Fundamentals of statistics with fuzzy data. — Berlin: Springer, 2006. 204 p.
17. De la Rosa de Saa S., Gil M. A., Gonzalez-Rodriguez G., Lopez M. T., Lubiano M. A. Fuzzy rating scale-based questionnaires and their statistical analysis // IEEE T. Fuzzy Syst., 2015. Vol. 23. P. 111–126.
18. Nguyen H. T. A Note on the extension principle for fuzzy sets // J. Math. Anal. Appl., 1978. Vol. 79. P. 369–380.
19. Джини К. Средние величины / Пер. с итал. — М.: Статистика, 1970. 447 с. (Gini C. Le medie. — Torino: UTET, 1958. 512 p.)

Поступила в редакцию 11.08.21

FUZZY AVERAGING OPERATORS IN THE PROBLEM OF AGGREGATING FUZZY INFORMATION

V. L. Khatskevich

N. E. Zhukovsky and Y. A. Gagarin Air Force Academy, 54a Old Bolsheviks Str., 394064 Voronezh, Russian Federation

Abstract: The problem of aggregating fuzzy information by constructing fuzzy averaging operators is considered. Weighted fuzzy averages of systems of fuzzy numbers are studied and a class of nonlinear fuzzy averages of systems of fuzzy numbers is introduced which is a modification to fuzzy numbers of the general class of dissipative numerical averages. The properties of the corresponding averaging operators which are “fuzzy” analogues of the characteristic properties of scalar aggregating functions, are established. This provides a justification for the use of the introduced fuzzy averaging operators in the problem of aggregation of fuzzy information. At the same time, the result of aggregation of fuzzy information given by a set of fuzzy numbers is understood as a fuzzy number that reflects the essential features of this set.

Keywords: averaging fuzzy operators; aggregation of fuzzy information

DOI: 10.14357/19922264220408

References

1. Mesiar, R., A. Kolesarova, T. Calvo, and M. Komornakova. 2008. A review of aggregation functions. *Fuzzy sets and their extensions: Representation, aggregation and models*. Eds. H. Bustince, F. Herrera, and J. Montero. Studies in fuzziness and soft computing ser. Springer. 220:121–144.
2. Ledeneva, T. M., and S. L. Podval'nyy. 2016. Agregirovaniye informatsii v otsenochnykh sistemakh [The aggregation of information in the evaluated system]. *Proceedings of Voronezh State University. Ser. Systems Analysis and Information Technologies* 4:155–164.
3. Lopez de Hierro, A. F. R., C. Roldin, H. Bustince, J. Fernandez, I. Rodriguez, H. Fardoun, and J. Lafuente. 2021. Affine construction methodology of aggregation functions. *Fuzzy Set. Syst.* 414:146–164.
4. Dubois, D., and H. Prade. 2004. On the use of aggregation operations in information fusion processes. *Fuzzy Set. Syst.* 142:143–161.
5. Belyakov, G., H. Bustince, and T. Calvo. 2016. *A practical guide to averaging functions*. Cham: Springer. 352 p.
6. Yager, R. 1996. Quantifier guided aggregation using OWA operators. *Int. J. Intell. Syst.* 11(1):49–73.
7. Torra, V. 2021. Andness directedness for operators of the OWA and WOWA families. *Fuzzy Set. Syst.* 144:28–37.
8. Dubois, D., and H. Prade. 1988. *Possibility theory*. New York, NY: Springer. 280 p.
9. Piegat, A. 2001. *Fuzzy modeling and control*. Springer Science & Business Media. 728 p.
10. Zadeh, L. A. 1965. Fuzzy sets. *Inform. Control* 8:338–353.
11. Tahani, H., and J. M. Keller. 1990. Information fusion in computer vision using the fuzzy integral. *IEEE T. Syst. Man Cyb.* 20(3):733–741.
12. Kwak, K., and W. Pedrycz. 2005. Face recognition: A study in information fusion using fuzzy integral. *Pattern Recogn. Lett.* 26:719–733.
13. Bustince, H., R. Mesiar, J. Fernandez, M. Galar, D. Pater-nain, A. Altalhi, G. P. Dimuro, B. Bedregal, and Z. Takaa. 2021. d-Choquet integrals: Choquet integrals based on dissimilarities. *Fuzzy Set. Syst.* 414:1–27.
14. Kaleva, O., and S. Seikkala. 1984. On fuzzy metric spaces. *Fuzzy Set. Syst.* 12:215–229.
15. Smolyak, S. A. 2002. *Otsenki effektivnosti investitsionnykh proektov v usloviyakh riska i neopredelennosti* [Evaluation of the effectiveness of investment projects under conditions of risk and uncertainty]. Moscow: Nauka. 182 p.
16. Nguyen, H. T., and B. Wu. 2006. *Fundamentals of statistics with fuzzy data*. Berlin: Springer. 204 p.
17. De la Rosa de Saa, S., M. A. Gil, G. Gonzalez-Rodriguez, M. T. Lopez, and M. A. Lubiano. 2015. Fuzzy rating scale-based questionnaires and their statistical analysis. *IEEE T. Fuzzy Syst.* 23:111–126.
18. Nguyen, H. T. 1978. A note on the extension principle for fuzzy sets. *J. Math. Anal. Appl.* 79:369–380.
19. Gini, C. 1958. *Le medie*. Torino: UTET. 512 p.

Received August 11, 2021

Contributor

Khatskevich Vladimir L. (b. 1951) — Doctor of Science in technology, professor, Department of Mathematics, N. E. Zhukovsky and Y. A. Gagarin Air Force Academy, 54a Old Bolsheviks Str., 394064 Voronezh, Russian Federation; vlkhats@mail.ru

О СЛОЖНОСТИ ОБУЧЕНИЯ ЛОГИЧЕСКИХ ПРОЦЕДУР КЛАССИФИКАЦИИ

Е. В. Дюкова¹, А. П. Дюкова²

Аннотация: Исследуются вопросы сложности логического анализа целочисленных данных. Для специальных задач поиска частых и нечастых элементов в данных, на решении которых базируется обучение логических процедур классификации, приведены асимптотики типичного числа решений. Технические основы получения указанных оценок опираются на методы получения аналогичных оценок для трудно-решаемой дискретной задачи построения (перечисления) тупиковых покрытий целочисленной матрицы, формулируемой в работе как задача поиска «минимальных» нечастых элементов. Новые результаты в основном касаются изучения метрических (количественных) свойств частых элементов. Полученные оценки типичного числа часто встречающихся фрагментов в описаниях прецедентов позволяют сделать вывод о перспективности применения алгоритмов поиска таких фрагментов на этапе обучения логических классификаторов типа «Кора».

Ключевые слова: атрибут; частый элементарный фрагмент; нечастый элементарный фрагмент; монотонная дуализация; тупиковое покрытие целочисленной матрицы; классификация по прецедентам; классификатор типа «Кора»

DOI: 10.14357/19922264220409

1 Введение

Рассматриваемые задачи поиска частых и нечастых элементов в данных формулируются следующим образом.

Исследуется множество объектов M . Каждый объект из M может быть представлен в виде числового вектора, полученного на основе наблюдения или измерения ряда его характеристик. Такие характеристики называют атрибутами. Предполагается, что каждый атрибут имеет ограниченное множество допустимых значений, которые кодируются целыми числами.

Пусть $X = \{x_1, \dots, x_n\}$ — заданное множество атрибутов; H — набор из r различных атрибутов вида $H = \{x_{j_1}, \dots, x_{j_r}\}$; $\sigma = (\sigma_1, \dots, \sigma_r)$ — набор, в котором σ_i — допустимое значение признака x_{j_i} , $i = 1, 2, \dots, r$. Пару (σ, H) назовем элементарным фрагментом (ЭФ) длины r . Через $W(M, X)$ обозначим множество всех ЭФ.

Пусть $S = (a_1, \dots, a_n)$ — объект из M (здесь a_j , $j \in \{1, 2, \dots, n\}$, — значение атрибута x_j для объекта S). Будем говорить, что S содержит ЭФ (σ, H) , если $a_{j_i} = \sigma_i$ при $i = 1, 2, \dots, r$.

Дана некоторая совокупность объектов D из M и задано число p , $0 < p \leq 1$. Объекты в D не обязательно различны. Через $|D|$ обозначается число объектов в D . Элементарный фрагмент (σ, H) , $(\sigma, H) \in W(M, X)$, называется (p, D) -частым, если не менее $p|D|$ объектов из D содержат (σ, H) .

Иначе ЭФ (σ, H) — (p, D) -нечастый. Элементарный фрагмент (σ, H) , $(\sigma, H) \in W(M, X)$, называется $(0, D)$ -нечастым, если ни один объект из D не содержит (σ, H) .

Элементарный фрагмент (σ, H) , являющийся (p, D) -частым в $W(M, X)$, называется *максимальным* (p, D) -частым в $W(M, X)$, если любой ЭФ (σ', H') из $W(M, X)$, такой что $\sigma' \supset \sigma$, $H' \supset H$, не является (p, D) -частым.

Элементарный фрагмент (σ, H) длины r называется *правильным* в $W(M, X)$, если (σ, H) — $(r/|D|, D)$ -частый в $W(M, X)$.

Элементарный фрагмент (σ, H) , являющийся (p, D) -нечастым в $W(M, X)$, называется *минимальным* (p, D) -нечастым в $W(M, X)$, если любой ЭФ (σ', H') из $W(M, X)$, такой что $\sigma' \subset \sigma$, $H' \subset H$, не является (p, D) -нечастым. Понятие *минимального* $(0, D)$ -нечастого ЭФ полностью аналогично введенному понятию минимального (p, D) -нечастого ЭФ для $p > 0$.

Возникают две отдельные задачи:

- (1) для заданного p , $0 < p \leq 1$, найти в $W(M, X)$ все (максимальные) (p, D) -частые ЭФ;
- (2) для заданного p , $0 \leq p \leq 1$, найти все (минимальные) (p, D) -нечастые ЭФ.

Иногда требуется совместное перечисление максимальных частых и минимальных нечастых ЭФ.

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

²Федеральный исследовательский центр «Информатика и управление». Российской академии наук, anastasia.d.95@gmail.com

Задачи поиска частых и нечастых элементов в данных входят в число центральных задач интеллектуального анализа данных и особенно важны в случае больших данных. Эти задачи актуальны для многих прикладных областей, среди которых следует выделить нахождение в данных ассоциативных правил и машинное обучение.

В первом случае D называют базой данных, а каждый объект из D называют транзакцией. Ассоциативное правило (АП) устанавливает зависимость между двумя частыми ЭФ, согласно которой один частый ЭФ X (посылка) с некоторой «достоверностью» влечет другой частый ЭФ Y . При этом ЭФ X и Y порождаются одним общим частым ЭФ, обозначаемым (X, Y) . Наиболее информативными считаются те АП, которые порождаются максимальными частыми ЭФ (X, Y) с «минимальной» посылкой X . Вопросы поиска ассоциативных правил возникли в связи с анализом потребительской корзины и наиболее изучены для случая бинарных данных [1].

Одна из главных задач машинного обучения — задача классификации на основе прецедентов. В этом случае D — обучающая выборка (некоторая заданная совокупность примеров объектов из M), а каждый объект из D — обучающий объект, или прецедент. Подлежащие измерению или наблюдению свойства исследуемых объектов называются признаками. В самом простом случае прецеденты делятся на два класса (класс положительных и класс отрицательных примеров). В общем случае число классов может быть больше двух. Требуется по признаковому описанию предъявленного объекта, о котором заранее не известно, к какому классу он относится, определить (распознать) этот класс.

Хорошие результаты показывают логические классификаторы, при конструировании которых используются основные идеи как алгоритма «Кора» [2], так и алгоритмов вычисления оценок [3]. Эти классификаторы впервые предложены в [4]. В дальнейшем подход развивался в работах [5–7] и др. В алгоритмах типа «Кора» анализ прецедентной информации проводится в предположении, что признаковые описания любых двух обучающих объектов, принадлежащих разным классам, не совпадают. На этапе обучения для каждого класса K ищутся так называемые (p, q) -представительные элементарные классификаторы, представляющие собой специальные ЭФ из $W(M, X)$.

Пусть $Q(K)$ и $Q(\overline{K})$ — множества прецедентов из класса K и не из класса K соответственно и $p > 0$, $q < p$. Тогда (p, q) -представительный элементарный классификатор является одновременно (максимальным) $(p, Q(K))$ -частым ЭФ и (минимальным) $(q, Q(\overline{K}))$ -нечастым ЭФ в $W(M, X)$. Как

правило, сначала строятся минимальные $(q, Q(\overline{K}))$ -нечастые ЭФ, а затем из них отбираются те, которые являются $(p, Q(K))$ -частыми. На следующем этапе найденные (p, q) -представительные элементарные классификаторы класса K участвуют в процедуре «голосования» за отнесение распознаваемого объекта к этому классу. Материал обучения безошибочно классифицируется при $q = 0$. Однако нахождение $(0, Q(\overline{K}))$ -нечастых ЭФ требует больших вычислительных затрат. В простейшем случае бинарных данных, когда требуется найти все минимальные $(0, Q(\overline{K}))$ -нечастые ЭФ вида $(0, \dots, 0)$, это известная труднорешаемая перечислительная задача, называемая монотонной дуализацией [8, 9].

Отметим, что к задаче совместного перечисления максимальных частых и минимальных нечастых ЭФ сводится задача расшифровки монотонной функции [10].

В целях исследования скорости решения рассматриваемых задач в случае больших данных представляет интерес получение асимптотик типичного числа частых и нечастых ЭФ, а также типичной длины таких ЭФ. В работах [5, 6] искомые оценки приведены для множества минимальных $(0, Q(\overline{K}))$ -нечастых ЭФ, называемых в этих работах тупиковыми покрытиями целочисленной матрицы. В настоящей работе так же, как и в [5, 6], рассмотрен случай большого числа атрибутов. Требуемые оценки получены для правильных ЭФ в предположении, что каждый атрибут имеет k , $k \geq 2$, допустимых значений. При доказательстве основной теоремы использовано матричное представление данных. Сравнение полученных в настоящей работе оценок с оценками из [5, 6] свидетельствует о перспективности применения методов поиска частых ЭФ для построения $(p, 0)$ -представительных элементарных классификаторов.

2 Асимптотика типичного числа правильных элементарных фрагментов в случае большого числа атрибутов

Пусть $L = (a_{ij})$, $i = 1, \dots, m$, $j = 1, \dots, n$, — матрица с элементами из $\{0, 1, \dots, k-1\}$, $k \geq 2$; E_k^r , $r \leq n$, $k \geq 2$, — множество наборов $(\sigma_1, \dots, \sigma_r)$, $\sigma_i \in \{0, 1, \dots, k-1\}$, $i = 1, 2, \dots, r$; W_r^n , $r \leq n$, — множество всех наборов вида $\{j_1, \dots, j_r\}$, где $j_t \in \{1, 2, \dots, n\}$ при $t = 1, 2, \dots, r$ и $j_1 < \dots < j_r$.

Положим $\sigma \in E_k^r$, $\sigma = (\sigma_1, \dots, \sigma_r)$, $H \in W_r^n$, $H = \{j_1, \dots, j_r\}$. Будем говорить, что строка (a_{i1}, \dots, a_{in}) матрицы L содержит фрагмент (σ, H) ,

если $a_{ij_t} = \sigma_t$ при $t = 1, 2, \dots, r$. Число r назовем длиной фрагмента (σ, H) . Фрагмент (σ, H) длины r назовем *правильным* в L , если не менее r строк в L содержат (σ, H) .

Набор столбцов матрицы L с номерами из H называется *тупиковым-покрытием* длины r , если ни одна строка в L не содержит фрагмент (σ, H) и для любого $t \in 1, 2, \dots, r$ в L найдется строка, содержащая фрагмент (γ, H) , где $\gamma = (\sigma_1, \dots, \sigma_{t-1}, \bar{\sigma}_t, \sigma_{t+1}, \dots, \sigma_r)$, $\bar{\sigma}_t \in \{0, 1, \dots, k-1\}$, $\bar{\sigma}_t \neq \sigma_t$.

Введем обозначения: \mathfrak{M}_{mn}^k — множество всех матриц размера $m \times n$ с элементами из $\{0, 1, \dots, k-1\}$, $k \geq 2$; $R(L)$ — множество правильных фрагментов в матрице L из \mathfrak{M}_{mn}^k ; $T(L, \sigma)$ — множество тупиковых-покрытий матрицы L из \mathfrak{M}_{mn}^k ; $T(L) = \bigcup_{r=1}^n \bigcup_{\sigma \in E_k^r} T(L, \sigma)$; $b_n \sim c_n$, $n \rightarrow \infty$, означает, что $\lim_{n \rightarrow \infty} b_n/c_n = 1$; $b_n \leq_n c_n$ означает, что $b_n \leq c_n$ при всех достаточно больших n ; $|N|$ — мощность множества N ; $\phi_k(m)$ — интервал $(0,5 \log_k mn - 0,5 \log_k \log_k mn - \log_k \log_k \log_k n, 0,5 \log_k mn - 0,5 \log_k \log_k mn + \log_k \log_k \log_k n)$.

В настоящей работе получены асимптотики типичных значений величины $|R(L)|$ и длины фрагмента из $R(L)$. Выявление типичной ситуации связано с высказыванием типа «для почти всех матриц L из \mathfrak{M}_{mn}^k при $n \rightarrow \infty$ выполнено $F_1(L) \sim F_2(L)$ » (здесь $F_1(L)$ и $F_2(L)$ — два функционала, заданные на матрицах из \mathfrak{M}_{mn}^k). Данное высказывание означает, что существуют две положительные бесконечно убывающие функции $\alpha(n)$ и $\beta(n)$ такие, что для всех достаточно больших n имеет место $1 - |\mathfrak{M}|/|\mathfrak{M}_{mn}^k| \leq \alpha(n)$, где \mathfrak{M} — множество таких матриц L в \mathfrak{M}_{mn}^k , для которых $1 - \beta(n) < |F_1(L)|/|F_2(L)| < 1 + \beta(n)$.

Аналогичные результаты для множества $T(L)$ получены в [5, 6] и приведены ниже (теорема 1) для полноты изложения.

Теорема 1 (см. [5, 6]). *Если $m^a \leq n \leq k^m$, $a > 1$, то для почти всех матриц L из \mathfrak{M}_{mn}^k справедливо*

$$|T(L)| \sim \sum_{r \in \phi_k(m)} C_n^r C_m^r r! (k-1)^r k^{r-r^2}, \quad n \rightarrow \infty,$$

и длины почти всех тупиковых покрытий из $T(L)$ принадлежат интервалу $\phi_k(m)$.

Теорема 2. *Если $m^a \leq n \leq k^m$, $a > 1$, $k \geq 2$, то для почти всех матриц L из \mathfrak{M}_{mn}^k справедливо*

$$|R(L)| \sim \sum_{r \in \phi_k(m)} C_n^r C_m^r k^{r-r^2}, \quad n \rightarrow \infty,$$

и длины почти всех фрагментов из $R(L)$ принадлежат интервалу $\phi_k(m)$.

Доказательство теоремы 2 опирается на ряд приводимых ниже лемм.

Обозначим через V_r^m ($r \leq m$) множество всех упорядоченных наборов вида (i_1, \dots, i_r) , где $i_t \neq i_l$ при $t, l = 1, 2, \dots, r$.

Пусть $v \in V_r^m$, $H \in W_r^n$, $\sigma \in E_k^r$. Матрица $L \in \mathfrak{M}$ называется (v, σ, H) -матрицей, если каждая строка с номером v содержит фрагмент (σ, H) . Матрица $L \in \mathfrak{M}_{mn}^k$, являющаяся (v, σ, H) -матрицей, называется *правильной*, если из условия $v_1 \in V_r^m$, $v_1 \neq v$, следует, что L не является (v_1, σ, H) -матрицей. Обозначим через $N_{(v, \sigma, H)}$ совокупность (v, σ, H) -матриц в \mathfrak{M}_{mn}^k , через $N_{(v, \sigma, H)}^*$ — совокупность всех правильных (v, σ, H) -матриц в \mathfrak{M}_{mn}^k .

Лемма 1. *Если $v \in V_r^m$, $H \in W_r^n$, $\sigma \in E_k^r$, то*

$$|N_{(v, \sigma, H)}| = k^{mn-r^2}.$$

Доказательство. Оценим, сколькими способами можно построить матрицу L из $N_{(v, \sigma, H)}$. Однозначным образом определяются те элементы матрицы L , которые расположены на пересечении строк с номерами из v и столбцов с номерами из H . Остальные элементы этой матрицы могут быть выбраны произвольным образом. Таким образом, элементы матрицы L , расположенные в столбцах с номерами, не входящими в H , могут быть выбраны $k^{m(n-r)}$ способами, а строки подматрицы матрицы L , образованной столбцами из H , можно выбрать $k^{r(m-r)}$ способами. Отсюда получаем требуемую оценку. Лемма доказана.

Лемма 2. *Если $v \in V_r^m$, $H \in W_r^n$, $\sigma \in E_k^r$, то*

$$|N_{(v, \sigma, H)}^*| = (1 - k^{-r})^{m-r} k^{mn-r^2}.$$

Доказательство. Оценим, сколькими способами можно построить матрицу L из $N_{(v, \sigma, H)}^*$. Элементы этой матрицы, расположенные в столбцах с номерами, не входящими в H , могут быть выбраны произвольным образом ($k^{m(n-r)}$ способов). Отсюда, учитывая, что строки в подматрице матрицы L , образованной столбцами из H , можно выбрать $(k^r - 1)^{m-r}$ способами, получаем требуемую оценку. Лемма доказана.

Лемма 3. *Пусть $v_1 \in V_r^m$, $v_2 \in V_l^m$, $H_1 \in W_r^n$, $H_2 \in W_l^n$, $\sigma' \in E_k^r$, $\sigma'' \in E_k^l$ и наборы v_1 и v_2 пересекаются по a ($a \geq 0$) элементам, а наборы w_1 и w_2 пересекаются по b ($b \geq 0$) элементам. Тогда*

$$|N_{(v_1, \sigma', H_1)} \cap N_{(v_2, \sigma'', H_2)}| \leq k^{mn-r^2-l^2+ab}.$$

Доказательство леммы 3 не приводится в силу ее очевидности.

Лемма 4. Если $m^a \leq n \leq k^m$, $a > 1$, $k \geq 2$, то имеет место

$$\sum_{r=1}^m C_n^r C_m^r k^{r-r^2} \sim \sum_{r \in \phi_k(m)} C_n^r C_m^r k^{r-r^2}, \quad n \rightarrow \infty.$$

Доказательство леммы 4 аналогично доказательству леммы 4 из работы [6].

Лемма 5. Если $r, l \leq c \log_k n$, $c < 1$, то имеет место

$$\sum_{m=0}^{\min(r,l)} k^{lb} C_n^r C_r^b C_{n-r}^{l-b} < C_n^r C_n^l (1 + \delta(n)),$$

где $\delta(n) \rightarrow 0$ при $n \rightarrow \infty$.

Доказательство леммы 5 аналогично доказательству леммы 5 из работы [6].

Будем считать $\mathfrak{M}_{mn}^k = \{L\}$ пространством элементарных событий, в котором каждое событие L происходит с вероятностью $1/|\mathfrak{M}_{mn}^k|$. Математическое ожидание случайной величины $X(L)$, определенной на множестве \mathfrak{M}_{mn}^k , будем обозначать через $MX(L)$, дисперсию — через $DX(L)$.

Лемма 6 (см. [11]). Пусть для случайных величин $X_1(L)$ и $X_2(L)$, определенных на \mathfrak{M}_{mn}^k , выполнено $X_1(L) \geq X_2(L) \geq 0$ и при $n \rightarrow \infty$ верно $MX_1(L) \sim MX_2(L)$, $DX_2(L)/(MX_2(L))^2 \rightarrow 0$. Тогда для почти всех матриц L из \mathfrak{M}_{mn}^k имеет место $X_1(L) \sim X_2(L) \sim MX_2(L)$, $n \rightarrow \infty$.

Пусть $v \in V_r^m$, $H \in W_r^n$, $\sigma \in E_k^r$. На $\mathfrak{M}_{mn}^k = \{L\}$ рассмотрим случайную величину $\zeta_{(\sigma,H)}(L)$, равную 1, если (σ, H) — правильный фрагмент в L , и равную 0 в противном случае. Положим

$$\zeta_1(L) = \sum_{r=1}^{\min(m,n)} \sum_{\substack{v \in V_r^m \\ H \in W_r^n}} \sum_{\sigma \in E_k^r} \zeta_{(\sigma,H)}(L);$$

$$\zeta_2(L) = \sum_{r \in \phi_k(m)} \sum_{\substack{v \in V_r^m \\ H \in W_r^n}} \sum_{\sigma \in E_k^r} \zeta_{(\sigma,H)}(L).$$

Нетрудно видеть, что $\zeta_1(L) = |R(L)|$, а $\zeta_2(L)$ — число тех фрагментов в $R(L)$, длины которых принадлежат интервалу $\phi_k(m)$.

Оценим вероятность события $\zeta_{(\sigma,H)}(L) = 1$, обозначаемую далее через $P(\zeta_{(\sigma,H)}(L) = 1)$. Очевидно, в силу леммы 1

$$P(\zeta_{(\sigma,H)}(L) = 1) \leq \frac{|N_{(v,\sigma,H)}|}{|\mathfrak{M}_{mn}^k|} = k^{-r^2}. \quad (1)$$

С другой стороны, в силу леммы 2 имеем

$$P(\zeta_{(\sigma,H)}(L) = 1) \geq \frac{|N_{(v,\sigma,H)}^*|}{|\mathfrak{M}_{mn}^k|} \geq (1 - k^{-r})^{m-r} k^{-r^2}.$$

В случае $r \in \phi_k(m)$ получаем

$$(1 - k^{-r})^{m-r} \geq (1 - k^{-r})^m \geq 1 - mk^{-r};$$

$$mk^{-r} \leq \frac{c_1 \log_k^2 n}{n^{c_2}}, \quad c_1, c_2 > 0.$$

Следовательно,

$$P(\zeta_{(\sigma,H)}(L) = 1) \geq F(n)k^{-r^2}, \quad (2)$$

где $F(n) \rightarrow 1$ при $n \rightarrow \infty$.

Лемма 7. Если $m^a \leq n \leq k^m$, $a > 1$, $k \geq 2$, то

$$M\zeta_1(L) \sim M\zeta_2(L) \sim \sum_{r \in \phi_k(m)} C_n^r C_m^r k^{r-r^2}, \quad n \rightarrow \infty.$$

Доказательство. Имеем

$$M\zeta_1(L) = \sum_{r=1}^m \sum_{\substack{v \in V_r^m \\ H \in W_r^n}} \sum_{\sigma \in E_k^r} P(\zeta_{(\sigma,H)}(L) = 1);$$

$$M\zeta_2(L) = \sum_{r \in \phi_k(m)} \sum_{\substack{v \in V_r^m \\ H \in W_r^n}} \sum_{\sigma \in E_k^r} P(\zeta_{(\sigma,H)}(L) = 1).$$

Следовательно, в силу (1)

$$M\zeta_2(L) \leq M\zeta_1(L) \leq \sum_{r=1}^m C_n^r C_m^r k^{r-r^2}. \quad (3)$$

В силу (2)

$$M\zeta_1(L) \geq M\zeta_2(L) \geq F(n) \sum_{r \in \phi_k(m)} C_n^r C_m^r k^{r-r^2}, \quad (4)$$

где $F(n) \rightarrow 1$ при $n \rightarrow \infty$.

Из (3), (4) и леммы 4 сразу следует утверждение доказываемой леммы.

Лемма 8. Имеет место

$$\frac{D\zeta_2(L)}{(M\zeta_2(L))^2} \rightarrow 0, \quad n \rightarrow \infty.$$

Доказательство. Имеем

$$D\zeta_2(L) = M(\zeta_2(L))^2 - (M\zeta_2(L))^2. \quad (5)$$

Нетрудно видеть, что

$$M(\zeta_2(L))^2 \leq \sum_{r,l \in \phi_k(m)} \sum_{\substack{v_1 \in V_r^m, v_2 \in V_l^m \\ H_1 \in W_r^n, H_2 \in W_l^n}} \sum_{\substack{\sigma' \in E_k^r \\ \sigma'' \in E_k^l}} \frac{|N|}{k^{mn}},$$

где $N = N_{(v_1, \sigma', H_1)} \cap N_{(v_2, \sigma'', H_2)}$. Отсюда, пользуясь леммами 3 и 5, получаем

$$\begin{aligned}
 & M(\zeta_2(L))^2 \leq \\
 & \leq \sum_{r,l \in \phi_k(m)} \sum_{b=0}^{\min(r,l)} k^{r+l} k^{-r^2-l^2+lb} C_n^r C_r^b C_{n-r}^{l-b} C_m^r C_m^l \leq \\
 & \leq \sum_{r,l \in \phi_k(m)} C_n^r C_n^l C_m^r C_m^l k^{r+l} k^{-r^2-l^2} (1 + \delta(n)), \quad (6)
 \end{aligned}$$

где $\delta(n) \rightarrow 0$ при $n \rightarrow \infty$.

С другой стороны, в силу леммы 7

$$(M\zeta_2(L))^2 \sim \sum_{r,l \in \phi_k(m)} C_n^r C_n^l C_m^r C_m^l k^{r+l} k^{-r^2-l^2}, \quad n \rightarrow \infty. \quad (7)$$

Из (5)–(7) следует утверждение доказываемой леммы.

Утверждение теоремы 2 следует непосредственно из лемм 6–8.

Замечание. В задаче классификации по прецедентам рассмотрим класс K и представим множества прецедентов $Q(K)$ и $Q(\bar{K})$ в виде матриц, имеющих размеры $m_1 \times n$ и $m_2 \times n$ соответственно. Тогда, заменяя m на m_1 , в качестве следствия из теоремы 2 получим асимптотики типичных значений числа и длины правильных ЭФ в $Q(K)$ и, заменяя m на m_2 , в качестве следствия из теоремы 1 получим аналогичные характеристики для $(0, Q(\bar{K}))$ -нечастых ЭФ. Сравнение этих оценок свидетельствует об эффективности в плане вычислительных затрат методов поиска частых ЭФ для построения представительных элементарных классификаторов в алгоритмах типа «Кора» в случае, когда m_2 не меньше m_1 . Полученные оценки согласуются с приведенными в [12] результатами экспериментальных исследований.

3 Заключение

Для множества объектов, описываемых в виде наборов значений атрибутов (измеряемых или наблюдаемых характеристик объектов), исследованы метрические (количественные) свойства часто и нечасто встречающихся частей описаний объектов (ЭФ). Рассмотрены специальные виды частых и нечастых ЭФ и приведены асимптотики типичных значений их числа и длины. Подобные оценки ранее были известны исключительно для множества решений дискретной перечислительной задачи, называемой монотонной дуализацией, а также некоторых обобщений этой задачи. Имеются в виду задачи построения тупиковых покрытий булевых

и целочисленных матриц. Полученные результаты свидетельствуют о целесообразности (в плане сокращения временных затрат) применения методов поиска частых ЭФ при построении логических процедур классификации по прецедентам, базирующихся на решении задачи монотонной дуализации.

Литература

1. Aggarwal C. Frequent pattern mining. — Heidelberg: Springer, 2014. 467 p.
2. Вайнцивайг М. Н. Алгоритм обучения распознаванию образов «Кора» // Алгоритмы обучения распознаванию образов / Под ред. В. Н. Вапника. — М.: Советское радио, 1973. С. 110–116.
3. Журавлёв Ю. И. Об алгебраическом подходе к решению задач распознавания и классификации // Проблемы кибернетики, 1978. Вып. 33. С. 5–68.
4. Баскакова Л. В., Журавлёв Ю. И. Модель распознающих алгоритмов с представительными наборами и системами опорных множеств // Ж. вычисл. матем. матем. физ., 1981. Т. 21. № 5. С. 1264–1275.
5. Дюкова Е. В. Алгоритмы распознавания типа «Кора»: сложность реализации и метрические свойства // Распознавание, классификация, прогноз. — М.: Наука, 1989. Вып. 2. С. 99–125.
6. Дюкова Е. В., Журавлев Ю. И. Дискретный анализ признаков описаний в задачах распознавания большой размерности // Ж. вычисл. матем. матем. физ., 2000. Т. 40. № 8. С. 1264–1278.
7. Дюкова Е. В., Песков Н. В. Поиск информативных фрагментов описаний объектов в дискретных процедурах распознавания // Ж. вычисл. матем. матем. физ., 2002. Т. 42. № 5. С. 741–753.
8. Fredman M. L., Khachiyan L. On the complexity of dualization of monotone disjunctive normal forms // J. Algorithm., 1996. Vol. 21. No. 3. P. 618–628.
9. Дюкова Е. В., Прокофьев П. А. Об асимптотически оптимальных алгоритмах дуализации // Ж. вычисл. матем. матем. физ., 2015. Т. 55. № 5. С. 895–910.
10. Алексеев В. Б. О расшифровке некоторых классов монотонных многозначных функций // Ж. вычисл. матем. матем. физ., 1976. Т. 16. № 1. С. 189–198.
11. Носков В. Н., Слепян В. А. О числе тупиковых тестов для одного класса таблиц // Кибернетика, 1972. № 1. С. 60–65.
12. Dragunov N., Djukova E., Djukova A. Supervised classification and finding frequent elements in data // 8th Conference (International) on Information Technology and Nanotechnology Proceedings. — Piscataway, NJ, USA: IEEE, 2022. Art. 9848521. 5 p. doi: 10.1109/ITNT55410.2022.9848521.

Поступила в редакцию 30.09.22

ON THE COMPLEXITY OF LOGICAL CLASSIFICATION LEARNING PROCEDURES

E. V. Djukova and A. P. Djukova

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

Abstract: The issues of integer data logical analysis complexity are investigated. For special tasks of searching in data for frequent and infrequent elements, on the solution of which logical supervised classification procedures are based, asymptotics of a typical number of solutions are given. The technical foundations for obtaining these estimates are based on methods for obtaining similar estimates for intractable discrete problem of constructing (enumerating) irredundant coverings of integer matrix formulated in the paper as the problem of finding “minimal” infrequent elements. The new results mainly concern the study of metric (quantitative) properties of frequent elements. The obtained estimates for the typical number of frequently occurring fragments in precedent descriptions allow one to conclude that the use of algorithms for finding such fragments at the stage of training logical classifiers of the “Kora” type is promising.

Keywords: attribute; frequent elementary fragment; infrequent elementary fragment; monotone dualization; irredundant covering of integer matrix; supervised classification; classifier of “Kora” type

DOI: 10.14357/19922264220409

References

1. Aggarwal, C. 2014. *Frequent pattern mining*. Heidelberg: Springer. 467 p.
2. Weinzweig, M. N. 1973. Algoritm obucheniya raspoznavaniyu obrazov “Kora” [Algorithm for learning pattern recognition “Kora”]. *Algoritmy obucheniya raspoznavaniyu obrazov* [Algorithms for learning pattern recognition]. Ed. V. N. Vapnik. Moscow: Sovetskoe radio. 110–116.
3. Zhuravlev, Yu. I. 1978. Ob algebraicheskom podkhode k resheniyu zadach raspoznavaniya i klassifikatsii [On the algebraic approach to solving recognition and classification tasks]. *Problemy kibernetiki* [Problems of Cybernetics] 33:5–68.
4. Baskakova, L., and Yu. Zhuravlev. 1981. A model of recognition algorithms with representative samples and systems of supporting sets. *USSR Comp. Math. Math.* 21(5):189–199.
5. Djukova, E. V. 1989. Algoritmy raspoznavaniya tipa “Kora”: slozhnost’ realizatsii i metricheskie svoystva [Kora-type recognition algorithms: Implementation complexity and metric properties]. *Raspoznavanie, klassifikatsiya, prognoz* [Recognition, classification, and prediction]. Moscow: Nauka. 2:99–125.
6. Djukova, E., and Y. Zhuravlev. 2000. Discrete analysis of feature descriptions in recognition problems of high dimensionality. *Comp. Math. Math. Phys.* 40(8):1214–1227.
7. Djukova, E., and N. Peskov. 2002. Search for informative fragments of object descriptions in discrete recognition procedures. *Comp. Math. Math. Phys.* 42(5):711–723.
8. Fredman, M., and L. Khachiyan. 1996. On the complexity of dualization of monotone disjunctive normal forms. *J. Algorithm.* 21(3):618–628.
9. Djukova, E., and P. Prokofyev. 2015. Asymptotically optimal dualization algorithms. *Comp. Math. Math. Phys.* 55(5):891–905.
10. Alekseev, V. B. 1976. Deciphering of some classes of monotonic many-valued functions. *Comp. Math. Math. Phys.* 16(1):180–189.
11. Noskov, V. N., and V. A. Slepyan. 1972. Number of dead-end tests for a certain class of tables. *Cybern. Syst. Anal.* 8(1):64–71.
12. Dragunov, N., E. Djukova, and A. Djukova. 2022. Supervised classification and finding frequent elements in data. *8th Conference (International) on Information Technology and Nanotechnology Proceedings*. Piscataway, NJ: IEEE. 9848521. 5 p. doi: 10.1109/ITNT55410.2022.9848521.9848521.

Received September 30, 2022

Contributors

Djukova Elena V. (b. 1945) — Doctor of Science in physics and mathematics, principal scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; associate professor, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskie Gory, GSP-1, Moscow 119991, Russian Federation; edjukova@mail.ru

Djukova Anastasia P. (b. 1995) — PhD student, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; anastasia.d.95@gmail.com

ТЕХНОЛОГИЯ КЛАССИФИКАЦИИ ТИПОВ КОНТЕНТА ЭЛЕКТРОННОГО УЧЕБНИКА*

А. В. Босов¹, А. В. Иванов²

Аннотация: Решается задача классификации контента электронной обучающей системы (ЭОС), относящегося к блокам задач или практических примеров. Необходимость автоматизированного классификатора обоснована перспективным направлением развития ЭОС — оценением качества обучающего контента. Ключевая идея — моделирование контента объектом с двумя свойствами: текстовым описанием на традиционном языке и набором формульных выражений на языке научной компьютерной верстки Т_ЭX. На основе электронного учебника по теории функций комплексного переменного выполнено формирование размеченных в соответствии с данной моделью образцов задач. Проведено обучение четырех алгоритмов классификации текстового контента — наивного байесовского классификатора, логистической регрессии, однослойной и многослойной нейронных сетей прямого распространения. Проведен комплекс сравнительных экспериментов на реализованных классификаторах как в условиях ограниченных моделей, содержащих только контент одного типа, текстовый или формульный, так и полной модели. В результате эксперимента выполнено не только формальное сравнение алгоритмов, но и показано принципиальное преимущество полной модели: при учете обоих свойств контента, текстового и формульного, качество классификации существенно превосходит однофакторные алгоритмы и обеспечивает показатели, подтверждающие готовность технологии к практическому применению.

Ключевые слова: электронная обучающая система; обучающий контент; задача и алгоритмы классификации; оценка качества контента; машинное обучение

DOI: 10.14357/19922264220410

1 Введение

Учебный процесс сегодня невозможно представить без средств дистанционного обучения, на что, несомненно, повлияла пандемия последних лет [1]. Формально эти средства могут быть различны: от простого электронного учебника для школьника до интегрированной электронно-обучающей среды вуза. И, конечно, любая форма представления учебного материала должна обеспечивать его высокое качество. Значит, нужны определенные правила его формирования, стандарты, позволяющие контролировать качество контента ЭОС.

На данный момент довольно ограниченное внимание профессионального сообщества привлекал только анализ качества традиционных учебников [2]. Электронные обучающие системы за счет средств автоматизации имеют гораздо более разнообразный контент, обеспечиваемый в значительной степени функциональными инструментами. Анализ качества их контента в этой связи представляется, с одной стороны, более сложной задачей, но с другой стороны, электронный контент по своей сущности лучше подготовлен для исследования

техническими методами, которые могут быть автоматизированы и потому не требуют привлечения экспертов. Некоторые принципиальные рекомендации для оценивания контента ЭОС приведены в [3, 4]. Говорить о возможности автоматизации таких методик в полном объеме пока преждевременно, но выделить некоторые задачи можно. Одной из них посвящена данная статья.

Важным инструментом системы анализа качества контента ЭОС должна стать технология классификации элементов контента по тематикам изучаемой дисциплины. Средство автоматического определения раздела курса, к которому относится задача, тест, пример, вопрос экзаменационного билета, будет ключевой частью как методик анализа качества контента, так и самого процесса обучения. Это, например, анализ полноты и достаточности ЭОС, определение сложности комплекта заданий (контрольной работы, экзаменационного билета), генерации заданий необходимого уровня сложности и/или тематической направленности и пр. В статье рассматривается наиболее типовая постановка задачи тематической классификации контента, когда имеется фиксированный перечень

* Исследование выполнено за счет гранта Российского научного фонда (проект 22-28-00588). Работа выполнялась с использованием инфраструктуры Центра коллективного пользования «Высокопроизводительные вычисления и большие данные» (ЦКП «Информатика» ФИЦ ИУ РАН, Москва).

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

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

разделов курса (оглавление учебника), образцы заданий (примеров, тестов, задач) для этих разделов и требуется обучить автоматический классификатор, который будет относить задания к разделам. Если бы задания имели исключительно текстовые (традиционно языковые) формулировки, то проблема сводилась бы к типовой классификации текстов. Но классифицировать требуется математический контент, т.е. формульные выражения. Формальное использование техник текстовой классификации приемлемых результатов не даст. Можно ли и каким образом применять эти техники, какая потребуется адаптация — это вопросы исследования, результаты которого представлены в статье.

Когда речь заходит об автоматических средствах обработки формульных выражений, то первыми в их числе обнаруживаются средства распознавания набранных и рукописных математических выражений. Среди многих работ на эту тему можно отметить [5], где система распознавания образов использует естественное для формул представление в виде древовидных структур. Подробности есть в обзорах [6, 7]. Для темы статьи важно то, что конечная задача в таких проектах состоит в получении формул в формате электронных редакторов текстов, например в \LaTeX . Потребителю здесь нужны в основном инструменты распознавания в области рукописного ввода.

Наиболее близка к рассматриваемой тематика поиска математических знаний (хороший обзор дан в [8]). Заметим, что и на сегодняшний день поисковые системы не имеют функций поиска с использованием формул, хотя этой проблемой много лет занимаются специалисты и выполнено много вполне содержательных проектов. Наиболее продуктивными оказывались те, что опирались на формульное представление на языке научной разметки \TeX [9]. В таких проектах речь идет, как правило, о полнотекстовом поиске с использованием математических выражений. В некоторый период такая система \LaTeXSearch была представлена издательством Springer, поиск формул реализовывался в своей библиотеке (к сожалению, сейчас проект недоступен и сведений о нем больше нет).

Можно выделить два вида стратегий реализации математического поиска. Первая заключается в расширении существующей системы текстового поиска путем преобразования двумерных формул в линейные строки, такие как \LeActiveMath [10, 11], \MathDex [12], \EgoMath [13], \DLMFSearch [14] и \MIaS [15]. Все эти методы направлены на то, чтобы превратить формулу в символьный поток и включить его в полнотекстовой поиск. Вторая стратегия заключается в реализации неко-

торой модели извлечения знаний для математических выражений и осуществлена, например, в \MathWebSearch [16] и \WikiMirs [17]. Еще один пример более-менее актуального исследования представлен в [18, 19], где реализовано онтологическое описание формул \TeX , основанное на присущей математическим формулам иерархии. На ее основе строится набор признаков, которые индексируются и участвуют в поиске. Использование иерархической структуры гораздо лучше передает семантику, но используется это все равно для построения полнотекстового индекса. Еще одна любопытная идея представлена в [20, 21], где предлагается объединить в поиске и формульные представления, и текстовые данные, извлекаемые из текста с этой формулой. Кроме того, обсуждается выполнение запросов на естественном языке.

Применить процитированные работы и другие результаты в области поиска математических знаний в постановках по классификации математических текстов вряд ли получится. Так, могло бы быть интересным определять по формулам область знаний, хотя бы в форме математика—физика—химия. Более глубокий вопрос мог бы звучать так: к какой области математики относится формула — алгебра, дифференциальные уравнения, стохастический анализ и т.д. Такого рода задачи тематически ближе к представленному в статье исследованию. Наверное, как и в представленном далее результате, можно обучить такой классификации подходящую нейронную сеть, но проблемой будет не столько технология, сколько создание репрезентативного обучающего набора. В рассматриваемой задаче ситуация проще, так как ограничена тематической классификацией текстово-формульного контента в рамках конкретной математической дисциплины. Эту роль выполняет теория функций комплексного переменного, выбранная по той причине, что авторы участвуют в общем проекте с авторами актуального учебника по данной дисциплине, разработанного для применения в составе действующего ЭСО. Это обеспечило работу содержательным электронным контентом, отсутствие которого в любой другой постановке сделало бы задачу неподъемной, поскольку реальных ресурсов для создания обучающего набора требуется гораздо больше, чем на само исследование. В данном случае благодаря доступности учебного материала в электронной форме задачу удалось не только поставить и предложить для нее модель и алгоритм, но и выполнить реальные расчеты для нескольких вариантов решения и сравнить результаты.

Статья организована следующим образом. В разд. 2 дано описание использованного учебного контента и процесса его предварительной подготов-

ки для формирования обучающего набора. В разд. 3 описана модель для выделения признаков из формульной части контента. В разд. 4 приведены результаты классификации, выполненной для текстовой, формульной и текстово-формульной моделей четырьмя типовыми алгоритмами классификации.

2 Исходные данные и предварительная подготовка

Основной использованный учебный материал (рис. 1) — это пособие [22], предоставленное для обработки его авторами в файле формата MS Word.

В пособии по дисциплине «Теория функций комплексного переменного» материал сгруппирован по 9 разделам (главам). Это оглавление и было использовано для классификации контента. В каждом разделе есть задания для аудиторной работы (с решениями) и задания для самостоятельного решения. Суммарно из [22] были получены 174 задачи. При этом число задач в разделах составило от 6 до 42, т.е. задачи оказались распределены по разделам довольно неравномерно. Поэтому к задачам, выделенным из [22] автоматически, были вручную добавлены задачи в наиболее «бедные» разделы. Источником этих «дополнительных» задач стали пособия [23, 24]. В итоге удалось сформировать достаточный обучающий блок из 200 задач, распределенных по имеющимся 9 разделам следующим образом.

1. Комплексные числа и действия — 42.
2. Функции комплексного переменного — 31.
3. Дифференцируемость. Аналитические функции — 18.
4. Интегрирование — 31.
5. Ряды — 15.
6. Нули — 16.
7. Вычеты — 16.
8. Преобразование Лапласа — 15.
9. Применение операционного метода — 16.

С учетом объема формульного материала данное распределение оказалось практически равномерным.

Задачи из каждого раздела сохранялись как отдельный документ. Далее документы были конвертированы в формат TeXc помощью приложения Pandoc. Каждая задача состоит из текстовой формулировки и нескольких формул, поэтому модель задач представлена двумя наборами признаков, которые используются для обучения классификаторов. Первый набор признаков формируется на основе текстовой части задачи с использованием типовых средств обработки текстов на естественном языке. Второй набор формируется на основе формул, для чего предложены специальные (кастомные) процедуры.

Процедура формирования описаний задач из файлов формата TeX включает следующие этапы:

- разбиение документа на отдельные задачи;
- выделение текстовой части задачи (удаляются все формулы);
- выделение и объединение всех формул, относящихся к задаче;
- уточнение формульной семантики.

В качестве основного инструмента для реализации этих этапов использовались регулярные выражения. Сами процедуры реализованы на языке Python. Для этапа выделения задач нужно идентифицировать отличительные признаки текста, позволяющие принять решение о том, является ли фрагмент текста номером задачи, формулой или описанием. Например, уникальным признаком номера задачи служит наличие символа «№». Как правило, данный символ выделен жирным шрифтом с помощью тега «`\textbf{}`». В ряде заданий имеются отсылки к номерам других заданий, которые следует исключать. Результирующее регулярное выражение имеет вид:

```
[^ ](\textbf{?}(\#){(.*)})?
```

Для выделения и/или исключения формул использовались следующие регулярные выражения:

```
\\((([sS]+?)\\) )
\\[[([sS]+?)\\]
\\$([sS]+?)\\$
```

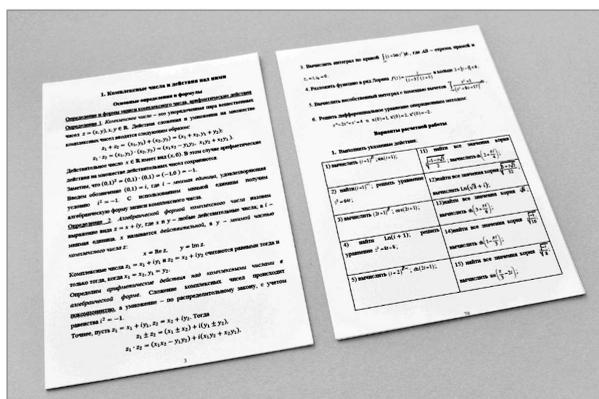


Рис. 1 Учебный материал

Эти выражения позволяют выделить типичные для формата \TeX способы определения формул — парные теги «\($\)$ » и «\($\)$ », «\($\)$ » и «\($\)$ », «\$» и «\$». Кроме того, при формировании описаний задач пришлось учитывать возможность наличия нескольких вариантов для одной формулировки. Каждое задание может включать несколько вариантов, обозначенных «a)», «б)», «в)» или «1)», «2)», «3)». В некоторых заданиях эта структура имеет два уровня, например нужно вычислить интеграл для нескольких областей. В ряде заданий все варианты объединены в одну формулу через символ «;». Из таких заданий выделялось столько задач, сколько вариантов представлено фактически. В итоговое описание каждой задачи попадают все формулы, относящиеся отдельно к каждому варианту и имеющиеся в общей формулировке задания. Выделение всех вариантов из задания обеспечивают:

- (1) регулярные выражения для последовательности символов вида «1)», «2)», «3)»:

```
\sd+\)(?=[\s\*]*\\\(|[\s\*]*\\\(|[\s\*]*\)$)
```

- (2) регулярные выражения для последовательности символов вида «a)», «б)», «в)»:

```
\s[a-я]\)(?=[\s\*~]*.)
```

- (3) процедура разбиения формул для случая, когда все варианты объединены в одну формулу через символ «;».

Наконец, отдельных преобразований потребовало сохранение семантики формулы, например запись функции косинус, которая может быть реализована тегом «\cos» или последовательностью символов «cos». Также надо было учитывать возможные пропуски некоторых операций, например умножения, и скобок вокруг аргументов функций. Если не учитывать такие ситуации при подготовке описаний задач, то это приведет к неоднозначностям на следующем этапе подготовки формул. Так, при токенизации будут выделены знаки операций, теги \TeX , последовательности латинских букв и цифр. Соответственно, такие элементы, как «xe», «iy», «zdz», «cost», должны содержать два токена. Для этого требуется предварительная подготовка формул перед их помещением в набор описаний задач: для учета особенностей формульной семантики выполняется процедура, включающая

- (1) стандартизацию форматирования (удаляются незначимые символы, такие как неразрывные пробелы, точки и запятые в конце формулы; группы из нескольких пробелов заменяются одним; символы перевода строки и возврата каретки заменяются на пробелы; символ «;» меняется на «,»);

- (2) разбиение формулы на базовые токены с использованием следующего регулярного выражения:

```
(\d+|\d+\.?|\d+|[\s,/+\-|\:\.\_\^{\}\{\<>]|\{\|\[a-zA-Z\-\.\]++)
```

Данное выражение позволяет выделить следующие типы токенов: пробельные символы, знаки пунктуации «,», «.», «:»; знаки операций «+», «-», «\», «<», «>»; знаки нижнего и верхнего индекса «_», «^»; скобки («(», «)», «{», «}», «|»; теги \TeX ; целые и вещественные числа;

- (3) разбиение элементов типа «последовательность латинских букв», в результате которого выделяются и стандартизируются элементы, содержащие тригонометрические, гиперболические функции и логарифмы, например элемент «cosh y», а также обрабатываются специальные имена «Re», «Im», «matrix», «res» с использованием следующего регулярного выражения:

```
^(d[a-z]|Re|Im|matrix|res)
```

Данное выражение используется процедурой, позволяющей, например, заменить элементы « x » на « $x dx$ », « Imz » на « $Im z$ » и т. д. Элементы, не относящиеся к специальным именам, дробятся побуквенно и образуют отдельные токены.

Каждое сформированное описание задачи, естественно, включает и правильный результат классификации — номер раздела учебного пособия, из которого было выделено задание.

3 Выделение признаков в модели задачи

Обе части описания задач, текстовая и формульная, далее использовались для построения двух векторов признаков.

Вектор признаков из текстового описания задачи — это типичное представление текстовых данных на естественном языке в машинном обучении, а именно: модель по методу «мешок слов» [25]. Тексты задач разбиваются на токены, из токенов формируется словарь, и модель текста получается как гистограмма — вектор весов всех элементов словаря, входящих в описание задачи. В качестве весов слов использовалась мера $tf-idf$ [26], т. е. частота слова в описании, умноженная на инверсную частоту слова во всех описаниях (в документе). Для выполнения токенизации текста (лексического анализа) и лемматизации (нормализации формы слова) использовалась библиотека `spaCy`, <https://spacy.io/>,

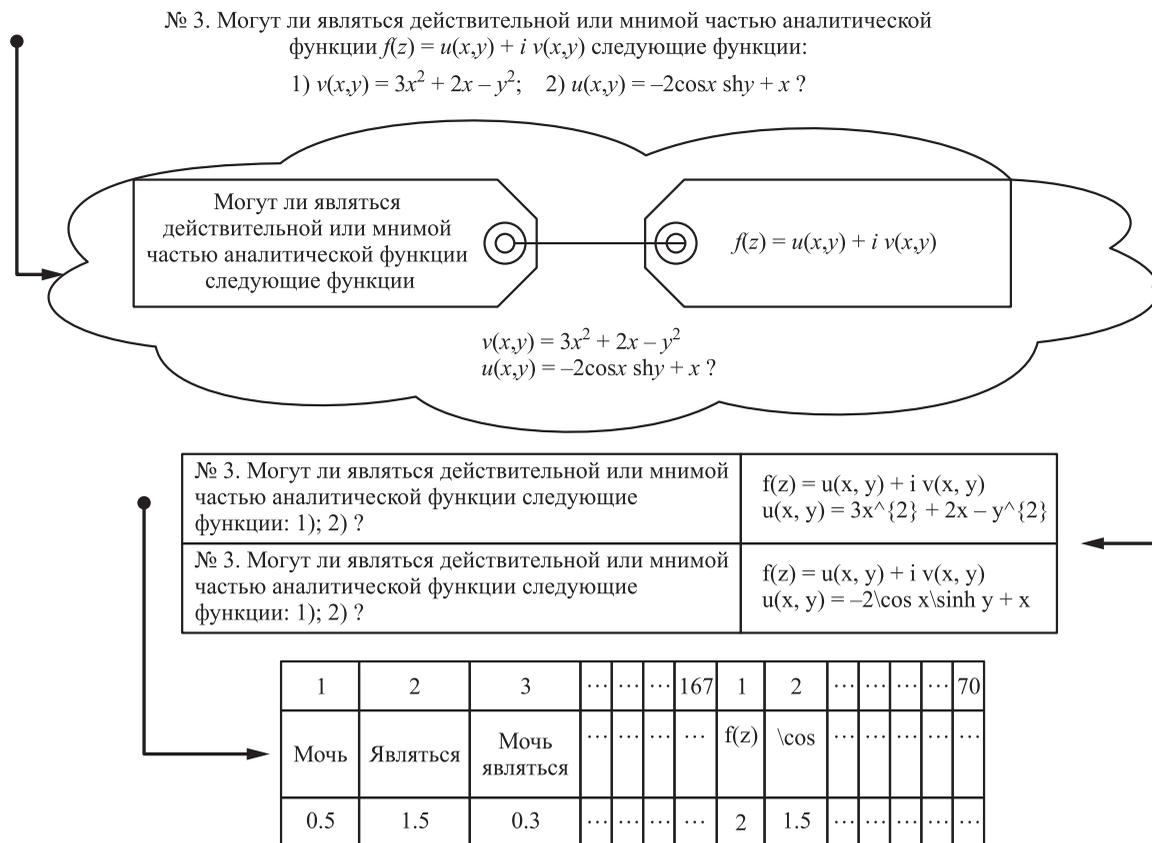


Рис. 2 Модель описания задач

для формирования словаря и векторов — библиотека scikit-learn, <https://scikit-learn.org/stable/>. Из формируемого списка токенов исключались пробелы, стоп-слова, знаки пунктуации, цифры. В качестве токенов использовались леммы слов. В состав словаря не включались слова, встречающиеся более чем в 70% описаний задач. Поскольку представление типа «мешок слов» нечувствительно к порядку слов в предложении, дополнительно задействовался механизм n -грамм. Наилучший результат дали биграммы, т.е. в словарь помимо слов включались и пары соседних слов.

Вектор признаков из формульного описания задачи также формировался методом «мешок слов», но большая часть операций с токенами потребовала существенной кастомизации, словарь и вектор-гистограмма с мерой tf-idf получались полностью аналогично текстовой части.

Для разбиения текста формул на токены уже на этапе предварительной подготовки использовалось регулярное выражение, которое подготовило формульное описание для выделения базовых токенов (см. разд. 2). Далее для формирования результирующего списка токенов из него исключались пробельные символы, знак «.», теги \TeX «\left»,

«\right», «\left.» и «\right.», которые влияют только на визуальное представление формулы. Кроме того, в процессе токенизации формул все элементы, соответствующие целым и вещественным числам, заменялись на токен «[number]», что позволило сократить объем словаря и повысить точность классификации.

В состав словаря включались все токены, вне зависимости от частоты их появления в описаниях. Попытка применить механизм n -грамм была сделана и для формульного описания задачи, но она не привела к успеху: результат классификации со словарем, содержащим n -граммы ($n = 2, \dots, 9$) оказался хуже. Также был опробован механизм формирования полиномиальных признаков, оказавшийся таким же нерезультативным.

Два вектора признаков конкатенируются (рис. 2) и передаются алгоритму обучения классификатора.

4 Результаты практической классификации

Сформированный в предыдущем разделе набор описаний задач использовался для обучения следующих классификаторов.

Анализ качества классификации

Метод	C_Text	Примечание	C_Form	Примечание	C_Comm	Примечание
Наивный байесовский	0,63	Раздел 3 $f_1 = 0,0$ Раздел 7 $f_1 = 0,18$	0,46	Разделы 3, 5, 6, 7 имеют $f_1 = 0,0$	0,76	Раздел 6 $f_1 = 0,29$
Логит	0,75	Раздел 3 $f_1 = 0,0$	0,6	Разделы 3, 6 имеют $f_1 = 0,0$	0,89	Раздел 6 $f_1 = 0,7$
Перцептрон	0,87	Раздел 8 имеет минимальную $f_1 = 0,71$	0,74	Раздел 6 имеет минимальную $f_1 = 0,5$	0,93	Раздел 6 $f_1 = 0,8$
Многослойная сеть	0,89	Раздел 8 имеет минимальную $f_1 = 0,71$	0,79	Раздел 7 имеет минимальную $f_1 = 0,57$	0,95	Раздел 7 $f_1 = 0,89$

1. Наивный байесовский классификатор [27, 28] — самая простая модель классификации, упрощающая исследуемую зависимость вход–выход предположением о независимости входных переменных. Поскольку классифицировались векторные данные, то применялся вариант классификатора, реализующий метод максимального правдоподобия в предположении полиномиального распределения входа. Кроме того, меры tf-idf сглаживались по методу Лапласа (описание реализации доступно по адресу https://scikit-learn.org/stable/modules/naive_bayes.html#multinomial-naive-bayes).
2. Логистическая регрессия, точнее мультиномиальная логистическая регрессия, или модель логит [29–32], возможно, самый распространенный метод классификации. Детали реализации есть по адресу https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression. Использовался вариант с l_2 -регуляризацией и приближенным алгоритмом оптимизации Бройдена–Флетчера–Гольдфарба–Шанно.
3. Однослойная нейронная сеть прямого распространения [33, 34] с пороговой функцией активации (перцептрон, https://scikit-learn.org/stable/modules/linear_model.html#perceptron).
4. Многослойная нейронная сеть прямого распространения [33,34] с двумя скрытыми слоями с числом нейронов 80 и 20 и функцией активации $f(x) = \max(0, x)$ (https://scikit-learn.org/stable/modules/neural_networks_supervised.html#neural-networks-supervised).

Имеющийся набор описаний задач разбивался случайным образом на две равные части: первые 100 задач использовались для обучения всех четырех классификаторов, оставшиеся задачи — для контроля качества классификации. Построенная модель задач дала размер словаря 167 для текстового описания и 70 — для формульного.

Качество классификации оценивается типовой метрикой «точность–отзыв» (precision–recall) [35],

т. е. гармоническим средним отношений результатов правильных классификаций к их сумме с числом ложноправильных результатов (точность) и отношений результатов правильных классификаций к их сумме с числом ложноотрицательных результатов (отзывчивость). Эта метрика, известная как f_1 -score, вычисляется для каждого из заданных классов классификации и общая — для всех разделов и всех 100 контрольных примеров. В приведенной таблице результаты расчетов характеризует общая метрика, а некоторые комментарии к величине f_1 -score для отдельных классов даны в форме примечаний. Суммарно выполнены три группы расчетов:

- (1) классификация только по текстовому описанию задачи (C_Text);
- (2) классификация только по формульному описанию задачи (C_Form);
- (3) классификация по общему описанию задачи (C_Comm).

В дополнение к этим расчетам выполнялся эксперимент по увеличению размера обучающей выборки, в котором 200 имеющихся задач распределялись так: 150 — для обучения; 50 — для контроля. Этот эксперимент дал ожидаемый результат — качество всех алгоритмов выросло на 10%–15%, в том числе многослойная сеть для модели C_Comm дала f_1 -score = 1. Но этот результат уже вызывает сомнения с точки зрения достоверности оценки качества и эскалирует проблему кардинального увеличения объема обучающей информации, которую в таких задачах решать практически невозможно.

5 Выводы

Основной полученный в статье результат — это доказательство принципиальной возможности автоматизировать классификацию математических задач и подтвердить принципиальную возможность практической интеграции такой технологии в действующие ЭОС. В качестве таковой авторам представляется система дистанционного обучения

CLASS.NET [36]. Анализируя результаты расчетов, можно отметить ожидаемое превосходство многослойной нейронной сети прямого распространения. Но важнее отметить тот факт, что приемлемый результат дает только модель описания задач, включающая оба набора признаков: и текстовый, и формульный.

Литература

1. *Боголепова С. В.* Анализ опыта обучения (в российском вузе) в онлайн-формате в период пандемии // Отечественная и зарубежная педагогика, 2021. Т. 1. № 4(77). С. 107–122.
2. *Кусаинов А. К.* Оценка качества современной учебной литературы // Ценности и смыслы, 2019. № 1(59). С. 8–19.
3. *Мартюшова Я. Г.* Теоретические основания конструирования электронных учебников для студентов технических университетов // Отечественная и зарубежная педагогика, 2018. Т. 1. № 5(54). С. 151–165.
4. *Босов А. В., Мартюшова Я. Г., Наумов А. В.* Выбор направлений оценивания качества электронных средств обучения для организации учебного процесса вуза // Сибирский педагогический ж., 2022. № 2. С. 54–63.
5. *Zanibbi R., Blostein D., Cordy J. R.* Recognizing mathematical expressions using tree transformation // IEEE T. Pattern Anal., 2002. Vol. 24. No. 11. P. 1455–1467.
6. *Blostein D., Grbavec A.* Recognition of mathematical notation // Handbook of character recognition and document image analysis / Eds. P. S. P. Wang, H. Bunke. — World Scientific Publishing Co., 1997. P. 557–582.
7. *Chan K. F., Yeung D. Y.* Mathematical expression recognition: A survey // Int. J. Doc. Anal. Recog., 2000. Vol. 3. No. 1. P. 3–15.
8. *Guidi F., Sacerdoti Coen C.* A survey on retrieval of mathematical knowledge // Mathematics Computer Science, 2016. Vol. 10. No. 4. P. 409–427.
9. *Knuth D. E.* The TeXbook. — Reading, MA, USA: Addison-Wesley, 1984. 483 p.
10. *Libbrecht P., Melis E.* Methods to access and retrieve mathematical content in ActiveMath // Congress (International) on Mathematical Software. — Berlin, Heidelberg: Springer, 2006. P. 331–342.
11. *Libbrecht P., Melis E.* Semantic search in LeActiveMath // 1st WebALT Conference and Exhibition Proceedings. — Eindhoven, Holland, 2006. P. 97–109.
12. *Miner R., Munavalli R.* An approach to mathematical search through query formulation and data normalization // Workshop (International) on Mathematical Knowledge Management. — Heidelberg: Springer, 2007. P. 342–355.
13. *Mišutka J., Galambos L.* System description: Egomath2 as a tool for mathematical searching on wikipedia.org // Conference (International) on Intelligent Computer Mathematics. — Berlin, Heidelberg: Springer, 2011. P. 307–309.
14. *Miller B. R., Youssef A.* Technical aspects of the digital library of mathematical functions // Ann. Math. Artif. Intel., 2003. Vol. 38. No. 1. P. 121–136.
15. *Sojka P., Liška M.* Indexing and searching mathematics in digital libraries // Conference (International) on Intelligent Computer Mathematics. — Berlin, Heidelberg: Springer, 2011. P. 228–243.
16. *Kohlhase M., Anca S., Jucovschi C., Palomo A. G., Sucan I. A.* MathWebSearch 0.4: A semantic search engine for mathematics, 2008. <http://mathweb.org/projects/mws/pubs/mkm08.pdf>.
17. *Hu X., Gao L. C., Lin X. Y., Zhi T., Lin X. F., Baker J. B.* Wikimirs: A mathematical information retrieval system for Wikipedia // 13th ACM/IEEE-CS Joint Conference on Digital Libraries Proceedings, 2013. P. 11–20.
18. *Liu H., Tian X., Tian B., Yang F., Li X.* An improved indexing and matching method for mathematical expressions based on inter-relevant successive tree // J. Computer Communications, 2016. Vol. 4. No. 15. P. 63–78.
19. *Tian X.* A mathematical indexing method based on the hierarchical features of operators in formulae // 2nd Conference (International) on Automatic Control and Information Engineering. — Atlantis Press, 2017. P. 49–52.
20. *Биряльцев Е. В., Гусенков А. М., Жибрик О. Н.* Некоторые подходы к разметке естественнонаучных текстов, содержащих математические выражения // Ученые записки Казанского университета. Сер. Физико-математические науки, 2014. Т. 156. № 4. С. 133–148.
21. *Gusenkov A., Gusenkova P., Palacheva Y., Zhibrik O.* Extended functionality of mathematical formulae search service // 12th Conference (International) on Advances in Semantic Processing / Eds. M. Spranger, P. Lorenz. — IARIA XPS Press, 2018. P. 35–41.
22. *Битюков Ю. И., Мартюшова Я. Г.* Решение задач по теории функций комплексного переменного. — М.: МАИ, 2022. 87 с.
23. *Краснов М. Л., Киселев А. И., Макаренко Г. И.* Функции комплексного переменного: задачи и примеры с подробными решениями. — М.: Либроком, 2012. 208 с.
24. *Краснов М. Л., Киселев А. И., Макаренко Г. И.* Операционное исчисление. Теория устойчивости: задачи и примеры с подробными решениями. — М.: Либроком, 2014. 176 с.
25. *McTear M. F., Callejas Z., Griol D.* The conversational interface. — Cham: Springer, 2016. 422 p.
26. *Salton G., McGill M. J.* Introduction to modern information retrieval. — New York, NY, USA: McGraw-Hill, 1983. 448 p.
27. *Minsky M.* Steps toward artificial intelligence // P. IRE, 1961. Vol. 49. No. 1. P. 8–30.
28. *McCallum A., Nigam K.* A comparison of event models for naive Bayes text classification // Workshop on Learning for Text Categorization Proceedings, 1998. Vol. 752. No. 1. P. 41–48.

29. Cox D. R. Some procedures connected with the logistic qualitative response curve // Research papers in probability and statistics / Ed. F. N. David. — London: Wiley, 1966. P. 55–71.
30. Theil H. A multinomial extension of the linear logit model // Int. Econ. Rev., 1969. Vol. 10. No. 3. P. 251–259.
31. Hosmer D. W., Lemeshow S. Applied logistic regression. — New York, NY, USA: Wiley, 1989. 307 p.
32. Hastie T., Tibshirani R., Friedman J. The elements of statistical learning. — 2nd ed. — New York, NY, USA: Springer, 2009. 533 p.
33. Галушкин А. И. Синтез многослойных систем распознавания образов. — М.: Энергия, 1974. 368 с.
34. Haykin S. Neural networks and learning machines. — 3rd ed. — Upper Saddle River, NJ, USA: Pearson Education, 2009. 906 p.
35. Van Rijsbergen C. J. Information retrieval. — 2nd ed. — Butterworth-Heinemann, 1979. 208 p.
36. Наумов А. В., Джумурат А. С., Иноземцев А. О. Система дистанционного обучения математическим дисциплинам CLASS.NET // Вестник компьютерных и информационных технологий, 2014. № 10. С. 36–44.

Поступила в редакцию 15.09.22

TECHNOLOGY FOR CLASSIFICATION OF CONTENT TYPES OF E-TEXTBOOKS

A. V. Bosov and A. V. Ivanov

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

Abstract: The problem of automatic classification of the educational content of the e-learning system, represented by tasks or practical examples, is being solved. A promising direction in the development of e-learning systems is the assessment of the quality of educational content. Carrying out such an assessment is the rationale for the need to create an automated classifier. The main idea is to model the content with an object with two properties — a textual description in natural language and a set of formulas in the language of scientific computer layout \TeX . Using tasks from the electronic textbook on the theory of functions of a complex variable, a data set was prepared and labeled in accordance with this model. Four text classification algorithms were trained — naive Bayes classifier, logistic regression, single-layer and multilayer feedforward neural networks. For these classifiers, a number of comparative experiments were carried out comparing the classification accuracy using text content only, formula content only, and the full model. As a result of the experiment, not only a formal comparison of the algorithms was carried out but also the fundamental advantage of the full model was shown. That is, when using both textual description and representation of formulas in the \TeX language, the classification accuracy significantly exceeds one-factor algorithms and confirms the readiness of the technology for practical application.

Keywords: e-learning system; training content; classification tasks and algorithms; content quality assessment; machine learning

DOI: 10.14357/19922264220410

Acknowledgments

The research was supported by the Russian Science Foundation (project No. 22-28-00588). The research was carried out using the infrastructure of the Shared Research Facilities “High Performance Computing and Big Data” (СКР “Informatics”) of FRC CSC RAS (Moscow).

References

1. Bogolepova, S. V. 2021. Analiz opyta obucheniya (v rossiyskom vuze) v onlayn-formate v period pandemii [Analysis of students’ online learning experience (at university in Russia) at the time of the pandemic]. *Otechestvennaya i zarubezhnaya pedagogika* [Domestic and Foreign Pedagogy] 1(4-77):107–122.
2. Kusainov, A. K. 2019. Otsenka kachestva sovremennoy uchebnoy literatury [Quality assessment of modern educational literature]. *Tsennosti i smysly* [Values and Meanings] 1(59):8–19.
3. Martyushova, Ya. G. 2018. Teoreticheskie osnovaniya konstruirovaniya elektronnykh uchebnikov dlya studentov tekhnicheskikh universitetov [Theoretical foundations of designing electronic textbooks for students of technical

- universities]. *Otechestvennaya i zarubezhnaya pedagogika* [Domestic and Foreign Pedagogy] 1(5-54):151–165.
4. Bosov, A. V., Ya. G. Martyushova, and A. V. Naumov. 2022. Vybory napravleniy otsenivaniya kachestva elektronnykh sredstv obucheniya dlya organizatsii uchebnogo protsesa vuza [Directions selection for assessing the quality of electronic learning tools for the organization of the educational process of the higher educational process]. *Sibirskiy pedagogicheskiy zh.* [Siberian Pedagogical J.] 2:54–63.
 5. Zanibbi, R., D. Blostein, and J. R. Cordy. 2002. Recognizing mathematical expressions using tree transformation. *IEEE T. Pattern Anal.* 24(11):1455–1467.
 6. Blostein, D., and A. Grbavec. 1997. Recognition of mathematical notation. *Handbook of character recognition and document image analysis*. Eds. P. S. P. Wang and H. Bunke. World Scientific Publishing Co. 557–582.
 7. Chan, K. F., and D. Y. Yeung. 2000. Mathematical expression recognition: A survey. *Int. J. Doc. Anal. Recog.* 3(1):3–15.
 8. Guidi, F., and C. Sacerdoti Coen. 2016. A survey on retrieval of mathematical knowledge. *Mathematics Computer Science* 10(4):409–427.
 9. Knuth, D. E. 1984. *The TeXbook*. Reading, MA: Addison-Wesley. 483 p.
 10. Libbrecht, P., and E. Melis. 2006. Methods to access and retrieve mathematical content in activemath. *Congress (International) on Mathematical Software Proceedings*. Berlin, Heidelberg: Springer. 331–342.
 11. Libbrecht, P., and E. Melis. 2006. Semantic search in LeActiveMath. *1st WebALT Conference and Exhibition Proceedings*. Eindhoven, Holland. 97–109.
 12. Miner R., and R. Munavalli. 2007. An approach to mathematical search through query formulation and data normalization. *Workshop (International) on Mathematical Knowledge Management Proceedings*. Heidelberg: Springer. 342–355.
 13. Mišutka, J., and L. Galamboš. 2011. System description: Egomath2 as a tool for mathematical searching on wikipedia.org. *Conference (International) on Intelligent Computer Mathematics Proceedings*. Berlin, Heidelberg: Springer. 307–309.
 14. Miller, B. R., and A. Youssef. 2003. Technical aspects of the digital library of mathematical functions. *Ann. Math. Artif. Intel.* 38(1):121–136.
 15. Sojka, P., and M. Liška. 2011. Indexing and searching mathematics in digital libraries. *Conference (International) on Intelligent Computer Mathematics Proceedings*. Berlin, Heidelberg: Springer. 228–243.
 16. Kohlhasse, M., S. Anca, S. Jucovschi, A. G. Palomo, and I. A. Sucan. 2008. MathWebSearch 0.4: A semantic search engine for mathematics. Available at: <http://mathweb.org/projects/mws/pubs/mkm08.pdf> (accessed November 17, 2022).
 17. Hu, X., L. C. Gao, X. Y. Lin, T. Zhi, X. F. Lin, and J. B. Baker. 2013. Wiki-mirs: A mathematical information retrieval system for Wikipedia. *13th ACM/IEEE-CS Joint Conference on Digital Libraries Proceedings*. 11–20.
 18. Liu, H., X. Tian, B. Tian, F. Yang, and X. Li. 2016. An improved indexing and matching method for mathematical expressions based on inter-relevant successive tree. *J. Computer Communications* 4(15):63–78.
 19. Tian, X. 2017. A mathematical indexing method based on the hierarchical features of operators in formulae. *2nd Conference (International) on Automatic Control and Information Engineering Proceedings*. Atlantis Press. 49–52.
 20. Biryaltsev, E. V., A. M. Gusenkov, and O. N. Zhibrik. 2014. Nekotorye podkhody k razmetke estestvennonauchnykh tekstov, sodержashchikh matematicheskie vyrazheniya [Some approaches to the markup of natural science texts containing mathematical expressions]. *Uchenye zapiski Kazanskogo universiteta. Ser. Fiziko-matematicheskie nauki* [Proceedings of Kazan University. Physics and mathematics ser.] 156(4):133–148.
 21. Gusenkov, A., P. Gusenkova, Y. Palacheva, and O. Zhibrik. 2018. Extended functionality of mathematical formulae search service. *12th Conference (International) on Advances in Semantic Processing*. Eds. M. Spranger and P. Lorenz. IARIA XPS Press. 35–41.
 22. Bityukov, Yu. I., and Ya. G. Martyushova. 2022. *Reshenie zadach po teorii funktsiy kompleksnogo peremennogo* [Solving problems on the theory of functions of a complex variable]. Moscow: MAI. 87 p.
 23. Krasnov, M. L., A. I. Kiselev, and G. I. Makarenko. 2012. *Funktsii kompleksnogo peremennogo: zadachi i primery s podrobnymi resheniyami* [Functions of a complex variable: Problems and examples with detailed solutions]. Moscow: Librokom. 208 p.
 24. Krasnov, M. L., A. I. Kiselev, and G. I. Makarenko. 2012. *Operatsionnoe ischislenie. Teoriya ustoychivosti: zadachi i primery s podrobnymi resheniyami* [Operational calculus. Stability theory: Problems and examples with detailed solutions]. Moscow: Librokom. 176 p.
 25. McTear, M. F., Z. Callejas, and D. Griol. 2016. *The conversational interface*. Cham: Springer. 422 p.
 26. Salton, G., and M. J. McGill. 1983. *Introduction to modern information retrieval*. New York, NY: McGraw-Hill. 448 p.
 27. Minsky, M. 1961. Steps toward artificial intelligence. *P. IRE* 49(1):8–30.
 28. McCallum, A., and K. Nigam. 1998. A comparison of event models for naive bayes text classification. *Workshop on Learning for Text Categorization Proceedings*. 752(1):41–48.
 29. Cox, D. R. 1966. Some procedures connected with the logistic qualitative response curve. *Research papers in probability and statistics*. Ed. F. N. David. London: Wiley. 55–71.
 30. Theil, H. 1969. A multinomial extension of the linear logit model. *Int. Econ. Rev.* 10(3):251–259.
 31. Hosmer, D. W., and S. Lemeshow. 1989. *Applied logistic regression*. New York, NY: Wiley. 307 p.

32. Hastie, T., R. Tibshirani, and J. Friedman. 2009. *The elements of statistical learning*. 2nd ed. New York, NY: Springer. 533 p.
33. Galushkin, A. I. 1974. *Sintez mnogosloynnykh sistem raspoznavaniya obrazov* [Synthesis of multilayer image recognition systems]. Moscow: Energiya. 368 p.
34. Haykin, S. 2009. *Neural networks and learning machines*. 3rd ed. Upper Saddle River, NJ: Pearson Education. 906 p.
35. Van Rijsbergen, C. J. 1979. *Information retrieval*. 2nd ed. Butterworth-Heinemann. 208 p.
36. Naumov, A. V., A. S. Dzhumurat, and A. O. Inozemtsev. 2014. Sistema distantsionnogo obucheniya matematicheskimi distsiplinami CLASS.NET [Distance learning system for mathematical disciplines CLASS.NET]. *Vestnik komp'yuternykh i informatsionnykh tekhnologiy* [Herald of Computer and Information Technologies] 10:36–44.

Received September 15, 2022

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

Ivanov Alexey V. (b. 1976) — 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; AIvanov@ipiran.ru

О НАУЧНОЙ ПАРАДИГМЕ ИНФОРМАТИКИ: ВЕРХНИЙ УРОВЕНЬ КЛАССИФИКАЦИИ ОБЪЕКТОВ ЕЕ ПРЕДМЕТНОЙ ОБЛАСТИ*

И. М. Зацман¹

Аннотация: Рассматривается подход А. Соломоника к структурированию научной парадигмы «зрелой» науки. Согласно его подходу, описание такой науки должно включать четыре составляющих (философские основания; аксиоматика; классификация исследуемых объектов; система терминов), которые могут разрабатываться отдельно, но объединяются в единую и цельную конструкцию. В рамках этого подхода предлагается начать описание парадигмы информатики с уточнения ее позиционирования в системе научного познания и построения верхнего уровня классификации исследуемых в ней объектов. Для позиционирования информатики предлагается развитие идеи Деннинга и Розенблюма о группировке научных дисциплин в четырех отраслях знания. Для построения верхнего уровня классификации используется идея Кристена Ньюгора о различении в предметной области информатики объектов ментальной природы (концепты знания человека) и сенсорно воспринимаемых объектов. Цель статьи состоит в попытке начать описание научной парадигмы информатики на основе подхода А. Соломоника и развития идей Деннинга, Ньюгора и Розенблюма с построения верхнего уровня классификации объектов, исследуемых в информатике.

Ключевые слова: научная парадигма; составляющие научной парадигмы; система научного познания; классификация объектов информатики

DOI: 10.14357/19922264220411

1 Введение

В работах Деннинга, Розенберга и Кари [1–4] дан анализ основных этапов эволюции информатики, на протяжении которых расширялась ее предметная область по сравнению с первоначальной трактовкой ее содержания². Кроме компьютерных кодов и кодируемых ими информационных объектов, например текстов, в информатике предметом исследования стали информационные процессы, происходящие в живой природе. По словам Деннинга, «первоначальное определение информатики [как компьютерной науки]. . . сейчас устарело. Она изучает и естественные, и искусственные информационные процессы» [1]. Появление работ по изучению влияния информационных систем на социум еще больше расширило предметную область информатики [6], и эта область стала охватывать широкий спектр феноменов разной природы, связанных с информационными процессами в технических, живых и социальных системах.

Проблема описания научной парадигмы информатики усложнялась с ростом многообразия объек-

тов ее расширяющейся предметной области. Это расширение было во многом обусловлено разработкой и применением технологий, охватывающих информационные процессы в технических, живых и социальных системах, что было отмечено в отчете «Глубинное изменение — технологические переломные моменты и социальное воздействие» [7]³.

В качестве исходной позиции для создания научной парадигмы информатики, учитывающей расширение ее предметной области, в статье предлагается начать описание этой парадигмы с позиционирования информатики в системе научного познания и построения верхнего уровня классификации исследуемых в ней объектов. Для позиционирования информатики предлагается развить идею Деннинга и Розенблюма о группировке научных дисциплин в четырех отраслях знания. Для построения верхнего уровня классификации используется идея Кристена Ньюгора о различении в предметной области информатики объектов ментальной природы (концепты знания человека) и сенсорно воспринимаемых объектов. Цель статьи

* Исследование выполнено с использованием ЦКП «Информатика» ФИЦ ИУ РАН

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

² В 1967 г. А. Ньюэлл, А. Дж. Перлис и Х. А. Саймон написали в журнале *Science*: «Объектам и явлениям соответствуют те науки, которые их изучают. Появились компьютеры. Следовательно, назначение информатики — это изучение компьютеров» [5]. Такой подход к определению ее содержания доминировал долгие годы в информатике как компьютерной науке.

³ Данный отчет был подготовлен под эгидой Всемирного экономического форума (Давос, Швейцария). В подготовке материалов, использованных в отчете, приняли участие около 800 экспертов и руководителей отрасли информационных и коммуникационных технологий. Документ содержит перечень новых информационных технологий, которые определяют кардинальный характер преобразования общества и экономики, получившего название «Четвертая промышленная революция» [8].

состоит в попытке начать описание научной парадигмы информатики на основе подхода А. Соломоники к структурированию научной парадигмы «зрелой» науки и развития идей Деннинга, Ньюгора и Розенблюма.

2 Информатика в системе научного познания

Согласно А. Соломонике, научная парадигма «зрелой» науки состоит из следующих четырех составляющих, которые могут разрабатываться отдельно, но объединяются в единую и цельную конструкцию [9, с. 23–24]:

- (1) философские основания;
- (2) аксиоматика;
- (3) классификация исследуемых объектов, процессов и явлений;
- (4) система терминов.

Сам термин «научная парадигма» трактуется им в соответствии с теорией Т. Куна, которая описывает процесс смены научных парадигм [10]. При этом А. Соломоник отмечает тот факт, что в книге Куна мы не находим ответа на вопрос: «Из чего должна состоять парадигма любой «зрелой» науки?» [9, с. 23].

О необходимости смены научной парадигмы информатики в 2011 г. написал Марк Смир, обосновывая это тем, что быстрая эволюция информационных технологий (ИТ) стимулирует периодический пересмотр теоретического фундамента их разработки и что именно сейчас настало время провести очередной пересмотр. Он пишет, что «раньше, когда с компьютерами имел дело ограниченный круг лиц, можно было рассчитывать на то, что те, кто профессионально общаются с компьютерами, могут к ним адаптироваться. В наши дни, когда значительно большее число людей взаимодействуют с цифровыми устройствами и информационными системами, ИТ тесно вплетены во многие когнитивные и социальные процессы. В этих условиях нельзя игнорировать эти процессы при создании ИТ» [11].

Наиболее радикальная попытка смены парадигмы информатики была предложена в 2013 г. Полом Розенблюмом [12]: она основывалась на новом позиционировании информатики в системе научного познания ($\text{united science} = \text{science} \cup \text{humanities}$), предложенном Деннингом и Розенблюмом в 2009 г. [13]. В предлагаемом ими позиционировании информатики рассматриваются только естественные науки (*science*). Они группируются

в три естественно-научные отрасли: физические науки (куда отнесены химия, геология и другие науки, изучающие неживую материю), науки о жизни и социальные науки. Математику Деннинг и Розенблюм не относят к естественным наукам. Для целей данной статьи добавим математику в первую отрасль и назовем ее «физико-математические науки». Так как методы и средства информатики широко используются и в гуманитарных науках (*humanities*), то их добавим в третью отрасль и назовем ее «социогуманитарные науки».

При таком подходе к группировке наук в рамках системы научного познания информатика плохо вписывается в три перечисленные отрасли (это утверждение остается справедливым и после добавления в них математики и гуманитарных наук). Хотя методы и средства информатики широко используются в системе научного познания, ни одна из этих отраслей не изучает информационные трансформации как таковые: информация не является в общем случае ни физико-математической, ни живой, ни социогуманитарной сущностью. Поэтому авторы работы [13] и предлагают рассматривать информатику как четвертую отрасль, тесно взаимодействующую с тремя остальными.

Развивая идеи, изложенные в [13], Розенблюм в своей монографии [12] предложил обозначить четыре отрасли науки следующими литерами, которые будем использовать с учетом расширения первой и четвертой отраслей: физико-математические науки (P), науки о жизни (L), социогуманитарные науки (S) и информатика (C). Далее он рассматривает информационные трансформации, охватывающие две, три или четыре отрасли, включая информатику, обозначая соответствующие им предметные области сочетаниями этих литер и давая им следующие названия: диадический компьютеринг (C + P, C + L, C + S); триадический компьютеринг (C + P + S, C + P + L, C + L + S); тетрадический компьютеринг (P + S + L + C), который в данной статье трактуется как синоним информатики.

Фундаментальное расширение предметной области информатики, предложенное Розенблюмом, включает не только информационные трансформации, охватывающие все отрасли науки, но и феномены разной природы, связанные с этими трансформациями. Эта связанность иллюстрируется им на примере интерфейса «мозг–компьютер» следующим образом [12]: «Для описания взаимосвязей между мозгом и компьютером, нам необходимо рассмотреть информационные трансформации, относящиеся ко всей предметной области тетрадического компьютеринга (P + S + L + C). Эти интерфейсы включают не только трансформации элект-

трических потенциалов, генерируемых мозгом (L), и формирование компьютерных кодов (C), но также физические устройства (P), которые обеспечивают преобразование электрических потенциалов в коды компьютера, плюс концепты знания человека (S), представленные электрическими потенциалами, которые в конечном итоге и будут использоваться для управления компьютером или другими устройствами».

3 Верхний уровень классификации

Если исходить из идеи Пола Розенблюма о смене парадигмы информатики и предложенного ее развития на основе деления всей системы научного познания (united science) на четыре отрасли, с неизбежностью приходится констатировать, что феномены одной и той же природы изучаются в разных научных отраслях и дисциплинах, а классификации объектов и парадигмы их исследования могут существенно отличаться [14]. Например, концепты знания человека изучаются в задачах представления знания в информатике [15–17], они являются объектами исследований в науках о жизни [18] и социогуманитарных науках [19]. При этом они относятся к среде одной и той же природы: ментальной среде знания человека. Поэтому описание верхнего уровня классификации исследуемых в информатике объектов как феноменов разной природы было предложено начать с выделения в ее предметной области сред, включающих объекты одной и той же природы [20].

Отметим, что еще в 1986 г. Кристен Ньюгор в предметной области информатики стал различать объекты ментальной природы (концепты знания человека) и сенсорно воспринимаемые объекты (например, тексты на естественных языках) [21], что можно считать прообразом предлагаемого выделения сред из предметной области информатики как тетрадического компьютеринга, где каждая из сред включает объекты одной природы (так, ментальная среда включает концепты знания человека, цифровая — компьютерные коды, информационная — сенсорно воспринимаемые объекты, например тексты на естественных языках, таблицы, графики и рисунки). Кристен Ньюгор в 1986 г. предложил следующую дефиницию термина «информатика» [21]: «Информатика — это наука, которая имеет своей областью [исследований] информационные процессы и связанные с ними феномены в артефактах, обществе и природе».

¹ Например, модели трансляции естественных ДНК, созданные микробиологами, используются при разработке методов записи и хранения данных с использованием синтезированных цепочек ДНК.

Затем Ньюгор дал расширенное толкование термина «феномен» [21]: «Важными примерами феноменов являются: живые организмы, неодушевленные объекты (включая артефакты, например машины), события и процессы (например, выполнение компьютерных программ). Мы также можем говорить о когнитивных феноменах, происходящих в сознании людей, в отличие от явных [сенсорно воспринимаемых] феноменов, находящихся вне сознания».

В развитие идеи Ньюгора совокупность когнитивных феноменов, формируемых в процессах познания, происходящих в сознании людей (далее — концепты), предлагается назвать *ментальной средой* предметной области информатики. Совокупность сенсорно воспринимаемых феноменов, находящихся вне сознания, предлагается назвать *информационной средой* предметной области информатики. Кроме того, предлагается выделить еще как минимум три среды:

- (1) *цифровая среда* — совокупность компьютерных кодов;
- (2) *нейросреда* электрических потенциалов и магнитных полей, генерируемых мозгом, которые используются в ИТ управления роботизированной рукой [22] и других ИТ на основе интерфейсов «мозг—компьютер»;
- (3) *ДНК-среда* — совокупность цепочек РНК и ДНК¹.

Согласно перечисленным средам, выделенным из предметной области информатики как тетрадического компьютеринга, на верхнем уровне классификации объектов ее исследований предлагается указать как минимум пять сред, каждая из которых включает объекты одной природы: *ментальная, информационная, цифровая, нейро- и ДНК-среда*. При этом с ростом разнообразия природы объектов предметной области информатики верхний уровень классификации может пополняться новыми средами, природа объектов которых отличается от природы сред, ранее включенных в этот уровень классификации, если при проектировании ИТ встретятся объекты, которые по своей природе не относятся ни к одной из ранее уже выделенных сред [17]. В будущей парадигме информатики его предлагается сделать открытым, что обусловлено разнообразием природы ее объектов и возможным включением в ее предметную область объектов новой природы (открытость уровня классификации). Отметим, что разнообразие природы объектов предметной области информатики служит объектом пристального изучения в сфере компьютерного образования [23].

Разнообразие природы объектов, сложившееся к настоящему времени, существенно усложняет разработку стратегии преподавания информатики на всех уровнях образования в части описания ее теоретических оснований. Например, в стратегии Европейского компьютерного образования не описана ни ее предметная область, ни классификация объектов ее исследований [24]. Отсутствие такого описания мотивируется его большей сложностью по сравнению с естественными науками [25, с. 7]: «В то время как естественные науки определяются применительно к миру, в котором мы живем, информатику как научную дисциплину определить сложнее; у нее нет эмпирических основ, как у естественных наук; это нечто большее, чем использование только логических выводов (как в математике); и это далеко не только сочетание инженерных принципов и технологий».

Описание сред разной природы на верхнем уровне классификации объектов исследований в информатике дает возможность увидеть весь спектр теоретически возможных интерфейсов между объектами разной природы и при проектировании ИТ, и в процессе преподавания информатики. В работе [20] дано описание 11 видов интерфейсов разного порядка: второго (6 видов), третьего (4 вида) и четвертого (1 вид), которые могут использоваться в ИТ, охватывающих объекты четырех сред: ментальной, информационной, цифровой и нейросреды.

Из 11 видов интерфейсов, которые теоретически могут встретиться при проектировании ИТ, охватывающих сущности этих четырех сред, наиболее широко используется интерфейс второго порядка на границе между *информационной и цифровой средами*. В информационных системах применяются кодовые таблицы, например Unicode, которые служат наиболее распространенным способом *симметричной реализации* этого интерфейса, т. е. каждому символу соответствует один компьютерный код и наоборот. Однако есть примеры и *асимметричной реализации* этого интерфейса [26, с. 222–227].

Теоретически есть еще пять интерфейсов второго порядка на границах между следующими средами:

- (1) ментальной и информационной;
- (2) ментальной и цифровой;
- (3) ментальной и нейросредой;
- (4) информационной и нейросредой;
- (5) цифровой и нейросредой.

Первый из вышеперечисленных интерфейсов (между ментальной и информационной средами) реализуется знаковыми системами, которые далеко

не всегда обеспечивают симметричную его реализацию. Известно, что вербальным знаковым системам естественных языков свойственна асимметрия, когда одно слово имеет несколько смысловых значений (омонимия и полисемия), а одно значение может быть выражено разными словами (синонимия) [27, с. 47]. Асимметрия вербальных знаковых систем существенно усложняет разработку систем обработки текстов на естественном языке и проектирование ИТ представления знания в информационных системах из-за необходимости разрешения лексической неоднозначности [28–30].

4 Заключение

Согласно А. Соломонику, описание научной парадигмы «зрелой» науки должно включать четыре составляющих (философские основания; аксиоматика; классификация исследуемых объектов, процессов и явлений; система терминов). В статье рассмотрена только одна составляющая (классификация) и дано описание только верхнего уровня классификации объектов информатики. При этом использовалось позиционирование информатики в системе научного познания, которое развивает идею Деннинга и Розенблюма о группировке научных дисциплин в четырех отраслях знания, так как объединяет математику, естественные и гуманитарные науки.

Предлагаемый вариант формирования верхнего уровня классификации объектов информатики рассмотрен на примере пяти сред разной природы, выделенных из ее предметной области: ментальной, информационной, цифровой, нейро- и ДНК-сред. Верхний уровень классификации объектов исследований в информатике дает возможность увидеть весь спектр теоретически возможных интерфейсов между объектами разной природы и при проектировании ИТ, и в процессе преподавания информатики. Для случая четырех сред разной природы он включает 11 видов интерфейсов второго, третьего и четвертого порядка для ИТ, охватывающих объекты этих сред. Если в будущем появится необходимость в проектировании ИТ, охватывающих объекты шести или более видов разной природы, то предлагаемый подход к выделению сред покажет весь спектр интерфейсов, которые теоретически может понадобиться реализовать в этих ИТ.

В заключение отметим, что после построения верхнего уровня классификации объектов информатики планируется дать дефиниции тех терминов, которые использовались при его описании, и затем построить следующий уровень классификации, выделяя из каждой среды составляющие ее домены.

Например, для цифровой среды планируется описать четыре домена (электрический, магнитный, оптический и генетический, содержащий *синтезированные цепочки ДНК*), а также интерфейсы между ними. Также планируется дать повторное описание верхнего уровня классификации на основе другого подхода к объединению научных дисциплин — в три отрасли знания: естественные науки, социальные и гуманитарные науки, формальные науки.

Литература

1. *Denning P.* Computing is a natural science // Commun. ACM, 2007. Vol. 50. No. 7. P. 13–18.
2. *Rozenberg G.* Computer science, informatics, and natural computing — personal reflections // New computational paradigms: Changing conceptions of what is computable / Eds. S. B. Cooper, B. Löwe, A. Sorbi. — New York, NY, USA: Springer Science + Business Media LLC, 2008. P. 373–379.
3. *Kari L., Rozenberg G.* The many facets of natural computing // Commun. ACM, 2008. Vol. 51. No. 10. P. 72–83.
4. *Denning P.* The science in computer science // Commun. ACM, 2013. Vol. 56. No. 5. P. 35–38.
5. *Newell A., Perlis A., Simon H.* Computer science // Science, 1967. Vol. 157. No. 3795. P. 1373–1374.
6. *Nygaard K., Håndlykken P.* The system development process — its setting, some problems and needs for methods // Software Engineering Environments Symposium Proceedings. — Amsterdam, 1981. P. 157–172.
7. Deep shift: Technology tipping points and societal impact. — Geneva, Switzerland: World Economic Forum, 2015. http://www3.weforum.org/docs/WEF_GAC15_Technological_Tipping_Points_report_2015.pdf.
8. *Шваб К.* Четвертая промышленная революция / Пер. с англ. — М.: Эксмо, 2018. 288 с. (*Schwab K.* The fourth industrial revolution. — Geneva, Switzerland: World Economic Forum, 2016. 172 p.)
9. *Соломоник А.* Парадигма семиотики. — Минск: МЕТ, 2006. 335 с.
10. *Кун Т.* Структура научных революций / Пер. с англ. — М.: Прогресс, 1977. 302 с. (*Kuhn T.* The structure of scientific revolutions. — Chicago, IL, USA: University of Chicago Press, 1962. 264 p.)
11. *Snir M.* Computer and information science and engineering: One discipline, many specialties // Commun. ACM, 2011. Vol. 54. No. 3. P. 38–43.
12. *Rosenbloom P. S.* On computing: The fourth great scientific domain. — Cambridge, MA, USA: MIT Press, 2013. 308 p.
13. *Denning P., Rosenbloom P.* Computing: The fourth great domain of science // Commun. ACM, 2009. Vol. 52. No. 9. P. 27–29.
14. *Tedre M., Pajunen J.* Grand theories or design guidelines? Perspectives on the role of theory in computing education research // ACM T. Comput. Educ., 2022 (in press). doi: 10.1145/3487049.
15. *Зацман И. М., Косарик В. В., Курчавова О. А.* Задачи представления личностных и коллективных концептов в цифровой среде // Информатика и её применения, 2008. Т. 2. Вып. 3. С. 54–69.
16. *Zatsman I.* Tracing emerging meanings by computer: Semiotic framework // 13th European Conference on Knowledge Management Proceedings. — Reading, U.K.: Academic Publishing International Ltd., 2012. Vol. 2. P. 1298–1307.
17. *Зацман И. М.* Таблица интерфейсов информатики как информационно-компьютерной науки // Научно-техническая информация. Сер. 1: Организация и методика информационной работы, 2014. № 11. С. 1–15.
18. *Baars B., Gage N.* Cognition, brain, and consciousness: Introduction to cognitive neuroscience. — Amsterdam, Netherlands: Academic Press/Elsevier, 2010. 677 p.
19. *Eco U.* A theory of semiotics. — Bloomington, IL, USA: Indiana University Press, 1976. 356 p.
20. *Зацман И. М.* Теоретические основания компьютерного образования: среды предметной области информатики как основание классификации ее объектов // Системы и средства информатики, 2022. Т. 32. № 4. С. 77–89.
21. *Nygaard K.* Program development as a social activity // Information Processing: 10th World Computer Congress Proceedings / Ed. H.-J. Kugler. — Amsterdam, Netherlands: Elsevier Science Pubs. B. V. (North Holland), IFIP, 1986. P. 189–198.
22. *Зацман И. М.* Интерфейсы третьего порядка в информатике // Информатика и её применения, 2019. Т. 13. Вып. 3. С. 82–89.
23. *Kaya E., Newley A., Yesilyurt E., Deniz H.* Nature of computer science: Identification of K-12 accessible nature of computer science tenets and development of an open-ended nature of computer science instrument // 17th Conference on International Computing Education Research Proceedings. — New York, NY, USA: ACM, 2021. P. 426. doi: 10.1145/3446871.3469784.
24. *Caspersen M. E., Gal-Ezer J., McGettrick A., Nardelli E.* Informatics for all: The strategy. — New York, NY, USA: ACM, 2018. 16 p.
25. Committee on European Computing Education. Informatics education in Europe: Are we all in the same boat? — New York, NY, USA: ACM, 2017. Technical Report. 251 p. doi: 10.1145/3106077.
26. *Зацман И. М.* Концептуальный поиск и качество информации. — М.: Наука, 2003. 272 с.
27. Лингвистический энциклопедический словарь / Под ред. В. Н. Ярцевой. — М.: Советская энциклопедия, 1990. 685 с.
28. *Bolshina A., Loukachevitch N.* Generating training data for word sense disambiguation in Russian // Компьютерная лингвистика и интеллектуальные технологии:

- По мат-лам ежегодной Междунар. конф. «Диалог». — М.: РГГУ, 2020. Вып. 19(26). С. 119–132.
29. Bolshina A., Loukachevitch N. All-words word sense disambiguation for Russian using automatically generated text collection // *Cybernetics Information Technologies*, 2020. Vol. 20. No. 4. С. 90–107.
30. Bolshina A., Loukachevitch N. Automatic labelling of genre-specific collections for word sense disambiguation in Russian // *Artificial intelligence / Eds. S. O. Kuznetsov, A. I. Panov, K. S. Yakovlev. — Lecture notes in computer science ser. — Cham, Switzerland: Springer, 2020. Vol. 12412. P. 215–227.*

Поступила в редакцию 14.10.22

ON THE SCIENTIFIC PARADIGM OF INFORMATICS: THE CLASSIFICATION HIGH LEVEL OF ITS OBJECTS

I. M. Zatsman

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

Abstract: The approach of A. Solomonik to structuring the scientific paradigm of “mature” science is considered. According to his approach, the description of such a science should include four components (philosophical foundations; axiomatics; classification of its objects; and system of terms) which can be developed separately but combined into a single and integral structure. Within the framework of this approach, it is proposed to begin the description of the paradigm of informatics by clarifying its positioning in united science (= science \cup humanities) and constructing the classification high level of its objects. To position informatics, it is proposed to develop the idea of Denning and Rosenbloom about grouping scientific disciplines in Four Great Scientific Domains. To build the high level of classification, Kristen Nygaard’s idea of distinguishing objects of mental nature (concepts of human knowledge) and sensory-perceived objects is used. The purpose of the paper is to attempt to begin the description of the scientific paradigm of informatics based on the approach of A. Solomonik and the development of the ideas of Denning, Nygaard, and Rosenbloom with the construction of the high level of classification.

Keywords: scientific paradigm; scientific paradigm components; united science; classification informatics objects

DOI: 10.14357/19922264220411

Acknowledgments

The research was carried out using infrastructure of shared research facilities CKP “Informatics” of FRC CSC RAS.

References

1. Denning, P. 2007. Computing is a natural science. *Commun. ACM* 50(7):13–18.
2. Rozenberg, G. 2008. Computer science, informatics, and natural computing — personal reflections. *New computational paradigms: Changing conceptions of what is computable*. Eds. S. B. Cooper, B. Löwe, and A. Sorbi. New York, NY: Springer Science + Business Media LLC. 373–379.
3. Kari, L., and G. Rozenberg. 2008. The many facets of natural computing. *Commun. ACM* 51(10):72–83.
4. Denning, P. 2013. The science in computer science. *Commun. ACM* 56(5):35–38.
5. Newell, A., A. Perlis, and H. Simon. 1967. Computer science. *Science* 157(3795):1373–1374.
6. Nygaard, K., and P. Håndlykken. 1981. The system development process — its setting, some problems and needs for methods. *Software Engineering Environments Symposium Proceedings*. Amsterdam. 157–172.
7. Deep shift: Technology tipping points and societal impact. 2015. World Economic Forum. Geneva, Switzerland. Available at: http://www3.weforum.org/docs/WEF_GAC15_Technological_Tipping_Points_report_2015.pdf (accessed October 31, 2022).
8. Schwab, K. 2016. *The fourth industrial revolution*. Geneva, Switzerland: World Economic Forum. 172 p.
9. Solomonik, A. 2006. *Paradigma semiotiki* [The paradigm of semiotics]. Minsk: MET Pubs. 335 p.
10. Kuhn, T. 1962. *The structure of scientific revolutions*. Chicago, IL: University of Chicago Press. 264 p.
11. Snir, M. 2011. Computer and information science and engineering: One discipline, many specialties. *Commun. ACM* 54(3):38–43.
12. Rosenbloom, P. 2013. *On computing: The fourth great scientific domain*. Cambridge, MA: MIT Press. 308 p.

13. Denning, P., and P. Rosenbloom. 2009. Computing: The fourth great domain of science. *Commun. ACM* 52(9): 27–29.
14. Tedre, M., and J. Pajunen. 2022 (in press). Grand theories or design guidelines? Perspectives on the role of theory in computing education research. *ACM T. Comput. Educ.* doi: 10.1145/3487049.
15. Zatsman, I., V. Kosarik, and O. Kurchavova. 2008. Zadachi predstavleniya lichnostnykh i kollektivnykh konseptov v tsifrovoy srede [Problems of representation of personal and collective concepts in the digital medium]. *Informatika i ee Primeneniya — Inform. Appl.* 2(3):54–69.
16. Zatsman, I. 2012. Tracing emerging meanings by computer: Semiotic framework. *13th European Conference on Knowledge Management Proceedings*. Reading, U.K.: Academic Publishing International Ltd. 2:1298–1307.
17. Zatsman, I. 2014. Table of interfaces of informatics as computer and information science. *Scientific Technical Information Processing* 41(4):233–246.
18. Baars, B., and N. Gage. 2010. *Cognition, brain, and consciousness: Introduction to cognitive neuroscience*. Amsterdam: Academic Press/Elsevier. 677 p.
19. Eco, U. 1976. *A theory of semiotics*. Bloomington, IL: Indiana University Press. 356 p.
20. Zatsman, I. 2022. Teoreticheskie osnovaniya komp'yuternogo obrazovaniya: sredi predmetnoy oblasti informatiki kak osnovanie klassifikatsii ee ob"ektov [Theoretical foundations of digital education: Subject domain media of informatics as the base of its objects' classification]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 32(4):77–89.
21. Nygaard, K. 1986. Program development as a social activity. *Information Processing: 10th World Computer Congress Proceedings*. Ed. H.-J. Kugler. Amsterdam, Netherlands: Elsevier Science Publs. B. V. (North Holland), IFIP. 189–198.
22. Zatsman, I. M. 2019. Interfeysy tret'ego poryadka v informatike [Third-order interfaces in informatics]. *Informatika i ee Primeneniya — Inform. Appl.* 13(3):82–89.
23. Kaya, E., A. Newley, E. Yesilyurt, and H. Deniz. 2021. Nature of computer science: Identification of K-12 accessible nature of computer science tenets and development of an open-ended nature of computer science instrument. *17th Conference on International Computing Education Research Proceedings*. New York, NY: ACM. 426. doi: 10.1145/3446871.3469784.
24. Caspersen, M. E., J. Gal-Ezer, A. McGettrick, and E. Nardelli. 2018. *Informatics for all: The strategy*. New York, NY: ACM. 16 p.
25. Committee on European Computing Education. 2017. Informatics education in Europe: Are we all in the same boat? New York, NY: ACM. Technical Report. 251 p. doi: 10.1145/3106077.
26. Zatsman, I. 2003. *Kontseptual'nyy poisk i kachestvo informatsii* [Conceptual retrieval and quality of information]. Moscow: Nauka. 272 p.
27. Yartseva, V. N., ed. 1990. *Lingvisticheskiy entsiklopedicheskiy slovar'* [Linguistic encyclopedic dictionary]. Moscow: Soviet Encyclopedia. 685 p.
28. Bolshina, A., and N. Loukachevitch. 2020. Generating training data for word sense disambiguation in Russian. *Computer Linguistic and Intellectual Technologies: Conference (International) "Dialog" Proceedings*. Moscow. 119–132.
29. Bolshina, A., and N. Loukachevitch. 2020. All-words word sense disambiguation for Russian using automatically generated text collection. *Cybernetics Information Technologies* 20(4):90–107.
30. Bolshina, A., and N. Loukachevitch. 2020. Automatic labelling of genre-specific collections for word sense disambiguation in Russian. *Artificial intelligence*. Eds. S. O. Kuznetsov, A. I. Panov, and K. S. Yakovlev. Lecture notes in computer science ser. Cham, Switzerland: Springer. 12412:215–227.

Received October 14, 2022

Contributor

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

МОДЕЛЬ И ТЕХНОЛОГИЯ ИЗВЛЕЧЕНИЯ НОВЫХ ТЕРМИНОВ ИЗ МЕДИЦИНСКИХ ТЕКСТОВ*

И. М. Зацман¹, О. В. Золотарев², А. Х. Хакимова³, Гу Дунсяо⁴

Аннотация: Рассматривается модель информационной технологии (ИТ) извлечения новых терминов из медицинских текстов, которая относится к ранее определенному классу средовых моделей информатики. В проведенном эксперименте для определения новизны терминов используется словарь MeSH (Medical Subject Headings), который создан и обновляется Национальной медицинской библиотекой США. Появление новых терминов обусловлено представлением в медицинских статьях и других научных текстах нового знания об исследуемых болезнях, методах их лечения и применяемых медикаментах, которое еще не нашло отражения в медицинских словарях и тезаурусах. В информационных системах медицинских учреждений и институтов предлагаемая технология позволяет регулярно актуализировать новыми терминами профили исследуемых болезней, соответствующих их предметной области. Цель статьи состоит в описании средовой модели ИТ актуализации терминологических профилей болезней.

Ключевые слова: средовые модели информатики; медицинские тексты; терминологический профиль; извлечение новых терминов из текстов

DOI: 10.14357/19922264220412

1 Введение

Одна из задач российско-китайского проекта, финансируемого по гранту № 21-57-53018, состоит в моделировании и создании экспериментальной ИТ регулярного и целенаправленного извлечения новых терминов из медицинских текстов и описания экспертами их значений в терминологических профилях болезней. В целом проект направлен на актуализацию терминологических профилей исследуемых болезней⁵ как компонентов корпоративной базы знаний медицинского учреждения (института) в его предметной области. Речь не идет о терминологическом дублировании известных медицинских словарей, тезаурусов и других лингвистических ресурсов [1–7]. В рамках проекта решается задача извлечения тех новых терминов, которые появляются в медицинских статьях и других научных текстах, но при этом еще не вошли в медицинские лингвистические ресурсы (МЛР). Эти термины используются для актуализации терминологических профилей исследуемых болезней в корпоративной базе знаний.

Для решения указанной выше задачи проектируется экспериментальная ИТ целенаправленного извлечения новых терминов с использованием средовой модели актуализации терминологических профилей болезней новыми терминами. Цель статьи состоит в описании этой модели, относящейся к классу средовых моделей информатики.

2 Средовая модель актуализации

Частные случаи средовой модели актуализации, рассмотренные в работах [8–10], предназначены для описания процесса извлечения новых терминов из текстов и итерационного обновления терминологического портрета болезни. В течение первого года выполнения проекта опыт применения этих частных случаев при проектировании экспериментальных ИТ и решении задач проекта показал возможность их обобщения. Средовая модель актуализации в предлагаемой обобщенной форме даст возможность использовать одну и ту же модель при решении более широкого спектра задач. Ключевой

* Исследование выполнено при поддержке РФФИ и Государственного фонда естественных наук (ГФЕН) Китая (проект 21-57-53018).

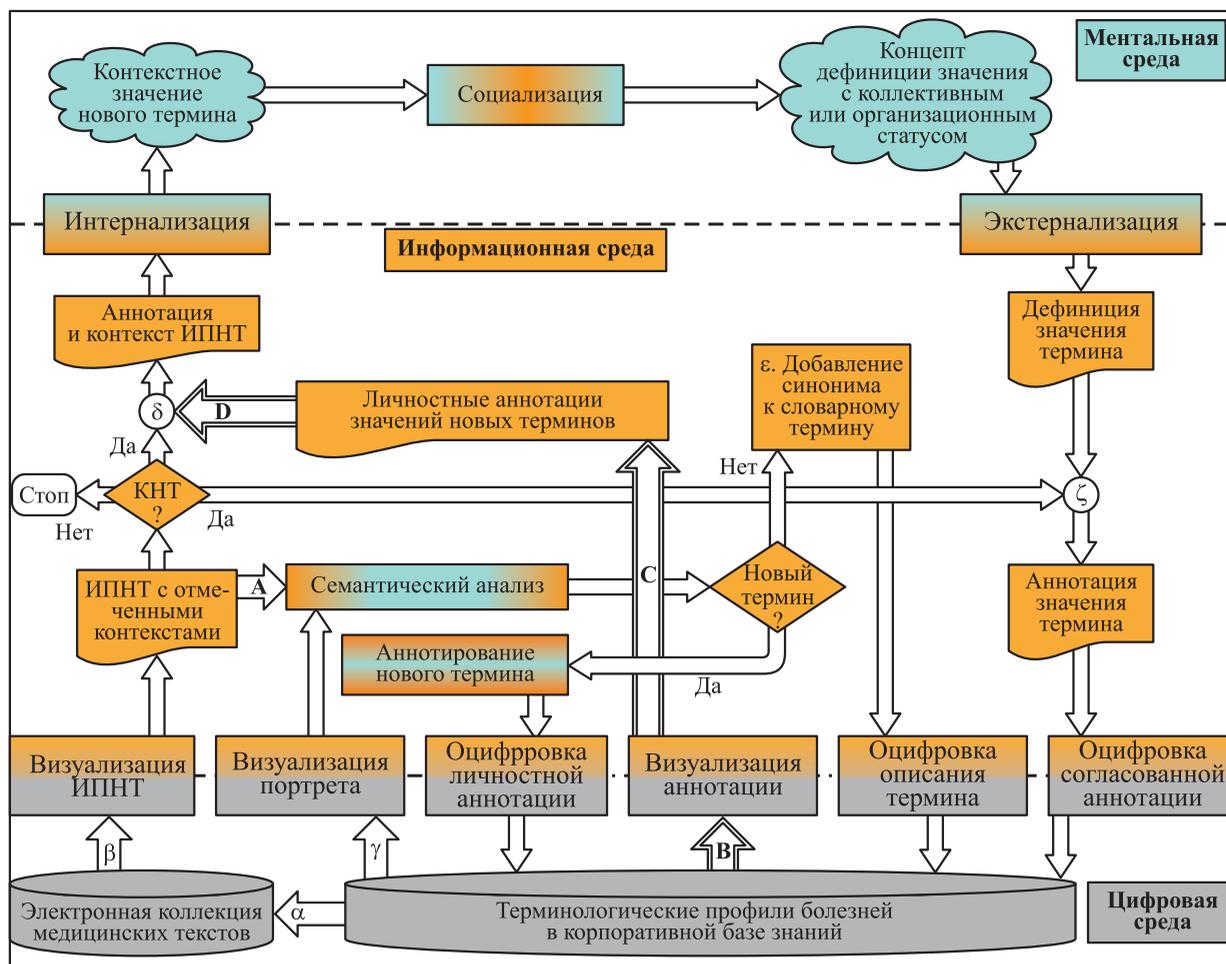
¹Федеральный исследовательский центр «Информатика и управление» Российской академии наук, izatsman@yandex.ru

²Институт информационных систем и инженерно-компьютерных технологий Российского нового университета, ol-zolot@yandex.ru

³Институт информационных систем и инженерно-компьютерных технологий Российского нового университета, aida_khatif@mail.ru

⁴Технологический университет г. Хэфэй (КНР), dongxiaogu@yeah.net

⁵Терминологический портрет болезни включает отобранные экспертами ключевые термины ее описания, дефиниции их значений, отношения между значениями терминов в портрете, контексты использования терминов и аннотации контекстов тех научных статей и других медицинских текстов, которые использовались экспертами при формировании дефиниций значений терминов, а также степень конвенциональности (социализации) значений в рамках корпоративной базы знаний.



Средовая модель актуализации терминологического портрета болезни в корпоративной базе знаний

аспект обобщения состоит в разделении новых терминов на четыре категории с точки зрения статуса их конвенциональности. Далее будут рассмотрены первые итерации обновления и статусы конвенциональности терминов.

До начала первой итерации обновления терминологического портрета исследуемой болезни экспертами формируется его начальная версия на основе терминов и их синонимов, используемых при описании этой болезни, а также их дефиниций в существующих МЛР [1–7]. Этим терминам, отобраным для начальной версии портрета, присваивается максимальный статус конвенциональности — *словарный*.

На первой итерации обновления начальная версия терминологического портрета используется как критерий новизны при поиске (см. стрелку с буквой α на рисунке). Из электронной коллекции из-

влекаются статьи с описанием исследуемой болезни (стрелка с буквой β) в качестве источников потенциально новых терминов (ИПНТ)¹. После извлечения найденных статей в них выделяются контексты ИПНТ, которые визуализируются перед операцией их семантического анализа (стрелка с буквой γ).

В процессе семантического анализа контекстов ИПНТ эксперт использует текущую (на первой итерации — начальную) версию терминологического портрета (стрелка с буквой γ). По завершении анализа эксперт разделяет найденные новые термины и новые синонимы словарных терминов (ромб «Новый термин?»), если он их обнаружил. Значения новых терминов аннотируются и оцифровываются, а дефиниции значений и аннотации контекстов новых терминов (КНТ) с этими значениями загружаются в терминологический профиль болезни. Новые синонимы оцифровываются и добавляются

¹ Если термин статьи с описанием болезни отсутствует в начальной версии терминологического портрета, то это говорит только о возможной (потенциальной) его новизне, так как при дальнейшем семантическом анализе его контекста он может оказаться синонимом словарного термина, который отсутствует в существующих МЛР. В таком случае речь идет о новом синониме, а не о новом термине.

к соответствующим словарным терминам. После обработки текстов эксперты формируют первую обновленную версию терминологического портрета. В процессе формирования каждому из этих терминов присваивается один из четырех используемых в проекте статусов конвенциональности [10–12]:

- (1) *личный*, если дефиниция значения (= концепта) термина, сформированная одним экспертом, не была согласована им с другими экспертами корпоративной базы знаний;
- (2) *коллективный*, если дефиниция значения термина, сформированная одним экспертом, была согласована им с другими экспертами корпоративной базы знаний, но не со всеми;
- (3) *организационный*, если дефиниция значения термина, сформированная одним экспертом, была согласована им со всеми экспертами корпоративной базы знаний;
- (4) *словарный*, если во время первой итерации в существующие МЛР был добавлен новый термин, включенный в портрет (если в существующие МЛР были добавлены термины, отсутствующие в обработанных текстах, то они также включаются в обновленную версию портрета в статусе словарных).

На завершающем этапе итерации предусмотрена возможность социализации значений новых терминов первых двух статусов конвенциональности (личностных и коллективных) с использованием их контекстов, т. е. КНТ. На рисунке показан случай социализации только *личностных значений терминов*. На этап их социализации передаются КНТ (см. ромб «КНТ?»), а не все контексты ИПНТ.

На второй и последующих итерациях используются уже обновленные портреты для извлечения новых терминов из новых поступлений текстов в коллекцию. В проекте используются новые поступления текстов в базу данных PubMed, содержащую более 34 млн записей и обеспечивающую развитые возможности поиска и формирования статистических данных, включая получение распределения статей с заданным термином по годам их публикации [13].

В процессе построения средовой модели актуализации использовались теоретические основания их построения, сформулированные в работе [14]. Далее будут рассмотрены три основных положения теоретических оснований построения всего класса средовых моделей (выделены ниже курсивом) в контексте их применения в проекте (см. рисунок) [10, 15].

1. *Выделение в предметной области информатики тех сред разной природы, которые необходимы для мо-*

делирования проектируемой ИТ, определяется целью ее создания.

В проекте цель проектирования ИТ сформулирована так: регулярное и целенаправленное извлечение новых терминов из медицинских текстов и описание экспертами их значений в терминологических профилях болезней корпоративной базы знаний.

Из формулировки этой цели следует, что из пяти сред предметной области информатики (ментальная, информационная, цифровая, нейро- и ДНК-среда), рассмотренных в работах [16–18], для проектируемой ИТ достаточно рассмотреть три следующих среды (см. рисунок), поскольку объекты нейро- и ДНК-сред не используются в этой ИТ:

- (1) *ментальная среда* — это совокупность значений терминов как концептов знания экспертов (верхняя часть рисунка);
- (2) *информационная среда* — это совокупность описаний значений терминов, подготовленных экспертами (средняя часть);
- (3) *цифровая среда* — это совокупность компьютерных кодов терминологических профилей болезней корпоративной базы знаний (нижняя часть).

2. *Распределение моноприродных этапов ИТ и полиприродных этапов ИТ с входами и выходами одной природы по средам согласно их природе.*

Рисунок иллюстрирует это достаточно общее теоретическое основание построения всего класса средовых моделей на примере следующих *шести* моноприродных этапов ИТ, обозначенных греческими буквами α – ζ :

- (1) чтение в цифровой среде терминологического портрета для его использования в качестве критерия потенциальной новизны термина при поиске (этап α на рисунке);
- (2) чтение из электронной коллекции статей с описанием исследуемой болезни как ИПНТ (этап β);
- (3) чтение терминологического портрета для разделения новых терминов и новых синонимов словарных терминов (этап γ);
- (4) объединение КНТ с аннотацией значения нового термина, включающей дефиницию этого значения (этап δ);
- (5) добавление синонима к словарному термину (этап ϵ , стрелка чтения словарного термина на рисунке не показана, чтобы упростить его);
- (6) объединение в форме аннотации дефиниции концепта нового термина с его контекстом, т. е. с КНТ (этап ζ).

Кроме шести моноприродных этапов на рисунке показаны *три полиприродных этапа* этой ИТ с входами и выходами одной природы, но с использованием внутри этапов сущностей другой среды. Распределение этих трех этапов по средам (двух этапов в информационной среде и одного этапа в ментальной среде) следует из этого же теоретического основания, так как:

- на этапе «Семантический анализ» эксперт изучает контекст ИПНТ (информационная среда), принимает решение о новизне термина и его значения на основе своего понимания смысла контекста термина, т. е. своего личного концепта (ментальная среда), на основе анализа которого проставляется метка «новый термин» или «синоним словарного термина» (информационная среда); таким образом, вход и выход этого этапа ИТ принадлежат одной среде — *информационной*;
- на этапе «Аннотирование нового термина» эксперт анализирует его контекст (информационная среда) и на основе своего понимания смысла контекста термина, т. е. своего личного концепта (ментальная среда), формирует аннотацию (информационная среда); таким образом, вход и выход этого этапа ИТ также принадлежат одной среде — *информационной*;
- на этапе «Социализация» эксперты на основе собственного понимания контекстного значения нового термина (ментальная среда) предлагают варианты дефиниций (информационная среда) и пытаются сформировать согласованный между ними концепт дефиниции значения с коллективным или организационным статусом (ментальная среда); таким образом, вход и выход этого этапа ИТ принадлежат одной среде — *ментальной* (возможный случай отсутствия согласования концепта экспертами на рисунке не показан).

3. *Распределение этапов ИТ по границам между средами, если их вход и выход принадлежат средам разной природы.*

На рисунке показаны *восемь этапов*, входы и выходы которых принадлежат средам разной природы, включая шесть этапов ИТ на границе между информационной и цифровой средами и два этапа на границе между информационной и ментальной средами:

- (1) этап «Визуализация ИПНТ», компьютерные коды которого (цифровая среда) преобразуются в текст ИПНТ (информационная среда);
- (2) этап «Визуализация [терминологического] портрета», компьютерные коды которого

(цифровая среда) преобразуются в текст портрета (информационная среда);

- (3) этап «Оцифровка личностной аннотации», текст которой (информационная среда) преобразуется в компьютерные коды терминологического портрета корпоративной базы знаний (цифровая среда);
- (4) этап «Визуализация [личностной] аннотации», компьютерные коды которой (цифровая среда) преобразуются в текст (информационная среда);
- (5) этап «Оцифровка описания [словарного] термина», текст которого (информационная среда) преобразуется в компьютерные коды терминологического портрета корпоративной базы знаний (цифровая среда);
- (6) этап «Оцифровка согласованной аннотации», текст которой (информационная среда) преобразуется в компьютерные коды терминологического портрета корпоративной базы знаний (цифровая среда);
- (7) этап «Интернализация», на котором у эксперта формируется личностное контекстное значение нового термина (ментальная среда) на основе аннотации с дефиницией значения термина и его контекста (информационная среда);
- (8) этап «Экстернализация», на котором эксперты формируют дефиницию значения термина (информационная среда) на основе согласованного ими концепта с коллективным или организационным статусом (ментальная среда).

Таким образом, согласно средовой модели актуализации терминологических портретов, 17 перечисленных этапов проектируемой ИТ можно распределить по трем категориям. К первой из них относятся шесть моноприродных этапов ИТ, обозначенных греческими буквами α – ζ ; ко второй — три этапа с входами и выходами одной природы, но с использованием внутри этапов сущностей среды другой природы; к третьей — восемь этапов, входы и выходы которых принадлежат средам разной природы.

На этапах первой категории отсутствуют интерфейсы между сущностями разной природы. При этом сами сущности могут быть только цифровыми или информационными. Каждый этап второй категории включает пару интерфейсов между сущностями разной природы. При этом сущности ментальной природы могут быть как на входе/выходе этапа (см. этап социализации), так и внутри этапа

(см. этапы семантического анализа и аннотирования). Шесть из восьми этапов третьей категории включают традиционный интерфейс между сущностями информационной и цифровой природы на границе между соответствующими средами. Этот интерфейс в проектируемых ИТ может быть реализован, например, с помощью таблиц Unicode. В процессе проектирования ИТ наибольшую сложность представляет разработка этапов второй категории, а также двух этапов на границе между ментальной и информационной средами, так как они включают сущности ментальной природы. Вопрос их компьютеризации заслуживает отдельного рассмотрения и выходит за рамки данной статьи.

Рассмотренная средовая модель актуализации позволила при выполнении проекта до начала проектирования ИТ определить природу сущностей каждого этапа и локализовать этапы с ментальными сущностями на их входе/выходе или внутри этапов, проблема компьютеризации которых относится в информатике к категории проблем *когнитивной сложности* [14, 19].

3 Эксперимент

В первый год выполнения проекта оставался открытым вопрос о степени полноты наборов терминов МЛР [1–7], т. е. насколько полно наборы терминов, используемых в научных статьях и других медицинских текстах при описании исследуемых болезней, представлены в МЛР. Поэтому на втором году выполнения проекта был проведен эксперимент, краткое описание которого приведено ниже, для проверки гипотезы неполноты наборов терминов МЛР и необходимости регулярной актуализации терминологических портретов болезней в корпоративной базе знаний.

В рамках проведенного эксперимента анализировались термины медицинских статей, содержащих описание кластеров кальцификации¹. В базе электронных биомедицинских публикаций PubMed был осуществлен поиск этих статей по запросу: (calcification[Abstract] AND (“neoplasms”[MeSH Terms] OR “neoplasms”[All Fields] OR “cancer”[All Fields])) AND (“breast”[MeSH Terms] OR “breast”[All Fields]) AND (“2003/01/01”[PubMed Date] : “2021/12/31”[PubMed Date]). За период 2003–2021 гг. найдены 844 статьи, которые распределились по трем временным периодам следующим образом: 2003–2009 гг. — 57 статей; 2010–2015 гг. — 225 статей; 2016–2021 гг. — 562 статьи.

При поиске использовались заголовки и тексты аннотаций статей. После автоматического поиска

вхождений термина calcification (всего были найдены 2658 его вхождений, но не каждое вхождение порождает терминоподобное словосочетание) экспертами был выполнен семантический анализ их контекстов, чтобы найти терминоподобные словосочетания, включающие термин calcification.

В результате семантического анализа 2658 вхождений экспертно были выделены 93 терминоподобных словосочетания за период 2003–2021 гг. Из них 23 были обнаружены в статьях 2003–2009 гг. (например, Amorphous Calcifications, Arterial Calcification, Aortic Calcification), 54 — в статьях 2010–2015 гг., включая 31 терминоподобное словосочетание, которого не было в статьях 2003–2009 гг. (например, Arteriosclerotic Calcification), и 93 — в статьях 2016–2021 гг., включая ранее не встречавшиеся 39 терминов (например, Abnormal Calcification, Arterial Media(L) Calcification, Calcification Crystallite).

Для проверки на словарный статус 93 терминоподобных словосочетаний, найденных экспертами в статьях, использовался тезаурус MeSH [1]. В результате проверки лишь 3 из 93 терминоподобных словосочетаний совпали с терминами тезауруса MeSH, т. е. имеют словарный статус.

4 Заключение

Проведенный эксперимент подтвердил гипотезу о существенной неполноте наборов терминов МЛР. При этом отсутствующие терминологические словосочетания уже используются в научных статьях и других медицинских текстах при описании исследуемой болезни. Проведенный эксперимент показал необходимость регулярной актуализации терминологических портретов болезней в корпоративной базе знаний, что, в свою очередь, подтверждает важность решения задачи моделирования и создания ИТ регулярного и целенаправленного извлечения новых терминов из медицинских текстов и описания экспертами их значений в терминологических профилях болезней.

Литература

1. Medical subject headings (MeSH) Thesaurus. <https://www.nlm.nih.gov/mesh/meshhome.html>.
2. Systematized nomenclature of medicine clinical terms (SNOMED CT). <https://bioportal.bioontology.org/ontologies/SNOMEDCT>.
3. Unified Medical Language System (UMLS). <https://www.nlm.nih.gov/research/umls/index.html>.

¹Наличие кластеров кальцификации может быть использовано для диагностики заболеваний молочных желез.

4. National Cancer Institute Thesaurus. <https://ncithesaurus.nci.nih.gov/ncitbrowser/pages/home.jsf?version=22.08e>.
5. Medical dictionary of health terms. — Cambridge, MA, USA: Harvard Health Publishing, 2011. <https://www.health.harvard.edu/a-through-c>.
6. Dorland's illustrated medical dictionary. — 33rd ed. — Philadelphia, PA, USA: Elsevier Saunders, 2019. 2144 p.
7. MedTerms Medical Dictionary. <https://www.medicinenet.com/medterms-medical-dictionary/article.htm>.
8. *Зацман И. М.* Проблемно-ориентированная актуализация словарных статей двуязычных словарей и медицинской терминологии: сопоставительный анализ // Информатика и её применения, 2021. Т. 15. Вып. 1. С. 94–101.
9. *Зацман И. М.* Формы представления нового знания, извлеченного из текстов // Информатика и её применения, 2021. Т. 15. Вып. 3. С. 83–90.
10. *Zatsman I., Khakimova A.* New knowledge discovery for creating terminological profiles of diseases // 22nd European Conference on Knowledge Management Proceedings. — Reading, U.K.: Academic Publishing International Ltd., 2021. P. 837–846.
11. *Nissen M. E.* Harnessing knowledge dynamics: Principled organisational knowing and learning. — London: IRM Press, 2006. 287 p.
12. *Зацман И. М.* Компьютерная и экономическая модели генерации нового знания: сопоставительный анализ // Системы и средства информатики, 2021. Т. 31. № 4. С. 84–96.
13. PubMed. National Library of Medicine. <https://pubmed.ncbi.nlm.nih.gov>.
14. *Зацман И. М.* Средовые модели информационных технологий: теоретические основания построения // Информатика и её применения, 2022. Т. 16. Вып. 3. С. 59–67.
15. *Зацман И. М. Золотарев О. В., Хакимова А. Х.* Средовые модели извлечения из текста новых терминов и индикаторов настроений // Информатика и её применения, 2022. Т. 16. Вып. 2. С. 60–67.
16. *Зацман И. М.* Интерфейсы третьего порядка в информатике // Информатика и её применения, 2019. Т. 13. Вып. 3. С. 82–89.
17. *Зацман И. М.* Кодирование концептов в цифровой среде // Информатика и её применения, 2019. Т. 13. Вып. 4. С. 97–106.
18. *Зацман И. М.* Теоретические основания компьютерного образования: среды предметной области информатики как основание классификации ее объектов // Системы и средства информатики, 2022. Т. 32. № 4. С. 77–89.
19. *Harel D.* Algorithmics — the spirit of computing. — Reading, MA, USA: Addison-Wesley, 1987. 514 p.

Поступила в редакцию 14.10.22

MODEL AND TECHNOLOGY FOR DISCOVERING NEW TERMS IN MEDICAL TEXTS

I. M. Zatsman¹, O. V. Zolotarev², A. K. Khakimova², and Gu Dongxiao³

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

²Russian New University, 22 Radio Str., Moscow 105005, Russian Federation

³Hefei University of Technology, 193 Tunxi Road, Hefei, Anhui 230009, P.R. China

Abstract: The model of information technology for discovering new terms in medical texts, which belongs to the previously defined class of informatics' medium models, is considered. In the conducted experiment, the MeSH (Medical Subject Headers) dictionary is used to determine the novelty of terms which was created and updated by the National Library of Medicine (USA). The emergence of new terms is due to the representation (in medical papers and other scientific texts) of new knowledge about the studied diseases, methods of their treatment, and medicines used which has not yet been reflected in medical dictionaries and thesauri. In information systems of medical institutions, the proposed technology makes it possible to regularly update the profiles of the studied diseases corresponding to their subject domain. The aim of the paper is to describe the medium model of information technology for updating terminological profiles of diseases.

Keywords: medium models in informatics; medical texts; terminological profile; discovering new terms in texts

DOI: 10.14357/19922264220412

Acknowledgments

The reported study was funded by RFBR and NSFC, project number 21-57-53018.

References

1. MeSH Thesaurus. Available at: <https://www.nlm.nih.gov/mesh/meshhome.html> (accessed November 21, 2022).
2. Systematized nomenclature of medicine clinical terms. Available at: <https://bioportal.bioontology.org/ontologies/SNOMEDCT> (accessed November 21, 2022).
3. Unified medical language system. Available at: <https://www.nlm.nih.gov/research/umls/index.html> (accessed November 21, 2022).
4. National Cancer Institute Thesaurus. Available at: <https://ncit.nci.nih.gov/ncitbrowser> (accessed November 21, 2022).
5. *Medical dictionary of health terms*. Cambridge, MA: Harvard Health Publishing. Available at: <https://www.health.harvard.edu/a-through-c> (accessed November 21, 2022).
6. *Dorland's illustrated medical dictionary*. 2019. 33rd ed. Philadelphia, PA: Elsevier Saunders. 2144 p.
7. MedTerms medical dictionary. Available at: <https://www.medicinenet.com/medterms-medical-dictionary/article.htm> (accessed November 21, 2022).
8. Zatsman, I. 2021. Problemno-orientirovannaya aktualizatsiya slovarnykh statey dvuyazychnykh slovarey i meditsinskoy terminologii: sopostavitel'nyy analiz [Problem-oriented updating of dictionary entries of bilingual dictionaries and medical terminology: Comparative analysis]. *Informatika i ee Primeneniya — Inform. Appl.* 15(1):94–101.
9. Zatsman, I. 2021. Formy predstavleniya novogo znaniya, izvlechnnogo iz tekstov [Forms representing new knowledge discovered in texts]. *Informatika i ee Primeneniya — Inform. Appl.* 15(3):83–90.
10. Zatsman, I., and A. Khakimova. 2021. New knowledge discovery for creating terminological profiles of diseases. *22nd European Conference on Knowledge Management Proceedings*. Reading, U.K.: Academic Publishing International Ltd. 837–846.
11. Nissen, M. E. 2006. *Harnessing knowledge dynamics: Principled organizational knowing & learning*. London: IRM Press. 287 p.
12. Zatsman, I. 2021. Komp'yuternaya i ekonomicheskaya modeli generatsii novogo znaniya: sopostavitel'nyy analiz [Computer and economic models of new knowledge generation: A comparative analysis]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 31(4):84–96.
13. National library of medicine. PubMed. Available at: <https://pubmed.ncbi.nlm.nih.gov> (accessed November 21, 2022).
14. Zatsman, I. 2022. Sredovye modeli informatsionnykh tekhnologiy: teoreticheskie osnovaniya postroeniya [Informatics' medium models of information technology: Theoretical foundations for their creating]. *Informatika i ee Primeneniya — Inform. Appl.* 16(3):59–67.
15. Zatsman, I., O. Zolotarev, and A. Khakimova. 2022. Sredovye modeli izvlecheniya iz teksta novykh terminov i indikatorov nastroeniya [Medium models for discovering novel terms and sentiments from texts]. *Informatika i ee Primeneniya — Inform. Appl.* 16(2):60–67.
16. Zatsman, I. M. 2019. Interfeysy tret'ego poryadka v informatike [Third-order interfaces in informatics]. *Informatika i ee Primeneniya — Inform. Appl.* 13(3):82–89.
17. Zatsman, I. M. 2019. Kodirovanie kontseptov v tsifrovoy srede [Digital encoding of concepts]. *Informatika i ee Primeneniya — Inform. Appl.* 13(4):97–106.
18. Zatsman, I. 2022. Teoreticheskie osnovaniya komp'yuternogo obrazovaniya: sredi predmetnoy oblasti informatiki kak osnovanie klassifikatsii ee ob'ektov [Theoretical foundations of digital education: Subject domain media of informatics as the base of its objects' classification]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 32(4):77–89.
19. Harel, D. 1987. *Algorithmics — the spirit of computing*. Reading, MA: Addison-Wesley. 514 p.

Received October 14, 2022

Contributors

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

Zolotarev Oleg V. (b. 1959) — Candidate of Science (PhD) in technology, head of department, Russian New University, 22 Radio Str., Moscow 105005, Russian Federation; ol-zolot@yandex.ru

Khakimova Aida Kh. (b. 1963) — Candidate of Science (PhD) in biology, leading scientist, Russian New University, 22 Radio Str., Moscow 105005, Russian Federation; aida_khatif@mail.ru

Dongxiao Gu (b. 1980) — Candidate of Science (PhD) in informatics, professor, Hefei University of Technology, 193 Tunxi Road, Hefei, Anhui 230009, P.R. China; dongxiaogu@yeah.net

О БЕЗОПАСНОЙ АРХИТЕКТУРЕ ВЫЧИСЛИТЕЛЬНОЙ СИСТЕМЫ НА ОСНОВЕ МИКРОСЕРВИСОВ

А. А. Грушо¹, Н. А. Грушо², М. И. Забейло³, Д. В. Смирнов⁴, Е. Е. Тимонина⁵,
С. Я. Шоргин⁶

Аннотация: Рассматривается сетевая система микросервисной архитектуры, в которой для простоты все микросервисные вычислители (МВ) одинаковы. В каждом МВ может произойти сбой или он может получить вредоносный код (ВК). Максимальное негативное воздействие на МВ заключается в ошибке вычислений и предоставлении потребителю неправильного результата. Рассмотрены задачи выявления сбойных МВ и МВ с ВК. В решении поставленных задач используются элементы обучения. Правильно решенные задачи (условия, исходные данные и правильные ответы) накапливаются в памяти управляющей системы. Это означает, что есть возможность повторного запуска любой задачи с уже известным правильным результатом. При этом в статье используются идеи и результаты авторов по обеспечению информационной безопасности с помощью метаданных. В зависимости от предположений о возможных действиях ВК построены два класса алгоритмов безопасных вычислений в условиях его возможного воздействия на промежуточные результаты в потоке решаемых задач. Второй класс алгоритмов работает в предположении, что ВК может с вероятностью p правильно вычислять решение текущей задачи и с вероятностью $1 - p$ вносить в результат искажение. Рассмотрены три вида искажений, которые может вносить ВК и которые позволяют находить истинное решение точно или с некоторой, сколь угодно малой вероятностью ошибки.

Ключевые слова: информационная безопасность; безопасные вычисления в условиях вредоносного кода; микросервисная архитектура

DOI: 10.14357/19922264220413

1 Введение

Безопасным архитектурам вычислительных систем посвящено много исследований. Основную часть научной литературы в области безопасных архитектур составляют руководства по внедрению различных наборов средств и механизмов защиты в распределенных информационных системах. Сюда относятся большинство книг и статей, таких как фундаментальные труды [1–3] и др. Один из основных принципов в создании безопасных архитектур — изоляция бизнес-функций и связанная с ней изоляция информационной поддержки этих функций [4, 5].

В последнее время наблюдается изменение архитектур вычислительных систем, позволяющее полностью или частично изолировать поддержку различных функций информационных технологий (ИТ). Здесь необходимо выделить два базовых направления: виртуализация и контейнеризация.

Виртуализация базируется на создании множества виртуальных компьютеров на базе одного вы-

числительного устройства. Создание виртуальных компьютеров основано на гипервизоре, который, опираясь на основную операционную систему (ОС) вычислительного устройства, реализует виртуальные компьютеры и их функционирование на том же сервере [6, 7]. Виртуальные машины с различными ОС работают на одном физическом сервере.

Контейнеризация — это технология, которая позволяет запускать приложения изолированно от основной ОС. Приложение упаковывается в специальный контейнер, внутри которого развернута среда, необходимая для работы. Контейнеры устанавливаются на физический сервер и его ОС. Вместе с тем проверена возможность удаленной контейнеризации, когда необходимые дополнения для создания облегченной ОС, необходимой для работы конкретного приложения, достаточно быстро передаются по сети и обеспечивают полную изоляцию работы контейнера и приложения в нем.

Одной из перспективных тенденций представляется введение микросервисных архитектур, поз-

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

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

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

⁴ПАО Сбербанк России, dvlsmirnov@sberbank.ru

⁵Федеральный исследовательский центр «Информатика и управление» Российской академии наук, eltimon@yandex.ru

⁶Федеральный исследовательский центр «Информатика и управление» Российской академии наук, sshorgin@ipiran.ru

воляющих диверсифицировать и ускорить решения многих задач и ИТ [8–10]. Среда для выполнения микросервисов состоит из системы контейнеризованных вычислителей. В этом случае каждый из микросервисов, как правило, изолируется в отдельный контейнер или небольшую группу контейнеров, доступную по сети, контролируется и управляется системой оркестрации или центром управления.

В работе рассматривается сетевая архитектура микросервисной архитектуры, в которой для простоты все МВ одинаковы.

Центр управления состоит из двух блоков I_1 и I_2 . Блок I_1 имеет прямые соединения со всеми МВ, случайно выбирает свободный МВ и направляет ему задачу с исходными данными. Каждый МВ может быть занят решением только одной задачи и освобождается только после передачи блоку I_1 результатов решения своей очередной задачи. После этого ему может быть назначена следующая задача. Блок I_2 содержит метаданные о порядке решения задач и использовании полученных от I_1 результатов их решения [11–14].

В каждом МВ может произойти сбой или МВ может получить ВК. Предположим, что максимальное негативное воздействие на МВ состоит в ошибке вычислений и предоставлении I_1 неправильного результата. Задача состоит в выявлении сбойных МВ и МВ с ВК.

Пусть правильно решенные задачи накапливаются в памяти I_2 (условия, исходные данные и правильные ответы). Это означает, что имеется возможность повторного запуска любой задачи с уже известным правильным результатом.

2 Алгоритм A_1 поиска сбойного микросервисного вычислителя

Рассмотрим следующую модель. Пусть имеются M МВ $\{\beta_1, \beta_2, \dots, \beta_M\}$, из которых m МВ оказываются сбойными или пораженными ВК. Система решает задачи $\alpha_1, \alpha_2, \dots, \alpha_N$ в потоке.

Рассмотрим решение задачи в ситуации, когда сбой МВ неустраним, а ВК портит решение каждой задачи. Тогда можно воспользоваться результатами правильно решенных задач из памяти I_2 следующим образом.

Алгоритм A_1 берет любую правильно решенную задачу и последовательно пробует решать ее каждым МВ. При этом считается, что сложность решения любой задачи равна 1. Если сбой или поражение ВК происходит только в одном МВ, то среднее число шагов поиска сбоя равно $(M - 1)/2$.

Если число сбойных МВ или МВ, содержащих ВК, неизвестно, то число шагов до выявления их всех алгоритмом A_1 равно M . После фильтрации всех сбойных МВ или МВ, зараженных ВК, система будет решать все задачи правильно.

Предположим, что ВК способен маскироваться в том смысле, что любую решаемую задачу он решает правильно с вероятностью p независимо от других задач и с вероятностью $(1 - p)$ решает ее заведомо неправильно. Рассмотрим возможности применения алгоритма A_1 для выявления ВК.

Для фиксированного МВ пробуем решить известную задачу, но получение правильного ответа в этом случае не гарантирует отсутствие ВК в данном МВ. Если предполагать, что проведены n испытаний одного МВ, то вероятность невыявления ВК равна p^n . Это означает, что вероятность выявления ВК за n шагов равна $1 - p^n$.

Эта вероятность может быть сделана близкой к 1. Однако может возникнуть эффект больших отклонений за счет потока задач. Вероятность выявления ВК на данной задаче равна $1 - p^n$, но вероятность выявления алгоритмом A_1 всех МВ с ВК равна $(1 - p^n)^M$. Тогда вероятность пропустить хотя бы одну МВ с ВК равна $1 - (1 - p^n)^M$, что может оказаться не совсем малой величиной, т. е. алгоритм A_1 не гарантирует правильного решения всех задач в потоке.

3 Алгоритм A_2 обеспечения защиты от вредоносного кода в микросервисном вычислителе

Можно построить алгоритм A_2 , гарантирующий правильное решение всех задач и выявление ВК в сформулированных выше предположениях при допущении случайного включения повреждения результата с помощью ВК. Для этого введем понятие «отказ», которое означает, что алгоритм A_2 не может точно указать наличие ВК в трех выбранных МВ.

Алгоритм A_2 строится следующим образом. Каждый раз, когда I_2 передает для I_1 очередную задачу, I_1 случайно и независимо выбирает не одну МВ, а сразу три различных МВ. Далее I_1 передает всем трем МВ одну и ту же задачу, которую он получил от I_2 . В результате с учетом возможности наличия ВК получаем три результата решения одной и той же задачи. Если все три результата одинаковые, то считаем, что получено правильное решение, а все три МВ не портят результат, даже если в них есть ВК.

Предположим, что на одном МВ получен ответ, не совпадающий с двумя одинаковыми ответами двух других МВ. Это означает, что правильный ответ дают два МВ, так как случайное совпадение двух ложных ответов, полученных независимо друг от друга, считаем невозможным событием. Одновременно выявляется МВ, в котором сработал механизм ВК, порождающий неправильный ответ.

Рассмотрим случай, когда все три ответа на одну и ту же поставленную задачу различны. Этот случай соответствует понятию отказ, так как в этом случае невозможно выделить правильный ответ, даже если он существует в полученных результатах. Возможен также случай, когда сработал ВК во всех трех МВ. В случае получения отказа алгоритм A_2 повторно запускает независимый выбор трех новых МВ и поручает им заново решать ту же задачу. Если опять получены разные результаты, то A_2 сравнивает их с предыдущей тройкой. Если получены два одинаковых результата, по одному в первой и второй тройке, то этот результат считается правильным, а все несовпадающие с ним определяют МВ, зараженные ВК. Если одинаковых результатов нет, то опять выбирается новая тройка МВ и т. д.

Отметим, что в алгоритме A_2 выявление МВ с ВК происходит непосредственно в процессе решения потока задач и допускает безопасную работу в условиях наличия небезопасных МВ. При этом не ставится задача выявления всех зараженных МВ.

4 Вероятностный анализ алгоритма A_2

Еще раз отметим, что в основе эффективной работы алгоритма A_2 лежит предположение о малой вероятности случайного совпадения неправильных ответов, порожденных ВК, между собой и между неправильными и настоящими ответами.

В сделанных предположениях найдем оценки отказов. Случайность выбора МВ для решения конкретной задачи позволяет допустить гипергеометрическое распределение вариантов появления МВ. Для двух МВ, содержащих ВК, которые породили две ошибки, и одном правильном решении получаем следующую вероятность отказа:

$$P_0(2) = (1 - p)^2 \frac{(M - m) \binom{m}{2}}{\binom{M}{3}} + p(1 - p)^2 \frac{\binom{m}{3}}{\binom{M}{3}}.$$

Если все три МВ дали неправильные результаты, то вероятность этого события равна

$$P_0(3) = (1 - p)^3 \frac{\binom{m}{3}}{\binom{M}{3}}.$$

Тогда суммарная вероятность отказа при решении данной задачи составит

$$P_0 = P_0(2) + P_0(3).$$

Среднее число отказов в потоке задач длины N равно NP_0 .

Оказывается, что алгоритм A_2 с большой вероятностью не сможет находить правильное решение одной задачи, если число зараженных МВ велико. В самом деле, если $m = M$, то

$$P_0 = (1 - p)^2.$$

Тогда вероятность получить правильное решение данной задачи равна $1 - (1 - p)^2$ и среднее число правильно решенных задач в потоке длины N равно $N(1 - (1 - p)^2)$.

5 Работа алгоритма A_2 в различных предположениях о вредоносном коде

Рассмотрим наихудший случай, когда в результате решения МВ очередной задачи получается числовой вектор, а ВК осуществляет одинаковую замену некоторого значения в этом векторе на фиксированное неизвестное значение γ . Пусть все остальные предположения, рассмотренные ранее, остаются в силе. В этом случае алгоритм A_2 перестает работать. В самом деле, рассмотрим различные сочетания результатов применения алгоритма A_2 .

Если получены все три одинаковых результата решения одной задачи тремя МВ, то все три из них могут оказаться одинаково испорченными с вероятностью $(1 - p)^3$. Два одинаковых результата и один отличный от них создают неопределенность: либо два ложных решения, либо одно ложное. Три различных результата в указанных условиях возникнуть не могут.

В то же время наличие базы правильно решенных задач позволяет найти неизвестное значение γ . Для этого надо запустить несколько ранее правильно решенных задач и дождаться нескольких различий правильных и неправильных ответов. Тогда значение γ восстанавливается однозначно.

При известном значении γ сочетания ответов по алгоритму A_2 по новой, ранее не решенной задаче оцениваются по-другому.

Если получены все три одинаковых ответа без значения γ в них, то полученное решение — правильное. Если получены два ответа без значения γ в них и один ответ со значением γ в нем, то два одинаковых ответа определяют истинное решение. Если получены два ответа со значением γ в них и один ответ без значения γ , то этот один ответ определяет истинное решение. Если все три ответа со значением γ одинаковые, то получаем отказ, так как как все три могут быть получены с помощью ВК. Вероятность такого события равна

$$P_0(3) = (1 - p)^3 \frac{\binom{m}{3}}{\binom{M}{3}}.$$

В этом случае необходимо повторять решение этой же задачи до тех пор, пока не появится ответ без значения γ . Поскольку вероятность ошибки известна, то, аппроксимируя повторные испытания геометрическим распределением, придется принять решение с присутствием значения γ с заданной вероятностью ошибки.

Пусть ВК может использовать несколько значений $\Gamma = \{\gamma_1, \gamma_2, \dots, \gamma_k\}$ для нарушения правильности ответа. Тогда, используя базу данных правильных решений различных задач, можно восстановить Γ . При этом все три решения с помощью алгоритма A_2 , не содержащие элементов Γ , определяют правильное решение текущей задачи.

Если одно решение не содержит элементов Γ , а два содержат, то истинным считается решение без элементов Γ . Если два одинаковых решения не содержат элементов Γ , то это истинное решение. Если все три содержат элементы Γ , то это отказ и надо повторять процедуру алгоритма A_2 . Если при повторениях возникают другие элементы Γ , то их можно сразу отсеять. Постоянно повторяющееся значение элемента из Γ определяет истинное решение с заданной вероятностью ошибки.

6 Заключение

В работе рассмотрены все варианты решения задачи преодоления ошибки вычисления, порожденного ВК, в сетевых архитектурах вычислительной системы. Построены два класса алгоритмов решения задачи поиска сбоя и ВК в такой архитектуре. В отличие от алгоритма A_1 , алгоритм A_2 предоставляет больше примеров однознач-

ного выявления истинного решения каждой текущей задачи. В тех случаях когда истинное решение определяется с ненулевой вероятностью ошибки, рассчитаны все необходимые вероятности.

Построен пример архитектуры, позволяющей обеспечивать безопасные решения задач в условиях наличия небезопасных компонентов в вычислительной системе.

Литература

1. *Ramachandran J.* Designing security architecture solutions. — John Wiley & Sons Singapore Pte. Ltd., 2002. 452 p.
2. *Sherwood J., Clark A., Lynas D.* Enterprise security architecture, 2009. 25 p. <https://sabsacourses.com/wp-content/uploads/2021/02/TSI-W100-SABSA-White-Paper.pdf>.
3. *Wang S., Ledley R. S.* Computer architecture and security fundamentals of designing secure computer systems. — John Wiley & Sons Singapore Pte. Ltd., 2013. 342 p.
4. *Тимонина Е. Е.* Анализ угроз скрытых каналов и методы построения гарантированно защищенных распределенных автоматизированных систем: Дисс. . . . докт. техн. наук. — М.: РГГУ, 2004. 204 с.
5. *Грушо А. А., Грушо Н. А., Тимонина Е. Е., Шоргин С. Я.* Возможности построения безопасной архитектуры для динамически изменяющейся информационной системы // Системы и средства информатики, 2015. Т. 25. № 3. С. 78–93.
6. *Lacoste M.* Architecture for secure computation infrastructure and self-management of VM security // Zenodo, 2015. 97 p. doi: 10.5281/zenodo.49743.
7. *Grusho A., Nikolaev A., Piskovski V., Sentchilo V., Timonina E.* Endpoint cloud terminal as an approach to secure the use of an enterprise private cloud // Scientific and Technical Conference (International) on Modern Computer Network Technologies Proceedings. — Piscataway, NJ, USA: IEEE, 2020. Art. 9258244. 4 p. doi: 10.1109/MoNeTeC49726.2020.9258244.
8. *Düllmann T. F.* Performance anomaly detection in microservice architectures under continuous change: Master's Thesis, 2017. <https://elib.uni-stuttgart.de/bitstream/11682/9083/1/MScThesis-TFDuellmann.pdf>.
9. *Mayer B., Weinreich R.* Dashboard for microservice monitoring and management // Conference (International) on Software Architecture Workshops. — Piscataway, NJ, USA: IEEE, 2017. P. 66–69. doi: 10.1109/ICSAW.2017.44.
10. *Brandón Á., Solé M., Huélamo A., Solans D., Pérez M. S., Muntés-Mulero V.* Graph-based root cause analysis for service-oriented and microservice architectures // J. Syst. Software, 2020. Vol. 159. Art. 110432. 17 p.
11. *Grusho A. A., Timonina E. E., Shorgin S. Ya.* Modelling for ensuring information security of the distributed information systems // 31st European Conference on Modelling

- and Simulation Proceedings. — Dudweiler, Germany: Digitaldruck Pirrot GmbH, 2017. P. 656–660.
12. Grusho A., Grusho N., Zabezhalo M., Zatsarinny A., Timonina E. Information security of SDN on the basis of meta data // 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.
 13. Грушо А. А., Забейхайло М. И., Грушо Н. А., Тимонина Е. Е. Информационная безопасность на основе метаданных в компонентно-интеграционных архитектурах информационных систем // Системы и средства информатики, 2018. Т. 28. № 2. С. 34–41.
 14. Grusho A., Grusho N., Zabezhalo M., Timonina E. Generation of metadata for network control // Distributed computer and communication networks / Eds. V. M. Vishnevskiy, K. E. Samouylov, D. V. Kozyrev. — Lecture notes in computer science ser. — Cham: Springer, 2020. Vol. 12563. P. 723–735. doi: 10.1007/978-3-030-66471-8_55.

Поступила в редакцию 11.10.22

ABOUT THE SECURE ARCHITECTURE OF A MICROSERVICE-BASED COMPUTING SYSTEM

A. A. Grusho¹, N. A. Grusho¹, M. I. Zabezhalo¹, D. V. Smirnov², E. E. Timonina¹, and S. Ya. Shorgin¹

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

²Sberbank of Russia, 19 Vavilov Str., Moscow 117999, Russian Federation

Abstract: The paper discusses a network-centric microservice architecture system in which all microservice computers are the same for simplicity. Each microservice computer may fail or receive malicious code. The maximum negative impact on the microservice computer is a calculation error and providing the consumer with the wrong result. The tasks of detecting failed microservice computers and detecting microservice computers with malicious code are considered. In solving the set tasks, elements of training are used. Correctly solved problems (conditions, source data, and correct answers) are accumulated in the memory of the control system. This means that one can restart any task with an already known correct result. At the same time, the article uses the ideas and results of the present authors to ensure information security while using metadata. Depending on the assumptions about the possible actions of malicious code, two classes of secure computing algorithms are built in the context of its possible impact on intermediate results in the flow of solved problems. The second class of algorithms works in the assumption that malicious code can correctly calculate the solution to the current problem with probability p and introduce distortion into the result with probability $1 - p$. The authors consider three types of distortions that malicious code can introduce and which allow one to either find the true solution accurately or with low probability of error.

Keywords: information security; secure computing under malicious code conditions; microservice architecture

DOI: 10.14357/19922264220413

References

1. Ramachandran, J. 2002. *Designing security architecture solutions*. John Wiley & Sons Singapore Pte. Ltd. 452 p.
2. Sherwood, J., A. Clark, and D. Lynas. 2009. Enterprise security architecture. 25 p. Available at: <https://sabsacourses.com/wp-content/uploads/2021/02/TSI-W100-SABSA-White-Paper.pdf> (accessed November 16, 2022).
3. Wang, S., and R. S. Ledley. 2013. *Computer architecture and security fundamentals of designing secure computer systems*. John Wiley & Sons Singapore Pte. Ltd. 342 p.
4. Timonina, E. E. 2004. Analiz ugroz skrytykh kanalov i metody postroeniya garantirovanno zashchishchennykh raspredelennykh avtomatizirovannykh sistem [The analysis of threats of covert channels and methods of creation of guaranteed protected distributed automated systems]. Moscow: Russian State University for the Humanities. D.Sc. Diss. 204 p.
5. Grusho, A. A., N. A. Grusho, E. E. Timonina, and S. Ya. Shorgin. 2015. Vozможности postroeniya bezopasnoy arkhitektury dlya dinamicheskoi izmenyayushchey informatsionnoy sistemy [Regarding capabilities of secured architecture creation for dynamic changing information system]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 25(3):78–93.
6. Lacoste, M., ed. 2015. *Architecture for secure computation infrastructure and self-management of VM security*. Zenodo. 97 p. doi: 10.5281/zenodo.49743.
7. Grusho, A., A. Nikolaev, V. Piskovski, V. Sentchilo, and E. Timonina. 2020. Endpoint cloud terminal as

- an approach to secure the use of an enterprise private cloud. *Scientific and Technical Conference (International) on Modern Computer Network Technologies Proceedings*. Piscataway, NJ: IEEE. 9258244. 4 p. doi: 10.1109/MoNeTeC49726.2020.9258244.
8. Düllmann, T.F. 2017. Performance anomaly detection in microservice architectures under continuous change. Master's Thesis. Available at: <https://elib.uni-stuttgart.de/bitstream/11682/9083/1/MScThesis-TFDuellmann.pdf> (accessed November 16, 2022).
 9. Mayer, B., and R. Weinreich. 2017. Dashboard for microservice monitoring and management. *Conference (International) on Software Architecture Workshops Proceedings*. Piscataway, NJ: IEEE. 66–69. doi: 10.1109/ICSAW.2017.44.
 10. Brandon, Á., M. Solé, A. Huélamo, D. Solans, M. S. Pérez, and V. Muntés-Mulero. 2020. Graph-based root cause analysis for service-oriented and microservice architectures. *J. Syst. Software* 159:110432. 17 p.
 11. Grusho, A. A., E. E. Timonina, and S. Ya. Shorgin. 2017. Modelling for ensuring information security of the distributed information systems. *31st European Conference on Modelling and Simulation Proceedings*. Dudweiler, Germany: Digitaldruck Pirrot GmbH. 656–660.
 12. Grusho, A., N. Grusho, M. Zabezhailo, A. Zatsarinny, and E. Timonina. 2017. Information security of SDN on the basis of meta data. *Computer network security*. Eds. J. Rak, J. Bay, I. V. Kotenko, et al. Lecture notes in computer science ser. Springer. 10446:339–347.
 13. Grusho, A. A., M. I. Zabezhailo, N. A. Grusho, and E. E. Timonina. 2018. Informatsionnaya bezopasnost' na osnove metadannykh v komponentno-integratsionnykh arkhitekturakh informatsionnykh sistem [Information security on the basis of meta data in enterprise application integration architecture of information systems]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(2):34–41.
 14. Grusho, A., N. Grusho, M. Zabezhailo, and E. Timonina. 2020. Generation of metadata for network control. *Distributed computer and communication networks*. Eds. V. M. Vishnevskiy, K. E. Samouylov, and D. V. Kozyrev. Lecture notes in computer science ser. Cham: Springer. 12563:723–735. doi: 10.1007/978-3-030-66471-8.55.

Received October 11, 2022

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 119333, Russian Federation; grusho@yandex.ru

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

Zabezhailo Michael I. (b. 1956) — Doctor of Science in physics and mathematics, principal 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; m.zabezhailo@yandex.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; elimon@yandex.ru

Smirnov Dmitry V. (b. 1984) — business partner for IT security department, Sberbank of Russia, 19 Vavilov Str., Moscow 117999, Russian Federation; dvlsmirnov@sberbank.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

ЛОГИЧЕСКАЯ РЕЛЯЦИОННАЯ МОДЕЛЬ СТРУКТУР ДАННЫХ ДЛЯ РЕШЕНИЯ ЗАДАЧ В ПРЕДМЕТНОЙ ОБЛАСТИ УПРАВЛЕНИЯ ЗЕМЛЕПОЛЬЗОВАНИЕМ*

Д. О. Брюхов¹, С. А. Ступников²

Аннотация: Статья относится к теоретическим основам агроинформатики, а более конкретно — к области разработки методов и средств автоматизации управления землепользованием. Предлагается логическая модель структур данных для решения задач в упомянутой предметной области. В виде набора отношений реляционной модели данных и связей между ними формализованы: реестр агроэкологических групп и видов земель, упорядочивающий разнообразие их геоморфологических, почвенных и агрохимических условий; реестр сортов сельскохозяйственных культур и регистр агротехнологий как комплексов технологических операций по управлению производственными процессами сельскохозяйственных культур. Применение логической модели демонстрируется на примерах нескольких обобщенных задач управления землепользованием. Решение каждой из задач реализовано в виде декларативного выражения (запроса) над логической моделью. Модель реализована в реляционной системе управления базами данных (СУБД).

Ключевые слова: реляционная логическая модель; управление землепользованием; решение задач

DOI: 10.14357/19922264220414

1 Введение

Разработка методов и средств автоматизации управления землепользованием — это важное направление современной агроинформатики. В частности, одной из насущных оптимизационных задач стала задача планирования севооборота (на каких участках поля и в какой последовательности следует выращивать сельскохозяйственные культуры). Для автоматизации решения этой задачи в мире активно разрабатываются информационные методы и системы поддержки умного землепользования [1].

В качестве входных данных для проектирования севооборотов используются характеристики сортов культур, климатические данные, данные о доступной технике, данные о полях (включая размеры, форму, характеристики почвы) [2, 3]. В частности, подход, разрабатываемый в Почвенном институте им. В. В. Докучаева, опирается на методологию комплексной оценки сельскохозяйственных земель [4, 5] и предполагает использование данных, накопленных в реестре агроэкологических групп и видов земель (упорядочивающем разнообразие их геоморфологических, почвенных и агрохимических условий), реестре сортов сельскохозяйственных культур (содержащем характеристики урожайности, качества и устойчивости сортов к различным

факторам) и регистре базовых агротехнологий (как комплексов технологических операций по управлению производственными процессами сельскохозяйственных культур).

Данная статья посвящена теоретическому обоснованию подхода с точки зрения структур данных.

Предлагается логическая реляционная модель для решения задач в области управления землепользованием, формализующая упомянутые выше реестры в виде набора отношений реляционной модели данных и связей между ними. Применение логической модели демонстрируется на примерах нескольких обобщенных задач управления землепользованием. Решение каждой из задач реализовано в виде декларативного выражения (запроса) над логической моделью (в тексте статьи приведен один запрос).

Модель реализована в СУБД PostgreSQL, а соответствующая база данных рассматривается как составная часть системы «Умное землепользование», реализуемой в рамках проекта «Актуальные научные задачи стратегии адаптации потенциала землепользования России в условиях беспрецедентных вызовов (экономический кризис, изменения климата, кризис глобальных тенденций природопользования)» (далее — Проект).

* Исследование выполнено при финансовой поддержке Российской Федерации (соглашение с Минобрнауки России № 075-15-2020-805 от 02 октября 2020 г.).

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

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

2 Логическая реляционная модель для реестров агроэкологических групп земель, сортов, агротехнологий

2.1 Логическая модель реестра агроэкологических групп и видов земель

Реестр агроэкологических групп и видов земель упорядочивает разнообразие их геолого-геоморфологических, гидрогеоморфологических, почвенных и агрохимических условий.

На рис. 1 представлен фрагмент логической модели реестра агроэкологических групп и видов земель, содержащий структуры данных о группах земель (отношение AE_GROUP) и видах земель (отношение AE_TYPE), представленных в конкретных провинциях (отношение PROVINCE) природных зон (отношение ZONE). Модель также содержит

связи провинций с конкретными административными областями и районами и структуру характеристик почв каждого вида земель (отношение AE_TYPE_PROPERTY).

2.2 Логическая модель реестра сортов культур

На рис. 2 представлен фрагмент логической модели реестра сортов, содержащий структуру основных данных о сортах культур (отношение CROP_CULTIVAR), включающую название культуры (crop_name) и сорта (crop_cultivar), год районирования (zoning_year), разновидность (variety) и назначение использования сорта (usage). Отношение CROP_CULTIVAR_PROPERTY позволяет связывать с сортом различные дополнительные свойства (indicator), такие как длительность периода вегетации, морозостойкость и зимостойкость, засухоустойчивость, требования к условиям возделывания и др., и указывать их значения (value).

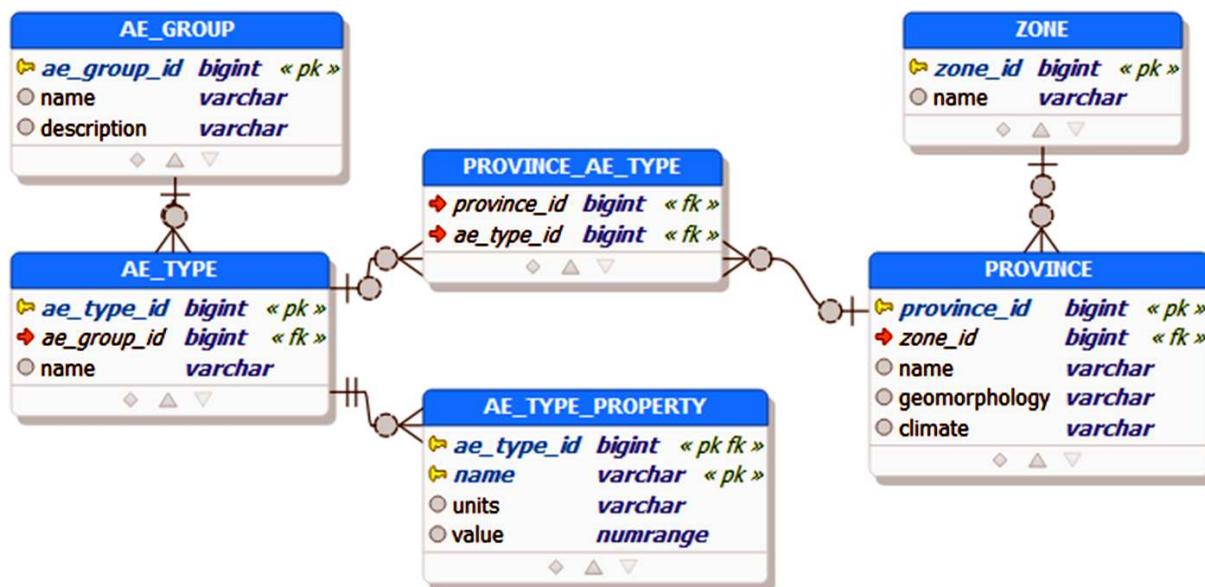


Рис. 1 Фрагмент модели реестра агроэкологических групп и видов земель

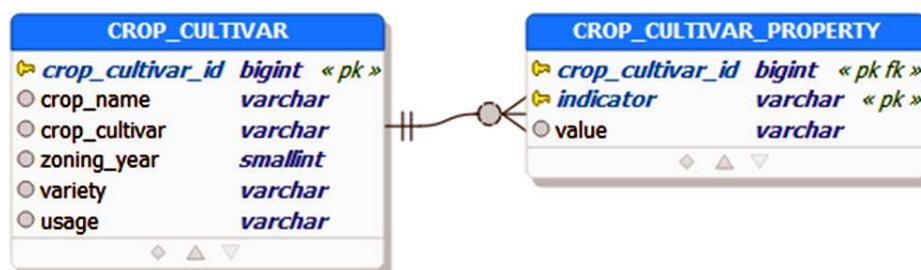


Рис. 2 Фрагмент модели реестра сортов культур

2.3 Логическая модель регистра базовых агротехнологий

Современные агротехнологии представляют собой комплексы технологических операций по управлению производственными процессами сельскохозяйственных культур в агроценозах с целью достижения планируемой урожайности и качества продукции при обеспечении экологической безопасности и экономической эффективности [6].

На рис. 3 представлен фрагмент логической модели регистра агротехнологий, содержащий структуру общих данных об агротехнологии (отношение AGROTECHNOLOGY), применяемые ресурсы интенсификации производства — удобрения и химические средства защиты растений (отношение TECHNOLOGY_INTENSIFICATION_RESOURCES), список процессов и операций агротехнологии (отношение TECHNOLOGY_OPERATIONS). Структура общих данных включает название выращиваемой культуры (crop_name) в провинции (province_name) зоны (zone_name),

применяемый уровень технологии (technology_level), список подходящих агроэкологических видов земель, список рекомендуемых сортов, показатели качества, гарантированную урожайность и др. Структура данных об операциях включает технологические параметры (technology_parameters), сроки (time), используемую технику (equipment).

2.4 Логическая модель структуры хозяйственно-экономических параметров

В отдельную схему модели выделены структуры хозяйственно-экономических параметров использования агроэкологических видов земель.

На рис. 4 представлен фрагмент модели хозяйственно-экономических параметров. Отношение SOIL_TYPE_TECHNOLOGY содержит данные об урожайности (harvest) и себестоимости за тонну (cost) выращивания культур в зависимости от провинции, агроэкологического вида земли,

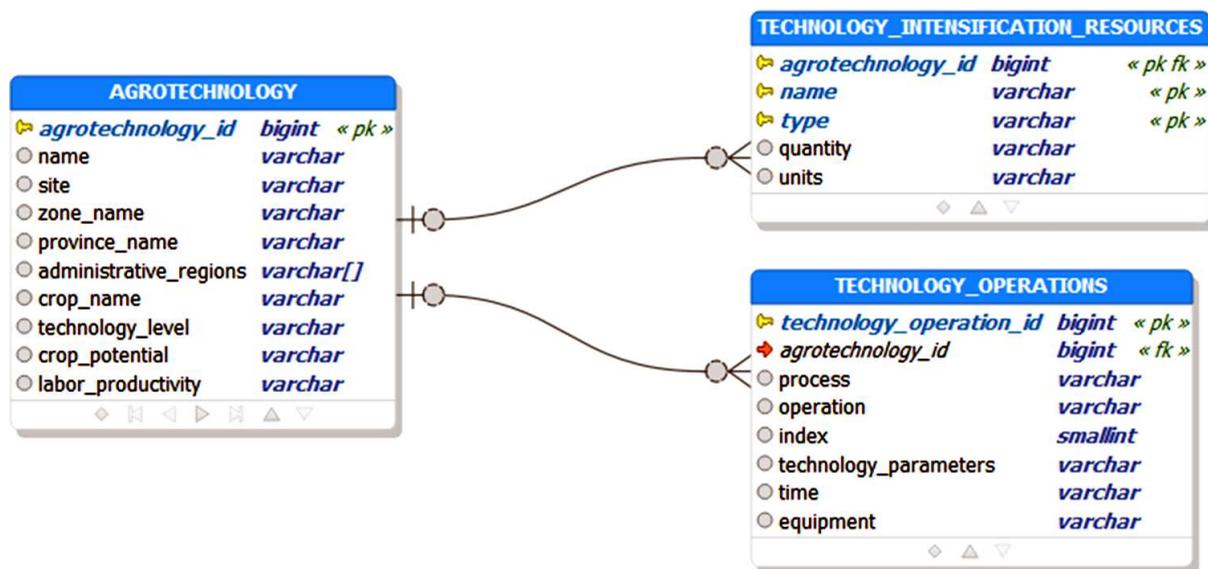


Рис. 3 Фрагмент модели регистра базовых агротехнологий

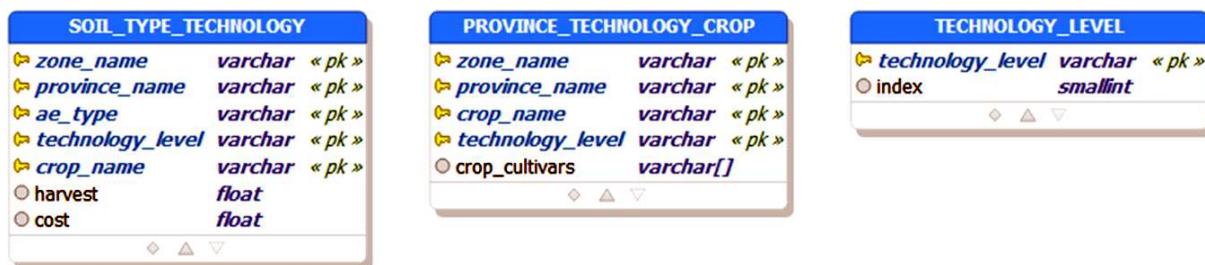


Рис. 4 Фрагмент модели хозяйственно-экономических параметров

уровня интенсивности технологий. Отношение PROVINCE_TECHNOLOGY_CROP содержит данные о подходящих сортах (crop_cultvars) для выращивания на землях конкретных агроэкологических видов и в зависимости от уровня технологий. Отношение TECHNOLOGY_LEVEL представляет собой словарь существующих уровней интенсивности технологий (экстенсивный, нормальный, интенсивный, точный) [6].

2.5 Логическая модель данных о конкретных участках

При планировании землепользования в конкретном хозяйстве одним из этапов является разделение территории на участки с выделением их агроэкологических групп и видов земель. Эти данные выделяются в отдельное отношение FIELD_AE_TYPE (рис. 5), содержащее название конкретного хозяйства (agrostation_name), список его полей/участков (field_name), агроэкологический вид земель каждого участка (ae_type) и площадь каждого участка (area).

Рассмотренная логическая модель базы данных реестров была реализована в СУБД PostgreSQL. В рамках Проекта были собраны и загружены в базу данные о группах и видах земель для Среднерусской провинции южно-таежно-лесной зоны, данные о сортах зерновых и зернобобовых культур и данные о базовых агротехнологиях возделывания этих культур для Центрального региона Нечерноземной зоны.

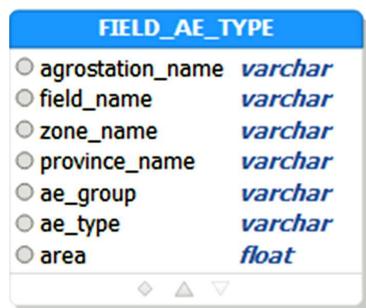


Рис. 5 Фрагмент модели данных о конкретных участках

Полная реализация модели в SQL (structured query language), данные, использовавшиеся для экспериментов, и скрипты их загрузки в базу данных выложены в репозитории github¹.

3 Применение логической модели на примерах обобщенных задач управления землепользованием

Ниже приведены некоторые варианты задач, возникающих при планировании землепользования и решаемых с использованием данных из реестров.

1. Для поля, состоящего из n участков с видами земель ae_1, \dots, ae_n , определить минимальный уровень интенсивности технологий, при котором культура $crop$ даст среднюю урожайность по полю не менее $yield$.
2. Для участка площадью S с видом земли ae определить максимальную по интенсивности технологию для культуры $crop$, при которой себестоимость производства не превышает $cost$.
3. Для участка с видом земли ae определить технологию, обеспечивающую максимальную рентабельность выращивания культуры $crop$.
4. Для участка с видом земли ae определить технологию, обеспечивающую максимальный по объему урожай культуры $crop$.

Реализация решения подобных задач осуществляется с помощью декларативных реляционных запросов над базой данных реестров, реализующей логическую модель из разд. 3. Ввиду ограниченного объема статьи рассмотрим только запрос, реализующий пример задачи 2².

Для примера задачи 2 был выбран участок 1 площадью 0,792 га в хозяйстве Козлово, находящемся в среднерусской провинции южно-таежно-лесной зоны (табл. 1). В качестве культуры выбрана озимая рожь, при этом ограничение на себестоимость выращивания составляет 20 000 руб.

Таблица 1 Данные об участке

Хозяйство	Участок	Вид земель	Площадь, га
Козлово	1	1.1. Дерново-подзолистые средне- и легкосуглинистые культурные	0,792

¹ <https://github.com/sstupnikov/Smart-Landuse>.

² Запросы, реализующие примеры для всех четырех задач, выложены в файле https://github.com/sstupnikov/Smart-Landuse/blob/main/db_queries_ed.txt.

Таблица 2 Результат запроса

Хозяйство	Участок	Культура	Уровень технологий	Себестоимость за гектар, руб.	Площадь, га	Общая себестоимость, руб.
Козлово	1	Озимая рожь	Нормальный	24990	0,792	19 792,08

Реляционный запрос на языке SQL, реализующий задачу, основан на соединении трех отношений: FIELD_AE_TYPE, SOIL_TYPE_TECHNOLOGY и TECHNOLOGY_LEVEL (см. разд. 3). Запрос выглядит следующим образом:

```
select
  d.agrostation_name
  , d.field_name
  , st.crop_name
  , st.technology_level
  , st.harvest*st.cost cost_per_ha
  , d.area
  , st.harvest*st.cost*d.area cost_for_field
from
  economic_parameters."SOIL_TYPE_TECHNOLOGY" st
  , data."FIELD_AE_TYPE" d
  , economic_parameters."TECHNOLOGY_LEVEL" t1
where
  st.zone_name = d.zone_name
    and st.province_name = d.province_name
  and st.ae_type = d.ae_type
  and st.technology_level = t1.technology_level
  and d.agrostation_name = 'Козлово'
  and d.field_name = '1'
  and st.crop_name = 'Озимая рожь'
  and st.harvest*st.cost*d.area < 20000
order by
  t1.index desc
limit 1
```

Себестоимость вычисляется как произведение урожайности (harvest), себестоимости за тонну (cost) и площади (area). Максимальный уровень интенсивности выбирается при помощи упорядочения (order by) индекса уровня технологии (index) по убыванию (desc) и выбора первого результата (limit 1).

В табл. 2 представлен результат этого запроса, содержащий данные об удовлетворяющей заданному бюджету себестоимости выращивания на участке озимой ржи для максимального уровня технологии.

4 Заключение

Направление разработки методов и средств автоматизации управления землепользованием — одно из важнейших в современной агроинформатике. В работе рассмотрена логическая модель структур данных, формализующая данные об агроэкологических группах земель, сортах и агротехнологиях. Показано, что эта модель может служить основой для решения задач планирования землепользования.

Направление дальнейшей работы — встраивание реализованной базы данных в систему «Умное землепользование» и реализация полного набора задач планирования землепользования в виде реляционных запросов.

Литература

1. *Schöning J., Wachter P., Trautz D.* Crop rotation and management tools for every farmer? The current status on crop rotation and management tools for enabling sustainable agriculture worldwide // *Smart Agricultural Technology*, 2023. Vol. 3. Art. 100086. 15 p. doi: 10.1016/j.atech.2022.100086.
2. *Pahmeyer C., Kuhn T., Britz W.* “FruchtFolge”: A crop rotation decision support system for optimizing cropping choices with big data and spatially explicit modeling // *Comput. Electron. Agr.*, 2021. Vol. 181. Art. 105948. 11 p.
3. *Kim J. S., Kisekka I.* FARMS: A geospatial crop modeling and agricultural water management system // *ISPRS Int. J. Geo-Inf.*, 2021. Vol. 10. No. 8. Art. 553. 17 p. doi: 10.3390/ijgi10080553.
4. *Кирюшин В. И.* Методология комплексной оценки сельскохозяйственных земель // *Почвоведение*, 2020. № 7. С. 871–879.
5. *Кирюшин В. И., Дубачинская Н. Н., Юрова А. Ю.* Комплексная оценка сельскохозяйственных земель на примере Южного Урала // *Почвоведение*, 2021. № 11. С. 1363–1375.
6. *Кирюшин В. И., Кирюшин С. В.* Агротехнологии. — СПб.: Лань, 2022. 464 с.

Поступила в редакцию 29.08.22

LOGICAL RELATIONAL MODEL OF DATA STRUCTURES FOR PROBLEM SOLVING IN LAND USE MANAGEMENT

D. O. Briukhov and S. A. Stupnikov

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

Abstract: The paper belongs to theoretical foundations of agroinformatics. To be more precise, it concerns development of methods and tools for automatization of land use management. A logical model for data structures intended for problem solving in the subject area is proposed. Registry for agroecological groups and types of lands organizing diversity of their geomorphological, soil, and agrochemical conditions, registry for crops and cultivars, and registry for agrotechnologies as complexes of technological operations for management of crop production processes are formalized as sets of relations of the relational data model and relationships between them. Application of the logical model is demonstrated via several generic problems of land use management. Each problem is implemented as a declarative query expression over the logical model. The model is implemented in a relational database management system.

Keywords: logical relational model; land use management; problem solving

DOI: 10.14357/19922264220414

Acknowledgments

The study was financially supported by the Ministry of Education and Science of the Russian Federation (agreement No. 075-15-2020-805 dated October 2, 2020).

References

1. Schöning, J., P. Wächter, and D. Trautz. 2023. Crop rotation and management tools for every farmer? The current status on crop rotation and management tools for enabling sustainable agriculture worldwide. *Smart Agricultural Technology* 3:100086. 15 p. doi: 10.1016/j.atech.2022.100086.
2. Pahlmeyer, C., T. Kuhn, and W. Britz. 2021. “FruchtFolge”: A crop rotation decision support system for optimizing cropping choices with big data and spatially explicit modeling. *Comput. Electron. Agr.* 181:105948. 11 p.
3. Kim, J. S., and I. Kisekka. 2021. FARMS: A geospatial crop modeling and agricultural water management system. *ISPRS Int. J. Geo-Inf.* 10(8):553. 17 p. doi: 10.3390/ijgi10080553.
4. Kiryushin, V. I. 2020. Methodology for integrated assessment of agricultural land. *Eurasian Soil Sci.* 53(7):960–967.
5. Kiryushin, V. I., N. N. Dubachiskaya, and A. Yu. Yurova. 2021. Comprehensive assessment of agricultural land by the example of the Southern Urals. *Eurasian Soil Sci.* 54(11):1721–1731.
6. Kiryushin, V. I., and S. V. Kiryushin. 2022. *Agrotekhnologii* [Agrotechnologies]. St. Petersburg: Lan?. 464 p.

Received August 29, 2022

Contributors

Briukhov Dmitry O. (b. 1971) — 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; dbriukhov@ipiran.ru

Stupnikov Sergey A. (b. 1978) — 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; sstupnikov@ipiran.ru

ЕДИНАЯ МОДЕЛЬ ГОСУДАРСТВЕННЫХ ДАННЫХ: СЦЕНАРИИ РАЗВИТИЯ

А. П. Сучков¹

Аннотация: Обсуждается проблема информационного взаимодействия разнородных информационных систем, которая должна решаться путем создания, внедрения и поддержания единых моделей данных (ЕМД) в рамках основных разделов предметных областей и в перспективе в рамках всей предметной области информационного взаимодействия в национальном масштабе. На основе онтологического подхода исследуется проблема поиска эффективных и оптимальных путей формирования моделей государственных данных. Также рассматриваются сценарии интеграции ведомственных систем.

Ключевые слова: информационное взаимодействие; единая модель данных; онтология; сценарии интеграции

DOI: 10.14357/19922264220415

1 Введение

Мировой и отечественный опыт — как положительный, так и отрицательный — говорит о том, что проблема информационного взаимодействия должна решаться путем создания, внедрения и поддержания ЕМД в рамках основных разделов предметных областей и в перспективе в рамках всей предметной области информационного взаимодействия в национальном масштабе [1]. Когда речь идет о национальном масштабе, основную роль в этом процессе должно играть государство, так как оно заинтересовано в повышении эффективности создания, сбора и использования *государственных данных* как для предоставления государственных и муниципальных *услуг* и осуществления государственных и муниципальных *функций*, так и для обеспечения потребности физических и юридических лиц в доступе к информации [2].

Так, с 2006 г. развивается National Information Exchange Model (NIEM) в США [3], в 2019 г. положено начало созданию Национальной системы управления данными (НСУД) РФ [2], в 2020 г. Еврокомиссия представила “A European strategy for data” [4].

Текущая версия NIEM 5.0 включает 15 доменов, более 11 000 взаимосвязанных информационных объектов (сущностей), для описания которых используются свыше 22 000 атрибутов, и объединяет более 70 информационно-лингвистических систем (словарей терминов, классификаторов, рубрикаторов и т. п.) различных ведомств. Однако модель пока не охватывает все разделы национальных данных, далеко не все ведомства ее используют.

По проекту НСУД осуществлена разработка и введение в опытную эксплуатацию информаци-

онно-технологических компонентов НСУД, включая систему защиты информации, необходимых для проведения эксперимента по интеграции крупнейших информационных систем на уровне метаданных (Росреестр, Федеральное казначейство, Федеральная налоговая служба, Пенсионный фонд России и др.). Эксперимент выявил существенные пробелы в нормативно-правовой базе.

Европейская стратегия представляет собой часть инициативы по цифровой трансформации Европы и находится на начальной стадии, определены 10 основных доменов, выделено финансирование (6 млрд евро на три года) и принят «Скользкий план стандартизации информационно-коммуникационных технологий 2020» [5].

Помимо этих проектов можно рассмотреть достаточное число промежуточных паллиативных решений по созданию систем межведомственного взаимодействия в России — это Система межведомственного электронного взаимодействия (СМЭВ) [6], Система межведомственного электронного документооборота федеральных органов исполнительной власти (МЭДО) [7], которые входят в инфраструктуру электронного правительства «Безопасный город» [8], и, конечно, Система распределенных ситуационных центров, взаимодействующих по единому регламенту [9]. Как очень верно утверждается в [2]: «Одна из основных задач НСУД — установление правил создания *модели государственных данных*, основанной на принципах постоянного развития, постепенного наполнения, консистентности и непротиворечивости, включая разработку описаний, взаимосвязей сущностей и форматов, и, по сути, является онтологией. . .» Такая постановка проблемы у нас

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

впервые появляется в официальном нормативно-правовом документе. На основе такого онтологического подхода в статье исследуется проблема поиска эффективных и оптимальных путей формирования моделей государственных данных, тем более что такое исследование все еще актуально в силу того, что отечественные разработки находятся в начальной стадии. Также далее рассматриваются сценарии интеграции ведомственных систем.

2 Единая модель государственных данных

Информационную модель данных некоторой предметной области можно трактовать как *онтологию*. Онтология описывается совокупностью базовых объектов, классов (концептов), атрибутов, отношений и правил.

Рассматривая обобщенную онтологию всего спектра данных, потенциально охватывающих все виды национального межведомственного информационного взаимодействия (МВИВ), можно определить так называемую *верхнюю* онтологию, которая охватывает все доменные онтологии различных предметных областей народного хозяйства.

Из необходимости интеграции разнородных баз данных возникла задача *отображения* онтологий, разработанных независимо друг от друга и имеющих, таким образом, свое собственное представление базовых объектов. Отображение двух онтологий означает, что для каждого понятия, отношения или объекта одной онтологии подыскиваются соответствующие элементы в другой онтологии [10]. Эта проблема стала важнейшей задачей в области искусственного интеллекта.

Основные свойства единой модели государственных данных (ЕМГД):

- ЕМГД должна формироваться в виде онтологии, характеризующейся совокупностью базовых объектов (сущностей), классов, атрибутов и отношений;
- ЕМГД должна быть функционально ориентированной на реализацию законодательно закрепленных государственных функций и услуг, а также обеспечивать информационную поддержку процессов управления на всех уровнях;
- ЕМГД должна быть независимой от поставщиков операционных систем, носителей информации и приложений, что может быть достигнуто путем применения языков описания XML или JSON;
- ЕМГД должна быть оснащена методическими, технологическими и организационными сред-

ствами поддержания ее в актуальном состоянии.

Единая модель данных должна иметь доменную структуру, отражающую все разделы предметной области ЕМГД.

Организационно, по мнению автора, центральный орган ЕМГД должен быть государственным и вневедомственным. Для организации коллективной работы, поддержания методологической базы и обучения экспертов, разработчиков и потребителей необходимо создать информационный портал ЕМГД. С технологической точки зрения необходимо обеспечить оптимальные по соотношению «цена–качество» сценарии создания и ведения ЕМГД, к обоснованию которых и переходим.

3 Математическая модель процесса формирования единой модели государственных данных

Отображение онтологий в процессе интеграции различных информационных систем может проходить по разным сценариям. Однако у всех сценариев есть одна общая черта: по мере построения отображений онтологий происходит формирование *ядра интеграционной модели данных*, реализующей интегральную онтологию данных. Способы развития этого ядра в процессе межведомственной интеграции и определяют ее основные сценарии. Для того чтобы оценить эффективность процесса формирования ЕМГД, необходимо оценить ориентировочную стоимость интеграции автоматизированной информационной системы (АИС) при различных сценариях в рамках формирования ЕМГД.

Означим $\Omega_{\text{в}}$ — верхнюю, т.е. гипотетическую всеохватывающую онтологию ЕМГД, $\Omega_{\text{н}}$ — текущую онтологию ядра интеграционной модели данных ЕМГД, $\Omega_{\text{с}}$ — онтологию интегрируемой АИС. Сценарии развертывания во времени процесса формирования ЕМГД в виде онтологии подразделяются в зависимости от способов пополнения интегрального ядра и получаемого соотношения объемов автоматического и автоматизированного взаимодействия между АИС (рис. 1).

Выделим четыре сценария формирования ЕМГД: первые два связаны со способами формирования первоначального и окончательного интеграционного ядра; третий и четвертый — с доменным подходом.

Сценарий I. Осуществляется начальное формирование небольшой онтологии интеграционного

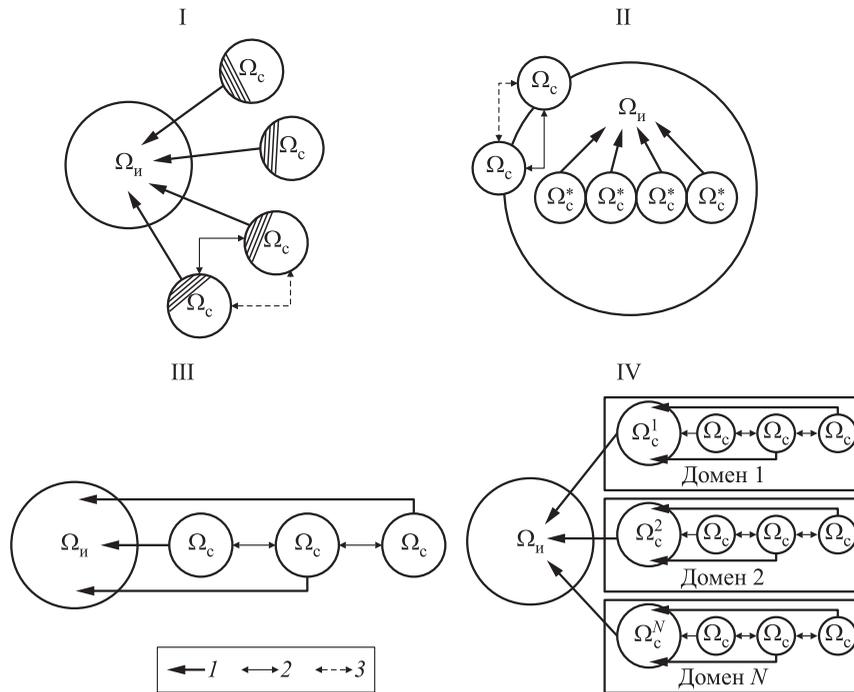


Рис. 1 Сценарии формирования ЕМД: $\Omega_{и}$ — интегральная онтология МВИВ; $\Omega_{с}$ — онтология интегрируемой АИС; $\Omega_{с}^*$ — онтология первоначального пула интегрируемой АИС; $\Omega_{с}^i$ — онтология i -го домена интегрируемой АИС; 1 — пополнение интегральной онтологии МВИВ; 2 — автоматическое взаимодействие АИС; 3 — взаимодействие АИС с помощью адаптеров

ядра ЕМГД, содержащей часто встречающиеся в $\Omega_{с}$ базовые сущности. Интеграция АИС осуществляется с учетом существования этого ядра, но без его изменения. Автоматическое взаимодействие АИС осуществляется только в рамках онтологии интеграционного ядра относительно небольшого размера. Остальные взаимодействия происходят с помощью адаптеров. Пример такой интеграции — СМЭВ, портал госуслуг.

Сценарий II. Осуществляется начальное формирование онтологии интеграционного ядра ЕМГД на основе представительной выборки онтологий из всего состава АИС (первоначального пула), представляющей собой совокупность базовых элементов в $\Omega_{в}$. Интеграция АИС осуществляется с учетом существования этого ядра, но без его изменения. Автоматическое и адаптерное взаимодействие аналогично первому сценарию, но на существенно большем ядре. Примеры — GJXDM (Global Justice XML Data Model — Глобальная XML-модель данных правоохранительной сферы на базе веб-технологий, первая версия НИЕМ).

Сценарий III. Формирование ядра ЕМГД осуществляется по мере интеграции АИС путем последовательного присоединения онтологий $\Omega_{с}$

к $\Omega_{и}$ без создания начального интеграционного ядра. Все взаимодействия интегрированных АИС — автоматизированные. Такой сценарий реализуется при формировании доменов НИЕМ и в НСУД (в части формирования метаданных).

Сценарий IV. Формирование ядра ЕМГД осуществляется по третьему сценарию, но интегрируются не отдельные АИС, а целые домены предметной области, которые, в свою очередь, формируются также по третьему сценарию из АИС каждого домена. Этот сценарий реализован в НИЕМ.

Так как построение отображений онтологий в процессе их интеграции происходит поэлементно, затраты на интеграцию базового набора объектов онтологии пропорциональны их числу. Введем понятие μ — «мощности» онтологии, которую положим равной числу содержащихся в ней базовых элементов, и обозначим $\mu_{в} = \Omega_{в}$, $\mu_{и} = \Omega_{и}$ и $\mu_{с} = \Omega_{с}$.

Мощность верхней онтологии составляет более десятка тысяч базовых элементов, поэтому для моделирования процессов интеграции вполне применимы вероятностные методы. Процесс интеграции представим следующим образом. В общем случае онтология $\Omega_{с}$ интегрируемой АИС содержит часть объектов онтологии ядра, для которых нет надоб-

ности поиска соответствий, и часть объектов верхней онтологии, не входящих в $\Omega_{и}$. Если предположить, что состав Ω_c случаен, математическое ожидание числа объектов Ω_c , не принадлежащих $\Omega_{и}$, составляет $\mu_c(\mu_B - \mu_{и})/\mu_B$.

В [11] осуществлен анализ соотношения этих характеристик и показано, что среди первых трех сценариев наилучшим оказался третий. Теперь включим в рассмотрение четвертый сценарий. Итак, по первому и второму сценарию для подключения N АИС затраты пропорциональны мощности пополнения ядра и составляют [11]:

$$C_1 = \mu_{и1} + N\mu_c \frac{\mu_B - \mu_{и1}}{\mu_B};$$

$$C_2 = \mu_{и2} + N\mu_c \frac{\mu_B - \mu_{и2}}{\mu_B}.$$

Соотношение затрат по четырем сценариям представлено на рис. 2. При $N = \mu_B/\mu_c$ имеем $C_1 = C_2$ и с ростом N C_1 превышает C_2 .

Для третьего сценария затраты для N АИС составляют [11]:

$$C_3 = \mu_B \left(1 - \left(1 - \frac{\mu_c}{\mu_B} \right)^{N-1} \right),$$

т. е. C_3 стремится в геометрической прогрессии к μ_B и при достаточно большом N $\mu_B - C_3 < 1$, что означает $\Omega_{и} = \Omega_B$ в силу дискретности процесса.

Оценим границу G_3 значения N , при котором $\mu_B - C_3 < 1$. Имеем

$$\mu_B - C_3 = \mu_B - \mu_B \left(1 - \left(1 - \frac{\mu_c}{\mu_B} \right)^{N-1} \right) < 1.$$

Отсюда следует, что верхняя онтология достигается при

$$N > \frac{\ln \mu_B}{\ln(\mu_B/(\mu_B - \mu_c))} = G_3.$$

Покажем, что четвертый сценарий более оптимален, чем третий. Для четвертого сценария общие затраты суть сумма затрат на создание и присоединение к ядру ЕМГД отдельных доменов. Пусть N_i , $i = 1, \dots, R$, — число АИС в i -м домене мощностью μ_B^i , R — число доменов. Тогда

$$C_4 = \sum_{i=1}^R f_i(N),$$

где

$$f_1(N) = \begin{cases} 0, & N \leq \sum_1^i N_{j-1}, N_0 = 0; \\ \mu_B^i \left(1 - \left(1 - \frac{\mu_c}{\mu_B^i} \right)^{N - \sum_2^i N_{j-1} - 1} \right), & N > \sum_1^i N_{j-1}. \end{cases}$$

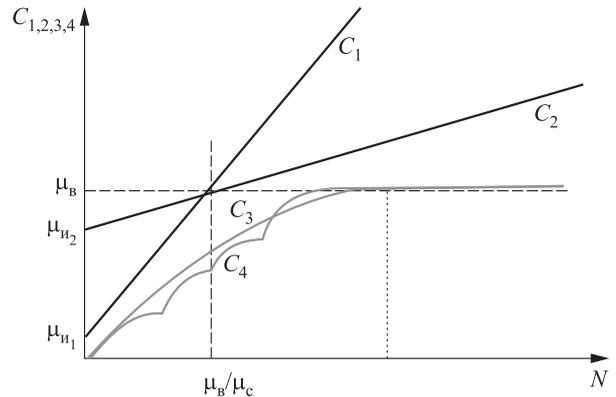


Рис. 2 Соотношение затрат по четырем сценариям интеграции

По четвертому сценарию затраты на формирование доменных онтологий уменьшаются по сравнению с третьим сценарием из-за большей скорости пополнения ядра доменов на начальных стадиях.

Оценим границу G_4 значения N , при котором $\mu_B - C_4 < 1$. Для упрощения вычислений предположим, что μ_B^i одинаковы для всех доменов и равны μ_B^d , при этом $\mu_B^d = \mu_B/R$. Имеем для каждого домена

$$\mu_B^d - \mu_B^d \left(1 - \left(1 - \frac{\mu_c}{\mu_B^d} \right)^{N-1} \right) < 1.$$

Отсюда

$$N > R \frac{\ln(\mu_B/R)}{\ln(\mu_B/(\mu_B - R\mu_c))} = G_4.$$

Обозначим

$$G(x) = x \frac{\ln(\mu_B/x)}{\ln(\mu_B/(\mu_B - x\mu_c))}, \quad x > \frac{\mu_B}{\mu_c}.$$

Очевидно, $G(1) = G_3$, $G(R) = G_4$. Можно показать, что на интервале $1 < x < \mu_B/\mu_c$ функция $G(x)$ имеет отрицательную производную и, значит, монотонно убывает (рис. 3) (опускаем вычисления производной в силу их громоздкости). Следовательно, $G(1) = G_3 > G(R) = G_4$. Это означает, что по четвертому сценарию, предполагающему подомненную интеграцию ЕМГД, верхняя онтология ЕМГД достигается существенно быстрее, чем по третьему сценарию.

Самым плохим вариантом по суммарным и постоянно растущим затратам является первый сценарий. Во втором сценарии затраты также постоянно растут, но медленнее. Третий сценарий обеспечивает меньшие суммарные затраты, которые к тому же после интеграции определенного числа АИС вообще не растут, так как не требуют дополнительных расходов. Четвертый сценарий дает такой же результат, однако существенно улучшает скорость

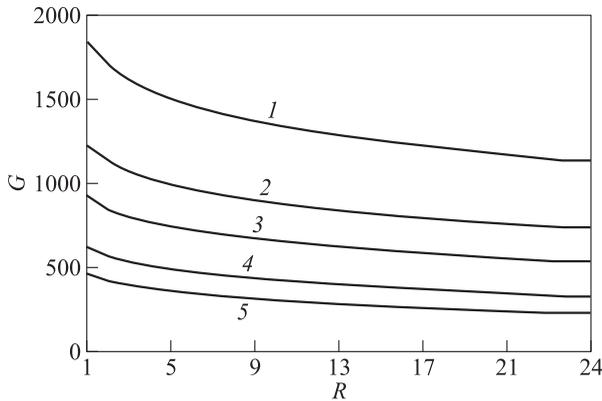


Рис. 3 Семейство функций $G(x)$: 1 — $\mu_c = 50$; 2 — 75; 3 — 100; 4 — 150; 5 — $\mu_c = 200$

и качество интеграции и уменьшает организационные затраты, так как не требует концентрации универсальных экспертов *всей* верхней онтологии ЕМГД. Это показывает, что, насколько бы ни была малозатратной разработка отдельного адаптера, с ростом числа подключаемых систем расходы на это превысят расходы на создание единой модели верхней онтологии (ЕМГД).

Следует заметить, что оценка затрат проведена лишь с учетом отображения онтологий без учета расходов на создание и поддержание ЕМГД. Сюда входят расходы на создание организационной системы ЕМГД, проведение научно-исследовательских и опытно-конструкторских работ (НИОКР) по формированию структуры и правил создания единой XML-модели. Однако это не влияет на корректность сделанных выводов, так как эти расходы не зависят от числа подключаемых систем.

4 Оптимальный сценарий интеграции ведомственной системы

Рассмотрим гипотетическую ведомственную распределенную систему управления (ВРСУ) со своей информационной моделью данных. Регламент информационного взаимодействия (РИВ) в ВРСУ можно рассматривать как правовой акт и сопутствующие соглашения, позволяющие осуществлять информационное, в том числе межведомственное, взаимодействие, обеспечивающее функциональные потребности ВРСУ для реализации своих функций и услуг. Если предположить, что будет взят курс на создание ЕМГД в полном объеме и в национальном масштабе, то каковы могут быть пути адаптации РИВ в этом процессе? Несомненно, магистральным направлением развития

станет создание ЕМД ВРСУ, совместимой с национальной системой ЕМГД. Рассмотрим возможные варианты такого развития.

1. Участие в создании ЕМГД:

- (а) создание и ведение ЕМГД — интегрального ядра — и присоединение доменов;
- (б) создание первоначального центрального ядра на основе созданного первоначального домена.

2. Присоединение к ЕМГД:

- (а) интеграция в ЕМГД путем создания адаптеров;
- (б) использование ЕМГД для создания и ведения его подмножества — ЕМД для ВРСУ.

Рассмотрим перечисленные варианты.

Вариант 1а. Полномасштабное создание и ведение ЕМГД предполагает:

- разработку методологии, структуры и правил формирования XML-модели государственных данных;
- разработку программного инструментария по созданию, ведению, модернизации элементов XML-модели государственных данных;
- тестирование этих элементов на одном из доменов предметной области (по аналогии с NIEM, где первоначально была отработана GJXDM);
- создание и ведение общедоступного портала для привлечения к обсуждению, тестированию, обучению широкого круга экспертов всех предметных областей ЕМГД, а также инструментария для формирования подмоделей для начала работы в доменах;
- организацию сети доменов с определением состава экспертных центров в доменах;
- создание технологий и интеграцию моделей доменов в общую XML-модель государственных данных.

Вариант 1б. Создание первоначального центрального ядра на основе созданного домена предполагает:

- проведение НИОКР по формированию методологии, структуры и правил формирования и развития единой XML-модели;
- создание на этой основе первоначального домена;
- первоначальное наполнение ядра ЕМГД на основе общей части созданного домена;
- передачу наработанных технологий в центральный орган ЕМГД для внедрения.

Сравнительная экспертная оценка вариантов развития РИВ ВРСУ

Варианты развития РИВ ВРСУ		p_1	p_2	p_3	P
1	Участие в создании ЕМГД:				
(а)	создание и ведение ЕМГД — интегрального ядра и присоединение доменов	2	4	4	7
(б)	создание первоначального центрального ядра на основе созданного домена	1	3	2	6
2	Присоединение к ЕМГД:				
(а)	интеграция в ЕМГД путем создания адаптеров	3	1	1	8
(б)	использование ЕМГД для создания и ведения подмножества ЕМГД для ВРСУ	4	2	2	9

Вариант 2а. Интеграция в ЕМГД путем создания адаптеров — это первый этап, позволяющий подстроиться к изменяющейся структуре данных, поступающих извне от центрального звена ВРСУ. Такие процессы, по всей видимости, апробированы и отлажены в настоящее время, так как это основа существующей технологии запроса и получения данных из внешних источников, однако существенно увеличивается период реализации.

Вариант 2б. Преобразование информационной модели данных ВРСУ в форматы ЕМГД связано с формированием подмножества ЕМГД для целей функционирования ВРСУ и ее центрального звена. Как уже упоминалось, один из необходимых инструментариев ЕМГД — формирование подмножеств общей XML-модели по различным срезам предметной области, т. е. формирование ЕМД ВРСУ. После формирования ЕМД ВРСУ необходимо провести корректировку структур данных информационных систем центрального звена и взаимодействующих систем всех уровней.

Для обоснования оптимального сценария развития РИВ в ЕМД ВРСУ необходимо провести сравнительный анализ всех выбранных вариантов. Для сравнения этих вариантов определим показатели, способы их оценки и критерии для анализа.

Первый показатель (p_1) — сравнительное время, необходимое для реализации варианта (от 1 до 4): $p_1 \rightarrow \min$. Второй показатель (p_2) — сравнительные затраты на реализацию (от 1 до 4): $p_2 \rightarrow \min$. Третий показатель (p_3) — качество реализации государственных функций и услуг (от 1 до 4): $p_3 \rightarrow \max$. Интегральный показатель $P = p_1 + p_2 + (5 - p_3) \rightarrow \min$.

В таблице приведены экспертные оценки показателей на основе изложенных выше соображений. Оценки показателей осуществлялись следующим образом: для каждого показателя варианты ранжировались по возрастанию и затем им присваивались порядковые номера от 1 до 4. Возможны и другие способы оценки интегрального показателя, например с установлением весов частных показателей. В соответствии с выбранной методологией оценки оптимальным путем развития РИВ оказался вариант 1б — создание и включение в РИВ первоначаль-

ного центрального ядра ЕМГД на основе созданного домена (см. таблицу). Конечно, эти оценки относительно и зависят еще от многих факторов и обстоятельств, не исключая даже амбициозность ставящихся руководством ведомства задач. Несомненно, данный вариант позволяет ведомству наиболее быстро подключиться и даже возглавить процессы решения проблемы информационного взаимодействия, включая межведомственное в целом.

Дальнейший ход этой деятельности, по мнению автора, должен осуществляться в рамках созданной отдельной организационной системы ЕМГД федерального уровня при активном участии и ведущей роли ведомства и обеспечения эволюции РИВ в ЕМД ВРСУ.

Литература

1. Зацаринный А. А., Сучков А. П. Информационное взаимодействие в распределенных системах ситуационного управления. — М.: ФИЦ ИУ РАН, 2021. 256 с.
2. Концепция создания и функционирования национальной системы управления данными: Распоряжение Правительства РФ от 3 июня 2019 г. № 1189-р. <http://static.government.ru/media/files/jYh27VlwiZs44qa0IXJIZCa3uu7qqLzl.pdf>.
3. National Information Exchange Model. <https://www.niem.gov>.
4. A European strategy for data. <https://www.tadviser.ru/index.php>.
5. Rolling Plan for ICT Standardisation. <https://joinup.ec.europa.eu/collection/rolling-plan-ict-standardisation/rolling-plan-2020>.
6. Технологический портал СМЭВ. <http://smev.gosuslugi.ru>.
7. Система межведомственного электронного документооборота федеральных органов исполнительной власти. <https://cs-consult.ru/products/delo/medo>.
8. Концепция построения и развития аппаратно-программного комплекса «Безопасный город»: Постановление Правительства РФ от 3 декабря 2014 г. № 2446-р.
9. О формировании системы распределенных ситуационных центров, работающих по единому регламенту.

- ту взаимодействия: Указ Президента РФ № 648 от 25.07.2013 г.
10. *Лапшин В. А.* Онтологии в компьютерных системах. — М.: Научный мир, 2010. 224 с.
11. *Сучков А. П.* Анализ процессов межведомственного информационного взаимодействия // Системы и средства информатики, 2018. Т. 28. № 3. С. 118–130.

Поступила в редакцию 15.12.21

UNIFIED MODEL OF NATIONAL DATA: DEVELOPMENT SCENARIOS

A. P. Suchkov

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

Abstract: The problem of information interaction between heterogeneous information systems is considered. Such point of view is taken that this problem must be solved by creating, implementing, and maintaining unified data models within the main sections of the subject areas. In the perspective, the scope must be enlarged so as to encompass the entire subject area of information interaction on a national scale. Based on the ontological approach, the author proposes the solution to the problem of finding effective as well as optimal ways to form national data models. The scenarios of integration of departmental systems are considered as well.

Keywords: information interaction; unified data model; ontology; integration scenarios

DOI: 10.14357/19922264220415

References

- Zatsarinny, A. A., and A. P. Suchkov. 2021. *Informatsionnoe vzaimodeystvie v raspredelennykh sistemakh situatsionnogo upravleniya* [Information interaction in distributed situational management systems]. Moscow: FRC CSC RAS. 256 p.
- Kontseptsiya sozdaniya i funktsionirovaniya natsional'noy sistemy upravleniya dannymi: Rasporyazhenie Pravitel'stva RF ot 3 iyulya 2019 g. No. 1189-r [The concept of construction and development of the hardware and software complex “Safe City”: Government order No. 1189-r dated 03.06.2019]. Available at: <http://static.government.ru/media/files/jYh27VlwiZs44qa0lXJIZCa3uu7qqLzl.pdf> (accessed November 1, 2022)
- National Information Exchange Model. Available at: <https://www.niem.gov> (accessed November 1, 2022).
- A European strategy for data. Available at: <https://www.tadviser.ru/index.php> (accessed November 1, 2022).
- Rolling Plan for ICT Standardization. Available at: <https://joinup.ec.europa.eu/collection/rolling-plan-ict-standardisation/rolling-plan-2020> (accessed November 1, 2022).
- Tekhnologicheskii portal SMEV [SMEV Technology Portal]. Available at: <http://smev.gosuslugi.ru> (accessed November 1, 2022).
- Sistema mezhhvedomstvennogo elektronnoho dokumentooborota federal'nykh organov ispolnitel'noy vlasti [System of interdepartmental electronic document management of the federal executive authorities]. Available at: <https://cs-consult.ru/products/delo/medo> (accessed November 1, 2022).
- Kontseptsiya postroeniya i razvitiya apparatno-programmnogo kompleksa “Bezopasnyy gorod”: Postanovlenie Pravitel'stva RF ot 3 dekabrya 2014 g. No. 2446-r [The concept of construction and development of the hardware and software complex “Safe City”: Government decree No. 2446-r dated 03.12.2014]. Available at: <http://static.government.ru/media/files/OapBppc8jyA.pdf> (accessed November 1, 2022).
- O formirovaniy sistemy raspredelennykh situatsionnykh tsevtrov, rabotayushchikh po edinomu reglamentu vzaimodeystviya: Ukaz Prezidenta RF No. 648 ot 25.07.2013 [On the formation of a system of distributed situational centers operating under the single rules of interaction. Presidential Decree No. 648 dated 25.07.2013]
- Lapshin, V. A. 2010. *Ontologii v komp'yuternykh sistemakh* [Ontology in complex systems]. Moscow: Nauchnyy mir. 224 p.
- Suchkov, A. P. 2018. Analiz protsessov mezhhvedomstvennogo informatsionnogo vzaimodeystviya [Analysis of processes of interdepartmental information interaction]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(3):118–130.

Received December 15, 2021

Contributor

Suchkov Alexander P. (b. 1954) — Doctor of Science in technology, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; ASuchkov@frccsc.ru

Агаларов Явер Мирзабекович (р. 1952) — кандидат технических наук, доцент, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Базилевский Михаил Павлович (р. 1987) — кандидат технических наук, доцент Иркутского государственного университета путей сообщения

Бесчастный Виталий Александрович (р. 1992) — кандидат физико-математических наук, старший преподаватель кафедры прикладной информатики и теории вероятностей Российского университета дружбы народов

Борисов Андрей Владимирович (р. 1965) — доктор физико-математических наук, главный научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Босов Алексей Вячеславович (р. 1969) — доктор технических наук, главный научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Брюхов Дмитрий Олегович (р. 1971) — кандидат технических наук, старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Гайдамака Юлия Васильевна (р. 1971) — доктор физико-математических наук, профессор кафедры прикладной информатики и теории вероятностей Российского университета дружбы народов; старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Грушо Александр Александрович (р. 1946) — доктор физико-математических наук, профессор, главный научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Грушо Николай Александрович (р. 1982) — кандидат физико-математических наук, старший научный сотрудник Института проблем информатики

Федерального исследовательского центра «Информатика и управление» Российской академии наук

Дунсяо Гу (р. 1980) — кандидат технических наук, профессор технологического университета г. Хэфэй (КНР)

Дьяченко Юрий Георгиевич (р. 1958) — кандидат технических наук, старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Дюкова Анастасия Петровна (р. 1995) — аспирант Федерального исследовательского центра «Информатика и управление» Российской академии наук

Дюкова Елена Всеволодовна (р. 1945) — доктор физико-математических наук, главный научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук; доцент факультета вычислительной математики и кибернетики Московского государственного университета имени М. В. Ломоносова

Забжайло Михаил Иванович (р. 1956) — доктор физико-математических наук, доцент, главный научный сотрудник Вычислительного центра им. А. А. Дородницына 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

Золотарев Олег Васильевич (р. 1959) — кандидат технических наук, заведующий кафедрой информационных систем в экономике и управлении Института информационных систем и инженерно-компьютерных технологий Российского нового университета

Иванов Алексей Владимирович (р. 1976) — кандидат технических наук, старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Мачнев Егор Александрович (р. 1996) — аспирант кафедры прикладной информатики и теории вероятностей Российского университета дружбы народов

Острикова Дарья Юрьевна (р. 1988) — кандидат физико-математических наук, доцент кафедры прикладной информатики и теории вероятностей Российского университета дружбы народов

Пешкова Ирина Валерьевна (р. 1975) — кандидат физико-математических наук, доцент Петрозаводского государственного университета; старший научный сотрудник Федерального исследовательского центра «Карельский научный центр Российской академии наук»

Рождественский Юрий Владимирович (р. 1952) — кандидат технических наук, ведущий научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Смирнов Дмитрий Владимирович (р. 1984) — бизнес-партнер по ИТ департамента безопасности ПАО «Сбербанк России»

Соколов Игорь Анатольевич (р. 1954) — доктор технических наук, академик РАН, директор Федерального исследовательского центра «Информатика и управление» Российской академии наук

Степченков Юрий Афанасьевич (р. 1951) — кандидат технических наук, заведующий отделом, ведущий научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Ступников Сергей Александрович (р. 1978) — кандидат технических наук, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Сучков Александр Павлович (р. 1954) — доктор технических наук, ведущий научный сотрудник

Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Тимонина Елена Евгеньевна (р. 1952) — доктор технических наук, профессор, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Хакимова Аида Хатифовна (р. 1963) — кандидат биологических наук, ведущий научный сотрудник Института информационных систем и инженерно-компьютерных технологий Российского нового университета

Хацкевич Владимир Львович (р. 1951) — доктор технических наук, профессор, профессор кафедры математики Военно-воздушной академии им. профессора Н. Е. Жуковского и Ю. А. Гагарина

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

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

АВТОРСКИЙ УКАЗАТЕЛЬ ЗА 2022 г.

	Вып.	Стр.
Абгарян К. К., Гаврилов Е. С. Программный комплекс для многомасштабного моделирования структурных свойств композиционных материалов	1	88–97
Аблаев Ф. М. см. Андрианов С. Н.		
Агаларов Я. М. Оптимальное управление подключением резервного прибора в системе массового обслуживания $G/M/1$	4	34–41
Агаларов Я. М. Оптимизация порогового управления переключением скорости обслуживания в системе массового обслуживания $G/M/1$	1	73–81
Агасандян Г. А. Многомерные бинарные рынки и CC-VaR	2	2–10
Алию Б., Мачнев Е. А., Мокров Е. В. Гистерезисное управление нагрузкой в беспроводных сенсорных сетях	3	83–89
Андрианов С. Н., Андрианова Н. С., Аблаев Ф. М., Кочнева Ю. Ю. Контекстный поиск на фотонах с использованием тестов Белла	1	20–24
Андрианова Н. С. см. Андрианов С. Н.		
Базилевский М. П. Обобщение метода выпрямления искаженных из-за мультикол-линейности коэффициентов для регрессионных моделей с различной степенью корреляции объясняющих переменных	4	20–25
Бесчастный В. А., Острикова Д. Ю., Шоргин С. Я., Молчанов Д. А., Гайдамака Ю. В. Анализ плотности базовых станций 5G NR для предоставления услуг виртуальной и дополненной реальности	2	102–108
Бесчастный В. А. см. Мачнев Е. А.		
Битюков Ю. И. см. Босов А. В.		
Борисов А. В. Общий порядок аппроксимации оценок фильтрации состояний марковских скачкообразных процессов по дискретизованным наблюдениям	4	8–13
Босов А. В. Применение самоорганизующихся нейронных сетей к процессу формирования индивидуальной траектории обучения	3	7–15
Босов А. В. Управление линейным выходом марковской цепи по квадратичному критерию. Случай полной информации	2	19–26
Босов А. В., Битюков Ю. И., Денискина Г. Ю. О поиске оптимальной схемы 3D-печати конструкций из композиционных материалов	1	10–19
Босов А. В., Иванов А. В. Технология классификации типов контента электронного учебника	4	63–72
Брюхов Д. О., Ступников С. А. Логическая реляционная модель структур данных для решения задач в предметной области управления землепользованием	4	93–98
Бурцева С. А. см. Власкина А. С.		
Васильев Н. С. О достаточных условиях экстремума в многомерных вариационных задачах	3	39–44
Власкина А. С., Бурцева С. А., Кочеткова И. А., Шоргин С. Я. Управляемая система массового обслуживания с эластичным трафиком и сигналами для анализа нарезки ресурсов в сети радиодоступа	3	90–96
Гаврилов Е. С. см. Абгарян К. К.		
Гайдамака Ю. В. см. Бесчастный В. А.		
Гайдамака Ю. В. см. Мачнев Е. А.		
Горшенин А. К., Гусейнова Е. И. Повышение доходности торговли на FOREX с помощью LSTM-идентификации свечных паттернов и индикатора тиковых объемов	3	26–38
Григорьев О. Г. см. Смирнов И. В.		

	Вып.	Стр.
Грушо А. А., Грушо Н. А., Забейайло М. И., Зацаринный А. А., Тимонина Е. Е., Шоргин С. Я. Анализ цепочек причинно-следственных связей	2	68–74
Грушо А. А., Грушо Н. А., Забейайло М. И., Смирнов Д. В., Тимонина Е. Е., Шоргин С. Я. О безопасной архитектуре вычислительной системы на основе микросервисов	4	87–92
Грушо А. А., Грушо Н. А., Тимонина Е. Е. Метаданные в защищенном электронном документообороте	3	97–102
Грушо Н. А. см. Грушо А. А.		
Грушо Н. А. см. Грушо А. А.		
Грушо Н. А. см. Грушо А. А.		
Гусейнова Е. И. см. Горшенин А. К.		
Денискина Г. Ю. см. Босов А. В.		
Драгунов Н. А., Дюкова Е. В. О поиске максимальных частых и минимальных нечастых наборов произведения частичных порядков	1	82–87
Дубанов А. А., Нефедова В. А. Кинематические модели задач преследования на плоскости методами параллельного сближения и погони	3	103–109
Дунсяо Гу см. Зацман И. М.		
Дурново А. А., Инькова О. Ю., Попкова Н. А. Принципы описания показателей логико-семантических отношений и их иерархии	2	52–59
Дьяченко Ю. Г. см. Соколов И. А.		
Дюкова А. П. см. Дюкова Е. В.		
Дюкова Е. В., Дюкова А. П. О сложности обучения логических процедур классификации	4	57–62
Дюкова Е. В. см. Драгунов Н. А.		
Забейайло М. И. см. Грушо А. А.		
Забейайло М. И. см. Грушо А. А.		
Зацаринный А. А. см. Грушо А. А.		
Зацман И. М. О научной парадигме информатики: верхний уровень классификации объектов ее предметной области	4	73–79
Зацман И. М. Средовые модели информационных технологий: теоретические основания построения	3	59–67
Зацман И. М., Золотарев О. В., Хакимова А. Х. Средовые модели извлечения из текста новых терминов и индикаторов настроений	2	60–67
Зацман И. М., Золотарев О. В., Хакимова А. Х., Дунсяо Гу. Модель и технология извлечения новых терминов из медицинских текстов	4	80–86
Зейфман А. И. см. Ковалёв И. А.		
Зейфман А. И. см. Сатин Я. А.		
Золотарев О. В. см. Зацман И. М.		
Золотарев О. В. см. Зацман И. М.		
Иванов А. В. см. Босов А. В.		
Инькова О. Ю. см. Дурново А. А.		
Кириков И. А. см. Листопад С. В.		
Кириков И. А. см. Румовская С. Б.		
Киселёв Г. А. см. Смирнов И. В.		
Ковалёв И. А., Сатин Я. А., Синицина А. В., Зейфман А. И. Об одном подходе к оцениванию скорости сходимости нестационарных марковских моделей систем обслуживания	3	75–82
Ковалёв С. П. Алгебраическая спецификация графовых вычислительных структур	1	2–9
Коновалов М. Г., Разумчик Р. В. Синтез управления двумерным случайным блужданием с эталонным стационарным распределением	2	109–117
Кочеткова И. А. см. Власкина А. С.		
Кочнева Ю. Ю. см. Андрианов С. Н.		
Кравцова О. А. Использование критериев стационарности для настройки моделей при прогнозировании временных рядов	2	11–18
Кривенко М. П. Выбор модели при факторизации матрицы данных с пропусками	3	52–58
Крюкова А. Л. см. Сатин Я. А.		

	Вып.	Стр.
Курузов И. А. см. Смирнов И. В.		
Листопад С. В., Кириков И. А. Разрешение конфликтов в гибридных интеллектуальных многоагентных системах	1	54–60
Малашенко Ю. Е. Метрические оценки угловых точек множества достижимых межузловых потоков многопользовательской сети	1	25–31
Малашенко Ю. Е. Последовательный анализ и метрические оценки предельных распределений межузловых потоков в многопользовательской сети	3	45–51
Мачнев Е. А., Бесчастный В. А., Острикова Д. Ю., Гайдамака Ю. В., Шоргин С. Я. Об оптимальном расположении антенн для V2X-соединений в субтерагерцевом диапазоне	4	42–50
Мачнев Е. А. см. Алию Б.		
Мигуля М. А. см. Шнурков П. В.		
Мокров Е. В. см. Алию Б.		
Молчанов Д. А. см. Бесчастный В. А.		
Нефедова В. А. см. Дубанов А. А.		
Нуриев В. А. Переводческий анализ текста с применением информационных ресурсов: редуцирование спектра моделей перевода в надкорпусных базах данных	3	68–74
Острикова Д. Ю. см. Бесчастный В. А.		
Острикова Д. Ю. см. Мачнев Е. А.		
Ошушкова В. С. см. Сатин Я. А.		
Палионная С. И., Шестаков О. В. Использование FDR-метода множественной проверки гипотез при обращении линейных однородных операторов	2	44–51
Панов А. И. см. Смирнов И. В.		
Пешкова И. В. Границы экстремального индекса времени ожидания в системе $M/G/1$ с распределением времени обслуживания в виде конечной смеси	4	26–33
Пешкова И. В. Сравнение экстремальных индексов времен ожидания в системах обслуживания $M/G/1$	1	61–67
Попкова Н. А. см. Дурново А. А.		
Разумчик Р. В. см. Коновалов М. Г.		
Рождественский Ю. В. см. Соколов И. А.		
Румовская С. Б., Кириков И. А. Метод визуализации снижения интенсивности и разрешения конфликтов в гибридных интеллектуальных многоагентных системах	2	94–101
Сатин Я. А., Крюкова А. Л., Ошушкова В. С., Зейфман А. И. О монотонности некоторых классов марковских цепей	2	27–34
Сатин Я. А. см. Ковалёв И. А.		
Синицина А. В. см. Ковалёв И. А.		
Синицын И. Н. Нормализация систем, стохастически не разрешенных относительно производных	1	32–38
Синицын И. Н. Совместная фильтрация и распознавание нормальных процессов в стохастических системах, не разрешенных относительно производных	2	85–93
Смирнов Д. В. см. Грушо А. А.		
Смирнов И. В., Панов А. И., Чуганская А. А., Суворова М. И., Киселёв Г. А., Курузов И. А., Григорьев О. Г. Персональный когнитивный ассистент: планирование поведения на основе сценариев деятельности	1	46–53
Соколов И. А., Степченков Ю. А., Дьяченко Ю. Г., Рождественский Ю. В. Оценка надежности синхронного и самосинхронного конвейеров	4	2–7
Степченков Ю. А. см. Соколов И. А.		
Ступников С. А. см. Брюхов Д. О.		
Суворова М. И. см. Смирнов И. В.		
Сучков А. П. Единая модель государственных данных: сценарии развития	4	99–105
Тимонина Е. Е. см. Грушо А. А.		
Тимонина Е. Е. см. Грушо А. А.		
Тимонина Е. Е. см. Грушо А. А.		

	Вып.	Стр.
Торшин И. Ю. О применении топологического подхода к анализу плохо формализуемых задач для построения алгоритмов виртуального скрининга квантово-механических свойств органических молекул I: Основы проблемно ориентированной теории	1	39–45
Торшин И. Ю. О применении топологического подхода к анализу плохо формализуемых задач для построения алгоритмов виртуального скрининга квантово-механических свойств органических молекул II: Сопоставление формализма с конструктами квантовой механики и экспериментальная апробация предложенных алгоритмов	2	35–43
Хакимова А. Х. см. Зацман И. М.		
Хакимова А. Х. см. Зацман И. М.		
Хацкевич В. Л. Нечеткие усредняющие операторы в задаче агрегирования нечеткой информации	4	51–56
Чуганская А. А. см. Смирнов И. В.		
Шведов А. С. Критерий непустоты эpsilon-ядер для нечетких игр с нетрансферабельной полезностью и вычислительные процедуры	3	2–6
Шестаков О. В. Несмещенная оценка риска пороговой обработки с двумя пороговыми значениями	4	14–19
Шестаков О. В. см. Палионная С. И.		
Шихиев Ф. Ш. см. Шихиев Ш. Б.		
Шихиев Ш. Б., Шихиев Ф. Ш. Упрощенный язык зрительных образов	1	68–72
Шнурков П. В. Об аналитической структуре некоторых видов целевых функционалов, связанных с задачами управления полумарковскими случайными процессами	2	75–84
Шнурков П. В., Мигуля М. А. Некоторые результаты анализа процесса изменения цены бивалютной корзины на основе методов статистики случайных процессов	3	16–25
Шоргин С. Я. см. Бесчастный В. А.		
Шоргин С. Я. см. Власкина А. С.		
Шоргин С. Я. см. Грушо А. А.		
Шоргин С. Я. см. Грушо А. А.		
Шоргин С. Я. см. Мачнев Е. А.		

2022 AUTHOR INDEX

	Issue	Page
Abgaryan K. K. and Gavrilo E. S. Software package for multiscale modeling of structural properties of composite materials	1	88–97
Ablaev F. M. see Andrianov S. N.		
Agalarov Ya. M. Optimal control of a queue-length dependent additional server in GI/M/1 queue	4	34–41
Agalarov Ya. M. Optimization of the threshold service speed control in the G/M/1 queue	1	73–81
Agasandyan G. A. Multidimensional binary markets and CC-VaR	2	2–10
Aliyu B., Machnev E. A., and Mokrov E. V. Hysteretic congestion control in wireless cloud sensor networks	3	83–89
Andrianov S. N., Andrianova N. S., Ablaev F. M., and Kochneva Yu. Yu. Context query on photons with the use of Bell tests	1	20–24
Andrianova N. S. see Andrianov S. N.		
Bazilevskiy M. P. Generalization of a method for straightening coefficients distorted due to multicollinearity in regression models with different degrees of explanatory variables correlation	4	20–25
Beschastnyi V. A., Ostrikova D. Yu., Shorgin S. Ya., Moltchanov D. A., and Gaidamaka Yu. V. Density analysis of mmWave NR deployments for delivering scalable AR/VR video services	2	102–108
Beschastnyi V. A. see Machnev E. A.		
Bitjukov Yu. I. see Bosov A. V.		
Borisov A. V. Total approximation order for Markov jump process filtering given discretized observations	4	8–13
Bosov A. V. Application of self-organizing neural networks to the process of forming an individual learning path	3	7–15
Bosov A. V. Linear output control of Markov chain by square criterion. Complete information case	2	19–26
Bosov A. V., Bitjukov Yu. I., and Deniskina G. Yu. About searching for the optimal 3D printing scheme of structures from composite materials	1	10–19
Bosov A. V. and Ivanov A. V. Technology for classification of content types of e-textbooks	4	63–72
Briukhov D. O. and Stupnikov S. A. Logical relational model of data structures for problem solving in land use management	4	93–98
Burtseva S. A. see Vlaskina A. S.		
Chuganskaya A. A. see Smirnov I. V.		
Deniskina G. Yu. see Bosov A. V.		
Diachenko Yu. G. see Sokolov I. A.		
Djukova A. P. see Djukova E. V.		
Djukova E. V. and Djukova A. P. On the complexity of logical classification learning procedures	4	57–62
Djukova E. V. see Dragunov N. A.		
Dongxiao Gu see Zatsman I. M.		
Dragunov N. A. and Djukova E. V. Finding maximal frequent and minimal infrequent sets in partially ordered data	1	82–87
Dubanov A. A. and Nefedova V. A. Kinematic models of pursuit problems on the plane by the methods of parallel approach and pursuit	3	103–109
Durnovo A. A., Inkova O. Yu., and Popkova N. A. Principles of describing markers of logical-semantic relations and their hierarchy	2	52–59
Gaidamaka Yu. V. see Beschastnyi V. A.		
Gaidamaka Yu. V. see Machnev E. A.		
Gavrilo E. S. see Abgaryan K. K.		

	Issue	Page
Gorshenin A. K. and Guseynova E. I. Increasing FOREX trading profitability with LSTM candlestick pattern recognition and tick volume indicator	3	26–38
Grigoriev O. G. see Smirnov I. V		
Grusho A. A., Grusho N. A., and Timonina E. E. Metadata in secure electronic document management	3	97–102
Grusho A. A., Grusho N. A., Zabezhailo M. I., Smirnov D. V., Timonina E. E., and Shorgin S. Ya. About the secure architecture of a microservice-based computing system	4	87–92
Grusho A. A., Grusho N. A., Zabezhailo M. I., Zatsarinny A. A., Timonina E. E., and Shorgin S. Ya. Cause-and-effect chain analysis	2	68–74
Grusho N. A. see Grusho A. A.		
Grusho N. A. see Grusho A. A.		
Grusho N. A. see Grusho A. A.		
Guseynova E. I. see Gorshenin A. K.		
Inkova O. Yu. see Durnovo A. A.		
Ivanov A. V. see Bosov A. V.		
Khakimova A. K. see Zatsman I. M.		
Khakimova A. K. see Zatsman I. M.		
Khatskevich V. L. Fuzzy averaging operators in the problem of aggregating fuzzy information	4	51–56
Kirikov I. A. see Listopad S. V.		
Kirikov I. A. see Rumovskaya S. B.		
Kiselev G. A. see Smirnov I. V		
Kochetkova I. A. see Vlaskina A. S.		
Kochneva Yu. Yu. see Andrianov S. N.		
Konovalov M. G. and Razumchik R. V. Controlling a bounded two-dimensional Markov chain with a given invariant measure	2	109–117
Kovalev I. A., Satin Y. A., Sinitcina A. V., and Zeifman A. I. On an approach for estimating the rate of convergence for nonstationary Markov models of queueing systems	3	75–82
Kovalyov S. P. Algebraic specification of graph computational structures	1	2–9
Kravtsova O. A. Model setting using stationarity criteria for time series forecasting	2	11–18
Krivenko M. P. Model selection for matrix factorization with missing components	3	52–58
Kryukova A. L. see Satin Y. A.		
Kuruzov I. A. see Smirnov I. V		
Listopad S. V. and Kirikov I. A. Conflict resolution in hybrid intelligent multiagent systems	1	54–60
Machnev E. A., Beschastnyi V. A., Ostrikova D. Yu., Gaidamaka Yu. V., and Shorgin S. Ya. On the optimal antenna deployment for subterahertz V2X communications	4	42–50
Machnev E. A. see Aliyu B.		
Malashenko Yu. E. Metric evaluations of the angular points of the set of attainable internodal flows of multiuser network	1	25–31
Malashenko Yu. E. Sequential analysis and metric estimates of peak load flows in the multiuser network	3	45–51
Migulya M. A. see Shnurkov P. V.		
Mokrov E. V. see Aliyu B.		
Moltchanov D. A. see Beschastnyi V. A.		
Nefedova V. A. see Dubanov A. A.		
Nuriev V. A. Computer-assisted textual analysis in translation: Reducing the spectrum of translation models in supracorpora databases	3	68–74
Oshushkova V. S. see Satin Y. A.		
Ostrikova D. Yu. see Beschastnyi V. A.		
Ostrikova D. Yu. see Machnev E. A.		
Palionnaya S. I. and Shestakov O. V. The use of the FDR method of multiple hypothesis testing when inverting linear homogeneous operators	2	44–51
Panov A. I. see Smirnov I. V		
Peshkova I. V. On bounds of the stationary waiting time extremal index in $M/G/1$ system with mixture service times	4	26–33

	Issue	Page
Peshkova I. V. The comparison of waiting time extremal indexes in $M/G/1$ queueing systems	1	61–67
Popkova N. A. see Durnovo A. A.		
Razumchik R. V. see Konovalov M. G.		
Rogdestvenski Yu. V. see Sokolov I. A.		
Rumovskaya S. B. and Kirikov I. A. Visual representation of the decrease in conflict intensity and its resolution in hybrid intelligent multiagent systems	2	94–101
Satin Y. A., Kryukova A. L., Oshushkova V. S., and Zeifman A. I. On monotonicity of some classes of Markov chains	2	27–34
Satin Y. A. see Kovalev I. A.		
Shestakov O. V. Unbiased thresholding risk estimate with two threshold values	4	14–19
Shestakov O. V. see Palionnaya S. I.		
Shihiev F. Sh. see Shihiev Sh. B.		
Shihiev Sh. B. and Shihiev F. Sh. Simplified language for visual images	1	68–72
Shnurkov P. V. On the analytical structure of some kinds of target functionals associated with the control problems of semi-Markov stochastic processes	2	75–84
Shnurkov P. V. and Migulya M. A. Some results of the analysis of the process of changing the price of a dual currency basket based on random process statistics methods	3	16–25
Shorgin S. Ya. see Beschastnyi V. A.		
Shorgin S. Ya. see Grusho A. A.		
Shorgin S. Ya. see Grusho A. A.		
Shorgin S. Ya. see Machnev E. A.		
Shorgin S. Ya. see Vlaskina A. S.		
Shvedov A. S. A condition for non-emptiness of the epsilon-core of a nontransferable utility fuzzy game and computational schemes	3	2–6
Sinitcina A. V. see Kovalev I. A.		
Sinitsyn I. N. Joint filtration and recognition of normal processes in stochastic systems with unsolved derivatives	2	85–93
Sinitsyn I. N. Normalization of systems with stochastically unsolved derivatives	1	32–38
Smirnov D. V. see Grusho A. A.		
Smirnov I. V., Panov A. I., Chuganskaya A. A., Suvorova M. I., Kiselev G. A., Kuruzov I. A., and Grigoriev O. G. Personal cognitive assistant: Planning activity with scripts	1	46–53
Sokolov I. A., Stepchenkov Yu. A., Diachenko Yu. G., and Rogdestvenski Yu. V. Synchronous and self-timed pipeline's reliability estimation	4	2–7
Stepchenkov Yu. A. see Sokolov I. A.		
Stupnikov S. A. see Briukhov D. O.		
Suchkov A. P. Unified model of national data: Development scenarios	4	99–105
Suvorova M. I. see Smirnov I. V.		
Timonina E. E. see Grusho A. A.		
Timonina E. E. see Grusho A. A.		
Timonina E. E. see Grusho A. A.		
Torshin I. Yu. On the application of a topological approach to analysis of poorly formalized problems for constructing algorithms for virtual screening of quantum-mechanical properties of organic molecules I: The basics of the problem-oriented theory	1	39–45
Torshin I. Yu. On the application of a topological approach to analysis of poorly formalized problems for constructing algorithms for virtual screening of quantum-mechanical properties of organic molecules II: Comparison of formalism with constructions of quantum mechanics and experimental approbation of the proposed algorithms	2	35–43
Vasilyev N. S. On extremum sufficient conditions in multidimensional variation calculus problems	3	39–44
Vlaskina A. S., Burtseva S. A., Kochetkova I. A., and Shorgin S. Ya. Controllable queuing system with elastic traffic and signals for analyzing network slicing	3	90–96
Zabekhailo M. I. see Grusho A. A.		
Zabekhailo M. I. see Grusho A. A.		
Zatsarinny A. A. see Grusho A. A.		

	Issue	Page
Zatsman I. M. Informatics' medium models of information technology: Theoretical foundations for their creating	3	59–67
Zatsman I. M. On the scientific paradigm of informatics: The classification high level of its objects	4	73–79
Zatsman I. M., Zolotarev O. V., and Khakimova A. K. Medium models for discovering novel terms and sentiments from texts	2	60–67
Zatsman I. M., Zolotarev O. V., Khakimova A. K., and Dongxiao Gu. Model and technology for discovering new terms in medical texts	4	80–86
Zeifman A. I. see Kovalev I. A.		
Zeifman A. I. see Satin Y. A.		
Zolotarev O. V. see Zatsman I. M.		
Zolotarev O. V. see Zatsman I. M.		

Правила подготовки рукописей для публикации в журнале «Информатика и её применения»

Журнал «Информатика и её применения» публикует теоретические, обзорные и дискуссионные статьи, посвященные научным исследованиям и разработкам в области информатики и ее приложений.

Журнал издается на русском языке. По специальному решению редколлегии отдельные статьи могут печататься на английском языке.

Тематика журнала охватывает следующие направления:

- теоретические основы информатики;
- математические методы исследования сложных систем и процессов;
- информационные системы и сети;
- информационные технологии;
- архитектура и программное обеспечение вычислительных комплексов и сетей.

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 предоставленной 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 ЛАТЭХ (.tex), дополнительно — в формате .pdf, на дискете, лазерном диске или электронной почтой. Предоставление бумажной рукописи необязательно.

7. При подготовке рукописи в MS Word рекомендуется использовать следующие настройки.

Параметры страницы: формат — А4; ориентация — книжная; поля (см): внутри — 2,5, снаружи — 1,5, сверху — 2, 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/journal/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 почтового адреса организации и электронного адреса каждого автора;
- сведения об авторах, в соответствии с форматом, образцы которого представлены на страницах:
http://www.ipiran.ru/journal/issues/2013_07_01/authors.asp и
http://www.ipiran.ru/journal/issues/2013_07_01_eng/authors.asp;
- аннотация (не менее 100 слов на каждом из языков). Аннотация — это краткое резюме работы, которое может публиковаться отдельно. Она является основным источником информации в информационных системах и базах данных. Английская аннотация должна быть оригинальной, может не быть дословным переводом русского текста и должна быть написана хорошим английским языком. В аннотации не должно быть ссылок на литературу и, по возможности, формул;
- ключевые слова — желательно из принятых в мировой научно-технической литературе тематических тезаурусов. Предложения не могут быть ключевыми словами;
- источники финансирования работы (ссылки на гранты, проекты, поддерживающие организации и т. п.).

9. Требования к спискам литературы.

Ссылки на литературу в тексте статьи нумеруются (в квадратных скобках) и располагаются в каждом из списков литературы в порядке первых упоминаний.

Списки литературы представляются в двух вариантах:

- (1) **Список литературы к русскоязычной части.** Русские и английские работы — на языке и в алфавите оригинала;
- (2) **References.** Русские работы и работы на других языках — в латинской транслитерации с переводом на английский язык; английские работы и работы на других языках — на языке оригинала.

Необходимо для составления списка “References” пользоваться размещенной на сайте <http://www.translit.net/ru/bgn/> бесплатной программой транслитерации русского текста в латиницу.

Список литературы “References” приводится полностью отдельным блоком, повторяя все позиции из списка литературы к русскоязычной части, независимо от того, имеются или нет в нем иностранные источники. Если в списке литературы к русскоязычной части есть ссылки на иностранные публикации, набранные латиницей, они полностью повторяются в списке “References”.

Ниже приведены примеры ссылок на различные виды публикаций в списке “References”.

Описание статьи из журнала:

Zagurenko, A. G., V. A. Korotovskikh, A. A. Kolesnikov, A. V. Timonov, and D. V. Kardymon. 2008. Tekhniko-ekonomicheskaya optimizatsiya dizayna gidrorazryva plasta [Technical and economic optimization of the design of hydraulic fracturing]. *Neftyanoe hozyaystvo [Oil Industry]* 11:54–57.

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 статьи из продолжающегося издания (сборника трудов):

Astakhov, M. V., and T. V. Tagantsev. 2006. Eksperimental'noe issledovanie prochnosti soedineniy “stal’–kompozit” [Experimental study of the strength of joints “steel–composite”]. *Trudy MGTU “Matematicheskoe modelirovanie slozhnykh tekhnicheskikh sistem” [Bauman MSTU “Mathematical Modeling of Complex Technical Systems” Proceedings]*. 593:125–130.

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 povysheniya 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.

Latyshev, V. N. 2009. *Tribologiya rezaniya. Kn. 1: Friksionnye protsessy pri rezanii metallov* [Tribology of cutting. Vol. 1: Frictional processes in metal cutting]. Ivanovo: Ivanovskii State Univ. 108 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: Wiley. 521 p.

Описание неопубликованного документа:

Laturov, A. R., M. M. Khasanov, and V. A. Baikov. 2004 (unpubl.). *Geologiya i dobycha (NGT GiD)* [Geology and production (NGT GiD)]. Certificate on official registration of the computer program No. 2004611198.

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 slovarya informatsionnogo monitoringa* [Technology of development of semantic dictionary of information monitoring system]. Moscow: IPI RAN. PhD Thesis. 23 p.

Описание ГОСТа:

GOST 8.586.5-2005. 2007. *Metodika vypolneniya izmereniy. Izmerenie raskhoda i kolichestva zhidkostey i gazov s pomoshch'yu standartnykh suzhayushchikh ustroystv* [Method of measurement. Measurement of flow rate and volume of liquids and gases by means of orifice devices]. Moscow: Standardinform Publs. 10 p.

Описание патента:

Bolshakov, M. V., A. V. Kulakov, A. N. Lavrenov, and M. V. Palkin. 2006. *Sposob orientirovaniya po krenu letatel'nogo apparata s opticheskoy 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 «Информатика и её применения» является некоммерческим изданием. Плата за публикацию не взимается, гонорар авторам не выплачивается.

Адрес редакции журнала «Информатика и её применения»:
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: iiep@frccsc.ru (Стригина Светлана Николаевна)
<http://www.ipiran.ru/journal/issues/>

Requirements for manuscripts submitted to Journal “Informatics and Applications”

Journal “Informatics and Applications” (Inform. Appl.) publishes theoretical, review, and discussion articles on the research and development in the field of informatics and its applications.

The journal is published in Russian. By a special decision of the editorial board, some articles can be published in English.

The topics covered include the following areas:

- theoretical fundamentals of informatics;
- mathematical methods for studying complex systems and processes;
- information systems and networks;
- information technologies; and
- architecture and software of computational complexes and networks.

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 “Informatics and Applications” 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).

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 the authors receive their article for correction after reviewing, it does not mean that the article is approved for publication. 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/journal/template.doc>.

8. The 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/issues/2013_07_01/authors.asp and

http://www.ipiran.ru/journal/issues/2013_07_01_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 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).
- Article from the continuing publication (collection of works, proceedings):**
Astakhov, M. V., and T. V. Tagantsev. 2006. Eksperimental’noe issledovanie prochnosti soedineniy “stal’–kompozit” [Experimental study of the strength of joints “steel–composite”]. *Trudy MGTU “Matematicheskoe modelirovanie slozhnykh tekhnicheskikh sistem” [Bauman MSTU “Mathematical Modeling of Complex Technical Systems” Proceedings]*. 593:125–130.
- 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 povysheniya neftegazooitdachi” [6th Symposium (International) “New Energy Saving Subsoil Technologies and the Increasing of the Oil and Gas Impact” Proceedings]*. Moscow. 267–272.
- Books and other monographs:**
Lindorf, L. S., and L. G. Mamikonians, eds. 1972. *Ekspluatatsiya turbogeneratorov s neposredstvennym okhlazhdeniem [Operation of turbine generators with direct cooling]*. Moscow: Energy Publs. 352 p.
- Dissertation and Thesis:**
Kozhunova, O. S. 2009. Tekhnologiya razrabotki semanticheskogo slovarya informatsionnogo monitoringa [Technology of development of semantic dictionary of information monitoring system]. Moscow: IPI RAN. PhD Thesis. 23 p.
- State standards and patents:**
GOST 8.586.5–2005. 2007. Metodika vypolneniya izmereniy. Izmerenie raskhoda i kolichestva zhidkostey i gazov s pomoshch’yu standartnykh suzhayushchikh ustroystv [Method of measurement. Measurement of flow rate and volume of liquids and gases by means of orifice devices]. M.: Standardinform Publs. 10 p.
Bolshakov, M. V., A. V. Kulakov, A. N. Lavrenov, and M. V. Palkin. 2006. Sposob orientirovaniya po krenu letatel’nogo apparata s opticheskoy 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. “Informatics and Applications” 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: iiep@frccsc.ru (to Svetlana Strigina)
<http://www.ipiran.ru/english/journal.asp>