

Информатика и её применения

Том 10 Выпуск 4 Год 2016

СОДЕРЖАНИЕ

Анализ характеристик интерференции в модели взаимодействия устройств с учетом среды распространения сигнала <i>Ю. В. Гайдамака, С. Д. Андреев, Э. С. Сопин, К. Е. Самуйлов, С. Я. Шоргин</i>	2
Теорема Пуассона для схемы испытаний Бернулли со случайной вероятностью успеха и дискретный аналог распределения Вейбулла <i>В. Ю. Королев, А. Ю. Корчагин, А. И. Зейфман</i>	11
Несимметричные распределения Линника как предельные законы для случайных сумм независимых случайных величин с конечными дисперсиями <i>В. Ю. Королев, А. И. Зейфман, А. Ю. Корчагин</i>	21
Вычисление асимптотического дефекта некоторых статистических процедур, основанных на выборках случайного объема <i>В. Е. Бенинг</i>	34
Метод кумулятивных сумм для поиска смены режима в процессе Орнштейна–Уленбека на основе процесса Леви <i>А. В. Черток, А. И. Каданер, Г. Т. Хазеева, И. А. Соколов</i>	46
О размещении заданий на двух серверах при неполном наблюдении <i>М. Г. Коновалов, Р. В. Разумчик</i>	57
Байесовские модели массового обслуживания и надежности: вырожденно-вейбулловский случай <i>А. А. Кудрявцев, А. И. Титова</i>	68
Аналитическое решение задачи оптимального управления полумарковским процессом с конечным множеством состояний <i>П. В. Шнурков, А. К. Горшенин, В. В. Белоусов</i>	72
Обобщенный статистический метод анализа текстов, основанный на расчете распределений вероятностей значений статистик <i>А. К. Мельников, А. Ф. Ронжин</i>	89
Об одном способе сокращения вычислений при формировании замыканий Галуа <i>А. А. Грушо, М. И. Забейхайло, А. А. Зацаринный</i>	96
Системотехнические подходы к созданию системы поддержки принятия решений на основе ситуационного анализа <i>А. А. Зацаринный, А. П. Сучков</i>	105
Неформальная аксиоматическая теория ролевых визуальных моделей <i>А. В. Колесников, С. В. Листопад, С. Б. Румовская, В. И. Данишевский</i>	114
Классификация временных рядов в пространстве параметров порождающих моделей <i>М. Е. Карасиков, В. В. Стрижов</i>	121
База данных безличных глагольных конструкций русского языка <i>Анна А. Зализняк, М. Г. Кружков</i>	132
Об авторах	142
Авторский указатель за 2016 г.	144
2016 Author Index	148
Правила подготовки рукописей	152
Requirements for manuscripts	155

АНАЛИЗ ХАРАКТЕРИСТИК ИНТЕРФЕРЕНЦИИ В МОДЕЛИ ВЗАИМОДЕЙСТВИЯ УСТРОЙСТВ С УЧЕТОМ СРЕДЫ РАСПРОСТРАНЕНИЯ СИГНАЛА*

Ю. В. Гайдамака¹, С. Д. Андреев², Э. С. Сопин³, К. Е. Самуйлов⁴, С. Я. Шоргин⁵

Аннотация: Современные беспроводные сети четвертого (4G) и пятого (5G) поколения допускают для обеспечения связи внутри зданий размещение отдельной точки беспроводного доступа в каждом помещении здания. В статье проведен анализ отношения сигнал/интерференция для беспроводных систем, работающих в смежных помещениях, разделенных стенами из различных материалов, с учетом потери мощности при прохождении сигнала сквозь различные среды распространения. Предложен метод оценки плотности распределения случайной величины отношения сигнал/интерференция путем аппроксимации мощности интерферирующего сигнала несколькими известными распределениями. Численный анализ показал, что эффективность таких систем характеризуется как параметром распространения сигнала в пространстве, так и параметром затухания сигнала, зависящим от материала, из которого изготовлены стены. Получены результаты для различных материалов и толщины стен.

Ключевые слова: беспроводная сеть; отношение сигнал/интерференция; прямое взаимодействие устройств; распространение сигнала; среда распространения

DOI: 10.14357/19922264160401

1 Введение

Прогнозируемое увеличение объема трафика в беспроводных сетях создает высокие требования к пятому поколению мобильных систем [1, 2]. Согласно ежегодному отчету компании Cisco «Наглядный индекс развития сетевых технологий» [1], ежегодный прирост трафика, создаваемый мобильными беспроводными устройствами, составляет 53% и по сравнению с 6,2 экзбайт в месяц ($6,2 \cdot 10^{18}$) в 2016 г. достигнет 30,6 экзбайт в месяц к 2020 г.

Одной из причин является рост числа пользовательских мобильных устройств — смартфонов, коммуникаторов, планшетов, носимых (wearable) устройств, таких как умные очки, браслеты, датчики слежения за физическим состоянием. Кроме того, прирост трафика связан с активным продвижением концепции Интернета вещей [3], которая предусматривает автоматическое или с минимальным вовлечением человека производство и передачу данных межмашинного взаимодействия

(machine-to-machine, M2M), собранных системами телеметрии и телеконтроля в жилищно-коммунальном хозяйстве (умный дом, умный город), сельском хозяйстве (сенсоры контроля технологических параметров в агрономии, датчики отслеживания и мониторинга в животноводстве), в интеллектуальных системах безопасности. По данным текущего отчета компании Ericsson «Ericsson Mobility Report» [2], в котором исследуются различные аспекты мобильной индустрии, на июнь 2016 г. на 5 млрд чел., проживающих в зоне действия беспроводных сетей, приходится 7,4 млрд мобильных подключений [2], а к 2020 г. эти цифры вырастут до 7,8 млрд и 11,6 млрд соответственно [1].

Таким образом, уже сегодня во многих странах число абонентов мобильной связи превышает численность населения, на душу населения приходится в среднем 1,5 мобильных устройства.

Беспроводные сети последующих поколений должны не только покрывать огромные территории с многомиллиардным населением, но и быть

* Исследование выполнено при финансовой поддержке Российского научного фонда (проект 16-11-10227).

¹ Российский университет дружбы народов; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, ygaidamaka@sci.pfu.edu.ru

² Российский университет дружбы народов; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, serge.andreev@gmail.com

³ Российский университет дружбы народов; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, eduard.sopin@gmail.com

⁴ Российский университет дружбы народов; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, ksam@sci.pfu.edu.ru

⁵ Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, sshorgin@ipiran.ru

способны решить задачи социально-технологической эволюции следующего десятилетия, обеспечить разнообразие и масштабируемость услуг для пользователей. Сегодня действуют сети четвертого поколения 4G (International Mobile Telecommunications Advanced, IMT-Advanced), работающие на технологиях LTE Advanced (LTE-A) и WiMAX 2 (WMAN-Advanced), в России сотовые операторы осуществляют пользовательский доступ в Интернет на скорости передачи данных до 300 Мбит/с. Коммерческий запуск сетей стандарта 5G прогнозируется в 2020 г. В 2016 г. тестовые пилотные испытания мобильной сети пятого поколения на оборудовании компании Nokia показали пиковые скорости свыше 30 Гбит/с и более 1 млн одновременных подключений в соте [4], а в России в сети компании «Мегафон» на этом оборудовании удалось добиться скорости на уровне 5 Гбит/с [5]. Стандартизацией сетей пятого поколения 5G занимается созданная в 2015 г. Оперативная группа по сетевым аспектам IMT-2020, которая опубликовала так называемую «дорожную карту» развития мобильной сети 5G, где обозначена задача разработки точных и подробных требований к новой сети для достижения сверхвысокой скорости передачи, ультранизкой задержки, высокой энергоэффективности и экологической безопасности.

Эти требования охватывают технологии радиоинтерфейсов, входящие в семейство IMT-2020, которые стали кандидатами для стандарта сетей пятого поколения, в том числе новые методы модуляции, кодирования, применение адаптивных антенных решеток с MIMO (Multiple Input Multiple Output) с узкой направленностью сигнала, а также ряд сетевых решений, обеспечивающих значительное увеличение производительности, включая использование малых (микро/пико/фемто) сот [6], ретрансляции через клиента (client-relays) [7], прямые соединения между конечными устройствами (device-to-device, D2D) [8]. Эти технологии основываются на повторном пространственном использовании частот, механизмы которого учитывают возникающую при этом повышенную интерференцию. Особенно заметно интерференция влияет на производительность беспроводной сети в гетерогенных сотовых сетях, где одновременное применение нескольких технологий беспроводного доступа сочетается с высокой плотностью пользователей. Универсальной метрикой определения производительности беспроводных систем является отношение сигнал/интерференция (SIR, Signal-to-Interference Ratio) в канале от передатчика к приемнику радиосигнала [9], при этом интерференция зависит от мощностей передатчиков, расстояний между интерферирующими устройствами

и от среды распространения сигнала. Отношение сигнал/интерференция влияет на показатели качества передачи в радиоканале, например коэффициент ошибок на бит (bit error rate, BER), максимальную пропускную способность и спектральную эффективность канала связи, которые для услуг сетей последующих поколений определены международными стандартами.

Таким образом, задача нахождения отношения сигнал/интерференция для типичных сетевых конфигураций имеет особое значение, поскольку ее решение позволяет оценить применимость сценариев взаимодействия устройств в современных и перспективных беспроводных сетях. При анализе интерференции для моделирования расположения мобильных устройств традиционно применяются пространственные точечные процессы [10], которые позволяют для устройства-приемника в одной из точек процесса оценить искомую метрику — отношение сигнал/интерференция — с учетом интерференции от остальных устройств-передатчиков — соседних точек процесса.

Так, в [11] был исследован пространственный пуассоновский точечный процесс (Poisson Point Process, PPP), для которого известна формула расстояния до n -ближайшего соседа, а также процесс твердого ядра Матерна (Matern Hardcore Process, MHC), который позволяет учесть ограничения на расстояние внутри пары приемник–передатчик. Для этих процессов показана сложность получения распределения величины отношения сигнал/интерференция в замкнутой аналитической форме, поэтому актуальной является задача подбора аппроксимации для нахождения отношения сигнал/интерференция.

В статье проблема нахождения отношения сигнал/интерференция исследуется для случая прямого взаимодействия устройств (D2D-взаимодействия) внутри помещений, например в офисных зданиях или торгово-развлекательных центрах, где для обеспечения покрытия на каждом этаже размещается несколько точек беспроводного доступа (Access Point, AP), имеющих относительно небольшую зону покрытия. При этом предполагается, что соседние точки располагаются в смежных помещениях. Особенностью этой задачи является фиксированное расположение передатчиков, что не позволяет при анализе интерференции воспользоваться в полной мере аппаратом стохастической геометрии — дисциплины, изучающей взаимоотношения между геометрией и теорией вероятностей, которая развилась из классической интегральной геометрии и задач о геометрической вероятности с привнесением идей и методов теории случайных процессов, в особенности теории точечных

процессов. Для исследованного случая аналитические формулы, учитывающие интерференцию от нескольких передатчиков из смежных помещений, получены в предположении об идентичности основных характеристик интерференции для всех источников и являются развитием результатов для случая одного interfering источника, известных из [12].

Также получен метод оценки плотности распределения случайной величины (с.в.) отношения сигнал/интерференция с помощью аппроксимации мощности interfering сигнала несколькими классическими распределениями, который учитывает потерю мощности при прохождении сигнала сквозь различные среды распространения (пространство внутри помещений, стены, межэтажные перекрытия). Проведено сравнение аппроксимации для случая нормального и гиперэкспоненциального распределений с разработанной имитационной моделью. Приведена оценка отношения сигнал/интерференция для различных материалов стен.

2 Сценарий взаимодействия беспроводных устройств

Исследуется сценарий прямого взаимодействия устройств внутри смежных помещений, имеющих прямоугольную или квадратную форму (рис. 1).

Построение модели проведем в обозначениях статьи [12]. В каждом помещении имеется собственная точка доступа $Rx_i, i = \overline{0, 4}$, которая размещена в геометрическом центре помещения. Техника повторного пространственного использования частот предполагает назначение одного и того же радиоканала для пар приемник–передатчик, находя-

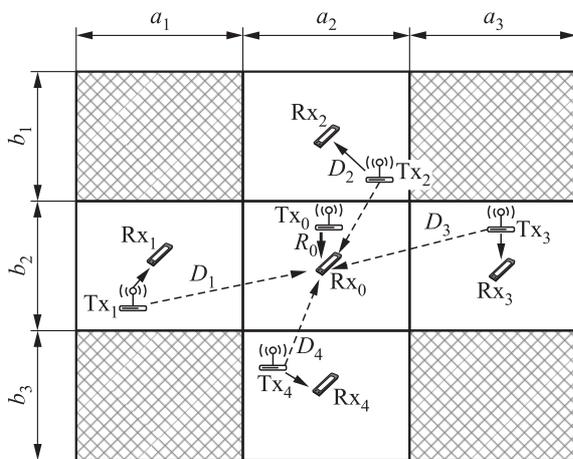


Рис. 1 Схема взаимодействия устройств в смежных помещениях

щихся в смежных помещениях. Исследуется интерференция на приемнике (точке доступа) Rx_0 для восходящего радиоканала (uplink) в центральном помещении. Предполагается, что в каждом из четырех смежных помещений имеется один interfering передатчик (мобильное устройство пользователя) $Tx_i, i = \overline{1, 4}$. При этом каждая координата передатчиков имеет равномерное распределение по длине стороны прямоугольника. Расстояние между передатчиком Tx_0 и приемником Rx_0 в центральном помещении обозначим R_0 . Мощность interfering сигнала от передатчика Tx_i зависит от расстояния между передатчиком Tx_i и приемником Rx_0 , которое обозначим $D_i, i = \overline{1, 4}$. Затухание сигнала при прохождении сквозь различные среды учитывается в модели с помощью поправочных коэффициентов B_i . При построении модели использован подход [13] — при оценке интерференции учитывается только вклад слагаемых, соответствующих устройствам, расположенным в помещениях, смежных с центральным, и пренебрегается interfering сигналом от остальных устройств, затухающим при прохождении через два и более экрана (стены, перегородки).

3 Модель системы и метод анализа показателей качества

Для сценария с четырьмя interfering устройствами, расположенными в помещениях, смежных с центральным, как показано на рис. 1, величина SIR отношения сигнал/интерференция может быть рассчитана по формуле:

$$SIR = \frac{S(R_0)}{\sum_{i=1}^4 B_i I_i(D_i)}, \quad (1)$$

где $S(R_0)$ — мощность «полезного» сигнала от передатчика в центральном помещении, $\sum_{i=1}^4 B_i I_i(D_i)$ — мощность суммарного interfering сигнала от других передатчиков. При этом величины $S(R_0)$ и $I_i(D_i)$ определяются сходным образом:

$$S(R_0) = g_0 R_0^{-\gamma_0}; \quad (2)$$

$$\sum_{i=1}^4 B_i I_i(D_i) = \sum_{i=1}^4 B_i g_i D_i^{-\gamma_i}. \quad (3)$$

Здесь g_i — базовые мощности сигнала передатчика в центральном помещении ($i = 0$) и сигналов interfering передатчиков в смежных помещениях ($i = \overline{1, 4}$); γ_i — коэффициент затухания

(path loss exponent) сигнала в центральном ($i = 0$) и смежных ($i = \overline{1, 4}$) помещениях; B_i — коэффициент потерь сигнала при прохождении сквозь стены, $i = \overline{1, 4}$.

Дальнейший анализ проводим для квадратных помещений со стороной $c = a_j = b_j$, $j = \overline{1, 3}$, в предположении о равенстве мощностей сигналов передатчиков $g_i = g$, одинаковых коэффициентах затухания $\gamma_i = \gamma$ во всех помещениях, $i = \overline{0, 4}$, а также одинаковых коэффициентах потерь сигнала при прохождении сквозь стены $B_i = B$, $i = \overline{1, 4}$. Введем случайные величины $\xi = SIR$, $\eta_1 = R^{-\gamma}$ и $\eta_2 = D^{-\gamma}$. Из [14] известна функция плотности $W_{\eta_2}(y_2)$ с.в. $\eta_2 = D^{-\gamma}$:

$$W_{\eta_2}(y_2) = \left(\frac{2}{\gamma c^2} y_2^{-2/\gamma-1} \right) \times \begin{cases} 0, & y_2 \geq \left(\frac{c}{2} \right)^\gamma; \\ \arcsin \left[\frac{c}{2y_2^{-1/\gamma}} \right] - \\ \quad - \arcsin \left[\frac{\sqrt{-9c^2 + 4y_2^{-2/\gamma}}}{2y_2^{-1/\gamma}} \right], & \left(c\sqrt{\frac{5}{2}} \right)^{-\gamma} < y_2 \leq \left(\frac{3c}{2} \right)^{-\gamma}; \\ \arcsin \left[\frac{\sqrt{-c^2 + 4y_2^{-2/\gamma}}}{2y_2^{-1/\gamma}} \right], & \left(\frac{c}{\sqrt{2}} \right)^{-\gamma} < y_2 \leq \left(\frac{c}{2} \right)^{-\gamma}; \\ \arcsin \left[\frac{c}{2y_2^{-1/\gamma}} \right], & \left(\frac{3c}{2} \right)^{-\gamma} < y_2 \leq \left(\frac{c}{\sqrt{2}} \right)^{-\gamma}. \end{cases} \quad (4)$$

Формула (4) служит для оценки характеристик интерференции одного из слагаемых в знаменателе выражения (1) для отношения сигнал/интерференция. Для оценки основного исследуемого показателя, а именно отношения сигнал/интерференция, необходимо для заданных распределений с.в. расстояния R_0 от приемника Rx_0 до передатчика Tx_0 и с.в. расстояний D_i от приемника Rx_0 до передатчика Tx_i , $i = \overline{1, 4}$, найти распределение с.в. $S(R_0)$ по формуле (2) и $I_i(D_i)$ по формуле (3), а затем получить совместное распределение с.в. SIR по формуле (1).

4 Аппроксимация нормальным распределением

Особенность рассматриваемой задачи заключается в том, что плотность суммарного интерферирующего сигнала (3) представляет собою свертку выражений вида (4). Исследование ее в аналитическом виде нецелесообразно, поэтому далее в статье решается задача нахождения аппроксимирующего распределения для суммарного интерферирующего сигнала. Сделаем упрощающее предположение о том, что основные показатели интерференции идентичны для всех слагаемых в (3). Тогда плотность распределения суммарного интерферирующего сигнала можно аппроксимировать нормальным законом $N(\tilde{\mu}, \tilde{\sigma})$, параметры которого в предположении о том, что материал и толщина стен одинаковы, имеют вид:

$$\tilde{\mu} = 4B\mu(D); \quad \tilde{\sigma}^2 = 4B\sigma^2(D),$$

где $\mu(D)$ и $\sigma^2(D)$ определяются из (4).

В сделанных предположениях плотность распределения $W_{\eta_1}(y_1)$ мощности полезного сигнала определяется формулой:

$$W_{\eta_1}(y_1) = \left(\frac{2}{\gamma c^2} y_1^{-2/\gamma-1} \right) \times \begin{cases} \pi, & \left(\frac{c}{2} \right)^{-\gamma} < y_1 < \infty; \\ 2 \left(\arcsin \left[\frac{c}{2y_1^{-1/\gamma}} \right] - \right. \\ \quad \left. - \arcsin \left[\frac{\sqrt{-c^2 + 4y_1^{-2/\gamma}}}{2y_1^{-1/\gamma}} \right] \right), & \left(\frac{c}{\sqrt{2}} \right)^{-\gamma} < y_1 \leq \left(\frac{c}{2} \right)^{-\gamma}. \end{cases} \quad (5)$$

Воспользовавшись известным из [15] преобразованием с.в., получим плотность распределения с.в. $\xi = SIR$ как совместную плотность мощности полезного сигнала и суммы мощностей интерферирующих сигналов. Учитывая вид (1) с.в. ξ , введем вспомогательные переменные $y_3 = f(x_3, x_4) = x_3/x_4$ и $y_4 = x_4$, а также обратное преобразование $x_3 = \varphi(y_3, y_4) = y_3 y_4$, заметив, что $\partial \varphi(y_3, y_4) / \partial y_3 = y_4$. Искомую функцию получим, используя соотношение:

$$W_{\eta_1, \eta_2}(y_3, y_4) = w_{\chi_1, \chi_2}(\varphi(y_3, y_4), y_4) \left| \frac{\partial \varphi(y_3, y_4)}{\partial y_3} \right|, \quad (6)$$

где $w_{\chi_1, \chi_2}(y_3, y_4)$ — совместная плотность с.в. $R_0^{-\gamma}$ и $\sum_{i=1}^4 D_i^{-\gamma_i}$. После интегрирования (6) по y_4 получаем плотность с.в. ξ в виде:

$$W_{\xi}(y_3) = \int_{\mathbf{Y}} w_{\chi_1, \chi_2}(\varphi(y_3, y_4), y_4) \left| \frac{\partial \varphi(y_3, y_4)}{\partial y_3} \right| dy_4,$$

где \mathbf{Y} — область значений переменной y_4 .

Вследствие ограничений выражений (4) и (5) множество \mathbf{Y} имеет вид:

$$\mathbf{Y} = \left\{ \left(\frac{c}{\sqrt{2}} \right)^{-\gamma} < y_1 \leq \left(\frac{c}{2} \right)^{-\gamma}, -\infty < y_2 < \infty \right\} \cup \left\{ \left(\frac{c}{2} \right)^{-\gamma} < y_1 < \infty, -\infty < y_2 < \infty \right\}.$$

Таким образом,

$$\mathbf{Y} = \mathbf{Y}^1 \cup \mathbf{Y}^2;$$

$$\mathbf{Y}_1 =$$

$$= \left\{ y_3 < 0, \frac{e^{\gamma \ln 2 - \gamma \ln c}}{y_3} < y_4 < \frac{e^{(1/2)\gamma \ln 2 - \gamma \ln c}}{y_3} \right\} \cup \left\{ y_3 > 0, \frac{e^{(1/2)\gamma \ln 2 - \gamma \ln c}}{y_3} < y_4 < \frac{e^{\gamma \ln 2 - \gamma \ln c}}{y_3} \right\};$$

$$\mathbf{Y}_2 = \left\{ y_3 < 0, y_4 < \frac{e^{\gamma \ln 2 - \gamma \ln c}}{y_3} \right\} \cup \left\{ y_3 > 0, y_4 > \frac{e^{\gamma \ln 2 - \gamma \ln c}}{y_3} \right\}.$$

Утверждение. Плотность $W_{\xi}(y_3)$ с.в. ξ отношения сигнал/интерференция имеет вид:

$$W_{\xi}(y_3) = \begin{cases} \int_{M_1} I_1(y_3, y_4) dy_4 + \int_{M_3} I_1(y_3, y_4) dy_4, & y_3 < 0; \\ \int_{M_2} I_2(y_3, y_4) dy_4 + \int_{M_4} I_2(y_3, y_4) dy_4, & y_3 \geq 0 \end{cases}$$

с областями интегрирования

$$M_1 = \left\{ (y_3, y_4) : \frac{e^{\gamma \ln 2 - \gamma \ln c}}{y_3} < y_4 < \frac{e^{(1/2)\gamma \ln 2 - \gamma \ln c}}{y_3} \right\};$$

$$M_2 =$$

$$= \left\{ (y_3, y_4) : \frac{e^{(1/2)\gamma \ln 2 - \gamma \ln c}}{y_3} < y_4 < \frac{e^{\gamma \ln 2 - \gamma \ln c}}{y_3} \right\};$$

$$M_3 = \left\{ (y_3, y_4) : y_4 < \frac{e^{\gamma \ln 2 - \gamma \ln c}}{y_3} \right\};$$

$$M_4 = \left\{ (y_3, y_4) : y_4 > \frac{e^{\gamma \ln 2 - \gamma \ln c}}{y_3} \right\}$$

и подынтегральными выражениями

$$I_1(y_3, y_4) = \left(\arcsin \left[\frac{c}{2(y_3 y_4)^{-1/\gamma}} \right] - \arcsin \left[\frac{\sqrt{-c^2 + 4(y_3 y_4)^{-2/\gamma}}}{2(y_3 y_4)^{-1/\gamma}} \right] \right) / \left(\frac{(4y_4/(\gamma c^2))(y_3 y_4)^{-2/\gamma-1}}{\tilde{\sigma} \sqrt{2\pi}} e^{(y_4 - \bar{\mu})^2 / (2\tilde{\sigma}^2)} \right);$$

$$I_2(y_3, y_4) = \frac{2\sqrt{2\pi^3} y_4 \tilde{\sigma} (y_3 y_4)^{-2/\gamma-1}}{\gamma c^2 e^{-(y_4 - \bar{\mu})^2 / (2\tilde{\sigma}^2)}}.$$

Точность предложенной аппроксимации проверена с помощью имитационного моделирования методом Монте-Карло случайных переменных, а именно координат расположения устройств в помещениях, с последующей оценкой характеристик исследуемого показателя — отношения сигнал/интерференция. Разработанная в [13] имитационная программа позволяет варьировать входные параметры модели, представляющие интерес для данного исследования, в том числе размеры помещений, коэффициенты потерь и коэффициенты затухания, число экспериментов и частоту дискретизации при построении гистограммы плотностей распределения основных случайных величин. Выходными данными имитационной программы являются функции эмпирической плотности, значения среднего и среднеквадратического отклонения, а также оценки квантилей основного исследуемого показателя — отношения сигнал/интерференция.

5 Численные примеры

Результаты численного анализа, которые позволяют оценить отношение сигнал/интерференция на приемнике в центральном помещении, представлены на рис. 2 и 3. Моделирование проводилось для квадратных помещений с длиной стороны $c = 10$ для двух значений коэффициента затухания $\gamma = 2$ и 3. Число экспериментов по моделированию положения мобильных устройств в помеще-

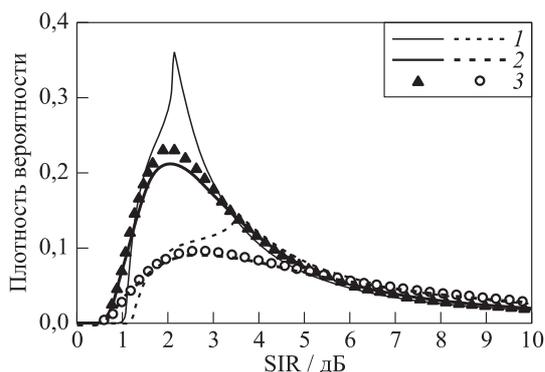


Рис. 2 Отношение сигнал/интерференция: аппроксимация (1 — нормальное распределение; 2 — гиперэкспоненциальное распределение) и имитационное моделирование (3). Сплошные кривые и черные значки — $\gamma = 2$; пунктирные и штриховые кривые и пустые значки — $\gamma = 3$

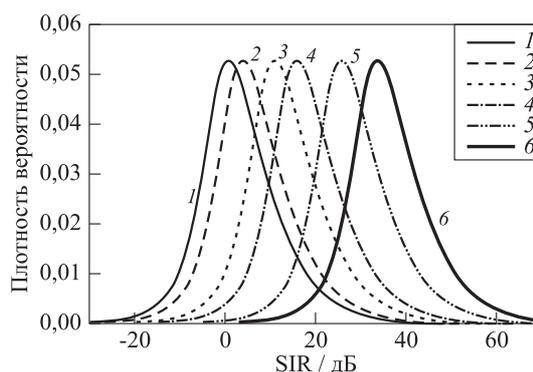


Рис. 3 Отношение сигнал/интерференция: среда распространения сигнала: 1 — стекло 13 мм, $B = 2$ дБ; 2 — кирпич 178 мм, $B = 5$ дБ; 3 — бетон 102 мм, $B = 12$ дБ; 4 — каменная кладка 406 мм, $B = 17$ дБ; 5 — железобетон 203 мм, $B = 27$ дБ; 6 — бетон 305 мм, $B = 35$ дБ

ниях для построения эмпирической функции плотности с.в. SIR отношения сигнал/интерференция составило 10^6 .

Форма графика для аппроксимации с помощью нормального закона на рис. 2 повторяет форму кривой имитационного моделирования. Однако существенный сдвиг не позволяет рекомендовать эту аппроксимацию для оценки отношения сигнал/интерференция, особенно при малых значениях этого показателя, которые важны для приложений. При $SIR < 1$ мощность суммарного интерферирующего сигнала превосходит мощность полезного сигнала, что дает скорость передачи данных в канале, неприемлемую для большинства услуг, предоставляемых в сетях последующих поколений. Для приложений важно оценить квантили нижних уровней с.в. отношения сигнал/интерференция, показывающие долю мобильных пользователей, которые не смогут получить запрошенную услугу из-за низкого качества передачи данных в канале.

В ходе численного эксперимента для оценки функции плотности отношения сигнал/интерференция была предложена аппроксимация мощности интерферирующего сигнала от одного источника с помощью усеченного трехпараметрического гиперэкспоненциального распределения, параметры которого подобраны так, что относительная погрешность аппроксимации не превышает 4%. Плотность распределения $W_\chi(x)$ отношения сигнал/интерференция χ от каждого источника имеет вид:

$$W_\chi(x) = \begin{cases} 0, & x < 0,5; \\ q\alpha_1 e^{\alpha_1 x} + (1 - q)\alpha_2 e^{\alpha_2 x}, & x \geq 0,5, \end{cases}$$

где для $\gamma = 2$ $q = 0,517$, $\alpha_1 = 8571,284$ и $\alpha_2 = 1516,775$, а для $\gamma = 3$ $q = 0,534$, $\alpha_1 = 1295,618$ и $\alpha_2 = 280,362$.

Эта аппроксимация, график которой также приведен на рис. 2 и с высокой степенью точности повторяет кривую имитационного моделирования, может быть рекомендована для оценки отношения сигнал/интерференция снизу.

Рисунок 3 иллюстрирует зависимость характеристик получаемого приемником интерферирующего сигнала мобильных устройств, работающих в смежных помещениях, от материала, из которого изготовлены межкомнатные перегородки.

Показаны полученные с помощью имитационного моделирования графики плотности распределения с.в. SIR отношения сигнал/интерференция с коэффициентом затухания $\gamma = 3$ для нескольких значений коэффициента потерь B при прохождении сигнала сквозь различные среды, взятых из стандарта Международного союза электросвязи [16]. На практике показатель SIR часто измеряют в децибелах, поэтому на рис. 3 по оси абсцисс отложено отношение сигнал/интерференция $S_{дБ}$ в децибелах, которое получено из величины SIR дополнительным преобразованием $S_{дБ} = 10 \lg(SIR)$. Здесь величина 0 дБ означает, что мощности полезного и интерферирующего сигналов равны. Заметим, что, например, для беспроводных сетей четвертого поколения LTE при $S_{дБ} \geq 13$ дБ качество канала считается хорошим, а при $S_{дБ} \geq 20$ дБ — отличным. Современные методы модуляции и кодирования в сетях последующих поколений, в частности мультиплексирование с разделением по ортогональным частотам (Orthogonal Frequency Division Multiplexing, OFDM), позволяют передавать

данные даже при отрицательном значении S_{dB} , т. е. в случае, когда в получаемом приемником сигнале помех больше, чем полезной части. Основываясь на результатах проведенного эксперимента, можно сделать вывод, что экранирование заметно влияет на значение отношения сигнал/интерференция, при этом существенную роль играют как материал, из которого изготовлен экран, так и толщина этого экрана. Если стеклянные перегородки между помещениями практически не гасят интерферирующий сигнал мобильных устройств, работающих в смежных помещениях на той же или близких радиочастотах, то наличие перегородки в виде бетонной стены толщиной 30 см обеспечивает отличное качество передачи данных ($S_{dB} \geq 20$ дБ) для подавляющего большинства (свыше 97%) мобильных устройств.

6 Заключение

В статье представлен аналитический метод оценки плотности распределения с.в. отношения сигнал/интерференция при взаимодействии устройств внутри смежных помещений, имеющих прямоугольную форму. Исследована интерференция в восходящем радиоканале от мобильного устройства к точке доступа, при этом учитывается потеря мощности при прохождении сигнала сквозь различные среды (пространство внутри помещений, стены, межэтажные перекрытия). Заметим, что интерференция в нисходящем радиоканале для описанной задачи взаимодействия мобильного устройства с точкой доступа исследована в [13]. При этом, как и в исследованном в настоящей статье случае, распределение с.в. отношения сигнал/интерференция имеет достаточно громоздкое представление.

Предложен метод оценки отношения сигнал/интерференция с помощью аппроксимирующего распределения для интерферирующего сигнала от одного источника. Сравнение аппроксимации для случая нормального и гиперэкспоненциального распределений с результатами имитационного моделирования выявило, что более точную аппроксимацию дает полученное численно гиперэкспоненциальное распределение, которое можно использовать для оценки снизу отношения сигнал/интерференция.

Приведена оценка отношения сигнал/интерференция для различных материалов стен, которая позволяет сделать вывод о существенной зависимости производительности беспроводных соединений в рассмотренном сценарии взаимодействия беспроводных устройств от потерь мощности за счет

распространения сигнала в пространстве и материала, из которого изготовлены стены между помещениями.

Авторы выражают благодарность магистру кафедры прикладной информатики и теории вероятностей РУДН Р. Ковальчукову за вычисления и оформление графиков в статье.

Литература

1. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2015–2020 White Paper. — Cisco, February 01, 2016. <http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html>.
2. Ericsson mobility report. — Ericsson, 2016. <https://www.ericsson.com/res/docs/2016/ericsson-mobility-report-2016.pdf>.
3. Overview of Internet of Thing: Recommendation Y.2060 (06/2012). — ITU-T, 2012-06-15. <https://www.itu.int/rec/T-REC-Y.2060-201206-I>.
4. Nokia demonstrates world's first 5G-ready network. — Nokia, June 28, 2016. <http://company.nokia.com/en/news/press-releases/2016/06/28/nokia-demonstrates-worlds-first-5g-ready-network>.
5. МегаФон показала мобильный 5G-интернет на скорости 5 Гбит/с. *ferra.ru*, 21.09.2016. <http://www.ferra.ru/ru/mobile/news/2016/09/21/megafon-nn-5g>.
6. *Andrews J., Claussen H., Dohler M., Rangan S.* Femtocells: Past, present, and future // *IEEE J. Sel. Areas Commun.*, 2012. Vol. 30. No. 3. P. 497–508.
7. *Lee J., Wang H., Andrews J., Hong D.* Outage probability of cognitive relay networks with interference constraints // *IEEE Trans. Wir. Comm.*, 2011. Vol. 10. No. 2. P. 390–395.
8. *Fodor G., Parkvall S., Sorrentino S., Wallentin P., Lu Q., Brahma N.* Device-to-device communications for national security and public safety // *IEEE Access*, 2014. Vol. 2. P. 1510–1520.
9. *Sathya V., Ramamurthy A., Kumar S., Tamma B.* On improving SINR in LTE HetNets with D2D relays // *Comput. Commun.*, 2016. Vol. 83. P. 27–44.
10. *Baccelli F., Błaszczyszyn B.* Stochastic geometry and wireless networks. // *Foundations and Trends® in Networking*, 2010. Vol. 3. No. 3-4. P. 249–449; Vol. 4. No. 1-2. P. 1–312. doi: 10.1561/1300000006.
11. *Abaev P., Samouylov K., Gaidamaka Yu., Shorgin S.* Tractable distance distribution approximations for hardcore processes // *Distributed Computer and Communication Networks: Control, Computation, Communications Conference Proceedings*. — *Communications in Computer and Information Science*. Vol. 678. — Springer, 2016 (in press).
12. *Гайдамака Ю. В., Самуйлов А. К.* Метод расчета характеристик интерференции двух взаимодействующих

- устройств в беспроводной гетерогенной сети // Информатика и её применения, 2015. Т. 9. Вып. 1. С. 10–15.
13. Samuylov A., Moltchanov D., Gaidamaka Yu., Begishev V., Kovalchukov R., Abaev P., Shorgin S. SIR analysis in square-shaped indoor premises // 30th European Conference on Modelling and Simulation Proceedings. — Regensburg, Germany: Digitaldruck Pirrot GmbH. P. 692–697.
 14. Samuylov A., Ometov A., Begishev V., Kovalchukov R., Moltchanov D., Gaidamaka Yu., Samouylov K., Andreev S., Koucheryavy Y. Analytical performance estimation of network-assisted D2D communications in urban scenarios with rectangular cells // Trans. Emerging Telecommunications Technol., 12 November, 2015. P. 1–14. doi: 10.1002/ett.2999A.
 15. Левин Б. П. Теоретические основы статистической радиотехники. — 3-е изд. — М.: Радио и связь, 1989. 656 с.
 16. Propagation data and prediction methods for the planning of indoor radiocommunication systems and radio local area networks in the frequency range 300 MHz to 100 GHz: Recommendation P.1238-8 (07/2015). — ITU-R, 2015-07-29. <http://www.itu.int/rec/R-REC-P.1238-8-201507-1>.

Поступила в редакцию 14.10.16

INTERFERENCE ANALYSIS OF THE DEVICE-TO-DEVICE COMMUNICATIONS MODEL WITH REGARD TO A SIGNAL PROPAGATION ENVIRONMENT

Yu. V. Gaidamaka^{1,2}, S. D. Andreev^{1,2}, E. S. Sopin^{1,2}, K. E. Samouylov^{1,2}, and S. Ya. Shorgin²

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

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

Abstract: Modern dense 4G and 5G networks allow placing a wireless access point in every indoor location to optimize wireless coverage. The paper analyzes the signal-to-interference ratio (SIR) for wireless systems operating in adjacent premises, and takes into account the loss of signal power during penetration of a signal through different environments. As analytical formulas for SIR are cumbersome, two approximations of the probability density function of interference are developed: an analytical expression for the total interfering signal in the form of the normal distribution and a three-parameter hyperexponential distribution with numerically assorted parameters for simulation of the interfering signal from each device. As a result, the last one gives the most accurate approximation, which can be used for lower-bound estimation of SIR. The authors analyzed SIR in the case of the uniform distribution of devices in the premises of the rectangular shape for five wall materials (glass, brick, concrete, masonry blocks, and reinforced concrete) and varying wall thickness, and found that the presence of walls between rooms fundamentally changes the structure of SIR density.

Keywords: wireless networks; signal-to-interference ratio; indoor propagation; wall penetration

DOI: 10.14357/19922264160401

Acknowledgments

This work was financially supported by the Russian Science Foundation (grant No. 16-11-10227).

References

1. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2015–2020. Available at: www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html (accessed November 1, 2016).
2. Ericsson mobility report. Available at: www.ericsson.com/res/docs/2016/ericsson-mobility-report-2016.pdf (accessed November 1, 2016).
3. Recommendation ITU-T Y.2060. 2012. Overview of Internet of Thing. Available at: www.itu.int/rec/T-REC-Y.2060-201206-1 (accessed November 1, 2016).
4. Nokia demonstrates world's first 5G-ready network. Available at: company.nokia.com/en/news/press-releases/2016/06/28/nokia-demonstrates-worlds-first-5g-ready-network (accessed November 1, 2016).
5. Megafon pokazala mobil'nyy 5G-internet na skorosti 5 Gbit/s [MegaFon showed 5G mobile-Internet at the

- speed of 5 Gbps]. Available at: <http://www.ferra.ru/ru/mobile/news/2016/09/21/megafon-nn-5g> (accessed November 1, 2016).
6. Andrews, J., H. Claussen, M. Dohler, and S. Rangan. 2012. Femtocells: Past, present, and future. *IEEE JSAC* 30:497–508.
 7. Lee, J., H. Wang, J. Andrews, and D. Hong. 2011. Outage probability of cognitive relay networks with interference constraints. *IEEE Trans. Wir. Comm.* 10(2):390–395.
 8. Fodor, G., S. Parkvall, S. Sorrentino, P. Wallentin, Q. Lu, and N. Brahmī. 2014. Device-to-device communications for national security and public safety. *IEEE Access* 2:1510–1520.
 9. Sathya, V., A. Ramamurthy, S. Kumar, and B. Tamma. 2015. On improving SINR in LTE hetnets with D2D relays. *Comput. Commun.* 83:27–44.
 10. Baccelli, F., and B. Błaszczyszyn. 2010. Stochastic geometry and wireless networks: Vol. I Theory. *Foundations and Trends[®] in Networking* 3(3-4):249–449; 4(1-2):1–312. doi: 10.1561/1300000006.
 11. Abaev, P., K. Samouylov, Yu. Gaidamaka, and S. Shorgin. 2016 (in press). Tractable distance distribution approximations for hardcore processes. *DCCN 2016, CCIS 678*. Switzerland: Springer International Publishing.
 12. Gaidamaka, Yu., and A. Samuylov. 2015. Metod rascheta kharakteristik interferentsii dvukh vzaimodeystvuyushchikh ustroystv v besprovodnoy geterogennoy seti [Method for calculating numerical characteristics of two devices interference for device to device communications in a wireless heterogeneous network]. *Informatika i ee Primeneniya — Inform. Appl.* 9(1):10–15.
 13. Samuylov, A., D. Moltchanov, Yu. Gaidamaka, V. Begishev, R. Kovalchukov, P. Abaev, and S. Shorgin. 2016. SIR analysis in square-shaped indoor premises. *30th European Conference on Modelling and Simulation Proceedings*. Regensburg, Germany: Digitaldruck Pirrot GmbH. 692–697.
 14. Samuylov, A., A. Ometov, V. Begishev, R. Kovalchukov, D. Moltchanov, Yu. Gaidamaka, K. Samouylov, S. Andreev, and Y. Koucheryavy. 2015. Analytical performance estimation of network-assisted D2D communications in urban scenarios with rectangular cells. *Trans. Emerging Telecommunications Technol.* 1–14. doi: 10.1002/ett.2999A.
 15. Levin, B. R. 1989. *Teoreticheskie osnovy statisticheskoy radiotekhniki* [Theoretical basis of statistical radiotechnics]. 3rd ed. Moscow: Radio and Communications. 656 p.
 16. ITU-R. 2015. Propagation data and prediction methods for the planning of indoor radiocommunication systems and radio local area networks in the frequency range 300 MHz to 100 GHz. Recommendation ITU-R P.1238-8 (07-2015).

Received October 14, 2016

Contributors

Gaidamaka Yuliya V. (b. 1971) — Candidate of Science (PhD) in physics and mathematics, associate professor, Peoples' Friendship University of Russia, 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; ygaidamaka@sci.pfu.edu.ru

Andreev Sergey D. (b. 1984) — Candidate of Sciences (PhD) in technology; associate professor, Peoples' Friendship University of Russia, 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; serge.andeev@gmail.com

Sopin Eduard S. (b. 1986) — Candidate of Science (PhD) in physics and mathematics, associate professor, Peoples' Friendship University of Russia, 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; esopin@sci.pfu.edu.ru

Samouylov Konstantin E. (b. 1955) — Doctor of Science in technology, professor, Head of Department, Peoples' Friendship University of Russia, 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; 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; ksam@sci.pfu.edu.ru

Shorgin Sergey Ya. (b. 1952) — Doctor of Science in physics and mathematics, professor, Deputy Director, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; sshorgin@ipiran.ru

ТЕОРЕМА ПУАССОНА ДЛЯ СХЕМЫ ИСПЫТАНИЙ БЕРНУЛЛИ СО СЛУЧАЙНОЙ ВЕРОЯТНОСТЬЮ УСПЕХА И ДИСКРЕТНЫЙ АНАЛОГ РАСПРЕДЕЛЕНИЯ ВЕЙБУЛЛА*

В. Ю. Королев¹, А. Ю. Корчагин², А. И. Зейфман³

Аннотация: Рассматривается задача, связанная с испытаниями Бернулли со случайной вероятностью успеха. Сначала в результате «предварительного» эксперимента определяется значение случайной величины $\pi \in (0, 1)$, которое принимается в качестве вероятности успеха в испытаниях Бернулли. Затем случайная величина N определяется как число успехов в $k \in \mathbb{N}$ испытаниях Бернулли с так определенной вероятностью успеха π . Распределение случайной величины N называется π -смешанным биномиальным. В рамках такой схемы испытаний Бернулли со случайной вероятностью успеха формулируется «случайный» аналог классической теоремы Пуассона для π -смешанных биномиальных распределений, в котором предельным законом оказывается смешанное пуассоновское распределение. Особое внимание уделено случаю, в котором смешивающим распределением является распределение Вейбулла. Соответствующее смешанное пуассоновское распределение — пуассон-вейбулловское распределение — предложено в качестве дискретного аналога распределения Вейбулла. Обсуждаются некоторые свойства пуассон-вейбулловского распределения. В частности, показано, что это распределение является смешанным геометрическим. Предложен двухэтапный сеточный алгоритм оценивания параметров смешанных пуассоновских распределений и, в частности, пуассон-вейбулловского распределения. Построены статистические оценки верхней границы сетки. Приведены примеры вычислений по предложенному алгоритму.

Ключевые слова: испытания Бернулли со случайной вероятностью успеха; смешанное биномиальное распределение; теорема Пуассона; смешанное пуассоновское распределение; распределение Вейбулла; пуассон-вейбулловское распределение; смешанное геометрическое распределение; EM-алгоритм

DOI: 10.14357/19922264160402

1 Введение

Исследование, некоторые результаты которого излагаются в данной статье, мотивировано несколькими обстоятельствами. Во-первых, в последнее время получил серьезное развитие метод прогнозирования временных характеристик катастроф в неоднородных потоках экстремальных событий (см., например, [1]). Этот метод можно считать глубокой модернизацией метода превышений порога (POT-method, POT — Peaks Over Threshold). В рамках этого метода исходный ряд (маркированный точечный процесс) прореживается таким образом, что все наблюдения (точки), «марки» которых меньше указанного порога, выбрасываются. При этом, как правило, величина порога определя-

ется статистически, и потому вероятность, с которой очередное наблюдение отбрасывается, является случайной.

Во-вторых, схема простого прореживания процессов восстановления (см., например, [2, 3]) предоставляет естественный подход к определению редкого события, в рамках которого это понятие связывается с хорошо известной конструкцией пуассоновского процесса, характеризующегося показателем распределения интервалов времени между событиями («восстановлениями») в классе процессов восстановления. Было бы желательно иметь столь же простой и основанный на прореживании подход к конструкции смешанного пуассоновского процесса. В рамках такого подхода, построив смешанный пуассоновский процесс как

* Работа выполнена при поддержке Российского научного фонда (проект 14-11-00397).

¹ Факультет вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, vkorolev@cs.msu.ru

² Факультет вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, sasha.korchagin@gmail.com

³ Вологодский государственный университет; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук; Институт социально-экономического развития территорий Российской академии наук, a_zeifman@mail.ru

асимптотическую аппроксимацию для дважды стохастически прореженных процессов восстановления, можно надеяться получить дополнительное понимание структуры многих смежных математических моделей, в частности популярных ныне байесовских моделей и методов, и более осмысленно подойти в выборе соответствующего смешивающего («априорного») распределения.

В дальнейшем будет удобнее вести изложение не в терминах распределений, а в терминах случайных величин (с.в.) (предполагая, что все они заданы на одном вероятностном пространстве $(\Omega, \mathfrak{A}, P)$).

Символы $\stackrel{d}{=}$ и \implies будут соответственно обозначать совпадение распределений и сходимость по распределению.

Функция распределения (ф.р.) и плотность строго устойчивого распределения с характеристическим показателем α и параметром формы θ , определяемого характеристической функцией

$$f_{\alpha, \theta}(t) = \exp \left\{ -|t|^\alpha \exp \left\{ -\frac{1}{2} i\pi\theta\alpha \operatorname{sign} t \right\} \right\}, \quad t \in \mathbb{R},$$

где $0 < \alpha \leq 2$, $|\theta| \leq \min\{1, 2/\alpha - 1\}$, будут соответственно обозначаться $G_{\alpha, \theta}(x)$ и $g_{\alpha, \theta}(x)$ (см., например, [4]). Любую с.в. с ф.р. $G_{\alpha, \theta}(x)$ будем обозначать $S_{\alpha, \theta}$.

Симметричным строго устойчивым распределениям соответствует значение $\theta = 0$ и х.ф. $f_{\alpha, 0}(t) = e^{-|t|^\alpha}$, $t \in \mathbb{R}$. Односторонним строго устойчивым законам, сосредоточенным на неотрицательной полуоси, соответствуют значения $\theta = 1$ и $0 < \alpha \leq 1$. Пары $\alpha = 1, \theta = \pm 1$ отвечают распределениям, вырожденным в ± 1 соответственно. Остальные устойчивые распределения абсолютно непрерывны. Явные выражения устойчивых плотностей в терминах элементарных функций отсутствуют за четырьмя исключениями (нормальный закон ($\alpha = 2, \theta = 0$), распределение Коши ($\alpha = 1, \theta = 0$), распределение Леви ($\alpha = 1/2, \theta = 1$) и распределение, симметричное к распределению Леви ($\alpha = 1/2, \theta = -1$)). Выражения устойчивых плотностей в терминах функций Фокса (обобщенных G -функций Мейера) можно найти в [5, 6].

2 Смешанные биномиальные распределения и их асимптотическое поведение

Рассмотрим задачу, связанную с испытаниями Бернулли со случайной вероятностью успеха. Сначала в результате «предварительного» эксперимента

определяется значение с.в. $\pi \in (0, 1)$. Это значение принимается в качестве вероятности успеха в испытаниях Бернулли. Затем случайная величина N определяется как число успехов в $k \in \mathbb{N}$ испытаниях Бернулли с так определенной вероятностью успеха π . Чтобы описать бесконечную малость вероятности успеха π , снабдим последнюю и (для общности) параметр k , а также, соответственно, с.в. N «бесконечно большим» индексом n , позволяющим проследить сходимость последовательности с.в. $\pi = \pi_n$ к нулю при $n \rightarrow \infty$. В свою очередь, бесконечная малость π_n означает, что успехи являются редкими событиями в рамках рассматриваемой последовательности испытаний Бернулли.

В рамках схемы испытаний Бернулли со случайной вероятностью успеха, описанной выше, можно сформулировать и доказать «случайный» аналог классической теоремы Пуассона (так называемого «закона малых чисел») для π_n -смешанных биномиальных распределений со случайной вероятностью успеха и неограниченно возрастающим целочисленным параметром k_n («числом испытаний»). В известных вариантах «случайного» аналога теоремы Пуассона (см., к примеру, [7]), наоборот, случайным считалось число испытаний, а вероятность успеха оставалась неслучайной.

Пусть $k_n \in \mathbb{N}$, $k = 1, 2, \dots$. Будем говорить, что с.в. Q_n имеет π_n -смешанное биномиальное распределение с параметром k_n , если

$$P(Q_n = j) = C_{k_n}^j \int_0^1 z^k (1-z)^{k_n-j} dP(\pi_n < z), \quad j = 0, 1, \dots, k_n. \quad (1)$$

Для $x \in \mathbb{R}$ обозначим $B_n(x) = P(Q_n < x)$. Пусть N — положительная с.в. Смешанная пуассоновская ф.р. со структурной с.в. N (по терминологии, принятой в [8]) будет обозначаться $\Pi^{(N)}(x)$:

$$\Pi^{(N)}(x+0) = \sum_{j=0}^{[x]} \frac{1}{j!} \int_0^\infty e^{-z} z^j dP(N < z), \quad x \in \mathbb{R}.$$

В [9] доказана следующая теорема.

Теорема 1. Пусть $(k_n)_{n \geq 1}$ — неограниченно возрастающая последовательность натуральных чисел. Пусть Q_n — с.в. с π_n -смешанным биномиальным распределением (1) с целочисленным параметром k_n и ф.р. $B_n(x)$. Предположим, что в (1) с.в. π_n бесконечно малы в том смысле, что существует с.в. N такая, что $P(0 < N < \infty) = 1$ и выполнено условие $k_n \pi_n \implies N$ при $n \rightarrow \infty$. Тогда $B_n(x) \implies \Pi^{(N)}(x)$ ($n \rightarrow \infty$).

Необходимо отметить, что если прореживание процесса происходит по «независимой» схе-

ме, в которой имеется двойной массив $\{\pi_{n,j}\}_{j \geq 1}$, $n = 1, 2, \dots$, независимых в каждой серии с.в., причем $\pi_{n,j} \stackrel{d}{=} \pi_n$, $j = 1, 2, \dots$, при каждом $n \geq 1$, так что j -я точка исходного процесса удаляется с вероятностью $1 - \pi_{n,j}$, то, как несложно убедиться, процесс прорезживания сводится к классическому варианту, в котором предельный процесс является «чистым» пуассоновским.

3 Пуассон-вейбулловское распределение

Приведем два хорошо известных примера смешанных пуассоновских распределений. Во-первых, это *геометрическое распределение* как дискретный аналог непрерывного показательного распределения, который получается, если в смешанной пуассоновской модели «случайный» параметр пуассоновского распределения имеет показательное распределение.

Во-вторых, это *отрицательное биномиальное распределение* как дискретный аналог непрерывного гамма-распределения, который получается, если в смешанной пуассоновской модели «случайный» параметр пуассоновского распределения имеет гамма-распределение.

Пусть $W_{\gamma,\mu}$ — с.в., имеющая распределение Вейбулла с параметром масштаба $\mu > 0$ и параметром формы $\gamma > 0$: $P(W_{\gamma,\mu} < x) = 1 - \exp\{-\mu x^\gamma\}$, $x \geq 0$, и $P(W_{\gamma,\mu} < 0) = 0$. Распределение Вейбулла играет важную роль во многих прикладных задачах в качестве популярной и адекватной модели распределения времени жизни или безотказной работы. Известен дискретный аналог распределения Вейбулла, который получается формальным «квантованием» непрерывного распределения Вейбулла: если классический дискретный аналог с.в. $W_{\gamma,\mu}$ с распределением Вейбулла обозначить $\widetilde{W}_{\gamma,\mu}$, то обычно полагают

$$P(\widetilde{W}_{\gamma,\mu} = k) = e^{-\mu k^\gamma} - e^{-\mu(k+1)^\gamma}, \quad k = 0, 1, \dots$$

(см., например, [10]). У такого формального подхода есть существенный недостаток: крайне сложно (если вообще возможно) сформулировать предельную теорему в более-менее простой предельной схеме (например, суммирования или взятия экстремумов с.в.), в которой такое распределение было бы предельным. А стало быть, весьма проблематичным становится теоретическое обоснование возможности использования такого распределения в качестве асимптотической аппроксимации.

Используя аппарат смешанных пуассоновских распределений (являющихся предельными, напри-

мер, в теореме 1 и сходных с ней), можно предложить альтернативный аналог дискретного распределения Вейбулла как пуассон-вейбулловское распределение, т.е. смешанное пуассоновское распределение, в котором смешивание происходит по распределению Вейбулла.

Рассмотрим с.в. $V_{\gamma,\mu}$, имеющую смешанное пуассоновское распределение вида

$$\begin{aligned} P(V_{\gamma,\mu} = k) &= \frac{1}{k!} \int_0^\infty e^{-z} z^k dP(W_{\gamma,\mu} < z) = \\ &= \frac{\mu^\gamma}{k!} \int_0^\infty z^{k+\gamma-1} \exp\{-(z + \mu z^\gamma)\} dz, \quad k = 0, 1, 2, \dots \end{aligned}$$

Такое распределение будем называть *пуассон-вейбулловским*. Возможность считать пуассон-вейбулловское распределение дискретным аналогом распределения Вейбулла обусловлена не только формальным сходством с отрицательным биномиальным или геометрическим распределениями, являющимися смешанными пуассоновскими распределениями, в которых смешивание происходит именно по тем непрерывным распределениям (соответственно гамма- и показательному), дискретными аналогами которых они являются. Дополнительным аргументом можно считать следующее асимптотическое свойство пуассон-вейбулловского распределения.

Теорема 2. *Справедливо следующее асимптотическое соотношение: при $\mu \rightarrow 0$*

$$P(\mu^{1/\gamma} V_{\gamma,\mu} < x) \implies P(W_{\gamma,1} < x).$$

Доказательство этого утверждения основано на том, что, во-первых, $V_{\gamma,\mu} \stackrel{d}{=} P_1(W_{\gamma,\mu})$, где $P_1(t)$, $t \geq 0$, — стандартный пуассоновский процесс, независимый от с.в. $W_{\gamma,\mu}$, во-вторых, $W_{\gamma,\mu} \stackrel{d}{=} \mu^{-1/\gamma} W_{\gamma,1}$ и, в-третьих, $z^{-1} P_1(z\Lambda) \implies \Lambda$ при $z \rightarrow \infty$, где Λ — независимая от $P_1(t)$ неотрицательная с.в. (см., например, [7]).

Особый интерес представляет случай $0 < \gamma \leq 1$. Распределения Вейбулла с такими параметрами формы называются *растянутыми показательными* (stretched exponential) [11–13]. Они занимают промежуточное место между распределениями с хвостами, убывающими степенным образом, и распределениями с экспоненциально убывающими хвостами, играющими важную роль при математическом моделировании явлений и процессов на финансовых рынках. Оказывается, что пуассон-вейбулловское распределение с.в. $V_{\gamma,\mu}$ с $0 < \gamma \leq 1$

является смешанным геометрическим, т. е. описывает распределение числа испытаний до первого успеха в описанной выше схеме испытаний Бернулли со случайной вероятностью успеха, имеющей специальное распределение. Для простоты без ограничения общности в следующем утверждении считаем, что $\mu = 1$.

Теорема 3. Пуассон-вейбулловское распределение с.в. $V_{\gamma,1}$ с $0 < \gamma \leq 1$ является смешанным геометрическим:

$$P(V_{\gamma,1} = k) = \int_0^1 z(1-z)^k p_{\gamma,1}(z) dz, \quad k = 0, 1, \dots,$$

где плотность $p_{\gamma,1}(z)$ имеет вид:

$$p_{\gamma,1}(z) = \frac{1}{(1-z)^2} g_{\gamma,1}\left(\frac{z}{1-z}\right), \quad 0 \leq z \leq 1.$$

Доказательство. В работе [14] показано, что если $0 < \gamma \leq 1$, то $W_{\gamma,1} \stackrel{d}{=} W_{1,1} S_{\gamma,1}^{-1}$, где с.в. $W_{1,1}$ и $S_{\gamma,1}$ независимы. Тогда для $k \in \{0\} \cup \mathbb{N}$

$$\begin{aligned} P(V_{\gamma,1} = k) &= P(P_1(W_{\gamma,1}) = k) = \\ &= \int_0^\infty e^{-\lambda} \frac{\lambda^k}{k!} \left(\int_0^\infty z e^{-\lambda z} g_{\gamma,1}(z) dz \right) d\lambda = \\ &= \frac{1}{k!} \int_0^\infty z g_{\gamma,1}(z) \left(\int_0^\infty e^{-\lambda(z+1)} \lambda^k d\lambda \right) dz = \\ &= \frac{\Gamma(k+1)}{k!} \int_0^\infty \frac{z g_{\gamma,1}(z)}{(z+1)^{k+1}} dz = \\ &= \int_0^\infty \frac{z}{z+1} \left(1 - \frac{z}{z+1} \right)^k g_{\gamma,1}(z) dz = \\ &= \int_0^1 z(1-z)^k g_{\gamma,1}\left(\frac{z}{1-z}\right) \frac{dz}{(1-z)^2} = \\ &= \int_0^1 z(1-z)^k p_{\gamma,1}(z) dz, \end{aligned}$$

что и требовалось доказать.

Как известно, производящая функция моментов с.в. $W_{\gamma,\mu}$ имеет вид:

$$\Psi_{\gamma,\mu}(t) = E \exp\{tW_{\gamma,\mu}\} = \sum_{n=0}^\infty \frac{t^n}{\mu^{n/\gamma} n!} \Gamma\left(1 + \frac{n}{\gamma}\right), \quad t \in \mathbb{R}.$$

В общем случае производящая функция $\psi_{\gamma,\mu}(s) \equiv E s^{V_{\gamma,\mu}}$ пуассон-вейбулловского распределения имеет вид:

$$\begin{aligned} \psi_{\gamma,\mu}(s) &= \sum_{k=0}^\infty s^k P(V_{\gamma,\mu} = k) = \\ &= \sum_{k=0}^\infty \frac{s^k}{k!} \int_0^\infty e^{-z} z^k dP(W_{\gamma,\mu} < z) = \\ &= \int_0^\infty e^{-z} \left[\sum_{k=0}^\infty \frac{(sz)^k}{k!} \right] dP(W_{\gamma,\mu} < z) = \\ &= \int_0^\infty e^{z(s-1)} dP(W_{\gamma,\mu} < z) = E \exp\{(s-1)W_{\gamma,\mu}\} = \\ &= \Psi_{\gamma,\mu}(s-1) = \sum_{n=0}^\infty \frac{(s-1)^n}{\mu^{n/\gamma} n!} \Gamma\left(1 + \frac{n}{\gamma}\right), \quad s \in [0, 1]. \end{aligned}$$

Отсюда имеем:

$$\begin{aligned} P(V_{\gamma,\mu} = 0) &= \psi_{\gamma,\mu}(s)|_{s=0} = \Psi_{\gamma,\mu}(-1) = \\ &= \sum_{n=0}^\infty \frac{\Gamma(1+n/\gamma)(-1)^n}{\mu^{n/\gamma} n!}; \\ P(V_{\gamma,\mu} = 1) &= \frac{d\psi_{\gamma,\mu}(s)}{ds} \Big|_{s=0} = \\ &= \sum_{n=0}^\infty \frac{\Gamma(1+n/\gamma)}{\mu^{n/\gamma} n!} \frac{d}{ds} (s-1)^n \Big|_{s=0} = \\ &= \sum_{n=1}^\infty \frac{\Gamma(1+n/\gamma)(-1)^{n-1}}{\mu^{n/\gamma} (n-1)!}; \\ P(V_{\gamma,\mu} = 2) &= \frac{d^2\psi_{\gamma,\mu}(s)}{2ds^2} \Big|_{s=0} = \\ &= \frac{1}{2} \sum_{n=0}^\infty \frac{\Gamma(1+n/\gamma)}{\mu^{n/\gamma} n!} \frac{d^2}{ds^2} (s-1)^n \Big|_{s=0} = \\ &= \frac{1}{2} \sum_{n=2}^\infty \frac{\Gamma(1+n/\gamma)(-1)^{n-2}}{\mu^{n/\gamma} (n-2)!}; \dots; \\ P(V_{\gamma,\mu} = k) &= \frac{d^k\psi_{\gamma,\mu}(s)}{k! ds^k} \Big|_{s=0} = \\ &= \frac{1}{k!} \sum_{n=k}^\infty \frac{\Gamma(1+n/\gamma)(-1)^{n-k}}{\mu^{n/\gamma} (n-k)!}, \quad k = 3, 4, \dots \end{aligned}$$

Как видно, это распределение весьма громоздко и «прямое» оценивание его параметров представляется нетривиальной задачей.

4 Двухэтапный сеточный EM-алгоритм для оценивания параметров смешанных пуассоновских распределений и, в частности, параметров пуассон-вейбулловского распределения

По сути оценивание параметров смешанных пуассоновских моделей сводится к оцениванию смешивающего распределения. Традиционно с этой целью используется классический EM (expectation-maximization) алгоритм [15]. Однако иногда, в частности в пуассон-вейбулловском случае, классический EM-алгоритм оказывается менее эффективным, чем альтернативный двухэтапный сеточный EM-алгоритм оценивания параметров смешанных пуассоновских распределений.

Следует отметить, что сеточные методы разделения смесей довольно эффективны не только при оценивании параметров смешанных пуассоновских распределений, но и при разделении конечных или произвольных дисперсионно-сдвиговых смесей нормальных законов [16].

Рассмотрим следующий двухэтапный метод разделения смешанных пуассоновских распределений на примере оценивания параметров γ, μ пуассон-вейбулловского распределения $\Pi^{(N)}(x) = \Pi^{(W_{\gamma,\mu})}(x)$.

На первом этапе на положительной полупрямой выделим основную часть носителя смешивающего распределения, т. е. ограниченный интервал, вероятность которого, вычисленная в соответствии со смешивающим распределением, практически равна единице. На этот интервал накинём конечную сетку, содержащую (возможно, очень большое число) $K \in \mathbb{N}$ известных узлов $\lambda_1, \dots, \lambda_K$. Приближим искомое смешанное пуассоновское распределение конечной смесью пуассоновских законов:

$$\Pi^{(W_{\gamma,\mu})}(x+0) \approx \sum_{j=0}^{\lfloor x \rfloor} \frac{1}{j!} \sum_{i=1}^K p_i e^{-\lambda_i} \lambda_i^j, \quad x \in \mathbb{R}. \quad (2)$$

В смеси, стоящей в правой части соотношения (2), неизвестными являются только параметры p_1, \dots, p_K . Пусть x_1, \dots, x_n — анализируемая выборка значений случайной величины с оцениваемым смешанным пуассоновским распределением. Итерационный процесс, определяющий сеточный EM-алгоритм для данной задачи, задается следующим образом. Пусть $p_1^{(m)}, \dots, p_{K-1}^{(m)}$ — оценки

параметров p_1, \dots, p_{K-1} на m -й итерации, $p_K^{(m)} = 1 - p_1^{(m)} - \dots - p_{K-1}^{(m)}$. Для $i = 1, \dots, K, j = 1, \dots, n$ обозначим $\phi_{ij} = e^{-\lambda_i + x_j \ln \lambda_i}$. Тогда, используя стандартные рассуждения, определяющие вычислительные формулы EM-алгоритма для параметров конечной смеси вероятностных распределений (см, например, [17, разд. 5.3.7–5.3.8]), следует положить:

$$p_i^{(m+1)} = \frac{p_i^{(m)}}{n} \sum_{j=1}^n \frac{\phi_{ij}}{\sum_{r=1}^K p_r^{(m)} \phi_{rj}}, \quad i = 1, \dots, K. \quad (3)$$

Как видим, итерационный процесс, задаваемый соотношением (3), очень прост. В силу монотонности классического EM-алгоритма справедливо следующее утверждение.

Теорема 4. Пусть узлы $\lambda_1, \dots, \lambda_K$ сетки различны, неотрицательны и известны. Тогда итерационный процесс (3) является монотонным, т. е. каждая его итерация не уменьшает целевую сеточную функцию правдоподобия

$$L(p_1, \dots, p_K; x_1, \dots, x_n) = \prod_{j=1}^n \left[\sum_{i=1}^K p_i \phi_{ij} \right].$$

Заметим, что, как показано в [17, разд. 5.7.4], сеточная функция правдоподобия $L(p_1, \dots, p_K; x_1, \dots, x_n)$ вогнута по аргументам p_1, \dots, p_K . Поэтому на каждом шаге итерационного процесса вместо соотношения (3) можно использовать любой более быстрый алгоритм максимизации функции $L(p_1, \dots, p_K; x_1, \dots, x_n)$ по переменным p_1, \dots, p_K . Например, оценки весов p_1, \dots, p_K можно искать методом условного градиента [17, 18].

Таким образом, на первом этапе получаются оценки весов p_i всех узлов $\lambda_i, i = 1, \dots, K$, конечной сетки, накинутой на носитель смешивающего распределения.

На втором этапе остается применить какой-либо стандартный метод подгонки распределения Вейбулла $P(W_{\gamma,\mu} < x)$ к эмпирическим данным типа гистограммы $(\lambda_1, p_1), \dots, (\lambda_K, p_K)$, полученным на первом этапе. Например, параметры γ и μ можно оценить, минимизируя соответствующую статистику хи-квадрат. Или же, например, можно решить задачу наименьших квадратов:

$$\begin{aligned} (\gamma^*, \mu^*) &= \\ &= \arg \min_{\gamma, \mu} \sum_{i=1}^K \left[p_i - \exp \left\{ -\frac{\mu}{2} (\lambda_{i-1} + \lambda_i)^\gamma \right\} + \right. \\ &\quad \left. + \exp \left\{ -\frac{\mu}{2} (\lambda_i + \lambda_{i+1})^\gamma \right\} \right]^2, \end{aligned}$$

где $\lambda_0 = 0, \lambda_{K+1} = \infty$.

На практике хорошие результаты показал подход с решением задачи наименьших квадратов. Для поиска параметров использовался алгоритм ns2sol, описанный в книге [19]. Указанный алгоритм доступен во многих статистических пакетах, отличается высоким быстродействием и возможностью при желании задавать разумные интервалы для поиска параметров.

Также хорошие результаты показал метод поиска наилучшего распределения в смысле минимизации расстояния Кульбака–Лейблера, который в данном случае эквивалентен максимизации правдоподобия полученной гистограммы в классе распределений Вейбулла.

При фиксированной сетке двухэтапный метод дает лишь приближенные оценки параметров смешанных пуассоновских распределений, причем точность приближения зависит от успешного выбора сетки. Некоторые аспекты этого выбора будут рассмотрены в следующем разделе. Говорить о состоятельности получаемых оценок при фиксированной сетке нельзя. Но если объем выборки неограниченно возрастает и вместе с ним согласованно увеличивается число узлов, то вопрос о состоятельности получаемых оценок приобретает смысл и будет рассмотрен в одной из следующих публикаций.

5 О практическом выборе сетки на первом этапе двухэтапного сеточного ЕМ-алгоритма для разделения смесей пуассоновских распределений

Естественно, что при использовании указанного двухэтапного метода в динамическом режиме крайне важным становится вопрос о выборе наиболее эффективных и быстродействующих численных процедур и их параметров. В частности, исключительную важность приобретает правильный выбор границ сетки на первом этапе. Рассмотрим этот вопрос подробнее.

Формально рассматриваемая задача выглядит так: по наблюдаемым значениям x_1, \dots, x_n требуется построить статистическую оценку верхней границы квантилей заданного порядка смешивающего закона так, чтобы как можно точнее оценить носитель смешивающего распределения.

В дальнейшем будем считать, что x_1, \dots, x_n — независимые реализации с.в. $X \stackrel{d}{=} P_1(\Lambda)$, где $P_1(t)$ — стандартный пуассоновский процесс, $t \geq 0$, Λ — независимая от $P_1(t)$ неотрицательная с.в. В случае

пуассон-вейбулловского распределения $X \stackrel{d}{=} V_{\gamma, \mu}$, $\Lambda \stackrel{d}{=} W_{\gamma, \mu}$.

Сначала рассмотрим более общий случай и предположим, что $E\Lambda^2 < \infty$. Тогда $EX = E\Lambda$, $DX = E\Lambda + D\Lambda = EX + D\Lambda$. Следовательно, $D\Lambda = DX - EX$. Но $D\Lambda = E\Lambda^2 - (E\Lambda)^2 = E\Lambda^2 - (EX)^2$. Поэтому $E\Lambda^2 = D\Lambda + (EX)^2 = DX - EX + (EX)^2 = EX^2 - EX$. Таким образом, для $\lambda > 0$ по неравенству Маркова имеем:

$$P(\Lambda \geq \lambda) \leq \frac{E\Lambda^2}{\lambda^2} = \frac{EX^2 - EX}{\lambda^2}. \quad (4)$$

Отсюда, задав произвольно малое $\varepsilon > 0$, можно найти приближенную верхнюю оценку $(1-\varepsilon)$ -квантили $\lambda^{(1-\varepsilon)}$ с.в. Λ . С этой целью положим:

$$\lambda_\varepsilon = \sqrt{\frac{EX^2 - EX}{\varepsilon}}.$$

Тогда из (4) вытекает, что $P(\Lambda \geq \lambda_\varepsilon) \leq \varepsilon$, т.е. можно положить $\lambda_K = \lambda^{(1-\varepsilon)}$, причем

$$\lambda^{(1-\varepsilon)} \leq \lambda_\varepsilon = \sqrt{\frac{EX^2 - EX}{\varepsilon}} \approx \sqrt{\frac{1}{n\varepsilon} \sum_{j=1}^n x_j(x_j - 1)}.$$

Теперь отдельно рассмотрим случай, когда $Ee^X < \infty$. Этот случай, в частности, имеет место для пуассон-вейбулловского распределения $X \stackrel{d}{=} V_{\gamma, \mu}$, $\Lambda \stackrel{d}{=} W_{\gamma, \mu}$, если $\gamma > 1$ или $\gamma = 1$ и $\mu > e - 1$. Имеем

$$\begin{aligned} Ee^X &= \sum_{k=0}^{\infty} \frac{e^k}{k!} \int_0^{\infty} e^{-\lambda} \lambda^k dP(\Lambda < \lambda) = \\ &= \int_0^{\infty} e^{-\lambda} \left(\sum_{k=0}^{\infty} \frac{(e\lambda)^k}{k!} \right) dP(\Lambda < \lambda) = \\ &= \int_0^{\infty} e^{\lambda(e-1)} dP(\Lambda < \lambda) = Ee^{\Lambda(e-1)}. \end{aligned}$$

Тогда

$$\begin{aligned} P(\Lambda \geq \lambda) &= P(\Lambda(e-1) \geq \lambda(e-1)) \leq \\ &\leq \frac{Ee^{\Lambda(e-1)}}{e^{\lambda(e-1)}} = \frac{Ee^X}{e^{\lambda(e-1)}}. \end{aligned} \quad (5)$$

Для произвольно малого положительного ε положим:

$$\lambda_\varepsilon = \frac{\ln Ee^X - \ln \varepsilon}{e - 1}.$$

Тогда из (5) вытекает, что $P(\Lambda \geq \lambda_\varepsilon) \leq \varepsilon$, т. е. можно положить $\lambda_K = \lambda^{(1-\varepsilon)}$, причем

$$\lambda_K = \lambda^{(1-\varepsilon)} \leq \lambda_\varepsilon = \frac{\ln Ee^X - \ln \varepsilon}{e - 1} \approx \frac{1}{e - 1} \ln \left(\frac{1}{n\varepsilon} \sum_{j=1}^n e^{x_j} \right).$$

6 Примеры

В этом разделе приведены результаты тестовых вычислений по описанному выше алгоритму (рис. 1–4). Моделировались искусственные выборки из пуассон-вейбулловского распределения, к которым применялся описанный выше двухэтап-

ный метод. Необходимо отметить, что во всех случаях размер сетки был относительно небольшим ($K = 15$), тем не менее достигнута приемлемая точность. В каждой паре рисунков слева — графики «истинной» смешивающей плотности и ее оценки, полученной сеточным методом, справа — графики соответствующего «истинного» пуассон-вейбулловского распределения и его статистической оценки.

Авторы выражают благодарность И. Г. Шевцовой и А. В. Дорофеевой за участие в работе по тестированию алгоритма, проведению вычислений и построению графиков, проведенную в рамках выполнения проекта № 14-11-00397 Российского научного фонда.

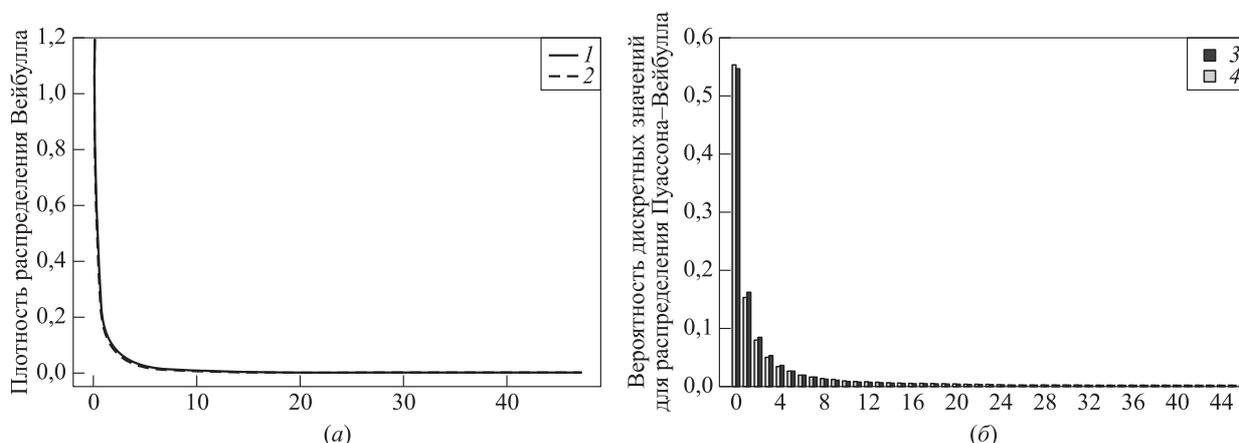


Рис. 1 Графики «истинной» смешивающей плотности с $\mu = 1$ и $\gamma = 1/2$ (1) и ее оценки, полученной двухэтапным сеточным методом (2) (а) и графики соответствующего «истинного» пуассон-вейбулловского распределения (3) и его статистической оценки (4) (б). Объем выборки $n = 500$

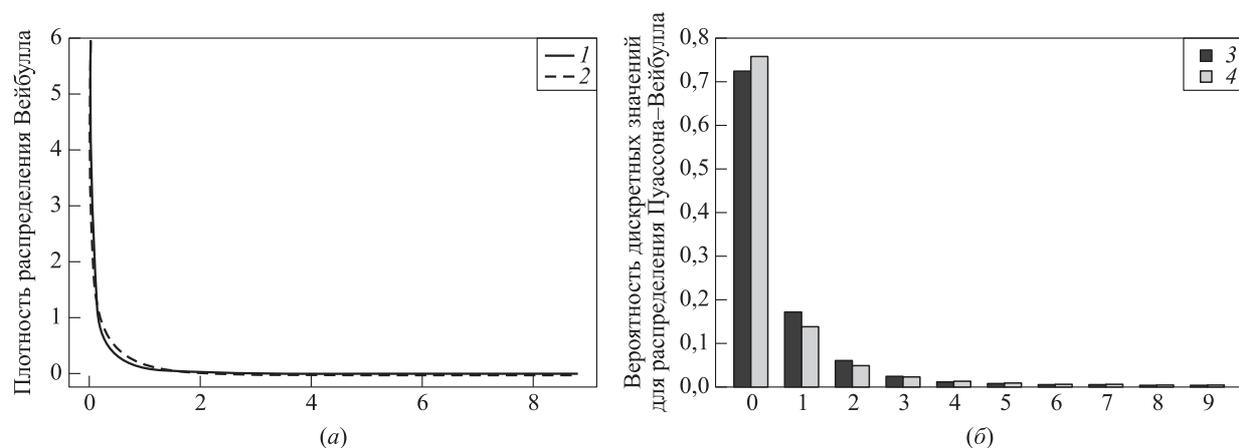


Рис. 2 Графики «истинной» смешивающей плотности с $\mu = 2$ и $\gamma = 1/2$ (1) и ее оценки, полученной двухэтапным сеточным методом (2) (а) и графики соответствующего «истинного» пуассон-вейбулловского распределения (3) и его статистической оценки (4) (б). Объем выборки $n = 500$

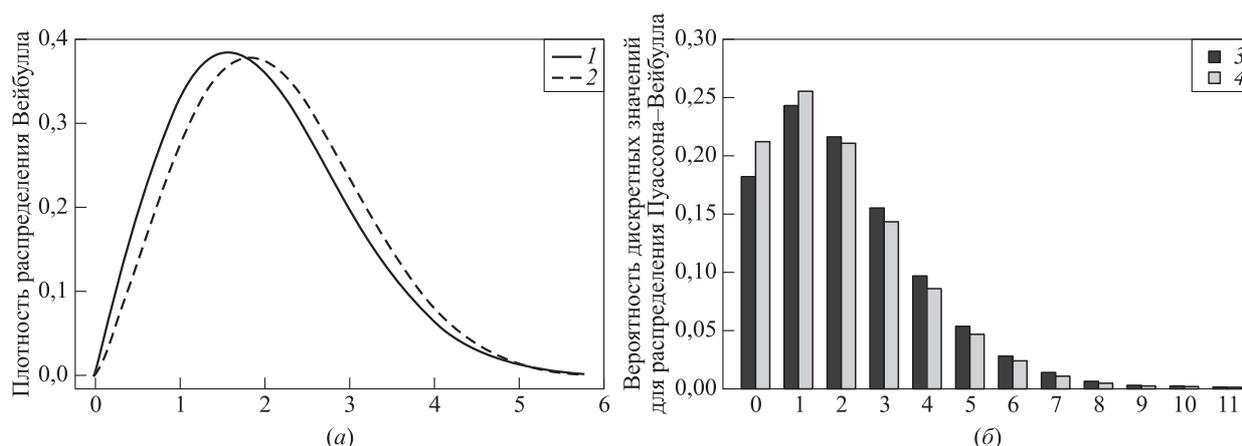


Рис. 3 Графики «истинной» смешивающей плотности с $\mu = 0,2$ и $\gamma = 2$ (1) и ее оценки, полученной двухэтапным сеточным методом (2) и графики соответствующего «истинного» пуассон-вейбулловского распределения (3) и его статистической оценки (4) (б). Объем выборки $n = 1000$

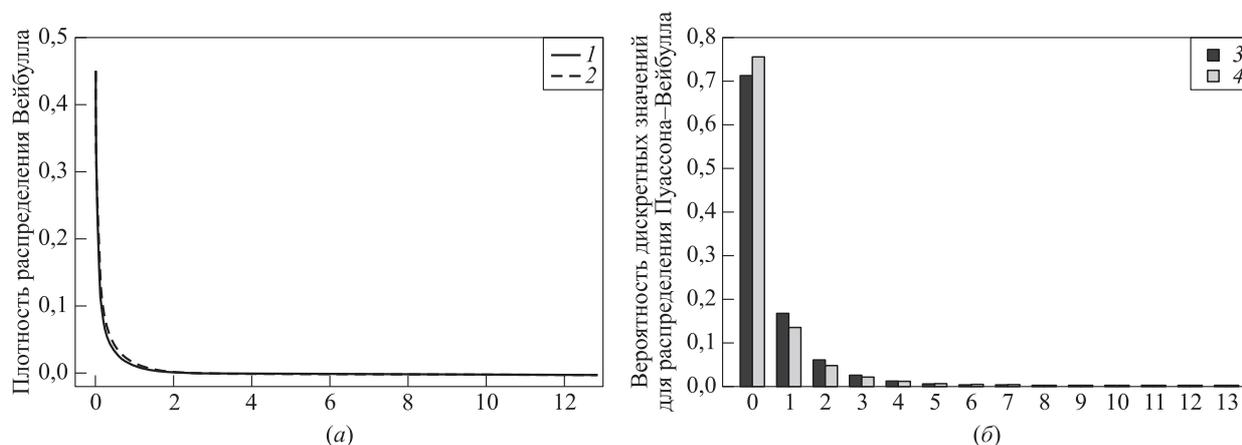


Рис. 4 Графики «истинной» смешивающей плотности с $\mu = 2$ и $\gamma = 1/2$ (1) и ее оценки, полученной двухэтапным сеточным методом (2) и графики соответствующего «истинного» пуассон-вейбулловского распределения (3) и его статистической оценки (4) (б). Объем выборки $n = 1000$

Литература

1. Григорьева М. Е., Королев В. Ю., Соколов И. А. Предельная теорема для геометрических сумм независимых неодинаково распределенных случайных величин и ее применение к прогнозированию вероятности катастроф в неоднородных потоках экстремальных событий // Информатика и её применения, 2013. Т. 7. Вып. 4. С. 11–19.
2. Rényi A. A Poisson-folyamat egy jellemzese // Magyar Tud. Acad. Mat. Int. Kozl., 1956. Vol. 1. P. 519–527.
3. Mogyorodi J. Some notes on thinning recurrent flows // Litovsky Math. Sbornik, 1971. Vol. 11. P. 303–315.
4. Золотарев В. М. Одномерные устойчивые распределения. — М.: Наука, 1983. 304 с.
5. Schneider W. R. Stable distributions: Fox function representation and generalization // Stochastic processes in classical and quantum systems / Eds. S. Albeverio, G. Casati, D. Merlini. — Berlin: Springer, 1986. P. 497–511.
6. Uchaikin V. V., Zolotarev V. M. Chance and stability. — Utrecht: VSP, 1999. 570 p.
7. Королев В. Ю., Бенинг В. Е., Шоргин С. Я. Математические основы теории риска. — 2-е изд. — М.: Физматлит, 2011. 591 с.
8. Grandell J. Mixed Poisson processes. — London: Chapman and Hall, 1997. 268 p.
9. Korolev V. Yu., Korchagin A. Yu., Zeifman A. I. On doubly stochastic rarefaction of renewal processes // 14th Conference (International) of Numerical Analysis and Applied Mathematics Proceedings. — American Institute of Physics Proceedings, 2017 (in press).
10. Nakagawa T., Osaki Sh. The discrete Weibull distribution // IEEE Trans. Reliab., 1975. Vol. 24. P. 300–301.

11. *Laherrère J., Sornette D.* Stretched exponential distributions in nature and economy: “Fat tails” with characteristic scales // *Eur. Phys. J. B*, 1998. Vol. 2. P. 525–539.
12. *Malevergne Y., Pisarenko V., Sornette D.* Empirical distributions of stock returns: Between the stretched exponential and the power law? // *Quant. Financ.*, 2005. Vol. 5. P. 379–401.
13. *Malevergne Y., Pisarenko V., Sornette D.* On the power of generalized extreme value (GEV) and generalized Pareto distribution (GDP) estimators for empirical distributions of stock returns // *Appl. Financ. Econ.*, 2006. Vol. 16. P. 271–289.
14. *Korolev V. Yu.* Product representations for random variables with the Weibull distributions and their applications // *J. Math. Sci.*, 2016. Vol. 218. No. 3. P. 298–313.
15. *Karlis D.* An EM algorithm for mixed Poisson distributions // *ASTIN Bull.*, 2005. Vol. 35. P. 3–24.
16. *Королев В. Ю., Корчагин А. Ю.* Модифицированный сеточный метод разделения дисперсионно-сдвиговых смесей нормальных законов // *Информатика и её применения*, 2014. Т. 8. Вып. 4. С. 11–19.
17. *Королев В. Ю.* Вероятностно-статистические методы декомпозиции волатильности хаотических процессов. — М.: Изд-во Московского ун-та, 2011. 510 с.
18. *Королев В. Ю., Назаров А. Л.* Разделение смесей вероятностных распределений при помощи сеточных методов моментов и максимального правдоподобия // *Автоматика и телемеханика*, 2010. Вып. 3. С. 98–116.
19. *Dennis J. E., Schnabel R. B.* Numerical methods for unconstrained optimization and nonlinear equations. — Englewood Cliffs: Prentice-Hall, 1983. 375 p.

Поступила в редакцию 15.10.16

THE POISSON THEOREM FOR BERNOULLI TRIALS WITH A RANDOM PROBABILITY OF SUCCESS AND A DISCRETE ANALOG OF THE WEIBULL DISTRIBUTION

V. Yu. Korolev^{1,2}, A. Yu. Korchagin^{1,2}, and A. I. Zeifman^{2,3,4}

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

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

³Vologda State University, 15 Lenin Str., Vologda 160000, Russian Federation

⁴ISEDT RAS, 56-A Gorky Str., Vologda 16001, Russian Federation

Abstract: A problem related to the Bernoulli trials with a random probability of success is considered. First, as a result of the preliminary experiment, the value of the random variable $\pi \in (0, 1)$ is determined that is taken as the probability of success in the Bernoulli trials. Then, the random variable N is determined as the number of successes in $k \in \mathbb{N}$ Bernoulli trials with the so determined success probability π . The distribution of the random variable N is called π -mixed binomial. Within the framework of these Bernoulli trials with the random probability of success, a “random” analog of the classical Poisson theorem is formulated for the π -mixed binomial distributions, in which the limit distribution turns out to be the mixed Poisson distribution. Special attention is paid to the case where mixing is performed with respect to the Weibull distribution. The corresponding mixed Poisson distribution called Poisson–Weibull law is proposed as a discrete analog of the Weibull distribution. Some properties of the Poisson–Weibull distribution are discussed. In particular, it is shown that this distribution can be represented as the mixed geometric distribution. A two-stage grid algorithm is proposed for estimation of parameters of mixed Poisson distributions and, in particular, of the Poisson–Weibull distribution. Statistical estimators for the upper bound of the grid are constructed. The examples of practical computations performed by the proposed algorithm are presented.

Keywords: Bernoulli trials with a random probability of success; mixed binomial distribution; Poisson theorem; mixed Poisson distribution; Weibull distribution; Poisson–Weibull distribution; mixed geometric distribution; EM-algorithm

DOI: 10.14357/19922264160402

Acknowledgments

This work was financially supported by the Russian Science Foundation (grant No. 14-11-00397).

References

1. Grigoryeva, M. E., V. Yu. Korolev, and I. A. Sokolov. 2013. Predel'naya teorema dlya geometricheskikh summ nezavisimyykh neodnakovo raspredelennykh sluchaynykh velichin i ee primeneniye k prognozirovaniyu veroyatnosti katastrof v neodnorodnykh potokakh ekstremal'nykh sobytii [A limit theorem for geometric sums of independent nonidentically distributed random variables and its application to the prediction of the probabilities of catastrophes in nonhomogeneous flows of extremal events]. *Informatika i ee Primeneniya — Inform. Appl.* 7(4):11–19.
2. Rényi, A. 1956. A Poisson-folyamat egy jellemzese. *Magyar Tud. Acad. Mat. Int. Kozl.* 1:519–527.
3. Mogyorodi, J. 1971. Some notes on thinning recurrent flows. *Litovskiy Math. Sbornik* 11:303–315.
4. Zolotarev, V. M. 1983. *Odnomernye ustoychivye raspredeleniya* [One-dimensional stable distributions]. Moscow: Nauka. 304 p.
5. Schneider, W. R. 1986. Stable distributions: Fox function representation and generalization. *Stochastic processes in classical and quantum systems*. Eds. S. Albeverio, G. Casati, and D. Merlini. Berlin: Springer. 497–511.
6. Uchaikin, V. V., and V. M. Zolotarev. 1999. *Chance and stability*. Utrecht: VSP. 570 p.
7. Korolev, V. Yu., V. E. Bening, and S. Ya. Shorgin. 2011. *Matematicheskie osnovy teorii riska* [Mathematical fundamentals of risk theory]. 2nd ed. Moscow: Fizmatlit. 591 p.
8. Grandell, J. 1997. *Mixed Poisson processes*. London: Chapman and Hall. 268 p.
9. Korolev, V. Yu., A. Yu. Korchagin, and A. I. Zeifman. 2017 (in press). On doubly stochastic rarefaction of renewal processes. *14th Conference (International) of Numerical Analysis and Applied Mathematics Proceedings*. American Institute of Physics Proceedings.
10. Nakagawa, T., and Sh. Osaki. 1975. The discrete Weibull distribution. *IEEE Trans. Reliab.* 24:300–301.
11. Laherrère, J., and D. Sornette. 1998. Stretched exponential distributions in nature and economy: “Fat tails” with characteristic scales. *Eur. Phys. J. B* 2:525–539.
12. Malevergne, Y., V. Pisarenko, and D. Sornette. 2005. Empirical distributions of stock returns: Between the stretched exponential and the power law? *Quant. Financ.* 5:379–401.
13. Malevergne, Y., V. Pisarenko, and D. Sornette. 2006. On the power of generalized extreme value (GEV) and generalized Pareto distribution (GPD) estimators for empirical distributions of stock returns. *Appl. Financ. Econ.* 16:271–289.
14. Korolev, V. Yu. 2016. Product representations for random variables with the Weibull distributions and their applications. *J. Math. Sci.* 218(3):298–313.
15. Karlis, D. 2005. An EM algorithm for mixed Poisson distributions. *ASTIN Bull.* 35:3–24.
16. Korolev, V. Yu., and A. Yu. Korchagin. 2014. Modifitsirovannyi setochnyy metod razdeleniya dispersionnosdvigovykh smesey normal'nykh zakonov [Modified grid method for decomposition of mean-variance normal mixtures]. *Informatika i ee Primeneniya — Inform. Appl.* 8(4):11–19.
17. Korolev, V. Yu. 2011. *Veroyatnostno-statisticheskie metody dekompozitsii volatil'nosti khaoticheskikh protsessov* [Probability-based method for volatility decomposition of chaotic processes]. Moscow: Moscow University Press. 510 p.
18. Korolev, V. Yu., and A. L. Nazarov. 2010. Separating mixtures of probability distributions with the grid maximum likelihood method. *Avtomat. Rem. Contr.* 71(3):455–472.
19. Dennis, J. E., and R. B. Schnabel. 1983. *Numerical methods for unconstrained optimization and nonlinear equations*. Englewood Cliffs: Prentice-Hall. 375 p.

Received October 15, 2016

Contributors

Korolev Victor Yu. (b. 1954) — Doctor of Science in physics and mathematics, professor, Head of the Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; 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; vkorolev@cs.msu.su

Korchagin Alexander Yu. (b. 1989) — junior scientist, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; 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; sasha.korchagin@gmail.com

Zeifman Alexander I. (b. 1954) — Doctor of Science in physics and mathematics, professor, Head of Department, Vologda State University, 15 Lenin Str., Vologda 160000, 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; principal scientist, ISEDT RAS, 56-A Gorky Str., Vologda 16001, Russian Federation; a_zeifman@mail.ru

НЕСИММЕТРИЧНЫЕ РАСПРЕДЕЛЕНИЯ ЛИННИКА КАК ПРЕДЕЛЬНЫЕ ЗАКОНЫ ДЛЯ СЛУЧАЙНЫХ СУММ НЕЗАВИСИМЫХ СЛУЧАЙНЫХ ВЕЛИЧИН С КОНЕЧНЫМИ ДИСПЕРСИЯМИ*

В. Ю. Королев¹, А. И. Зейфман², А. Ю. Корчагин³

Аннотация: Распределения Линника (симметричные геометрически устойчивые распределения) находят широкое применение в финансовой математике, телекоммуникационных системах, астрофизике, генетике. Такие распределения являются предельными для геометрических сумм независимых одинаково распределенных случайных величин (с.в.), распределения которых принадлежат области нормального притяжения симметричного строго устойчивого распределения. В статье рассматриваются три несимметричных обобщения распределения Линника. Традиционный (и формальный) подход к несимметричному обобщению распределения Линника заключается в рассмотрении геометрических сумм случайных слагаемых, распределения которых притягиваются к *несимметричному* строго устойчивому распределению. Дисперсии таких слагаемых бесконечны. Поскольку при моделировании реальных явлений, как правило, нет веских причин отвергать предположение о конечности дисперсии элементарных слагаемых, в качестве альтернатив традиционному подходу в статье предложены несимметричные обобщения, основанные на представлении распределения Линника в виде смеси нормальных распределений и смеси распределений Лапласа. Приведены примеры предельных теорем для сумм случайного числа независимых с.в. с *конечными дисперсиями*, в которых предложенные несимметричные распределения Линника выступают в качестве предельных законов.

Ключевые слова: распределение Линника; распределение Лапласа; распределение Миттаг–Леффлера; нормальное распределение; масштабная смесь; дисперсионно-сдвиговая смесь нормальных законов; устойчивое распределение; геометрически устойчивое распределение

DOI: 10.14357/19922264160403

1 Введение

Распределения Линника (симметричные геометрически устойчивые распределения) имеют довольно широкое применение в финансовой математике, телекоммуникационных системах, астрофизике, генетике. Разнообразные приложения распределений Линника описаны в работах [1, 2]. В частности, распределения Линника возникают при изучении механизма синтеза мелатонина в человеческом организме, солнечных нейтринных потоков в космосе, явлений роста-упадка в природе, эконометрических явлений и т. п.

Геометрически устойчивые законы и только они могут быть предельными распределениями для геометрических сумм случайных сумм независимых одинаково распределенных с.в. Поэтому традиционно несимметричное обобщение распределения Линника достигается за счет того, что в *геометрической* случайной сумме рассматриваются скошенные слагаемые. При этом распределения слагаемых принадлежат области нормального притяжения несимметричного устойчивого закона с некоторым показателем $\alpha \in (0, 2]$ и, значит, при $0 < \alpha < 2$ имеют бесконечные моменты порядков, больших или равных α . Что касается случая $\alpha = 2$, ко-

*Работа выполнена при финансовой поддержке Российского научного фонда (проект 14-11-00364).

¹Факультет вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, vkorolev@cs.msu.ru

²Вологодский государственный университет; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук; Институт социально-экономического развития территории Российской академии наук, Факультет вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова, a_zeifman@mail.ru

³Факультет вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова, sasha.korchagin@gmail.com

гда конечна дисперсия, то в рамках схемы геометрического суммирования он неизбежно приводит к единственно возможному распределению — распределению Лапласа.

Однако при использовании распределений Линника в качестве моделей реальных явлений нельзя не задуматься над вопросом о том, что если используется аддитивная структурная модель реального процесса типа случайно остановленного случайного блуждания, то какая комбинация условий встречается чаще:

- распределение числа слагаемых (числа скачков) является геометрическим (асимптотически экспоненциальным), но слагаемые (скачки) имеют столь тяжелые хвосты, что как минимум у них бесконечна дисперсия, или
- вторые моменты (дисперсии) слагаемых (скачков) конечны, но число слагаемых отличается нерегулярным поведением, допускающим иногда возможность очень больших значений?

Поскольку, как правило, при моделировании реальных явлений нет веских причин отвергать предположение о конечности дисперсии скачков, вторая комбинация как минимум заслуживает внимательного изучения.

Оказывается, что распределения Линника допускают представление в виде масштабных смесей нормальных законов. Это означает, что они могут быть предельными в аналогах центральной предельной теоремы для случайных сумм независимых с.в. с *конечными* дисперсиями [3, 4], что открывает пути альтернативных несимметричных обобщений этих распределений, которым и посвящена данная статья.

2 Вспомогательные сведения

В дальнейшем иногда будет удобнее вести изложение не в терминах распределений, а в терминах с.в., предполагая, что все они заданы на одном вероятностном пространстве $(\Omega, \mathfrak{A}, P)$.

Случайная величина со стандартной показательной функцией распределения (ф.р.) будет обозначаться $W_1: P(W_1 < x) = [1 - e^{-x}] \mathbf{1}(x \geq 0)$ (здесь и далее символ $\mathbf{1}(C)$ обозначает индикатор множества C). Случайная величина со стандартной нормальной ф.р. $\Phi(x)$ будет обозначаться X ,

$$P(X < x) = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-z^2/2} dz, \quad x \in \mathbb{R}.$$

Функция распределения и плотность строго устойчивого распределения с характеристическим пока-

зателем α и параметром формы θ , определяемого характеристической функцией (х.ф.)

$$f_{\alpha,\theta}(t) = \exp \left\{ -|t|^\alpha \exp \left\{ -\frac{1}{2} i\pi\theta\alpha \operatorname{sign} t \right\} \right\}, \quad t \in \mathbb{R},$$

где $0 < \alpha \leq 2$, $|\theta| \leq \min\{1, (2/\alpha) - 1\}$, будут соответственно обозначаться $G_{\alpha,\theta}(x)$ и $g_{\alpha,\theta}(x)$ (см., например, [5]). Любую с.в. с ф.р. $G_{\alpha,\theta}(x)$ будем обозначать $S_{\alpha,\theta}$. Симметричным строго устойчивым распределениям соответствует значение $\theta = 0$ и х.ф.

$$f_{\alpha,0}(t) = e^{-|t|^\alpha}, \quad t \in \mathbb{R}. \quad (1)$$

Отсюда несложно видеть, что $S_{2,0} \stackrel{d}{=} \sqrt{2}X$.

Односторонним строго устойчивым законам, сосредоточенным на неотрицательной полуоси, соответствуют значения $\theta = 1$ и $0 < \alpha \leq 1$. Пары $\alpha = 1, \theta = \pm 1$ отвечают распределениям, вырожденным в ± 1 соответственно. Остальные устойчивые распределения абсолютно непрерывны. Явные выражения устойчивых плотностей в терминах элементарных функций отсутствуют за четырьмя исключениями (нормальный закон ($\alpha = 2, \theta = 0$), распределение Коши ($\alpha = 1, \theta = 0$), распределение Леви ($\alpha = 1/2, \theta = 1$) и распределение, симметричное к распределению Леви ($\alpha = 1/2, \theta = -1$)). Выражения устойчивых плотностей в терминах функций Фокса (обобщенных G -функций Мейера) можно найти в [6, 7].

Хорошо известно, что если $0 < \alpha < 2$, то $E|S_{\alpha,\theta}|^\beta < \infty$ для любого $\beta \in (0, \alpha)$, при этом моменты с.в. $S_{\alpha,\theta}$ порядков $\beta > \alpha$ не существуют (см., например, [5]). Несмотря на отсутствие явных выражений плотностей устойчивых распределений в терминах элементарных функций, можно показать [8], что для $0 < \beta < \alpha < 2$

$$E|S_{\alpha,0}|^\beta = \frac{2^\beta}{\sqrt{\pi}} \frac{\Gamma((\beta+1)/2)\Gamma(1-\beta/\alpha)}{\Gamma(2/\beta-1)}$$

и для $0 < \beta < \alpha \leq 1$

$$ES_{\alpha,1}^\beta = \frac{\Gamma(1-\beta/\alpha)}{\Gamma(1-\beta)}.$$

Символы $\stackrel{d}{=}$ и \implies будут соответственно обозначать совпадение распределений и сходимость по распределению.

Говорят, что распределение с.в. Y принадлежит к области нормального притяжения строго устойчивого закона $G_{\alpha,\theta}, \mathcal{L}(Y) \in \text{DNA}(G_{\alpha,\theta})$, если существует конечная положительная постоянная c такая, что

$$\frac{c}{n^{1/\alpha}} \sum_{j=1}^n X_j \implies S_{\alpha,\theta} \quad (n \rightarrow \infty),$$

где X_1, X_2, \dots — независимые копии с.в. Y . В дальнейшем будем рассматривать случай стандартного масштаба и полагаем $c = 1$. В работе [9] было показано, что если $\mathcal{L}(Y) \in \text{DNA}(G_{\alpha, \theta})$, то $E|Y|^\beta = \infty$ для любого $\beta > \alpha$.

Распределение H с.в. Q называется геометрически устойчивым, если оно является слабым пределом геометрических случайных сумм независимых одинаково распределенных с.в., а именно: если существует последовательность независимых одинаково распределенных с.в. X_1, X_2, \dots и с.в. V_p , имеющая геометрическое распределение

$$P(V_p = n) = p(1-p)^{n-1}, \quad n = 1, 2, \dots, \quad p \in (0, 1),$$

при каждом $p \in (0, 1)$ независимая от X_1, X_2, \dots , и положительные константы $a_p > 0$ такие, что

$$a_p (X_1 + \dots + X_{V_p}) \implies Q$$

при $p \rightarrow 0$ (см., например, [1, 10–12]). В работе [10] (также см., например, [13, 14]) показано, что распределение H является геометрически устойчивым тогда и только тогда, когда соответствующая ему х.ф. $h(t)$ допускает представление

$$h(t) = (1 - \log f_{\alpha, \theta}(t))^{-1}, \quad t \in \mathbb{R}, \quad (2)$$

при некоторых $\alpha \in (0, 2]$ и $\theta \in [-\min\{1, 2/\alpha - 1\}, \min\{1, 2/\alpha - 1\}]$.

Некоторые результаты данной работы будут существенно опираться на следующее вспомогательное утверждение. Рассмотрим последовательность с.в. Y_1, Y_2, \dots . Пусть N_1, N_2, \dots — натуральнозначные с.в. такие, что при каждом n с.в. N_n независима от последовательности Y_1, Y_2, \dots . Всюду далее сходимость подразумевается при $n \rightarrow \infty$.

Лемма 1. *Предположим, что существуют неограниченно возрастающая (убывающая к нулю) последовательность положительных чисел $\{b_n\}_{n \geq 1}$ и с.в. Y такие, что*

$$\frac{Y_n}{b_n} \implies Y.$$

Если существуют неограниченно возрастающая (убывающая к нулю) последовательность положительных чисел $\{d_n\}_{n \geq 1}$ и с.в. V такие, что

$$\frac{b_{N_n}}{d_n} \implies V, \quad (3)$$

то

$$\frac{Y_{N_n}}{d_n} \implies YV, \quad (4)$$

причем случайные сомножители в правой части (4) независимы. Если дополнительно $N_n \rightarrow \infty$ по вероятности и семейство масштабных смесей ф.р. с.в. Y

идентифицируемо, то условие (3) не только достаточно для (4), но и необходимо.

Доказательство см. в [15] (случай $b_n, d_n \rightarrow \infty$), [16] (случай $b_n, d_n \rightarrow 0$) или [17], теорема 3.5.5.

3 Распределения Линника

Распределения с х.ф.

$$f_\alpha^L(t) = (1 + |t|^\alpha)^{-1}, \quad t \in \mathbb{R},$$

где $0 < \alpha \leq 2$, принято называть *распределениями Линника* (в работе [18] предложено альтернативное менее употребительное название *α -Laplace distribution*). Они были введены Ю. В. Линником в 1953 г. [19]. При $\alpha = 2$ распределение Линника превращается в распределение Лапласа, соответствующее плотности

$$f^\Lambda(x) = \frac{1}{2} e^{-|x|}, \quad x \in \mathbb{R}. \quad (5)$$

Лапласовская с.в. с плотностью (5) и ее ф.р. будут соответственно обозначаться Λ и $F^\Lambda(x)$.

Случайная величина, имеющая распределение Линника с параметром α , ее ф.р. и плотность будут соответственно обозначаться L_α , F_α^L и f_α^L . При этом $F_2^L(x) \equiv F^\Lambda(x)$, $x \in \mathbb{R}$.

Относительно недавно эти распределения и их обобщения вновь привлекли внимание исследователей как вполне адекватные модели многих реальных явлений. Распределения Линника обладают многими интересными свойствами. Лаха [20] (также см. [21]) доказал унимодальность распределений Линника. В работах [22, 23] исследованы свойства плотности f_α^L . Показано, что для f_α^L справедливо интегральное представление

$$f_\alpha^L = \frac{\sin(\pi\alpha/2)}{\pi} \int_0^\infty \frac{z^\alpha e^{-z|x|} dz}{1 + z^{2\alpha} + 2 \cos(\pi\alpha/2)}, \quad x \in \mathbb{R}. \quad (6)$$

Сабу и Пиллаи [24] получили представление плотности f_α^L в терминах обобщенных G -функций Мейера. Лин [25] доказал саморазложимость F_α^L . Существование моментов с.в. L_α обсуждается в работе [26]. Абсолютные моменты порядков $\beta < \alpha$ с.в. L_α имеют вид:

$$E|L_\alpha|^\beta = \frac{2^\beta}{\sqrt{\pi}} \frac{\Gamma(1 + \beta/\alpha) \Gamma((1 + \beta)/2) \Gamma(1 - \beta/\alpha)}{\Gamma(1 - \beta/2)}.$$

Распределения Линника безгранично делимы [27], имеют бесконечный пик плотности в нуле при $\alpha \leq 1$ [27]. В работе [28] показано, что при $0 < \alpha < 2$

хвосты распределения Лапласа убывают степенным образом:

$$\lim_{x \rightarrow \infty} x^\alpha [1 - F_\alpha^L(x)] = \frac{\Gamma(\alpha)}{\pi} \sin \frac{\pi\alpha}{2}.$$

Из представлений (1) и (3) вытекает, что *распределения Линника и только они являются симметричными геометрически устойчивыми законами*.

В работах [3, 4, 29, 30] получены разнообразные представления распределений Линника в виде смесей. Некоторые из этих представлений будут приведены и использованы ниже. Другие аналитические и асимптотические свойства распределения Линника рассмотрены в [22, 23].

4 Распределения Миттаг–Леффлера

Пусть $\alpha \in (0, 1)$ и M_α — неотрицательная с.в. с преобразованием Лапласа–Стилтьеса (п. Л.–С.)

$$E \exp\{-sM_{\alpha 1}\} = (1 + s^\alpha)^{-1}, \quad s \geq 0. \quad (7)$$

Распределения с п. Л.–С. (7) принято называть *распределениями Миттаг–Леффлера*. Происхождение этого названия связано с тем, что плотность, соответствующая п. Л.–С. (7), имеет вид:

$$f_\alpha^M(x) = \frac{1}{x^{1-\alpha}} \sum_{n=0}^{\infty} \frac{(-1)^n x^{\alpha n}}{\Gamma(\alpha n + 1)} = -\frac{d}{dx} E_\alpha(-x^\alpha), \quad x \geq 0, \quad (8)$$

где $E_\alpha(z)$ — функция Миттаг–Леффлера индекса α , определяемая как степенной ряд

$$E_\alpha(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(\alpha n + 1)}, \quad \alpha > 0, \quad z \in \mathbb{Z}.$$

Функция распределения, соответствующая плотности (8), будет обозначаться $F_\alpha^M(x)$.

Для ф.р. $F_\alpha^M(x)$ при $x > 0$ справедливо интегральное представление:

$$\begin{aligned} F_\alpha^M(x) &= 1 - \int_0^\infty e^{-xz} f_{\alpha,1}^Q(z) dz = \\ &= 1 - \frac{\sin(\pi\alpha)}{\pi} \int_0^\infty \frac{z^{\alpha-1} e^{-zx} dz}{1 + z^{2\alpha} + 2z^\alpha \cos(\pi\alpha)}. \quad (9) \end{aligned}$$

При $\alpha = 1$ распределение Миттаг–Леффлера превращается в стандартное показательное распределение: $M_1 \stackrel{d}{=} W_1$. Но при $\alpha < 1$ плотность (8)

имеет хвост, убывающий степенным образом: если $0 < \alpha < 1$, то

$$\lim_{x \rightarrow \infty} x^{\alpha+1} f_\alpha^M(x) = \frac{\Gamma(\alpha+1)}{\pi} \sin \pi\alpha$$

(см., например, [31]).

Моменты с.в. с распределением Миттаг–Леффлера порядков $\beta \geq \alpha$ бесконечны, но при $0 < \beta < \alpha < 1$

$$EM_\alpha^\beta = \Gamma\left(1 + \frac{\beta}{\alpha}\right) \Gamma\left(1 - \frac{\beta}{\alpha}\right).$$

Распределение Миттаг–Леффлера геометрически устойчиво. Еще в 1965 г. И. Н. Коваленко [32] показал, что распределения с п. Л.–С. (7) и только они являются возможными предельными распределениями для надлежащим образом нормированных геометрических сумм вида $a_p(X_1 + \dots + X_{V_p})$ независимых неотрицательных с.в. при $p \rightarrow 0$. Доказательства этого результата были воспроизведены в книгах [14, 33, 34], где вместо термина «распределения Миттаг–Леффлера» класс распределений с п. Л.–С. (7) был назван *классом К* в честь И. Н. Коваленко.

Спустя 25 лет упомянутое предельное свойство распределений с п. Л.–С. (7) было переоткрыто Р. Пиллаи [35, 36], который предложил для них использовать термин *распределения Миттаг–Леффлера*, ставший общепринятым.

Распределения Миттаг–Леффлера используются при описании аномальной диффузии или эффектов релаксации (см. [37, 38] и дальнейшие ссылки в этих работах).

При каждом $\alpha \in (0, 1]$ распределение Миттаг–Леффлера является смешанным показательным распределением:

$$M_\alpha \stackrel{d}{=} W_1 \frac{S_{\alpha,1}}{S'_{\alpha,1}},$$

где $S'_{\alpha,1} \stackrel{d}{=} S_{\alpha,1}$ и все с.в. в правой части независимы. Доказательство этого факта можно найти в работах [3, 4, 29], где, в частности, показано, что плотность $p_\alpha(x)$ отношения $S_{\alpha,1}/S'_{\alpha,1}$ двух независимых односторонних строго устойчивых с.в. с характеристическим показателем $\alpha \in (0, 1)$ имеет вид:

$$p_\alpha(x) = \frac{\sin(\pi\alpha)x^{\alpha-1}}{\pi[1 + x^{2\alpha} + 2x^\alpha \cos(\pi\alpha)]}, \quad x > 0. \quad (10)$$

Для дальнейшего важно подчеркнуть, что *распределения Миттаг–Леффлера и только они являются геометрически устойчивыми законами, сосредоточенными на неотрицательной полуоси*.

5 Традиционный подход к определению несимметричных распределений Линника

Канонический вид (3) х.ф. геометрически устойчивого распределения обусловлен определением последнего и теоремой переноса Гнеденко–Фахима: пусть $\{X_{n,j}\}_{j \in \mathbb{N}}, n \in \mathbb{N}$, — последовательность серий независимых и одинаково в каждой серии распределенных с.в., $\{N_n\}_{n \in \mathbb{N}}$ — последовательность неотрицательных целочисленных с.в., при каждом $n \geq 1$ независимых от $X_{n,1}, X_{n,2}, \dots$. Для $k \in \mathbb{N} \cup \{0\}$ обозначим $S_{n,k} = X_{n,1} + \dots + X_{n,k}$ ($S_{n,0} = 0$). Согласно теореме переноса Гнеденко–Фахима [39] (также см., например, [14]), если существует последовательность $\{k_n\}_{n \in \mathbb{N}}$ натуральных чисел и с.в. Y и N такие, что при $n \rightarrow \infty$ $S_{n,k_n} \Rightarrow Y$ и

$$\frac{N_n}{k_n} \Rightarrow N, \quad (11)$$

то $S_{n,N_n} \Rightarrow Z$, где Z — с.в. с х.ф.

$$f(t) = \int_0^\infty (h(t))^u dA(u), \quad t \in \mathbb{R}.$$

Здесь $h(t)$ — х.ф. с.в. Y ; $A(u) = P(N < u)$.

Если п. Л.–С. Ee^{-sN} с.в. N обозначить $\psi_N(s)$, $s \geq 0$, то х.ф. $f(t)$ можно записать в виде:

$$f(t) = \psi_N(-\log h(t)), \quad t \in \mathbb{R}.$$

При $k_n = n$, $N_n \stackrel{d}{=} V_{1/n}$ в силу теоремы Реньи [40] в (11) имеем $N \stackrel{d}{=} W_1$. При этом $\psi_N(s) = (1+s)^{-1}$, так что в таком случае

$$f(t) = \psi(-\log h(t)) = (1 - \log h(t))^{-1}, \quad t \in \mathbb{R}.$$

Если, более того, $X_{n,j} \stackrel{d}{=} n^{-1/\alpha} X_j$ для всех $n, j \in \mathbb{N}$, где X_1, X_2, \dots — независимые одинаково распределенные с.в. с $\mathcal{L}(X_1) \in \text{DNA}(G_{\alpha,\theta})$ при некоторых допустимых значениях α и θ , то х.ф. $f(t)$, предельная для геометрических случайных сумм, имеет вид (3). Поэтому базирующаяся на схеме *геометрического* случайного суммирования и традиционном представлении о распределениях Линника как симметричных *геометрически устойчивых* законах идея несимметричного их обобщения заключается в замене $g_{\alpha,0}(t) = e^{-|t|^\alpha}$ в (3) на $g_{\alpha,\theta}(t)$ с произвольным допустимым θ , что приводит к распределениям с х.ф. вида:

$$\tilde{f}_\alpha^L(t) = \left(1 + |t|^\alpha \exp \left\{ -\frac{1}{2} i\pi\theta\alpha \operatorname{sign} t \right\} \right)^{-1}, \quad t \in \mathbb{R}$$

(см., например, [22, 23]). Такие распределения будем называть *несимметричными распределениями Линника первого рода*. Более-менее полная библиография по этой теме приведена в [41].

Другими словами, традиционно несимметричное обобщение распределений Линника достигается за счет того, что в *геометрической* случайной сумме рассматриваются скошенные слагаемые, распределения которых принадлежат области нормального притяжения несимметричного устойчивого закона.

6 Представление распределений Линника в виде масштабных смесей нормальных или лапласовых законов

Распределения Линника допускают представление в виде масштабных смесей нормальных законов. Это означает, что они могут быть предельными в аналогах центральной предельной теоремы для случайных сумм независимых с.в. с *конечными* дисперсиями [3, 4], что открывает пути альтернативных несимметричных обобщений этих распределений.

В работах [4, 29] установлена интересная связь между распределениями Линника, Лапласа и Миттаг–Леффлера и показано, что

$$L_\alpha \stackrel{d}{=} X \sqrt{2M_{\alpha/2}} \stackrel{d}{=} \Lambda \sqrt{\frac{S_{\alpha/2,1}}{S'_{\alpha/2,1}}}, \quad (12)$$

где все сомножители независимы и $S'_{\alpha/2,1} \stackrel{d}{=} S_{\alpha/2,1}$.

Как уже отмечалось, все распределения Линника и только они являются симметричными геометрически устойчивыми законами. Все распределения Миттаг–Леффлера и только они являются односторонними геометрически устойчивыми законами. С учетом упоминавшегося выше соотношения $S_{2,0} \stackrel{d}{=} \sqrt{2}X$ левое равенство (12) можно переписать в виде $L_\alpha \stackrel{d}{=} S_{2,0} \sqrt{M_{\alpha/2}}$, являющемся частным случаем более общего утверждения, доказанного в работе [4]: если $\alpha \in (0, 2]$ и $\alpha' \in (0, 1]$, то справедливо соотношение

$$L_{\alpha\alpha'} \stackrel{d}{=} S_{\alpha,0} M_{\alpha'}^{1/\alpha},$$

представляющее собой аналог «теоремы умножения» устойчивых с.в. (см., например, теорему 3.3.1. в [5]), в классе геометрически устойчивых распределений. Это интересная иллюстрация изоморфизма

класса геометрически устойчивых распределений классу устойчивых распределений, установленного в работе [10].

Представление (12) означает, что схема геометрического случайного суммирования отнюдь не исчерпывает все возможные предельные постановки задач для случайных сумм и других последовательностей с независимыми случайными индексами, в которых распределение Линника может выступать в качестве предельного. Соответствующие примеры предельных теорем приведены в [4].

7 Несимметричные распределения Линника как масштабные смеси несимметричных распределений Лапласа

В этом разделе будет реализован формальный подход к несимметричному обобщению распределения Линника, для чего будет использовано правое равенство (12). В результате будет получено несимметричное распределение, каждая ветвь которого (положительная и отрицательная) будут копиями соответствующих ветвей *разных* распределений Лапласа.

Пусть a_1 и a_2 — два положительных числа. Будем говорить, что с.в. Λ_{a_1, a_2} имеет *несимметричное распределение Лапласа с параметрами a_1 и a_2* , если ее ф.р. имеет вид:

$$F_{a_1, a_2}^\Lambda(x) \equiv P(\Lambda_{a_1, a_2} < x) = \begin{cases} \frac{a_1}{a_1 + a_2} e^{-a_2|x|}, & x \leq 0; \\ 1 - \frac{a_2}{a_1 + a_2} e^{-a_1x}, & x > 0. \end{cases}$$

Несложно видеть, что плотность $f_{a_1, a_2}^\Lambda(x)$, соответствующая ф.р. $F_{a_1, a_2}^\Lambda(x)$, имеет вид:

$$f_{a_1, a_2}^\Lambda(x) = \begin{cases} \frac{a_1 a_2}{a_1 + a_2} e^{a_2 x}, & x \leq 0; \\ \frac{a_1 a_2}{a_1 + a_2} e^{-a_1 x}, & x > 0. \end{cases}$$

Несимметричное распределение Лапласа является популярной моделью, широко используемой в разных областях (см., например, [2]). Следующая лемма, доказательство которой можно найти, например, в [42], утверждает, что это распределение является специальной дисперсионно-сдвиговой смесью нормальных законов.

Лемма 2. Пусть $\mu \in \mathbb{R}$, $\sigma^2 \in (0, \infty)$, $\lambda \in (0, \infty)$. Предположим, что с.в. Y допускает представление

$$Y \stackrel{d}{=} \frac{\sigma}{\sqrt{\lambda}} X \sqrt{W_1} + \mu \frac{W_1}{\lambda},$$

где с.в. X имеет стандартное нормальное распределение, с.в. W_1 имеет стандартное показательное распределение (т.е. с.в. W_1/λ имеет показательное распределение с параметром λ), причем с.в. X и W_1 независимы. Тогда $Y \stackrel{d}{=} \Lambda_{a_1, a_2}$, т.е.

$$P(Y < x) = E\Phi\left(\frac{\lambda x - \mu W_1}{\sigma \sqrt{\lambda W_1}}\right) = F_{a_1, a_2}^\Lambda(x), \quad x \in \mathbb{R},$$

где

$$a_1 = \frac{1}{\sqrt{\mu^2 + 2\lambda\sigma^2 + \mu}}; \quad a_2 = \frac{1}{\sqrt{\mu^2 + 2\lambda\sigma^2 - \mu}}.$$

Обозначим $Q_{\alpha/2} \stackrel{d}{=} \sqrt{S_{\alpha/2, 1}/S'_{\alpha/2, 1}}$, где с.в. $S_{\alpha/2, 1}$ и $S'_{\alpha/2, 1}$ независимы и имеют одинаковое одностороннее устойчивое распределение с характеристическим показателем $\alpha/2$. Легко видеть, что $Q_{\alpha/2} \stackrel{d}{=} Q_{\alpha/2}^{-1}$. Теперь из правого равенства (12), (10) и (6) вытекает, что для любых $\alpha \in (0, 2]$ и $y \geq 0$

$$P(\Lambda_{a_1, a_2} Q_{\alpha/2} > y) = P(\Lambda_{a_1, a_2} > y Q_{\alpha/2}) = \frac{a_2 \sin(\pi\alpha/2)}{\pi(a_1 + a_2)} \int_0^\infty \frac{z^\alpha e^{-a_1 y z} dz}{1 + z^{2\alpha} + 2z^\alpha \cos(\pi\alpha/2)} = 1 - \frac{a_2}{a_1 + a_2} F_\alpha^L(a_1 y),$$

а для $y < 0$

$$P(\Lambda_{a_1, a_2} Q_{\alpha/2} < y) = P(\Lambda_{a_1, a_2} < y Q_{\alpha/2}) = \frac{a_1 \sin(\pi\alpha/2)}{\pi(a_1 + a_2)} \int_0^\infty \frac{z^\alpha e^{-a_2 |y| z} dz}{1 + z^{2\alpha} + 2z^\alpha \cos(\pi\alpha/2)} = \frac{a_1}{a_1 + a_2} F_\alpha^L(a_2 y).$$

Таким образом, естественно получено формальное несимметричное обобщение распределения Линника.

Определение 1. Пусть a_1 и a_2 — два положительных числа, $\alpha \in (0, 2]$. Будем говорить, что с.в. $\hat{L}_{\alpha; a_1, a_2}$ имеет *несимметричное распределение Линника второго рода с параметрами α , a_1 и a_2* , если ее ф.р. имеет вид:

$$\widehat{F}_{\alpha; a_1, a_2}^L(x) \equiv P(\widehat{L}_{\alpha; a_1, a_2} < x) = \begin{cases} \frac{a_1}{a_1 + a_2} F_{\alpha}^L(a_2 x), & x \leq 0; \\ 1 - \frac{a_2}{a_1 + a_2} F_{\alpha}^L(a_1 x), & x > 0. \end{cases}$$

Необходимо особо отметить, что с формальной точки зрения несимметричное распределение Линника второго рода является специальной сдвиг-масштабной смесью нормальных законов. Более того, справедливо следующее утверждение.

Теорема 1. Пусть $\alpha \in (0, 1]$, $\mu \in \mathbb{R}$, $\sigma^2 \in (0, \infty)$, $\lambda \in (0, \infty)$. Предположим, что с.в. Z допускает представление

$$Z \stackrel{d}{=} \left(\frac{\sigma}{\sqrt{\lambda}} X \sqrt{W_1} + \frac{\mu W_1}{\lambda} \right) Q_{\alpha/2},$$

где с.в. X , W_1 , $Q_{\alpha/2}$ независимы, с.в. X имеет стандартное нормальное распределение, с.в. W_1 имеет стандартное показательное распределение (т.е. с.в. W_1/λ имеет показательное распределение с параметром λ). Тогда $Z \stackrel{d}{=} \widehat{L}_{\alpha; a_1, a_2}$, т.е.

$$P(Z < x) = E\Phi\left(\frac{\lambda x - \mu W_1 Q_{\alpha/2}}{\sigma \sqrt{\lambda W_1} Q_{\alpha/2}}\right) = P(\widehat{L}_{\alpha; a_1, a_2} < x) = \widehat{F}_{\alpha; a_1, a_2}^L(x), \quad x \in \mathbb{R},$$

где a_1 и a_2 имеют вид (7).

8 Сходимость распределений дважды случайных сумм независимых одинаково распределенных случайных величин к несимметричному распределению Линника второго рода

Приведем пример предельной схемы типа «рандомизированного» закона больших чисел для сумм независимых с.в. с конечными математическими ожиданиями, в которой в качестве предельных возникают несимметричные распределения Линника второго рода. Из-за непростой связи смешивающих с.в. в случайном сдвиге и случайном изменении масштаба в соответствующей смеси (см. теорему 1), обычной схемы случайного суммирования для этой цели недостаточно и приходится вводить

в модель дополнительный источник случайности и рассматривать так называемые *дважды случайные* суммы.

Пусть a_1 и a_2 — два конечных положительных числа. Пусть $X_1^{(1)}, X_2^{(1)}, \dots$ — независимые одинаково распределенные с.в. такие, что $EX_1^{(1)} = a_2^{-1}$, $X_1^{(2)}, X_2^{(2)}, \dots$ — независимые одинаково распределенные с.в. такие, что $EX_1^{(2)} = a_1^{-1}$. Тогда по закону больших чисел

$$\frac{1}{n} \sum_{j=1}^n X_j^{(1)} \Rightarrow \frac{1}{a_2}; \quad \frac{1}{n} \sum_{j=1}^n X_j^{(2)} \Rightarrow \frac{1}{a_1} \quad (13)$$

при $n \rightarrow \infty$.

Пусть $V_{1/n}^{(1)}$ и $V_{1/n}^{(2)}$ — с.в. с одинаковым геометрическим распределением (2) с параметром $p = 1/n$. Будем считать, что с.в. $V_{1/n}^{(1)}, V_{1/n}^{(2)}, X_1^{(1)}, X_2^{(1)}, \dots, X_1^{(2)}, X_2^{(2)}, \dots$ независимы. Для каждого $n \in \mathbb{N}$ введем с.в.

$$Y_n = \begin{cases} \sum_{j=1}^{V_{1/n}^{(1)}} X_j^{(1)} & \text{с вероятностью } \frac{a_2}{a_1 + a_2}; \\ - \sum_{j=1}^{V_{1/n}^{(2)}} X_j^{(2)} & \text{с вероятностью } \frac{a_1}{a_1 + a_2}. \end{cases}$$

Тогда по теореме Реньи из (13) вытекает, что при $n \rightarrow \infty$

$$\frac{Y_n}{n} \Rightarrow \Lambda_{a_1, a_2}. \quad (14)$$

Пусть теперь $\alpha \in (0, 2]$ и N_n — целочисленная неотрицательная с.в., независимая от последовательности Y_1, Y_2, \dots и такая, что

$$\frac{N_n}{n} \Rightarrow Q_{\alpha/2} \quad (15)$$

при $n \rightarrow \infty$. Такая с.в. может быть построена, например, следующим образом. Пусть $P(t)$, $t \geq 0$, — стандартный пуассоновский процесс (пуассоновский процесс с единичной интенсивностью), независимый от с.в. $Q_{\alpha/2}$. Положим $N_n = P(nQ_{\alpha/2})$. Несложно убедиться, что такие с.в. N_n удовлетворяют (15).

Тогда по лемме 1 из (14) и (15) вытекает, что

$$\frac{Y_{N_n}}{n} \Rightarrow \Lambda_{a_1, a_2} Q_{\alpha/2} \stackrel{d}{=} \widehat{L}_{\alpha; a_1, a_2}.$$

9 Несимметричные распределения Линника как дисперсионно-сдвиговые смеси нормальных законов

Хотя несимметричное распределение Линника второго рода $\tilde{F}_{\alpha; a_1, a_2}^L(x)$, введенное в предыдущем разделе, является специальной сдвиг-масштабной смесью нормальных законов, пример предельной схемы, скажем, для сумм случайного числа независимых с.в., в которой такое распределение возникает в качестве предельного, не так прост, поскольку параметры, по которым происходит смешивание, связаны нетривиальным образом.

Однако возможен еще один подход к определению несимметричного распределения Линника, для которого подобная предельная схема строится довольно просто. Этот подход основан на левом равенстве (12) — представлении распределения Линника в виде масштабной смеси нормальных законов, в которой смешивающим является распределение Миттаг–Леффлера.

В рамках описываемого подхода несимметричное обобщение достигается за счет рассмотрения дисперсионно-сдвиговых смесей нормальных законов вместо чисто масштабных смесей.

Вероятностные модели типа дисперсионно-сдвиговых смесей нормальных законов рассматриваются в качестве базовых во многих практических задачах.

Подобные модели уже хорошо себя зарекомендовали во многих исследованиях, где они продемонстрировали очень высокую адекватность. Последнее обстоятельство можно легко объяснить довольно большим числом настраиваемых параметров в указанных моделях. Однако на самом деле их адекватность имеет гораздо более глубокие теоретические обоснования, а именно: дисперсионно-сдвиговые смеси нормальных законов являются предельными законами в довольно простых предельных теоремах для случайно остановленных случайных блужданий.

Такие теоремы позволяют однозначно связать конкретный смешивающий закон в дисперсионно-сдвиговых смесях с поведением интенсивности потока информативных событий, в результате которых накапливаются данные, характеризующие анализируемый случайный процесс. Тем самым эти теоремы как бы позволяют разделить вклады внешних и внутренних факторов в случайность поведения анализируемого процесса.

Понятие дисперсионно-сдвиговой смеси нормальных законов (normal variance-mean mixture) введено в 1970–1980-х гг. в работах О.-Е. Барндорфф-Нильсена и его коллег [43–45] как довольно гибкое обобщение нормального распределения.

Пусть $\beta \in \mathbb{R}$, $\alpha \in \mathbb{R}$, $0 < \sigma < \infty$, $A(x)$ — функция распределения, все точки роста которой сосредоточены на \mathbb{R}_+ . Дисперсионно-сдвиговой смесью нормальных законов называется ф.р.

$$F(x) = \int_0^\infty \Phi\left(\frac{x - \beta - \alpha z}{\sigma\sqrt{z}}\right) dA(z), \quad x \in \mathbb{R}. \quad (16)$$

Обратим внимание, что в соотношении (16) смешивание происходит одновременно и по параметру сдвига, и по параметру масштаба, но так как эти параметры в (16) связаны жесткой зависимостью, при которой параметры положения (сдвига) смешиваемых нормальных законов пропорциональны их дисперсиям, то фактически смесь (16) является однопараметрической. Именно поэтому смеси вида (16) называются *дисперсионно-сдвиговыми*.

Класс дисперсионно-сдвиговых смесей нормальных законов обширен и содержит, в частности, обобщенные гиперболические законы [43–45] и обобщенные дисперсионные гамма-распределения [46, 47], демонстрирующие отличное согласие со статистическими данными из самых разных областей — от атмосферной турбулентности до финансовых рынков.

Без существенного ограничения общности для простоты далее будем считать, что $\beta = 0$.

Определение 2. Пусть $\mu \in \mathbb{R}$, $\sigma > 0$, $\alpha \in (0, 2]$. Будем говорить, что с.в. $\tilde{L}_{\alpha; \mu, \sigma}$ имеет несимметричное распределение Линника третьего рода с параметрами α , μ и σ , если ее ф.р. $\tilde{F}_{\alpha; \mu, \sigma}^L(x)$ является дисперсионно-сдвиговой смесью нормальных законов вида:

$$\tilde{F}_{\alpha; \mu, \sigma}^L(x) = \int_0^\infty \Phi\left(\frac{x - \mu z}{\sigma\sqrt{z}}\right) dF_{\alpha/2}^M(z), \quad x \in \mathbb{R},$$

где смешивающая ф.р. Миттаг–Леффлера $F_{\alpha/2}^M$ имеет вид (9).

Из представления (12) вытекает, что при $\mu = 0$ несимметричное распределение Линника третьего рода превращается в обычное симметричное распределение Линника.

10 Сходимость распределений случайных сумм независимых одинаково распределенных случайных величин к несимметричному распределению Линника третьего рода

Пусть $\{X_{n,j}\}_{j \geq 1}$, $n = 1, 2, \dots$, — последовательность серий одинаково в каждой серии распределенных с.в. Пусть $\{N_n\}_{n \geq 1}$ — последовательность неотрицательных целочисленных с.в. таких, что при каждом $n \geq 1$ с.в. $N_n, X_{n,1}, X_{n,2}, \dots$ независимы. Напомним, что используется обозначение $S_{n,k} = X_{n,1} + \dots + X_{n,k}$, $n, k \in \mathbb{N}$. В работе [48] доказано следующее утверждение.

Лемма 3. *Предположим, что существуют последовательность натуральных чисел $\{k_n\}_{n \geq 1}$ и числа $\mu \in \mathbb{R}$ и $\sigma \in (0, \infty)$ такие, что*

$$P(S_{n,k_n} < x) \implies \Phi\left(\frac{x - \mu}{\sigma}\right). \quad (17)$$

Предположим, что $N_n \rightarrow \infty$ по вероятности. Тогда распределения случайных сумм S_{N_n} независимых одинаково распределенных с.в. слабо сходятся к некоторой ф.р. $F(x)$:

$$P(S_{n,N_n} < x) \implies F(x),$$

если и только если существует ф.р. $H(x)$ такая, что $H(0) = 0$,

$$F(x) = \int_0^\infty \Phi\left(\frac{x - \mu z}{\sigma \sqrt{z}}\right) dH(z);$$

$$P(N_n < x k_n) \implies H(x).$$

Из леммы 3 и определения 2 непосредственно вытекает следующее утверждение, устанавливающее необходимые и достаточные условия сходимости распределений случайных сумм независимых одинаково распределенных с.в. с конечными дисперсиями к несимметричному распределению Линника третьего рода.

Теорема 2. *Предположим, что существуют последовательность натуральных чисел $\{k_n\}_{n \geq 1}$ и числа $\mu \in \mathbb{R}$ и $\sigma > 0$ такие, что имеет место сходимость (17). Предположим, что $N_n \rightarrow \infty$ по вероятности. Тогда распределения случайных сумм S_{N_n} независимых одинаково распределенных с.в. слабо сходятся к несимметричному распределению Линника третьего рода $\tilde{F}_{\alpha;\mu,\sigma}^L$ при некотором $\alpha \in (0, 2]$:*

$$P(S_{n,N_n} < x) \implies \tilde{F}_{\alpha;\mu,\sigma}^L(x),$$

если и только если

$$P(N_n < x k_n) \implies F_{\alpha/2}^M(x),$$

где ф.р. Миттаг–Леффлера $F_{\alpha/2}^M$ имеет вид (9).

Примеры индексов N_n , удовлетворяющих условию теоремы 2, приведены в [4].

Литература

1. Mittnik S., Rachev S. Modeling asset returns with alternative stable models // *Economet. Rev.*, 1993. Vol. 12. P. 261–330.
2. Kotz S., Kozubowski T. J., Podgorski K. The Laplace distribution and generalizations: A revisit with applications to communications, economics, engineering, and finance. — Boston: Birkhauser, 2001. 349 p.
3. Korolev V. Yu., Zeifman A. I. A note on mixture representations for the Linnik and Mittag–Leffler distributions and their applications // *J. Math. Sci.*, 2016. Vol. 218. P. 314–327.
4. Korolev V. Yu., Zeifman A. I. Convergence of random sums and statistics constructed from samples with random sizes to the Linnik and Mittag–Leffler distributions and their generalizations // *J. Korean Stat. Soc.*, 2016. arXiv:1602.02480v1.
5. Золотарев В. М. Одномерные устойчивые распределения. — М.: Наука, 1983. 304 с.
6. Schneider W. R. Stable distributions: Fox function representation and generalization // *Stochastic processes in classical and quantum systems* / Eds. S. Albeverio, G. Casati, D. Merlini. — Berlin: Springer, 1986. P. 497–511.
7. Uchaikin V. V., Zolotarev V. M. Chance and stability. — Utrecht: VSP, 1999. 570 p.
8. Korolev V. Yu. Product representations for random variables with the Weibull distributions and their applications // *J. Math. Sci.*, 2016. Vol. 218. No. 3. P. 298–313.
9. Tucker H. On moments of distribution functions attracted to stable laws // *Houston J. Math.*, 1975. Vol. 1. No. 1. P. 149–152.
10. Клебанов Л. Б., Маня Г. М., Меламед И. А. Одна задача В. М. Золотарева и аналоги безгранично делимых и устойчивых распределений в схеме суммирования случайного числа случайных величин // *Теория вероятностей и ее применения*, 1984. Т. 29. Вып. 4. С. 791–794.
11. Klebanov L. B., Rachev S. T. Sums of a random number of random variables and their approximations with ε -accompanying infinitely divisible laws // *Serdica*, 1996. Vol. 22. P. 471–498.
12. Bunge J. Compositions semigroups and random stability // *Ann. Probab.*, 1996. Vol. 24. P. 1476–1489.
13. Rachev S. T. Probability metrics and the stability of stochastic models. — Chichester–New York: Wiley, 1991. 494 p.

14. Gnedenko B. V., Korolev V. Yu. Random summation: Limit theorems and applications. — Boca Raton: CRC Press, 1996. 267 p.
15. Королев В. Ю. Сходимость случайных последовательностей с независимыми случайными индексами. I // Теория вероятностей и ее применения, 1994. Т. 39. Вып. 2. С. 313–333.
16. Королев В. Ю. Сходимость случайных последовательностей с независимыми случайными индексами. II // Теория вероятностей и ее применения, 1995. Т. 40. Вып. 4. С. 907–910.
17. Bening V. E., Korolev V. Yu. Generalized Poisson models and their applications in insurance and finance. — Utrecht: VSP, 2002. 434 p.
18. Pillai R. N. Semi- α -Laplace distributions // Commun. Stat. Theor. Meth., 1985. Vol. 14. P. 991–1000.
19. Линник Ю. В. Линейные формы и статистические критерии. I, II // Украинский математический журнал, 1953. Т. 5. Вып. 2. С. 207–243; Вып. 3. С. 247–290.
20. Laha R. G. On a class of unimodal distributions // Proc. Am. Math. Soc., 1961. Vol. 12. P. 181–184.
21. Лукач Е. Характеристические функции. — М.: Наука, 1979. 424 с.
22. Kotz S., Ostrovskii I. V., Hayfavi A. Analytic and asymptotic properties of Linnik's probability densities, I // J. Math. Anal. Appl., 1995. Vol. 193. P. 353–371.
23. Kotz S., Ostrovskii I. V., Hayfavi A. Analytic and asymptotic properties of Linnik's probability densities, II // J. Math. Anal. Appl., 1995. Vol. 193. P. 497–521.
24. Sabu G., Pillai R. N. Multivariate α -Laplace distributions // J. Nat. Acad. Math., 1987. Vol. 5. P. 13–18.
25. Lin G. D. Characterizations of the Laplace and related distributions via geometric compound // Sankhya, A1, 1994. Vol. 56. P. 1–9.
26. Anderson D. N. A multivariate Linnik distribution // Stat. Probabil. Lett., 1992. Vol. 14. P. 333–336.
27. Devroye L. A note on Linnik's distribution // Stat. Probabil. Lett., 1990. Vol. 9. P. 305–306.
28. Jacques C., Rémillard B., Theodorescu R. Estimation of Linnik law parameters // Stat. Decision, 1999. Vol. 17. No. 3. P. 213–236.
29. Kotz S., Ostrovskii I. V. A mixture representation of the Linnik distribution // Stat. Probabil. Lett., 1996. Vol. 26. P. 61–64.
30. Pakes A. G. Mixture representations for symmetric generalized Linnik laws // Stat. Probabil. Lett., 1998. Vol. 37. P. 213–221.
31. Gorenflo R., Kilbas A. A., Mainardi F., Rogosin S. V. Mittag-Leffler functions, related topics and applications. — Berlin – New York: Springer, 2014. 420 p.
32. Kovalenko I. N. On the class of limit distributions for rarefied flows of homogeneous events // Lith. Math. J., 1965. Vol. 5. No. 4. P. 569–573.
33. Gnedenko B. V., Kovalenko I. N. Introduction to queueing theory. — Jerusalem: Israel Program for Scientific Translations, 1968. 281 p.
34. Gnedenko B. V., Kovalenko I. N. Introduction to queueing theory. — 2nd ed. — Boston: Birkhauser, 1989. 314 p.
35. Pillai R. N. Harmonic mixtures and geometric infinite divisibility // J. Indian Stat. Ass., 1990. Vol. 28. P. 87–98.
36. Pillai R. N. On Mittag-Leffler functions and related distributions // Ann. Inst. Stat. Math., 1990. Vol. 42. P. 157–161.
37. Weron K., Kotulski M. On the Cole–Cole relaxation function and related Mittag-Leffler distributions // Physica A, 1996. Vol. 232. P. 180–188.
38. Gorenflo R., Mainardi F. Continuous time random walk, Mittag-Leffler waiting time and fractional diffusion: Mathematical aspects // Anomalous transport: Foundations and applications / Eds. R. Klages, G. Radons, I. M. Sokolov. — Weinheim, Germany: Wiley-VCH, 2008. P. 93–127. <http://arxiv.org/abs/0705.0797>.
39. Гнеденко Б. В., Фахим Х. Об одной теореме переноса // Докл. АН СССР, 1969. Т. 187. Вып. 1. С. 15–17.
40. Rényi A. A Poisson-folyamat egy jellemzése // Magyar Tud. Acad. Mat. Int. Kozl., 1956. Vol. 1. P. 519–527.
41. Lim S. C., Teo L. P. Analytic and asymptotic properties of multivariate generalized Linnik's probability densities // J. Fourier Anal. Appl., 2010. Vol. 16. P. 715–747.
42. Korolev V. Yu., Kurmangazjeva L., Zeifman A. I. On asymmetric generalization of the Weibull distribution by scale-location mixing of normal laws // J. Korean Stat. Soc., 2016. Vol. 45. P. 238–249. [arXiv:1506.06232](https://arxiv.org/abs/1506.06232).
43. Barndorff-Nielsen O.-E. Exponentially decreasing distributions for the logarithm of particle size // Proc. Roy. Soc. Lond. A, 1977. Vol. A(353). P. 401–419.
44. Barndorff-Nielsen O.-E., Kent J., Sørensen M. Normal variance-mean mixtures and z -distributions // Int. Stat. Rev., 1977. Vol. 50. No. 2. P. 145–159.
45. Barndorff-Nielsen O.-E. Hyperbolic distributions and distributions of hyperbolae // Scand. J. Stat., 1978. Vol. 5. P. 151–157.
46. Королев В. Ю., Соколов И. А. Скошенные распределения Стьюдента, дисперсионные гамма-распределения и их обобщения как асимптотические аппроксимации // Информатика и её применения, 2012. Т. 6. Вып. 1. С. 2–10.
47. Закс Л. М., Королев В. Ю. Обобщенные дисперсионные гамма-распределения как предельные для случайных сумм // Информатика и её применения, 2013. Т. 7. Вып. 1. С. 105–115.
48. Королев В. Ю. Обобщенные гиперболические распределения как предельные для случайных сумм // Теория вероятностей и ее применения, 2013. Т. 58. Вып. 1. С. 117–132.
49. Erdoĝan M. B., Ostrovskii I. V. Analytic and asymptotic properties of generalized Linnik probability densities // J. Math. Anal. Appl., 1998. Vol. 217. P. 555–578.
50. Erdoĝan M. B., Ostrovskii I. V. On mixture representation of the Linnik density // J. Aust. Math. Soc. A, 1998. Vol. 64. P. 317–326.

51. Kalashnikov V. V. Geometric sums: Bounds for rare events with applications. — Dordrecht: Kluwer Academic Publ., 1997. 270 p.
52. Pakes A. G. A characterization of gamma mixtures of stable laws motivated by limit theorems // Stat. Neerl., 1992. Vol. 2-3. P. 209–218.
53. Kozubowski T. J. Mixture representation of Linnik distribution revisited // Stat. Probabil. Lett., 1998. Vol. 38. P. 157–160.
54. Kozubowski T. J. Exponential mixture representation of geometric stable distributions // Ann. Inst. Stat. Math., 1999. Vol. 52. No. 2. P. 231–238.

Поступила в редакцию 14.10.16

ASYMMETRIC LINNIK DISTRIBUTIONS AS LIMIT LAWS FOR RANDOM SUMS OF INDEPENDENT RANDOM VARIABLES WITH FINITE VARIANCES

V. Yu. Korolev^{1,2}, A. I. Zeifman^{1,2,3,4}, and A. Yu. Korchagin¹

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

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

³Vologda State University, 15 Lenin Str., Vologda 160000, Russian Federation

⁴ISEDT RAS, 56-A Gorky Str., Vologda 160001, Russian Federation

Abstract: Linnik distributions (symmetric geometrically stable distributions) are widely applied in financial mathematics, telecommunication systems modeling, astrophysics, and genetics. These distributions are limiting for geometric sums of independent identically distributed random variables whose distribution belongs to the domain of normal attraction of a symmetric strictly stable distribution. In the paper, three asymmetric generalizations of the Linnik distribution are considered. The traditional (and formal) approach to the asymmetric generalization of the Linnik distribution consists in the consideration of geometric sums of random summands whose distributions are attracted to an asymmetric strictly stable distribution. The variances of such summands are infinite. Since in modeling real phenomena, as a rule, there are no solid reasons to reject the assumption of the finiteness of the variances of elementary summands, in the paper, two alternative asymmetric generalizations are proposed based on the representability of the Linnik distribution as a scale mixture of normal laws or a scale mixture of Laplace laws. Examples are presented of limit theorems for sums of a random number of independent random variables with finite variances in which the proposed asymmetric Linnik distributions appear as limit laws.

Keywords: Linnik distribution; Laplace distribution; Mittag–Leffler distribution; normal distribution; scale mixture; normal variance–mean mixture; stable distribution; geometrically stable distribution

DOI: 10.14357/19922264160403

Acknowledgments

This work was financially supported by the Russian Science Foundation (grant No. 14-11-00364).

References

- Mittnik, S., and S. Rachev. 1993. Modeling asset returns with alternative stable models. *Economet. Rev.* 12:261–330.
- Kotz, S., T. J. Kozubowski, and K. Podgorski. 2001. *The Laplace distribution and generalizations: A revisit with applications to communications, economics, engineering, and finance*. Boston: Birkhauser. 349 p.
- Korolev, V. Yu., and A. I. Zeifman. 2016. A note on mixture representations for the Linnik and Mittag–Leffler distributions and their applications. *J. Math. Sci.* 218(3):314–327.
- Korolev, V. Yu., and A. I. Zeifman. 2016. Convergence of random sums and statistics constructed from samples with random sizes to the Linnik and Mittag–Leffler distributions and their generalizations. *J. Korean Stat. Soc.* Available at: arXiv:1602.02480v1 (accessed December 10, 2016).
- Zolotarev, V. M. 1986. *One-dimensional stable distributions*. Providence: AMS. 284 p.

6. Schneider, W. R. 1986. Stable distributions: Fox function representation and generalization. *Stochastic processes in classical and quantum systems*. Eds. S. Albeverio, G. Casati, and D. Merlini. Berlin: Springer. 497–511.
7. Uchaikin, V. V., and V. M. Zolotarev. 1999. *Chance and stability*. Utrecht: VSP. 570 p.
8. Korolev, V. Yu. 2016. Product representations for random variables with the Weibull distributions and their applications. *J. Math. Sci.* 218(3):298–313.
9. Tucker, H. 1975. On moments of distribution functions attracted to stable laws. *Houston J. Math.* 1(1):149–152.
10. Klebanov, L. B., G. M. Maniya, and I. A. Melamed. 1985. A problem of Zolotarev and analogs of infinitely divisible and stable distributions in a scheme for summing a random number of random variables. *Theor. Probab. Appl.* 29(4):791–794.
11. Klebanov, L. B., and S. T. Rachev. 1996. Sums of a random number of random variables and their approximations with ε -accompanying infinitely divisible laws. *Serdica* 22:471–498.
12. Bunge, J. 1996. Compositions semigroups and random stability. *Ann. Probab.* 24:1476–1489.
13. Rachev, S. T. 1991. *Probability metrics and the stability of stochastic models*. Chichester – New York: Wiley. 494 p.
14. Gnedenko, B. V., and V. Yu. Korolev. 1886. *Random summation: Limit theorems and applications*. Boca Raton: CRC Press. 267 p.
15. Korolev, V. Yu. 1995. Convergence of random sequences with the independent random indices I. *Theor. Probab. Appl.* 39(2):282–297.
16. Korolev, V. Yu. 1996. Convergence of random sequences with the independent random indices II. *Theor. Probab. Appl.* 40(4):770–772.
17. Bening, V. E., and V. Yu. Korolev. 2002. *Generalized Poisson models and their applications in insurance and finance*. Utrecht: VSP. 434 p.
18. Pillai, R. N. 1985. Semi- α -Laplace distributions. *Commun. Stat. Theor. Meth.* 14:991–1000.
19. Linnik, Yu. V. 1953. Lineynye formy i statisticheskiye kriterii. I, II [Linear forms and statistical criteria. I, II]. *Ukr. Math. J.* 5(2):207–243; 3:247–290.
20. Laha, R. G. 1961. On a class of unimodal distributions. *Proc. Am. Math. Soc.* 12:181–184.
21. Lukacs, E. 1970. *Characteristic functions*. 2nd ed. London: Griffin. 350 p.
22. Kotz, S., I. V. Ostrovskii, and A. Hayfavi. 1995. Analytic and asymptotic properties of Linnik’s probability densities, I. *J. Math. Anal. Appl.* 193:353–371.
23. Kotz, S., I. V. Ostrovskii, and A. Hayfavi. 1995. Analytic and asymptotic properties of Linnik’s probability densities, II. *J. Math. Anal. Appl.* 193:497–521.
24. Sabu, G., and R. N. Pillai. 1987. Multivariate α -Laplace distributions. *J. Nat. Acad. Math.* 5:13–18.
25. Lin, G. D. 1994. Characterizations of the Laplace and related distributions via geometric compound. *Sankhya, AI* 56:1–9.
26. Anderson, D. N. 1992. A multivariate Linnik distribution. *Stat. Probabil. Lett.* 14:333–336.
27. Devroye, L. 1990. A note on Linnik’s distribution. *Stat. Probabil. Lett.* 9:305–306.
28. Jacques, C., B. R emillard, and R. Theodorescu. 1999. Estimation of Linnik law parameters. *Stat. Decision* 17(3):213–236.
29. Kotz, S., and I. V. Ostrovskii. 1996. A mixture representation of the Linnik distribution. *Stat. Probabil. Lett.* 26:61–64.
30. Pakes, A. G. 1998. Mixture representations for symmetric generalized Linnik laws. *Stat. Probabil. Lett.* 37:213–221.
31. Gorenflo, R., A. A. Kilbas, F. Mainardi, and S. V. Rogosin. 2014. *Mittag–Leffler functions, related topics and applications*. Berlin – New York: Springer. 420 p.
32. Kovalenko, I. N. 1965. On the class of limit distributions for rarefied flows of homogeneous events. *Lith. Math. J.* 5(4):569–573.
33. Gnedenko, B. V., and I. N. Kovalenko. 1968. *Introduction to queueing theory*. Jerusalem: Israel Program for Scientific Translations. 281 p.
34. Gnedenko, B. V., and I. N. Kovalenko. 1989. *Introduction to queueing theory*. 2nd ed. Boston: Birkhauser. 314 p.
35. Pillai, R. N. 1990. Harmonic mixtures and geometric infinite divisibility. *J. Indian Stat. Ass.* 28:87–98.
36. Pillai, R. N. 1990. On Mittag–Leffler functions and related distributions. *Ann. Inst. Stat. Math.* 42:157–161.
37. Weron, K., and M. Kotulski. 1996. On the Cole–Cole relaxation function and related Mittag–Leffler distributions. *Physica A* 232:180–188.
38. Gorenflo, R., and F. Mainardi. 2008. Continuous time random walk, Mittag–Leffler waiting time and fractional diffusion: Mathematical aspects. *Anomalous transport: Foundations and applications*. Eds. R. Klages, G. Radons, and I. M. Sokolov. Weinheim, Germany: Wiley-VCH. 93–127.
39. Gnedenko, B. V., and H. Fahim. 1969. Ob odnoy teoreme perenosa [About one transfer theorem]. *Dokl. USSR Akad. Nauk* 187(1):15–17.
40. R enyi, A. 1956. A Poisson-folyamat egy jellemzese. *Magyar Tud. Acad. Mat. Int. Kozl.* 1:519–527.
41. Lim, S. C., and L. P. Teo. 2010. Analytic and asymptotic properties of multivariate generalized Linnik’s probability densities. *J. Fourier Anal. Appl.* 16:715–747.
42. Korolev, V. Yu., L. Kurmangazieva, and A. I. Zeifman. 2016. On asymmetric generalization of the Weibull distribution by scale-location mixing of normal laws. *J. Korean Stat. Soc.* 45:238–249. arXiv:1506.06232.
43. Barndorff-Nielsen, O. E. 1977. Exponentially decreasing distributions for the logarithm of particle size. *Proc. Roy. Soc. Lond. A* 353:401–419.
44. Barndorff-Nielsen, O. E., J. Kent, and M. S orensen. 1982. Normal variance-mean mixtures and z -distributions. *Int. Stat. Rev.* 50(2):145–159.

45. Barndorff-Nielsen, O. E. 1978. Hyperbolic distributions and distributions of hyperbolae. *Scand. J. Stat.* 5:151–157.
46. Korolev, V. Yu., and I. A. Sokolov. 2012. Skoshennye raspredeleniya St'yudenta, dispersionnye gamma-raspredeleniya i ikh obobshcheniya kak asimptoticheskie approksimatsii [Skewed Student's distributions, variance gamma distributions, and their generalizations as asymptotic approximations]. *Informatika i ee Primeneniya — Inform. Appl.* 6(1):2–10.
47. Zaks, L. M., and V. Yu. Korolev. 2013. Obobshchennye dispersionnye gamma-raspredeleniya kak predel'nye dlya sluchaynykh summ [Generalized variance gamma distributions as limiting for random sums]. *Informatika i ee Primeneniya — Inform. Appl.* 7(1):105–115.
48. Korolev, V. Yu. 2014. Generalized hyperbolic laws as limit distributions for random sums. *Theor. Probab. Appl.* 58(1):63–75.
49. Erdoğan, M. B., and I. V. Ostrovskii. 1998. Analytic and asymptotic properties of generalized Linnik probability densities. *J. Math. Anal. Appl.* 217:555–578.
50. Erdoğan, M. B., and I. V. Ostrovskii. 1998. On mixture representation of the Linnik density. *J. Aust. Math. Soc. A* 64:317–326.
51. Kalashnikov, V. V. 1997. *Geometric sums: Bounds for rare events with applications*. Dordrecht: Kluwer Academic Publs. 270 p.
52. Pakes, A. G. 1992. A characterization of gamma mixtures of stable laws motivated by limit theorems. *Stat. Neerl.* 2-3:209–218.
53. Kozubowski, T. J. 1998. Mixture representation of Linnik distribution revisited. *Stat. Probabil. Lett.* 38:157–160.
54. Kozubowski, T. J. 1999. Exponential mixture representation of geometric stable distributions. *Ann. Inst. Stat. Math.* 52(2):231–238.

Received October 14, 2016

Contributors

Korolev Victor Yu. (b. 1954) — Doctor of Science in physics and mathematics, professor, Head of the Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; 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; vkorolev@cs.msu.su

Zeifman Alexander I. (b. 1954) — Doctor of Science in physics and mathematics, professor, Head of Department, Vologda State University, 15 Lenin Str., Vologda 160000, 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; principal scientist, ISEDT RAS, 56-A Gorky Str., Vologda 160001, Russian Federation; Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; a_zeifman@mail.ru

Korchagin Alexander Yu. (b. 1989) — junior scientist, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; sasha.korchagin@gmail.com

ВЫЧИСЛЕНИЕ АСИМПТОТИЧЕСКОГО ДЕФЕКТА НЕКОТОРЫХ СТАТИСТИЧЕСКИХ ПРОЦЕДУР, ОСНОВАННЫХ НА ВЫБОРКАХ СЛУЧАЙНОГО ОБЪЕМА*

В. Е. Бенинг¹

Аннотация: Рассматривается случай, когда число наблюдений случайно. Это приводит к возникновению распределений с тяжелыми хвостами и к изменению эффективности обычно используемых статистических процедур. Проведено асимптотическое сравнение статистических процедур, основанных на выборках случайного и неслучайного объема. Для этого используется понятие «асимптотический дефект», которое имеет смысл добавочного числа наблюдений, необходимого данной процедуре для достижения того же качества, что и оптимальной процедуре. С помощью этого понятия сравниваются оценки, доверительные множества и статистические критерии в случае, когда число наблюдений случайно.

Ключевые слова: доверительное множество; статистическая гипотеза; асимптотический дефект; выборка случайного объема; распределение Пуассона; биномиальное распределение

DOI: 10.14357/19922264160404

1 Введение

В работе развивается подход, предложенный в работах [1, 2]. Напомним кратко постановку задачи и основные обозначения. Рассмотрим сначала задачу статистического оценивания известной параметрической функции $g(\theta)$, зависящей от неизвестного параметра θ , и обозначим через $m(n)$ необходимое число наблюдений, которое требуется оценке $\delta_{m(n)}(X_1, \dots, X_{m(n)})$ для достижения такого же качества (например, среднеквадратичного отклонения или дисперсии), что и оценке $\delta_n^*(X_1, \dots, X_n)$, основанной на n наблюдениях X_1, \dots, X_n . Рассматривается асимптотический подход, означающий, что $n \rightarrow \infty$. Под асимптотической относительной эффективностью (АОЭ) оценки $\delta_n(X_1, \dots, X_n)$ по отношению к оценке $\delta_n^*(X_1, \dots, X_n)$ понимается предел (в случае его существования и независимости от последовательности $m(n)$) вида (см., например, [3, с. 305]):

$$e \equiv \lim_{n \rightarrow \infty} \frac{n}{m(n)}.$$

Вместо отношения необходимого числа наблюдений, естественно, можно было бы рассматривать разность вида $m(n) - n$, которая тоже имеет наглядный смысл необходимого дополнительного числа наблюдений, требующихся оценке $\delta_n(X_1, \dots, X_n)$. Однако исторически сложилось так, что многие ав-

торы сначала исследовали асимптотические свойства отношения $n/m(n)$ (возможно, в силу относительной простоты его поведения).

Впервые общее асимптотическое исследование поведения разности $m(n) - n$ было предпринято в 1970 г. Ходжесом и Леманом [4]. Они назвали разность $m(n) - n$ дефектом (deficiency) конкурирующей оценки δ_n относительно оценки δ_n^* и предложили обозначение:

$$d_n = m(n) - n. \quad (1)$$

Если предел $\lim_{n \rightarrow \infty} d_n$ существует, то он называется *асимптотическим дефектом* оценки δ_n относительно оценки δ_n^* и обозначается символом d . Часто d называют просто дефектом δ_n относительно δ_n^* . Заметим, что если АОЭ $e \neq 1$, то $d = \infty$, и этот случай малоинтересен. В работе [4] также было отмечено, что существуют статистические задачи, в которых типичным образом возникает случай $e = 1$ (см., например, книгу [5]), т. е. в этом случае понятие АОЭ не дает ответа на вопрос, какая оценка лучше, и понятие дефекта проясняет эту ситуацию, поскольку в этом случае асимптотический дефект может, в принципе, быть любым.

Обозначим функции риска оценок δ_n и δ_n^* соответственно через

$$R_n(\theta) = E_\theta (\delta_n(X_1, \dots, X_n) - g(\theta))^2; \\ R_n^*(\theta) = E_\theta (\delta_n^*(X_1, \dots, X_n) - g(\theta))^2,$$

*Работа выполнена при поддержке Российского научного фонда (проект 14-11-00364).

¹Факультет вычислительной математики и кибернетики Московского государственного университета имени М. В. Ломоносова; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, bening@yandex.ru

где $g(\theta)$ — оцениваемая функция, а θ — неизвестный параметр (произвольной природы), тогда по определению величины $d_n(\theta) \equiv d_n = m(n) - n$ для каждого n должно выполняться равенство:

$$R_n^*(\theta) = R_{m(n)}(\theta). \quad (2)$$

При решении уравнения (2) целочисленную величину $m(n)$ следует рассматривать как переменную, принимающую произвольные действительные значения. Для этого можно определить функцию риска $R_{m(n)}(\theta)$ для нецелых значений $m(n)$ по формуле:

$$R_{m(n)}(\theta) = (1 - m(n) + [m(n)]) R_{[m(n)]}(\theta) + (m(n) - [m(n)]) R_{[m(n)]+1}(\theta)$$

(см. работу [4]).

Типичным образом функции риска $R_n^*(\theta)$ и $R_n(\theta)$ не известны точно, и используются их аппроксимации. Предположим, что для функций риска $R_n^*(\theta)$ и $R_n(\theta)$ справедливы асимптотические разложения вида:

$$R_n^* = \frac{a(\theta)}{n^r} + \frac{b(\theta)}{n^{r+s}} + o(n^{-r-s}); \quad (3)$$

$$R_n = \frac{a(\theta)}{n^r} + \frac{c(\theta)}{n^{r+s}} + o(n^{-r-s}), \quad (4)$$

где $a(\theta)$, $b(\theta)$ и $c(\theta)$ — некоторые постоянные, не зависящие от n , а $r > 0$, $s > 0$ — некоторые константы, определяющие порядок убывания по n этих функций риска. Первый член в этих асимптотических разложениях одинаков, и это отражает тот факт, что АОЭ этих оценок равна единице. Из соотношений (1)–(4) легко получить, что (см. работу [4] или книгу [3, с. 310])

$$d_n(\theta) \equiv \frac{c(\theta) - b(\theta)}{ra(\theta)} n^{1-s} + O(n^{1-s}). \quad (5)$$

Таким образом, асимптотический дефект имеет вид:

$$d(\theta) \equiv d = \begin{cases} \pm\infty, & 0 < s < 1; \\ \frac{c(\theta) - b(\theta)}{ra(\theta)}, & s = 1; \\ 0, & s > 1. \end{cases} \quad (6)$$

Случай, когда выполняется равенство $s = 1$, представляется наиболее интересным, поскольку при этом асимптотический дефект конечен. Ходжес и Леман в работе [4] привели ряд простых примеров, показывающих естественность возникновения этого случая в математической статистике.

Совершенно аналогично определяется дефект и в общем случае асимптотического сравнения двух статистических процедур, соответственно с мерами

качества π_n и π_n^* . В этом случае необходимое число наблюдений k_n для первой процедуры определяется из равенства (считая k_n непрерывной переменной)

$$\pi_{k_n} = \pi_n^*,$$

а предел вида (в случае его существования)

$$d = \lim_{n \rightarrow \infty} (k_n - n)$$

называется асимптотическим дефектом первой процедуры относительно второй. Если для π_n^* и π_n выполняются формулы типа (3) и (4), то для асимптотического дефекта d справедливы соотношения типа (5) и (6).

В настоящей работе приведены примеры вычисления асимптотического дефекта в задачах доверительного оценивания и проверки статистических гипотез с помощью статистик, основанных на выборках случайного объема. Кратко рассмотрена байесовская постановка задачи статистического оценивания. Рассмотрены также задачи проверки статистических гипотез в случае выборок случайного объема. Используя асимптотический дефект, проведено асимптотическое сравнение качества этих процедур, основанных на выборках случайного и неслучайного объема.

2 Статистическое оценивание параметров распределения

Рассмотрим случайные величины (с.в.) N_1, N_2, \dots и X_1, X_2, \dots , заданные на одном и том же вероятностном пространстве $(\Omega, \mathcal{A}, \rho)$. При этом с.в. X_1, X_2, \dots, X_n имеют смысл статистических наблюдений, а с.в. N_n трактуется как случайный объем выборки, зависящий от натурального параметра $n \in \mathbb{N}$. Типичным образом будем предполагать, что

$$EN_n = n,$$

т. е. в среднем объем случайной выборки равен размеру выборки неслучайного размера. При нахождении дефектов статистических процедур, основанных на выборках случайного объема $N_{m(n)}$, и соответствующей процедуры, основанной на выборке неслучайного объема n , фактически сравнивается средний объем случайной выборки $m(n)$ и n с помощью величины $d_n = m(n) - n$ и ее предела.

Предположим, что для каждого $n \geq 1$ с.в. N_n принимает только натуральные значения (т. е. $N_n \in \mathbb{N}$) и не зависит от последовательности с.в. X_1, X_2, \dots . Всюду далее предполагается, что с.в. X_1, X_2, \dots независимы, одинаково распределены и имеют распределение, зависящее от неизвестного параметра $\theta \in \Theta$, при этом множество Θ может иметь произвольную природу.

Для каждого $n \geq 1$ обозначим через $T_n = T_n(X_1, \dots, X_n)$ некоторую статистику, т.е. действительную измеримую функцию, зависящую от наблюдений X_1, \dots, X_n . Для каждого $n \geq 1$ определим статистику T_{N_n} , зависящую от выборки случайного объема, как

$$T_{N_n}(\omega) \equiv T_{N_n(\omega)}(X_1(\omega), \dots, X_{N_n(\omega)}(\omega)), \quad \omega \in \Omega.$$

Далее под статистикой в этом разделе будем понимать оценку (см., например, книгу [6]) действительной известной функции $g(\theta)$, зависящей от неизвестного параметра $\theta \in \Theta$, и будем обозначать ее символами типа $\delta_n(X_1, \dots, X_n)$.

В этом разделе иллюстрируются следствия 4.3, 4.6, 4.7 и теорема 3.2 из работы [1] рассмотрением случая, когда исходная выборка X_1, \dots, X_n представляет собой независимые одинаково нормально распределенные с параметрами (θ, σ^2) с.в. Для удобства ссылок приведем без доказательства необходимые результаты из работы [1].

Следствие 2.1. Пусть существуют числа $a(\theta)$, $b(\theta)$ и $r > 0$, $s > 0$ такие, что

$$R_n^*(\theta) = \frac{a(\theta)}{n^r} + \frac{b(\theta)}{n^{r+s}},$$

тогда

$$R_n(\theta) = a(\theta)EN_n^{-r} + b(\theta)EN_n^{-r-s}.$$

Теорема 3.1. Пусть существуют числа $a(\theta)$, $b(\theta)$ и k_1, k_2 такие, что справедливы соотношения:

$$\begin{aligned} R_n^*(\theta) &= E_\theta (\delta_n(X_1, \dots, X_n) - g(\theta))^2 = \\ &= \frac{a(\theta)}{n} + \frac{b(\theta)}{n^2} + o(n^{-2}) \end{aligned}$$

и

$$EN_n^{-1} = \frac{1}{n} + \frac{k_1}{n^2} + o(n^{-2});$$

$$EN_n^{-2} = \frac{k_2}{n^2} + o(n^{-2}); \quad EN_n^{-3} = o(n^{-2}),$$

тогда для асимптотического дефекта оценки $\delta_{N_n}(X_1, \dots, X_{N_n})$ относительно оценки $\delta_n(X_1, \dots, X_n)$ справедливо равенство:

$$d(\theta) = \frac{k_1 a(\theta) + b(\theta)(k_2 - 1)}{a(\theta)}.$$

Следствие 3.1. Пусть в условиях теоремы 3.1 $k_2 = 1$, тогда асимптотический дефект оценки $\delta_{N_n}(X_1, \dots, X_{N_n})$ относительно оценки $\delta_n(X_1, \dots, X_n)$ равен k_1 , т.е. он не зависит от вида оценки и имеет вид:

$$d(\theta) = k_1.$$

Теорема 3.2. Пусть существуют числа $a(\theta)$, $b(\theta)$ такие, что для функции риска оценки $\delta_n(X_1, \dots, X_n)$ справедливо соотношение:

$$R_n^*(\theta) = E_\theta (\delta_n(X_1, \dots, X_n) - g(\theta))^2 = \frac{a(\theta)}{n} + \frac{b(\theta)}{n^2}.$$

Пусть случайные величины N_{ni} , $i = 1, 2$, принимают натуральные значения и не зависят от наблюдений X_1, X_2, \dots . Предположим, что для некоторых чисел k_{1i}, k_{2i} , $i = 1, 2$, справедливы равенства:

$$EN_{ni}^{-1} = \frac{1}{n} + \frac{k_{1i}}{n^2} + o(n^{-2}),$$

$$EN_{ni}^{-2} = \frac{k_{2i}}{n^2} + o(n^{-2}), \quad i = 1, 2.$$

Тогда для асимптотического дефекта оценки $\delta_n^{(2)} \equiv \delta_{N_{n2}}(X_1, \dots, X_{N_{n2}})$ относительно оценки $\delta_n^{(1)} \equiv \delta_{N_{n1}}(X_1, \dots, X_{N_{n1}})$ справедливо равенство:

$$d_{21}(\theta) = \frac{a(\theta)(k_{12} - k_{11}) + b(\theta)(k_{22} - k_{21})}{a(\theta)}.$$

Следствие 4.2. Пусть с.в. M_n имеет биномиальное распределение с параметрами $m(n-1)$, $n \geq 2$, и $p = 1/m$, где $m \geq 2$ — фиксированное натуральное число. Тогда для с.в.

$$N_n = M_n + 1$$

справедливы равенства:

$$EN_n = n;$$

$$\begin{aligned} EN_n^{-1} &= \frac{m \left(1 - (1 - 1/m)^{m(n-1)+1} \right)}{m(n-1) + 1} = \\ &= \frac{1}{n} + \frac{m-1}{mn^2} + O(n^{-3}); \end{aligned}$$

$$EN_n^{-2} = \frac{1}{n^2} + O(n^{-3}), \quad n \rightarrow \infty.$$

Следствие 4.3. Пусть с.в. M_n имеет биномиальное распределение с параметрами $m(n-1)$, $n \geq 2$, и $p = 1/m$, где $m \geq 2$ — фиксированное натуральное число и

$$N_n = M_n + 1.$$

Предположим также, что существуют числа $a(\theta)$ и $b(\theta)$ такие, что справедливо соотношение:

$$R_n^*(\theta) = E_\theta (\delta_n(X_1, \dots, X_n) - g(\theta))^2 = \frac{a(\theta)}{n} + \frac{b(\theta)}{n^2}.$$

Тогда для асимптотического дефекта оценки $\delta_{N_n}(X_1, \dots, X_{N_n})$ относительно оценки $\delta_n(X_1, \dots, X_n)$ справедливо равенство:

$$d(\theta) = \frac{m-1}{m}.$$

Следствие 4.4. Пусть с.в. N_n имеет геометрическое распределение с параметром $p = 1/n$, $n \geq 2$, $n \in \mathbb{N}$. Тогда справедливы равенства:

$$\begin{aligned} \mathbb{E}N_n &= n; \\ \mathbb{E}N_n^{-1} &= \frac{\log n}{n-1} = \frac{\log n}{n} + \frac{\log n}{n^2} + O\left(\frac{\log n}{n^3}\right); \\ \mathbb{E}N_n^{-2} &= \frac{\pi^2}{6n} - \frac{\log n}{n^2} + \frac{1}{n^2} \left(\frac{\pi^2}{6} - 1\right) + O\left(\frac{\log n}{n^3}\right), \\ & n \rightarrow \infty. \end{aligned}$$

Следствие 4.5. Пусть с.в. M_n имеет распределение Пуассона с параметром $\lambda = n-1$, $n \geq 2$, $n \in \mathbb{N}$, и

$$N_n = M_n + 1.$$

Тогда справедливы равенства:

$$\begin{aligned} \mathbb{E}N_n &= n; \quad \mathbb{E}N_n^{-1} = \frac{1}{n} + \frac{1}{n^2} + O\left(\frac{1}{n^3}\right); \\ \mathbb{E}N_n^{-2} &= \frac{1}{n^2} + O\left(\frac{1}{n^3}\right), \quad n \rightarrow \infty. \end{aligned}$$

Следствие 4.6. Пусть с.в. M_n имеет распределение Пуассона с параметром $\lambda = n-1$, $n \geq 2$, и

$$N_n = M_n + 1.$$

Предположим также, что существуют числа $a(\theta)$ и $b(\theta)$ такие, что справедливо соотношение:

$$R_n^*(\theta) = \mathbb{E}_\theta (\delta_n(X_1, \dots, X_n) - g(\theta))^2 = \frac{a(\theta)}{n} + \frac{b(\theta)}{n^2}.$$

Тогда для асимптотического дефекта оценки $\delta_{N_n}(X_1, \dots, X_{N_n})$ относительно оценки $\delta_n(X_1, \dots, X_n)$ справедливо равенство:

$$d(\theta) = 1.$$

Следствие 4.7. Пусть с.в. N_n имеет геометрическое распределение с параметром $p = 1/n$, $n \geq 2$. Предположим также, что существуют числа $a(\theta)$ и $b(\theta)$ такие, что для функции риска оценки $\delta_n(X_1, \dots, X_n)$ справедливо соотношение:

$$R_n^*(\theta) = \mathbb{E}_\theta (\delta_n(X_1, \dots, X_n) - g(\theta))^2 = \frac{a(\theta)}{n} + \frac{b(\theta)}{n^2}.$$

Тогда для функции риска оценки $\delta_{N_n}(X_1, \dots, X_{N_n})$ справедливо равенство:

$$\begin{aligned} R_n(\theta) &= \mathbb{E}_\theta (\delta_{N_n}(X_1, \dots, X_{N_n}) - g(\theta))^2 = \\ &= a(\theta)\mathbb{E}N_n^{-1} + b(\theta)\mathbb{E}N_n^{-2} = \\ &= \frac{a(\theta)\log n}{n} + \frac{\pi^2 b(\theta)}{6n} + \frac{(a(\theta) - b(\theta))\log n}{n^2} + \\ &+ \frac{b(\theta)}{n^2} \left(\frac{\pi^2}{6} - 1\right) + O\left(\frac{\log n}{n^3}\right), \quad n \rightarrow \infty. \end{aligned}$$

В этом разделе рассмотрена также байесовская постановка задачи оценивания. Предположим сначала, что требуется оценить функцию $g(\theta) = \theta^2$ в присутствии мешающего параметра σ^2 . Оценка максимального правдоподобия в этом случае имеет вид:

$$\delta_{0n}(X_1, \dots, X_n) = (\bar{X}_n)^2,$$

где

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i.$$

Если σ^2 известно, то несмещенная оценка с минимальной дисперсией для функции $g(\theta)$, основанная на n наблюдениях, есть

$$\delta_{1n}(X_1, \dots, X_n) = (\bar{X}_n)^2 - \frac{\sigma^2}{n}.$$

Обе эти оценки являются частными ($c = 0$, $c = 1$) случаями оценок вида:

$$\delta_{cn}(X_1, \dots, X_n) = (\bar{X}_n)^2 - \frac{c\sigma^2}{n}, \quad c \in \mathbb{R}. \quad (7)$$

Функция риска этих оценок имеет вид (см. [3, с. 302]):

$$\begin{aligned} R_{cn}^*(\theta) &= \mathbb{E}_\theta (\delta_{cn}(X_1, \dots, X_n) - \theta^2)^2 = \\ &= \frac{4\sigma^2\theta^2}{n} + \frac{(c^2 - 2c + 3)\sigma^4}{n^2}. \quad (8) \end{aligned}$$

Если σ^2 неизвестно, то несмещенная оценка с минимальной дисперсией для функции $g(\theta)$, основанная на n наблюдениях, есть

$$\delta_n(X_1, \dots, X_n) = (\bar{X}_n)^2 - \frac{S_n^2}{n},$$

где

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2.$$

Используя независимость \bar{X}_n и S_n^2 , нетрудно получить, что функция риска оценки δ_n имеет вид:

$$\begin{aligned} R_n^*(\theta) &= \mathbb{E}_\theta (\delta_n(X_1, \dots, X_n) - \theta^2)^2 = \\ &= D_\theta \bar{X}_n^2 + \frac{1}{n^2} D_\theta S_n^2 = \frac{4\sigma^2\theta^2}{n} + \frac{4\sigma^4}{n^2}. \end{aligned}$$

Рассмотрим теперь оценки $\delta_{cn}(X_1, \dots, X_n)$ и $\delta_n(X_1, \dots, X_n)$, основанные на выборке случайного объема N_n . Согласно следствию 2.1 из работы [1] имеем:

$$R_{cn}(\theta) = E_{\theta} (\delta_{cN_n}(X_1, \dots, X_{N_n}) - \theta^2)^2 = 4\sigma^2\theta^2 EN_n^{-1} + (c^2 - 2c + 3)\sigma^4 EN_n^{-2};$$

$$R_n(\theta) = E_{\theta} (\delta_{N_n}(X_1, \dots, X_{N_n}) - \theta^2)^2 = 4\sigma^2\theta^2 EN_n^{-1} + 4\sigma^4 EN_n^{-2}.$$

Лемма 2.1. Пусть с.в. M_n имеет биномиальное распределение с параметрами $m(n-1)$, $n \geq 2$, и $p = 1/m$, где $m \geq 2$ — фиксированное натуральное число, и

$$N_n = M_n + 1.$$

Тогда асимптотические дефекты оценок $\delta_{cN_n}(X_1, \dots, X_{N_n})$, $\delta_{N_n}(X_1, \dots, X_{N_n})$ относительно оценок $\delta_{cn}(X_1, \dots, X_n)$, $\delta_n(X_1, \dots, X_n)$ соответственно равны:

$$d_c(\theta) = \frac{m-1}{m}; \quad d(\theta) = \frac{m-1}{m}.$$

Асимптотический дефект оценки $\delta_{cn}(X_1, \dots, X_n)$ относительно оценки $\delta_n(X_1, \dots, X_n)$ равен:

$$d(c, \theta) = \frac{(c^2 - 2c - 1)\sigma^2}{4\theta^2}.$$

Асимптотический дефект оценки $\delta_{cN_n}(X_1, \dots, X_{N_n})$ относительно оценки $\delta_{N_n}(X_1, \dots, X_{N_n})$ также равен:

$$d(c, \theta) = \frac{(c^2 - 2c - 1)\sigma^2}{4\theta^2}.$$

Доказательство непосредственно вытекает из следствия 4.3 и формул (1.6), (5.5), (5.6), теоремы 3.1 и следствия 4.2 из работы [1].

Лемма 2.2. Пусть с.в. M_n имеет распределение Пуассона с параметром $\lambda = n - 1$, $n \geq 2$, $n \in \mathbb{N}$, и

$$N_n = M_n + 1.$$

Тогда асимптотические дефекты оценок $\delta_{cN_n}(X_1, \dots, X_{N_n})$ и $\delta_{N_n}(X_1, \dots, X_{N_n})$ относительно оценок $\delta_{cn}(X_1, \dots, X_n)$ и $\delta_n(X_1, \dots, X_n)$ соответственно равны:

$$d_c(\theta) = 1; \quad d(\theta) = 1.$$

Асимптотический дефект оценки $\delta_{cN_n}(X_1, \dots, X_{N_n})$ относительно оценки $\delta_{N_n}(X_1, \dots, X_{N_n})$ равен:

$$d(c, \theta) = \frac{(c^2 - 2c - 1)\sigma^2}{4\theta^2}.$$

Доказательство непосредственно вытекает из следствия 4.6 и формул (5.5), (5.6), теоремы 3.1, следствия 3.1 и следствия 4.4 из работы [1]. Непосредственным следствием теоремы 3.2 и следствий 4.2, 4.5 [1] и формулы (8) является следующая теорема.

Теорема 2.1. Пусть с.в. M_{n1} имеет биномиальное распределение с параметрами $m(n-1)$, $n \geq 2$, и $p = 1/m$, где $m \geq 2$ — фиксированное натуральное число, и

$$N_{n1} = M_{n1} + 1.$$

Пусть также с.в. M_{n2} имеет распределение Пуассона с параметром $\lambda = n - 1$, $n \geq 2$, $n \in \mathbb{N}$, и

$$N_{n2} = M_{n2} + 1.$$

Предположим, что с.в. N_{ni} , $i = 1, 2$, не зависят от наблюдений X_1, X_2, \dots . Тогда для асимптотического дефекта оценки (см. (7)) $\delta_{cn}^{(2)} \equiv \delta_{cN_{n2}}(X_1, \dots, X_{N_{n2}})$ относительно оценки $\delta_{cn}^{(1)} \equiv \delta_{cN_{n1}}(X_1, \dots, X_{N_{n1}})$ справедливо равенство:

$$d_{21}(\theta) = -\frac{1}{m}.$$

Теорема 2.2. Пусть с.в. N_n имеет геометрическое распределение с параметром $p = 1/n$, $n \geq 2$. Тогда для функции риска оценки $\delta_{cN_n}(X_1, \dots, X_{N_n})$ справедливо равенство:

$$R_{cn}(\theta) = E_{\theta} (\delta_{cN_n}(X_1, \dots, X_{N_n}) - \theta^2)^2 = \frac{4\sigma^2\theta^2 \log n}{n} + \frac{\pi^2(c^2 - 2c + 3)\sigma^4}{6n} + \frac{(4\sigma^2\theta^2 - (c^2 - 2c + 3)\sigma^4) \log n}{n^2} + \frac{(c^2 - 2c + 3)\sigma^4}{n^2} \left(\frac{\pi^2}{6} - 1 \right) + O\left(\frac{\log n}{n^3} \right), \quad n \rightarrow \infty.$$

Доказательство теоремы 2.2 вытекает из следствия 4.7 [1] и формулы (8).

Замечание 2.1. Рассмотрим теперь байесовскую постановку. Пусть исходная выборка X_1, \dots, X_n представляет собой независимые одинаково нормально распределенные с.в. с параметрами:

$$E_{\theta} X_1 = \theta; \quad D_{\theta} X_1 = \sigma^2.$$

Предположим, что дисперсия σ^2 известна, и рассмотрим оптимальную несмещенную оценку

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$$

функции

$$g(\theta) = \theta.$$

Ее функция риска имеет вид:

$$R_n^*(\theta) = E_\theta (\bar{X}_n - \theta)^2 = D_\theta \bar{X}_n = \frac{\sigma^2}{n}.$$

Предположим теперь, что параметр θ имеет априорное нормальное распределение с параметрами $\mu \in \mathbb{R}$ и $b^2 > 0$. Тогда смещенная байесовская оценка параметра θ имеет вид (см. [3, с. 222]):

$$\theta_n = \frac{nb^2 \bar{X}_n + \sigma^2 \mu}{nb^2 + \sigma^2}.$$

Нетрудно видеть, что функция риска этой оценки может быть записана в виде (см. [4, формулы (6.2) и (6.3)]):

$$\begin{aligned} R_n(\theta) &= E_\theta (\theta_n - \theta)^2 = \\ &= \frac{\sigma^2}{(nb^2 + \sigma^2)^2} (nb^4 + \sigma^2(\theta - \mu)^2) = \\ &= \frac{\sigma^2}{n} + \frac{\sigma^4}{n^2 b^4} ((\theta - \mu)^2 - 2b^2) + O(n^{-3}). \end{aligned}$$

Таким образом, при естественных условиях на случайный объем выборки N_n асимптотический дефект оценки θ_{N_n} относительно наилучшей оценки \bar{X}_n при неслучайном объеме выборки, например, в пуассоновском случае равен:

$$d = \frac{\sigma^2}{b^4} ((\theta - \mu)^2 - 2b^2) + 1.$$

Для биномиального случая соответственно имеем:

$$d = \frac{\sigma^2}{b^4} ((\theta - \mu)^2 - 2b^2) + \frac{m-1}{m}.$$

В геометрическом случае для функции риска оценки \bar{X}_{N_n} справедливо представление:

$$\begin{aligned} R_{N_n}^*(\theta) &= E_\theta (\bar{X}_{N_n} - \theta)^2 = \\ &= \frac{\sigma^2 \log n}{n} + \frac{\sigma^2 \log n}{n^2} + O\left(\frac{\log n}{n^3}\right). \end{aligned}$$

Замечание 2.2. Рассмотрим теперь байесовскую постановку для биномиального распределения. Пусть исходная выборка X_1, \dots, X_n представляет собой независимые одинаково распределенные с.в., принимающие значения 0 и 1 с вероятностями $\theta \in (0, 1)$ и $1 - \theta$ соответственно. Оптимальная несмещенная оценка для параметра θ есть

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i.$$

Ее функция риска имеет вид:

$$R_n^*(\theta) = E_\theta (\bar{X}_n - \theta)^2 = D_\theta \bar{X}_n = \frac{\theta(1-\theta)}{n}.$$

Предположим теперь, что параметр θ имеет априорное бета-распределение с параметрами (a, b) , $a > 0$, $b > 0$. Тогда байесовская оценка параметра θ имеет вид (см. [3, с. 221]):

$$\hat{\theta}_n = \frac{a + n\bar{X}_n}{a + b + n}.$$

Нетрудно видеть, что функция риска этой оценки может быть записана в виде:

$$\begin{aligned} R_n(\hat{\theta}) &= E_\theta (\hat{\theta}_n - \theta)^2 = \\ &= \frac{1}{(a + b + n)^2} (n\theta(1-\theta) + (a(1-\theta) - \theta b)^2) = \\ &= \frac{\theta(1-\theta)}{n} + \frac{(a(1-\theta) - \theta b)^2 - \theta(1-\theta)2(a+b)}{n^2} + \\ &\quad + O(n^{-3}). \end{aligned}$$

Таким образом, при естественных условиях на случайный объем выборки N_n асимптотический дефект оценки θ_{N_n} относительно наилучшей оценки \bar{X}_n при неслучайном объеме выборки, например, в пуассоновском случае равен:

$$d(\theta) = \frac{((a(1-\theta) - \theta b)^2 - \theta(1-\theta)2(a+b))}{\theta(1-\theta)} + 1.$$

Если $a = b$ и $\theta = 1/2$, то эта формула упрощается и

$$d\left(\frac{1}{2}\right) = 1 - 4a.$$

Если случайный объем выборки N_n имеет биномиальное распределение, то соответствующий асимптотический дефект равен:

$$d(\theta) = \frac{((a(1-\theta) - \theta b)^2 - \theta(1-\theta)2(a+b))}{\theta(1-\theta)} + \frac{m-1}{m}.$$

В случае $a = b$ и $\theta = 1/2$ имеем:

$$d\left(\frac{1}{2}\right) = \frac{m-1}{m} - 4a.$$

В геометрическом случае для функции риска оценки \bar{X}_{N_n} справедливо представление:

$$\begin{aligned} R_{N_n}^*(\theta) &= E_\theta (\bar{X}_{N_n} - \theta)^2 = \\ &= \frac{\theta(1-\theta) \log n}{n} + \frac{\theta(1-\theta) \log n}{n^2} + O\left(\frac{\log n}{n^3}\right). \end{aligned}$$

3 Доверительное оценивание

В этом разделе находятся асимптотические дефекты некоторых доверительных интервалов, основанных на выборках случайного объема.

Пусть исходные наблюдения X_1, \dots, X_n представляют собой независимые одинаково распределенные с.в. с плотностью вида:

$$p(x, \theta) = \begin{cases} e^{-x+\theta}, & x \geq \theta; \\ 0, & x < \theta, \end{cases} \quad \theta \in \mathbb{R}. \quad (9)$$

Нетрудно видеть, что для каждого $\alpha \in (0, 1)$ и каждого $n \in \mathbb{N}$ в качестве доверительного интервала с коэффициентом доверия $1 - \alpha$ для параметра θ можно взять интервал вида:

$$\left(X_{n:1} + \frac{\log \alpha}{n}, X_{n:1} \right), \quad (10)$$

где $X_{n:1} \equiv \min\{X_1, \dots, X_n\}$ — минимальная порядковая статистика, построенная по исходной выборке X_1, \dots, X_n . Длина этого доверительного интервала неслучайна и равна

$$-\frac{\log \alpha}{n}. \quad (11)$$

В качестве меры качества доверительного интервала (10) выберем его длину (11).

Рассмотрим теперь доверительный интервал для параметра θ , основанный на выборке случайного объема N_n вида:

$$\left(X_{N_n:1} + \frac{\log \alpha}{N_n}, X_{N_n:1} \right). \quad (12)$$

Нетрудно видеть, что он имеет также коэффициент доверия $1 - \alpha$ и случайную длину $-\log \alpha / N_n$.

Рассмотрим в качестве меры качества такого рода доверительных интервалов их среднюю длину, т. е. величину

$$-\log \alpha \mathbb{E}N_n^{-1} \quad (13)$$

при условии $\mathbb{E}N_n = n$. Найдем дефект случайного доверительного интервала (12) относительно неслучайного доверительного интервала (10).

Теорема 3.1. Пусть исходные наблюдения X_1, \dots, X_n представляют собой независимые одинаково распределенные с.в. с плотностью вида (9). Тогда

- (1) если с.в. M_n имеет биномиальное распределение с параметрами $m(n - 1)$, $n \geq 2$, и $p = 1/m$, где $m \geq 2$ — фиксированное натуральное число, и

$$N_n = M_n + 1,$$

то асимптотический дефект случайного доверительного интервала (12) относительно неслучайного доверительного интервала (10) равен

$$d = \frac{m - 1}{m};$$

- (2) если с.в. M_n имеет распределение Пуассона с параметром $\lambda = n - 1$, $n \geq 2$, $n \in \mathbb{N}$, и

$$N_n = M_n + 1,$$

то асимптотический дефект случайного доверительного интервала (12) относительно неслучайного доверительного интервала (10) равен

$$d = 1;$$

- (3) если с.в. N_n имеет геометрическое распределение с параметром $p = 1/n$, $n \geq 2$, то средняя длина случайного доверительного интервала (12) равна

$$\begin{aligned} -\log \alpha \mathbb{E}N_n^{-1} &= -\frac{\log \alpha \log n}{n - 1} = \\ &= -\log \alpha \left(\frac{\log n}{n} + \frac{\log n}{n^2} + O\left(\frac{\log n}{n^3}\right) \right). \end{aligned}$$

Доказательство теоремы 3.1 непосредственно следует из формул (11), (13), (6) и следствий 4.2, 4.5 и 4.4 работы [1].

Аналогично можно получить следующую теорему.

Теорема 3.2. Пусть исходные наблюдения X_1, \dots, X_n имеют плотность вида (9). Предположим, что с.в. M_{n1} имеет биномиальное распределение с параметрами $m(n - 1)$, $n \geq 2$, и $p = 1/m$, где $m \geq 2$ — фиксированное натуральное число, и

$$N_{n1} = M_{n1} + 1.$$

Пусть также с.в. M_{n2} имеет распределение Пуассона с параметром $\lambda = n - 1$, $n \geq 2$, $n \in \mathbb{N}$, и

$$N_{n2} = M_{n2} + 1.$$

Предположим, что с.в. N_{ni} , $i = 1, 2$, не зависят от наблюдений X_1, X_2, \dots . Тогда для асимптотического дефекта (с критерием качества (13)) доверительного интервала $(X_{N_{n1}:1} + \log \alpha / N_{n1}, X_{N_{n1}:1})$ относительно доверительного интервала $(X_{N_{n2}:1} + \log \alpha / N_{n2}, X_{N_{n2}:1})$ справедливо равенство:

$$d = -\frac{1}{m}.$$

Пусть теперь исходные наблюдения X_1, \dots, X_n представляют собой независимые одинаково распределенные нормальные с параметрами $\theta \in \mathbb{R}$, $\sigma^2 > 0$ случайные величины. Предположим, что дисперсия σ^2 известна. Для параметра θ рассмотрим доверительный интервал с коэффициентом доверия $1 - \alpha$ вида

$$\left(\bar{X}_n - \frac{\sigma u_\alpha}{\sqrt{n}}, \bar{X}_n + \frac{\sigma u_\alpha}{\sqrt{n}} \right), \quad (14)$$

где

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$$

а стандартная нормальная функция распределения

$$\Phi(u_\alpha) = 1 - \frac{\alpha}{2}.$$

Длина этого доверительного интервала неслучайна и равна $2\sigma u_\alpha / \sqrt{n}$.

В качестве меры качества доверительного интервала (14) выберем квадрат его длины, т. е. величину

$$\frac{4\sigma^2 u_\alpha^2}{n}. \quad (15)$$

Рассмотрим теперь доверительный интервал для параметра θ , основанный на выборке случайного объема N_n вида:

$$\left(\bar{X}_{N_n} - \frac{\sigma u_\alpha}{\sqrt{N_n}}, \bar{X}_{N_n} + \frac{\sigma u_\alpha}{\sqrt{N_n}} \right). \quad (16)$$

Он имеет также коэффициент доверия $1 - \alpha$ и случайную длину $2\sigma u_\alpha / \sqrt{N_n}$.

Аналогично в качестве меры качества этого доверительного интервала рассмотрим средний квадрат его длины, т. е. величину

$$4\sigma^2 u_\alpha^2 \mathbb{E} N_n^{-1} \quad (17)$$

при условии $\mathbb{E} N_n = n$. Найдем дефект случайного доверительного интервала (16) относительно неслучайного доверительного интервала (14).

Теорема 3.3. Пусть исходные наблюдения X_1, \dots, X_n представляют собой независимые одинаково нормально распределенные с.в. с параметрами $\theta \in \mathbb{R}$ и $\sigma^2 > 0$. Тогда

- (1) если с.в. M_n имеет биномиальное распределение с параметрами $m(n-1)$, $n \geq 2$, и $p = 1/m$, где $m \geq 2$ — фиксированное натуральное число, и

$$N_n = M_n + 1,$$

то асимптотический дефект случайного доверительного интервала (16) относительно неслучайного доверительного интервала (14) (с критериями качества (15) и (17)) равен

$$d = \frac{m-1}{m};$$

- (2) если с.в. M_n имеет распределение Пуассона с параметром $\lambda = n-1$, $n \geq 2$, $n \in \mathbb{N}$, и

$$N_n = M_n + 1,$$

то асимптотический дефект случайного доверительного интервала (16) относительно неслучайного доверительного интервала (14) равен

$$d = 1;$$

- (3) если с.в. N_n имеет геометрическое распределение с параметром $p = 1/n$, $n \geq 2$, то средний квадрат длины случайного доверительного интервала (16) равен

$$\begin{aligned} 4\sigma^2 u_\alpha^2 \mathbb{E} N_n^{-1} &= \frac{4\sigma^2 u_\alpha^2 \log n}{n-1} = \\ &= 4\sigma^2 u_\alpha^2 \left(\frac{\log n}{n} + \frac{\log n}{n^2} + O\left(\frac{\log n}{n^3}\right) \right). \end{aligned}$$

Доказательство теоремы 3.3 непосредственно следует из формул (15), (17), (6) и следствий 4.2, 4.5 и 4.4 из работы [1]. Аналогично можно получить следующую теорему.

Теорема 3.4. Пусть исходные наблюдения X_1, \dots, X_n имеют нормальное распределение с параметрами $\theta \in \mathbb{R}$ и $\sigma^2 > 0$. Предположим, что с.в. M_{n1} имеет биномиальное распределение с параметрами $m(n-1)$, $n \geq 2$, и $p = 1/m$, где $m \geq 2$ — фиксированное натуральное число, и

$$N_{n1} = M_{n1} + 1.$$

Пусть также с.в. M_{n2} имеет распределение Пуассона с параметром $\lambda = n-1$, $n \geq 2$, $n \in \mathbb{N}$, и

$$N_{n2} = M_{n2} + 1.$$

Предположим, что с.в. N_{ni} , $i = 1, 2$, не зависят от наблюдений X_1, X_2, \dots . Тогда для асимптотического дефекта (с критерием качества (17)) доверительного интервала $(\bar{X}_{N_{n1}} - \sigma u_\alpha / \sqrt{N_{n1}}, \bar{X}_{N_{n1}} + \sigma u_\alpha / \sqrt{N_{n1}})$ относительно доверительного интервала $(\bar{X}_{N_{n2}} - \sigma u_\alpha / \sqrt{N_{n2}}, \bar{X}_{N_{n2}} + \sigma u_\alpha / \sqrt{N_{n2}})$ справедливо равенство:

$$d = -\frac{1}{m}.$$

4 Проверка статистических гипотез

В этом разделе находится асимптотический дефект критерия Стьюдента относительно наиболее мощного критерия в случае выборки случайного объема из нормального распределения.

Пусть наблюдения X_1, \dots, X_n есть независимые одинаково распределенные нормальные с параметрами $\theta \in \mathbb{R}$, $\sigma^2 > 0$ случайные величины. Предположим, что дисперсия σ^2 известна. Рассмотрим задачу проверки простой гипотезы

$$H_0 : \theta = 0$$

против сложной односторонней альтернативы

$$H_1 : \theta > 0.$$

По лемме Неймана–Пирсона равномерно наиболее мощный критерий уровня значимости $\alpha \in (0, 1)$ существует и отвергает гипотезу H_0 , если

$$\frac{\sqrt{n}\bar{X}_n}{\sigma} > u_{1-\alpha}, \tag{18}$$

где

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$$

и

$$\Phi(u_{1-\alpha}) = 1 - \alpha.$$

Функция мощности этого критерия имеет вид:

$$\beta_n^*(\theta) = \Phi(\sqrt{n}\theta - u_{1-\alpha}).$$

Предположим, что исходная выборка имеет случайный объем N_n , и рассмотрим критерий, отвергающий гипотезу H_0 , если (см. (18))

$$\frac{\sqrt{N_n}\bar{X}_{N_n}}{\sigma} > u_{1-\alpha}. \tag{19}$$

Тогда этот критерий также имеет размер α и мощность

$$\begin{aligned} P_\theta \left(\frac{\sqrt{N_n}\bar{X}_{N_n}}{\sigma} > u_{1-\alpha} \right) &= E\beta_{N_n}^*(\theta) = \\ &= E\Phi(\sqrt{N_n}\theta - u_{1-\alpha}). \end{aligned}$$

Рассмотрим теперь для проверки гипотезы H_0 против альтернативы H_1 критерий Стьюдента, основанный на $k \in \mathbb{N}$ наблюдениях и имеющий критическую область вида

$$\frac{\sqrt{k}\bar{X}_k}{S_k} > c_{\alpha,k}, \tag{20}$$

где

$$S_k^2 = \frac{1}{k-1} \sum_{i=1}^k (X_i - \bar{X}_k)^2$$

и критическое значение $c_{\alpha,k}$ выбрано так, чтобы этот критерий имел уровень значимости α . В работе [4, формула (5.6)] для функции мощности этого критерия получено асимптотическое разложение вида

$$\begin{aligned} \beta_k(\theta) &= \Phi\left(\sqrt{k}\theta\left(1 - \frac{u_{1-\alpha}^2}{4k}\right) - u_{1-\alpha}\right) + O(k^{-2}) = \\ &= \Phi\left(\sqrt{k}\theta\left(1 - \frac{u_{1-\alpha}^2}{4k}\right) - u_{1-\alpha} + O(k^{-2})\right), \end{aligned}$$

причем это равенство выполнено равномерно по θ . Приравнявая выражение $\beta_{k_n}(\theta)$ к $\beta_n^*(\theta)$, найдем асимптотический дефект критерия Стьюдента (20) относительно равномерно наиболее мощного критерия (18):

$$\frac{n}{k_n} = 1 - \frac{u_{1-\alpha}^2}{2k_n} + O(k_n^{-2});$$

$$d_n = k_n - n \rightarrow d = \frac{u_{1-\alpha}^2}{2}, \quad n \rightarrow \infty.$$

Предположим, что исходная выборка имеет случайный объем N_n , и рассмотрим аналог критерия Стьюдента, отвергающий гипотезу H_0 , если (см. (20))

$$\frac{\sqrt{N_n}\bar{X}_{N_n}}{S_{N_n}} > c_{\alpha,N_n}. \tag{21}$$

Тогда нетрудно видеть, что этот критерий также имеет размер α и мощность

$$\begin{aligned} P_\theta \left(\frac{\sqrt{N_n}\bar{X}_{N_n}}{S_{N_n}} > c_{\alpha,N_n} \right) &= E\beta_{N_n}(\theta) = \\ &= E\Phi\left(\sqrt{N_n}\theta\left(1 - \frac{u_{1-\alpha}^2}{4N_n}\right) - u_{1-\alpha}\right) + O(EN_n^{-2}) = \\ &= E\Phi\left(\sqrt{N_n}\theta\left(1 - \frac{u_{1-\alpha}^2}{4N_n}\right) - u_{1-\alpha} + O(EN_n^{-2})\right). \end{aligned}$$

Для нахождения дефектов критериев, основанных на выборках случайного объема, рассмотрим в качестве меры качества критериев с критическими областями (18) и (20) соответственно величины вида

$$\pi_n^* = (\Phi^{-1}(\beta_n^*(\theta)) + u_{1-\alpha})^2 = n\theta^2; \tag{22}$$

$$\begin{aligned} \pi_n &= (\Phi^{-1}(\beta_n(\theta)) + u_{1-\alpha})^2 = \\ &= \left(\sqrt{n}\theta\left(1 - \frac{u_{1-\alpha}^2}{4n}\right) + O(n^{-2})\right)^2. \end{aligned} \tag{23}$$

В качестве меры качества критериев с критическими областями (19) и (21), основанных на выборках случайного объема N_n , рассмотрим соответственно усредненные величины (22) и (23):

$$\begin{aligned} E\pi_{N_n}^* &= E(\Phi^{-1}(\beta_{N_n}^*(\theta)) + u_{1-\alpha})^2 = \theta^2 EN_n; \\ E\pi_n &= E(\Phi^{-1}(\beta_{N_n}(\theta)) + u_{1-\alpha})^2 = \\ &= E\left(\sqrt{N_n}\theta\left(1 - \frac{u_{1-\alpha}^2}{4N_n}\right) + O(N_n^{-2})\right)^2. \end{aligned}$$

При условии

$$N_n = n$$

эти выражения соответственно приобретают вид:

$$\begin{aligned} E\pi_{N_n}^* &= \theta^2 n; \quad (24) \\ E\pi_n &= \\ &= \theta^2 n - \frac{\theta^2 u_{1-\alpha}^2}{2} + \frac{\theta^2 u_{1-\alpha}^2}{16} EN_n^{-1} + O(EN_n^{-3/2}). \quad (25) \end{aligned}$$

Теперь для нахождения асимптотического дефекта из равенства

$$E\pi_{N_{k_n}} = E\pi_{N_n}^*$$

получаем соотношение:

$$d_n = k_n - n = \frac{u_{1-\alpha}^2}{2} + O(EN_{k_n}^{-1}), \quad k_n \rightarrow \infty.$$

Из этой формулы получаем следующую теорему.

Теорема 4.1. Пусть исходные наблюдения X_1, \dots, X_n имеют нормальное распределение с параметрами $\theta \in \mathbb{R}$ и $\sigma^2 > 0$. Предположим, что с.в. N_n не зависит от наблюдений X_1, \dots, X_n и удовлетворяет следующим условиям:

$$N_n = n, \quad EN_n^{-1} \rightarrow 0, \quad n \rightarrow \infty.$$

Тогда асимптотический дефект критерия (21) относительно критерия (19) (с мерами качества (25) и (24)) равен

$$d = \frac{u_{1-\alpha}^2}{2}.$$

Следствие 4.1. Пусть с.в. N_n удовлетворяет условиям теоремы 3.3 (т.е. имеет соответственно пуассоновское, биномиальное или геометрическое распределение), тогда условия теоремы 4.1 выполнены и асимптотический дефект равен

$$d = \frac{u_{1-\alpha}^2}{2}.$$

Найдем теперь дефекты критериев с критическими областями (19) и (21) относительно критериев, основанных на выборках неслучайного объема

и имеющих соответственно критические области вида (18) и (20). При этом используем меры качества (см. (22), (23)):

$$\pi_n^* = n\theta^2, \quad E\pi_{N_n}^* = \theta^2 EN_n; \quad (26)$$

$$\begin{aligned} \pi_n &= n\theta^2 - \frac{\theta^2 u_{1-\alpha}^2}{2} + \frac{\theta^2 u_{1-\alpha}^4}{16n} + O(n^{-3/2}), \\ E\pi_n &= \theta^2 EN_n - \frac{\theta^2 u_{1-\alpha}^2}{2} + \frac{\theta^2 u_{1-\alpha}^4 EN_n^{-1}}{16} + \\ &+ O(EN_n^{-3/2}). \quad (27) \end{aligned}$$

Теперь, находя k_n из равенств

$$E\pi_{N_{k_n}}^* = \pi_n^*; \quad E\pi_{N_{k_n}} = \pi_n,$$

получаем следующую теорему.

Теорема 4.2. Пусть исходные наблюдения X_1, \dots, X_n имеют нормальное распределение с параметрами $\theta \in \mathbb{R}$ и $\sigma^2 > 0$. Предположим, что с.в. N_n не зависит от наблюдений X_1, \dots, X_n и удовлетворяет следующим условиям:

$$N_n = n, \quad EN_n^{-3/2} \rightarrow 0, \quad n \rightarrow \infty.$$

Тогда асимптотический дефект критерия (19) относительно критерия (18) (с мерой качества (26)) равен

$$d_n = k_n - n = 0, \quad d = \lim_{n \rightarrow \infty} d_n = 0.$$

Асимптотический дефект критерия (21) относительно критерия (20) (с мерой качества (27)) равен

$$\begin{aligned} d_n = k_n - n &= \frac{u_{1-\alpha}^4}{16} \left(EN_{k_n}^{-1} - \frac{1}{n} \right) + \\ &+ O(EN_{k_n}^{-3/2} + n^{-3/2}), \quad d = \lim_{n \rightarrow \infty} d_n = 0. \end{aligned}$$

Следствие 4.2. Пусть с.в. N_n удовлетворяет условиям теоремы 3.3 (т.е. имеет соответственно пуассоновское, биномиальное или геометрическое распределение), тогда условия теоремы 4.2 выполнены и соответствующие асимптотические дефекты равны нулю. Более того, в этих случаях справедливы соответственно асимптотические представления для величины d_n в случае критериев (21) и (20)

$$d_n = k_n - n = \frac{u_{1-\alpha}^4}{16n^2} + o(n^{-2});$$

$$d_n = k_n - n = \frac{u_{1-\alpha}^4(m-1)}{16mn^2} + o(n^{-2});$$

$$d_n = k_n - n = \frac{u_{1-\alpha}^4}{16} \left(\frac{\log n}{n-1} - \frac{1}{n} \right) + o\left(\frac{\log n}{n}\right).$$

5 Заключение

Таким образом, в работе рассмотрены проблемы точечного оценивания, доверительного оценивания и проверки гипотез, касающиеся неизвестных параметров распределения. При этом предполагается, что статистики, используемые в статистических процедурах, основаны на выборках случайного объема.

Рассмотрена также байесовская постановка задачи оценивания. С помощью понятия дефекта проведено асимптотическое сравнение качества таких процедур.

Получены явные формулы для асимптотического дефекта, имеющего смысл необходимого добавочного числа наблюдений.

Подробно рассмотрены случаи, когда случайный объем выборки имеет распределение Пуассона, биномиальное и геометрическое распределение. Случай, когда исходные наблюдения распределены по нормальному закону с неизвестными параметрами, приводится в качестве примера, иллюстрирующего общие результаты. Находятся асимптотические дефекты доверительных интервалов, а также некоторых статистических критериев (например,

критерия Стьюдента), построенных по выборкам случайного объема.

Литература

1. Бенинг В. Е. О дефекте некоторых оценок, основанных на выборках случайного объема // Вестник Тверского гос. ун-та. Сер.: Прикладная математика, 2015. Вып. 1. С. 5–14.
2. Бенинг В. Е., Королев В. Ю., Савушкин В. А. О сравнении качества оценок, основанных на выборках случайного объема, с помощью понятия дефект // Статистические методы оценивания и проверки гипотез, 2015. Вып. 26. С. 26–42.
3. Леман Э. Теория точечного оценивания / Пер. с англ. — М.: Наука, 1994. 444 с.
4. Hodges J. L., Lehmann E. L. Deficiency // Ann. Math. Stat., 1970. Vol. 41. No. 5. P. 783–801.
5. Bening V. E. Asymptotic theory of testing statistical hypotheses: Efficient statistics, optimality, power loss, and deficiency. — Utrecht: VSP, 2000. 277 p.
6. Крамер Г. Математические методы статистики / Пер. с англ. — М.: Мир, 1976. 648 с. (*Cramer H. Mathematical methods of statistics*. Princeton, NJ, USA: Princeton University Press, 1946. 656 p.)

Поступила в редакцию 15.10.16

CALCULATION OF THE ASYMPTOTIC DEFICIENCY OF SOME STATISTICAL PROCEDURES BASED ON SAMPLES WITH RANDOM SIZES

V. E. Bening^{1,2}

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

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

Abstract: Statistical regularities of information flows in contemporary communication, computational and other information systems are characterized by the presence of the so-called “heavy tails.” The random character of the intensity of the flow of informative events results in that the available sample size (traditionally, this is the number of observations registered within a certain time interval) is random. The randomness of the sample size crucially changes the asymptotic properties of the statistical procedures (e. g., estimators). The present paper consists of a number of applications of the deficiency concept, i. e., the number of additional observations required by the less effective procedure and, thereby, provides a basis for deciding whether or not the price is too high. The deficiency was introduced by Hodges and Lehmann in 1970. In the paper, asymptotic deficiencies of statistical procedures based on samples with random sizes are considered. Three examples concerning testing statistical hypotheses, point, and confidence estimation are presented.

Keywords: confidence set; statistical hypothesis; asymptotic deficiency; sample with random size; Poisson distribution; binomial distribution

DOI: 10.14357/19922264160404

Acknowledgments

This work was financially supported by the Russian Science Foundation (grant No. 14-11-00364).

References

1. Bening, V. E. 2015. O defekte nekotorykh otsenok, osnovannykh na vyborkakh sluchaynogo ob"ema [On deficiencies of some estimators based on samples with random sizes]. *Vestnik Tverskogo gos. un-ta. Ser. Prikladnaya matematika* [Herald of Tver State University, Ser. Applied Mathematics] 1:5–14.
2. Bening, V. E., V. Yu. Korolev, and V. A. Savushkin. 2015. O sravnenii kachestva otsenok, osnovannykh na vyborkakh sluchaynogo ob"ema, s pomoshch'yu ponyatiya defekt [On the comparison of statistical estimators based on samples with random sizes with the help of deficiency concept]. *Statisticheskie metody otsenivaniya i proverki gipotez* [Statistical Methods of Estimation and Testing Hypotheses] 26:26–42.
3. Lehmann, E. L., and G. Casella. 1998. *Theory of point estimation*. 2nd ed. New York, NY: Springer Verlag. 470 p.
4. Hodges, J. L., and E. L. Lehmann. 1970. Deficiency. *Ann. Math. Stat.* 41(5):783–801.
5. Bening, V. E. 2000. *Asymptotic theory of testing statistical hypotheses: Efficient statistics, optimality, power loss, and deficiency*. Utrecht: VSP. 277 p.
6. Cramer, H. 1946. *Mathematical methods of statistics*. Princeton, NJ: Princeton University Press. 656 p.

Received October 15, 2016

Contributor

Bening Vladimir E. (b. 1954) — Doctor of Science in physics and mathematics, professor, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; bening@yandex.ru

МЕТОД КУМУЛЯТИВНЫХ СУММ ДЛЯ ПОИСКА СМЕНЫ РЕЖИМА В ПРОЦЕССЕ ОРНШТЕЙНА–УЛЕНБЕКА НА ОСНОВЕ ПРОЦЕССА ЛЕВИ*

А. В. Черток¹, А. И. Каданер², Г. Т. Хазеева³, И. А. Соколов⁴

Аннотация: Рассматривается процесс Орнштейна–Уленбека (ОУ) с трендом на основе процесса Леви для моделирования финансовых временных рядов. Продемонстрировано, что использование процесса Леви в основе процесса ОУ дает больше гибкости для описания финансовых временных рядов по сравнению с классической гауссовой моделью. В частности, процесс Леви позволяет моделировать остатки с тяжелыми хвостами, что является распространенным свойством реальных временных рядов. Приводятся эффективные решения для оценивания параметров модели с использованием таких методов, как OLS (ordinary least squares) и RLS (regularized least squares). Решается задача поиска моментов смены режима в модели при условии поступления данных в режиме реального времени. Приведен алгоритм, основанный на CUSUM (CUmulative SUM) методах, способный последовательно обрабатывать смены режима и поддерживать параметры модели актуальными для каждого момента времени. Решение задачи поиска разладки модели и соответствующих смен режима имеет важное прикладное значение, поскольку в большинстве случаев параметры моделей, описывающих динамику реальных систем, меняются во времени под действием внешних факторов.

Ключевые слова: случайные процессы; процессы со свойством возвратности к среднему; процесс Орнштейна–Уленбека, управляемый процессом Леви; процесс Орнштейна–Уленбека с трендом; смена режима; CUSUM-алгоритмы

DOI: 10.14357/19922264160405

1 Введение

Процессы со свойством возвратности к среднему играют важную роль в моделировании динамики явлений из различных областей человеческой деятельности. В частности, эти процессы привлекательны для моделирования различных явлений в эконометрике, таких как процентные ставки, курсы обмена валют и цены на сырьевые товары, где свойство возвратности к среднему имеет фундаментальную природу.

В работе [1] рассмотрено несколько видов случайных процессов со свойством возвратности к среднему и описаны их основные характеристики. В настоящей статье в качестве такого процесса рассматривается процесс ОУ, управляемый процессом Леви.

Классическая версия процесса была впервые представлена в совместной работе голландских фи-

зиков Л. С. Орнштейна и Дж. Е. Уленбека [2] в качестве модели, которая способна описать данные с гауссовской и диффузионной структурой. В экономике же классический процесс ОУ известен как модель Васичека благодаря фундаментальной работе [3], где автор предлагает использовать ее для моделирования временного ряда процентной ставки. Ее основной недостаток заключается в том, что существует ненулевая вероятность появления отрицательных значений, нереалистичных для экономических процессов. Для решения данной проблемы позднее была разработана экспоненциальная модель Васичека, а также модель процесса Кокса–Ингерсолла–Росса, также называемая «моделью с квадратным корнем», в которой процентная ставка принимает только неотрицательные значения и имеет гамма-распределение [4].

В настоящей статье подтверждается тот факт, что предположение нормальности в классической

* Работа выполнена при частичной финансовой поддержке РФФИ (проект 14-07-00041).

¹ Факультет вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова; Сбербанк России, avchertok.sbt@sberbank.ru

² Механико-математический факультет Московского государственного университета им. М. В. Ломоносова; Сбербанк России, aikadanner.sbt@sberbank.ru

³ Факультет вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова, gelana.khazeyeva@gmail.com

⁴ Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, isokolov@ipiran.ru

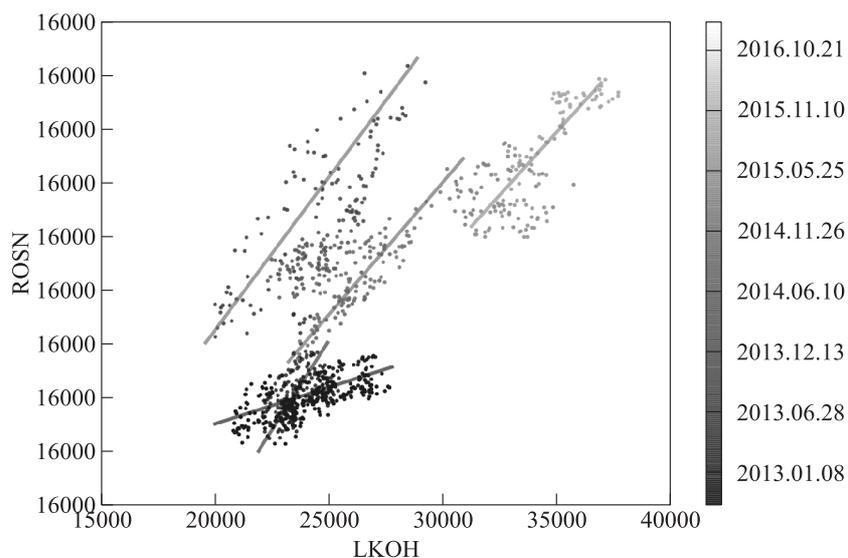


Рис. 1 График соотношения цен для фьючерсов компаний «Лукойл» и «Роснефть»

версии процесса ОУ не описывает реальную структуру данных, и поэтому рассматривается обобщение классического процесса — процесс ОУ, управляемый процессом Леви. Некоторые его модификации изучены в работе [5]. Предложено рассматривать нормальный обратный гауссовский и дисперсионный гамма-процесс для описания динамики его остатков. Распределения приращений этих процессов имеют хвосты тяжелее, чем у нормального распределения, что часто встречается в реальных данных.

Дополнительная мотивация в использовании именно этих распределений исходит из приложений в финансах. Например, дисперсионное гамма-распределение используется для моделирования цен акций, как это делается в работе [6], а нормальное обратное гауссовское распределение хорошо описывает логарифмические приращения цен активов, например в работе А. В. Кузьминой [7] это подтверждается на примере данных о цене фьючерса RTS.

Для более общего механизма построения моделей к классической модели ОУ добавляется линейная составляющая, или тренд. Такой подход позволяет моделировать большее число явлений, не выходя за рамки одной модели.

Как известно, финансовые рынки являются динамическими и нестационарными системами. Поэтому отношения, связывающие различные факторы рынка, склонны меняться во времени. Пример данного явления продемонстрирован на рис. 1. По оси x отложены цены фьючерса на акции компании «Роснефть» (ROSN), а по оси y — цены фьючерса на акции компании «Лукойл» (LKOH) с 08.01.2013

по 28.10.2016. Видно, что параметры этой зависимости являются также изменяющимися во времени на протяжении дня, так как можно отчетливо выделить области, где точки группируются в окрестностях прямых с разными параметрами.

Все это ставит задачу определения моментов, в которые предложенная для описания данных модель с определенными параметрами перестает работать, после чего процесс начинает следовать той же самой модели, но уже с другими параметрами. В данной статье эта проблема решена для модели ОУ. Более того, предлагается процедура оценивания параметров и обнаружения смен режима в реальном времени с использованием RLS или рекурсивного метода наименьших квадратов для оценивания параметров, а также алгоритм, основанный на CUSUM-процедурах для обнаружения смен режима. В конце статьи предложенная процедура применяется на различных данных.

2 Моделирование временного ряда

2.1 Одномерный процесс Орнштейна–Уленбека

Процесс ОУ с трендом, управляемый процессом Леви, определяется как решение стохастического дифференциального уравнения (СДУ):

$$d(X_t - \mu - \nu t) = -\alpha(X_t - \mu - \nu t) dt + dL_{\lambda t}, \quad \forall t > 0;$$

$$X(0) = X_0,$$

где $\alpha, \mu \in \mathbb{R}$; L_t — процесс Леви; X_0 — некоторая случайная величина, независимая от $\{L_t\}$; ν определяет постоянный на всем промежутке времени линейный тренд. Параметр μ здесь означает долгосрочное среднее, а α определяет скорость стремления процесса возвращаться к своему среднему — тренду.

Как показано в [8], данное СДУ имеет следующее решение:

$$X(t) = \nu t + \mu + \exp(-\alpha t) \times \left((X_0 - \mu) + \int_0^t \exp(\alpha s) dL_{\lambda s} \right), \quad X_0 = X(0),$$

или

$$X(t + \tau) = \mu + \nu(t + \tau) + \exp(-\alpha \tau) \times \left((X(t) - \mu - \nu t) + \exp(-\alpha t) \int_t^{t+\tau} \exp(\alpha s) dL_{\lambda s} \right). \quad (1)$$

Отсюда, в частности, следует, что X_t — марковский процесс. Еще стоит заметить, что данное решение единственно с точностью до неотличимости [9]. Для более подробного рассмотрения процессов Леви см. [8, 10].

Для удобства обозначим через $Y_t = X_t - \mu - \nu t$ соответствующий приведенный процесс ОУ без тренда и имеющий нулевое среднее.

Свойство возвратности процесса Y_t к нулевому уровню при $\alpha > 0$ может быть получено из (1): если Y_t стал больше 0 в момент времени t , то коэффициент при dt отрицательный и Y_t будет стремиться немедленно вернуться к 0; аналогично происходит, если случайный процесс становится меньше 0.

2.1.1 Авторегрессия и оценка параметров

Пусть $X^* = (X_{t_i}^*)_{i=1, \dots, N}$ — наблюдения с интервалом $\Delta = 1$ процесса, описываемого определенной выше моделью ОУ с трендом. В дискретном случае уравнение процесса (1) выглядит следующим образом:

$$X_{i+1} = \mu + \mu_0 (1 - e^{-\lambda}) + \mu (1 - e^{-\lambda}) i + e^{-\lambda} X_i + l_i, \quad (2)$$

где l_i — некоторая случайная величина с нулевым средним.

Соотношение (2) описывается регрессионной моделью. Запишем его в виде:

$$X_{i+1} = c + bt_i + aX_i + l_i.$$

Чтобы оценить параметры a , b и c регрессии, можно воспользоваться методом наименьших квадратов и получить оценки \hat{a} , \hat{b} и \hat{c} . Тогда параметры исходного процесса ОУ с трендом можно получить из соотношений:

$$\hat{\lambda} = -\frac{\ln \hat{a}}{\tau}; \quad \hat{\mu} = \frac{\hat{b}}{1 - \hat{a}}; \quad \hat{\mu}_0 = \frac{\hat{c} - \hat{\mu}\tau}{1 - \hat{a}}.$$

Из независимости приращений также можно явно посчитать логарифмическую функцию правдоподобия:

$$L(X^*, \theta) = \sum_{k=2}^n \ln f_{Y_i|Y_{i-1}}(X_i, \theta) = \sum_{k=2}^n \ln f(Y_i - aY_{i-1} - c, \theta),$$

где θ — параметры модели.

2.1.2 Симуляция

Используя соотношения авторегрессионного вида процесса ОУ, можно смоделировать процесс ОУ итеративно, задав некоторую начальную точку X_0 . На рис. 2 проиллюстрирован построенный итеративно процесс ОУ с положительным трендом.

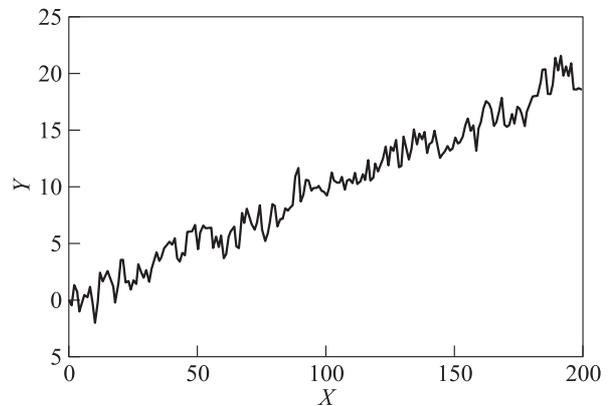


Рис. 2 Пример процесса ОУ с трендом ($\alpha = 0,5$, $\nu = 0,1$, $\mu_0 = 0$ и $\sigma = 1$)

2.2 Частные случаи моделирования остатков процесса Орнштейна–Уленбека

В работе [11] авторы приводят быстрые и эффективные методики для симуляции различных ОУ-процессов, управляемых процессом Леви, а также описание множества различных частных его случаев. Будем рассматривать три типа процессов Леви: винеровский процесс, дисперсионный гамма-процесс (VG), а также нормальный обратный

гауссовский процесс (NIG). Оба последних процесса моделируют тяжелые хвосты и принадлежат классу обобщенных гиперболических распределений. Они часто применяются в финансах и эконометрике (для VG см. [12, 13]), для NIG см. [10, 14–16]).

2.2.1 Дисперсионный гамма-процесс

Определение 2.1. Случайная величина ξ имеет дисперсионное гамма-распределение, если ее плотность распределения имеет вид:

$$f_{\xi}(x) = \int_0^{\infty} \frac{1}{\sigma\sqrt{2\pi g}} \exp\left(-\frac{(x-\theta g)^2}{2\sigma^2 g}\right) \times \frac{g^{1/\nu-1} \exp(-g/\nu)}{\nu^{1/\nu} \Gamma(1/\nu)} dg, \quad x \in \mathbb{R}, \quad (3)$$

где $\Gamma(x)$, $x > 0$, — гамма-функция, а $\sigma > 0$, $\nu > 0$, $\theta \in \mathbb{R}$.

Обозначение: $\xi \sim V(\sigma, \nu, \theta)$.

Определение 2.2. Случайный процесс $V = (V_t)_{t \geq 0}$ с параметрами $\sigma > 0$, $\nu > 0$, $\theta \in \mathbb{R}$, заданный на вероятностном пространстве (Ω, F, \mathbb{P}) со значениями в \mathbb{R} , называется дисперсионным гамма-процессом, если $V_0 \stackrel{p.n.}{=} 0$, V имеет независимые приращения и для любых $s \geq 0$, $t \geq 0$ V имеет стационарные приращения с дисперсионным гамма-распределением (3) с параметрами $\sigma\sqrt{t} > 0$, $\nu/t > 0$ и $t\theta > 0$, т. е.

$$V_{t+s} - V_s \stackrel{d}{=} V_t - V_0 \sim V\left(\sigma\sqrt{t}, \frac{\nu}{t}, t\theta\right).$$

Характеристики дисперсионного гамма-процесса $V = (V_t)_{t \geq 0}$ с параметрами σ , ν и θ представлены в табл. 1.

В работе [13] показано, что плотность дисперсионного гамма-процесса $V = (V_t)_{t \geq 0}$ выражается

аналитически с использованием модифицированной функции Бесселя второго рода с индексом ν .

2.2.2 Нормальный обратный гауссовский процесс

Определение 2.3. Случайная величина η имеет нормальное обратное гауссовское распределение с параметрами α , β , δ и μ ($\eta \sim \text{NIG}(\alpha, \beta, \delta, \mu)$), если ее плотность распределения имеет вид:

$$f_{\eta}(x, \alpha, \beta, \delta, \mu) = \frac{\alpha\delta}{\pi} \exp\left(\delta\sqrt{\alpha^2 - \beta^2} + \beta(x - \mu)\right) \times \frac{K_1\left(\alpha\sqrt{\delta^2 + (x - \mu)^2}\right)}{\sqrt{\delta^2 + (x - \mu)^2}},$$

где $K_1(z) = (1/2) \int_0^{\infty} \exp(-(1/2)z(u + u^{-1})) du$, $z > 0$, — модифицированная функция Бесселя второго рода с индексом 1, $\alpha > 0$, $-\alpha < \beta < \alpha$, $\delta > 0$, $\mu \in \mathbb{R}$, $x > 0$.

Параметры α , β , δ и μ являются параметрами формы, асимметрии, масштаба и расположения соответственно.

Определение 2.4. Случайный процесс $N = (N_t)_{t \geq 0}$ с параметрами α , β , δ и μ , заданный на вероятностном пространстве (Ω, F, P) со значениями в \mathbb{R} , называется нормальным обратным гауссовским процессом, если $N_0 \stackrel{p.n.}{=} 0$, N имеет независимые приращения и для любых $s \geq 0$, $t \geq 0$ N имеет стационарные приращения с нормальным обратным гауссовским распределением:

$$N_{t+s} - N_s \stackrel{d}{=} N_t - N_0 \sim \text{NIG}(\alpha, \beta, \delta t, \mu t)$$

с параметрами $\alpha > 0$, $-\alpha < \beta < \alpha$, $\delta t > 0$ и $\mu t \in \mathbb{R}$.

Плотность нормального обратного гауссовского распределения может быть представлена в аналитической форме.

Характеристики нормального обратного гауссовского распределения представлены в табл. 2.

Таблица 1 Характеристики дисперсионного гамма-процесса $V = (V_t)_{t \geq 0}$

Математическое ожидание	Дисперсия	Асимметрия	Эксцесс
θt	$(\sigma^2 + \nu\theta^2) t$	$\frac{\theta\nu(3\sigma^2 + 2\nu\theta^2)}{t^{1/2}} (\sigma^2 + \nu\theta^2)^{3/2}$	$3\left(1 + \frac{2\nu}{t} - \nu\theta^4 t (\sigma^2 + \nu\theta^2)^{-2}\right)$

Таблица 2 Характеристики нормального обратного гауссовского распределения

Математическое ожидание	Дисперсия	Асимметрия	Эксцесс
$\mu + \delta\tau$	$\frac{\delta^2(1 + \tau^2)}{\xi}$	$\frac{3}{\tau\sqrt{\xi(1 + \tau^2)}}$	$\frac{3}{\xi} \left(1 + 4\frac{\tau^2}{1 + \tau^2}\right)$

3 Оценивание параметров и поиск смен режима в реальном времени

В данной разделе рассмотрено моделирование и описание данных при условии их поступления в режиме реального времени, когда значения выборки данных поступают одно за другим. Специфика данной задачи заключается в высокой скорости поступления данных в ее приложениях и их большом объеме, поэтому любые приводимые алгоритмы должны быть достаточно быстрыми и эффективно использовать компьютерную память.

3.1 Оценивание параметров

Без каких-либо ограничений на компьютерные мощности самым очевидным решением для оценивания параметров было бы на каждом шаге использовать метод наименьших квадратов (OLS). Чтобы удовлетворить необходимость обрабатывать потоковые данные, воспользуемся рекурсивным методом наименьших квадратов (RLS). Данный алгоритм на каждом шаге обновляет рекурсивно оценку параметра θ и ковариационную матрицу $X^T X$ вместо того, чтобы насчитываться с нуля каждый раз. Данный алгоритм и его реализация хорошо известны и могут быть найдены в [17].

3.2 Постановка простейшей смены режима

В реальной жизни некоторые явления могут быть связаны отношениями, например линейными, параметры которых изменяются во времени. Самым простым подобным примером является процесс, описываемый следующим образом:

$$M_t = \begin{cases} M_t^1, & t \leq t^*; \\ M_t^2, & t > t^*, \end{cases}$$

где $t \in [1, \dots, T]$ обозначает время; t^* — критическое значение внешней переменной t , или момент смены режима (change point, regime switch); $M^{1,2}$ — это две различные модели, соответствующие разным временным промежуткам: до и после. В общем случае нельзя точно определить значение t^* . Задача состоит в том, чтобы наилучшим образом оценить ее значение, имея на входе выборку наблюдений, при условии что на данном временном промежутке произошла ровно одна смена режима (рис. 3), а также оценить параметры старой и новой модели. В данном случае рассматривается модель ОУ и исследуемый процесс выглядит следующим образом:

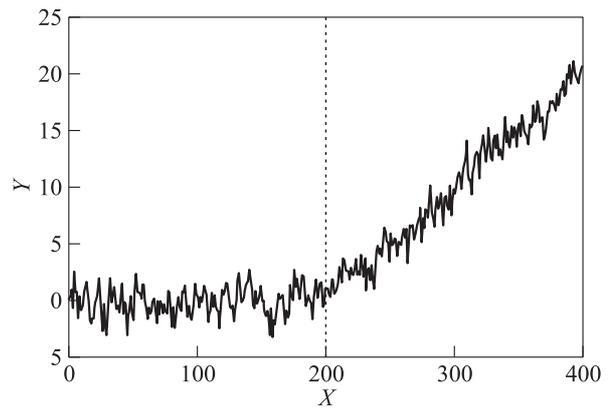


Рис. 3 Пример смены режима

$$X_t = \begin{cases} OU_t^1, & t \leq t^*; \\ OU_t^2, & t > t^*, \end{cases}$$

где OU^i — процессы ОУ, описанные выше.

3.3 Постановка задачи для потоковых данных

В случае потока данных значения выборки $X_1, X_2, \dots, X_n, \dots$ поступают последовательно. В этом случае нет предпосылок для того, чтобы в какой-то момент произошла смена режима, а сами смены могут происходить последовательно много раз. Задача состоит в том, чтобы последовательно их обнаруживать и предоставлять оценку для параметров новой модели. Эффективность методов определяется тем, как часто метод ошибочно определяет режимы и как быстро он способен обнаруживать смену режима.

3.4 Решение задачи

В современной литературе можно найти множество методов для определения смен режима. Основным применяемым подходом является построение по наблюдаемой системе некоторого детектора (change-detector), который сигнализирует, когда параметры модели перестают соответствовать выборке наблюдений и предположительно сменился режим. Одним из таких методов является CUSUM-тест, или метод кумулятивных сумм, который рассматривается в данной работе для обнаружения смен режима. Стоит отметить другие известные методы, такие как метод GLT (generalized likelihood test) [18] и MLT (marginalized likelihood text) [19].

3.4.1 Краткое описание CUSUM-методов

В данной статье будут рассмотрены два базовых CUSUM-метода: CUSUM-метод, основанный на максимизации правдоподобия выборки наблюдений, и CUSUM-метод, определяющий смену среднего значения выборки наблюдений.

CUSUM-метод максимизации правдоподобия

Пусть есть некоторый поток данных $X^* = X_1^*, \dots, X_n^*, \dots$, элементы которого являются выборкой независимых одинаково распределенных случайных величин. Обозначим плотность соответствующей случайной величины через $p_\theta(x)$. Предполагается, что в какой-то момент времени r^* может произойти смена режима модели. Это означает, что до r^* в модели действует параметр θ_0 , а после — θ_1 . Введем соответствующие гипотезы: гипотезу «неизменности» модели H_0 (разладки не произошло) и гипотезу H_1 об «одноразовой разладке». Одним из самых известных и простых методов для нахождения разладки модели является тест отношения правдоподобий [20].

Алгоритм 3.1. Определим логарифмическое отношение правдоподобий для моделей H_0 и H_1 :

$$\text{LLR}(X, r^*) = \ln \frac{p_{H_1}(X)}{p_{H_0}(X)}.$$

Тогда принимается гипотеза H_1 , если $\text{LLR} > h$, где параметр h отвечает за чувствительность алгоритма: чем меньше h , тем быстрее будет происходить обнаружение разладки, но при этом тем больше будет срабатывать ложных сигналов.

К сожалению, в данном случае неизвестно значение r^* и поэтому явно посчитать $p_{H_1}(X)$ не представляется возможным. Данную проблему можно решить, перебрав все значения r^* и взяв то, которому соответствует максимальное значение LLR. Данный метод называется обобщенным методом максимального правдоподобия (GLT).

Алгоритм 3.2. Определим обобщенное логарифмическое отношение правдоподобий для выборки размера N :

$$\begin{aligned} \text{GLLR}(X) &= \max_{1 \leq r^* \leq N} \text{LLR}(X, r^*) = \\ &= \max_{1 \leq r^* \leq N} \ln \frac{p_{H_1}(X)}{p_{H_0}(X)} = \max_{1 \leq r^* \leq N} \sum_{i=r^*}^N \ln \frac{p_{\theta_1}(X_i)}{p_{\theta_0}(X_i)}. \end{aligned}$$

Тогда на каждом шаге n принимается гипотеза H_1 против гипотезы H_0 , если $\text{GLLR}(X) > h$.

Введем кумулятивную сумму точечных отношений правдоподобий:

$$S(n) = \sum_{i=1}^n \ln \frac{p_{\theta_1}(X_i)}{p_{\theta_0}(X_i)}.$$

Тогда

$$\begin{aligned} \text{LLR}(X, r^*) &= S(N) - S(r^*); \\ \text{GLLR}(X, N) &= S(N) - \min_{1 \leq r^* \leq N} S(r^*), \end{aligned}$$

где

$$\hat{r}^* = \operatorname{argmin}_{1 \leq r^* \leq N} S(r^*).$$

Заметим, что для того чтобы использовать алгоритм 3.2, достаточно считать кумулятивную сумму S . Более того, так как уровень h берется положительным, вместо того, чтобы явно насчитывать значение GLLR на каждом шаге, достаточно рекурсивно считать функцию

$$G(N) = \max \left(G(N-1) + \ln \frac{p_{\theta_1}(X_N)}{p_{\theta_0}(X_N)}, 0 \right),$$

которая совпадает с GLLR там, где последняя положительна. В этом и заключается метод кумулятивных сумм. Из вышесказанного вытекает следующий алгоритм, эквивалентный алгоритму 3.2:

Алгоритм 3.3. На каждом шаге N принимается гипотеза H_1 против гипотезы H_0 , если $G(N) > h$.

На практике значение θ_0 можно оценить, а значение θ_1 неизвестно. Поэтому берут $\theta_1 = \theta_0 + \delta$, где δ — минимальная величина, изменение которой хотят детектировать.

CUSUM-метод определения смены среднего выборки

Алгоритм 3.4. Определим рекурсивно

$$S_n^+ = \max \left(S_{n-1}^+ + \frac{X_n - \mu_0}{\sigma} - k, 0 \right),$$

где μ_0 — среднее текущей модели; σ — среднее текущей выборки; k — уровень чувствительности метода к разбросам. Тогда считается, что принимается гипотеза H_1 , если $S_n^+ > h$.

Подробнее с алгоритмом можно ознакомиться, например, в [21].

3.5 Общий алгоритм для процесса Орнштейна–Уленбека

Как было замечено ранее, приведенный процесс ОУ с трендом обладает авторегрессионным свойством AR(1). Поэтому можно применять CUSUM-метод максимального правдоподобия для приращений l_i данного процесса. Данный метод будем

применять для детектирования смены волатильности модели, в то время как для обнаружения смены среднего, или тренда, будем применять CUSUM-метод поиска смены среднего. Применить первый алгоритм для детектирования среднего оказывается сложно из-за большого числа параметров, которые нельзя адекватно оценить, в частности параметров μ_0 и ν .

Таким образом, общий алгоритм следующий:

Алгоритм 3.5.

1. На каждом шаге оцениваем наиболее вероятные параметры выборки θ_0 для выбранной модели с помощью метода RLS.
2. На каждом шаге считаем значения детекторов CUSUM смены волатильности и смены тренда.
3. В случае, когда детекторы сигнализируют о смене режима, проходимся общим методом обобщенного отношения правдоподобий (GLT) по выборке и находим наиболее вероятную точку смены режима r^* с оценкой параметров θ_1 . Далее исключаем из выборки все ее элементы до r^* и продолжаем процедуру алгоритма с $\theta_0 := \theta_1$.

4 Анализ данных

В этом разделе описывается моделирование процессом ОУ реальных финансовых данных. В качестве данных были выбраны секундные данные по ценам фьючерсов на индекс RTS (x_t) и на акции компании «Газпром» (y_t) с MOEX за 07.10.2014. Если наблюдения сделаны через равные промежутки времени, то можно рассматривать их как временной ряд. Предполагается, что разность $z_t = x_t - \beta y_t$ обладает свойством стационарности и может быть описана с помощью процесса ОУ. Для того чтобы

ряд имел свойство авторегрессии, вычитаем из ряда его скользящее среднее с периодом 5 мин.

Тест Дики–Фуллера (с уровнем значимости 0,05) подтверждает предположение о стационарности: значение статистики Дики–Фуллера: $-25,374$; p -значение: 0,001. Тест отвергает нулевую гипотезу о существовании единичного корня с уровнем значимости 0,05, что подтверждает стационарность данного ряда. Для проверки наличия свойства AR(1) проанализируем вид автокорреляционной (ACF) и частичной автокорреляционной (PACF) функций.

Для модели AR(1) характерен следующий вид автокорреляционных и частичных автокорреляционных функций: график ACF экспоненциально убывает, а график PACF имеет пик при значении сдвига, равном 1, и практически равен 0 при значениях сдвига более высокого порядка.

На рис. 4 изображен график PACF для рассматриваемых данных. Заметно, что он имеет пик при сдвиге 1 и практически равен нулю для сдвигов более высокого порядка. На рис. 5 ACF для исходных данных убывает экспоненциально. Такое поведение графиков ACF и PACF соответствует модели AR(1).

Теперь можно говорить о том, что данные представляют собой процесс ОУ, и перейти к оценке его параметров.

Для начала оценим параметры θ и α с помощью метода наименьших квадратов, получаем оценки параметров $\hat{\theta} = 0$ и $\hat{\alpha} = 0,7506$.

Для того чтобы оценить качество полученных оценок для данного процесса, построим QQ-плот для оцененных параметров нормального распределения для остатков (рис. 6). Это график, где по оси x — квантили теоретического распределения, а по оси y — эмпирические квантили данных. Если теоретическое распределение хорошо описывает

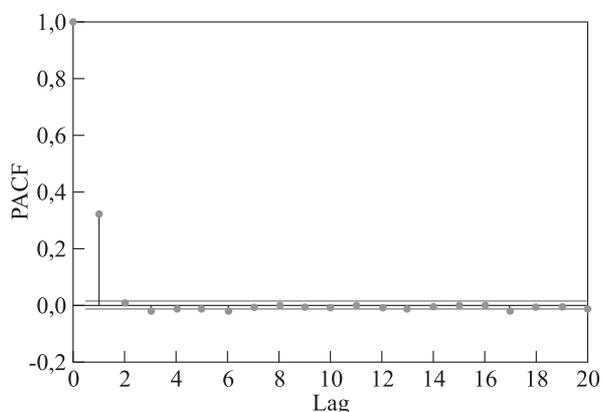


Рис. 4 Частичная автокорреляционная функция

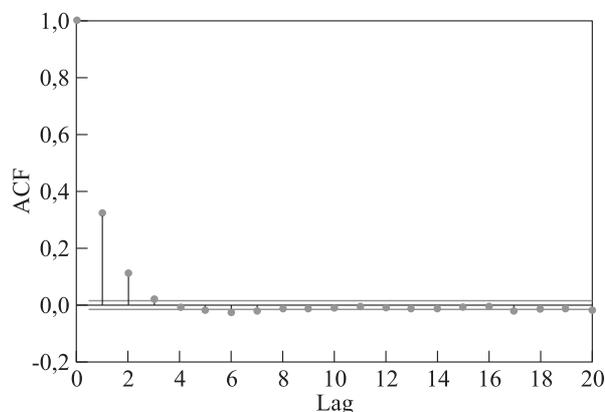


Рис. 5 Автокорреляционная функция

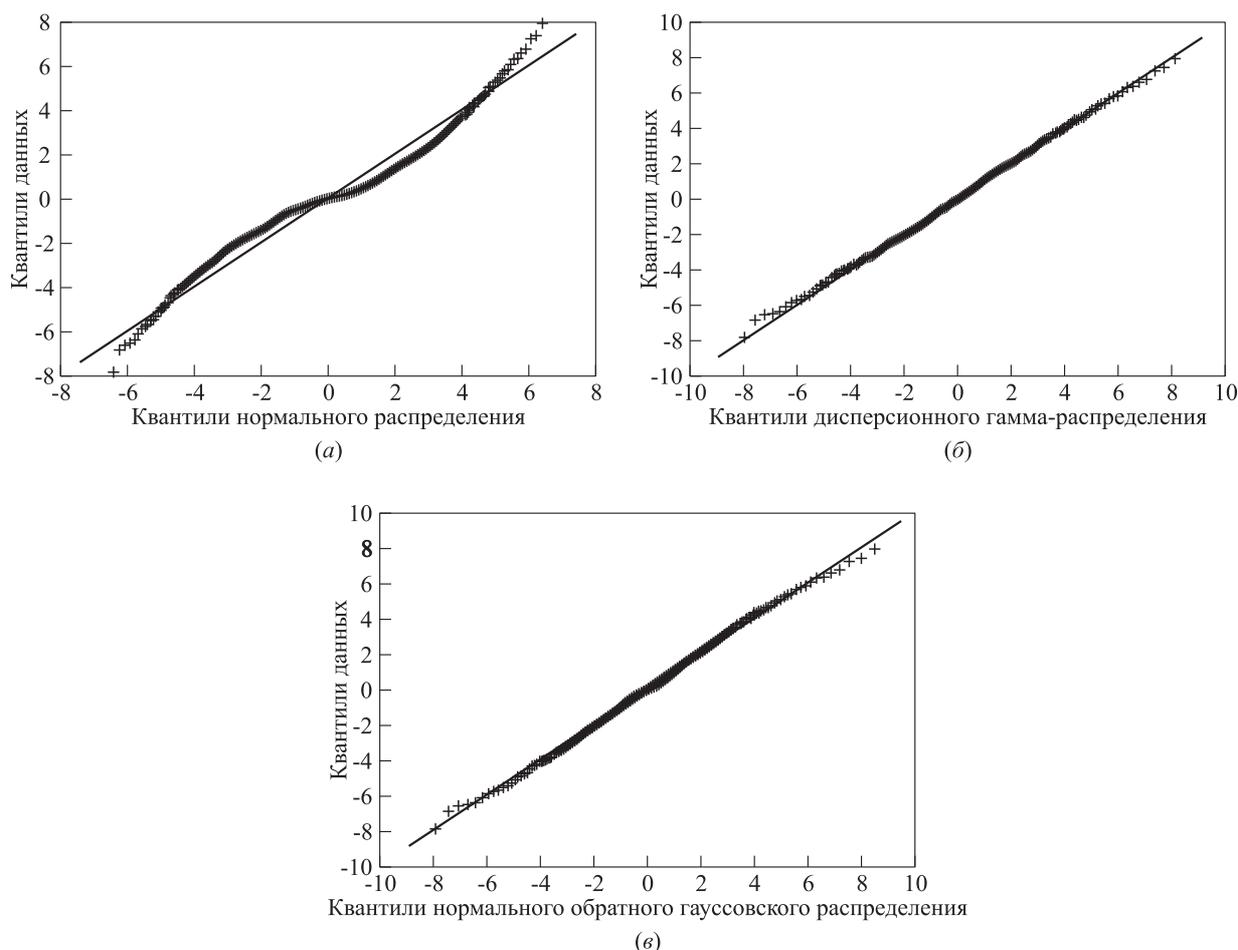


Рис. 6 Графики QQ-plot для оценивания параметров

реальные данные, то график «квантиль–квантиль» близок к прямой $y = x$.

Видно, что нормальное распределение не очень хорошо описывает распределение остатков. Попробуем вместо нормального распределение использовать распределение с более тяжелыми хвостами, например дисперсионное гамма-распределение и нормальное обратное гауссовское распределение. Для оценки параметров этих распределений используем метод максимального правдоподобия, описанный в п. 3.2.3.

Анализируя графики QQ-plot (см. рис. 6) для оцененных параметров дисперсионного гамма- и нормального обратного гауссовского распределений для остатков, можно прийти к выводу, что дисперсионное гамма- и нормальное обратное гауссовское распределение лучше описывают структуру независимых приращений в процессе ОУ.

Для того чтобы оценить качество полученных результатов, применим критерий согласия Колмогорова для новой выборки данных. Результат подсчета статистики представлен в табл. 3.

По результатам, представленным в табл. 3, можно заключить, что гипотеза о нормальном распределении остатков отвергнута при уровне значимости 0,01, гипотеза о дисперсионном гамма-распределении остатков и нормальном обратном гауссовском распределении остатков принята при уровне значимости 0,01.

Таблица 3 Оценки параметров с результатом критерия Колмогорова

Оценка параметра		Значение статистики	p-значение
$N(\mu, \sigma^2)$	$\hat{\mu}$	0	0,087731
	$\hat{\sigma}$	3,6885	
VG (σ, ν, θ)	$\hat{\sigma}$	3,6712	0,026806
	$\hat{\nu}$	1,5226	
	$\hat{\theta}$	0,0379	
NIG (θ, ξ, δ, μ)	$\hat{\theta}$	0,0592	0,034654
	$\hat{\xi}$	0,3714	
	$\hat{\delta}$	2,2690	
	$\hat{\mu}$	-0,0963	

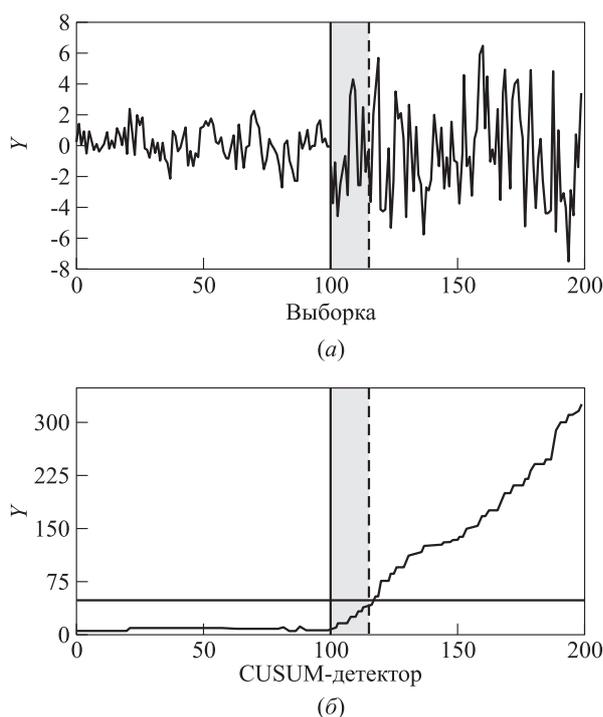


Рис. 7 Пример применения CUSUM-алгоритма

Полученный результат говорит о том, что структура реальных данных сложнее, чем может описать классический процесс ОУ. Целесообразнее использовать обобщенный процесс ОУ, где процесс броуновского движения заменен на процесс Леви.

4.1 Применение алгоритма для детектирования изменения волатильности

Будем рассматривать гауссовский процесс ОУ. Для этого были сгенерированы две выборки процесса размера 100 с $\sigma_1 = 1$ и $\sigma_2 = 3$. Для CUSUM-теста будем брать $\theta_1 = 2$, т. е. $\delta = 1$. Уровень $h = 45$. Результат применения алгоритма проиллюстрирован на рис. 7. На рис. 7, а изображена выборка сгенерированного процесса ОУ со сменой режима, в то время как на рис. 7, б построено значение CUSUM-детектора. Сплошная вертикальная линия обозначает фактическую смену режима, а пунктирная — время ее обнаружения. Заметим, что смена режима могла бы быть обнаружена быстрее при другом выборе уровня h .

4.2 Применение алгоритма для обнаружения тренда

Для обнаружения тренда также были сгенерированы две выборки гауссовского процесса ОУ, ко-

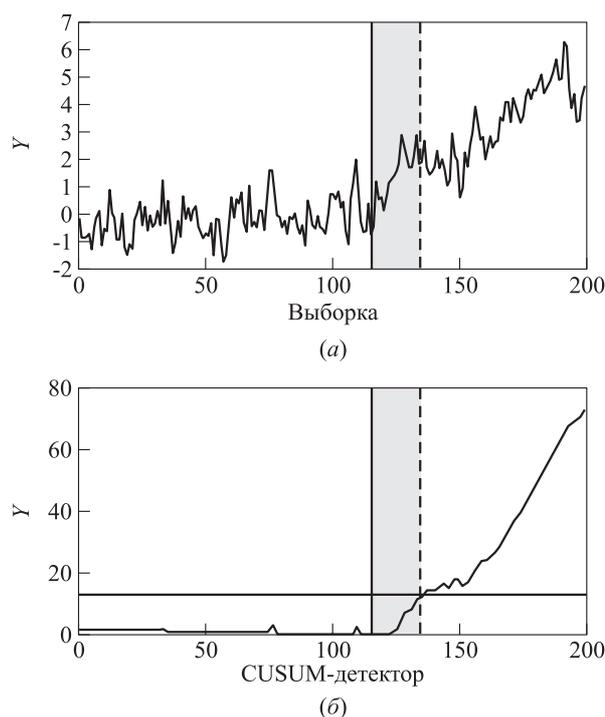


Рис. 8 Пример применения CUSUM-алгоритма

торые потом были склеены. Параметры процессов следующие: $\alpha_0 = 0,5$, $\nu_0 = 0$, $\mu_0^0 = 0$, $\sigma_0 = 0,6$, $\alpha_1 = 0,5$, $\nu_1 = 0,05$, $\mu_0^1 = 0$ и $\sigma_1 = 0,6$. Аналогично построена выборка и значения CUSUM-детектора. Уровень $h = 13$. Алгоритм успешно определил смену режима (рис. 8).

5 Заключение

В статье рассмотрен процесс ОУ с трендом, управляемый процессом Леви, для описания финансовых временных рядов. На реальных данных было показано, что дисперсионный гамма- и нормальный обратный гауссовский процессы в качестве процесса Леви способны гораздо точнее описывать реальные явления. Также были рассмотрены проблемы разладки модели и поиска смены режима в реальном времени. Была представлена процедура обнаружения разладки модели, а также определения параметров новой модели. Данный алгоритм способен детектировать многократные смены режима последовательно, сохраняя текущую модель актуальной для текущего потока данных.

Литература

1. Brigo D., Dalessandro A., Neugebauer M., Triki F. A stochastic processes toolkit for risk management. —

- London: King's College London, November 2007. Working paper. 48 p.
2. Ornstein L. S., Uhlenbeck G. E. On the theory of the Brownian motion // *Phys. Rev.*, 1930. Vol. 36. No. 5. P. 823.
 3. Vasicek O. An equilibrium characterization of the term structure // *J. Financ. Econ.*, 1977. Vol. 5. P. 177.
 4. Cox J. C., Ingersoll E., Jr., Ross S. A. A theory of the term structure of interest rates // *Econometrica*, 1985. Vol. 53. No. 2. P. 385–407.
 5. Garbaczewski P., Olkiewicz R. Ornstein–Uhlenbeck–Cauchy process // *J. Math. Phys.*, 2000. Vol. 41. P. 6843.
 6. Finlay R. The variance gamma (VG) model with long range dependence: A model for financial data incorporating long range dependence in squared returns. — Sydney, Australia: University of Sydney, School of Mathematics and Statistics, 2009. PhD Thesis. 144 p.
 7. Кузьмина А. В. Моделирование нормального обратного гауссовского процесса и оценивание его параметров // *Информатика*, 2011. № 2. С. 133–136.
 8. Protter P. Stochastic integration and differential equations. — Heidelberg: Springer-Verlag, 1990. 415 p.
 9. Sato K. I. Lévy processes and infinitely divisible distributions. — Cambridge: Cambridge University Press, 1999. 500 p.
 10. Barndorff-Nielsen O. E., Shephard N. Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics // *J. Roy. Stat. Soc. B*, 2001. Vol. 63. P. 167–241.
 11. Tauffer E., Leonenko N. Simulation of Lévy-driven Ornstein–Uhlenbeck processes with given marginal distribution // *Comput. Stat. Data An.*, 2008. Vol. 53. P. 2427–2437.
 12. Madan D. B., Seneta E. The VG model for share market returns // *J. Bus.*, 1990. Vol. 63. P. 511–524.
 13. Madan D. B., Carr P. P., Chang E. C. The variance gamma process and option pricing // *Eur. Finance Rev.*, 1998. Vol. 2. P. 79–105.
 14. Barndorff-Nielsen O. E. Normal inverse Gaussian distributions and stochastic volatility modelling // *Scand. J. Stat.*, 1997. Vol. 24. No. 1. P. 1–13.
 15. Rydberg H. The Normal inverse Gaussian Lévy process: Simulation and approximation // *Commun. Stat. Stochastic Models*, 1997. Vol. 13. No. 4. P. 887–910.
 16. Barndorff-Nielsen O. E. Processes of normal inverse Gaussian type // *Financ. Stoch.*, 1998. Vol. 2. P. 41–68.
 17. Haykin S. Adaptive filter theory. — 3rd ed. — Upper Saddle River, NJ, USA: Prentice Hall, 1996. 989 p.
 18. Appel U., Brandt A. V. Adaptive sequential segmentation of piecewise stationary time series // *Inform. Sci.*, 1983. Vol. 29. P. 27–56.
 19. Gustafsson F. The marginalized likelihood ratio test for detecting abrupt changes // *IEEE Trans. Automat. Contr.*, 1996. Vol. 41. P. 66–78.
 20. Kay S. Fundamentals of statistical signal processing. Vol. I. Estimation theory. — Upper Saddle River, NJ, USA: Prentice Hall, 1993. 625 p.
 21. Page E. S. Cumulative sum control chart // *Technometrics*, 1961. Vol. 3. P. 1–9.

Поступила в редакцию 20.10.16

REGIME SWITCHING DETECTION FOR THE LEVY DRIVEN ORNSTEIN–UHLENBECK PROCESS USING CUSUM METHODS

A. V. Chertok^{1,2}, A. I. Kadaner^{2,3}, G. T. Khazeeva¹, and I. A. Sokolov⁴

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

²Sberbank of Russia, 19 Vavilov Str., Moscow 117999, Russian Federation

³Faculty of Mechanics and Mathematics, M. V. Lomonosov Moscow State University, Main Building, Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation

⁴Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

Abstract: The article considers using a trending Ornstein–Uhlenbeck process, driven by a Levy process, for modeling financial time series. The authors demonstrate that the Levy driven model gives more flexibility to describe financial time series than the simple classical model. In particular, the Levy driven model allows modeling distributions with heavy tails, which is a common property of time series in real applications. The authors describe efficient methods for estimating model parameters using such methods as OLS (ordinary least squares) and RLS (regularized least squares). The article also solves the regime switching problem in a real time data stream. The authors built an algorithm based on CUSUM (CUmulative SUM) methods that is capable of determining regime switches consecutively as they happen online and keep model parameters up to date. Solution of the regime switching problem is important in real applications, since the dynamics of real systems tend to change over time under the influence of external factors.

Keywords: random process; mean-reverting process; Ornstein–Uhlenbeck process driven by Levy process; trending Ornstein–Uhlenbeck process; regime switch; change point detection; CUSUM algorithm

DOI: 10.14357/19922264160405

Acknowledgments

The research was partially supported by the Russian Foundation for Basic Research (project 14-07-00041).

References

1. Brigo, D., A. Dalessandro, M. Neugebauer, and F. Triki. 2007. A stochastic processes toolkit for risk management. London: King's College London. Working paper. 48 p.
2. Ornstein, L. S., and G. E. Uhlenbeck. 1930. On the theory of the Brownian motion. *Phys. Rev.* 36(5):823.
3. Vasicek, O. 1977. An equilibrium characterization of the term structure. *J. Financ. Econ.* 5(2):177–188.
4. Cox, J. C., E. Ingersoll, Jr., and S. A. Ross. 1985. A theory of the term structure of interest rates. *Econometrica* 53(2):385–407.
5. Garbaczewski, P., and R. Olkiewicz. 2000. Ornstein–Uhlenbeck–Cauchy process. *J. Math. Phys.* 41:6843–6860.
6. Finlay, R. 2009. The variance gamma (VG) model with long range dependence: A model for financial data incorporating long range dependence in squared returns. Sydney, Australia: University of Sydney, School of Mathematics and Statistics. PhD Thesis. 144 p.
7. Kuzmina, A. V. 2011. Modelirovanie normal'nogo obratnogo gaussovskogo protsesssa i otsenivanie ego parametrov [Normal inverse Gaussian distribution modeling and its parameters estimation]. *Vestnik Belorusskogo gosudarstvennogo universiteta. Ser. 1: Fizika. Matematika. Informatika* [Herald of the Belarusian State University. Ser. 1: Physics. Mathematics. Informatics] 2:133–136.
8. Protter, P. 1990. *Stochastic integration and differential equations*. Heidelberg: Springer-Verlag. 415 p.
9. Sato, K. I. 1999. *Lévy processes and infinitely divisible distributions*. Cambridge: Cambridge University Press. 500 p.
10. Barndorff-Nielsen, O. E., and N. Shephard. 2001. Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics. *J. Roy. Stat. Soc. B* 63:167–241.
11. Taufer, E., and N. Leonenko. 2008. Simulation of Lévy-driven Ornstein–Uhlenbeck processes with given marginal distribution. *Comput. Stat. Data An.* 53:2427–2437.
12. Madan, D. B., and E. Seneta. 1990. The VG model for share market returns. *J. Bus.* 63:511–524.
13. Madan, D. B., P. P. Carr, and E. C. Chang. 1998. The variance gamma process and option pricing. *Eur. Finance Rev.* 2:79–105.
14. Barndorff-Nielsen, O. E. 1997. Normal inverse Gaussian distributions and stochastic volatility modeling. *Scand. J. Stat.* 24(1):1–13.
15. Rydberg, H. 1997. The normal inverse Gaussian Lévy process: Simulation and approximation. *Commun. Stat. Stochastic Models* 13(4):887–910.
16. Barndorff-Nielsen, O. E. 1998. Processes of normal inverse Gaussian type. *Financ. Stoch.* 2:41–68.
17. Haykin, S. 1996. *Adaptive filter theory*. 3rd ed. Upper Saddle River, NJ: Prentice Hall. 989 p.
18. Appel, U., and A. V. Brandt. 1983. Adaptive sequential segmentation of piecewise stationary time series. *Inform. Sci.* 29:27–56.
19. Gustafsson, F. 1996. The marginalized likelihood ratio test for detecting abrupt changes. *IEEE Trans. Automat. Contr.* 41:66–78.
20. Kay, S. 1993. *Fundamentals of statistical signal processing. Vol. I. Estimation theory*. Upper Saddle River, NJ: Prentice Hall. 625 p.
21. Page, E. S. 1961. Cumulative sum control chart. *Technometrics* 3:1–9.

Received October 20, 2016

Contributors

Chertok Andrey V. (b. 1987) — junior scientist, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; Head of R&D, Data Science, Sberbank of Russia, 19 Vavilov Str., Moscow 117999, Russian Federation; avchertok.sbt@sberbank.ru

Kadaner Arsenii I. (b. 1995) — student, Faculty of Mechanics and Mathematics, M. V. Lomonosov Moscow State University, Main Building, Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; data scientist, Sberbank of Russia, 19 Vavilov Str., Moscow 117999, Russian Federation; aikadaner.sbt@sberbank.ru

Khazeyeva Gelana T. (b. 1993) — student, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; gelana.khazeyeva@gmail.com

Sokolov Igor A. (b. 1954) — Academician of the Russian Academy of Sciences, Doctor of Science in technology, Director, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; isokolov@ipiran.ru

О РАЗМЕЩЕНИИ ЗАДАНИЙ НА ДВУХ СЕРВЕРАХ ПРИ НЕПОЛНОМ НАБЛЮДЕНИИ*

М. Г. Коновалов¹, Р. В. Разумчик²

Аннотация: Рассматривается задача эффективного распределения заданий в системах обслуживания, состоящих из двух параллельно работающих серверов с очередями неограниченной емкости и одним диспетчером, немедленно распределяющим поступающие задания по серверам. При поступлении очередного задания диспетчер принимает решение, основываясь только на информации о производительности серверов, распределении времени между поступлениями заданий и распределении их размеров. Другая информация о состоянии системы (как, например, длина очередей) диспетчеру недоступна. Для таких ненаблюдаемых систем известно, что минимум среднего времени пребывания задания в системе обеспечивает детерминированная (периодическая) стратегия. Поскольку при принятии решения диспетчер неявно наблюдает и момент поступления задания в систему, возникает следующий вопрос: возможно ли уменьшить среднее время пребывания в системе, если учитывать предысторию моментов поступления заданий? В работе с помощью численных экспериментов дается положительный ответ. Диспетчеризация, которая позволяет это сделать, основывается помимо априорной информации о системе на оценках состояния очереди по результатам имеющихся наблюдений. Достижимый выигрыш составляет от 1,5% до 10% по сравнению с известной оптимальной стратегией и увеличивается при уменьшении дисперсии длины заданий. При малых значениях дисперсии предложенная стратегия является более эффективной, чем диспетчеризация по самой короткой очереди.

Ключевые слова: системы с параллельным обслуживанием; стратегии размещения заданий; диспетчеризация; управление при неполном наблюдении

DOI: 10.14357/19922264160406

1 Введение

Во многих технических, информационных и других системах имеются параллельно работающие обслуживающие ресурсы (серверы), на которые поступают заявки на выполнение заданий. При этом возникает проблема построения оптимальной процедуры выбора сервера для выполнения очередного задания (проблема диспетчеризации). Оптимальность может пониматься в разных смыслах, однако для распределенных компьютерных систем передачи, обработки и хранения данных наиболее популярным критерием оптимальности является среднее время пребывания задания в системе (активно используется также среднее значение для отношения времени выполнения задания к его размеру (в зарубежной литературе — *slow-down*)). Проблема диспетчеризации является сложной, и относящиеся к ней задачи чаще всего не удается решить исключительно математическими методами. Поэтому нередким явлением стало применение методов статистического моделирования,

эвристических идей и инженерного подхода для получения удовлетворительных решений. К настоящему времени по данной проблематике известно очень большое число работ теоретического и прикладного характера (см., например, обзоры [1, 2]). Несмотря на это, круг нерешенных или не вполне решенных вопросов остается широким и, как свидетельствуют регулярно появляющиеся публикации в ведущих периодических научных изданиях, исследования в данной области актуальны [3–5].

Следуя [3], напомним, что оптимальная процедура в рассматриваемых системах существенным образом зависит от:

- типа и количества информации, доступной при принятии решения;
- правил обработки заданий внутри серверов;
- критерия оптимальности.

В данной работе рассматривается проблема диспетчеризации в следующих предположениях:

* Работа выполнена при частичной поддержке РФФИ (проект 15-07-03406).

¹ Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, mkonov@ipiran.ru

² Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, Российский университет дружбы народов, grazumchik@ipiran.ru

- (1) при принятии решения доступна априорная информация о системе (распределение времени между поступлениями заданий, распределение размера заданий, производительности серверов);
- (2) наблюдаются моменты поступления заданий и известны принятые в эти моменты решения;
- (3) серверы работают параллельно и независимо друг от друга, переход заявок между серверами невозможен;
- (4) задания внутри сервера обслуживаются в порядке поступления;
- (5) критерием оптимальности является минимум среднего времени пребывания задания в системе.

Таким образом, в данной статье речь пойдет об оптимальных с точки зрения минимального среднего времени пребывания стратегиях в «почти» ненаблюдаемых системах параллельной обработки заданий. Основная отличительная особенность постановки задачи в том, что недоступна динамическая информация о состоянии системы (например, о числе заданий на серверах, об остаточных временах обслуживания, о размерах заданий и др.). С практической точки зрения данные системы находят применение при моделировании систем облачных вычислений, grid-систем¹ и ряда других систем (см., например, [6–8]).

2 Постановка задачи

Проблема диспетчеризации порождает большое число задач, многие из которых формулируются в рамках следующей схемы.

В систему из двух серверов поступает рекуррентный входной поток заданий, так что интервалы времени между поступлениями заданий образуют последовательность независимых случайных величин с одинаковым распределением. Задания имеют случайную длину (объем), которая определяется одним и тем же распределением. Каждое поступившее задание должно быть немедленно направлено на один из серверов. Каждый сервер имеет неограниченный бункер (очередь) для хранения заданий. Выполнение заданий происходит в порядке поступления, без перерывов между окончанием выполнения одного задания и началом следующего (при наличии непустой очереди). Серверы работают независимо, без обмена заданиями. Описанная схема приводит к следующей математической модели.

Обозначим через F функцию распределения длины интервала между последовательными поступлениями заданий, а через G — функцию распределения длины задания. Производительность (интенсивность обслуживания) одного из серверов полагаем равной 1. Производительность другого сервера (более быстрого) обозначим через $v (> 1)$. Серверы и связанные с ними величины будем индексировать цифрами 0 (более медленный сервер) и 1.

Система функционирует в непрерывном времени $t \geq 0$. Пусть $0 \leq t_1 < \dots < t_n < \dots$ — последовательность моментов поступления заданий в систему, а $\Delta_n = t_{n+1} - t_n$ — промежутки между этими моментами. Обозначим через x_n длину задания, поступившего в момент t_n . Обозначим также через $u_n^{(i)}$ время, необходимое для выполнения всех заданий, имеющихся на сервере $i (= 0 \vee 1)$ к моменту $t = t_n - 0$ (т.е. в момент, «непосредственно предшествующий» поступлению в систему очередного задания). Уточним, что величины $u_n^{(i)}$ включают, в том числе, время, необходимое для завершения заданий, уже находящихся в стадии непосредственного выполнения. Назовем эти величины «остаточными временами». Решение (действие), принимаемое в момент t_n относительно вновь поступившего задания, обозначим через y_n . Полагаем, что $y_n = 0$, если задание направлено на сервер 0, и $y_n = 1$, если решение противоположное. Очевидно, справедливы соотношения:

$$u_{n+1}^{(0)} = \left(u_n^{(0)} + (1 - y_n)x_n - \Delta_n \right)^+;$$

$$u_{n+1}^{(1)} = \left(u_n^{(1)} + y_n \frac{x_n}{v} - \Delta_n \right)^+,$$

где $a^+ = \max(a, 0)$.

Задание, поступившее в момент t_n и обслуженное согласно правилу y_n , проведет в системе время, равное

$$w_n = u_n^{(y_n)} + \frac{x_n}{1 - y_n + y_n v},$$

где первое слагаемое — это время, проведенное в очереди, а второе — время непосредственного выполнения. Таким образом, возникает общая задача отыскания такой стратегии $y = \{y_n, n = 1, 2, \dots\}$, которая минимизировала бы предельное среднее время пребывания заданий в системе, определяемое как

$$W = W(y; F, G) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N E_y w_n,$$

где E_y — интегрирование по мере, порождаемой последовательностью y , а существование предела

¹Строящихся на их основе систем добровольных вычислений (volunteer computing).

обусловлено структурой этой последовательности, которая обычно выбирается из «разумных соображений» достаточно регулярной. Стратегия y называется диспетчеризацией, а ее элементы y_n — правилами диспетчеризации.

Сформулированная общая постановка задачи фактически распадается на ряд задач в зависимости от того, на каком множестве допустимых диспетчеризаций осуществляется минимизация. В свою очередь, задание множества допустимых стратегий можно интерпретировать с точки зрения возможностей наблюдения за траекторией процесса. Если обозначить через h_n совокупность наблюдаемых параметров системы до момента принятия решения t_n , то допустимое правило диспетчеризации имеет вид:

$$y_n = f(h_n),$$

где f — рандомизированная или детерминированная функция со значениями 0 или 1, а h_n принимает значения из некоторого множества наблюдений H_n . Наиболее употребительными и исследованными являются следующие четыре варианта.

1. Крайний случай, приводящий к наиболее бедному множеству допустимых стратегий — это отсутствие вообще каких-либо наблюдений, $H_n = \emptyset$. В этом случае допустимая стратегия¹ и описывается одним-единственным параметром — вероятностью p выбора для очередного задания сервера 0. Известен ряд результатов, касающихся ее оптимальности, используя которые можно численно находить значения вероятностей p (см., например, [9–11]). Наиболее полно задача решается для полностью марковских систем и систем с входящим пуассоновским потоком заданий и серверов типа $M/G/1$. В общих случаях (например, когда система состоит из серверов типа $G/G/s$ и $MAP/G/s$, $s \geq 1$), известны различные приближенные и эвристические решения, которые получены с использованием математического программирования. Важным обстоятельством, которое позволяет упростить решение оптимизационной задачи в случае вероятностной стратегии, является то, что если поток поступающих в систему заявок является рекуррентным, то и «прожженный» поток на каждый сервер также явля-

ется рекуррентным. Тогда система из нескольких серверов декомпозируется на несколько систем из одного сервера, для которых можно использовать известные точные или приближенные формулы.

2. Больше разнообразие в выборе диспетчеризации получается, если допустить возможность наблюдения за траекторией принятых решений, что приводит к допустимым правилам диспетчеризации вида:

$$y_n = f(y_{n-k_n}, \dots, y_{n-1}),$$

где $1 \leq k_n \leq n-1$. В этом случае предыстория к моменту t_n определяется значением из множества $H_n = Y_k = \{0; 1\}^{k_n}$, где число k_n характеризует глубину предыстории, используемую в момент t_n . Большой интерес представляют программные (детерминированные) стратегии, т.е. стратегии, параметризуемые бесконечными последовательностями $\{a_1, a_2, \dots, a_{n-1}, a_n, \dots\}$, в которых a_i означает, что a -я заявка направляется на сервер с номером i . Внимание к ним связано с интуитивным представлением о том, что входящий поток на каждый сервер при детерминированной стратегии является более регулярным (т.е. «менее» случайным), чем при вероятностной, что может приводить к уменьшению значения среднего времени пребывания в системе (см., например, [12, 13]). В общем случае (для произвольного числа серверов) нахождение оптимальной детерминированной стратегии является сложной задачей. Если детерминированная стратегия предписывает направлять на каждый из $N \geq 2$ серверов заявки в соответствии с распределением (p_1, \dots, p_N) , то из [14] известно, что оптимальными являются так называемые сбалансированные последовательности² [15]. Однако сбалансированные последовательности для произвольного распределения (p_1, \dots, p_N) существуют лишь в редких случаях. Такими случаями являются система из произвольного числа одинаковых серверов³ (здесь оптимальной является циклическая стратегия⁴) и случай двух серверов (здесь при рациональном значении p оптимальной является так называемая последовательность Битти⁵). Заметим, что в послед-

¹В зарубежной литературе — PAP (Probabilistic Allocation Policy), или RND (Random), или BS (Bernoulli splitting).

²Если последовательность состоит только из нулей и единиц, то она называется сбалансированной, если число единиц в любых двух произвольно выделенных подпоследовательностях фиксированной длины отличается не более чем на 1.

³Например, серверов типа $\bullet/G/1$.

⁴ n -е поступающее задание направляется на сервер с номером $(n \bmod R) + 1$.

⁵В литературе встречаются и другие названия — последовательность Штурма, бильярдная последовательность (см. подробнее в [16, 17]). Заметим, что для реализации такой диспетчеризации вообще не требуются наблюдения за траекторией принятых решений, а для определения нужного сервера необходимо знать лишь порядковый номер поступающей заявки.

нем случае оптимальность стратегии зависит от значения p и способа нахождения точного значения до сих пор не предложено (см., например, [18, 19]). Тем не менее простой эвристический подход к нахождению значения p приводит к значениям целевой функции, которые не удается уменьшить, не привлекая при диспетчеризации дополнительную информацию о системе. При $N \geq 3$ сбалансированную последовательность удастся построить лишь в частных случаях (см. подробнее в [3, 15]). Поэтому действуют по-другому: для заданного распределения (p_1, \dots, p_N) ищут детерминированную последовательность, расстояние¹ которой от сбалансированной последовательности было бы минимальным. Эта задача является комбинаторной, и для нее известно несколько численных алгоритмов решения [17, 20, 21]. В наиболее важных случаях (в случаях рациональных значений p_i) результаты работы этих алгоритмов приводят к периодическим последовательностям и последовательностям специального вида — так называемым бильярдным последовательностям (см., например, [17]).

3. Частичное наблюдение за состоянием очереди означает, что постоянно известен ее размер. В этом варианте, как правило, рассматриваются правила диспетчеризации вида

$$y_n = f(l_n),$$

где l_n — размер очереди к моменту принятия решения в момент t_n . Таким образом, $H_n = N = \{0, 1, 2, \dots\}$. Основным примером такого рода является диспетчеризация JSQ (Join the Shortest Queue), направляющая задание на сервер с минимальной очередью. Она является оптимальной в случае пуассоновского входящего потока, экспоненциальных времен обслуживания с одинаковыми параметрами и с дисциплинами FCFS (first come — first served) на всех серверах (подробнее см., например, обзоры [1, 2]).

4. Наиболее полный вариант наблюдений предполагает возможность использовать при выборе сервера значения остаточных времен $u_n^{(0)}$ и $u_n^{(1)}$, а также длину x_n нового задания. В этом случае оптимальная диспетчеризация находится в множестве стратегий вида

$$y_n = f(u_n^{(0)}, u_n^{(1)}, x_n).$$

Среди диспетчеризаций такого вида простым и в то же время достаточно эффективным ре-

шением является стратегия *muорis*, когда вновь поступающее задание посылается на тот сервер, который минимизирует время, необходимое для освобождения от заданий всей системы целиком, в предположении, что в дальнейшем задания в систему не поступают. Более сложная и менее универсальная стратегия *deer* получена с помощью использования приближенного алгоритма из теории марковского процесса принятия решений [22]. Стратегия *deer*, возможно, является «квазиоптимальной» в том смысле, что она дает значение целевой функции, близкое к (неизвестному) оптимальному значению. Однако, несмотря на формальное сведение задачи диспетчеризации к марковскому процессу принятия решений в варианте с полным наблюдением, до сих пор не известно, как находить оптимальное значение функционала.

Следует подчеркнуть, что рассмотренные выше четыре основных примера не исчерпывают все возможности постановки задачи. Например, для вариантов с неполным наблюдением можно ставить вопрос о том, улучшится ли показатель W , если учитывать предысторию наблюдаемых компонент. Одна из таких постановок рассматривается далее.

3 Диспетчеризация по предыстории действий

Постановка задачи 2 из предыдущего раздела ограничивает доступные наблюдения совершенными действиями. Это весьма жесткие ограничения на допустимые стратегии диспетчеризации, которые влекут существенный проигрыш в целевой функции по сравнению со стратегиями, использующими максимальную текущую информацию. Далее эти ограничения сохраняются, и по-прежнему предполагается, что недоступны наблюдения, отражающие «состояние бункера» — количество заданий, объем нагрузки в очереди, моменты начала и окончания непосредственного выполнения задания. Неизвестной также считается длина поступающего задания. Однако информация о совершенных действиях понимается несколько более широко. Побудительный мотив можно выразить такими словами: зная, «что было сделано», естественно допустить, что известно также, «когда было сделано». Точнее говоря, помимо самих решений y_n предполагается использовать информацию о моментах времени t_n , в которые эти решения принимались. Таким образом, далее предполагается, что допустимыми являются диспетчеризации, правила которых основываются на предыстории принятых

¹Подробнее о том, как определяется расстояние между последовательностями, см. [17].

решений и моментов поступления заданий. Это означает, что допустимым считается правило диспетчеризации, которое представимо в виде (детерминированной или рандомизированной) функции вида:

$$y_n = f(y_1, \dots, y_{n-1}, t_1, \dots, t_n),$$

или

$$y_n = f(y_1, \dots, y_{n-1}, \Delta_1, \dots, \Delta_{n-1}).$$

Множество доступных наблюдений к моменту поступления n -го задания выражается как

$$H_n = Y^n \times R^n,$$

где $R = [0, \infty)$.

Обратим внимание, что сделанные предположения означают, что диспетчеризация происходит в ситуации, когда не наблюдаемы важные для решения задачи оптимизации характеристики. Более того, не наблюдается даже показатель, подлежащий минимизации, в данном случае это величина w_n — время, проведенное в системе заданием, поступившим в момент t_n . При этом остаются неизвестными обе компоненты этого показателя: и время ожидания в очереди, и время непосредственного выполнения задания.

Вопрос, который нас интересует: дает ли незначительная дополнительная информация (учет моментов поступления заданий) возможность улучшить показатель качества W ? Интуитивная предпосылка для возможного положительного ответа заключается в следующем простом замечании: большие промежутки времени между поступлениями заданий повышают вероятность того, что серверы находятся в состоянии с меньшим остаточным временем, и наоборот. Чтобы воспользоваться этим довольно расплывчатым соображением, можно попытаться на основании доступной информации получить хотя бы приближенную оценку остаточных времен. Будем обозначать такую «оценку» к моменту t_n принятия очередного решения через $\hat{u}_n^{(i)}$ и положим $\hat{u}_1^{(i)} = 0, i = 0 \vee 1$.

Для $n = 2$ наблюдаемая предыстория — это пара $h_1 = (y_1, \Delta_1 = t_2 - t_1)$. Оценку $\hat{u}_2^{(i)}$ определим как $E_{h_1} u_2^{(i)}$, где E_{h_1} — условное математическое ожидание при условии, что предыстория к моменту t_2 была h_1 . Продолжая аналогичным образом, для $n = 2, \dots, k$ определим оценки $\hat{u}_n^{(i)} = E_{h_n} u_n^{(i)}$, где $h_n = (y_1, \Delta_1, \dots, y_{n-1}, \Delta_{n-1})$.

Фиксированное натуральное число k будем называть глубиной памяти. Начиная с номера $n = k + 1$, будем строить оценки, исходя из «усеченной» предыстории:

$$h_{n,k} = (y_{n-k}, \Delta_{n-k}, \dots, y_{n-1}, \Delta_{n-1}).$$

Полагаем $\hat{u}_n^{(i)} = E_{h_{n,k}} u_n^{(i)}$, где $E_{h_{n,k}}$ — это условное математическое ожидание при условии, что (а) наблюдаемая часть предыстории на предшествующих k тактах была $h_{n,k}$ и (б) остаточные времена к моменту t_{n-k} равнялись $u_{n-k}^{(i)} = \hat{u}_{n-k}^{(i)}$.

Воспользуемся теперь оценками остаточных времен для построения диспетчеризации, которую обозначим МЕМ. За эвристическую основу возьмем алгоритм порогового типа [23], который является эффективной диспетчеризацией, а в случае, когда все задания имеют одинаковую длину, является даже оптимальным. Правило диспетчеризации определяем следующим образом: задание, поступившее в момент t_n ,

- направляется на сервер 0, если $\hat{u}_n^{(0)} + T < \hat{u}_n^{(1)}$;
- направляется на сервер 1, если $\hat{u}_n^{(0)} + T \geq \hat{u}_n^{(1)}$.

Пороговое значение T является фиксированной неотрицательной величиной.

Для реализации диспетчеризации МЕМ необходимо уметь вычислять оценки остаточных времен и определять пороговое значение. Оценки $\hat{u}_n^{(i)}$ можно получать, усредняя результаты многократной имитации отрезка траектории процесса. Для данного n длина имитируемого отрезка составляет $l = \min(k, n - k)$, начальное значение остаточных времен принимается равным $\hat{u}_{n-l}^{(i)}$, а значения действий и промежутков между ними фиксированы и совпадают с наблюдаемой предысторией $h_{n,k}$. Таким образом, одна имитационная итерация сводится к l -кратной реализации случайной величины с распределением длины задания G и соответствующему пересчету остаточных времен. Отметим, что, несмотря на конечную глубину предыстории, используемой при расчете оценок на каждом шаге, фактически диспетчеризация МЕМ учитывает, хотя и косвенно, всю предыторию.

Способ вычисления оптимального порога при детерминированных и одинаковых длинах заданий описан в [24]. Поскольку в рассматриваемом случае нет какого-либо теоретически обоснованного способа определения наилучшего порогового значения, так же как нет теоретического обоснования для возможного преимущества стратегии МЕМ, то в качестве параметра T предлагается брать значение, которое получается с помощью алгоритма из [24].

4 Результаты численных расчетов

4.1 Условия эксперимента

Рассматривается система из двух серверов: 0 и 1. Предполагается, что производительность одного сервера равна 1, а другого — 2. Входящий

Таблица 1 Оптимальные значения параметра для стратегий RND и PROG

Интенсивность входного потока	Распределение длины заданий									
	Экспоненциальное		Равномерное		Парето		Нормальное		Вырожденное	
	RND	PROG	RND	PROG	RND	PROG	RND	PROG	RND	PROG
0,6	0,01	0,85	0	1	0,01	0,99	0	1	0	1
0,9	0,14	0,75	0,07	0,78	0,08	0,78	0,04	0,83	0,03	0,86
1,2	0,21	0,73	0,18	0,72	0,20	0,75	0,17	0,73	0,16	0,75
1,5	0,25	0,71	0,24	0,70	0,25	0,73	0,23	0,69	0,23	0,68
1,8	0,28	0,70	0,27	0,68	0,28	0,70	0,27	0,67	0,27	0,67
2,1	0,30	0,69	0,30	0,68	0,30	0,68	0,29	0,67	0,30	0,67
2,4	0,31	0,68	0,31	0,68	0,32	0,67	0,31	0,67	0,31	0,67

в систему поток является пуассоновским с параметром λ .

Распределение длин поступающих заданий выбирается из следующих вариантов:

- (а) экспоненциальное распределение с параметром 1 (среднее и дисперсия равны 1);
- (б) равномерное распределение на отрезке $[0,1; 1,9]$ (среднее равно 1, дисперсия — 0,27);
- (в) распределение Парето с функцией распределения $F(t) = 1 - (b/t)^a, t \geq b$, где $b = 0,6, a = 2,5$ (среднее — 1, дисперсия — 0,8);
- (г) $\max(10^{-7}, \xi)$, где ξ имеет нормальное распределение со средним 1 и дисперсией 0,1;
- (д) вырожденное распределение — длина всех заданий равна 1.

Далее проводится сравнение значений целевой функции для четырех диспетчеризаций (RND, PROG, JSQ и MEM) при разных распределениях длины заданий и при разных интенсивностях входного потока.

Под стратегией RND будем понимать описанный в разд. 2 алгоритм, в котором серверы выбираются с фиксированной вероятностью. В нашем случае стратегия RND задается единственным параметром p — вероятностью выбора первого сервера. Предполагается, что этот параметр выбран оптимальным образом.

Стратегия PROG использует упомянутую в разд. 2 детерминированную последовательность Битти, состоящую из нулей и единиц. Правило этой стратегии заключается в том, чтобы направлять k -е по счету задание на сервер i , если $b_k = [(k+1)p + \varphi] - [kp + \varphi] = i, i = 0 \vee 1$. Здесь $-\infty < \varphi < \infty$ и $0 < p < 1$ — произвольные рациональные числа, а скобки означают целую часть числа. Параметр φ («фаза») обуславливает только сдвиг детерминированной последовательности и не влияет на значение целевой функции, поэтому полагаем $\varphi = 0$. А вот от

Таблица 2 Пороговые значения для алгоритма MEM

Интенсивность входного потока	Пороговое значение
0,6	0,38
0,9	0,33
1,2	0,28
1,5	0,25
1,8	0,22
2,1	0,19
2,4	0,17

параметра p значение показателя W зависит существенно, и снова предполагается, что это значение выбрано наилучшим способом.

Отыскание оптимальных значений параметра p для стратегий RND и PROG, по существу, представляет собой отдельные задачи, которые в данном случае решаются с помощью специального адаптивного алгоритма на имитационной модели. Найденные значения представлены для различных распределений длины задания и различных интенсивностей входного потока в табл. 1.

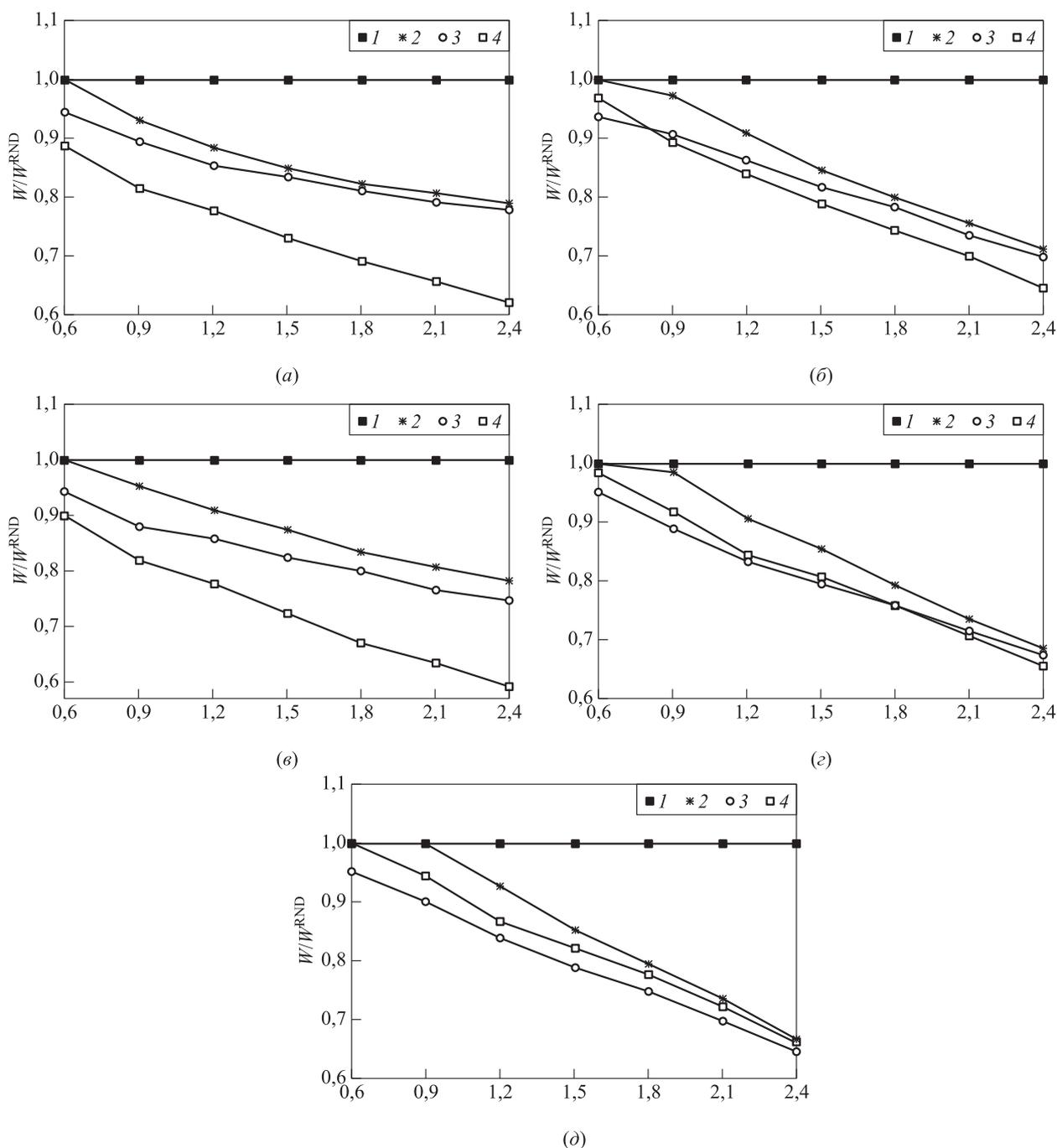
Диспетчеризация JSQ направляет задания каждый раз в наиболее короткую очередь.

Пороговые значения, необходимые для реализации алгоритма MEM, полученные в результате численного решения задачи о наилучшем пороге (см. разд. 3), указаны в табл. 2.

Глубина памяти k в алгоритме MEM динамически корректировалась в процессе имитации в диапазоне от 2 до 6.

4.2 Результаты эксперимента

Численные результаты сравнения диспетчеризаций получены методом Монте Карло и изображены на рисунке. Значения целевой функции для различных стратегий обозначаются соответственно $W^{MEM}, W^{RND}, W^{PROG}$ и W^{JSQ} . Представлены относительные значения функций W^{MEM}, W^{PROG} и W^{JSQ} к значениям функции W^{RND} .



Сравнение эффективности диспетчеризаций в системе из двух серверов: 1 — W^{RND}/W^{RND} ; 2 — W^{PROG}/W^{RND} ; 3 — W^{MEM}/W^{RND} ; 4 — W^{JSQ}/W^{RND} . Длина заданий: (а) экспоненциально распределена с параметром 1; (б) равномерно распределена на отрезке $[0,1; 1,9]$; (в) имеет распределение Парето с параметрами 2,5 и 0,6; (г) распределена как $\max(10^{-7}, \xi)$, где ξ — нормально распределенная случайная величина со средним 1 и дисперсией 0,1; (д) равна 1

В рассмотренных примерах стратегия RND, которая не использует никаких наблюдений, играет роль точки отсчета и заведомо менее эффективна по сравнению с остальными. Стратегия PROG формально также не использует наблюдений, но

эквивалентна периодической детерминированной стратегии, т.е. эквивалентна стратегии, использующей предысторию совершенных действий. Эта стратегия является основным объектом сравнения для стратегии MEM.

Таблица 3 Средний выигрыш в эффективности при стратегии MEM по сравнению со стратегией DET

Распределение	Дисперсия	Выигрыш, %
Экспоненциальное	1	2,8
Равномерное	0,8	4,0
Парето	0,27	5,5
Нормальное	0,1	5,3
Вырожденное	0	6,1

Во всех случаях диспетчеризация MEM с предысторией моментов поступления заявок позволяет улучшить значение показателя W по сравнению с детерминированной стратегией PROG. Достигаемый выигрыш лежит для отдельных значений в диапазоне от 1,5% до 10,0%. При этом наблюдается уменьшение выигрыша с ростом нагрузки.

Средний выигрыш при разных интенсивностях зависит от типа распределения длины задания и от дисперсии (табл. 3). Наблюдается рост относительного среднего выигрыша с уменьшением дисперсии длины заданий.

С влиянием дисперсии можно также связать результат сравнения стратегий MEM и JSQ. Диспетчеризация JSQ в контексте данного эксперимента интересна тем, что использует наблюдения, хотя и неполные, но связанные с состояниями очередей. Интуитивно можно предположить, что такие наблюдения должны давать существенные преимущества по сравнению с полностью ненаблюдаемыми характеристиками нагрузки. Это предположение, однако, подтверждается на приведенных рисунках только для больших значений дисперсии длины заданий (см. $a-v$ на рисунке). Для малых дисперсий (т.е. менее 0,2, см. z и d на рисунке), стратегия JSQ дает относительно меньший выигрыш по сравнению с оптимальной детерминированной диспетчеризацией PROG и, более того, проигрывает по сравнению со стратегией MEM. Можно предположить, что последнее обстоятельство связано с усилением влияния предыстории в ситуации, когда мала дисперсия и, соответственно, в меньшей степени присутствует эффект перемешивания.

5 Заключение

Неполнота наблюдений является частым ограничением во многих технических системах и связана как с недоступностью некоторых параметров для отслеживания, так и с невозможностью обработки слишком больших массивов наблюдений. Это порождает необходимость разрабатывать стратегии, основывающиеся на неполных или косвенных данных о состоянии процесса управления.

Данная работа посвящена проблеме управления распределением заданий в системе с параллельным обслуживанием при минимальном наблюдении, когда неизвестны длины заданий и состояния очередей. Известно, что если наблюдаема только последовательность номеров серверов, на которые направляются задания, то детерминированная стратегия (в частных случаях — периодическая) позволяет получать оптимальные (или близкие к оптимальным) значения среднего времени пребывания задания в системе. Такая стратегия оказывается зависящей от предыстории действий. Возникает вопрос, улучшится ли эффективность, если дополнительно учитывать предысторию моментов поступления заданий. В работе с помощью численных экспериментов дается положительный ответ на этот вопрос на примере системы из двух серверов и специальной эвристической диспетчеризации MEM, основанной на оценках состояния очереди по результатам имеющихся наблюдений и априорной информации о распределении длины заданий. Во всех экспериментах с различными интенсивностями пуассоновского входного потока и разными распределениями длины заданий стратегия MEM показала выигрыш (от 1,5% до 10%) по сравнению с оптимальной программной стратегией. Выигрыш оказался более существенным для малых значений дисперсии длины заданий. Для таких примеров стратегия MEM оказалась даже эффективнее, чем диспетчеризация по самой короткой очереди.

Результаты проведенного анализа указывают на принципиальную возможность повышения качества частично наблюдаемых систем обслуживания путем построения стратегий диспетчеризации, учитывающих предысторию. Дальнейшие усилия будут направлены на разработку таких стратегий и на повышение их практической ценности за счет уменьшения объема необходимой априорной информации о системе.

Литература

1. Semchedine F., Bouallouche-Medjkoune L., Aïssani D. Review: Task assignment policies in distributed server systems: A survey // J. Netw. Comput. Appl., 2011. Vol. 34. No. 4. P. 1123–1130.
2. Коновалов М. Г., Разумчик Р. В. Обзор моделей и алгоритмов размещения заданий в системах с параллельным обслуживанием // Информатика и её применения, 2015. Т. 9. Вып. 4. С. 56–67.
3. Anselmi J., Gaujal B., Netti T. Control of parallel non-observable queues: Asymptotic equivalence and optimality of periodic policies // Stochastic Syst., 2015. Vol. 5. P. 120–145.

4. Gardner K., Borst S., Harchol-Balter M. Optimal scheduling for jobs with progressive deadlines // IEEE Conference on Computer Communications (INFOCOM). — IEEE, 2015. P. 1113–1121.
5. Hyytiä E., Richter R. Routing jobs with deadlines to heterogeneous parallel servers // Oper. Res. Lett., 2016. Vol. 44. No. 4. P. 507–513.
6. Harchol-Balter M., Scheller-Wolf A., Young A. R. Surprising results on task assignment in server farms with high-variability workloads // ACM SIGMETRICS/Performance Proceedings. — Seattle: ACM, 2009. P. 287–298.
7. Javadi B., Kondo D., Vincent J.-M., Anderson D. P. Discovering statistical models of availability in large distributed systems: An empirical study of seti@home // IEEE Trans. Parall. Distr. Syst., 2011. Vol. 22. No. 11. P. 1896–1903.
8. Javadi B., Thulasiraman P., Buyya R. Cloud resource provisioning to extend the capacity of local resources in the presence of failures // IEEE 14th Conference (International) on High Performance Computing and Communication & IEEE 9th Conference (International) on Embedded Software and Systems. — IEEE Computer Society, 2012. P. 311–319.
9. Combe M. B., Boxma O. J. Optimization of static traffic allocation policies // Theor. Comput. Sci., 1994. Vol. 125. No. 1. P. 17–43.
10. Tang Ch. S., van Vliet M. Traffic allocation for manufacturing systems // Eur. J. Oper. Res., 1994. Vol. 75. No. 1. P. 171–185.
11. Sethuraman J., Squillante M. S. Optimal stochastic scheduling in multiclass parallel queues // ACM SIGMETRICS'99 Proceedings. — New York, NY, USA: ACM, 1999. P. 93–102.
12. Humblet P. Determinism minimizes waiting time in queues. The Laboratory for Information and Decision Systems Technical Report ser., 1982. LIDS-P/1207.
13. Hajek B. The proof of a folk theorem on queuing delay with applications to routing in networks // J. ACM, 1983. Vol. 30. No. 4. P. 834–851.
14. Altman E., Gaujal B., Hordijk A. Balanced sequences and optimal routing // J. ACM, 2000. Vol. 47. No. 4. P. 752–775.
15. Altman E., Gaujal B., Hordijk A. Balanced sequences and optimal routing // J. ACM, 2000. Vol. 47. No. 4. P. 752–775.
16. Van der Laan D. The structure and performance of optimal routing sequences. — Universiteit Leiden, 2003. PhD Thesis.
17. Hordijk A., van der Laan D. A. Periodic routing to parallel queues and billiard sequences // Math. Method. Oper. Res., 2004. Vol. 59. No. 2. P. 173–192.
18. Gaujal B., Hyon E. Optimal routing policy in two deterministic queues // Calculateurs Parallèles, 2001. Vol. 13. P. 601–634.
19. Van der Laan D. A. Routing jobs to servers with deterministic service times // Math. Oper. Res., 2005. Vol. 30. No. 1. P. 195–224.
20. Arian Y., Levy Y. Algorithms for generalized round robin routing // Oper. Res. Lett., 1992. Vol. 12. P. 313–319.
21. Sano S., Miyoshi N. Applications of m-balanced sequences to some network scheduling problems // Discrete Event System: Analysis and Control: 5th Workshop on Discrete Event Systems Proceedings. — Kluwer Academic Publisher, 2000. P. 317–325.
22. Hyytiä E. Lookahead actions in dispatching to parallel queues // Perform. Evaluation, 2013. Vol. 70. No. 10. P. 859–872.
23. Hyytiä E. Optimal routing of fixed size jobs to two parallel servers // INFOR: Inform. Syst. Oper. Res., 2013. Vol. 51. No. 4. P. 215–224.
24. Konovalov M., Razumchik R. Iterative algorithm for threshold calculation in the problem of routing fixed size jobs to two parallel servers // J. Telecommun. Inform. Technol., 2015. No. 3. P. 32–38.

Поступила в редакцию 19.09.16

DISPATCHING TO TWO PARALLEL NONOBSERVABLE QUEUES USING ONLY STATIC INFORMATION

M. G. Konovalov¹ and R. V. Razumchik^{1,2}

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

²Peoples’ Friendship University of Russia, 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation

Abstract: Consideration is given to the problem of dispatching independent jobs from one flow to two parallel single server queueing systems each with an infinite capacity queue. There is one dispatcher, which immediately

makes decisions where to route newly incoming jobs. In order to make the decision, the dispatcher uses only static information about the system, i. e., servers' speeds, job interarrival time distribution, and job size distribution. The dynamic information (for example, current queue sizes) is unavailable for the dispatcher. For such nonobservable queues, it is known that minimum mean sojourn time is achieved when the dispatcher uses the deterministic (periodic) policy. Even when using this optimal policy, the dispatcher also observes jobs' arrival instants but this information is left unused. The question posed in this paper is the following: Is it possible to reduce the mean sojourn time if one, in addition to the static information, also uses the historical information about jobs' arrival instants? Using simulation techniques, the authors show that the answer to the question is positive. Such policy uses static information and the estimates of the queue sizes based on multiple replications of the system's trajectory. Compared to the optimal policy, the achievable gain varies from 1,5% to 10%, and increases with the decrease of job size variance. When the job size variance is low, the proposed policy is even better than the dynamic join-the-shortest queue policy.

Keywords: dispatching; static information; parallel service; queueing system; nonobservable queues

DOI: 10.14357/19922264160406

Acknowledgments

The research was supported by the Russian Foundation for Basic Research (project 15-07-03406).

References

1. Semchedine, F., L. Bouallouche-Medjkoune, and D. Aïssani. 2011. Review: Task assignment policies in distributed server systems: A survey. *J. Netw. Comput. Appl.* 34(4):1123–1130.
2. Konovalov, M. G., and R. V. Razumchik. 2015. Obzor modeley i algoritmov razmeshcheniya zadaniy v sistemakh s parallel'nym obsluzhivaniem [Methods and algorithms for job scheduling in systems with parallel service: A survey]. *Informatika i ee primeneniya — Inform. Appl.* 9(4):56–67.
3. Anselmi, J., B. Gaujal, and T. Nesti. 2015. Control of parallel non-observable queues: Asymptotic equivalence and optimality of periodic policies. *Stochastic Syst.* 5:120–145.
4. Gardner, K., S. Borst, and M. Harchol-Balter. 2015. Optimal scheduling for jobs with progressive deadlines. *IEEE Conference on Computer Communications (INFOCOM)*. IEEE. 1113–1121.
5. Hyytiä, E., and R. Righter. 2016. Routing jobs with deadlines to heterogeneous parallel servers. *Oper. Res. Lett.* 44(4):507–513.
6. Harchol-Balter, M., A. Scheller-Wolf, and A. R. Young. 2009. Surprising results on task assignment in server farms with high-variability workloads. *ACM SIGMETRICS/Performance Proceedings*. Seattle: ACM. 287–298.
7. Javadi, B., D. Kondo, J.-M. Vincent, and D. P. Anderson. 2011. Discovering statistical models of availability in large distributed systems: An empirical study of seti@home. *IEEE Trans. Parall. Distr. Syst.* 22(11):1896–1903.
8. Javadi, B., P. Thulasiraman, and R. Buyya. 2012. Cloud resource provisioning to extend the capacity of local resources in the presence of failures. *IEEE 14th Conference (International) on High Performance Computing and Communication & IEEE 9th Conference (International) on Embedded Software and Systems Proceedings*. 311–319.
9. Combe, M. B., and O. J. Boxma. 1994. Optimization of static traffic allocation policies. *Theor. Comput. Sci.* 125(1):17–43.
10. Tang, Ch. S., and M. van Vliet. 1994. Traffic allocation for manufacturing systems. *Eur. J. Oper. Res.* 75(1):171–85.
11. Sethuraman, J., and M. S. Squillante. 1999. Optimal stochastic scheduling in multiclass parallel queues. *ACM SIGMETRICS'99 Proceedings*. New York, NY: ACM. P. 93–102.
12. Humblet, P. 1982. Determinism minimizes waiting time in queues. The Laboratory for Information and Decision Systems Technical Report ser. LIDS-P/1207.
13. Hajek, B. 1983. The proof of a folk theorem on queuing delay with applications to routing in networks. *J. ACM* 30(4):834–851.
14. Altman, E., B. Gaujal, and A. Hordijk. 2000. Balanced sequences and optimal routing. *J. ACM.* 47(4):752–775.
15. Altman, E., B. Gaujal, and A. Hordijk. 2000. Balanced sequences and optimal routing. *J. ACM* 47(4):752–775.
16. Van der Laan, D. 2003. The structure and performance of optimal routing sequences. Universiteit Leiden. PhD Thesis.
17. Hordijk, A., and D. A. van der Laan. 2004. Periodic routing to parallel queues and billiard sequences. *Math. Method. Oper. Res.* 59(2):173–192.
18. Gaujal, B., and E. Hyon. 2001. Optimal routing policy in two deterministic queues. *Calculateurs Parallèles* 13:601–634.
19. Van der Laan, D. A. 2005. Routing jobs to servers with deterministic service times. *Math. Oper. Res.* 30(1):195–224.
20. Arian, Y., and Y. Levy. 1992. Algorithms for generalized round robin routing. *Oper. Res. Lett.* 12:313–319.

21. Sano, S., and N. Miyoshi. 2000. Applications of m-balanced sequences to some network scheduling problems. *Discrete Event System: Analysis and Control: 5th Workshop on Discrete Event Systems Proceedings*. Kluwer Academic Publisher. 317–325.
22. Hyytiä, E. 2013. Lookahead actions in dispatching to parallel queues. *Perform. Evaluation* 70(10):859–872.
23. Hyytiä, E. 2013. Optimal routing of fixed size jobs to two parallel servers. *INFOR: Inform. Syst. Oper. Res.* 51(4):215–224.
24. Konovalov, M., and R. Razumchik. 2015. Iterative algorithm for threshold calculation in the problem of routing fixed size jobs to two parallel servers. *J. Telecommun. Inform. Technol.* 3:32–38.

Received September 19, 2016

Contributors

Konovalov Mikhail G. (b. 1950) — Doctor of Science in technology, principal scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; mkonovalov@ipiran.ru

Razumchik Rostislav V. (b. 1984) — Candidate of Science (PhD) in physics and mathematics, 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; assistant professor, Peoples’ Friendship University of Russia, 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; rrazumchik@ipiran.ru

БАЙЕСОВСКИЕ МОДЕЛИ МАССОВОГО ОБСЛУЖИВАНИЯ И НАДЕЖНОСТИ: ВЫРОЖДЕННО-ВЕЙБУЛЛОВСКИЙ СЛУЧАЙ

А. А. Кудрявцев¹, А. И. Титова²

Аннотация: Данная работа посвящена изучению байесовских моделей массового обслуживания и надежности. В рамках байесовского подхода для классических постановок задач предполагается, что основные параметры системы не являются заданными, а известны только их априорные распределения. За счет рандомизации таких параметров системы, как интенсивность входящего потока и интенсивность обслуживания, происходит рандомизация различных характеристик системы, например коэффициента загрузки. Байесовский подход является целесообразным при изучении больших совокупностей однотипных систем или одной системы, характеристики которой меняются в моменты времени, неизвестные исследователю. В работе представлены конкретные результаты для вероятностных характеристик коэффициента загрузки и вероятности «непотери» вызова в случае, когда в качестве пары априорных распределений параметров системы $M|M|1|0$ рассматриваются вырожденное распределение и распределение Вейбулла

Ключевые слова: байесовский подход; системы массового обслуживания; надежность; смешанные распределения; распределение Вейбулла; вырожденное распределение

DOI: 10.14357/19922264160407

1 Введение

Нередко в математических моделях жизненный цикл функционирования различных объектов зависит от параметров, «способствующих» и «препятствующих» функционированию. В моделях структур и систем массового обслуживания к параметрам, «способствующим» функционированию, можно отнести интенсивность обслуживания запросов, а к параметрам, «препятствующим» функционированию, — интенсивность входящего потока требований. Нетрудно заметить, что итоговые результаты работы системы зависят не столько от значений параметров, сколько от их соотношения.

Ниже будет рассмотрена система массового обслуживания $M|M|1|0$. Одним из основных показателей такой системы является ее коэффициент загрузки ρ . Значение ρ определяется отношением параметра входящего потока λ к параметру обслуживания μ . Многие характеристики разнообразных систем массового обслуживания зависят от величины ρ , в том числе вероятность «непотери» вызова $\pi = \mu/(\lambda + \mu) = 1/(1 + \rho)$.

В рамках байесовских моделей массового обслуживания и надежности в классических задачах предполагается, что значения параметров λ и μ неизвестны, однако имеется информация об их априорных распределениях. Подробное описа-

ние предпосылок для исследования, особенностей и библиографии байесовских моделей в теории массового обслуживания и надежности можно найти в книге [1].

Далее для модели $M|M|1|0$ приводятся вероятностные характеристики коэффициента загрузки ρ и вероятности «непотери» вызова π в случае, когда в качестве пары априорных распределений параметров системы λ и μ рассматриваются вырожденное распределение и распределение Вейбулла.

2 Основные результаты

Введем следующие обозначения: через $D(\lambda)$ обозначим вырожденное в точке λ распределение, а через $W(p, \alpha)$ — распределение Вейбулла с плотностью $w_{p,\alpha}(x)$, имеющей вид:

$$w_{p,\alpha}(x) = \frac{px^{p-1}e^{-(x/\alpha)^p}}{\alpha^p}, \quad x > 0, \quad p > 0, \quad \alpha > 0.$$

Обозначим через $\mathcal{L}_s\{f(x)\}$ преобразование Лапласа соответствующей функции:

$$\mathcal{L}_s\{f(x)\} = \int_0^{\infty} f(x)e^{-sx} dx. \quad (1)$$

¹Московский государственный университет им. М. В. Ломоносова, факультет вычислительной математики и кибернетики, nubigena@mail.ru

²Московский государственный университет им. М. В. Ломоносова, факультет вычислительной математики и кибернетики, onkelskroot@gmail.com

Теорема 1. Пусть параметр входящего потока λ имеет вырожденное распределение, а параметр обслуживания μ имеет распределение Вейбулла $W(p, \alpha)$. Тогда функция распределения, плотность и моменты коэффициента загрузки ρ имеют вид:

$$\begin{aligned} F_\rho(x) &= e^{-(\lambda/(\alpha x))^p}, \quad x > 0; \\ f_\rho(x) &= \frac{p\lambda^p e^{-(\lambda/(\alpha x))^p}}{\alpha^p x^{p+1}}, \quad x > 0; \\ E\rho^k &= \frac{\lambda^k}{\alpha^k} \Gamma\left(\frac{p-k}{p}\right), \quad k < p, \end{aligned} \quad (2)$$

а функция распределения, плотность и моменты вероятности «непотери» вызова π определяются соотношениями:

$$\begin{aligned} F_\pi(x) &= 1 - \exp\left\{-\left(\frac{\lambda x}{\alpha(1-x)}\right)^p\right\}, \quad x \in (0, 1); \\ f_\pi(x) &= \frac{p\lambda^p x^{p-1}}{\alpha^p(1-x)^{p+1}} \exp\left\{-\left(\frac{\lambda x}{\alpha(1-x)}\right)^p\right\}, \\ & \quad x \in (0, 1); \quad (3) \\ E\pi^k &= \mathcal{L}_1\left\{\left(\frac{\alpha x^{1/p}}{\lambda + \alpha x^{1/p}}\right)^k\right\}. \quad (4) \end{aligned}$$

Доказательство. Заметим, что

$$F_\rho(x) = 1 - \frac{p}{\alpha^p} \int_0^{\lambda/x} t^{p-1} e^{-(t/\alpha)^p} dt = e^{-(\lambda/(\alpha x))^p}, \quad x > 0.$$

Продифференцировав полученное выражение, получаем (2).

Для k -го момента случайной величины ρ имеем:

$$\begin{aligned} E\rho^k &= \frac{p\lambda^p}{\alpha^p} \int_0^\infty \frac{e^{-(\lambda/(\alpha x))^p}}{x^{p-k+1}} dx = \\ &= \frac{p\lambda^k}{\alpha^k} \int_0^\infty z^{p-k-1} e^{-z^p} dz = \\ &= \frac{\lambda^k}{\alpha^k} \int_0^\infty t^{(p-k)/p-1} e^{-t} dt = \frac{\lambda^k}{\alpha^k} \Gamma\left(\frac{p-k}{p}\right), \quad k < p. \end{aligned}$$

Теперь рассмотрим характеристики вероятности «непотери» вызова π . Для функции распределения имеем:

$$\begin{aligned} F_\pi(x) &= 1 - P\left(\rho < \frac{1-x}{x}\right) = \\ &= 1 - \exp\left\{-\left(\frac{\lambda x}{\alpha(1-x)}\right)^p\right\}, \quad x \in (0, 1). \end{aligned}$$

Продифференцировав последнее выражение, получаем (3).

Найдем $E\pi^k$. Имеем:

$$\begin{aligned} E\pi^k &= \int_0^1 \frac{px^{k-1}}{(1-x)} \left(\frac{\lambda x}{\alpha(1-x)}\right)^p \times \\ & \times \exp\left\{-\left(\frac{\lambda x}{\alpha(1-x)}\right)^p\right\} dx = \int_0^\infty \frac{p\alpha^k z^{p+k-1} e^{-z^p}}{(\lambda + \alpha z)^k} dz. \end{aligned}$$

Воспользовавшись заменой $t = z^p$ и соотношением (1), получим (4). Теорема доказана.

Теорема 2. Пусть параметр входящего потока λ имеет распределение Вейбулла $W(q, \theta)$, а параметр обслуживания μ имеет вырожденное распределение. Тогда функция распределения, плотность и моменты коэффициента загрузки ρ имеют вид:

$$\begin{aligned} F_\rho(x) &= 1 - e^{-(\mu x/\theta)^q}, \quad x > 0; \\ f_\rho(x) &= q \left(\frac{\mu}{\theta}\right)^q x^{q-1} e^{-(\mu x/\theta)^q}, \quad x > 0; \\ E\rho^k &= \left(\frac{\theta}{\mu}\right)^k \Gamma\left(\frac{k+q}{q}\right), \end{aligned} \quad (5)$$

а функция распределения, плотность и моменты вероятности «непотери» вызова π определяются соотношениями:

$$\begin{aligned} F_\pi(x) &= \exp\left\{-\left(\frac{\mu(1-x)}{\theta x}\right)^q\right\}, \quad x \in (0, 1); \\ f_\pi(x) &= \frac{q\mu}{\theta x^2} \left(\frac{\mu(1-x)}{\theta x}\right)^{q-1} \exp\left\{-\left(\frac{\mu(1-x)}{\theta x}\right)^q\right\}, \\ & \quad x \in (0, 1); \quad (6) \end{aligned}$$

$$E\pi^k = \mathcal{L}_1\left\{\left(\frac{\mu}{\mu + \theta x^{1/q}}\right)^k\right\}.$$

Доказательство. Аналогично доказательству теоремы 1 рассмотрим соотношение для функции распределения случайной величины ρ :

$$F_\rho(x) = \frac{q}{\theta^q} \int_0^{\mu x} t^{q-1} e^{-(t/\theta)^q} dt = 1 - e^{-(\mu x/\theta)^q}, \quad x > 0,$$

из которого вытекает (5) и цепочка равенств:

$$\begin{aligned} E\rho^k &= \frac{q\mu^q}{\theta^q} \int_0^\infty x^{q+k-1} e^{-(\mu x/\theta)^q} dx = \\ &= \frac{q\theta^k}{\mu^k} \int_0^\infty z^{q+k-1} e^{-z^q} dz = \\ &= \int_0^\infty \left(\frac{\theta}{\mu}\right)^k t^{k/q} e^{-t} dt = \left(\frac{\theta}{\mu}\right)^k \Gamma\left(\frac{k+q}{q}\right). \end{aligned}$$

Таблица 1 Частные значения $E\rho (\lambda \sim D(\lambda), \mu \sim W(p, \alpha))$

λ	$p; \alpha$									
	3; 3	3; 4	3; 5	3; 6	3; 7	3; 8	3; 9	3; 10	3; 11	3; 12
0,5	0,23	0,17	0,14	0,11	0,10	0,08	0,08	0,07	0,06	0,06
1,0	0,45	0,34	0,27	0,23	0,19	0,17	0,15	0,14	0,12	0,11
1,5	0,68	0,51	0,41	0,34	0,29	0,25	0,23	0,20	0,18	0,17
2,0	0,90	0,68	0,54	0,45	0,39	0,34	0,30	0,27	0,25	0,23
2,5	1,13	0,85	0,68	0,56	0,48	0,42	0,38	0,34	0,31	0,28
3,0	1,35	1,02	0,81	0,68	0,58	0,51	0,45	0,41	0,37	0,34
3,5	1,58	1,18	0,95	0,79	0,68	0,59	0,53	0,47	0,43	0,39
4,0	1,81	1,35	1,08	0,90	0,77	0,68	0,60	0,54	0,49	0,45
4,5	2,03	1,52	1,22	1,02	0,87	0,76	0,68	0,61	0,55	0,51
5,0	2,26	1,69	1,35	1,13	0,97	0,85	0,75	0,68	0,62	0,56

Таблица 2 Частные значения $E\pi (\lambda \sim W(q, \theta), \mu \sim D(\mu))$

$q; \theta$	μ									
	2,5	3,0	3,5	4,0	4,5	5,0	5,5	6,0	6,5	7,0
3; 1	0,74	0,60	0,50	0,43	0,38	0,34	0,31	0,28	0,26	0,24
3; 2	0,78	0,64	0,54	0,48	0,42	0,38	0,35	0,32	0,29	0,27
3; 3	0,80	0,67	0,58	0,51	0,46	0,42	0,38	0,35	0,33	0,30
3; 4	0,82	0,70	0,61	0,54	0,49	0,45	0,41	0,38	0,35	0,33
3; 5	0,84	0,72	0,64	0,57	0,52	0,48	0,44	0,41	0,38	0,36
3; 6	0,85	0,74	0,66	0,60	0,54	0,50	0,46	0,43	0,40	0,38
3; 7	0,86	0,76	0,68	0,62	0,57	0,52	0,49	0,45	0,43	0,40
3; 8	0,87	0,78	0,70	0,64	0,59	0,54	0,51	0,48	0,45	0,42
3; 9	0,88	0,79	0,72	0,66	0,61	0,56	0,53	0,49	0,47	0,44
3; 10	0,89	0,80	0,73	0,67	0,62	0,58	0,54	0,51	0,48	0,46

Аналогично для вероятности «непотери» вызова π имеем:

$$F_{\pi}(x) = 1 - P\left(\rho < \frac{1-x}{x}\right) = \exp\left\{-\left(\frac{\mu(1-x)}{\theta x}\right)^q\right\}, \quad x \in (0, 1).$$

Продифференцировав последнее выражение, получаем (6). Для k -го момента случайной величины π справедливо:

$$E\pi^k = \int_0^1 \frac{qx^{k-1}}{1-x} \left(\frac{\mu(1-x)}{\theta x}\right)^q \times \exp\left\{-\left(\frac{\mu(1-x)}{\theta x}\right)^q\right\} dx = \int_0^{\infty} \frac{q\mu^k z^{q-1} e^{-z^q}}{(\mu + \theta z)^k} dz =$$

$$= \int_0^{\infty} \frac{\mu^k e^{-t}}{(\mu + \theta t^{1/q})^k} dt = \mathcal{L}_1\left\{\left(\frac{\mu}{\mu + \theta t^{1/q}}\right)^k\right\}.$$

Теорема доказана.

3 Численные результаты

В качестве иллюстрации приведем табл. 1 и 2, содержащие частные значения характеристик, полученных в теоремах 1 и 2, для некоторых наборов параметров. Значения в таблицах приведены с точностью до сотых.

Литература

1. Кудрявцев А. А., Шоргин С. Я. Байесовские модели в теории массового обслуживания и надежности. — М.: ФИЦ ИУ РАН, 2015. 76 с.

Поступила в редакцию 29.06.16

BAYESIAN QUEUING AND RELIABILITY MODELS: DEGENERATE-WEIBULL CASE

A. A. Kudryavtsev and A. I. Titova

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

Abstract: This paper is devoted to Bayesian queuing and reliability models. In the framework of the Bayesian approach, it is assumed that the key parameters of classical systems are random and only their *a priori* distributions are known. By randomizing system's parameters such as the input flow intensity and the service intensity, one may randomize system's characteristics, for example, system's loading factor. The Bayesian approach can be used in the case of studying large groups of systems and devices or one system with variable characteristics. The results for probability characteristics of the system's loading factor and the probability that the claim received by the system will not be lost in the case of the system of the $M|M|1|0$ type where one of the system's parameters has a degenerate distribution and the other has the Weibull distribution are presented.

Keywords: Bayesian approach; queuing theory; reliability theory; mixed distributions; Weibull distribution; degenerate distribution

DOI: 10.14357/19922264160407

References

1. Kudryavtsev, A. A., and S. Ya. Shorgin. 2015. *Bayesovskie modeli v teorii massovogo obsluzhivaniya i nadezhnosti*

[Bayesian models in mass service and reliability theories]. Moscow: Federal Research Center "Computer Sciences and Control" of the Russian Academy of Sciences. 76 p.

Received June 29, 2016

Contributors

Kudryavtsev Alexey A. (b. 1978) — Candidate of Science (PhD) in physics and mathematics, associate professor, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; nubigena@mail.ru

Titova Anastasiia I. (b. 1995) — student, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; onkelskroot@gmail.com

АНАЛИТИЧЕСКОЕ РЕШЕНИЕ ЗАДАЧИ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ ПОЛУМАРКОВСКИМ ПРОЦЕССОМ С КОНЕЧНЫМ МНОЖЕСТВОМ СОСТОЯНИЙ*

П. В. Шнурков¹, А. К. Горшенин², В. В. Белоусов³

Аннотация: Настоящее исследование посвящено теоретическому обоснованию нового метода нахождения оптимальной стратегии управления полумарковским процессом с конечным множеством состояний. Рассматриваются марковские рандомизированные стратегии управления, определяемые конечным набором вероятностных мер, соответствующих каждому состоянию. Характеристикой качества управления служит стационарный стоимостной показатель. Данный показатель представляет собой дробно-линейный интегральный функционал от набора вероятностных мер, задающих стратегию управления. Для этого функционала известны явные аналитические представления подынтегральных функций числителя и знаменателя. Дальнейшие результаты основываются на новой усиленной и обобщенной форме теоремы об экстремуме дробно-линейного интегрального функционала. Доказывается, что проблемы существования оптимальной стратегии управления полумарковским процессом и ее нахождения сводятся к задаче численного исследования на глобальный экстремум заданной функции от конечного числа вещественных переменных.

Ключевые слова: оптимальное управление полумарковским процессом; стационарный стоимостной показатель качества управления; дробно-линейный интегральный функционал

DOI: 10.14357/19922264160408

1 Введение

Теория оптимального управления марковскими и полумарковскими случайными процессами интенсивно развивается с начала 1960-х гг. Еще в первых основополагающих исследованиях рассматривались не только проблемы существования оптимальных стратегий управления, но и способы нахождения этих стратегий.

Для решения таких проблем, имеющих алгоритмическое содержание, использовались открытые незадолго до этого мощные методы прикладной математики: линейное программирование и динамическое программирование. Отметим, прежде всего, классическую работу Р. Ховарда [1], в которой метод динамического программирования был применен для решения проблемы оптимального управления марковским процессом с непрерывным временем. В дальнейшем В. В. Рыков [2] доказал, что для аналогичной модели управления марковским процессом с учетом переоценки оптимальной стратегией также является стационарная.

Важную роль в развитии теории управления случайными процессами сыграла работа В. Джевелла [3], в которой были впервые рассмотрены полумарковские модели управления для вариантов с переоценкой и без переоценки. Данная работа была переведена на русский язык и послужила основой для многих последующих работ отечественных и зарубежных специалистов. В частности, Б. Фокс показал [4], что оптимальной стратегией управления полумарковским процессом в варианте без переоценки является стационарная; аналогичные результаты были получены Э. Денардо и для варианта с переоценкой [5].

Среди последующих исследований алгоритмической направленности отметим работы Р. Ховарда [6], Б. Фокса [4], а также С. Осаки и Х. Майна [7]. В этих работах для нахождения оптимальных стратегий управления полумарковскими процессами использовался метод линейного программирования.

В 1970 г. была опубликована фундаментальная монография Х. Майна и С. Осаки [8], переведенная

* Работа выполнена при частичной поддержке РФФИ (проект 15-07-05316).

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

²Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, agorshenin@frccsc.ru

³Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, vbelousov@ipiran.ru

на русский язык в 1977 г., в которой были систематизированы и изложены основные результаты по теории оптимального управления марковскими и полумарковскими случайными процессами. Фактически данная книга стала итогом исследований по проблемам стохастического управления за 10 лет. Отметим, что в этой монографии рассматривались марковские и полумарковские модели управления с конечными множествами состояний и допустимых решений, принимаемых в каждом состоянии. Были получены принципиальные теоретические результаты, заключающиеся в том, что оптимальные стратегии управления для основных видов рассматриваемых моделей с переоценкой и без переоценки являются детерминированными и стационарными. Были разработаны и обоснованы процедуры нахождения оптимальных стратегий управления. В частности, для модели управления полумарковским процессом без переоценки, когда множество состояний образует один эргодический класс, а показатель качества управления представляет собой стационарный средний удельный доход (см. [8, гл. 5, п. 5.5]), процедура поиска оптимальной рандомизированной стратегии осуществлялась методом линейного программирования. Обратим особое внимание на данный результат, поскольку аналогичная модель управления полумарковским процессом будет рассмотрена в настоящей работе.

Принципиальную роль в развитии теории стохастического управления сыграла монография И. И. Гихмана и А. В. Скорохода [9]. В этой книге были впервые систематически изложены основы теории оптимального управления случайными процессами с дискретным и непрерывным временем, включая теорию управления процессами, которые описываются стохастическими дифференциальными уравнениями. Отдельно были рассмотрены проблемы управления марковскими процессами с дискретным временем и скачкообразными марковскими процессами с непрерывным временем. Роли множеств состояний и допустимых управлений играли пространства весьма общей структуры. Для широких классов функционалов качества управления (так называемых эволюционных функционалов в марковских моделях с дискретным временем и интегральных функционалов накопления в марковских моделях с непрерывным временем) были доказаны теоремы о существовании и формах представления оптимальных стратегий управления. Было установлено, что для однородных марковских моделей оптимальные стратегии управления существуют, являются стационарными и детерминированными. Иначе говоря, такие стратегии задаются детерминированными функциями, аргументом которых является состояние системы

в момент принятия решения, и не зависящими от самого момента принятия решения. Что же касается важного вопроса о формах представления этих функций, то их можно охарактеризовать следующим образом. Были найдены функциональные уравнения, осложненные условием экстремума, которым удовлетворяют упомянутые функции. По существу эти соотношения представляют собой уравнения Беллмана для соответствующих динамических стохастических моделей.

Особо отметим, что в монографии [9] не рассматривались проблемы управления полумарковскими процессами. Однако дальнейшее развитие общей теории управления такими процессами шло по пути, идейно намеченному в указанной книге.

В последующие годы развитие теории управления полумарковскими процессами осуществлялось по направлению усложнения моделей и обобщения исходных предположений. Например, в работах [10, 11] рассмотрены управляемые полумарковские процессы при весьма общих предположениях относительно характера пространств состояний и управлений. Проблемы управления исследовались по отношению к различным видам целевых показателей, обобщающих упомянутый выше стационарный показатель средней удельной прибыли. В этих работах доказывается, что оптимальная стратегия управления по отношению к каждому из показателей существует и является одной и той же стационарной детерминированной стратегией, определяемой некоторой функцией, заданной на множестве состояний процесса. Об этой функции известно лишь то, что она удовлетворяет некоторому интегральному уравнению, которое по содержанию представляет собой уравнение Беллмана для соответствующей задачи управления.

Среди исследований, предшествовавших настоящему, отметим работу В. А. Каштанова [12, гл. 13]. В этом разделе коллективной монографии [12] автором была рассмотрена проблема оптимального управления полумарковским процессом с конечным множеством состояний и множеством возможных решений, которое представляет собой произвольный интервал множества вещественных чисел. Модель относится к виду моделей без переоценки, показателем качества управления служит стационарное значение среднего удельного дохода, определяемое аналогично классическим работам [3, 8]. Рандомизированное управление в каждом состоянии определяется в соответствии с вероятностным распределением, совокупность которых задает стратегию управления. В. А. Каштановым было сформулировано утверждение о том, что стационарное значение среднего удельного дохода представляет собой дробно-линейный интегральный

функционал от набора вероятностных распределений, образующих стратегию управления. При этом ранее [12, гл. 10; 13] было установлено, что дробно-линейный функционал достигает экстремума на вырожденных распределениях. Отсюда естественно следует, что оптимальная стратегия управления является детерминированной и должна определяться точкой экстремума функции, представляющей собой отношение подынтегральных функций числителя и знаменателя данного дробно-линейного функционала. Однако в [12] не были получены явные представления для указанных функций. Кроме того, приведенный в гл. 10 монографии [12] вариант теоремы об экстремуме дробно-линейного интегрального функционала требовал проверки выполнения условия существования этого экстремума. Такие условия указаны не были. В связи с этими обстоятельствами использовать полученные в [12] результаты для доказательства существования оптимальной детерминированной стратегии управления полумарковским процессом и для строгого обоснования способа нахождения такой стратегии оказалось невозможным.

Настоящее исследование посвящено теоретическому обоснованию нового метода нахождения оптимальной стратегии управления полумарковским процессом с конечным множеством состояний. Рассматриваются марковские рандомизированные стратегии управления, определяемые конечным набором вероятностных мер, соответствующих каждому состоянию. Показателем качества управления служит уже упоминавшийся классический показатель — стационарное значение средней удельной прибыли. Доказано, что этот показатель представляет собой дробно-линейный интегральный функционал от набора вероятностных мер, задающих стратегию управления. При этом, в отличие от [12], получены явные аналитические представления для подынтегральных функций числителя и знаменателя этого дробно-линейного функционала. Дальнейшие результаты основываются на новой усиленной и обобщенной форме теоремы об экстремуме дробно-линейного интегрального функционала, впервые опубликованной в работе П. В. Шнуркова [14]. Согласно утверждению этой теоремы, если существует глобальный экстремум так называемой основной функции дробно-линейного функционала, которая представляет собой отношение подынтегральных функций числителя и знаменателя, то существует безусловный экстремум самого дробно-линейного функционала, который достигается на наборе вырожденных вероятностных распределений, сосредоточенных в точке глобального экстремума. В этом случае оптимальная стратегия управления

существует, является стационарной и детерминированной и определяется точкой, в которой основная функция достигает глобального экстремума. Таким образом, проблемы существования оптимальной стратегии управления полумарковским процессом и ее нахождения сводятся к задаче численного исследования на глобальный экстремум заданной функции от конечного числа вещественных переменных.

2 Общее описание модели управления полумарковским случайным процессом

Построим модель управления полумарковским случайным процессом, следуя общему подходу, принятому в классических работах [3, 8]. Пусть $\xi(t)$ — случайный полумарковский процесс с конечным множеством состояний $X = \{1, 2, \dots, N\}$, $N < \infty$. Обозначим через t_n , $n = 0, 1, 2, \dots$, $t_0 = 0$, случайные моменты изменения состояний данного процесса, $\theta_n = t_{n+1} - t_n$, $n = 0, 1, 2, \dots$, $\xi_n = \xi(t_n) = \xi(t_n + 0)$, $n = 0, 1, 2, \dots$ (предполагается, что траектории процесса $\xi(t)$ непрерывны справа). Случайная последовательность $\{\xi_n\}$ образует цепь Маркова, вложенную в полумарковский процесс $\xi(t)$. Зададим набор измеримых пространств $(U_1, \mathcal{B}_1), (U_2, \mathcal{B}_2), \dots, (U_N, \mathcal{B}_N)$, где U_i — множество возможных допустимых управлений, \mathcal{B}_i — σ -алгебра подмножеств множества U_i , включающая в себя все одноточечные подмножества множества U_i , т.е. если $u_i \in U_i$, то $\{u_i\} \in \mathcal{B}_i$, $i = 1, 2, \dots, N$. Пусть Γ_i — некоторое множество всевозможных вероятностных мер $\Psi_i \in \Gamma_i$, заданных на σ -алгебре \mathcal{B}_i , $i = 1, 2, \dots, N$.

Поскольку идейное содержание и свойства вероятностных мер существенно используются в данной работе, укажем на некоторые фундаментальные издания, в которых изложена соответствующая теория. Понятие и основные свойства вероятностной меры определены и подробно проанализированы в книге А. Н. Ширяева [15, гл. II]. Глубокое изложение основ теории вероятностных мер имеется также в книге А. А. Боровкова [16]. Заметим попутно, что в книге [16] имеются разделы, посвященные изложению основ теории полумарковских и регенерирующих случайных процессов. Из зарубежных изданий отметим фундаментальную работу П. Хеннекена и А. Тортра [17], основная часть которой посвящена изложению математических основ теории вероятностей.

Введем специальное понятие вырожденной вероятностной меры, которое будет часто использо-

ваться в дальнейшем. Пусть (U_0, \mathcal{B}_0) — некоторое измеримое пространство, \mathcal{B}_0 — σ -алгебра подмножеств множества U_0 , включающая в себя все одноточечные подмножества этого множества.

Определение 1. Вероятностная мера Ψ^* , заданная на σ -алгебре \mathcal{B}_0 , называется вырожденной, если существует такой элемент $u^* \in U_0$, для которого выполняются условия $\Psi^*(\{u^*\}) = 1$, $\Psi^*(U_0 \setminus \{u^*\}) = 0$, где $\{u^*\} = u^*$ — множество, состоящее из единственной точки $u^* \in U_0$. Соответствующая точка $u^* \in U_0$ будет называться точкой сосредоточения вырожденной вероятностной меры Ψ^* . Таким образом, всякая вырожденная вероятностная мера Ψ^* определяется своей точкой сосредоточения u^* . В дальнейшем будем использовать обозначение $\Psi_{u^*}^*$, имея в виду, что вырожденная вероятностная мера Ψ^* сосредоточена в точке u^* . Отметим также, что вырожденная вероятностная мера $\Psi_{u^*}^*$ соответствует детерминированной величине, которая принимает фиксированное значение $u = u^*$ с вероятностью, равной единице.

Обозначим через Γ_0 множество всех вероятностных мер, заданных на измеримом пространстве (U_0, \mathcal{B}_0) , а через Γ_0^* — множество всех вырожденных вероятностных мер, заданных на этом пространстве, $\Gamma_0^* \in \Gamma_0$. Аналогичные обозначения будут использоваться и в дальнейшем. Заметим, что множество Γ_0^* находится во взаимно однозначном соответствии с множеством точек сосредоточения вырожденных вероятностных мер, т.е. с множеством U_0 .

Пусть Γ_i^* — множество всех вырожденных мер, заданных на σ -алгебре \mathcal{B}_i , $\Gamma_i^* \subset \Gamma_i$. Произвольная вероятностная мера Ψ_i описывает случайную величину, принимающую значения в U_i , а вырожденная мера Ψ_i^* , сосредоточенная в точке u_i^* , соответствует детерминированной величине $u_i^* \in U_i$. Предполагается, что соответствующие конструкции определены на всех измеримых пространствах управлений $(U_1, \mathcal{B}_1), (U_2, \mathcal{B}_2), \dots, (U_N, \mathcal{B}_N)$.

Предположим, что управления случайным полумарковским процессом $\xi(t)$ осуществляются в моменты времени $t_n, n = 0, 1, 2, \dots$, непосредственно после изменения состояния процесса. Если $\xi_n = \xi(t_n) = i \in X$, то значение управления представляет собой случайную величину u_n , принимающую значения в множестве допустимых управлений U_i и описываемую вероятностной мерой (распределением вероятностей) $\Psi_i \in \Gamma_i$. Будем предполагать, что при фиксированном условии $\xi_n = \xi(t_n) = i$ управление определяется независимо от прошлого поведения процесса $\xi(t)$ и вероятностная мера Ψ_i , описывающая стохастическое управление u_n , зависит только от состояния $i \in X$.

Тогда выбор управлений в моменты изменения состояний $\{t_n, n = 0, 1, 2, \dots\}$ описывается набором вероятностных мер (распределений вероятностей) $(\Psi_1, \Psi_2, \dots, \Psi_N)$, $\Psi_i \in \Gamma_i, i = 1, 2, \dots, N$. Назовем любой такой набор стратегией управления полумарковским процессом $\xi(t)$. По своим свойствам такая стратегия является марковской, однородной и рандомизированной.

Следуя классической монографии П. Халмша [18, гл. VII], рассмотрим декартово произведение пространств $U = U_1 \times U_2 \times \dots \times U_N$ и соответствующих σ -алгебр $\mathcal{B} = \mathcal{B}_1 \times \mathcal{B}_2 \times \dots \times \mathcal{B}_N$. Обозначим через $\Psi = \Psi_1 \times \Psi_2 \times \dots \times \Psi_N$ вероятностную меру на (U, \mathcal{B}) , определяемую как произведение мер $\Psi_1, \Psi_2, \dots, \Psi_N$, где $\Psi_i \in \Gamma_i, i = 1, 2, \dots, N$. Обозначим также через Γ множество вероятностных мер Ψ , заданных на (U, \mathcal{B}) , которые представляют собой произведение мер $\Psi_1, \Psi_2, \dots, \Psi_N$, где $\Psi_i \in \Gamma_i, i = 1, 2, \dots, N$. Множество Γ можно отождествить с множеством всех стратегий управления полумарковским процессом $\xi(t)$.

Проблема оптимального управления полумарковским процессом $\xi(t)$ будет в дальнейшем сформулирована в виде задачи безусловного экстремума некоторого функционала $I(\Psi) = I(\Psi_1, \Psi_2, \dots, \Psi_N)$, заданного на множестве допустимых стратегий управления. Содержание показателя качества управления $I(\Psi)$, аналитическое представление для него, а также описание множества допустимых стратегий управления будут приведены в последующих разделах данной работы.

Для получения дальнейших результатов потребуются различные вероятностные характеристики управляемого полумарковского процесса $\xi(t)$. Как известно из общей теории полумарковских процессов [19, 20], основной вероятностной характеристикой такого процесса является так называемая полумарковская функция. Определим эту функцию для процесса с управлением (см. [8, гл. 5]):

$$Q_{ij}(t, u) = P(\xi_{n+1} = j, \theta_n < t \mid \xi_n = i, u_n = u), \\ t \in [0, \infty), u \in U_i; i, j \in X = \{1, 2, \dots, N\}. \quad (1)$$

Используя полумарковские функции, можно получить вероятности перехода управляемой цепи Маркова $\{\xi_n\}$:

$$p_{ij}(u) = P(\xi_{n+1} = j \mid \xi_n = i, u_n = u) = \\ = \lim_{t \rightarrow \infty} Q_{ij}(t, u), u \in U_i; i, j \in X, \quad (2)$$

а также функции распределения длительностей пребывания полумарковского процесса $\xi(t)$ в соответствующих состояниях:

$$H_i(t, u) = P(\theta_n < t \mid \xi_n = i, u_n = u) = \sum_{j \in X} Q_{ij}(t, u), \quad t \in [0, \infty), u \in U_i; i \in X. \quad (3)$$

Обозначим через

$$T_i(u) = \mathbf{E}[\theta_n \mid \xi_n = i, u_n = u] = \int_0^\infty [1 - H_i(t, u)] dt, \quad u \in U_i, i \in X, \quad (4)$$

математические ожидания длительностей пребывания полумарковского процесса $\xi(t)$ в каждом из состояний.

Введенные выше характеристики (1)–(4) определены для случая, когда в момент изменения состояния t_n процесс оказывается в состоянии i и принимается решение $u \in U_i$. При заданной стратегии управления $\Psi = (\Psi_1, \Psi_2, \dots, \Psi_N)$ можно записать соответствующие вероятностные характеристики без условия на управление, а именно:

$$Q_{ij}(t) = P(\xi_{n+1} = j, \theta_n < t \mid \xi_n = i) = \int_{U_i} Q_{ij}(t, u) d\Psi_i(u), \quad t \in [0, \infty), i, j \in X;$$

$$p_{ij} = P(\xi_{n+1} = j \mid \xi_n = i) = \int_{U_i} p_{ij}(u) d\Psi_i(u), \quad i, j \in X; \quad (5)$$

$$T_i = \mathbf{E}[\theta_n \mid \xi_n = i] = \int_{U_i} T_i(u) d\Psi_i(u), \quad i \in X. \quad (6)$$

В дальнейшем будем предполагать, что для рассматриваемой полумарковской модели заданы вероятностные характеристики $p_{ij}(u)$, $u \in U_i$, $i, j \in X$, и $T_i(u)$, $u \in U_i$, $i \in X$, определяемые соотношениями (2) и (4). Для фиксированной стратегии управления $\Psi = (\Psi_1, \Psi_2, \dots, \Psi_N)$ соответствующие вероятностные характеристики p_{ij} и T_i , $i, j \in X$, определены равенствами (5) и (6) без условий на управление.

3 Стационарный стоимостной показатель качества управления

Определим некоторый стоимостной аддитивный функционал, связанный с рассматриваемым полумарковским процессом $\xi(t)$. По содержанию этот функционал представляет собой случайный доход или прибыль, накопленную за период времени $[0, t]$. Определения такого функционала приведены в основополагающих работах [3; 8, гл. 5].

Обозначим через $\tilde{v}(t)$, $t \geq 0$, значение этого аддитивного функционала в момент времени t ; $\tilde{v}_n = \tilde{v}(t_n + 0)$ — соответствующее значение непосредственно после очередного момента изменения состояния t_n , $n = 0, 1, 2, \dots$; $\tilde{v}_0 = v_0$ — заданное начальное значение в момент $t = 0$. Рассмотрим величину

$$d_i(u) = \mathbf{E}[\tilde{v}_{n+1} - \tilde{v}_n \mid \xi_n = i, u_n = u], \quad u \in U_i, i \in X, \quad (7)$$

представляющую собой математическое ожидание приращения стоимостного аддитивного функционала за период времени между последовательными изменениями состояния полумарковского процесса $\xi(t)$. Тогда соответствующее математическое ожидание, вычисляемое без условия на решение, принимаемое в момент времени t_n , представляется в виде:

$$d_i = \mathbf{E}[\tilde{v}_{n+1} - \tilde{v}_n \mid \xi_n = i] = \int_{U_i} d_i(u) d\Psi_i(u), \quad i \in X.$$

Предположим, что для заданной стратегии управления $\Psi = (\Psi_1, \Psi_2, \dots, \Psi_N)$ вложенная цепь Маркова $\{\xi_n\}$ имеет ровно один класс возвратных положительных состояний (по терминологии, принятой в [8], такое множество состояний называется эргодическим классом). Как известно [15, гл. VIII], данное условие является необходимым и достаточным для существования единственного стационарного распределения. Обозначим это стационарное распределение цепи Маркова $\{\xi_n\}$ через $\pi = (\pi_1, \pi_2, \dots, \pi_N)$. Заметим, что данное распределение зависит от стратегии управления $\Psi = (\Psi_1, \Psi_2, \dots, \Psi_N)$. При указанном условии имеет место следующий результат, называемый эргодической теоремой для аддитивного стоимостного функционала:

$$I = \lim_{t \rightarrow \infty} \frac{\mathbf{E}\tilde{v}(t)}{t} = \frac{\sum_{i=1}^N d_i \pi_i}{\sum_{i=1}^N T_i \pi_i}. \quad (8)$$

Соотношение (8) доказано в работе [8, гл. 5]. Заметим, что аналогичные результаты имеют место для гораздо более общих полумарковских моделей [10, 11].

По своему прикладному содержанию величина, определяемая соотношением (8), представляет собой среднюю удельную прибыль, связанную с эволюцией системы в стационарном режиме. Кроме того, величина I представляет собой функционал от набора вероятностных распределений Ψ_i , $i \in \{1, \dots, N\}$, определяющих стратегию управле-

ния системой. В дальнейшем будем рассматривать стационарный стоимостной функционал $I = I(\Psi_1, \Psi_2, \dots, \Psi_N)$ как показатель качества управления системой и построенным полумарковским процессом $\xi(t)$.

4 Представление стационарного показателя в форме дробно-линейного интегрального функционала

В данном разделе будет приведено утверждение об аналитическом представлении стационарного стоимостного функционала (8), служащего критерием качества управления в рассматриваемой задаче управления полумарковским процессом.

Теорема 1. *Стационарный стоимостной показатель, определяемый равенством (8), представляет собой дробно-линейный функционал от вероятностных распределений $\Psi_i(u_i)$, $i \in \{1, \dots, N\}$. Данный функционал задается аналитически следующей формулой:*

$$I = I(\Psi_1, \dots, \Psi_N) = \frac{\int_{U_1} \dots \int_{U_N} A(u_1, \dots, u_N) d\Psi_1(u_1) \dots d\Psi_N(u_N)}{\int_{U_1} \dots \int_{U_N} B(u_1, \dots, u_N) d\Psi_1(u_1) \dots d\Psi_N(u_N)}, \quad (9)$$

где подынтегральные функции числителя и знаменателя выражаются соотношениями:

$$A(u_1, \dots, u_N) = \sum_{i=1}^N d_i(u_i) \widehat{D}^{(i)}(u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_N); \quad (10)$$

$$B(u_1, \dots, u_N) = \sum_{i=1}^N T_i(u_i) \widehat{D}^{(i)}(u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_N). \quad (11)$$

В свою очередь, функции $\widehat{D}^{(i)}(u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_N)$, $i \in \{1, \dots, N\}$, входящие в правые части формул (10) и (11), определяются следующим образом:

$$\widehat{D}^{(i)}(u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_N) = (-1)^{N+i+2} \sum_{\alpha^{(N),i}} (-1)^{\delta(\alpha^{(N),i})} \widehat{D}_0^{(i)}(\alpha^{(N),i}, u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_N). \quad (12)$$

Здесь $\alpha^{(N),i} = (\alpha_1, \dots, \alpha_{i-1}, \alpha_{i+1}, \dots, \alpha_N)$ — произвольная перестановка чисел $(1, \dots, i-1, i+1, \dots, N)$; $\delta(\alpha^{(N),i})$ — число инверсий в перестановке $\alpha^{(N),i}$;

$$\widehat{D}_0^{(i)}(\alpha^{(N),i}, u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_N) = \widetilde{p}_{1,\alpha_1}(u_1) \dots \widetilde{p}_{i-1,\alpha_{i-1}}(u_{i-1}) \widetilde{p}_{i+1,\alpha_{i+1}}(u_{i+1}) \dots \widetilde{p}_{N,\alpha_N}(u_N), \quad (13)$$

где

$$\widetilde{p}_{k,\alpha_k}(u_k) = \begin{cases} p_{kk}(u_k) - 1, & \alpha_k = k; \\ p_{k,\alpha_k}(u_k), & \alpha_k \neq k, \end{cases} \quad k = 1, \dots, i-1, i+1, \dots, N. \quad (14)$$

Функции $p_{ij}(u_i)$, $T_i(u_i)$ и $d_i(u_i)$, $u_i \in U_i$, $i, j \in \{1, 2, \dots, N\}$, входящие в соотношения (10)–(14), определяются равенствами (2), (4) и (7) соответственно.

Доказательство теоремы 1 в весьма сжатой форме приведено в работе [21]. Читателю, интересующемуся более подробным обоснованием данного результата, порекомендуем обратиться к тексту кандидатской диссертации А. В. Иванова [22, гл. 3].

Итак, теорема 1 позволяет получить явное аналитическое представление для стационарного стоимостного показателя вида (8) в форме дробно-линейного интегрального функционала от набора вероятностных мер $\Psi = (\Psi_1, \Psi_2, \dots, \Psi_N)$, задающих стратегию управления полумарковским процессом $\xi(t)$. При этом подынтегральные функции числителя и знаменателя задаются формулами (10), (11) и вспомогательными равенствами (12)–(14). Таким образом, функция

$$C(u_1, u_2, \dots, u_N) = \frac{A(u_1, u_2, \dots, u_N)}{B(u_1, u_2, \dots, u_N)}, \quad (15)$$

которая в дальнейшем будет называться основной функцией дробно-линейного интегрального функционала (9) и которая будет играть важную роль в дальнейшем исследовании, также явно определяется формулами (15), (10), (11).

5 Формальная постановка оптимизационной задачи и условия существования оптимальной стратегии управления полумарковским процессом

Будем рассматривать проблему управления полумарковским процессом $\xi(t)$ в форме экстремальной задачи

$$I(\Psi) = I(\Psi_1, \Psi_2, \dots, \Psi_N) \rightarrow \text{extr}, \quad \Psi = (\Psi_1, \Psi_2, \dots, \Psi_N) \in \Gamma. \quad (16)$$

При этом показатель качества управления $I(\Psi)$ представляет собой дробно-линейный интегральный функционал вида (9).

Для решения экстремальной задачи (16) воспользуемся некоторым утверждением об экстремуме дробно-линейного интегрального функционала. Прежде чем сформулировать данное утверждение, отметим, что в теории оптимизации хорошо известны задачи, в которых целевая функция представляет собой отношение двух линейных отображений, а имеющиеся ограничения также линейны. Такой раздел называется дробно-линейным программированием. Основные теоретические результаты данного направления изложены в работе [23], там же приведена подробная библиография. В дальнейшем потребуются некоторый специальный результат о безусловном экстремуме дробно-линейного интегрального функционала вида (9), который был впервые сформулирован в работе [14]. Заметим, что для использования этого результата необходимо, чтобы выполнялись некоторые предварительные условия, которые в данном случае можно сформулировать следующим образом:

1. Интегральные выражения

$$I_1(\Psi) = I_1(\Psi_1, \Psi_2, \dots, \Psi_N) = \int_{U_1} \dots \int_{U_N} A(u_1, \dots, u_N) d\Psi_1(u_1) \dots d\Psi_N(u_N);$$

$$I_2(\Psi) = I_2(\Psi_1, \Psi_2, \dots, \Psi_N) = \int_{U_1} \dots \int_{U_N} B(u_1, \dots, u_N) d\Psi_1(u_1) \dots d\Psi_N(u_N)$$

определены для всех стратегий управления $\Psi = (\Psi_1, \dots, \Psi_N) \in \Gamma$.

2. Функционал $I_2(\Psi) = I_2(\Psi_1, \dots, \Psi_N) \neq 0$ для всех $\Psi = (\Psi_1, \dots, \Psi_N) \in \Gamma$.
3. Множество Γ включает в себя множество всех вырожденных вероятностных мер: $\Gamma^* \subset \Gamma$.

Сделаем несколько важных замечаний по поводу введенных предварительных условий.

Замечание 1. Из условия 2 следует, что функция $B(u_1, u_2, \dots, u_N)$ не может принимать значения разных знаков. С учетом условия 3 получаем, что указанная функция должна обладать свойством строгой знакопостоянности на всем множестве U . С другой стороны, если выполняется условие строгой знакопостоянности функции $B(u_1, u_2, \dots, u_N)$, $(u_1, u_2, \dots, u_N) \in U$, то условие 2 выполняется автоматически.

Замечание 2. Если рассматривать в качестве целевого функционала $I(\Psi_1, \Psi_2, \dots, \Psi_N)$ экстремальной задачи (16) стационарный стоимостной показатель (8), то функция $B(u_1, u_2, \dots, u_N)$ имеет

следующее теоретическое содержание. Данная функция представляет собой условное математическое ожидание длительности периода времени между соседними моментами изменения состояния полумарковского процесса $\xi(t)$ при условии, что стратегия его управления является детерминированной и задается набором значений аргументов (u_1, u_2, \dots, u_N) . Тогда условие строгой положительности функции $B(u_1, u_2, \dots, u_N)$ при всех $(u_1, u_2, \dots, u_N) \in U$ является естественным и фактически означает, что при любой заданной детерминированной стратегии управления процесс $\xi(t)$ не имеет мгновенных состояний, длительность пребывания в которых равна нулю.

Замечание 3. Сделаем некоторые замечания, связанные с подынтегральной функцией числителя дробно-линейного интегрального функционала (9). Как и ранее, будем рассматривать в качестве целевого функционала $I(\Psi_1, \Psi_2, \dots, \Psi_N)$ экстремальной задачи (16) стационарный стоимостной показатель (8). Тогда для любого фиксированного набора значений аргументов $(u_1, u_2, \dots, u_N) \in U$ значение функции $A(u_1, u_2, \dots, u_N)$ представляет собой условное математическое ожидание приращения рассматриваемого стоимостного функционала, происшедшее за время пребывания полумарковского процесса $\xi(t)$ в некотором фиксированном состоянии при условии, что стратегия управления является детерминированной и задается указанным набором $(u_1, u_2, \dots, u_N) \in U$. Отметим, что в теореме об экстремуме дробно-линейного интегрального функционала, доказанной в работе [12, гл. 10], на подынтегральную функцию числителя накладываются условия ограниченности на всем множестве значений аргумента. Для многих математических моделей и связанных с ними задач оптимального управления такое условие является излишне ограничительным. В качестве примера можно привести модели оптимального управления запасом непрерывного продукта, рассмотренные в работах [24, 25]. В настоящем исследовании на функцию $A(u_1, u_2, \dots, u_N)$ накладывается только условие интегрируемости по любому заданному набору вероятностных мер $\Psi = (\Psi_1, \Psi_2, \dots, \Psi_N)$, образующему стратегию управления полумарковским процессом $\xi(t)$ (условие 1 системы предварительных условий).

Замечание 4. Условия 1–3 являются необходимыми для корректной постановки задачи безусловного экстремума дробно-линейного интегрального функционала. Если этот функционал служит показателем качества в задаче оптимального управления случайным процессом, то необходимо добавить к этим условиям дополнительное, связанное

с некоторой регулярностью самого управляемого процесса, а именно: некоторый содержательный показатель, связанный с поведением этого процесса, должен существовать и быть представимым в виде дробно-линейного интегрального функционала. Если потребовать, чтобы выполнялось эргодическое соотношение (8), то можно использовать теорему 1 и сформулировать задачу оптимального управления в виде (16) для дробно-линейного интегрального функционала (9). Таким образом, необходимо ввести условие, обеспечивающее существование единственного стационарного распределения вложенной цепи Маркова и выполнение соотношения (8). По аналогии с [8, гл. 5] сформулируем это дополнительное условие в следующем виде:

4. Для любой рассматриваемой стратегии управления $\Psi = (\Psi_1, \Psi_2, \dots, \Psi_N) \in \Gamma$ вложенная цепь Маркова полумарковского процесса $\xi(t)$ имеет ровно один класс возвратных положительных состояний.

Теперь определим понятие допустимой стратегии управления полумарковским процессом с конечным множеством состояний.

Определение 2. Назовем стратегию управления $\Psi = (\Psi_1, \Psi_2, \dots, \Psi_N)$ допустимой в данной задаче, если она удовлетворяет условиям 1–4.

Замечание 5. Как следует из замечания 1, если потребовать, чтобы функция $B(u_1, u_2, \dots, u_N)$ являлась строго знакопостоянной при всех $(u_1, u_2, \dots, u_N) \in U$, то можно считать допустимыми стратегии $(\Psi_1, \Psi_2, \dots, \Psi_N)$, удовлетворяющие условиям 1, 3, 4. С учетом замечания 2 о естественном характере условия строгой знакопостоянности функции $B(u_1, u_2, \dots, u_N)$ при всех значениях аргументов $(u_1, u_2, \dots, u_N) \in U$ будем требовать выполнения этого условия в формулировке приводимой в дальнейшем основной теоремы об оптимальной стратегии управления полумарковским процессом.

Замечание 6. Ниже будет сформулирована и доказана основная теорема об оптимальной стратегии управления полумарковским процессом с конечным множеством состояний. Будем формулировать эту теорему по отношению к экстремальной задаче (16), в которой целевой функционал $I(\Psi_1, \Psi_2, \dots, \Psi_N)$ имеет вид дробно-линейного интегрального функционала. Это обстоятельство связано с тем, что целевой функционал в задаче оптимального управления необязательно должен иметь характер стационарного стоимостного показателя вида (8). В частности, еще в 1983 г. П. В. Шнурковым было установлено [26], что ряд показателей, связанных с временем пребывания управляемого

полумарковского процесса в заданном конечном подмножестве состояний, имеет структуру дробно-линейного интегрального функционала от набора вероятностных мер, определяющих стратегию управления. Таким образом, рассматриваемая задача управления имеет более общий характер, чем задача, в которой целевой функционал представляет собой стационарный стоимостной показатель вида (8).

Замечание 7. Если рассматривать задачу оптимального управления полумарковским процессом, в которой целевой функционал не совпадает со стационарным стоимостным показателем (8), то возможно, что могут потребоваться другие дополнительные условия, обеспечивающие существование этого показателя и его представление в форме (9). В связи с этим в формулировке основной теоремы будем использовать термин допустимые стратегии в широком смысле, имея в виду выполнение всех необходимых условий для каждого рассматриваемого показателя качества управления.

Замечание 8. Множество допустимых стратегий может не совпадать с множеством всех возможных стратегий управления. В частности, допустимые стратегии могут состоять только из дискретных вероятностных мер $\Psi_1, \Psi_2, \dots, \Psi_N$, т. е. таких, которые сосредоточены на дискретных множествах точек пространств U_1, U_2, \dots, U_N .

6 Теоретическое решение задачи оптимального управления

Перейдем к формулировке и доказательству теоремы об оптимальной стратегии управления полумарковским процессом с конечным множеством состояний.

Теорема 2. *Рассмотрим проблему оптимального управления полумарковским процессом $\xi(t)$ в виде экстремальной задачи (16), определенной на множестве допустимых стратегий Γ , для дробно-линейного функционала (9). Пусть функция $B(u_1, u_2, \dots, u_N)$, входящая в определение функционала (9), является строго знакопостоянной (строго положительной или строго отрицательной) при всех значениях аргументов $(u_1, u_2, \dots, u_N) \in U$. Тогда справедливы следующие утверждения:*

1. Если функция $C(u_1, u_2, \dots, u_N) = A(u_1, u_2, \dots, u_N) / B(u_1, u_2, \dots, u_N)$ ограничена сверху или снизу и достигает глобального экстремума на множестве $U = U_1 \times U_2 \times \dots \times U_N$ (максимума или минимума), то оптимальная стратегия управления полумарковским процессом $\xi(t)$ су-

существует, является детерминированной и определяется вырожденной вероятностной мерой $\Psi^* \in \Gamma^*$, сосредоточенной в точке, в которой достигается соответствующего экстремума функция $C(u_1, u_2, \dots, u_N)$, и при этом выполняются соотношения:

$$\begin{aligned} \max_{\Psi \in \Gamma} I(\Psi) &= \max_{\substack{\Psi_i \in \Gamma_i^*, \\ i=1, N}} I(\Psi_1, \Psi_2, \dots, \Psi_N) = \\ &= \max_{\substack{\Psi_i^* \in \Gamma_i^*, \\ i=1, N}} I(\Psi_1^*, \Psi_2^*, \dots, \Psi_N^*) = \\ &= \max_{(u_1, u_2, \dots, u_N) \in U} \frac{A(u_1, u_2, \dots, u_N)}{B(u_1, u_2, \dots, u_N)}; \end{aligned} \quad (17)$$

$$\begin{aligned} \min_{\Psi \in \Gamma} I(\Psi) &= \min_{\substack{\Psi_i \in \Gamma_i, \\ i=1, N}} I(\Psi_1, \Psi_2, \dots, \Psi_N) = \\ &= \min_{\substack{\Psi_i^* \in \Gamma_i^*, \\ i=1, N}} I(\Psi_1^*, \Psi_2^*, \dots, \Psi_N^*) = \\ &= \min_{(u_1, u_2, \dots, u_N) \in U} \frac{A(u_1, u_2, \dots, u_N)}{B(u_1, u_2, \dots, u_N)}. \end{aligned}$$

2. Если функция $C(u_1, u_2, \dots, u_N) = A(u_1, u_2, \dots, u_N)/B(u_1, u_2, \dots, u_N)$ ограничена сверху или снизу, но не достигает глобального экстремума на множестве $U = U_1 \times U_2 \times \dots \times U_N$, то для любого $\varepsilon > 0$ можно выбрать ε -оптимальную детерминированную стратегию управления полумарковским процессом $\xi(t)$, которая определяется вырожденной вероятностной мерой $\Psi^{*(+)}(\varepsilon) \in \Gamma^*$ или вырожденной вероятностной мерой $\Psi^{*(-)}(\varepsilon) \in \Gamma^*$, в зависимости от вида экстремума (максимума или минимума) в задаче (16). При этом вероятностная мера $\Psi^{*(+)}(\varepsilon) \in \Gamma^*$ может быть сосредоточена в любой точке $(u_1^{(+)}(\varepsilon), u_2^{(+)}(\varepsilon), \dots, u_N^{(+)}(\varepsilon))$, удовлетворяющей соотношению:

$$\begin{aligned} \sup_{(u_1, u_2, \dots, u_N) \in U} \frac{A(u_1, u_2, \dots, u_N)}{B(u_1, u_2, \dots, u_N)} - \varepsilon < \\ < \frac{A(u_1^{(+)}(\varepsilon), u_2^{(+)}(\varepsilon), \dots, u_N^{(+)}(\varepsilon))}{B(u_1^{(+)}(\varepsilon), u_2^{(+)}(\varepsilon), \dots, u_N^{(+)}(\varepsilon))} < \\ < \sup_{(u_1, u_2, \dots, u_N) \in U} \frac{A(u_1, u_2, \dots, u_N)}{B(u_1, u_2, \dots, u_N)} < \infty, \end{aligned} \quad (18)$$

если функция $C(u_1, u_2, \dots, u_N)$ ограничена сверху и экстремальная задача (16) представляет собой задачу на максимум. Аналогично вероятностная мера $\Psi^{*(-)}(\varepsilon) \in \Gamma^*$ может быть сосредоточена в любой точке $(u_1^{(-)}(\varepsilon), u_2^{(-)}(\varepsilon), \dots, u_N^{(-)}(\varepsilon))$, удовлетворяющей соотношению:

$$\begin{aligned} -\infty < \inf_{(u_1, u_2, \dots, u_N) \in U} \frac{A(u_1, u_2, \dots, u_N)}{B(u_1, u_2, \dots, u_N)} < \\ < \frac{A(u_1^{(-)}(\varepsilon), u_2^{(-)}(\varepsilon), \dots, u_N^{(-)}(\varepsilon))}{B(u_1^{(-)}(\varepsilon), u_2^{(-)}(\varepsilon), \dots, u_N^{(-)}(\varepsilon))} < \\ < \inf_{(u_1, u_2, \dots, u_N) \in U} \frac{A(u_1, u_2, \dots, u_N)}{B(u_1, u_2, \dots, u_N)} + \varepsilon, \end{aligned}$$

если функция $C(u_1, u_2, \dots, u_N)$ ограничена снизу и экстремальная задача (16) представляет собой задачу на минимум.

3. Если функция $C(u_1, u_2, \dots, u_N) = A(u_1, u_2, \dots, u_N)/B(u_1, u_2, \dots, u_N)$ не ограничена сверху или снизу, то оптимальной стратегии управления в смысле соответствующей экстремальной задачи не существует. При этом найдется такая последовательность вырожденных вероятностных мер $\Psi^{*(+)}(n)$, сосредоточенных в точках $(u_1^{(+)}(n), u_2^{(+)}(n), \dots, u_N^{(+)}(n))$, $n = 1, 2, \dots$, для которых выполняется соотношение:

$$\begin{aligned} I(\Psi^{*(+)}(n)) &= \\ &= I(\Psi_1^{*(+)}(n), \Psi_2^{*(+)}(n), \dots, \Psi_N^{*(+)}(n)) = \\ &= \frac{A(u_1^{(+)}(n), u_2^{(+)}(n), \dots, u_N^{(+)}(n))}{B(u_1^{(+)}(n), u_2^{(+)}(n), \dots, u_N^{(+)}(n))} \rightarrow \infty \\ &\text{при } n \rightarrow \infty, \end{aligned}$$

если функция $C(u_1, u_2, \dots, u_N)$ не ограничена сверху. Аналогично найдется такая последовательность вырожденных вероятностных мер $\Psi^{*(-)}(n)$, сосредоточенных в точках $(u_1^{(-)}(n), u_2^{(-)}(n), \dots, u_N^{(-)}(n))$, $n = 1, 2, \dots$, для которых выполняется соотношение:

$$\begin{aligned} I(\Psi^{*(-)}(n)) &= \\ &= I(\Psi_1^{*(-)}(n), \Psi_2^{*(-)}(n), \dots, \Psi_N^{*(-)}(n)) = \\ &= \frac{A(u_1^{(-)}(n), u_2^{(-)}(n), \dots, u_N^{(-)}(n))}{B(u_1^{(-)}(n), u_2^{(-)}(n), \dots, u_N^{(-)}(n))} \rightarrow -\infty \\ &\text{при } n \rightarrow \infty, \end{aligned}$$

если функция $C(u_1, u_2, \dots, u_N)$ не ограничена снизу.

При этом сформулированные утверждения каждого пункта теоремы 2 могут выполняться как по отдельности, для одного из двух видов экстремума, так и совместно, для обоих видов экстремума.

Прежде чем непосредственно доказывать теорему 2, докажем некоторые вспомогательные утверждения.

Лемма 1. *Рассмотрим дробно-линейный интегральный функционал $I(\Psi_1, \Psi_2, \dots, \Psi_N)$ вида (9), заданный на некотором множестве наборов вероятностных мер $\Psi = (\Psi_1, \Psi_2, \dots, \Psi_N) \in \Gamma$. Предположим, что на множестве Γ выполняется условие 1 из набора предварительных условий и функция $B(u_1, u_2, \dots, u_N)$ обладает свойством строгой знакопостоянности при всех $(u_1, u_2, \dots, u_N) \in U$. Тогда справедливы следующие утверждения:*

1. *Если основная функция $C(u_1, u_2, \dots, u_N) = A(u_1, u_2, \dots, u_N)/B(u_1, u_2, \dots, u_N)$ ограничена сверху, т. е. выполняется условие*

$$C(u_1, u_2, \dots, u_N) = \frac{A(u_1, u_2, \dots, u_N)}{B(u_1, u_2, \dots, u_N)} \leq c_0^{(+)} < \infty, \quad (u_1, u_2, \dots, u_N) \in U, \quad (19)$$

то имеет место неравенство:

$$I(\Psi_1, \Psi_2, \dots, \Psi_N) \leq c_0^{(+)} \quad (20)$$

для всех $(\Psi_1, \Psi_2, \dots, \Psi_N) \in \Gamma$.

2. *Если основная функция $C(u_1, u_2, \dots, u_N) = A(u_1, u_2, \dots, u_N)/B(u_1, u_2, \dots, u_N)$ ограничена снизу, т. е. выполняется условие*

$$C(u_1, u_2, \dots, u_N) = \frac{A(u_1, u_2, \dots, u_N)}{B(u_1, u_2, \dots, u_N)} \geq c_0^{(-)} > -\infty, \quad (u_1, u_2, \dots, u_N) \in U,$$

то имеет место неравенство:

$$I(\Psi_1, \Psi_2, \dots, \Psi_N) \geq c_0^{(-)}$$

для всех $(\Psi_1, \Psi_2, \dots, \Psi_N) \in \Gamma$.

Доказательство леммы 1. Докажем первое утверждение леммы. Предположим сначала, что функция $B(u_1, u_2, \dots, u_N)$ строго положительна:

$$B(u_1, u_2, \dots, u_N) > 0, \quad (u_1, u_2, \dots, u_N) \in U. \quad (21)$$

Заметим, что в таком случае по свойству интеграла [18, гл. V]

$$\int \dots \int_{U_1 \dots U_N} B(u_1, \dots, u_N) d\Psi_1(u_1) \dots d\Psi_N(u_N) > 0 \quad (22)$$

для любого фиксированного набора $\Psi = (\Psi_1, \dots, \Psi_N) \in \Gamma$. Из неравенства (19) с учетом (21) получаем:

$$A(u_1, \dots, u_N) \leq c_0^{(+)} B(u_1, \dots, u_N), \quad (u_1, \dots, u_N) \in U. \quad (23)$$

В свою очередь, из неравенства (23) и свойств интеграла следует:

$$\begin{aligned} & \int \dots \int_{U_1 \dots U_N} A(u_1, \dots, u_N) d\Psi_1(u_1) \dots d\Psi_N(u_N) \leq \\ & \leq c_0^{(+)} \int \dots \int_{U_1 \dots U_N} B(u_1, \dots, u_N) d\Psi_1(u_1) \dots d\Psi_N(u_N) \quad (24) \end{aligned}$$

для любого фиксированного набора $\Psi = (\Psi_1, \dots, \Psi_N) \in \Gamma$. Но тогда из (24) с учетом (22) получаем:

$$\begin{aligned} I(\Psi_1, \dots, \Psi_N) &= \\ &= \frac{\int \dots \int_{U_1 \dots U_N} A(u_1, \dots, u_N) d\Psi_1(u_1) \dots d\Psi_N(u_N)}{\int \dots \int_{U_1 \dots U_N} B(u_1, \dots, u_N) d\Psi_1(u_1) \dots d\Psi_N(u_N)} \leq \\ & \leq c_0^{(+)} \quad (25) \end{aligned}$$

для любого фиксированного набора $(\Psi_1, \dots, \Psi_N) \in \Gamma$.

Предположим теперь, что функция $B(u_1, \dots, u_N)$ строго отрицательна:

$$B(u_1, \dots, u_N) < 0 \quad (u_1, \dots, u_N) \in U. \quad (26)$$

Тогда

$$\int \dots \int_{U_1 \dots U_N} B(u_1, \dots, u_N) d\Psi_1(u_1) \dots d\Psi_N(u_N) < 0 \quad (27)$$

для любого фиксированного набора $(\Psi_1, \dots, \Psi_N) \in \Gamma$.

Как и ранее, будем исходить из неравенства (19). При выполнении условий (26) и (27) характер неравенств (23) и (24) меняется на противоположный, но характер неравенства (25) остается неизменным. Таким образом, для любой функции $B(u_1, u_2, \dots, u_N)$, обладающей свойством строгой знакопостоянности, из условия (19) следует выполнение неравенства (25), которое совпадает с (20). Первое утверждение леммы 1 доказано. Второе утверждение доказывается аналогично. Лемма 1 доказана.

Лемма 2. *Рассмотрим дробно-линейный интегральный функционал $I(\Psi_1, \Psi_2, \dots, \Psi_N)$ вида (9), заданный на некотором множестве наборов вероятностных мер $\Psi = (\Psi_1, \Psi_2, \dots, \Psi_N) \in \Gamma$. Предположим, что на множестве Γ выполняется условие 1 из набора предварительных условий и функция $B(u_1, u_2, \dots, u_N)$ обладает свойством строгой знакопостоянности при всех $(u_1, u_2, \dots, u_N) \in U$. Тогда справедливы следующие утверждения:*

1. Если основная функция $C(u_1, u_2, \dots, u_N) = A(u_1, u_2, \dots, u_N)/B(u_1, u_2, \dots, u_N)$ ограничена сверху, но не достигает своего максимального значения, то имеет место неравенство:

$$I(\Psi_1, \Psi_2, \dots, \Psi_N) < \sup_{(u_1, u_2, \dots, u_N) \in U} C(u_1, u_2, \dots, u_N) < \infty \quad (28)$$

для всех $(\Psi_1, \Psi_2, \dots, \Psi_N) \in \Gamma$.

2. Если основная функция $C(u_1, u_2, \dots, u_N) = A(u_1, u_2, \dots, u_N)/B(u_1, u_2, \dots, u_N)$ ограничена снизу, но не достигает своего минимального значения, то имеет место неравенство:

$$I(\Psi_1, \Psi_2, \dots, \Psi_N) > \inf_{(u_1, u_2, \dots, u_N) \in U} C(u_1, u_2, \dots, u_N) > -\infty$$

для всех $(\Psi_1, \Psi_2, \dots, \Psi_N) \in \Gamma$.

Доказательство леммы 2. Докажем первое утверждение леммы. Поскольку множество значений основной функции $C(u_1, u_2, \dots, u_N)$ ограничено сверху, оно имеет конечную верхнюю грань:

$$\exists \sup_{(u_1, u_2, \dots, u_N) \in U} C(u_1, u_2, \dots, u_N) < \infty$$

(см. [27, гл. 1, §3, п. 3.4, теорема 1]).

По условию функция $C(u_1, u_2, \dots, u_N)$ не достигает своего максимального значения. Следовательно, выполняется неравенство:

$$C(u_1, u_2, \dots, u_N) = \frac{A(u_1, u_2, \dots, u_N)}{B(u_1, u_2, \dots, u_N)} < \sup_{(u_1, u_2, \dots, u_N) \in U} C(u_1, u_2, \dots, u_N) < \infty, \quad (u_1, u_2, \dots, u_N) \in U. \quad (29)$$

Взяв за основу строгое неравенство (29), проведем рассуждения, аналогичные тем, которые были проведены в лемме 1 по отношению к неравенству (19). В результате получим строгое неравенство (28).

Второе утверждение леммы 2 доказывается аналогично. Лемма 2 доказана.

Доказательство теоремы 2. Начнем с доказательства утверждения 1. Предположим сначала, что основная функция $C(u_1, u_2, \dots, u_N) = A(u_1, u_2, \dots, u_N)/B(u_1, u_2, \dots, u_N)$ ограничена сверху и достигает глобального максимума на множестве U в некоторой точке $u^{(+)} = (u_1^{(+)}, u_2^{(+)}, \dots, u_N^{(+)}) \in U$, а именно:

$$\max_{(u_1, u_2, \dots, u_N) \in U} C(u_1, u_2, \dots, u_N) = C(u_1^{(+)}, u_2^{(+)}, \dots, u_N^{(+)}) < \infty.$$

Тогда выполняется соотношение:

$$C(u_1, u_2, \dots, u_N) = \frac{A(u_1, u_2, \dots, u_N)}{B(u_1, u_2, \dots, u_N)} \leq C(u_1^{(+)}, u_2^{(+)}, \dots, u_N^{(+)}) < \infty, \quad (u_1, u_2, \dots, u_N) \in U. \quad (30)$$

Условия леммы 1 выполнены, и можно воспользоваться ее утверждениями. Согласно первому из них, если выполняется неравенство (30), то имеет место соотношение:

$$I(\Psi_1, \Psi_2, \dots, \Psi_N) \leq C(u_1^{(+)}, u_2^{(+)}, \dots, u_N^{(+)}) < \infty$$

для всех стратегий управления $\Psi = (\Psi_1, \Psi_2, \dots, \Psi_N) \in \Gamma$.

Таким образом, множество значений дробно-линейного интегрального функционала $I(\Psi_1, \Psi_2, \dots, \Psi_N)$ ограничено сверху при всех $\Psi = (\Psi_1, \Psi_2, \dots, \Psi_N) \in \Gamma$. Тогда существует верхняя грань этого множества и выполняется неравенство:

$$\sup_{(\Psi_1, \Psi_2, \dots, \Psi_N) \in \Gamma} I(\Psi_1, \Psi_2, \dots, \Psi_N) \leq C(u_1^{(+)}, u_2^{(+)}, \dots, u_N^{(+)}) \quad (31)$$

Рассмотрим детерминированную стратегию управления $\Psi^{*(+)} = (\Psi_1^{*(+)}, \Psi_2^{*(+)}, \dots, \Psi_N^{*(+)})$, в которой каждая вероятностная мера $\Psi_i^{*(+)}$ является вырожденной и сосредоточена в точке $u_i^{(+)}$, $i = \overline{1, N}$. По свойству интеграла

$$I(\Psi_1^{*(+)}, \Psi_2^{*(+)}, \dots, \Psi_N^{*(+)}) = C(u_1^{(+)}, u_2^{(+)}, \dots, u_N^{(+)}) \quad (32)$$

Из соотношений (31) и (32) получаем:

$$\sup_{(\Psi_1, \Psi_2, \dots, \Psi_N) \in \Gamma} I(\Psi_1, \Psi_2, \dots, \Psi_N) \leq I(\Psi_1^{*(+)}, \Psi_2^{*(+)}, \dots, \Psi_N^{*(+)}) \quad (33)$$

Заметим дополнительно, что выполняются отношения принадлежности:

$$\Psi^{*(+)} = (\Psi_1^{*(+)}, \Psi_2^{*(+)}, \dots, \Psi_N^{*(+)}) \in \Gamma^* \subset \Gamma. \quad (34)$$

Из (34) и свойства верхней грани следует:

$$\sup_{(\Psi_1^*, \Psi_2^*, \dots, \Psi_N^*) \in \Gamma^*} I(\Psi_1^*, \Psi_2^*, \dots, \Psi_N^*) \leq \sup_{(\Psi_1, \Psi_2, \dots, \Psi_N) \in \Gamma} I(\Psi_1, \Psi_2, \dots, \Psi_N) \quad (35)$$

Объединяя (32), (33) и (35), получаем соотношение:

$$\begin{aligned} & \sup_{(\Psi_1^*, \Psi_2^*, \dots, \Psi_N^*) \in \Gamma^*} I(\Psi_1^*, \Psi_2^*, \dots, \Psi_N^*) \leq \\ & \leq \sup_{(\Psi_1, \Psi_2, \dots, \Psi_N) \in \Gamma} I(\Psi_1, \Psi_2, \dots, \Psi_N) \leq \\ & \leq I(\Psi_1^{*(+)}, \Psi_2^{*(+)}, \dots, \Psi_N^{*(+)}) = \\ & = \frac{A(u_1^{(+)}, u_2^{(+)}, \dots, u_N^{(+)})}{B(u_1^{(+)}, u_2^{(+)}, \dots, u_N^{(+)})}. \end{aligned} \quad (36)$$

Из соотношения (36) с учетом (34) получаем, что максимум функционала $I(\Psi_1, \Psi_2, \dots, \Psi_N)$ на множестве допустимых стратегий $\Psi = (\Psi_1, \Psi_2, \dots, \Psi_N) \in \Gamma$ существует и достигается на детерминированной стратегии $(\Psi_1^{*(+)}, \Psi_2^{*(+)}, \dots, \Psi_N^{*(+)})$.

Кроме того, выполняются соотношения (17). Таким образом, утверждение 1 в случае, когда основная функция $C(u_1, u_2, \dots, u_N)$ достигает глобального максимума, доказано. Соответствующее утверждение в случае, когда основная функция $C(u_1, u_2, \dots, u_N)$ достигает глобального минимума, доказывается аналогично. При этом используется второе утверждение леммы 1.

Перейдем к доказательству второго утверждения теоремы 2. Предположим, что основная функция $C(u_1, u_2, \dots, u_N) = A(u_1, u_2, \dots, u_N)/B(u_1, u_2, \dots, u_N)$ ограничена сверху, но не достигает глобального максимума на множестве $U = U_1 \times U_2 \times \dots \times U_N$. Тогда множество значений основной функции имеет конечную верхнюю грань:

$$\begin{aligned} C(u_1, u_2, \dots, u_N) &= \frac{A(u_1, u_2, \dots, u_N)}{B(u_1, u_2, \dots, u_N)} < \\ &< \sup_{(u_1, u_2, \dots, u_N) \in U} \frac{A(u_1, u_2, \dots, u_N)}{B(u_1, u_2, \dots, u_N)} < \infty, \\ & \quad (u_1, u_2, \dots, u_N) \in U. \end{aligned}$$

По определению верхней грани для любого фиксированного $\varepsilon > 0$ существует точка $(u_1^{(+)}(\varepsilon), u_2^{(+)}(\varepsilon), \dots, u_N^{(+)}(\varepsilon))$ такая, что выполняется двойное неравенство (18) (см. [27, гл. 1, § 3, п. 3.4]). Иначе говоря, значение основной функции в указанной точке лежит в левой ε -окрестности верхней грани. Рассмотрим детерминированную стратегию управления $\Psi^{*(+)}(\varepsilon) = (\Psi_1^{*(+)}(\varepsilon), \Psi_2^{*(+)}(\varepsilon), \dots, \Psi_N^{*(+)}(\varepsilon))$, компонентами которой являются вырожденные вероятностные меры $\Psi_1^{*(+)}(\varepsilon), \Psi_2^{*(+)}(\varepsilon), \dots, \Psi_N^{*(+)}(\varepsilon)$, причем вырожденная мера $\Psi_i^{*(+)}(\varepsilon)$ сосредоточена в точке $u_i^{(+)}(\varepsilon), i = 1, 2, \dots, N$.

По свойству интеграла

$$\begin{aligned} I(\Psi_1^{*(+)}(\varepsilon), \Psi_2^{*(+)}(\varepsilon), \dots, \Psi_N^{*(+)}(\varepsilon)) &= \\ &= C(u_1^{(+)}(\varepsilon), u_2^{(+)}(\varepsilon), \dots, u_N^{(+)}(\varepsilon)). \end{aligned} \quad (37)$$

Из соотношения (37) с учетом указанного свойства основной функции получаем:

$$\begin{aligned} & \sup_{(u_1, u_2, \dots, u_N) \in U} \frac{A(u_1, u_2, \dots, u_N)}{B(u_1, u_2, \dots, u_N)} - \varepsilon < \\ & < I(\Psi_1^{*(+)}(\varepsilon), \Psi_2^{*(+)}(\varepsilon), \dots, \Psi_N^{*(+)}(\varepsilon)) < \\ & < \sup_{(u_1, u_2, \dots, u_N) \in U} \frac{A(u_1, u_2, \dots, u_N)}{B(u_1, u_2, \dots, u_N)} < \infty. \end{aligned} \quad (38)$$

Заметим также, что в рассматриваемом случае выполнены условия леммы 2. Воспользуемся первым утверждением этой леммы, а именно соотношением (28):

$$\begin{aligned} I(\Psi_1, \Psi_2, \dots, \Psi_N) &< \\ &< \sup_{(u_1, u_2, \dots, u_N) \in U} \frac{A(u_1, u_2, \dots, u_N)}{B(u_1, u_2, \dots, u_N)} < \infty \end{aligned} \quad (39)$$

для всех $(\Psi_1, \Psi_2, \dots, \Psi_N) \in \Gamma$.

Из соотношений (38) и (39) следует, что детерминированная стратегия $\Psi^{*(+)}(\varepsilon) = (\Psi_1^{*(+)}(\varepsilon), \Psi_2^{*(+)}(\varepsilon), \dots, \Psi_N^{*(+)}(\varepsilon))$, определяемая набором вырожденных вероятностных мер, сосредоточенных в соответствующих точках $(u_1^{(+)}(\varepsilon), u_2^{(+)}(\varepsilon), \dots, u_N^{(+)}(\varepsilon))$, является ε -оптимальной. Вторая часть утверждения 2 теоремы 2, связанная со свойствами нижней грани, доказывается аналогично.

Докажем третье утверждение теоремы 2. Предположим, что множество значений основной функции $C(u_1, u_2, \dots, u_N) = A(u_1, u_2, \dots, u_N)/B(u_1, u_2, \dots, u_N)$ не является ограниченным сверху на множестве $U = U_1 \times U_2 \times \dots \times U_N$. Тогда существует последовательность точек $(u_1^{(+)}(n), u_2^{(+)}(n), \dots, u_N^{(+)}(n)) \in U, n = 1, 2, \dots$, для которой

$$\begin{aligned} C(u_1^{(+)}(n), u_2^{(+)}(n), \dots, u_N^{(+)}(n)) &= \\ &= \frac{A(u_1^{(+)}(n), u_2^{(+)}(n), \dots, u_N^{(+)}(n))}{B(u_1^{(+)}(n), u_2^{(+)}(n), \dots, u_N^{(+)}(n))} \rightarrow \infty, \\ & \quad n \rightarrow \infty. \end{aligned} \quad (40)$$

Зафиксируем некоторую последовательность точек $(u_1^{(+)}(n), u_2^{(+)}(n), \dots, u_N^{(+)}(n)) \in U, n = 1, 2, \dots$, обладающих указанным свойством, и рассмотрим последовательность детерминированных стратегий управления $\Psi^{*(+)}(n) = (\Psi_1^{*(+)}(n), \Psi_2^{*(+)}(n), \dots, \Psi_N^{*(+)}(n)), n = 1, 2, \dots$, определяемых набором вырожденных вероятностных мер, сосредоточенных в соответствующих точках $(u_1^{(+)}(n), u_2^{(+)}(n), \dots, u_N^{(+)}(n)), n = 1, 2, \dots$. По свойству интеграла для любого фиксированного значения $n = 1, 2, \dots$ выполняется равенство:

$$\begin{aligned} I(\Psi^{*(+)}(n)) &= \\ &= I(\Psi_1^{*(+)}(n), \Psi_2^{*(+)}(n), \dots, \Psi_N^{*(+)}(n)) = \\ &= \frac{A(u_1^{(+)}(n), u_2^{(+)}(n), \dots, u_N^{(+)}(n))}{B(u_1^{(+)}(n), u_2^{(+)}(n), \dots, u_N^{(+)}(n))}. \end{aligned} \quad (41)$$

Из соотношений (40) и (41) следует, что

$$\begin{aligned} I(\Psi^{*(+)}(n)) &= \\ &= I(\Psi_1^{*(+)}(n), \Psi_2^{*(+)}(n), \dots, \Psi_N^{*(+)}(n)) \longrightarrow \infty, \\ & \qquad \qquad \qquad n \rightarrow \infty. \end{aligned} \quad (42)$$

Соотношение (42) означает, что множество значений дробно-линейного интегрального функционала $I(\Psi_1, \Psi_2, \dots, \Psi_N)$ вида (9) не ограничено сверху на множестве наборов вырожденных вероятностных мер $(\Psi_1^{*(+)}(n), \Psi_2^{*(+)}(n), \dots, \Psi_N^{*(+)}(n)) \in \Gamma^*$, а следовательно, и на более широком множестве наборов вероятностных мер $(\Psi_1, \Psi_2, \dots, \Psi_N) \in \Gamma$. В таком случае решения экстремальной задачи (16) в форме задачи на максимум не существует. Соответствующее утверждение для варианта, когда множество значений основной функции $C(u_1, u_2, \dots, u_N) = A(u_1, u_2, \dots, u_N)/B(u_1, u_2, \dots, u_N)$ не является ограниченным снизу, доказывается аналогично. Третье утверждение теоремы 2 доказано. Тем самым теорема 2 доказана полностью.

Применим теорему 2 для решения поставленной задачи оптимального управления. Из утверждения этой теоремы следует, что для доказательства существования оптимального управления и его нахождения необходимо исследовать на глобальный экстремум основную функцию дробно-линейного интегрального функционала $C(u_1, u_2, \dots, u_N)$, определяемую формулой (15) с учетом равенств (10)–(14). В некоторых случаях, например когда основной процесс $\xi(t)$ является регенерирующим, а стоимостные характеристики модели

задаются линейными функциями, такое исследование можно провести аналитически. Однако для подавляющего большинства полумарковских моделей для этого необходимо использовать численные методы.

7 Заключение

В заключительной части работы приведем краткое описание теоретической основы метода решения задачи оптимального управления полумарковским процессом с конечным множеством состояний.

1. Исходная проблема оптимального управления формулируется в виде экстремальной задачи (16). Целевым показателем качества управления в данной задаче служит величина (8), которая имеет характер средней удельной прибыли.
2. Доказывается, что стационарный показатель (8) представим в виде дробно-линейного интегрального функционала (9), для которого явно определяются подынтегральные функции числителя и знаменателя, а следовательно, и основная функция данного функционала.
3. Используется теорема об экстремуме дробно-линейного интегрального функционала. На основании утверждений этой теоремы устанавливается, что исходная задача оптимального управления сводится к исследованию на глобальный экстремум основной функции этого функционала, для которой получено явное аналитическое представление.

Заметим, что такое исследование задач оптимального управления стохастическими системами фактически уже было проведено в ряде работ П. В. Шнуркова и его соавторов. В частности, в работе [28] была рассмотрена модель управления для обрывающегося процесса восстановления, описывающего функционирование некоторой технической системы. Задача управления решалась для различных показателей эффективности и надежности этой системы, имеющих структуру дробно-линейного интегрального функционала.

В работах [24, 25] рассматривались модели регенерирующих процессов для исследования систем управления запасами. Различные показатели качества управления были представлены в форме дробно-линейных интегральных функционалов. Основные функции этих функционалов были найдены в явной форме и исследовались на

глобальный экстремум. В работах [21, 29] рассматривалась достаточно сложная полумарковская модель с конечным множеством состояний, описывающая систему управления запасом непрерывного продукта. Показатели качества управления в этой модели также имели структуру дробно-линейных интегральных функционалов, для основных функций которых были найдены явные аналитические представления. Упомянем также работы [30, 31], в которых была исследована полумарковская модель с дискретно-непрерывным фазовым пространством. Показатели качества управления в этой модели были найдены в явной форме как функции от двух непрерывных параметров управления.

Фактически во всех упомянутых работах уже был использован метод решения задачи оптимального управления регенерирующим или полумарковским случайным процессом, основанный на исследовании экстремальных свойств основной функции соответствующего дробно-линейного интегрального функционала. Из соображений, изложенных во введении, следует, что в период написания и публикации этих работ данный метод не имел строгого обоснования. Однако после публикации работы [14] и настоящего исследования можно утверждать, что полученные в них результаты полностью теоретически обоснованы.

Таким образом, изложенный выше метод решения проблемы оптимального управления полумарковскими процессами с конечными множествами состояний может быть успешно реализован для многих задач, рассматриваемых в различных областях прикладной теории вероятностей.

Практическая реализация численной процедуры поиска оптимального решения на примере полумарковской модели управления запасом непрерывного продукта (подробнее см. [21, 29]), базирующаяся на изложенных выше результатах (в частности, теореме 1), была осуществлена А. К. Горшениным и соавторами в статье [32]. Коротко опишем наиболее важные аспекты этой работы.

Для решения поставленной задачи оптимального управления была создана специальная программа *Inventory* на встроенном языке программирования пакета *MATLAB*, ее возможности кратко представлены в упомянутой ранее статье [32]. В программе *Inventory* реализованы функции для оценивания через заданные исходные параметры вероятностных и стоимостных характеристик модели, которые в дальнейшем используются для поиска значений основной функции дробно-линейного функционала (15). Точка глобального экстремума этой функции и определяет оптимальное управление.

В качестве начальных данных необходимо задание следующих параметров:

- спрос и вместимость склада;
- разбиение множества значений объема запаса;
- вероятностные характеристики, описывающие модель пополнения запаса;
- условные математические ожидания длительностей задержек пополнения запаса;
- функции для характеристики затрат и доходов.

По итогам работы программы *Inventory* ряд вспомогательных функций представляется в аналитической форме (в частности, с использованием аппарата символьных вычислений *Symbolic Toolbox* пакета *MATLAB*), выводится точка глобального экстремума функции нескольких вещественных переменных (15), найденная с помощью применения численных и приближенно-аналитических аппроксимаций. Также формируются графики оценок значений вероятностно-стоимостных характеристик и основной функции дробно-линейного функционала (15), либо трехмерных сечений в случае наличия более трех параметров управления (переменных).

Функциональность пакета *Inventory* может быть расширена для практической реализации метода решения задачи поиска оптимального управления полумарковскими процессами с конечными множествами состояний, рассмотренного в данной статье.

Литература

1. Ховард Р. Динамическое программирование и марковские процессы / Пер. с англ. — М.: Сов. радио, 1964. 189 с. (*Howard R. A. Dynamic programming and Markov processes.* — Cambridge, MA, USA: MIT Press, 1960. 136 p.)
2. Рыков В. В. Управляемые марковские процессы с конечными пространствами состояний и управлений // Теория вероятностей и ее применения, 1966. Т. 11. Вып. 2. С. 343–351.
3. Джэвелл В. Управляемые полумарковские процессы // Кибернетич. сборник. — М.: Мир, 1967. Вып. 4. С. 97–134.
4. Fox B. Markov renewal programming by linear fractional programming // *SIAM J. Appl. Math.*, 1966. Vol. 14. P. 1418–1432.
5. Denardo E. V. Contraction mappings in the theory underlying dynamic programming // *SIAM Rev.*, 1967. Vol. 9. P. 165–177.
6. Howard R. A. Research in semi-Markovian decision structures // *J. Oper. Res. Soc. Japan*, 1963. Vol. 6. P. 163–199.

7. *Osaki S., Mine H.* Linear programming algorithms for Markovian decision processes // *J. Math. Anal. Appl.*, 1968. Vol. 22. P. 356–381.
8. *Майн Х., Осаки С.* Марковские процессы принятия решений / Пер. с англ. — М.: Наука, 1977. 176 с. (*Mine H., Osaki S.* Markovian decision processes. — New York, NY, USA: American Elsevier Publishing Co., 1970. 142 p.)
9. *Гихман И. И., Скороход А. В.* Управляемые случайные процессы. — Киев: Наукова думка, 1977. 251 с.
10. *Lique-Vasquez F., Hernandez-Lerma O.* Semi-Markov control models with average costs // *Appl. Math.*, 1999. Vol. 26. No. 3. P. 315–331.
11. *Vega-Amaya O., Lique-Vasquez F.* Sample-path average cost optimality for semi-Markov control processes on Borel spaces: Unbounded costs and mean holding times // *Appl. Math.*, 2000. Vol. 27. No. 3. P. 343–367.
12. Вопросы математической теории надежности / Под ред. Б. В. Гнеденко. — М.: Радио и связь, 1983. 376 с.
13. *Барзилович Е. Ю., Каштанов В. А.* Некоторые математические вопросы теории обслуживания сложных систем. — М.: Сов. радио, 1971. 272 с.
14. *Шнурков П. В.* О решении проблемы безусловного экстремума для дробно-линейного интегрального функционала на множестве вероятностных мер // Докл. РАН. Сер. Математика, 2016. Т. 470. № 4. С. 387–392.
15. *Ширяев А. Н.* Вероятность. — М.: МЦНМО, 2011. Кн. 1. 552 с. Кн. 2. 968 с.
16. *Боровков А. А.* Теория вероятностей. — М.: Либроком, 2009. 656 с.
17. *Хеннекен П. Л., Тортра А.* Теория вероятностей и некоторые ее приложения. — М.: Наука, 1974. 472 с.
18. *Халмош П.* Теория меры / Пер. с англ. — М.: ИЛ, 1953. 282 с. (*Halmos P.* Measure theory. — Litton Educational Publishing, Inc. 1950. 304 p.)
19. *Королюк В. С., Турбин А. Ф.* Полумарковские процессы и их приложения. — Киев: Наукова думка, 1976. 184 с.
20. *Janssen J., Manca R.* Applied semi-Markov processes. — New York, NY, USA: Springer, 2006. 309 p.
21. *Шнурков П. В., Иванов А. В.* Анализ дискретной полумарковской модели управления запасом непрерывного продукта при периодическом прекращении потребления // Дискретная математика, 2014. Т. 26. Вып. 1. С. 143–154.
22. *Иванов А. В.* Анализ дискретной полумарковской модели управления запасом непрерывного продукта при периодическом прекращении потребления. — М.: НИУ ВШЭ, 2014. Дисс. . . . канд. физ.-мат. наук. 120 с.
23. *Bajalinov E. B.* Linear-fractional programming. Theory, methods, applications and software. — Boston/Dordrecht/London: Kluwer Academic Publs., 2003. 423 p.
24. *Шнурков П. В., Мельников Р. В.* Оптимальное управление запасом непрерывного продукта в модели регенерации // Обзорные прикладной и промышленной математики, 2006. Т. 13. Вып. 3. С. 434–452.
25. *Шнурков П. В., Мельников Р. В.* Исследование проблемы управления запасом непрерывного продукта при детерминированной задержке поставки // Автоматика и телемеханика, 2008. Т. 10. С. 93–113.
26. *Шнурков П. В.* Методы исследования задач оптимального обслуживания в математической теории надежности. — М.: МИЭМ, 1983. Дисс. . . . канд. физ.-мат. наук.
27. *Кудрявцев Л. Д.* Курс математического анализа. Т. 1. — М.: Дрофа, 2006. 704 с.
28. *Шнурков П. В.* Оптимальное обслуживание на периоде до первого отказа системы // Применение аналитических методов в вероятностных задачах. — Киев: Институт математики АН УССР, 1986. С. 121–129.
29. *Шнурков П. В., Иванов А. В.* Исследование задачи оптимизации в дискретной полумарковской модели управления непрерывным запасом // Вестник МГТУ им. Н. Э. Баумана. Сер. Естественные науки, 2013. Т. 3. Вып. 50. С. 62–87.
30. *Shnourkoff P. V.* The two-element system with one restoring device optimum maintenance // *Stoch. Anal. Appl.*, 1997. Vol. 15. No. 5. P. 823–837.
31. *Shnourkoff P. V.* The two-element system optimum maintenance till the first fail // *Stoch. Anal. Appl.*, 2001. Vol. 19. No. 6. P. 1005–1024.
32. *Gorshenin A. K., Belousov V. V., Shnourkoff P. V., Ivanov A. V.* Numerical research of the optimal control problem in the semi-Markov inventory model // AIP Conference Proceedings, 2015. Vol. 1648. 250007. 4 p.

Поступила в редакцию 15.07.16

ANALYTICAL SOLUTION OF THE OPTIMAL CONTROL TASK OF A SEMI-MARKOV PROCESS WITH FINITE SET OF STATES

P. V. Shnurkov¹, A. K. Gorshenin², and V. V. Belousov²

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

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

Abstract: The theoretical verification of the new method of finding the optimal strategy of control of a semi-Markov process with finite set of states is presented. The paper considers Markov randomized strategies of control, determined by a finite collection of probability measures, corresponding to each state. The quality characteristic is the stationary cost index. This index is a linear-fractional integral functional, depending on collection of probability measures, giving the strategy of control. Explicit analytical forms of integrands of numerator and denominator of this linear-fractional integral functional are known. The basis of consequent results is the new generalized and strengthened form of the theorem about an extremum of a linear-fractional integral functional. It is proved that problems of existence of an optimal control strategy of a semi-Markov process and finding this strategy can be reduced to the task of numerical analysis of global extremum for the given function, depending on finite number of real arguments.

Keywords: optimal control of a semi-Markov process; stationary cost index of quality control; linear-fractional integral functional

DOI: 10.14357/19922264160408

Acknowledgments

The research was partially supported by the Russian Foundation for Basic Research (project 15-07-05316).

References

- Howard, R. A. 1960. *Dynamic programming and Markov processes*. Cambridge, MA: MIT Press. 136 p.
- Rykov, V. V. 1966. Upravlyaemye markovskie protsessy s konechnymi prostranstvami sostoyaniy i upravleniy [Controlled Markov processes with finite spaces of states and controls]. *Teoriya veroyatnostey i ee primeneniya* [Theory of Probability and Its Applications] 11(2):343–351.
- Jewell, W. S. 1963. Markov-renewal programming. *Oper. Res.* 11:938–971.
- Fox, B. 1966. Markov renewal programming by linear fractional programming. *SIAM J. Appl. Math.* 14:1418–1432.
- Denardo, E. V. 1967. Contraction mappings in the theory underlying dynamic programming. *SIAM Rev.* 9:165–177.
- Howard, R. A. 1963. Research in semi-Markovian decision structures. *J. Oper. Res. Soc. Japan* 6:163–199.
- Osaki, S., and H. Mine. 1968. Linear programming algorithms for Markovian decision processes. *J. Math. Anal. Appl.* 22:356–381.
- Mine, H., and S. Osaki. 1970. *Markovian decision processes*. New York, NY: Elsevier. 142 p.
- Gikhman, I. I., and A. V. Skorokhod. 1977. *Upravlyaemye sluchaynye protsessy* [Controlled random processes]. Kiev: Naukova Dumka. 251 p.
- Luque-Vasquez, F., and O. Herndandez-Lerma. 1999. Semi-Markov control models with average costs. *Appl. Math.* 26(3):315–331.
- Vega-Amaya, O., and F. Luque-Vasquez. 2000. Sample-path average cost optimality for semi-Markov control processes on Borel spaces: Unbounded costs and mean holding times. *Appl. Math.* 27(3):343–367.
- Gnedenko, B. V., ed. 1983. *Voprosy matematicheskoy teorii nadezhnosti* [Problems of the mathematical theory of reliability]. Moscow: Radio i svyaz'. 376 p.
- Barzilovich, E. Yu., and V. A. Kashtanov. 1971. *Nekotorye matematicheskie voprosy teorii obsluzhivaniya slozhnykh sistem* [Some mathematical questions in theory of complex systems maintenance]. Moscow: Sovetskoe radio. 272 p.
- Shnurkov, P. V. 2016. Solution of the unconditional extremum problem for a linear-fractional integral functional on a set of probability measures. *Dokl. Math.* 94(2):550–554.
- Shiryayev, A. N. 2016. *Probability-1*. Graduate texts in mathematics ser. New York, NY: Springer. Vol. 95. 503 p.; 2017. *Probability-2*. Vol. 900. 500 p.
- Borovkov, A. A. 2009. *Teoriya veroyatnostey* [Probability theory]. Moscow: Librokom. 656 p.
- Khenneken, P. L., and A. Tortra. 1974. *Teoriya veroyatnostey i nekotorye ee prilozheniya* [Probability theory and some of its applications]. Moscow: Nauka. 472 p.

18. Halmos, P. 1950. *Measure theory*. Litton Educational Publishing. 304 p.
19. Korolyuk, V. S., and A. F. Turbin. 1976. *Polumarkovskie protsessy i ikh prilozheniya* [Semi-Markov processes and their applications]. Kiev: Naukova Dumka. 184 p.
20. Janssen, J., and R. Manca. 2006. *Applied semi-Markov processes*. New York, NY: Springer. 309 p.
21. Shnurkov, P. V. and A. V. Ivanov. 2015. Analysis of a discrete semi-Markov model of continuous inventory control with periodic interruptions of consumption. *Discrete Math. Appl.* 25(1):59–67.
22. Ivanov, A. V. 2014. Analiz diskretnoy polumarkovskoy modeli upravleniya zapasom nepreryvnogo produkta pri periodicheskom prekrashchenii potrebleniya [Analysis of a discrete semi-Markov control model of continuous product inventory in a periodic cessation of consumption]. Moscow: Natsional'nyy Issledovatel'skiy Universitet "Vysshaya Shkola Ekonomiki." PhD Thesis. 120 p.
23. Bajalinov, E. B. 2003. *Linear-fractional programming. Theory, methods, applications and software*. Boston/Dordrecht/London: Kluwer Academic Publs. 423 p.
24. Shnurkov, P. V., and R. V. Mel'nikov. 2006. Optimal'noe upravlenie zapasom nepreryvnogo produkta v modeli regeneratsii [Optimal control of a continuous product inventory in the regeneration model]. *Obozrenie prikladnoy i promyshlennoy matematiki* [Rev. Appl. Ind. Math.] 13(3):434–452.
25. Shnurkov, P. V., and R. V. Mel'nikov. 2008. Analysis of the problem of continuous-product inventory control under deterministic lead time. *Automat. Rem. Contr.* 69(10):1734–1751.
26. Shnurkov, P. V. 1983. Metody issledovaniya zadach optimal'nogo obsluzhivaniya v matematicheskoy teorii nadezhnosti [Research methods of optimal service problems in the mathematical theory of reliability]. Moscow: Moskovskiy Institut Elektronnoy Mashinostroeniya. PhD Thesis.
27. Kudryavtsev, L. D. 2006. *Kurs matematicheskogo analiza* [A course of mathematical analysis]. Vol. 1. Moscow: Drofa. 704 p.
28. Shnurkov, P. V. 1986. Optimal'noe obsluzhivanie na periode do pervogo otkaza sistemy [The optimum service period until the first system failure]. *Primenenie analiticheskikh metodov v veroyatnostnykh zadachakh* [The application of analytical methods in probabilistic tasks]. Kiev: Institute of Mathematics of the Academy of Sciences of the USSR. 121–129.
29. Shnurkov, P. V., and A. V. Ivanov. 2013. Issledovanie zadachi optimizatsii v diskretnoy polumarkovskoy modeli upravleniya nepreryvnym zapasom [Study of the optimization problem in discrete semi-Markov model of continuous inventory control]. *Vestnik MGTU im. N. E. Baumana. Ser. Estestvennye nauki* [Vestnik of MSTU named after N. E. Bauman. Ser. Natural sciences] 3(50):62–87.
30. Shnurkov, P. V. 1997. The two-element system with one restoring device optimum maintenance. *Stoch. Anal. Appl.* 15(5):823–837.
31. Shnurkov, P. V. 2001. The two-element system optimum maintenance till the first fail. *Stoch. Anal. Appl.* 19(6):1005–1024.
32. Gorshenin, A. K., V. V. Belousov, P. V. Shnurkov, and A. V. Ivanov. 2015. Numerical research of the optimal control problem in the semi-Markov inventory model. *AIP Conference Proceedings* 1648:250007.

Received July 15, 2016

Contributors

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

Gorshenin Andrey K. (b. 1986) — Candidate of Science (PhD) in physics and mathematics, 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; associate professor, Federal State Budget Educational Institution of Higher Education "Moscow Technological University," 78 Vernadskogo Ave., Moscow 119454, Russian Federation; agorshenin@frccsc.ru

Belousov Vasily V. (b. 1977) — 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; VBelousov@ipiran.ru

ОБОБЩЕННЫЙ СТАТИСТИЧЕСКИЙ МЕТОД АНАЛИЗА ТЕКСТОВ, ОСНОВАННЫЙ НА РАСЧЕТЕ РАСПРЕДЕЛЕНИЙ ВЕРОЯТНОСТЕЙ ЗНАЧЕНИЙ СТАТИСТИК

А. К. Мельников¹, А. Ф. Ронжин²

Аннотация: Рассмотрены возможности применения точных и предельных распределений вероятностей значений статистик для построения критериев согласия в рамках анализа текстовой информации. Исследованы значения граничных параметров, для которых возможен расчет точных распределений на современном этапе. Разработан обобщенный статистический метод анализа (ОСМА) текстов, применимый при широком спектре значений параметров текстов.

Ключевые слова: вероятность; точное распределение; предельное распределение; статистика; критерий; частота; сложность алгоритма; производительность многопроцессорной вычислительной системы; метод анализа

DOI: 10.14357/19922264160409

1 Введение

Различные задачи автоматической обработки результатов наблюдений сводятся к обработке потоков данных, которые далее будем называть текстами.

Тексты представляют собой последовательности знаков длины n в некотором алфавите мощности N . В данных имеются тексты, которые являются реализациями случайных последовательностей с равномерным распределением на множестве всех своих значений, или, другими словами, случайными выборками длины n из равновероятного распределения на алфавите мощности N , и тексты с уникальными свойствами. Такие тексты, во-первых, не являются выбором из равномерного распределения; во-вторых, обладают структурными зависимостями внутри последовательности.

Алгоритмы определения структурных зависимостей сложны и требуют значительного вычислительного ресурса, поэтому предлагается двухэтапная процедура. На первом этапе выбираются самые неравновероятные тексты, а затем применяются алгоритмы поиска структурных зависимостей.

В ряде случаев, когда данные образуют непрерывный поток и имеются естественные ограничения на ресурс памяти, нет возможности накопить тексты, выбрать из них самые неравновероятные и обработать их по алгоритмам поиска структурных зависимостей. Тогда по критериям согласия с равновероятным распределением отсеивают тексты без уникальных свойств и к оставшимся тек-

стам применяют алгоритмы поиска структурных зависимостей. В этом случае размер критерия согласия определяет долю ложно принятых текстов без уникальных свойств или объем лишней работы по поиску текстов с уникальными свойствами. Для вычисления размера критерия согласия необходимо знать вероятностные распределения применяемых в критериях согласия статистик.

Целью данной работы является описание границ, для которых возможно вычисление точных значений вероятностных распределений на современных многопроцессорных вычислительных системах [1], и построение ОСМА текстов, применимого для всех значений их параметров.

2 Статистический анализ текстов, критерии согласия

Сосредоточимся на первом этапе анализа текстов.

Задача состоит в том, чтобы относительно ν -го текста $T_n(\nu)$ длины n

$$T_n(\nu) = \{t_1(\nu), \dots, t_n(\nu)\}$$

проверить простую гипотезу H_0 , состоящую в том, что все знаки текста (испытания) $t_i(\nu)$, $i = 1, \dots, n$, принимают значения из множества (алфавита) $A_N = \{a_1, \dots, a_N\}$ мощности N с равными вероятностями $P\{t_i(\nu) = a_j\} = 1/N$ для $i = 1, \dots, n$,

¹НТЦ ЗАО «ИнформИнвестГрупп», ak@iigroup.ru

²Институт точной механики и вычислительной техники им. С. А. Лебедева РАН, raf@zao-zt.ru

$j = 1, \dots, N$, т. е. наблюдается выборка из равномерного полиномиального распределения.

В качестве альтернативы выступает сложная гипотеза H , которая заключается в том, что текст $T_n(\nu)$ не является выборкой из равномерного полиномиального распределения в том смысле, что существует такое j , что $P\{t_i(\nu) = a_j\} \neq 1/N | i = 1, \dots, n$.

Будем отвергать гипотезу H_0 при $S_n(T_n(\nu)) \geq c$, где S_n — функция от текста $T_n(\nu)$ (статистика), а константа c выбирается исходя из заданного размера критерия α : $P\{S_n \geq c\} = \alpha$.

3 Расчет точных распределений значений статистик

Не ограничивая общности приводимых рассуждений, в качестве статистики S_n критерия будем рассматривать введенную более 100 лет назад Карлом Пирсоном [2, 3] статистику хи-квадрат χ_n :

$$\chi_n = \sum_{i=1}^N \frac{(h_i - np_i)^2}{np_i},$$

где h_i — частота встречаемости знака (исхода) a_i в первых n испытаниях; n — длина текста (объем выборки); N — число исходов полиномиальной схемы (мощность алфавита — A_N); p_i — вероятность a_i -го исхода.

Для расчета точного распределения вероятности $P\{S_n \geq c\}$ методом «полного перебора» необходимо проделать следующие шаги.

Шаг 1: сгенерировать множество $\bar{T}(n, A_N)$ всех неповторяющихся текстов длины n , состоящих только из знаков, принадлежащих алфавиту $A_N = \{a_1, \dots, a_N\}$ мощности N :

$$\bar{T}(n, A_N) = \{T_n(\nu) | \nu = 1, \dots, N^n\}.$$

Шаг 2: получить множество $H(n, N)$ массивов частот встречаемости $h(\nu)$ знаков алфавита A_N в текстах $T_n(\nu)$:

$$H(n, N) = \{h(\nu) = (h_1(\nu), \dots, h_N(\nu)) | \nu = 1, \dots, N^n\},$$

где $h_i(\nu)$ — частота встречаемости знака a_i алфавита A_N в ν -м тексте $T_n(\nu)$ массива текстов $H(n, N)$.

Шаг 3: получить множество значений статистик $\check{S}_n(N)$ на множестве массивов частот встречаемости знаков $H(n, N)$:

$$\check{S}_n(N) = \{S_n(h(\nu)) | \nu = 1, \dots, N^n\},$$

применяя формулу выбранной статистики S_n и учитывая равновероятность исходов полиномиальной схемы.

Шаг 4: получить распределение вероятностей значений статистики $S_n - P_T\{S_n \geq c\}$. Для этого надо провести маркировку множества значений статистик $\check{S}_n(N)$ в интервалы значений $[c, +\infty)$, где, например, $c = 1(1)100$.

Таблица 1 Максимальные значения параметров рассчитанных точных распределений статистик S_n

Объем выборки (длина текста) n	Число исходов полиномиальной схемы (мощность алфавита) N
50	2
30	3
27	4
22	5
19	6
16	7
15	8
14	9
11	10
8	16
6	26
4	32

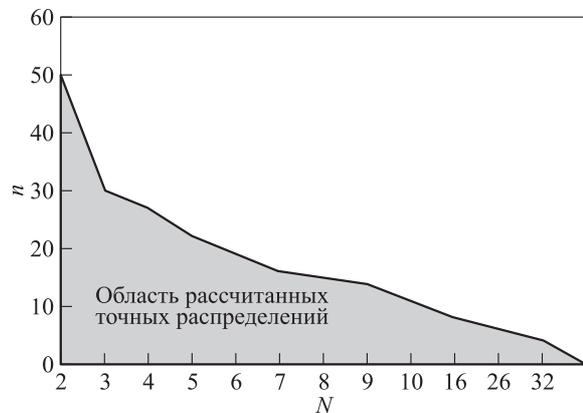


Рис. 1 Диаграмма параметров рассчитанных точных распределений статистик S_n

Методом «полного перебора» в XX в. были получены таблицы точного распределения статистики $S_n = \chi_n$ для n и N , максимальные значения которых приведены в табл. 1 и на рис. 1.

Необходимо указать, что, применяя некоторые свойства статистик (симметричность и пр.) и рекуррентные соотношения, математикам прошлых веков удалось рассчитать таблицы точных распределений, точность которых 10^{-5} , для следующих критических значений параметров: $N = 2(1)4$, $n =$

$= 26(1)50; N = 5, n = 26(1)40; N = 6(1)10, n = 26(1)30; N = 10, n = 1(1)30; N = 16, n = 1(1)30; N = 26, n = 1(1)30; N = 32, n = 1(1)30.$

Отметим некоторую универсальность алгоритма расчета точных распределений методом «полного перебора».

Алгоритм может применяться «последовательно» и «параллельно»:

- при «последовательном» применении алгоритма вначале генерируем все множество текстов $\bar{T}(n, A_N)$ (шаг 1), затем получаем все множество векторов частот встречаемости $H(n, N)$ (шаг 2), далее вычисляем все множество значений статистик $\check{S}_n(N)$ (шаг 3) и, обрабатывая все значения множества $\check{S}_n(N)$ (шаг 4), получаем точное распределение вероятностей значений статистики $S_n - P\{S_n \geq c\}$;
- при «параллельном» применении алгоритма на шаге 1 выполняем генерацию только одного текста $T_n(\nu)$ из множества $\bar{T}(n, A_N)$ всех неповторяющихся текстов, далее для сгенерированного текста вычисляем вектор встречаемости знаков $h(\nu)$ (шаг 2) и значение статистики $S_n(h(\nu))$ (шаг 3), полученное значение статистики заносим в массив интервалов значений $[c, +\infty)$ (шаг 4), после чего возвращаемся к генерации следующего неповторяющегося текста на шаге 1; алгоритм заканчивает свою работу тогда, когда на шаге 1 исчерпаны все неповторяющиеся тексты рассматриваемой длины.

«Последовательное» применение алгоритма требует достаточных ресурсов памяти для хранения множеств, «параллельное» применение не требует больших ресурсов памяти для хранения всех элементов множеств и позволяет распараллеливать выполнение алгоритма по данным, что дает возможность применять для его выполнения кластерные многопроцессорные вычислительные системы.

Предположим, что в течение 30 дней (1 месяц) доступен вычислительный ресурс производительностью 10^{15} операций в секунду, т. е. за отведенное время можно произвести для решения задачи расчета точных распределений $2,59 \cdot 10^{21}$ операций:

$$10^{15} \text{ оп./с} \cdot (30 \text{ дн.} \cdot 24 \text{ ч} \cdot 60 \text{ мин} \cdot 60 \text{ с}) = 2,59 \cdot 10^{21} \text{ оп.}$$

Предположим также, что для расчета одного значения статистики $S_n(h(\nu))$ и маркировки его значения в интервале $[c, +\infty)$ необходимо произвести $(5n + 3N + 50)$ операций, где $5n$ — число операций на маркировку текста длины n ; $3N$ — число

Таблица 2 Максимальные значения параметров, для которых на современном этапе могут быть рассчитаны точные распределения статистик

Объем выборки (длина текста) n	Число исходов полиномиальной схемы (мощность алфавита) N
58	2
29	4
19	8
17	10
12	26
11	36
9	64
8	128
7	256

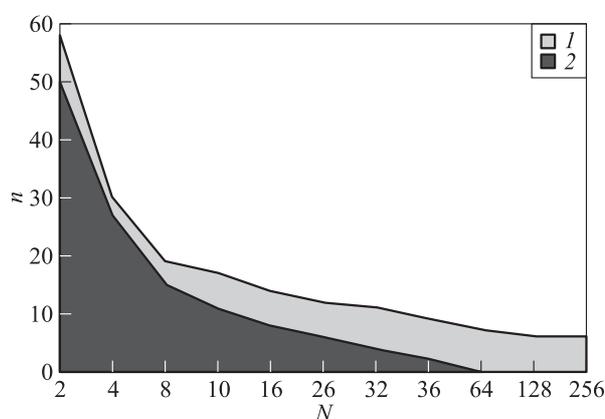


Рис. 2 Диаграмма параметров, для которых на современном этапе могут быть рассчитаны точные распределения статистик: 1 — область возможного расчета точных распределений на современном этапе; 2 — область рассчитанных точных распределений

операций на вычисление статистики S_n на равномерном распределении с N исходами; 50 — число вспомогательных операций. Тогда параметры n и N должны удовлетворять следующему соотношению: $N^n \leq 2,59 \cdot 10^{21} / (5n + 3N + 50)$.

Используя указанные выше предположения, были рассчитаны максимальные значения параметров, для которых на современном этапе за приемлемое время (1 мес.) могут быть рассчитаны точные распределения статистик. Значения параметров приведены в табл. 2 и на рис. 2.

Диаграмма на рис. 2 определяет область O_1 , в которой находятся параметры текстов (n, N) множеств $\bar{T}(n, A_N)$, для которых на современном этапе могут быть рассчитаны точные распределения статистик S_n . Линию, ограничивающую эту область сверху, обозначим через G_1 .

Определим множество $M_1 \bar{T}(n, A_N)$ как множество текстов, являющееся объединением непересе-

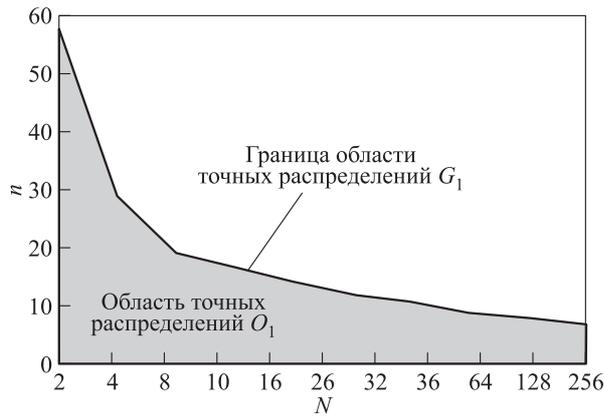


Рис. 3 Диаграмма параметров точных распределений статистик

кающихся подмножеств $\bar{T}(j, A_k)$ с параметрами j и k из области O_1 на рис. 2:

$$M_1 \bar{T}(n, A_N) = \{ \cup \bar{T}(j, A_k) \mid j = 1, \dots, n; k = 1, \dots, N \mid (n, N) \in O_1 \} .$$

Анализируя значения параметров на рис. 2, можно сделать вывод о том, что значительный рост производительности вычислительных средств не дает сколько-нибудь заметного увеличения значений параметров n и N в решении задачи расчета точных распределений и, соответственно, эффективного решения задач анализа текстов.

Назовем область O_1 областью точных распределений. Соответствующая диаграмма приведена на рис. 3.

4 Расчет Δ -точных распределений значений статистик

Для преодоления возникшего барьера и создания возможности анализа длинных текстов в алфавитах большой мощности был разработан метод расчета Δ -точных распределений $P_\Delta\{S_n \geq c\}$ — распределений, отличающихся от точных распределений не более чем на заранее заданную величину Δ .

Для $\Delta = 10^{-5}$ по методу расчета Δ -точных распределений были проведены расчеты $P_\Delta\{S_n \geq c\}$ для параметров n и N , значения которых приведены в табл. 3 и на рис. 4.

Диаграмма на рис. 4 определяет линию, ограничивающую область возможности расчета Δ -точных распределений сверху — G_2 . Область, заключенную между границами G_1 и G_2 , в которой находятся параметры текстов (n, N) множеств $\bar{T}(n, A_N)$,

Таблица 3 Значения параметров, для которых проведены расчеты Δ -точных распределений статистик S_n для $\Delta = 10^{-5}$

Объем выборки (длина текста) n	Число исходов полиномиальной схемы (мощность алфавита) N
150	2
80	4
70	8
60	10
50	26
50	32
45	128
40	256



Рис. 4 Значения параметров, для которых рассчитаны Δ -точные распределения S_n : 1 — область Δ -точных распределений O_2 ; 2 — область точных распределений O_1

для которых рассчитаны Δ -точные распределения статистик S_n , определим как O_2 .

По аналогии с $M_1 \bar{T}(n, A_N)$ определим множество $M_2 \bar{T}(n, A_N)$ как множество текстов, являющееся объединением непересекающихся подмножеств $\bar{T}(j, A_k)$ с параметрами j и k из области O_2 на рис. 4:

$$M_2(n, A_N) = \{ \cup \bar{T}(j, A_k) \mid j = 1, \dots, n; k = 1, \dots, N \mid (n, N) \in O_2 \} .$$

Сложность алгоритма расчета Δ -точных распределений статистик $P_\Delta\{S_n \geq c\} = C(P_\Delta\{S_n \geq c\})$ — определяется как число целочисленных решений системы уравнений:

$$\begin{aligned} \mu_0 + \mu_1 + \dots + \mu_k &= N; \\ 1\mu_1 + 2\mu_2 + \dots + k\mu_k &= n, \end{aligned}$$

в которой μ_ν — число значений частот исходов h_i , равных ν .

Вектор (μ_0, \dots, μ_N) получил название вектора вторых маркировок.

Значение k подбирается из условия $P\{M_n > k\} \leq \Delta$, в котором M_n — статистика максимальной частоты: $M_n = \max_{i=1}^N h_i$.

Вероятность $P\{M_n > k\}$ вычисляется с помощью рекуррентной формулы, предложенной Б. И. Селивановым:

$$P\{M_{n+1} < m\} = \sum_{\nu=0}^n \binom{n}{\nu} P\{M_{n-\nu} < m\} d_{\nu+1}^{(m)} \frac{1}{N^\nu}$$

с начальным условием $P\{M_0 < m\} = 1$, коэффициенты $d_{\nu+1}^{(m)}$ вычисляются по рекуррентной формуле

$$d_{h+1}^{(m)} = - \sum_{\nu=1}^{m-1} \binom{n}{\nu} d_{n-\nu+1}^{(m)}$$

с начальными условиями $d_1^{(m)} = 1, d_2^{(m)} = \dots = d_{n-1}^{(m)} = 0, d_m^{(m)} = -1, d_{m+1}^{(m)} = m$.

Сложность алгоритма расчета Δ -точных распределений статистик $C(P_\Delta\{S_n \geq c\})$ может быть оценена сверху как

$$C(P_\Delta\{S_n \geq c\}) \leq \binom{|n - N| + k}{|n - N|},$$

где $|x|$ обозначает модуль числа x .

5 Пределыные распределения значений статистик

Одним из замечательных свойств статистики χ_n является то, что при $m = \min_{i=1}^N np_i \rightarrow \infty$ ее предельное распределение не зависит от p_1, \dots, p_N и совпадает с χ^2 -распределением с $(N - 1)$ степенью свободы [4].

Вопрос о том, начиная с какого m можно пользоваться предельным распределением, задавали себе многие авторы, и не всегда их мнения совпадали. Фишер [5] рекомендует ограничение $m \geq 5$, Крамер [4] — $m \geq 30$, Кендалл и Стюарт [6] — $m \geq 20$.

Следует напомнить, что теоретическая оценка остаточного члена в предельной теореме, приведенная в [7], имеет вид $O(n^{-(N-1)/N})$ и мало что дает для практического использования. Поэтому расчет непределных распределений значений статистик для анализа текстов еще долго будет оставаться актуальной проблемой для специалистов в области математической статистики, вычислительной математики и программирования.

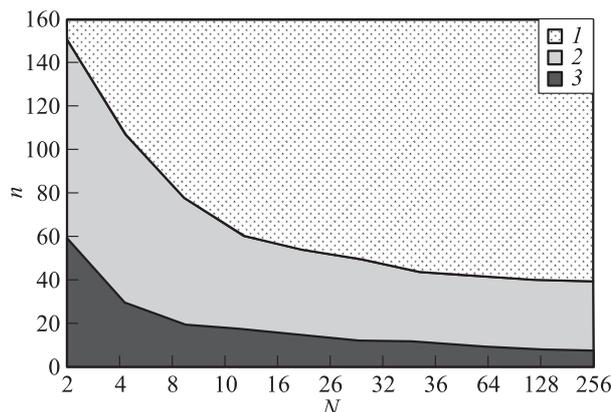


Рис. 5 Область предельных распределений, ограниченная снизу G_2 : 1 — область предельных распределений; 2 — область Δ -точных распределений O_2 ; 3 — область точных распределений O_1

Областью предельных распределений назовем область параметров (n, N) , для которой при современной производительности вычислительных средств невозможно рассчитать точные и Δ -точные распределения. Эта область ограничена снизу G_2 . Отметим, что граница G_2 рассчитанных Δ -точных распределений на рис. 4 совпадает с границей G_2 области предельных распределений на рис. 5.

6 Обобщенный статистический метод анализа текстов

После проведенного анализа возможностей расчета и применения точных, Δ -точных и предельных распределений вероятностей значений статистики S_n для построения статистического критерия согласия с равновероятным распределением, используемого для отбора текстов без уникальных свойств, построим обобщенный метод анализа текстов без ограничений на его параметры, такие как длина и мощность алфавита входящих в него знаков.

Пусть необходимо проанализировать текст $T_i(n, A_N)$ длины n из алфавита A_N мощности N и определить, обладает ли он уникальными свойствами. Тогда его стоит дальше исследовать с помощью трудоемких алгоритмов выявления структурных зависимостей. Либо данный текст не обладает уникальными свойствами и не представляет дальнейшего интереса.

Обобщенный статистический метод анализа текстов $T_i(n, A_N)$ с выводом ОСМА $\{T_i(n, A_N)\}$ будем строить следующим образом:

- (1) в зависимости от значений параметров (n, N) выберем распределение $P\{S_n \geq c\}$:

$$P\{S_n \geq c\} = \begin{cases} P_T\{S_n \geq c\}, & \text{если } (n, N) \in O_1, \\ & \text{ограниченной } G_1 \text{ сверху;} \\ P_\Delta\{S_n \geq c\}, & \text{если } (n, N) \in O_2, \\ & \text{ограниченной } G_1 \text{ снизу} \\ & \text{и } G_2 \text{ сверху;} \\ P\{\chi^2 \geq c\} \text{ с } (N-1) \text{ степенью свободы} & \text{в противном случае;} \end{cases}$$

(2) в соответствии с выбранным распределением $P\{S_n \geq c\}$ и задаваемым размером критерия α получаем разделяющую константу c : $P\{S_n \geq c\} = \alpha$;

(3) рассчитываем значение статистики $S_n(T_i(n, A_N))$ от анализируемого текста $T_i(n, A_N)$;

(4) сравниваем значение статистики $S_n(T_i(n, A_N))$ с разделяющей константой c и делаем вывод (ОСМА $\{T_i(n, A_N)\}$) о наличии или отсутствии у текста $T_i(n, A_N)$ уникальных свойств:

- если $S_n(T_i(n, A_N)) \geq c$, то ОСМА $\{T_i(n, A_N)\}$ = текст $T_i(n, A_N)$ обладает уникальными свойствами;
- если $S_n(T_i(n, A_N)) < c$, то ОСМА $\{T_i(n, A_N)\}$ = текст $T_i(n, A_N)$ не обладает уникальными свойствами.

Построение обобщенного метода анализа текстов окончено.

Необходимо отметить, что значения G_1 и G_2 по построению определяются алгоритмами расчета точных и Δ -точных распределений и производительностью имеющих на момент времени τ вычислительных средств $p(\tau)$, а следовательно, могут быть определены как функции от τ : $G_1(\tau)$ и $G_2(\tau)$.

Придерживаясь закона Мура, считаем функцию $p(\tau)$ возрастающей, т.е. при $\tau_1 < \tau_2$ $p(\tau_1) \leq p(\tau_2)$. Тогда функции $G_1(\tau)$ и $G_2(\tau)$ должны обладать следующими свойствами:

- $G_1(\tau) < G_2(\tau)$ по построению;
- если $\tau_1 < \tau_2$, то $G_1(\tau_1) \leq G_1(\tau_2)$;
- если $\tau_1 < \tau_2$, то $G_2(\tau_1) \leq G_2(\tau_2)$.

Увеличение значений $G_2(\tau)$ при росте $p(\tau)$ позволяет при использовании ОСМА для большего числа значений параметров (n, N) применять Δ -точные распределения вместо предельных, что в несколько раз уменьшает число ложно отобранных как обладающих уникальными свойствами текстов (ошибка первого рода).

Увеличение значений $G_1(\tau)$ при росте $p(\tau)$ позволяет при применении ОСМА для большего числа значений параметров (n, N) применить точные распределения вместо Δ -точных, что упрощает процедуру проведения анализа текста и уменьшает его трудоемкость.

Таким образом, предложенный обобщенный метод анализа текстов ОСМА при росте производительности $p(\tau)$ вычислительных средств позволяет увеличить точность анализа и снизить его трудоемкость.

Литература

1. Каляев В. А., Левин И. И., Семерников Е. А., Шмойлов В. И. Реконфигурируемые мультимедийные вычислительные структуры. — Ростов-на-Дону: ЮНЦ РАН, 2008. 397 с.
2. Pearson K. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling // Philos. Mag. Ser. 5, 1900. Vol. 50. No. 302. P. 157–175.
3. Smith P. F., Rae D. S., Manderscheid R. W., Silbergeld S. Exact and approximate distributions of the chi-squared statistic for equiprobability // Commun. Stat., 1979. Vol. 8. No. 2. P. 131–149.
4. Крамер Г. Математические методы статистики / Пер. с англ. — М.: Мир, 1975. 648 с. (Cramer G. Mathematical methods of statistics. — Princeton, NJ, USA: Princeton University Press, 1946. 592 p.)
5. Фишер Р. А. Статистические методы для исследователей / Пер. с англ. — М.: Госстатиздат, 1958. 73 с. (Fisher R. A. Statistical methods for research workers. — 12th ed. — Edinburgh: Oliver and Boyd, 1954. 356 p.)
6. Кендалл М. Г., Стьюарт А. Теория распределений. — М.: Наука, 1966. 302 с.
7. Hutchinson T. P. 1979. The validity of the chi-squared test when expected frequencies are small: A list of recent research references // Commun. Stat. A Theor. Vol. 8. No. 4. P. 327–335.

Поступила в редакцию 02.02.16

GENERALIZED STATISTICAL METHOD OF TEXT ANALYSIS BASED ON CALCULATION OF PROBABILITY DISTRIBUTIONS OF STATISTICAL VALUES

A. K. Melnikov¹ and A. F. Ronzhin²

¹STC CLSC “InformInvestGroup,” 125, Bld. 17 Varshavskoye Shosse, Moscow 117587, Russian Federation

²S. A. Lebedev Institute of Precision Mechanics and Computer Engineering of the Russian Academy of Sciences, 51 Leninsky Prosp., Moscow 119991, Russian Federation

Abstract: A lot of data streams are a mixture of random and unique data. One of the properties of unique data is the nonuniform distribution of probability of encountering the data on the set of the values. The procedure of two steps is implemented for distinguishing unique data. On the first step of candidate selection, the criterion of consensus with the uniform distribution is implemented. On the second step, resource-intensive calculation in a condition of indeterminacy is performed in order to check other unique attributes of the candidates. The choice of the size of the criterion depends on the amount of resources given for the second step. The accuracy of calculation determines the quantity of overhead of the second term for processing random data and, therefore, a part of unique data loss. The paper analyzes the values of boundary parameters for which at the current level of computer technology, one can calculate the exact distribution. A generalized statistical method of text analysis, which can be used for a wide spectrum of text parameters, is developed.

Keywords: probability; exact distribution; limit distribution; statistics; criterion; frequency; algorithm complexity; performance of multiprocessor computer system; analysis method

DOI: 10.14357/19922264160409

References

1. Kalyaev, I. A., I. I. Levin, E. A. Semernikov, and V. I. Shmoylov. 2008. *Rekonfiguriruyemye mul'tikonveyernye vychislitel'nye struktury* [Reconfigurable multiconference computational patterns]. Rostov-on-Don: YuNTs RAN. 397 p.
2. Pearson, K. 1900. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philos. Mag. Ser. 5* 50(302):157–175.
3. Smith, P. F., D. S. Rae, R. W. Manderscheid, and S. Silbergeld. 1979. Exact and approximate distributions of the chi-squared statistic for equiprobability. *Commun. Stat.* 8(2):131–149.
4. Cramer, G. 1946. *Mathematical methods of statistics*. Princeton, NJ: Princeton University Press. 592 p.
5. Fisher, R. A. 1954. *Statistical methods for research workers*. 12th ed. Edinburgh: Oliver and Boyd. 356 p.
6. Kendall, M. G., and A. Stuart. 1967. *The advanced theory of statistics. Vol. 1: Distribution theory*. 3rd ed. London: Charles Griffin Co. 439 p.
7. Hutchinson, T. P. 1979. The validity of the chi-squared test when expected frequencies are small: A list of recent research references. *Commun. Stat. A Theor.* 8(4):327–335.

Received February 2, 2016

Contributors

Melnikov Andrey K. (b. 1956) — Candidate of Science (PhD) in technology, associate professor, principal scientist, STC CLSC “InformInvestGroup,” 125, Bld. 17 Varshavskoye Shosse, Moscow 117587, Russian Federation; ak@iigroup.ru

Ronzhin Alexander F. (b. 1954) — Doctor of Science in physics and mathematics, professor, S. A. Lebedev Institute of Precision Mechanics and Computer Engineering of the Russian Academy of Sciences, 51 Leninsky Prosp., Moscow 119991, Russian Federation; raf@zao-zt.ru

ОБ ОДНОМ СПОСОБЕ СОКРАЩЕНИЯ ВЫЧИСЛЕНИЙ ПРИ ФОРМИРОВАНИИ ЗАМКНИЙ ГАЛУА*

А. А. Грушо¹, М. И. Забейло², А. А. Зацаринный³

Аннотация: Уточнение понятия сходимости как алгебраической операции — процедурная основа многих современных методов интеллектуального анализа данных (ИАД). Однако в ряде важных приложений (в частности, при управлении информационными потоками в компьютерных сетях, обеспечении информационной безопасности (ИБ) в облачных средах и др.) применение подобной математической техники оказывается ограниченным объемами вычислений при обработке больших объемов данных в режиме реального времени. Пример — многократные вычисления пересечений множеств булевских строк большой длины в процессе поиска неподвижных точек замыканий Галуа. В статье предложен алгоритм формирования замыканий Галуа, позволяющий ускорить проверку замкнутости множеств. Приведены примеры использования предлагаемой техники в ряде задач управления потоками сообщений в компьютерных сетях и контроля содержимого пересылаемых пакетов.

Ключевые слова: интеллектуальный анализ данных; сходимость как алгебраическая операция; методы сокращения перебора; скорость управления и безопасность информационных потоков в компьютерных сетях

DOI: 10.14357/19922264160410

1 Общее представление о рассматриваемой проблеме

Искусственный интеллект как область исследований сегодня наиболее ярким образом представлен разнообразием математических моделей, методов и компьютерных систем так называемого *интеллектуального анализа данных*. В компьютерном анализе данных как целом ИАД можно выделить «инструментально» как область анализа данных, выполняемого средствами *интеллектуальных компьютерных систем*. Основная отличительная особенность таких систем — использование *формальных моделей рассуждений*, в том числе корректных *формализаций эвристик* применяемых специализирующимися на анализе данных экспертами (например, выделение областей математически корректного использования соответствующих эвристик и контроль «невыхода» за границы таких областей в процессе ИАД, выполняемого реализующими эти эвристики «инструментальными» средствами).

Характерным для ИАД стало оценивание целевых параметров по косвенным признакам (см.,

например, [1–4]). При этом фактически постулируется принцип: *сходимость в описаниях прецедентов влечет сходимость в описаниях анализируемых* (по «косвенным» данным) *целевых свойств* изучаемых прецедентов. Так в центре внимания оказались формализация и анализ сходимости как математически корректной конструкции. В рамках ИАД предложены и активно развиваются различные варианты уточнения понятия сходимости — метрики, меры близости, формализация сходимости как алгебраической операции и др.

При вовлечении в ИАД данных комплексной «природы» (в том числе в работе с нечисловыми объектами, свойства которых определяются как особенностями их внутренней структуры, так и принимающими числовые значения⁴ параметрами) достаточно эффективным оказывается использование уточнения сходимости как *алгебраической операции*. Ключевым элементом задействованной здесь формальной математической конструкции оказывается поиск *всех тех* и *только тех объектов*, которые удовлетворяют заданному набору условий, т. е. использование специального *оператора замыкания* и поиска неподвижных точек некоторой спе-

*Работа поддержана РФФИ (проект 15-29-07981).

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

²Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, m.zabehailo@yandex.ru

³Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, alex250451@mail.ru

⁴Например, это могут быть физиологически активные химические соединения, целевые (в частности — *лекарственные*) свойства которых определяются как их структурными особенностями, так и целым рядом дополнительных параметров, описываемых определенными числовыми значениями (в том числе дозами, концентрациями, температурами и т. п.).

циальной версии замыкания Галуа [5]. Своего рода «платой» за богатые выразительные возможности применяемой математической техники оказывается необходимость оперировать достаточно сложными комбинаторными объектами — порождаемыми в процессе вычисления сходств диаграммами взаимной вложимости классов сходства и классов эквивалентности. Именно с этим связан пристальный интерес специалистов к любым возможностям «ускорения» вычислений (за счет тех или иных сокращений их объемов) при многократно выполняемом в процессе ИАД поиске сходств.

Процедурные конструкции рассматриваемого типа демонстрируют свою эффективность и в задачах обеспечения ИБ. Так, например, в ряде технологий DPI (Deep Packet Inspection), IDS/IPS (Intrusion Detection/Protection Systems) и др. критически важным компонентом оказывается выделение эмпирических (извлекаемых в процессе ИАД из описаний инцидентов) зависимостей, которые неявным образом представлены в накапливаемых данных. Цель порождения таких зависимостей — «обучение» НОРМЕ и идентификация АНОМАЛИЙ в поведении объекта защиты. Выделяемые в процессе обучения комбинации признаков дают возможность оперативно опознавать «аналогичные» явления (АНОМАЛИИ) во вновь анализируемых данных. При этом скорость анализа данных в задачах обеспечения ИБ (например, облачных сред или же обработки так называемых Big Data) оказывается критически важным фактором успешной работы соответствующих средств защиты. В практически значимых приложениях в первую очередь именно скоростью ведения ИАД определяются возможности обрабатывать в режиме реального времени огромные коммутационные таблицы в устройствах управления трафиком в компьютерных сетях, все более объемные реестры сигнатур вирусов и т. п. Представленная «интерполяционно-экстраполяционная» (порождения и «переноса» эмпирических зависимостей на описания новых ИБ-инцидентов) техника «диагностики» АНОМАЛИЙ фактически и есть один из механизмов ИАД, привнесенных в проблематику ИБ исследованиями в области искусственного интеллекта (в том числе разработки и экспериментального исследования интеллектуальных компьютерных систем).

2 Базовые понятия и определения

Пусть заданы два множества: $U = \{a_1, a_2, \dots, a_n\}$ и $\Omega = \{O_1, O_2, \dots, O_m\} \subseteq 2^U \setminus \emptyset$, первое из которых будем называть исходным алфавитом (множеством образующих элементов для анализируемых прецедентов из Ω), а второе — мно-

жеством описаний прецедентов (непустых множеств образующих), построенных над универсумом U , т. е. над множеством объектов, построенных из образующих a_1, a_2, \dots, a_n .

Располагая множествами U и Ω , определим два отображения f и φ :

- $\forall \xi \in 2^U$ (т. е. для каждого ξ — подмножества образующих из множества U) $f(\xi) = \{\text{множество всех таких } O_i \text{ из } \Omega, \text{ что } \forall a_j \in \xi \text{ имеет место } a_j \in O_i \text{ для каждого из этих выбранных } O_i \text{ (т. е. это множество всех прецедентов из } \Omega, \text{ в которые все } a_j \text{ из заданного } \xi \text{ входят одновременно)}\}$;
- $\forall \zeta \in 2^\Omega$ (т. е. для каждого ζ — подмножества прецедентов из множества Ω) $\varphi(\zeta) = \{\text{множество всех таких } a_j \text{ из } U, \text{ что } \forall O_i \in \zeta \text{ имеет место } a_j \in O_i \text{ для каждого из этих выбранных } a_j \text{ (т. е. это множество всех таких } a_j \text{ из } U, \text{ которые во все прецеденты из } \zeta \text{ входят одновременно)}\}$,

где 2^U — множество всех подмножеств множества U (исходного алфавита), а 2^Ω — множество всех подмножеств множества Ω (множества описаний прецедентов).

Можно показать, что пара отображений $\langle f, \varphi \rangle$ представляет собой соответствие Галуа (см., например, [5]), а их произведения $f(\varphi(\zeta))$ и $\varphi(f(\xi))$ — соответствующие замыкания Галуа (см., например, [5]), которые будем обозначать как $[-]_{U,\Omega}$ и $[-]_{\Omega,U}$.

Определение 1. Неподвижными точками замыкания Галуа $[-]_{U,\Omega}$ будем называть все такие $[X]_{U,\Omega}$, что $[X]_{U,\Omega} = X$.

Посредством $\mathbf{GC}_{f,\varphi}(\Omega)$ и $\mathbf{GC}_{\varphi,f}(U)$ будем обозначать множества неподвижных точек соответствующих замыканий Галуа: $\mathbf{GC}_{f,\varphi}(\Omega) = \{\zeta \in 2^\Omega \text{ таких, что } f(\varphi(\zeta)) = \zeta\}$ и $\mathbf{GC}_{\varphi,f}(U) = \{\xi \in 2^U \text{ таких, что } \varphi(f(\xi)) = \xi\}$. Каждое из этих множеств можно рассматривать как частично упорядоченное в соответствии со взаимным вложением как соответствующих множеств прецедентов (подмножеств для Ω), так и подмножеств образующих из U .

Таким образом, для заданного множества образующих $u = \{u_1, u_2, \dots, u_s\}$, где $u \subseteq U$, проверку его замкнутости (относительно заданного множества прецедентов Ω , представленных в алфавите U) может выполнить следующий

Алгоритм 1

Входные данные: множества U , Ω и u .

ШАГ 1. Для каждого i от 1 до s ($s \leq n$) в множестве Ω описаний прецедентов выделяется подмножество Ω_i таких, что образующая u_i входит в описание каждого из прецедентов O_j , входящих в соответствующее подмножество Ω_i .

ШАГ 2. Вычисляется \mathbf{u}^* — общая часть (пересечение соответствующих множеств образующих) всех Ω_i (по i от 1 до $s \leq n$).

ШАГ 3. Проверяется выполнимость равенства $\mathbf{u}^* = \mathbf{u}$. При положительном исходе \mathbf{u} замкнуто, в противном случае — нет.

ШАГ 4. СТОП.

К сожалению, для больших значений параметра m (числа описаний прецедентов в множестве Ω) при «массовой» проверке замкнутости большого числа «тестируемых» множеств \mathbf{u} (важный для приложений пример такой ситуации дает характерная для управления потоками данных в компьютерных сетях задача поиска заголовка пакета в таблице коммутации, где необходимо в режиме реального времени оперировать таблицами коммутации, содержащими десятки, а в ряде случаев и сотни тысяч строк) алгоритм 1 оказывается весьма ресурсоемким (в частности, шаг 2 требует массовых операций с булевскими последовательностями длины m). Этим обусловлен интерес к построению тех или иных сравнительно (с «прямой» проверкой) быстрых вычислимых локальных условий (где, например, было бы достаточно выявления совместного вхождения пары образующих u_i, u_j из алфавита \mathbf{U} , «непопадание» одной из которых в порожаемое замыкание \mathbf{u}^* сигнализировало бы о незамкнутости текущего тестируемого множества образующих \mathbf{u}). Один из вариантов подобного достаточного условия незамкнутости дает

Утверждение 1. Для того чтобы заданное подмножество $\mathbf{u} = \{u_1, u_2, \dots, u_s\}$ образующих из алфавита \mathbf{U} было незамкнуто относительно множества описаний прецедентов Ω , достаточно, чтобы для замыкающий одноэлементных подмножеств \mathbf{u} выполнялось строгое включение:

$$\mathbf{u} = \{u_1, u_2, \dots, u_s\} \subset \bigcup_{i=1}^s [\{u_i\}]_{\mathbf{U}, \Omega}.$$

К сожалению, представленное утверждением 1 условие незамкнутости не является необходимым, в чем нетрудно убедиться, принимая во внимание

Пример 1. Рассмотрим множество булевских векторов длины 4:

$$\mathbf{B} = \{\langle 0000 \rangle, \langle 0001 \rangle, \dots, \langle 1110 \rangle, \langle 1111 \rangle\}.$$

Положим $\mathbf{U} = \{a_1^0, a_1^1, a_2^0, a_2^1, a_3^0, a_3^1, a_4^0, a_4^1\}$, а множество описаний прецедентов $\Omega = \{O_1, O_2, \dots, O_{15}\}$ сформируем из первых 15 элементов множества \mathbf{B} по следующему правилу: в каждое множество $O_j = \{a_{j1}, a_{j2}, a_{j3}, a_{j4}\}$ в качестве элемента a_{jr} берется a_r^0 , если в i -м элементе множества \mathbf{B} на позиции номер r стоит 0, в противном

случае (когда в этой позиции в i -м элементе множества \mathbf{B} находится единица) на соответствующую позицию в O_j берем образующую a_r^1 . Несложно убедиться, что для последнего, 16-го элемента $\langle 1111 \rangle$ множества \mathbf{B} соответствующее множество $O_{16} = \{a_1^1, a_2^1, a_3^1, a_4^1\}$, не входящее в Ω , тем не менее, удовлетворяет условию:

$$O_{16} = \{a_1^1, a_2^1, a_3^1, a_4^1\} = \bigcup_{i=1}^4 [\{a_i^1\}]_{\mathbf{U}, \Omega}.$$

3 Процедура формирования замыкания и ее свойства

Пусть заданы: множество примеров (описаний прецедентов) $\Omega = \{O_1, O_2, \dots, O_m\}$, построенных на алфавите $\mathbf{U} = \{a_1, a_2, \dots, a_n\}$, и некоторое непустое подмножество $\mathbf{u} = \{a(u)_1, a(u)_2, \dots, a(u)_k\}$ исходного \mathbf{U} . Пусть α, α_1 и α_2 — перестановки индексов $\{1, 2, \dots, s\}$, а $\mathbf{u}_\alpha = \langle u_{i1}, u_{i2}, \dots, u_{is} \rangle$ — упорядочение входящих в множество \mathbf{U} образующих в соответствии с их перечислением в α .

Алгоритм 2

Входные данные: множества \mathbf{U}, Ω и \mathbf{u} , а также перестановка α , определяющая порядок рассмотрения образующих в тестовом множестве \mathbf{u} .

ШАГ 1. Выбираем в текущем \mathbf{u} в соответствии с α первый элемент — u_1 .

ШАГ 2. Строим по u_1, Ω и \mathbf{U} множество $[\{u_1\}]_{\mathbf{U}, \Omega}$.

ШАГ 3. Выделяем в Ω подмножество $\Omega(u_1)$ всех O_{u_1} , содержащих $[\{u_1\}]_{\mathbf{U}, \Omega}$.

ШАГ 4. Удаляем из каждого входящего в $\Omega(u_1)$ примера образующие из $[\{u_1\}]_{\mathbf{U}, \Omega}$ и формируем (этими удалениями) новое — модифицированное — множество $\Omega^*(u_1)$. Получаем «усеченное» множество примеров $\Omega^*(u_1)$ и использованный при его формировании «усеченный» алфавит $\mathbf{U}(u_1)$.

ШАГ 5. Удаляем из текущего \mathbf{u} все образующие, вошедшие в $[\{u_1\}]_{\mathbf{U}, \Omega}$.

ШАГ 6. Если в текущем \mathbf{u} уже нет (неудаленных) образующих, переходим на шаг 7. Иначе заменяем текущее Ω на вновь построенное множество $\Omega^*(u_1)$, а текущее \mathbf{U} — на «усеченный» алфавит, сформированный на шаге 4, и, наконец, заменяем имевшееся на шаге 1 множество \mathbf{u} — на его текущую (см. шаг 5) «усеченную» версию, после чего переходим на шаг 1.

ШАГ 7. СТОП.

Будем обозначать последовательность u_1, u_2, \dots, u_l (где $l \leq k$) формирующих \mathbf{u} элементов исходного \mathbf{U} , удаляемых на каждом из шагов принимаемой рекурсии, как $T_\alpha(\mathbf{u}, \Omega, \mathbf{U})$ и называть ее задаваемой упорядочением α траекторией «усечения» проверяемого на замкнутость множества \mathbf{u} .

Сократить объем вычислений алгоритма 2 можно, выбирая при порождении «сужений» Ω^* из имеющихся вариантов тот, где минимизируется размер Ω^* . Для этого достаточно рассмотреть на каждом шаге «сужений» замыкания всех одноэлементных подмножеств текущего алфавита. Максимальным по размеру элементам такого множества замыканий соответствуют минимальные по размеру подмножества исходного множества Ω .

Алгоритм 3

Входные данные: множества \mathbf{U} , Ω и \mathbf{u} .

ШАГ 1. Строим по Ω и \mathbf{U} множество замыканий всех одноэлементных подмножеств алфавита \mathbf{U} : $\mathbf{B_GC}_{s^*g}(\mathbf{U}, \Omega) = \{ \{a_1\}, \{a_2\}, \dots, \{a_n\} \}$.

ШАГ 2. Выбираем в \mathbf{u} образующую (обозначим ее u_1), которой соответствует максимальный элемент в $\mathbf{B_GC}_{s^*g}(\mathbf{U}, \Omega)$.

ШАГ 3. Выделяем в Ω подмножество $\Omega(u_1)$ всех O_{u_1} , содержащих $[u_1]_{\mathbf{U}, \Omega}$.

ШАГ 4. Удаляем из каждого входящего в $\Omega(u_1)$ примера образующие из $[u_1]_{\mathbf{U}, \Omega}$, формируя модифицированное множество примеров $\Omega^*(u_1)$. Результат — «усеченное» множество примеров $\Omega^*(u_1)$ и «усеченный» алфавит $\mathbf{U}(u_1)$.

ШАГ 5. Удаляем из текущего \mathbf{u} все образующие, вошедшие в $[u_1]_{\mathbf{U}, \Omega}$.

ШАГ 6. Если в текущей версии множества \mathbf{u} больше не осталось еще не удаленных образующих, переходим на шаг 7. В противном случае заменяем текущее Ω на вновь построенное множество $\Omega^*(u_1)$, а текущее множество \mathbf{U} — на «усеченный» алфавит, сформированный на шаге 4, после чего переходим на шаг 1 (т.е. возвращаемся к построению $\mathbf{B_GC}_{s^*g}$, но теперь уже на «усеченных» $\Omega^*(u')$ и $\mathbf{U}(u')$, т.е. переходим к построению множества замыканий всех одноэлементных подмножеств «усеченного» алфавита $\mathbf{U}(u')$ относительно «усеченного» множества примеров $\Omega^*(u')$, где текущее значение u' получено удалением из текущего \mathbf{u} всех образующих из текущего $[u_1]_{\mathbf{U}, \Omega}$).

ШАГ 7. СТОП.

Возможности представления средствами алгоритмов 2 и 3 «составных» (многоэлементных) неподвижных точек рассматриваемого замыкания Галуа

с помощью замыканий одноэлементных подмножеств алфавита \mathbf{U} демонстрирует

Утверждение 2 (Теорема о представлении). Множество образующих \mathbf{u} замкнуто относительно множества примеров Ω в алфавите \mathbf{U} тогда и только тогда, когда объединение всех замыканий $[u']_{\mathbf{U}^*, \Omega^*}$ вдоль сформированной траектории $T(\mathbf{u}, \Omega, \mathbf{U})$ совпадает с \mathbf{u} :

$$[\mathbf{u}]_{\mathbf{U}, \Omega} = \mathbf{u}$$

тогда и только тогда, когда

$$\mathbf{Un}(T) = \left(\bigcup_{u'_i \in T(\mathbf{u}, \Omega, \mathbf{U})} [u'_i] \right) = \mathbf{u}.$$

Доказательство. Из-за ограниченности объема данной публикации приведем лишь ключевые элементы используемой при доказательстве утверждения 2 системы аргументов.

- (i) Можно показать, что (вследствие ассоциативности операции пересечения множеств) для любых упорядочений α_1 и α_2 образующих из \mathbf{u} в случае, когда замыкание \mathbf{u} относительно \mathbf{U} и Ω не пусто (т.е. $\emptyset \neq [\mathbf{u}]_{\mathbf{U}, \Omega} \subseteq \mathbf{U}$), результаты вычисления алгоритмом 2 по каждому из этих упорядочений — по порождаемым ими траекториям $T_1(\mathbf{u}, \Omega, \mathbf{U})$ и $T_2(\mathbf{u}, \Omega, \mathbf{U})$, совпадают.
- (ii) Если тестовое \mathbf{u} не входит ни в один из элементов исходного Ω , то \mathbf{u} не замкнуто и его нет в множестве $\mathbf{Dom}(\Omega)$ всех пересечений примеров из Ω . При этом найдется $u_0 \in \mathbf{u}$ такая, что в соответствующем ей «усечении» Ω^* исходного Ω замыкание u_0 окажется пустым множеством: $[\{u_0\}]_{\mathbf{U}^*, \Omega^*} = \emptyset$.
- (iii) Если тестовое \mathbf{u} строго вкладывается ровно в один (пусть это будет $O_{\mathbf{u}}$) из элементов исходного Ω (т.е. $\mathbf{u} \subset O_{\mathbf{u}}$), то \mathbf{u} не замкнуто и $[\mathbf{u}]_{\mathbf{U}, \Omega} = O_{\mathbf{u}} = \mathbf{Un}(T)$. При совпадении \mathbf{u} ровно с одним (пусть это будет $O_{\mathbf{u}}$) из элементов исходного множества примеров Ω данное \mathbf{u} замкнуто, а $[\mathbf{u}]_{\mathbf{U}, \Omega} = O_{\mathbf{u}}$.
- (iv) Если тестовое \mathbf{u} вкладывается по крайней мере в два элемента исходного Ω , то $[\mathbf{u}]_{\mathbf{U}, \Omega}$ лежит в множестве $\mathbf{Dom}(\Omega)$ всех пересечений примеров из Ω .
- (v) Четыре возможные ситуации описывают взаимосвязи заданных \mathbf{u} и Ω :

(1) \mathbf{u} не входит ни в один из элементов Ω .

Следуя (ii), в данном случае \mathbf{u} не замкнуто, а по (iii) имеет место строгое вложение $\mathbf{Un}(T) \subset \mathbf{u}$, т.е. условия

$([u]_{U,\Omega} \neq u) \& (Un(T) = u)$ и $([u]_{U,\Omega} = u) \& (Un(T) \neq u)$ в данном случае выполняться не могут;

- (2) u как собственное подмножество входит ровно в один из элементов Ω .

Следуя (iii), имеем $Un(T) = O_u$, т.е. условия $([u]_{U,\Omega} \neq u) \& (Un(T) = u)$ и $([u]_{U,\Omega} = u) \& (Un(T) \neq u)$ в данном случае также выполняться не могут;

- (3) u входит ровно в один из элементов Ω , совпадая с этим примером.

Следуя (iv), множество u замкнуто: $[u]_{U,\Omega} = O_u$. По аналогии с (iii) нетрудно убедиться, что в данном случае также имеет место равенство $Un(T) = O_u$, т.е. формулируемый утверждением 2 критерий выполнен;

- (4) u входит по крайней мере в пару элементов Ω .

С учетом (iv) и алгоритма 1 здесь $[u]_{U,\Omega}$ принадлежит множеству $Dom(\Omega)$.

\Rightarrow . Пусть u замкнуто, тогда $u = \bigcup_{i=1}^k \{a_i\}$ и каждая u_i из u входит в результирующее объединение $Un(T)$; кроме того, каждое $\{a_i\}$ из множества $\{\{u_1\}, \{u_2\}, \dots, \{u_k\}\}$ не выходит за пределы множества u . Таким образом, выполняется и равенство $u = Un(T)$.

Пусть, наоборот, u не замкнуто, тогда в данном случае найдутся $u_{01} \in u$ и $u_{02} \in ([u]_{U,\Omega} \setminus u)$ такие, что $u_{02} \in \{u_{01}\}_{U,\Omega}$, иначе u здесь не может быть незамкнутым. Рассмотрим замыкание $\{u_{01}\}_{U^*,\Omega^*}$ в соответствующем «усечении» Ω^* исходного Ω . Как следствие (i) имеем $u_{02} \in \{u_{01}\}_{U^*,\Omega^*}$, т.е. $u \subset Un(T)$, так как u не содержит u_{02} , а $Un(T)$, наоборот, содержит u_{02} . Итак, в рассматриваемой ситуации формула $([u]_{U,\Omega} \neq u) \& (Un(T) = u)$ не выполняема.

\Leftarrow . Пусть $u \subset Un(T)$, т.е. в $Un(T)$ содержится некоторая образующая u_{02} , которая не входит в u : $u_{02} \notin u$. Тогда среди образующих u'_i , замыкаемых в «усечениях» Ω^* исходного Ω , найдется такая u_{01} , что $u_{02} \in \{u_{01}\}_{U^*,\Omega^*}$ — ее замыкание относительно соответствующего Ω^* содержит u_{02} . Тогда, следуя (i), в множестве $Dom(\Omega)$ можно указать некоторое u^* , которое замкнуто относительно исходного Ω и в которое входят не только все элементы из u , но также дополнительно и u_{02} . т.е. u в данном случае не замкнуто (этому «мешает» u_{02}), а формула $(Un(T) \neq u) \& ([u]_{U,\Omega} = u)$ не выполняема.

4 Примеры использования предлагаемого подхода

Рассмотрим несколько примеров использования предложенной техники при решении задач управления информационными потоками и обеспечения ИБ в компьютерных сетях. Контроль и управление информационными потоками в облачной среде — важные механизмы ИБ [6]. Ключевая проблема в реализации таких механизмов — скорость коммутации пакетов в соответствии с заданными правилами. Критичная «техническая» характеристика здесь — это скорость поиска в соответствующей таблице коммутации (КТ) заголовка пакета, поступающего на вход коммутатора. С формальной точки зрения это скорость поиска предъявляемого КТ заголовка пакета (булевского вектора фиксированной длины) в множестве строк этой таблицы (в реальных приложениях содержащем десятки, а иногда и сотни тысяч заголовков). Массовый приход пакетов и необходимость обрабатывать их в режиме процессно-реального времени (в том числе в условиях внесения динамических изменений в текущее состояние КТ) делают эту область весьма чувствительной к любым возможностям сократить трудоемкость соответствующих процедур проверки.

Примеры 2 и 3 демонстрируют возможности предлагаемой процедурной схемы. При этом каждый заголовок входного пакета (булевский вектор фиксированной длины), так же как и каждая строка соответствующей КТ по схеме из примера 1, могут быть взаимно однозначным образом представлены в виде множеств процедурой кодирования 0 и 1 в соответствующих разрядах отдельными переменными (образующими соответствующего алфавита).

Пример 2. Пусть дано множество примеров $\Omega = \{O_1, O_2, \dots, O_8\}$ (табл. 1). Требуется проверить, имеется ли в Ω пример O_9 , которому может быть сопоставлен вектор (заголовок пакета) (01110)? Для этого на первом шаге выберем столбец для «усечения» табл. 1, подсчитав число нулей и единиц в ее столбцах (табл. 2). Представляется естественным выбрать столбец номер 3 (на котором может быть выделено минимальное число строк табл. 1, имеющих тот же показатель в столбце номер 3, что и для O_9). Таким образом, оказывается, что исходное множество Ω отсекается до подмножества $\Omega^*(u_3) = \{O_3, O_5\}$ (табл. 3), где замыкание $\{u_3\}_{U,\Omega^*(u_3)}$ есть множество $\{u_3, u_5\}$, что исключает попадание O_9 в Ω , так как в O_9 единица в столбце 3, а в пятом — 0.

Пример 3. Пусть множество примеров (описаний прецедентов) $\Omega = \{O_1, O_2, \dots, O_8\}$ представлено

Таблица 1 Описания объектов (примеры 2 и 3)

O	Описание				
O_1	0	1	0	0	1
O_2	1	1	0	1	0
O_3	0	1	1	0	1
O_4	1	1	0	1	1
O_5	0	0	1	0	1
O_6	1	1	0	0	0
O_7	1	0	0	1	0
O_8	0	1	0	1	1

Таблица 2 Число нулей и единиц в столбцах (пример 2)

Параметр	Количество в столбцах				
0	4	2	6	4	3
1	4	6	2	4	5
$\langle 01110 \rangle$	4	6	2	4	3

Таблица 3 Результат усечения (пример 2)

Столбец	Описание				
3	0	1	1	0	1
5	0	0	1	0	1

табл. 1. Требуется проверить, имеется ли в Ω пример O_{10} , которому может быть сопоставлен вектор (заголовок пакета) $\langle 01011 \rangle$.

Для этого, подсчитав число нулей и единиц в столбцах табл. 1 и получив для O_{10} набор оценок (табл. 4 и 5), выберем первый столбец и построим множество $\Omega^*(u_1) = \{O_1, O_3, O_5, O_8\}$ (табл. 6), после чего подсчитаем требуемые 0–1 показатели для табл. 6 (табл. 7) и далее по столбцу номер 3 (с учетом корректности на O_{10} замыкания $\{u_1\}_{U, \Omega^*(u_1)} = \{u_1, u_5\}$) перейдем к новому $\Omega^{**}(u_1, u_3) = \{O_1, O_8\}$ (табл. 8), где далее, например по единственной единице в четвертой колонке (напомним, что примеру O_{10} сопоставлен вектор $\langle 01011 \rangle$), получаем $O_{10} = O_8$, т.е. «тестируемый» пример O_{10} присутствует в исходном Ω .

Еще одна интересная область приложения обсуждаемого подхода — задачи анализа содержимого информационных потоков в компьютерных сетях (см., например, уже упомянутую выше во введении проблематику DPI, IDS/IPS). Здесь одной из базовых процедур предпринимаемого интеллектуального анализа данных оказывается проверка вложимости тех или иных эталонных объектов в пересылаемые пакеты данных, которая может быть формализована, в частности, как задача проверки замкнутости предлагаемого для анализа множества «структурных компонентов» (например, фрагмен-

Таблица 4 Число нулей и единиц в столбцах (пример 3)

Параметр	Количество в столбцах				
0	2	1	0	2	0
1	0	1	2	0	2

Таблица 5 Оценки для примера O_{10}

Заголовок пакета	Оценки				
$\langle 01011 \rangle$	4	6	6	4	5

Таблица 6 Усеченное множество $\Omega^*(u_1) = \{O_1, O_3, O_5, O_8\}$

O	Описание				
O_1	0	1	0	0	1
O_3	0	1	1	0	1
O_5	0	0	1	0	1
O_8	0	1	0	1	1

Таблица 7 Показатели 0–1 для $\Omega^*(u_1)$

Параметр	Количество в столбцах				
0	4	1	2	3	0
1	0	3	2	10	4

Таблица 8 Усеченное множество $\Omega^{**}(u_1, u_3) = \{O_1, O_8\}$

Параметр	Описание				
1	0	1	0	0	1
8	0	1	0	1	1

тов исходного кода) относительно заданного множества прецедентов (т.е. как задача проверки: является ли заданное множество «структурных фрагментов» множеством *всех* тех и *только* тех компонентов, которые являются общими для того или иного подмножества описаний прецедентов). Примеры 4 и 5 демонстрируют возможности предлагаемой процедурной конструкции при решении задач этого типа.

Пример 4. Пусть множество $\Omega = \{O_1, O_2, \dots, O_8\}$ для алфавита $U = \{a_1, a_2, \dots, a_9\}$ сформировано так, как показано в табл. 9. Проверим, является ли замкнутым относительно Ω множество образующих $u = \{u_1, u_2, u_3\} = \{a_3, a_4, a_5\}$. Начнем с выбора образующей из U , по которой будет проведено «усечение» множества Ω (табл. 10). Имея одинаковые показатели во всех трех рассматриваемых столбцах, выберем для дальнейших расчетов столбец a_3 (учитывая, что образующие a_1, a_2 и $a_6–a_9$ следует исключить из рассмотрения как не «участвующие» в формировании множества u) (табл. 11). Принимая во внимание, что на «усе-

Таблица 9 Описания объектов (примеры 4 и 5)

O	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9
O_1						X	X		
O_2						X	X	X	
O_3						X	X	X	X
O_4	X	X	X	X		X	X	X	X
O_5	X	X		X	X				
O_6	X	X	X	X	X				
O_7	X	X	X	X	X				
O_8	X		X	X	X				

Таблица 10 Выбор образующей для усечения Ω (пример 4)

Множество	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9
$\{a_3, a_4, a_5\}$	—	—	4	4	4	—	—	—	—

Таблица 11 Усеченное множество $\Omega^* = \{O_4, O_6, O_7, O_8\}$

O	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9
O_4	X	X	X	X		X	X	X	X
O_6	X	X	X	X	X				
O_7	X	X	X	X	X				
O_8	X		X	X	X				

ченном» множестве $\Omega^*(u_1) = \{O_4, O_6, O_7, O_8\}$ замыкание $[\{u_1\}]_{\mathbf{U}, \Omega^*(u_1)} = \{a_1, a_3\}$ и то, что a_1 не входит в \mathbf{u} , получаем, что \mathbf{u} не является замкнутым относительно исходного Ω .

Пример 5. Пусть, как и в примере 4, множества $\Omega = \{O_1, O_2, \dots, O_8\}$ и $\mathbf{U} = \{a_1, a_2, \dots, a_9\}$ представлены табл. 9. Будет ли $\mathbf{u} = \{u_1, u_2, u_3, u_4\} = \{a_1, a_2, a_3, a_4\}$ замкнутым относительно Ω ? Оценим встречаемость образующих из \mathbf{u} в описаниях прецедентов из Ω (табл. 12). Проведем «усечение» Ω по a_2 (табл. 13) и убедимся, что на «суженном» множестве $\Omega^*(u_2) = \{O_4, O_5, O_6, O_7\}$ происходит замыкание $[\{u_2\}]_{\mathbf{U}, \Omega^*(u_2)} = \{a_1, a_2\}$ при выполнении включения $a_1 \in \{a_1, a_2, a_3, a_4\} = \mathbf{u}$. «Сузим» $\Omega^*(u_2)$ по a_3 (табл. 14), получив $\Omega^{**}(u_2, u_3) = \{O_4, O_6, O_7\}$, и (по a_4) — множество $\Omega^{***}(u_2, u_3, u_4) = \{O_4, O_6\}$, порождающее (сходством примеров O_4 и O_6 из исходного Ω — табл. 15) замкнутое подмножество $\mathbf{u} = \{u_1, u_2, u_3, u_4\} = \{a_1, a_2, a_3, a_4\}$.

Таблица 12 Показатели 0–1 для Ω (пример 5)

Множество	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9
$\{a_1, a_2, a_3, a_4\}$	5	4	4	4	—	—	—	—	—

Таблица 13 Усечение Ω по a_2

O	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9
O_4	X	X	X	X		X	X	X	X
O_5	X	X		X	X				
O_6	X	X	X	X	X				
O_7	X	X	X	X	X				

Таблица 14 Показатели 0–1 для $\Omega^*(u_2)$

Множество	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9
$\{a_1, a_2, a_3, a_4\}$	—	—	4	4	4	—	—	—	—

Таблица 15 Усеченное множество $\Omega^{***}(u_2, u_3, u_4) = \{O_4, O_6\}$

O	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9
O_4	X	X	X	X		X	X	X	X
O_6	X	X	X	X	X				

5 Заключение

Завершая обсуждение, обратим внимание на некоторые особенности представленной процедурной конструкции. Так, сравнивая «прямой» (см. алгоритм 1) и оптимизированный (см. алгоритм 3) способы проверки замкнутости, отметим возможность получить отрицательный ответ (т.е. продемонстрировать незамкнутость тестируемого множества образующих) в том числе и до полного завершения всей процедуры (например, в ситуации, когда замыкание по текущему столбцу (образующей) множества примеров выводит получаемый результат за пределы тестируемого множества). Сокращение объемов вычислений дает также переход от исходного множества описаний прецедентов к его все более «глубоко» усекаемым подмножествам.

При проверках встречаемости заголовка пакета в таблице коммутации КТ известен (см., например, [7]) быстрый алгоритм, использующий лексикографическое упорядочение строк КТ и дальнейшее преобразование этого упорядочения в дерево, которое позволяет диагностировать разряд за рядом булевский вектор заголовка текущего анализируемого пакета в процессе вычислений, объем которых линейно зависит от длины такого заголовка. Прямой перенос этой технологии на проверку замкнутости произвольных подмножеств исходного алфавита оказывается затруднительным. Это дополнительный аргумент в пользу предложенной здесь процедурной конструкции.

В условиях жестких ограничений реального времени выполнения ИАД предлагаемая техника формирования замыканий Галуа легко может быть реализована в режиме параллельных вычислений: в анализе замкнутости сразу нескольких множеств образующих каждое из них может быть рассмотрено в рамках отдельного вычислительного процесса (оперирующего копией исходного множества описаний прецедентов и собственно тестируемым множеством образующих). Более того,

представляющая исходные данные таблица [ОБЪЕКТЫ \times ОБРАЗУЮЩИЕ], как несложно показать, может быть разбита на «клетки» по подмножествам строк и подмножествам столбцов так, что каждая из «клеток» может обрабатываться независимо от других. При этом финальный результат ИАД (например, наличие заголовка входного пакета в таблице КТ или же замкнутость «тестового» множества образующих) будет характеризоваться достаточно простыми комбинациями результатов вычислений в порожденных «клетках» (например, наличием хотя бы одной «клетки», где зафиксирован эффект замкнутости соответствующего множества образующих; совместным рассмотрением всех демонстрирующих эффект замкнутости клеток и т. п.). Число «параллельных» процессов такого вида будет определяться лишь характеристиками соответствующей вычислительной установки. Дополнительные возможности для сокращения объемов перебора здесь можно получить, используя предложенную в работе [8] технику декомпозиции исходных описаний прецедентов на самостоятельно обрабатываемые псевдодеревья (диаграммы вложимости замыканий Галуа).

Наконец, при переходе к «усеченным» подмножествам исходного множества примеров можно останавливать дальнейшие шаги процедуры «усечения», получив текущее подмножество такого размера, что его можно полностью «упаковать» в то или иное схемное решение (TCAM (ternary content-addressable memory) или, например, FPGA (field-programmable gate array) соответствующего размера), где дальнейшие вычисления могут быть осуществлены аппаратными средствами за один «такт» работы такого устройства.

Рассмотренные особенности предложенной технологии позволяют говорить о ней как о еще одном инструментальном основании для разработки активно развиваемых в настоящее время специализированных компьютерных решений-«акселераторов» (см., например, обзор [9]), кото-

рые расширяют спектр возможностей увеличения производительности проблемно-ориентированных прикладных систем ИАД.

Литература

1. Финн В. К. О машинно-ориентированной формализации правдоподобных рассуждений в стиле Ф. Бэкона — Д. С. Милля // Семиотика и информатика, 1983. Вып. 20. С. 35–101.
2. McCarthy J. Phenomenal data mining: From data to phenomena. — Stanford, CA, USA: Computer Science Department, Stanford University, 2000. <http://jmc.stanford.edu/articles/phenomenal.html>.
3. McCarthy J. John McCarthy's home page. <http://www-formal.stanford.edu/jmc>.
4. Финн В. К. Индуктивные методы Д. С. Милля в системах искусственного интеллекта // Искусственный интеллект и принятие решений, 2010. Ч. I. № 3. С. 3–21; Ч. II. № 4. С. 14–40.
5. Кон П. М. Универсальная алгебра / Пер. с англ. — М.: Мир, 1968. 359 с. (Cohn P. M. Universal algebra. — New York, NY, USA: Harper and Row, 1965. 333 p.)
6. Грушо А. А., Забейжайло М. И., Зацаринный А. А. Контроль и управление информационными потоками в облачной среде // Информатика и её применения, 2015. Т. 9. Вып. 4. С. 95–101.
7. Кнут Д. Искусство программирования. Т. 3. Сортировка и поиск / Пер. с англ. — М: Вильямс, 2007. 832 с. (Knuth D. E. The art of computer programming. 3: Sorting and searching. — 2nd ed. — Reading, MA, USA: Addison-Wesley, 1998. 780 p.)
8. Забейжайло М. И. О некоторых возможностях управления перебором в ДСМ-методе // Искусственный интеллект и принятие решений, 2014. Ч. I. № 1. С. 95–110; Ч. II. № 3. С. 3–21.
9. Грушо А. А., Забейжайло М. И., Зацаринный А. А., Писковский В. О., Борохов С. В. О возможностях приложений интеллектуального анализа данных в задачах обеспечения информационной безопасности облачных сред // НТИ. Сер. 2: Информационные процессы и системы, 2015. № 11. С. 1–11.

Поступила в редакцию 15.10.16

ON THE ADVANCED PROCEDURE TO REDUCE CALCULATION OF GALOIS CLOSURES

A. A. Grusho, M. I. Zabezhailo, and A. A. Zatsarinny

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

Abstract: Formalization of similarity by algebraic operation is used as a key element of many modern intelligent data analysis methods. Nevertheless, in some important cases (e. g., in computer network traffic control, network

security policy control in cloud computing environment, and some other), direct implementation of this technique is limited by the necessity to process huge amount of data in the real time mode. For example, it is necessary to intersect elements of a large set of Boolean vectors of large length to find fixed points of so-called Galois closure. An advanced algorithm of Galois closure formation is introduced. The algorithm is used to optimize the process of set closeness checking. Some examples of applications of the presented technique in computer network traffic control and deep packet inspection are discussed.

Keywords: intelligent data analysis; object similarity formalized as algebraic operation; combinatorial search optimization; header analysis; traffic control in computer networks; information security in cloud computing environment

DOI: 10.14357/19922264160410

Acknowledgments

The paper was supported by the Russian Foundation for Basic Research (project 15-29-07981).

References

1. Finn, V. K. 1983. O mashinno-orientirovannoy formalizatsii pravdopodobnykh rassuzhdeniy v stile F. Bacona—J. S. Milla [On the computer-oriented formalization of plausible reasoning in the style of F. Bacon—J. S. Mill]. *Semiotika i informatika* [Semiotics and Informatics] 20:35–101.
2. McCarthy, J. 2000. *Phenomenal data mining: From data to phenomena*. Stanford, CA: Computer Science Department, Stanford University. Available at <http://jmc.stanford.edu/articles/phenomenal.html> (accessed October 15, 2016).
3. McCarthy, J. John McCarthy's home page. Available at <http://www-formal.stanford.edu/jmc/> (accessed October 15, 2016).
4. Finn, V. K. 2011–2012. J. S. Mill's inductive methods in artificial intelligence systems. *Scientific and Technical Information Processing*. Part I. 2011. 38(6):385–402; Part II. 2012. 39(5):241–261.
5. Cohn, P. M. 1965. *Universal algebra*. New York, NY: Harper and Row. 333 p.
6. Grusho, A. A., M. I. Zabezhailo, and A. A. Zatsarinny. 2015. Kontrol' i upravlenie informatsionnymi potokami v oblachnoy srede [Information flow monitoring and control in the cloud computing environment]. *Informatika i ee Primeneniya — Inform. Appl.* 9(4):95–101.
7. Knuth, D. E. 1998. *The art of computer programming. 3: Sorting and searching*. 2nd ed. Reading, MA: Addison-Wesley. 780 p.
8. Zabezhailo, M. I. 2014. O nekotorykh vozmozhnostyakh upravleniya pereborom v DSM-metode [On some new possibilities to control computational complexity of hypotheses generation in JSM-method]. *Iskusstvennyy intellekt i prinyatie resheniy* [Artificial Intelligence and Decision Making]. Part I: 1:95–110; Part II: 3:3–21.
9. Grusho, A. A., M. I. Zabezhailo, A. A. Zatsarinny, V. O. Piskovski, and S. V. Borokhov. 2015. O vozmozhnostyakh prilozheniy intellektual'nogo analiza dannykh v zadachakh obespecheniya informatsionnoy bezopasnosti oblachnykh sred [On the applications of intelligent data analysis in information security of cloud computing environment]. *Nauchno-tehnicheskaya informatsiya* [Scientific and Technical Information]. Ser. 2. 11:1–11.

Received October 15, 2016

Contributors

Grusho Alexander A. (b. 1946) — Doctor of Science in physics and mathematics, professor, Head of Laboratory, Institute of Informatics Problems, Federal Research Center “Computer Sciences and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; grusho@yandex.ru

Zabezhailo Michael I. (b. 1956) — Candidate of Science (PhD) in physics and mathematics, associate professor, Head of laboratory, Institute of Informatics Problems, Federal Research Center “Computer Sciences and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; m.zabezhailo@yandex.ru

Zatsarinny Alexander A. (b. 1951) — Doctor of Science in technology, professor, Deputy Director, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; azatsarinny@ipiran.ru

СИСТЕМОТЕХНИЧЕСКИЕ ПОДХОДЫ К СОЗДАНИЮ СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ НА ОСНОВЕ СИТУАЦИОННОГО АНАЛИЗА

А. А. Зацаринный¹, А. П. Сучков²

Аннотация: Обсуждаются вопросы создания систем поддержки принятия решений (СППР) на основе ситуационного анализа текущей и прогнозируемой обстановки в контролируемом пространстве органа управления. Как правило, такие системы управления в режиме реального времени опираются на ситуационные центры (СЦ) — совокупность информационных, программных и аппаратных средств, а также обслуживающего персонала, реализующих информационные технологии по мониторингу обстановки, ее ситуационному анализу для выработки решений и алгоритмов применения управляющих воздействий. Рассмотрены содержательные характеристики составляющих частей СППР, реализующих полный цикл управления от целеполагания до контроля исполнения принимаемых решений. Отмечается, что реализация СППР зависит от уровня системы управления — стратегического, оперативного, тактического, базового, приводятся функциональные особенности и способы анализа обстановки на различных уровнях системы управления.

Ключевые слова: ситуационный анализ; система поддержки принятия решений; система управления; ситуационный центр

DOI: 10.14357/19922264160411

1 Введение

В Стратегии национальной безопасности Российской Федерации (утверждена Указом Президента Российской Федерации от 31 декабря 2015 г. № 683) [1] определено, что информационную основу реализации Стратегии составляет федеральная информационная система стратегического планирования, включающая в себя информационные ресурсы органов государственной власти и органов местного самоуправления, системы распределенных СЦ и государственных научных организаций. В рамках такой системы должна быть реализована поддержка управленческих решений в интересах центральных органов исполнительной власти на основе организации взаимодействия региональных и ведомственных СЦ, а также других информационных систем. Для эффективного решения этой задачи необходимо создание СППР в составе СЦ и придания им принципиально новых качеств.

В связи с этим целью статьи является обоснование системотехнических и методических подходов к структурному и функциональному составу СППР и ее месту в составе СЦ, обеспечивающих информационно-аналитическую поддержку

принятия управленческих решений в рамках государственного управления, стратегического планирования и мониторинга реализации документов стратегического планирования в Российской Федерации.

2 Базовые понятия

При рассмотрении системных и методических вопросов создания СППР, основанных на ситуационном анализе, в статье используется ряд базовых понятий: событие, обстановка, ситуация, угроза, управление, цели управления и др. [2].

Ситуация определяется состоянием взаимосвязанных *элементов обстановки* в контролируемом пространстве; изменения обстановки определяются *событиями*, образующими некоторые разворачивающиеся во времени наблюдаемые и регистрируемые потоки. При этом под *управлением* понимается **целенаправленное** воздействие органа управления на подчиненные ему или взаимодействующие элементы обстановки (ресурсы).

Совокупность ситуаций в системе управления распадается на текущие, прогнозируемые и целевые ситуации. При этом текущие ситуации являют-

¹ Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, AZatsarinny@ipiran.ru

² Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, Asuchkov@ipiran.ru

ся результатом наблюдения и регистрации событий, прогнозируемые определяются методами ситуационного анализа, а целевые отражают краткосрочные, среднесрочные и долгосрочные цели управления. Последнее немаловажно, так как зачастую ситуационный анализ понимается как обеспечение реакций системы управления на чрезвычайные ситуации после того, как они сложились. Однако теория ситуационного подхода предполагает учет «планируемой и прогнозируемой обстановки», отражающей стратегические, тактические и оперативные цели управления, а также учет факторов самоорганизации управляющего сегмента системы, определяющих стимулы для достижения этих целей [2, 3]. Под *угрозой* в процессах управления понимается ситуация или совокупность ситуаций, развитие которых противоречит целям управления и отдаляет текущее состояние от целевого.

В конце 1970-х гг. была создана модель системы управления «наблюдение—ориентирование—решение—действие» (НОРД) для принятия решений при ведении боевых действий [4, 5]. В настоящее время эта модель активно используется во многих системах управления разных отраслей [6]. В рамках ситуационного подхода к управлению предложена модифицированная модель, включающая дополнительную стадию управляющего цикла — целеполагание [7].

Целеполагание (стадия Ц) — формализованное представление целевых показателей, установление количественных и временных критериев их достижения.

Мониторинг (стадия М) — это процесс сбора информации об окружающей среде в контролируемом пространстве, включая состояние целевых показателей. Стадия М также принимает внутренние инструкции от стадии анализа (А), так же как и поддержку от процессов Р и Д.

Анализ (стадия А) — оценка ситуации (типовая, нетиповая), анализ существующего опыта, пополнение опыта, обеспечивает внутреннюю поддержку М (корректировка фильтров).

Решение (стадия Р) — это процесс осуществления выбора среди гипотез о состоянии окружающей среды и возможной реакции на него. Процесс Р руководствуется прямой внутренней связью с процессом А и обеспечивает внутреннюю поддержку процесса М, возможна корректировка целевых показателей (стадия Ц).

Действие (стадия Д) — это процесс выполнения выбранной реакции путем взаимодействия с окружающей средой. Действие принимает внутренние руководства от процесса А, также оно напрямую связано с Р. Оно обеспечивает внутреннюю поддержку Ц и М.

Особенности реализации цикла управления в системе, реализующей процессы стратегического планирования и управления, заключаются в том, что содержательно стадии Ц, А и Р реализуются непосредственно высшими органами исполнительной власти. Это означает осуществление следующих основных функций:

- доведение до подчиненных органов данных целеполагания и стратегического планирования на основе их формализации и регламентации обмена (стадия Ц);
- регламентированный сбор данных о состоянии целевых показателей от органов исполнительной власти и об обстановке в контролируемом пространстве по определенному регламенту и в режиме реального времени (стадия М);
- обмен аналитическими данными участников стратегического планирования по целеполаганию, прогнозированию, планированию и программированию — федеральных органов исполнительной власти (ФОИВ), субъектов Российской Федерации и муниципальных образований, отраслей экономики и сфер государственного и муниципального управления (стадия А);
- доведение до подчиненных органов принимаемых решений по применению сил и средств и, возможно, по корректировке стратегических планов с целью достижения поставленных стратегических целей (стадия Р) и контроль исполнения решений (стратегических планов) на основе докладов (стадия Д).

На тактическом и базовом уровнях управления осуществляются, во-первых, реализация функций мониторинга контролируемого пространства и организации учета контролируемых объектов (стадия М), во-вторых, специальный анализ фактографических данных о конкретных элементах обстановки, формализованных в виде семантической сети, позволяющий выявлять неочевидные связи между элементами обстановки, определять схожие пространственно-событийные ситуации, выявлять ассоциативные связи и закономерности с целью поддержки процессов принятия решений (стадия А), в-третьих, процессы принятия решений по планированию применения сил и средств на период времени и по складывающейся обстановке в соответствии с указаниями вышестоящих органов (стадии Р и Д).

3 Ситуационный центр как составляющая современной системы управления

Определим СЦ системы управления как совокупность информационных, программных и аппаратных средств, а также обслуживающего персонала, реализующих информационные технологии по мониторингу обстановки, ее ситуационному анализу для выработки решений и алгоритмов применения управляющих воздействий с целью эффективной реализации функций управления и минимизации ущерба от угроз в зоне ответственности органа управления, доведения их до объектов управления и контроля исполнения,

По сути дела, СЦ является составной частью системы управления, осуществляющей автоматизацию ряда функций всего органа управления и отдельных должностных лиц.

Исходя из накопленного в Институте проблем информатики РАН опыта разработки крупных информационных систем в интересах органов государственной власти, в организационной структуре СЦ можно выделить четыре основных функциональных сегмента (рис. 1) [8]:

- (1) сегмент руководства (лиц, принимающих решения, ЛПР);
- (2) сегмент мониторинга состояния контролируемых объектов и окружающей среды и сбора информации;

- (3) сегмент ситуационного анализа и систематизации информации;
- (4) сегмент администрирования и эксплуатации.

При этом СППР базируется на ресурсах всех четырех сегментов. Вместе с тем центральным звеном СЦ и его СППР, обеспечивающим реализацию основной функции системы управления по эффективному управлению силами и средствами, является *сегмент ситуационного анализа и систематизации информации*. Он должен обеспечивать реализацию следующих функций:

- возможность визуализации результатов анализа обстановки на индивидуальных и коллективных средствах отображения;
- во взаимодействии с сегментом мониторинга получение данных о состоянии обстановки от собственных (субъективных и объективных средств наблюдения и контроля) и внешних по отношению к системе источников информации (ведомственных, межведомственных, международных, независимых и др.);
- извлечение фактов, структуризация и формализация разнородных данных о значимых событиях в соответствии с выбранной информационной моделью предметной области;
- формирование хранилищ ситуационных данных;
- формирование способов визуализации агрегированных данных о складывающейся обстановке для ЛПР и оперативного состава;

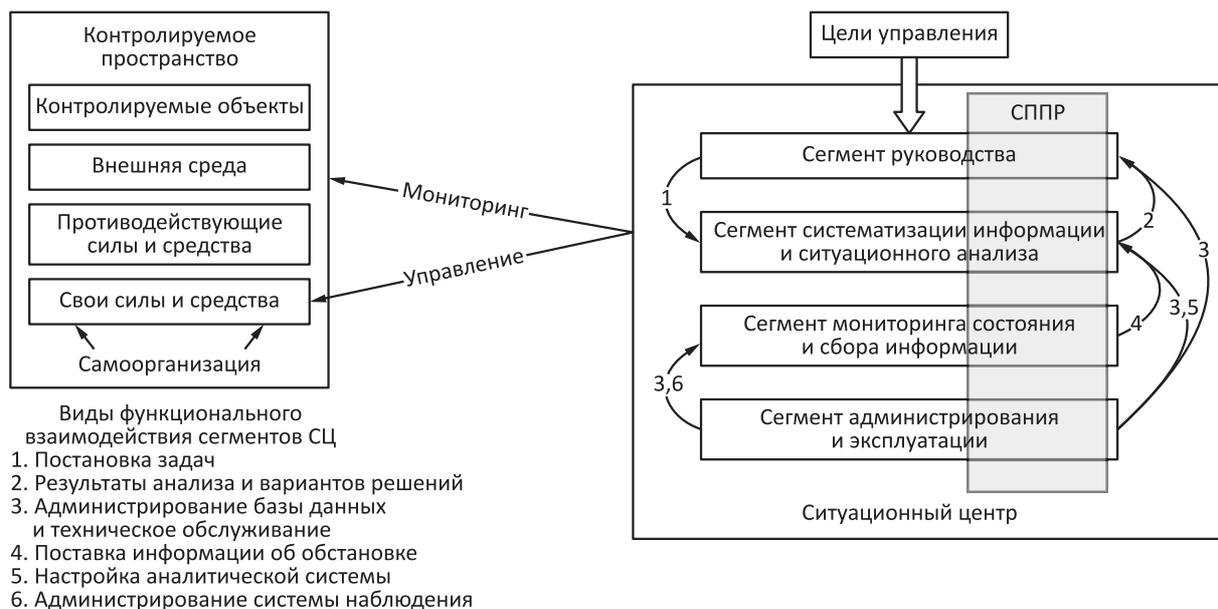


Рис. 1 Обобщенная функциональная структура СЦ

- формирование отчетности и служебной документации;
- расчет первичных и интегральных показателей обстановки, а также статистическая оценка характеристик ненаблюдаемых элементов обстановки;
- решение задач перспективного планирования, контроль исполнения решений по планированию;
- выявление значимых ситуаций, их ранжирование по степени важности, видам и типам, формирование текущего перечня аналитических задач по складывающейся обстановке и по поручениям руководства;
- выработка вариантов решений по применению управляющих воздействий для достижения целевых ситуаций, формирование способов наглядного представления вариантов решения для ЛПП (оперативное планирование);
- прогнозирование развития обстановки и процесса реализации целей системы управления на основе сформированных ситуационных моделей и моделей угроз, в том числе и с учетом применения выработанных вариантов решений;
- обеспечение процессов принятия решений комплексом информационно-расчетных задач (ИРЗ).

Наряду с перечисленными в СППР СЦ реализуются важнейшие функции администрирования аналитической подсистемы СЦ:

- формирование и корректировка системы целей управления;
- формирование, настройка и корректировка системы моделей целей управления, обстановки, ситуаций и угроз;
- формирование, настройка и корректировка системы расчетных показателей, характеризующих обстановку и ее элементы;
- формирование, настройка и корректировка системы критериев, пороговых значений, эвристик, параметров расчетных алгоритмов.

4 Целеполагание — определение целей системы управления

Под *целью ситуационного анализа* предлагается понимать поддержку процессов принятия решений для достижения поставленных целей путем

применения доступных в системе управления сил и средств (ресурсов).

Целесообразность деятельности системы управления определяется иерархической системой целей (подцелей). Для ФОИВ она задается законодательно, а также при определении приоритетов в организации деятельности системы управления первым лицом (руководителем). Формирование системы целей сопровождается формированием системы показателей реализации целей (подцелей) и критериев достижения целей. Показатели являются вычисляемыми величинами как функции обстановки или экспертно оцениваемые параметры. Критерии достижения обычно формулируются как некие пороговые плановые значения на временной шкале.

Эффективность системы управления в каждый момент времени определяется, во-первых, степенью достижения пороговых значений планируемых целевых показателей, во-вторых, объемом затрачиваемых ресурсов на единицу оптимизируемого целевого показателя.

Цели управления формируются на основании системного анализа нормативно-правовых основ функционирования системы управления. Цели управления образуют дерево целей, детализация которого (число уровней) определяется возможностью декомпозиции конкретной цели на значимые подцели. Цели и подцели должны обладать индикаторами состояния (как правило, %) и весовыми коэффициентами доли подцели в реализации всей цели. Цели могут включать ориентиры развития системы управления, установленные первым лицом.

Выбор структуры системы целей предлагается осуществлять с учетом следующих соображений.

1. Цели управления сложной управляющей системой определяются нормативно-правовыми документами, регламентирующими ее функционирование, и, как правило, образуют **иерархическую структуру** в соответствии со структурой направлений деятельности (рис. 2).
2. Ситуационный подход к управлению предполагает реагирование на складывающуюся обстановку в режиме реального времени. В силу этого, помимо фиксированных целей в системе управления необходим механизм формирования **динамических целей**, отражающих процесс нормализации складывающихся чрезвычайных ситуаций и присутствующих в системе целеполагания на период существования ситуации.
3. В концепции «управления по целям» эффективность целеполагания проверяется по критериям SMART [9]: цель должна быть конкретная, измеримая (подразумевает количественную изме-

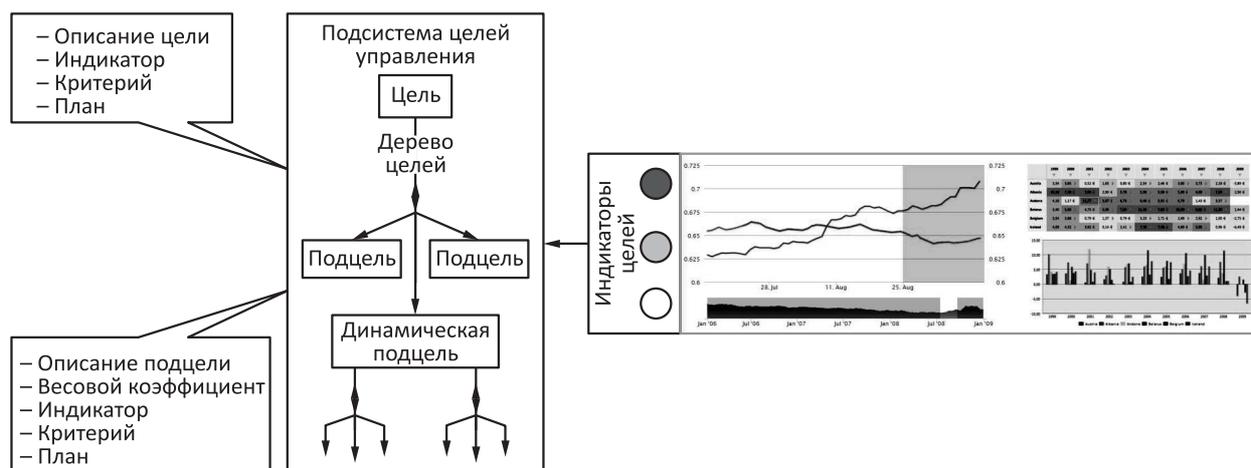


Рис. 2 Обобщенная структура системы целей

римость результата), достижимая (должна быть выполнимой), реалистичная (достижение цели должно быть обеспечено ресурсами), привязанная к точке/интервалу времени.

Данный подход накладывает **требования на атрибуты целей** в части формирования количественных характеристик их достижения, плановых характеристик, критериев достижения (см. рис. 2).

Основные атрибуты цели:

- описание — дает определение и конкретизацию цели;
- весовой коэффициент — определяет вклад подцели в вышестоящую цель;
- индикатор — задает количественный показатель достижения результата;
- критерий — задает способ определения достижения результата с помощью индикатора;
- план — определяет количественные значения критерия достижения цели и требуемые временные параметры.

5 Анализ обстановки и выработка вариантов решений

5.1 Мониторинг обстановки

В процессе мониторинга контролируемых элементов обстановки осуществляются (рис. 3):

- сбор данных о состоянии контролируемых объектов, анализ неструктурированной информации с целью извлечения фактов и знаний;
- постановка объектов на контроль (оператор, автоматически);

- отображение контролируемых объектов по шкале состояний и по критериям — соотношение текущего или прогнозируемого значения индикатора (интегрального показателя) и системы порогов, обеспечивающих градацию состояния («типовое», «чрезвычайное», «критическое» или другие подобные).

По данным мониторинга контролируемых элементов обстановки из различных источников формируется *хранилище* СППР, которое представляет собой совокупность взаимоувязанных на основе единого информационного и лингвистического обеспечения баз данных (БД): обстановки (события, ситуации, элементы окружающей среды), сил и средств (свои силы и средства, противодействующие силы и средства, тактико-технические характеристики), целевых показателей (первичные показатели, интегральные показатели, индикаторы, критерии), типовых решений (типовые решения, конкретные решения), ретроспективная (нормализованные исторические данные, архив обстановки), нормативных документов, библиотека математических моделей.

5.2 Поддержка процесса принятия решений

На основе мониторинга текущей обстановки и поступления событийной информации в хранилище осуществляется расчет заданных в системе первичных и интегральных показателей обстановки и целевых показателей в двух режимах: по регламенту (с определенной периодичностью) и по запросу пользователя с использованием блоков расчетов, блока первичного, краткосрочного, среднесрочного и долгосрочного анализа, блока визуализации и блока поддержки принятия решений (рис. 4).

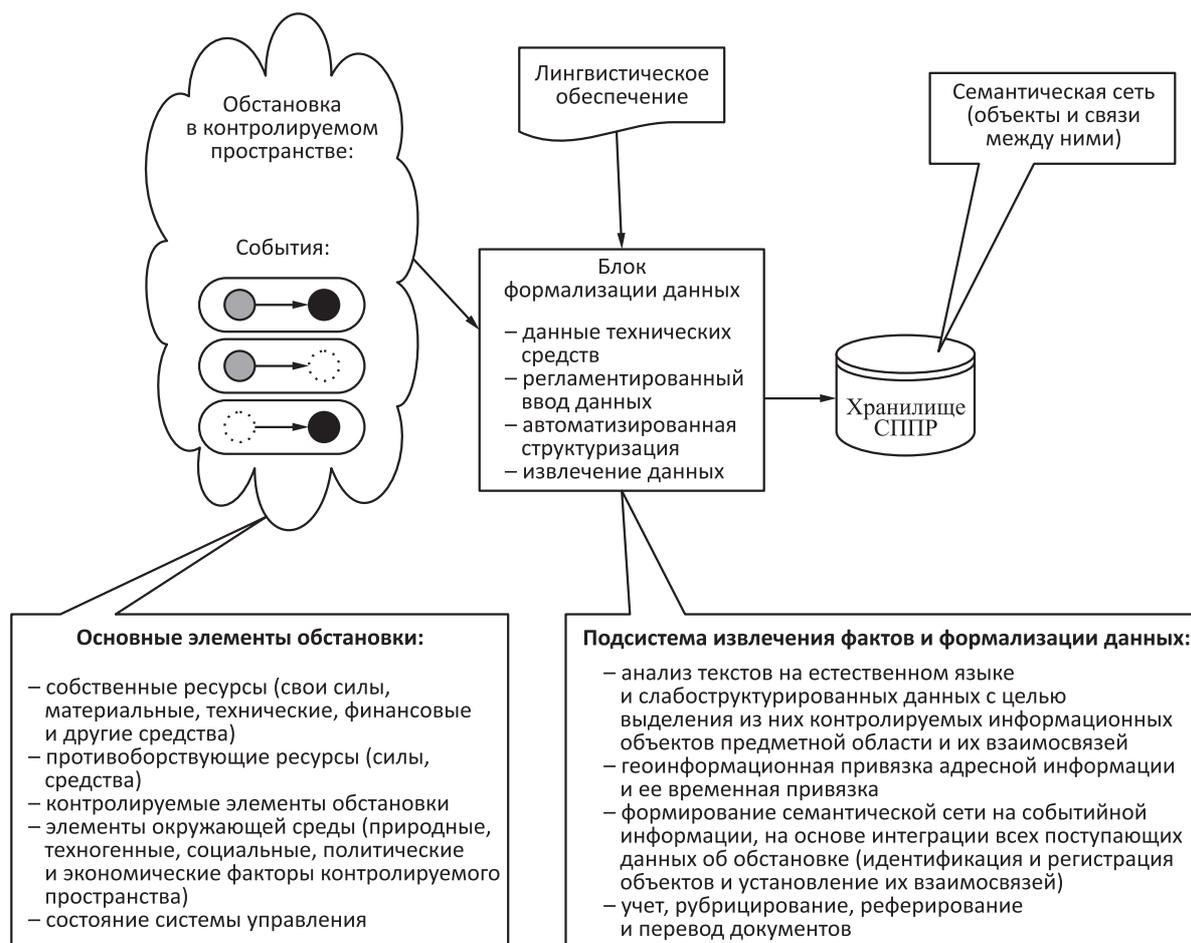


Рис. 3 Мониторинг обстановки

При этом реализуются следующие функции.

1. Создание (привязка существующих) динамических моделей обстановки:
 - моделей «нормальной» обстановки;
 - моделей для прогноза обстановки;
 - моделей для анализа трендов, циклов, аномалий обстановки.
2. Проведение оперативного анализа текущей обстановки с использованием математических методов (см. рис. 4):
 - анализ отклонения от «нормальной» текущей обстановки;
 - прогноз развития обстановки;
 - анализ трендов, циклов, аномалий обстановки;
 - выявление и идентификация значимых ситуаций на основе выявления типовых кон-

фигурций событий и правил идентификации, идентификация типа ситуации, формирование неотложных целей.

3. Визуализация и индикация состояний контролируемых объектов с использованием полученных результатов анализа (наглядное представление текущей с индикацией ситуаций, требующих принятия решения или применения типовых решений).
4. Выработка вариантов решений по складывающейся обстановке (решение содержит динамическую цель, перечень подцелей (с весами — доли подцели в реализации всей цели), сроки достижения подцелей, ответственных, совокупность типовых уведомлений и рапортов):
 - применение типовых решений по типовым ситуациям (привязка их к реальной обстановке);

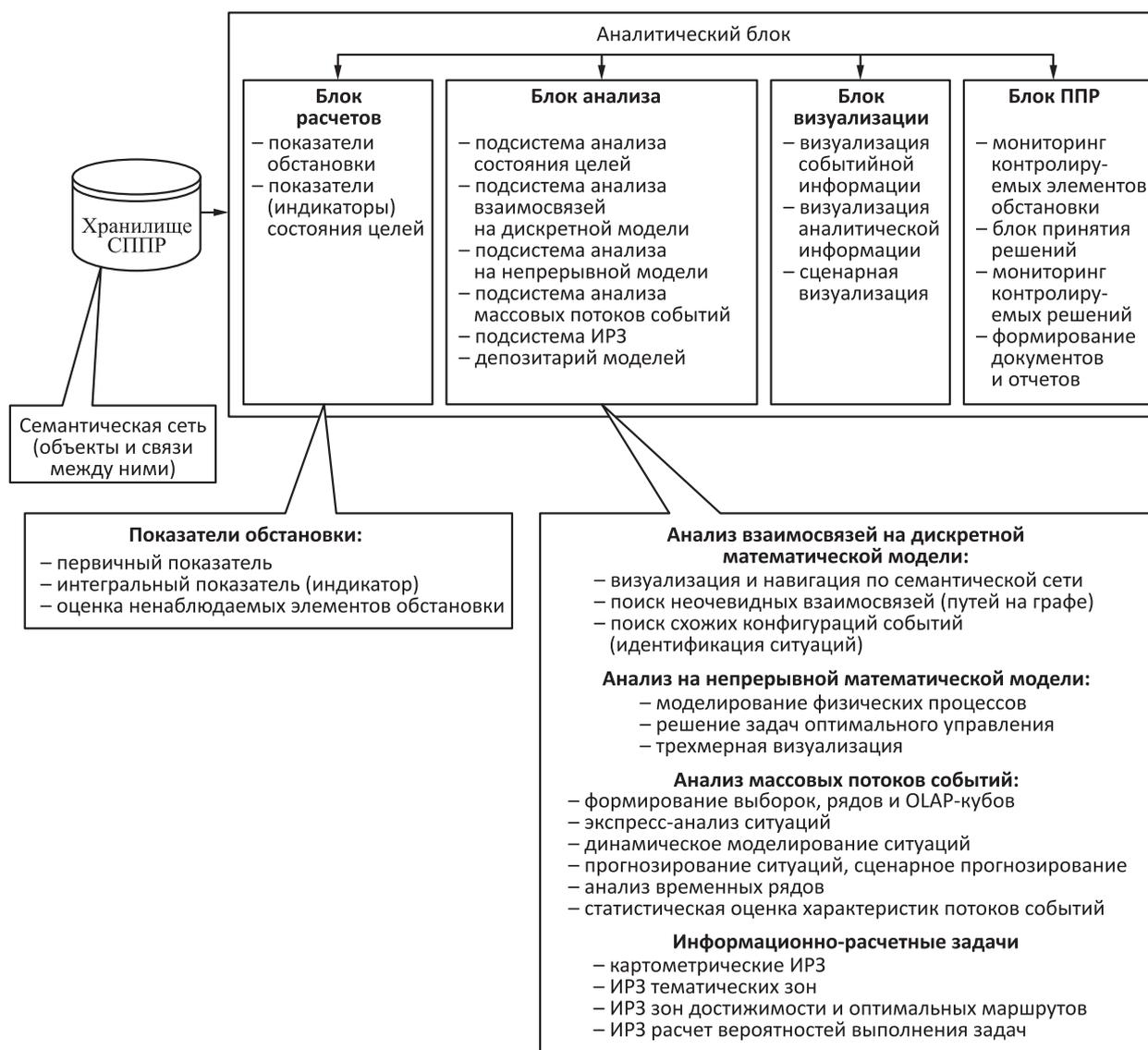


Рис. 4 Структура блока принятия решений

- выработка вариантов решения экспертным путем в случае критических и чрезвычайных ситуаций;
- анализ развития обстановки с учетом вариантов решений (прогноз благоприятного и неблагоприятного развития обстановки, расчет вероятностей выполнения задач, оценка вариантов решений).

5.3 Реализация решений

На данной стадии осуществляется мониторинг процессов реализации решений по краткосрочным, среднесрочным и долгосрочным планам (решение содержит цель, перечень подцелей (с весами — доли

подцели в реализации всей цели), сроки достижения подцелей, ответственных, виды отчетности):

- сбор информации по ходу выполнения плана (отчетность), визуализация хода исполнения, контроль исполнения;
- сравнительный анализ показателей плана по целям и подцелям и текущей обстановки, включая расчет степени реализации плана и прогнозирование возможности реализации плана;
- реализация обратной связи по уточнению решения по планированию с целью обеспечения выполнения плана;
- доведение уточненного решения (уведомления) и контроль исполнения.

Мониторинг реализации решений по ситуациям (решение содержит динамическую цель, перечень подцелей (с весами — доли подцели в реализации всей цели), сроки достижения подцелей, ответственных, совокупность типовых уведомлений и рапортов):

- сбор информации по ходу выполнения решения (рапорты), визуализация хода исполнения, контроль исполнения;
- сравнительный анализ показателей по целям и подцелям и текущей обстановки, включая расчет степени реализации решения и прогнозирование возможности реализации решения;
- реализация обратной связи по уточнению решения по ситуации с целью обеспечения выполнения плана.
- доведение уточненного решения (уведомления) и контроль исполнения.

6 Заключение

1. В современных условиях развития информационных систем особую значимость приобретает актуальность исследования системотехнических и технологических вопросов создания в составе СЦ СППР.
2. Важнейшей методологической и концептуальной основой СППР является полнофункциональный цикл управления, включающий стадии целеполагания, мониторинга обстановки, анализа обстановки, выработки вариантов решений и их реализации.
3. В СППР реализуются следующие функциональные задачи:
 - мониторинг контролируемых элементов обстановки;
 - расчет характеристик событийной информации (первичные и интегральные показатели текущей обстановки и состояния целей);
 - визуализация текущего состояния обстановки;
 - визуализация текущего состояния индикаторов целей;
 - блок анализа и принятия решений.
 - мониторинг контролируемых решений;
 - формирование документов и отчетов.
4. Важнейшим системообразующим компонентом СППР является хранилище, формируемое в автоматизированном режиме из различных источников в виде совокупности взаимосвязанных на основе единого информационного и лингвистического обеспечения БД (о событиях, силах и средствах, целевых показателях и критериях, типовых решений, ретроспективной информации, нормативных документов, математических моделей).
5. Предложенные в статье системотехнические подходы и решения апробированы в рамках нескольких проектов по созданию крупных территориально распределенных информационно-аналитических систем специального назначения.

Литература

1. Стратегия национальной безопасности Российской Федерации. Утверждена Указом Президента Российской Федерации от 31 декабря 2015 г. № 683.
2. Зацаринный А. А., Сучков А. П. Некоторые подходы к ситуационному анализу потоков событий // Открытое образование, 2012. № 1. С. 39–45.
3. Бир С. Э. Мозг фирмы / Пер. с англ. — М.: Радио и связь, 1993. 416 с. (*Beer S. Brain of the firm.* — Allen Lane, The Penguin Press, London; Herder and Herder, USA, 1972. 416 p.)
4. Grant T., Kooter B. Comparing OODA & other models as operational view C2 architecture // 10th Command and Control Research Technology Symposium (International) Proceedings. — McLean, VA, USA, 2005.
5. Ивлев А. А. Основы теории Бойда. Направления развития, применения и реализации. — SlideShare, 2008. 64 с. <http://www.slideshare.net/defensenetwork/ss-10380168>.
6. Босов А. В., Зацаринный А. А., Сучков А. П. Некоторые общие подходы к формированию функциональных требований к ситуационным центрам и их реализации // Системы и средства информатики, 2010. Вып. 20. № 3. С. 98–125.
7. Сучков А. П. Формирование системы целей для ситуационного управления // Системы и средства информатики, 2013. Т. 23. № 2. С. 171–182.
8. Зацаринный А. А., Сучков А. П., Козлов С. В. Особенности проектирования и функционирования системы ситуационных центров // Системы высокой доступности, 2012. Т. 8. № 1. С. 12–21.
9. Doran G. T. There's a S.M.A.R.T. way to write management's goals and objectives // Manag. Rev., 1981. Vol. 70. Iss. 11. P. 35–36.

Поступила в редакцию 23.08.16

SYSTEMS ENGINEERING APPROACHES TO THE ESTABLISHMENT OF A SYSTEM FOR DECISION SUPPORT BASED ON SITUATIONAL ANALYSIS

A. A. Zatsarinny and A. P. Suchkov

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

Abstract: The article discusses the issues of decision-making support systems (DMSS) creation based on the situational analysis of the current and projected situation in the controlled space. Typically, such control systems in real time are based on situational centers, which are sets of information, software, hardware, and staff implementing information technology to monitor the situation and its situational analysis to develop solutions and algorithms application of control actions. The paper considers characteristics of the DMSS components, implementing the full management cycle from goal setting to execution control decisions. It is noted that the implementation of the decision support system depends on the level of management — strategic, operational, tactical, basic, and functional features and methods of analysis of the situation at different levels of the control system.

Keywords: situational analysis; system of decision-making process support; management system; situational center

DOI: 10.14357/19922264160411

References

1. Strategiya natsional'noy bezopasnosti Rossiyskoy Federatsii [The National Security Strategy of the Russian Federation]. Approved by the Decree of the President of the Russian Federation No. 683, 31.12.2015.
2. Zatsarinny, A. A., and A. P. Suchkov. 2012. Nekotorye podkhody k situatsionnomu analizu potokov sobytiy [Some approaches to the situational analysis of the flows of events]. *Otkrytoe obrazovanie* [Open Education] 1:39–45.
3. Beer, S. 1972. *Brain of the firm*. Allen Lane, The Penguin Press, London; Herder and Herder, USA. 416 p.
4. Grant, T., and B. Koote. 2005. Comparing OODA & other models as operational view C2 architecture. *10th Command and Control Research Technology Symposium (International) Proceedings*. McLean, VA. USA.
5. Ivlev, A. A. 2008. *Osnovy teorii Boyda. Napravleniya razvitiya, primeneniya i realizatsii* [Fundamentals of the theory of Boyd. Areas of development, application, and implementation]. SlideShare. Available at: <http://www.slideshare.net/defensenetwork/ss-10380168> (accessed October 29, 2016).
6. Bosov, A. V., A. A. Zatsarinny, A. P. Suchkov. 2010. Nekotorye obshchie podkhody k formirovaniyu funktsional'nykh trebovaniy k situatsionnym tsestram i ikh realizatsii [Some common approaches to the formation of functional requirements for situation centers and their implementation]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 20(3):98–125.
7. Suchkov, A. P. 2013. Formirovanie sistemy tseley dlya situatsionnogo upravleniya [The formation of the objective system to situational management]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 23(2):171–182.
8. Zatsarinny, A. A., A. P. Suchkov, and S. V. Kozlov. 2012. Osobennosti proektirovaniya i funktsionirovaniya sistemy situatsionnykh tsestrov [Features of the design and functioning of the situational centers]. *Sistemy Vysokoy Dostupnosti* [High Availability Systems] 8(1):12–21.
9. Doran, G. T. 1981. There's a S.M.A.R.T. way to write management's goals and objectives. *Manag. Rev.* 70(11):35–36.

Received August 23, 2016

Contributors

Zatsarinny Alexander A. (b. 1951) — Doctor of Science in technology, professor, Deputy Director, Federal Research Center “Computer Sciences and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; AZatsarinny@ipiran.ru

Suchkov Alexander P. (b. 1954) — Doctor of Science in technology, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; Asuchkov@ipiran.ru

НЕФОРМАЛЬНАЯ АКСИОМАТИЧЕСКАЯ ТЕОРИЯ РОЛЕВЫХ ВИЗУАЛЬНЫХ МОДЕЛЕЙ*

А. В. Колесников¹, С. В. Листопад², С. Б. Румовская³, В. И. Данишевский⁴

Аннотация: Актуальность построения неформальной аксиоматической теории ролевых визуальных моделей обусловлена моделированием визуально-образных рассуждений в гибридных и синергетических интеллектуальных системах. Основные исследования визуально-образных рассуждений сосредоточены на специальных визуальных языках представления некоторых видов данных, информации и знаний. Отсутствие формализованных моделей визуальных языков — причина высокой наукоемкости разработки специальных сред манипулирования и обработки визуальных моделей. Построение неформальной аксиоматической теории ролевых визуальных моделей — шаг к новому классу интеллектуальных систем, релевантных реальным коллективам, принимающим решения, — гибридным интеллектуальным системам (ГиИС) с гетерогенным визуальным полем, имитирующим сотрудничество, относительность и дополнительность коллективного интеллекта, рассуждающим на символических и визуальных языках.

Ключевые слова: гибридная интеллектуальная система; гетерогенное визуальное поле; визуальный язык; семиотическая система

DOI: 10.14357/19922264160412

1 Введение

Принятие коллективных решений — сложное активное взаимодействие участников и обеспечение взаимопонимания между ними. Для интенсификации этих процессов применяются методы визуализации информации, своеобразные визуальные языки, наглядно описывающие структуру, свойства и отношения понятий предметной области. Правильно составленный график или диаграмму значительно проще анализировать, чем многостраничную таблицу с результатами измерений, а тем более текстовое их описание. Визуализация информации позволяет имитировать рассуждения на основе визуальных образов, обладающих большей конкретностью и интегрированностью, чем символические представления.

Рассуждения на визуальных образах рассматривались в работах Д. А. Поспелова, Г. П. Щедровицкого, Ю. Р. Валькмана, Б. А. Кобринского, О. П. Кузнецова, Г. С. Осипова, В. Б. Тарасова, И. Б. Фоминых, Т. А. Гавриловой, А. Е. Янковской. Визуальные языки разработаны для функционального программирования, программирования на примерах, для конечных автоматов, потоков

данных и других областей [1]. Реализация этих языков требует значительных усилий, разработки для каждого случая специальных сред создания, манипулирования и обработки визуальных моделей. Для их снижения предлагается неформальная аксиоматическая теория ролевых визуальных моделей на основе принципов теории систем и системного анализа.

2 Понятие неформальной аксиоматической теории ролевых визуальных моделей

Рассмотрим существо разработки и использования визуальных моделей на теоретическом уровне, построив неформальную аксиоматическую теорию языков визуального моделирования сложных систем. В качестве обобщения результатов работ по имитации рассуждений на визуальных образах [2–8] предлагается неформальная аксиоматическая теория визуального языка как семиотической системы:

$$v1 = \langle VT, VS, VA, VP, v\tau, v\sigma, v\alpha, v\pi \rangle, \quad (1)$$

* Работа выполнена при поддержке РФФИ (проект 16-07-00271а).

¹ Балтийский федеральный университет им. И. Канта, Калининградский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, avkolesnikov@yandex.ru

² Калининградский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, ser-list-post@yandex.ru

³ Калининградский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, sophiyabr@gmail.com

⁴ Балтийский федеральный университет им. И. Канта, danishevskii.v.i@mail.ru

где VT, VS и VA — множества основных символов, синтаксических правил и аксиом-знаний о предметной области (семантических правил) соответственно; VP — множество правил вывода решений (прагматических правил); $v\tau, v\sigma, v\alpha$ и $v\pi$ — правила изменения множеств VT, VS, VA и VP соответственно. Множества $VT, VS, VA, VP, v\tau, v\sigma, v\alpha$ и $v\pi$ из (1) определяются выражениями:

$$VT = \langle P, D, VR \rangle; \quad (2)$$

$$VS = \langle VT, VN, PRU \rangle; \quad (3)$$

$$VA = \langle DO, G^{RES}, G^{PR}, G^R \rangle, \quad (4)$$

где

$$DO = \langle RES, PR, R \rangle,$$

$$G^{RES} : RES \rightarrow P,$$

$$G^{PR} : PR \rightarrow D,$$

$$G^R : R \rightarrow VR;$$

$$VP = \{ \langle AG, act, M, W \rangle \}; \quad (5)$$

$$v\tau = \langle \Delta P, \Delta D, \Delta VR \rangle; \quad (6)$$

$$v\sigma = \langle v\tau, \Delta VN, \Delta PRU \rangle; \quad (7)$$

$$v\alpha = \langle \Delta DO, G^{\Delta RES}, G^{\Delta PR}, G^{\Delta R} \rangle; \quad (8)$$

где

$$\Delta DO = \langle \Delta RES, \Delta PR, \Delta R \rangle,$$

$$G^{\Delta RES} : \Delta RES \rightarrow \Delta P,$$

$$G^{\Delta PR} : \Delta PR \rightarrow \Delta D,$$

$$G^{\Delta R} : \Delta R \rightarrow \Delta VR;$$

$$v\pi = \{ \langle \Delta AG, \Delta act, \Delta M, \Delta W \rangle \}. \quad (9)$$

Здесь помимо ранее введенных обозначений P — множество визуальных примитивов; D — множество визуальных измерений, характеризующих визуальные примитивы; VR^n — множество визуальных отношений между одним и более примитивами [4]; VN — словарь нетерминальных символов; PRU — множество продукционных правил; RES, PR и R — множества ресурсов, свойств и отношений соответственно; AG — множество носителей языка (экспертов, элементов, агентов), которым адресована норма поведения (различные социальные запреты и ограничения, накладываемые сообществом на отдельного носителя); $act \in ACT$ — действие, определенное на множестве действий ACT и являющееся объектом нормативной регуляции (содержание нормы); M — множество систем модальностей, связанных с действием, например система норм, выраженных деонтическими модальностями: $M_N = \{O, P, B, Z\}$, где O — «обязательно», P —

«разрешено», B — «безразлично», Z — «запрещено»; W — множество миров, в которых применима норма (условия приложения, обстоятельства, в которых должно или не должно выполняться действие) [9]; $\Delta P, \Delta D, \Delta VR, \Delta VN, \Delta PRU, \Delta RES, \Delta PR, \Delta R, \Delta AG, \Delta M$ и ΔW — множества допустимых изменений множеств $P, D, VR, VN, PRU, RES, PR, R, AG, M$ и W соответственно; Δact — множество допустимых изменений содержания нормы act .

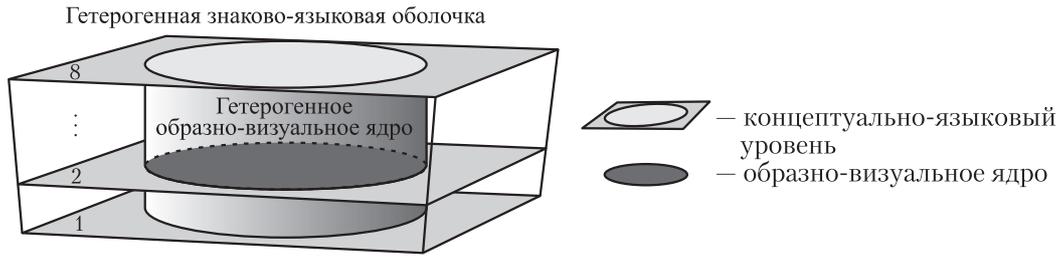
Как показано в [10], языки профессиональной деятельности, в том числе и визуальный язык, — полиязыки [11]. С одной стороны, это обусловлено присущей языку структурированностью внешнего мира. Это приводит к необходимости представления в языке знаков, обозначающих ресурсы, свойства, действия, структуры, ситуации, состояния, поведение, а с учетом деятельности субъекта управления — целей, задач, планов, оценок. В этом случае визуальный язык может рассматриваться как многослойная структура, описывающая решение сложной задачи комбинацией нескольких взаимосвязанных процессов рассуждений на разных языках.

С другой стороны, полиязыковой характер языка профессиональной деятельности — следствие эволюции систем управления в сторону много-модельных, гибридных и гибридных адаптивных систем управления [11]. Это приводит к узкой специализации управленцев-экспертов по профессиональным нишам и к тому, что информация в коллективе, принимающем решения, представляется на широком спектре языков профессиональной деятельности со своими относительно независимыми задачами, лексикой, данными, знаниями, принципами [11]. Стратификация по нишам при разработке функциональных ГИИС редуцирует сложность задач до простых элементов — подзадач в локальных подобластях мира управления — профессиональных нишах.

Рассмотрим подходы к представлению визуального языка как гетерогенной структуры.

3 Многослойная модель визуального языка

В работе [11] информационный язык представлен семейством языков описания ресурсов, операций, структур, ситуаций, состояния, поведения объекта управления, а также целей, планов и задач. В [12] выделено 8 уровней визуальных языков для реализации автоматизированных рассуждений в интеллектуальных системах: (1) концептуального и визуального базиса $v1^1$; (2) ресурсов, действий



Гетерогенное визуально-образное ядро в гетерогенной знаково-языковой оболочке. Уровни знаковых и графических высказываний: 1 — концептуального и визуального базиса; 2 — о ресурсах, действиях и свойствах; 3 — об иерархиях ресурсов, действий, свойств; 4 — о пространственных и производственных структурах; 5 — о состояниях, ситуациях и событиях; 6 — о задачах и проблемах; 7 — о моделях рассуждений экспертов; 8 — об интегрированной модели рассуждений коллективного интеллекта

и свойств vl^2 ; (3) иерархий ресурсов, действий, свойств vl^3 ; (4) пространственных и производственных структур vl^4 ; (5) состояний, ситуаций и событий vl^5 ; (6) задач и проблем vl^6 ; (7) моделей рассуждений экспертов vl^7 ; (8) интегрированных моделей рассуждений коллективного интеллекта vl^8 . В этом случае у разработчика есть набор средств-компонентов для конструирования из них метаязыка, описывающего решение сложной задачи комбинацией нескольких взаимосвязанных процессов рассуждений на разных языках. При этом в зависимости от требований поставленной задачи отдельные уровни могут отсутствовать. Таким образом, визуальный метаязык может быть представлен выражением

$$mvl = \langle vl^1, vl^2, vl^3, vl^4, vl^5, vl^6, vl^7, vl^8, VLR \rangle, \quad (10)$$

где VLR — множество отношений между элементами языков vl^k , $k \in \mathbb{N}$, $k \in [1, 8]$.

Метаязык визуализируется «слоеным пирогом» (см. рисунок). В его основании — словари понятий и отношений, концептуально-визуальный базис, над которым строится семейство упорядоченных по уровням языков описания.

Как показано на рисунке, на каждом языковом уровне выделяется гетерогенное образно-визуальное ядро базовых для данного уровня знаков VT^k . Визуальное ядро языков высшего уровня включает знаки ядра более низкого уровня $VT \subseteq VT^{k+1}$ и может содержать знаки, сформированные вне ядра на языке более низкого уровня $VT^{k+1} \cap VN^k \neq \emptyset$, $k \in \mathbb{N}$, $k \in [1, 7]$.

Рассмотрим отношения между языками различных уровней на примере их множеств синтаксических правил. В первом слое расположены словари понятий и отношений — концептуально-визуальный базис языка. Язык первого уровня vl^1 использует эвристические правила PRU^{-1} для построения

из P^1 , D^1 и VR^1 знаков производных (составных) отношений $vr^{n1} \in VR^{n1} \subseteq VT^1$:

$$vl^1(P^1, D^1, VR^1, PRU^1) = \{vr^{n1}\}.$$

В языке второго уровня VI^2 эвристики PRU^2 используются, чтобы сформировать графические образы ресурсов $res^2 \in RES^2 \subseteq VT^2$, действий $act^2 \in ACT^2 \subseteq VT^2$ и свойств $pr^2 \in PR^2 \subseteq VT^2$ без учета их иерархичности с помощью отношений определения $VR_1^{n1} \subseteq VR^{n1}$:

$$vl^2(P^1, D^1, VR_1^{n1}, PRU^2) = RES^2 \cup PR^2 \cup ACT^2.$$

На третьем уровне отношениями включения $VR_5^{n1} \subseteq VR^{n1}$ и эвристиками PRU^3 формализованы иерархии ресурсов $res^{n3} \in RES^{n3} \subseteq VT^3$, действий $act^{n3} \in ACT^{n3} \subseteq VT^3$ и свойств $pr^{n3} \in PR^{n3} \subseteq VT^3$:

$$vl^3(P^1, D^1, RES^2, PR^2, ACT^2, VR_5^{n1}, PRU^3) = RES^{n3} \cup PR^{n3} \cup ACT^{n3}.$$

Четвертый уровень на основе знаков предыдущих уровней, временных $VR_3^{n1} \subseteq VR^{n1}$, пространственных $VR_4^{n1} \subseteq VR^{n1}$ и причинно-следственных $VR_6^{n1} \subseteq VR^{n1}$ отношений, а также эвристик PRU^4 формализует пространственные $str_1^4 \in STR_1^4 \subseteq VT^4$, операционально-технологические $str_3^4 \in STR_3^4 \subseteq VT^4$ структуры:

$$vl^4(P^1, D^1, RES^{n3}, PR^{n3}, ACT^{n3}, VR_3^{n1}, VR_4^{n1}, VR_6^{n1}, PRU^4) = STR_1^4 \cup STR_3^4.$$

На пятом уровне эвристиками PRU^5 формализуют знаки-ситуации $sit^5 \in SIT^5 \subseteq VT^5$ и знаки-состояния $st^5 \in ST^5 \subseteq VT^5$:

$$vl^5(STR_1^4, STR_3^4, PRU^5) = SIT^5 \cup ST^5.$$

На шестом уровне на основе знаков предыдущих уровней и эвристик PRU^6 специфицируются знаки однородных $prb^{h6} \in PRB^{h6} \subseteq VT^6$ и неоднородных $prb^{u6} \in PRB^{u6} \subseteq VT^6$ задач:

$$vl^6(P^1, D^1, RES^{n3}, PR^{n3}, ACT^{n3}, VR^{n1}, ST^5, PRU^6) = PRB^{h6} \cup PRB^{u6}.$$

На седьмом уровне эвристиками PRU^7 формируются знаки автономных методов решения задач $met^{a7} \in MET^{a7} \subseteq VT^7$, имитирующих рассуждения отдельно взятого эксперта:

$$vl^7(P^1, D^1, RES^{n3}, PR^{n3}, ACT^{n3}, VR^{n1}, PRU^7) = MET^{a7}.$$

На восьмом уровне на основе знаков предыдущих уровней и эвристик PRU^8 специфицируются знаки интегрированных методов решения задач $met^{u8} \in MET^{u8} \subseteq VT^8$, имитирующих рассуждения коллектива экспертов:

$$vl^8(P^1, D^1, RES^{n3}, PR^{n3}, ACT^{n3}, VR^{n1}, ST^5, SIT^5, PRB^{h6}, PRB^{u6}, MET^{a7}, PRU^8) = MET^{u8}.$$

Такая многослойная модель визуального языка — инструмент сложного описания предметной области на различных уровнях обобщенности, формирующий его из набора более простых моделей.

Другой подход к снижению сложности построения моделей предметной области практических задач — стратификация языка по профессиональным нишам.

4 Модель гетерогенного визуального поля

Гетерогенность визуального поля, проявляющаяся в разнообразии информации, обусловлена отсутствием универсального описания любой предметной области. На практике используются более сотни методов визуального структурирования. Это обусловлено существенными различиями в природе, особенностях и свойствах знаний о различных предметных областях. В работах [13–15] проанализированы наиболее известные методы визуализации, определены критерии классификации и классы методов. Определив релевантность метода классам, а классов — классам задачам, можно разработать стратегию выбора визуальных языков для решения подзадач элементами ГиИС. Практические задачи требуют их комбинирования и установления соответствия элементов разных языков,

насколько это возможно. В результате при моделировании визуально-образных рассуждений в ГиИС формируется гетерогенное визуальное поле, обеспечивающее взаимодействие элементов ГиИС, рассуждающих на разных визуальных языках.

Формально гетерогенное визуальное поле может быть представлено в виде:

$$GVF = \langle MVL, COR^{VL} \rangle. \quad (11)$$

Здесь MVL — множество визуальных метаязыков гетерогенного визуального поля, построенных в соответствии с (10):

$$MVL = \{mvl_1, \dots, mvl_{N_{MVL}}\};$$

COR^{VL} — множество соответствий элементов визуальных языков, входящих в метаязыки $mvl_i \in MVL$:

$$COR^{VL} = \langle G^{VT}, G^{VS}, G^{VA}, G^{VP}, G^{v\tau}, G^{v\sigma}, G^{v\alpha}, G^{v\pi} \rangle,$$

где

$$G_{ij}^{VT} : VT_i \rightarrow VT_j, \quad G_{ij}^{VT} \subseteq G^{VT};$$

$$G_{ij}^{VS} : VS_i \rightarrow VS_j, \quad G_{ij}^{VS} \subseteq G^{VS};$$

$$G_{ij}^{VA} : VA_i \rightarrow VA_j, \quad G_{ij}^{VA} \subseteq G^{VA};$$

$$G_{ij}^{VP} : VP_i \rightarrow VP_j, \quad G_{ij}^{VP} \subseteq G^{VP};$$

$$G_{ij}^{v\tau} : v\tau_i \rightarrow v\tau_j, \quad G_{ij}^{v\tau} \subseteq G^{v\tau};$$

$$G_{ij}^{v\sigma} : v\sigma_i \rightarrow v\sigma_j, \quad G_{ij}^{v\sigma} \subseteq G^{v\sigma};$$

$$G_{ij}^{v\alpha} : v\alpha_i \rightarrow v\alpha_j, \quad G_{ij}^{v\alpha} \subseteq G^{v\alpha};$$

$$G_{ij}^{v\pi} : v\pi_i \rightarrow v\pi_j, \quad G_{ij}^{v\pi} \subseteq G^{v\pi},$$

$$i, j \in [1, N_{MVL}], \quad i \neq j.$$

Предлагаемая неформальная аксиоматическая теория ролевых визуальных моделей (1)–(9) с учетом моделей многослойной модели визуального языка (10) и гетерогенного визуального поля (11) будет положена в основу нового класса функциональных ГиИС, имитирующих работу коллективного интеллекта по поиску решений над гетерогенными модельным и визуальным полями. Сочетание символьных и визуально-образных рассуждений в таких системах обеспечит их релевантность реальным коллективам, принимающим решения в условиях сложных задач.

5 Заключение

Предложена неформальная аксиоматическая теория ролевых визуальных моделей — основа автоматизированного решения сложных задач на

основе визуальных образов, визуального управления. Предложена многослойная модель визуального языка и формализованная модель гетерогенного визуального поля.

Использование указанных моделей дает возможность реализовать ГиИС, способные динамически синтезировать интегрированную модель и метод над гетерогенными модельным и визуальным полями и имитировать сотрудничество, относительность и дополнительность коллективного интеллекта для поиска решений на символьных и визуальных языках. Гибридные интеллектуальные системы такого класса смогут управлять имитационным процессом в зависимости от неопределенности проблемной ситуации: когда область явлений формализована (частично формализована), подключать для поиска решений знания экспертов из гетерогенного модельного поля, а когда есть существенная неопределенность, не снимаемая точным анализом и логико-математическими рассуждениями, привести в действие механизмы визуально-пространственного, образного мышления, имитируя «скачки» в гибридном пространстве состояний функциональной ГиИС, соответствующие мгновенному интуитивному инсайту, озарению, прерывающему логико-математические рассуждения.

Литература

1. *Golin E. J., Reiss S. P.* The specification of visual language syntax // *J. Visual Lang. Comput.*, 1990. Vol. 1. P. 141–157.
2. *Bowman W. J.* Graphic communication. — New York, NY, USA: John Wiley, 1968. 210 p.
3. *Lakin F.* Visual grammars for visual languages // 6th National Conference on Artificial Intelligence Proceedings. — Menlo Park, CA, USA: AAAI Press, 1987. P. 683–688.
4. *Narayanan N. H., Hubscher R.* Visual language theory: Towards a human–computer interaction perspective // *Visual language theory.* — New York, NY, USA: Springer-Verlag, 1998. P. 81–128.
5. *Kremer R.* Visual languages for knowledge representation // 11th Workshop on Knowledge Acquisition, Modeling and Management, 1998. <http://ksi.cpsc.ucalgary.ca/KAW/KAW98/kremer>.
6. *Осипов Г. С.* От ситуационного управления к прикладной семиотике // *Новости искусственного интеллекта*, 2002. № 6(54). С. 3–7.
7. *Fitriani S., Rothkrantz L. J. M.* Two-dimensional visual language grammar. — Delft, The Netherlands: Delft University of Technology, 2008. http://mmi.tudelft.nl/pub/siska/TSD_2DVisLangGrammar.pdf.
8. *Sibbet D.* Visual leaders: New tools for visioning, management, and organization change. — Hoboken, NJ, USA: Wiley, 2013. 229 p.
9. *Тарасов В. Б.* Проблема понимания: настоящее и будущее искусственного интеллекта // *Открытые семантические технологии проектирования интеллектуальных систем: Мат-лы V Междунар. науч.-технич. конф.* — Минск: БГУИР, 2015. С. 25–42.
10. *Колесников А. В., Кириков И. А.* Методология и технология решения сложных задач методами функциональных гибридных интеллектуальных систем. — М.: ИПИ РАН, 2007. 387 с.
11. *Колесников А. В.* Гибридные интеллектуальные системы. Теория и технология разработки. — СПб.: СПбГТУ, 2001. 711 с.
12. *Колесников А. В., Листопад С. В.* Концептуально-визуальные основы виртуальных гетерогенных коллективов, поддерживающих принятие решений // *Гибридные и синергетические интеллектуальные системы: Мат-лы III Всеросс. Поспеловской конф. с междунар. участием.* — Калининград: БФУ им. И. Канта, 2016. С. 8–56.
13. *Mazza R.* Introduction to information visualization. — London: Springer-Verlag, 2009. 139 p.
14. *Lengler R., Eppler M.* A periodic table of visualization methods // *Visual literacy: An e-learning tutorial on visualization for communication, engineering and business.* http://www.visual-literacy.org/periodic-table/periodic_table.html.
15. *Li K., Tiwari A., Alcock J., Bermell-Garcia P.* Categorisation of visualisation methods to support the design of human–computer interaction systems // *Appl. Ergon.*, 2016. Vol. 55. P. 85–107.

Поступила в редакцию 16.10.16

INFORMAL AXIOMATIC THEORY OF THE ROLE VISUAL MODELS

A. V. Kolesnikov^{1,2}, S. V. Listopad², S. B. Rumovskaya², and V. I. Danishevsky¹

¹Immanuel Kant Baltic Federal University, 14 A. Nevskogo Str., Kaliningrad 236041, Russian Federation

²Kaliningrad Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 5 Gostinaya Str, Kaliningrad 236000, Russian Federation

Abstract: The relevance of creation of the informal axiomatic theory of the role visual models is caused by modeling visual-imaginative reasoning in hybrid and synergistic intelligent systems. Most of the research in the field of visual-imaginative reasoning is focused on developing special visual languages to represent certain kinds of data, information, and knowledge. The lack of formal models of the visual languages is the cause of high research and development intensity of special media for handling and processing of visual models. Creation of the informal axiomatic theory of the role visual models is a step to a new class of intelligent systems that are relevant to the real decision-making teams, i. e., hybrid intelligent systems with heterogeneous visual field, imitating cooperation, complementarity, and relativity of collective intelligence, reasoning using the symbolic and visual languages.

Keywords: hybrid intelligent system; heterogeneous visual field; visual language; semiotic system

DOI: 10.14357/19922264160412

Acknowledgments

The research was supported by the Russian Foundation for Basic Research (project No. 16-07-00271a).

References

- Golin, E. J., and S. P. Reiss. 1990. The specification of visual language syntax. *J. Visual Lang. Comput.* 1:141–157.
- Bowman, W. J. 1968. *Graphic communication*. New York, NY: John Wiley. 210 p.
- Lakin, F. 1987. Visual grammars for visual languages. *6th National Conference on Artificial Intelligence Proceedings*. Menlo Park, CA: AAAI Press. 683–688.
- Narayanan, N. H., and R. Hubscher. 1998. Visual language theory: Towards a human–computer interaction perspective. *Visual language theory*. New York, NY: Springer-Verlag. 81–128.
- Kremer, R. 1998. Visual languages for knowledge representation. *11th Workshop on Knowledge Acquisition, Modeling and Management*. Available at: <http://ksi.cpsc.ucalgary.ca/KAW/KAW98/kremer/> (accessed September 5, 2016).
- Osipov, G. S. 2002. Ot situatsionnogo upravleniya k prikladnoy semiotike [From situational management to applied semiotics]. *Novosti iskusstvennogo intellekta* [Artificial Intelligence News] 6(54):3–7.
- Fitriani, S., and L. J. M. Rothkrantz. 2008. *Two-dimensional visual language grammar*. Delft, The Netherlands: Delft University of Technology. Available at: <http://mmi.tudelft.nl/pub/siska/TSD2DVisLangGrammar.pdf> (accessed September 5, 2016).
- Sibbet, D. 2013. *Visual leaders: New tools for visioning, management, and organization change*. Hoboken, NJ: Wiley. 229 p.
- Tarasov, V. B. 2015. Problema ponimaniya: nastoyashchee i budushchee iskusstvennogo intellekta [The problem of understanding: The present and the future of artificial intelligence]. *5th Scientific and Technical Conference (International) “Open Semantic Technologies for Intelligent Systems” Proceedings*. Minsk: BSUIR. 25–42.
- Kolesnikov, A. V., and I. A. Kirikov. 2007. *Metodologiya i tekhnologiya resheniya slozhnykh zadach metodami funktsional’nykh gibridnykh intellektual’nykh sistem* [Methodology and technology of solving complex problems by the methods of functional hybrid intelligent systems]. Moscow: IPI RAN. 387 p.
- Kolesnikov, A. V. 2001. *Gibridnye intellektual’nye sistemy. Teoriya i tekhnologiya razrabotki* [Hybrid intelligent systems: Theory and technology of development]. St. Petersburg: SPbGTU Publ. 711 p.
- Kolesnikov, A. V., and S. V. Listopad. 2016. Kontseptual’no-vizual’nye osnovy virtual’nykh geterogennykh kolektivov, podderzhivayushchikh prinyatie resheniy [Conceptual and visual basics of virtual heterogeneous teams supporting decision-making]. *Gibridnye i sinergeticheskie intellektual’nye sistemy: mat-ly III Vseross. Pospelovskoy konf. s mezhdunar. uchastiem* [3rd All-Russia Pospelov Conference with International Participation “Hybrid and synergistic intelligent systems” Proceedings]. Kaliningrad: IKBFU Publ. 8–56.

13. Mazza, R. 2009. *Introduction to information visualization*. London: Springer-Verlag. 139 p.
14. Lengler, R., and M. Eppler. A periodic table of visualization methods. *Visual literacy: An e-learning tutorial on visualization for communication, engineering and business*. Available at: http://www.visual-literacy.org/periodic-table/periodic_table.html (accessed September 5, 2016).
15. Li, K., A. Tiwari, J. Alcock, and P. Bermell-Garcia. 2016. Categorisation of visualisation methods to support the design of human–computer interaction systems. *Appl. Ergon.* 55:85–107.

Received October 16, 2016

Contributors

Kolesnikov Alexander V. (b. 1948) — Doctor of Science in technology; professor, Department of Telecommunications, Immanuel Kant Baltic Federal University, 14 A. Nevskogo Str., Kaliningrad 236041, Russian Federation; senior scientist, Kaliningrad Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 5 Gostinaya Str, Kaliningrad 236000, Russian Federation, avkolesnikov@yandex.ru

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

Rumovskaya Sophiya B. (b. 1985) — programmer I, Kaliningrad Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 5 Gostinaya Str, Kaliningrad 236000, Russian Federation, sophiyabr@gmail.com

Danishevskii Vladislav I. (b. 1992) — PhD student, Immanuel Kant Baltic Federal University, 14 A. Nevskogo Str., Kaliningrad 236041, Russian Federation; danishevskii.v.i@mail.ru

КЛАССИФИКАЦИЯ ВРЕМЕННЫХ РЯДОВ В ПРОСТРАНСТВЕ ПАРАМЕТРОВ ПОРОЖДАЮЩИХ МОДЕЛЕЙ*

М. Е. Карасиков¹, В. В. Стрижов²

Аннотация: Работа посвящена задаче многоклассовой признаковой классификации временных рядов. Признаковая классификация временных рядов заключается в сопоставлении каждому временному ряду его краткого признакового описания и позволяет решать задачу классификации в пространстве признаков. Исследуются методы построения пространства признаков временных рядов, при этом временной ряд рассматривается как последовательность сегментов, аппроксимируемых некоторой параметрической моделью, параметры которой используются в качестве их признаков описаний. Построенное признаковое описание сегмента временного ряда наследует от модели аппроксимации такое полезное свойство, как инвариантность относительно сдвига. Для решения задачи классификации в качестве признаков описаний временных рядов предлагается использовать распределения параметров аппроксимирующих сегменты моделей, что обобщает базовые методы, использующие непосредственно сами параметры аппроксимирующих моделей. Проведен ряд вычислительных экспериментов на реальных данных, показавших высокое качество решения задачи многоклассовой классификации. Эксперименты показали превосходство предлагаемого метода над базовым и многими распространенными методами классификации временных рядов на всех рассмотренных наборах данных.

Ключевые слова: временные ряды; многоклассовая классификация; сегментация временных рядов; гиперпараметры аппроксимирующей модели; модель авторегрессии; дискретное преобразование Фурье

DOI: 10.14357/19922264160413

1 Введение

Временным рядом x будем называть конечную упорядоченную последовательность чисел

$$x = [x^{(1)}, \dots, x^{(t)}].$$

Временные ряды являются объектом исследования в таких задачах анализа данных, как прогнозирование, обнаружение аномалий, сегментация [1], кластеризация и классификация [1]. Обзор по задачам и методам анализа временных рядов дается в [2]. Последние годы связаны с ростом интереса к данной области, проявляющемся в непрекращающемся предложении новых методов анализа временных рядов — метрик, алгоритмов сегментации, кластеризации и др.

В данной работе рассматривается задача классификации временных рядов, возникающая во многих приложениях (медицинская диагностика по электрокардиограммам [3] и электроэнцефалограммам [4], классификация типов физической активности по данным акселерометра [5], верификация динамических подписей [6] и т. д.).

Формально задача классификации в общем виде ставится следующим образом. Пусть X — множество описаний объектов произвольной природы, Y — конечное множество меток классов. Предполагается существование целевой функции — отображения $y : X \rightarrow Y$, значения которого известны только на объектах обучающей выборки

$$\mathcal{D} = \{(x_1, y_1), \dots, (x_m, y_m)\} \subset X \times Y.$$

Требуется построить классификатор $a : X \rightarrow Y$ — отображение, приближающее целевую функцию y на множестве X . При $|Y| > 2$ задачу классификации будем называть многоклассовой. Задачей классификации временных рядов будем называть задачу классификации, в которой объектами классификации являются временные ряды.

Задание метрики, или функции расстояния [7], на парах временных рядов позволяет применять метрические методы классификации. При удачном выборе метрики классификация может производиться простейшими метрическими алгоритмами классификации, например методом ближайшего соседа [8]. Данный подход к решению задачи классификации временных рядов чрезвычайно распро-

*Работа выполнена при финансовой поддержке РФФИ (проект 16-37-00485).

¹Московский физико-технический институт, Сколковский институт науки и технологий, karasikov@phystech.edu

²Вычислительный центр им. А. А. Дородницына Федерального исследовательского центра «Информатика и управление» Российской академии наук, strizov@ccas.ru

странен в силу того, что позволяет свести исходную задачу классификации временных рядов к задаче выбора метрики.

Второй подход к решению задачи классификации состоит в построении для каждого временного ряда его информативного признакового описания $f : X \rightarrow \mathbb{R}^n$, позволяющего строить точные классификаторы с хорошей обобщающей способностью. Построение информативного пространства признаков исходных объектов множества X , позволяющего добиться заданной точности классификации и значительно упрощающего последующий анализ, является важнейшим этапом решения задачи классификации. Признаки могут задаваться экспертом. Так, в работе [9] предлагается использовать в качестве признаков статистические функции (среднее, отклонения от среднего, коэффициенты эксцесса и др.). Стоит заметить, что при таком подходе к построению пространства признаков часто удается добиться необходимого качества классификации путем выбора соответствующих конкретной задаче признаков (см., например, [10]), а сам выбор признаков становится важной технической задачей. Другой метод построения пространства признаков заключается в задании параметрической регрессионной или аппроксимирующей модели временного ряда. Тогда в качестве признаков временных рядов будут выступать параметры настроенной модели. В работе [11] в качестве признаков предлагается использовать коэффициенты дискретного преобразования Фурье (DFT) и дискретного вейвлет-преобразования (DWT), а в [12, 13] — модели авторегрессии. Таким образом, при данном методе построения признаковых описаний возникает задача выбора аппроксимирующей модели временного ряда.

В работе исследуются методы классификации временных рядов, использующие в качестве их признаковых описаний параметры аппроксимирующих моделей. Приводится сравнение моделей аппроксимации. Из временного ряда могут извлекаться сегменты — его подпоследовательности, для которых признаковые описания строятся так же, как и для исходных временных рядов. Использование подпоследовательностей позволяет обобщить алгоритмы классификации. Так, в работе [1] предлагается алгоритм классификации временных рядов методом голосования их случайных сегментов (непрерывных подпоследовательностей со случайным начальным элементом). В данной работе предлагается алгоритм классификации временных рядов в пространстве параметров распределений признаков их сегментов, который сравнивается с родственным ему алгоритмом голосования сегментов [1]. В разд. 7 приводятся результаты экс-

периментов на реальных данных, показывающие высокое качество и общность предлагаемого алгоритма в сочетании с методом признаковых описаний временных рядов параметрами аппроксимирующих их моделей.

2 Постановка задачи

Поставим задачу многоклассовой классификации временных рядов в общем виде. Пусть (X, ρ) — метрическое пространство временных рядов, Y — множество меток классов, $\mathcal{D} \subset X \times Y$ — конечная обучающая выборка.

Пусть S — процедура сегментации:

$$S(x) \subset 2^{\mathbf{S}(x)}, \quad (1)$$

где $\mathbf{S}(x)$ — множество всех сегментов временного ряда $x \in X$; $f(S(x)) \in \mathbb{R}^n$ — процедура построения признакового описания набора сегментов; b — алгоритм многоклассовой классификации:

$$b : \mathbb{R}^n \rightarrow Y. \quad (2)$$

Рассмотрим семейство $A = \{a : X \rightarrow Y\}$ алгоритмов классификации вида

$$a = b \circ f \circ S. \quad (3)$$

Пусть задана функция потерь $\mathcal{L} : X \times Y \times Y \rightarrow \mathbb{R}$ и функционал качества

$$Q(a, \mathcal{D}) = \frac{1}{|\mathcal{D}|} \sum_{(x,y) \in \mathcal{D}} \mathcal{L}(x, a(x), y). \quad (4)$$

В качестве методов обучения $\mu(\mathcal{D}) \in A$ будем использовать следующие:

$$\mu_{f,S}(\mathcal{D}) = \hat{b} \circ f \circ S,$$

где \hat{b} — минимизатор эмпирического риска:

$$\hat{b} = \arg \min_b Q(b \circ f \circ S, \mathcal{D}).$$

Оптимальный метод обучения определяется по скользящему контролю:

$$\mu^* = \arg \min_{f,S} \widehat{CV}(\mu_{f,S}, \mathcal{D}),$$

где $\widehat{CV}(\mu, \mathcal{D})$ — внешний критерий качества метода обучения μ ; при этом исходная обучающая выборка \mathcal{D} случайно разбивается r раз на обучающую и контрольную ($\mathcal{D} = \mathcal{L}_1 \sqcup \mathcal{T}_1 = \dots = \mathcal{L}_r \sqcup \mathcal{T}_r$),

$$\widehat{CV}(\mu, \mathcal{D}) = \frac{1}{r} \sum_{v=1}^r Q(\mu(\mathcal{L}_v), \mathcal{T}_v), \quad (5)$$

где

$$Q(a, \mathfrak{T}) = \frac{1}{|\mathfrak{T}|} \sum_{(x,y) \in \mathfrak{T}} \mathbf{1}\{a(x) = y\}. \quad (6)$$

Средняя точность (precision) классификации объектов класса $c \in Y$ оценивается функционалом скользящего контроля (5) с модифицированным функционалом качества Q :

$$Q_c(a, \mathfrak{T}) = \frac{|\{(x,y) \in \mathfrak{T} \mid a(x) = y = c\}|}{|\{(x,y) \in \mathfrak{T} \mid y = c\}|}. \quad (7)$$

3 Сегментация временных рядов

Определение 1. Сегментом временного ряда $x = [x^{(1)}, \dots, x^{(t)}]$ будем называть любую его непрерывную подпоследовательность $s = [x^{(i)}]_{i=t_0}^{t_1}$, $1 \leq t_0 \leq t_1 \leq t$.

Определение 2. Под сегментацией будем понимать отображение временных рядов во множество их сегментов (1).

Примеры

1. Тривиальная сегментация

$$S(x) = \{x\}, \forall x \in X. \quad (8)$$

2. Случайное выделение сегментов некоторой длины ℓ [1].
3. Важным является случай квазипериодичности временного ряда, когда сам ряд состоит из похожих в определенном смысле сегментов, называемых периодами:

$$x = \left[\underbrace{x^{(1)}, \dots, x^{(t_1)}}_{s^{(1)}}, \underbrace{x^{(t_1+1)}, \dots, x^{(t_2)}}_{s^{(2)}}, \dots, \dots, \underbrace{x^{(t_{p-1}+1)}, \dots, x^{(t)}}_{s^{(p)}} \right].$$

Тогда в качестве процедуры сегментации можно взять разбиение на периоды:

$$S(x) = \{s^{(1)}, \dots, s^{(p)}\}.$$

4 Аппроксимирующая модель сегмента временного ряда

Поскольку сегмент временного ряда сам является временным рядом, в этом разделе слово сегмент будем опускать.

Определение 3. Параметрической аппроксимирующей моделью временного ряда x будем называть отображение

$$g : \mathbb{R}^n \times X \rightarrow X. \quad (9)$$

В слово «аппроксимирующая» вкладывается тот смысл, что модель должна приближать временной ряд в пространстве (X, ρ) , т. е. для некоторого $\mathbf{w} \in \mathbb{R}^n$

$$g(\mathbf{w}, x) = \hat{x},$$

где

$$\rho(\hat{x}, x) < \varepsilon.$$

При этом естественно взять в качестве признакового описания временного ряда x вектор оптимальных параметров его модели.

Определение 4. Признаковым описанием временного ряда x , порожденным параметрической моделью $g(\mathbf{w}, x)$, назовем вектор оптимальных параметров этой модели:

$$\mathbf{w}_g(x) = \arg \min_{\mathbf{w} \in \mathbb{R}^n} \rho(g(\mathbf{w}, x), x). \quad (10)$$

В качестве аппроксимирующих моделей предлагается использовать следующие.

1. **Модель линейной регрессии.** Пусть задан r -компонентный временной ряд (например, время и три пространственные координаты):

$$x = [x^{(1)}, \dots, x^{(t)}],$$

где

$$x^{(k)} = [x_1^{(k)}, \dots, x_r^{(k)}]^T, \quad k = 1, \dots, t.$$

Рассмотрим модель линейной регрессии одной из компонент временного ряда на остальные компоненты как аппроксимирующую модель:

$$g(\mathbf{w}, x) = [\hat{x}^{(1)}, \dots, \hat{x}^{(t)}],$$

где

$$\hat{x}^{(k)} = [x_1^{(k)}, \dots, x_{r-1}^{(k)}, \hat{x}_r^{(k)}]^T, \quad k = 1, \dots, t,$$

$$\begin{bmatrix} \hat{x}_r^{(1)} \\ \vdots \\ \hat{x}_r^{(t)} \end{bmatrix} = \begin{bmatrix} x_1^{(1)} & \dots & x_{r-1}^{(1)} \\ \vdots & \ddots & \vdots \\ x_1^{(t)} & \dots & x_{r-1}^{(t)} \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_{r-1} \end{bmatrix}.$$

Тогда, выбрав в качестве ρ евклидово расстояние, по определению 4 получим признаковое описание объекта x :

$$\begin{aligned} \mathbf{w}_g(x) &= \arg \min_{\mathbf{w} \in \mathbb{R}^n} \|\mathbf{x}_r - \hat{\mathbf{x}}_r\|_2^2 = \\ &= \arg \min_{\mathbf{w} \in \mathbb{R}^n} \|\mathbf{x}_r - \mathbf{X}\mathbf{w}\|_2^2 = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{x}_r. \quad (11) \end{aligned}$$

2. Модель авторегрессии AR(p).

Задан временной ряд

$$x = [x^{(1)}, \dots, x^{(t)}], \quad x^{(k)} \in \mathbb{R}, \quad k = 1, \dots, t.$$

Выберем в качестве модели аппроксимации авторегрессионную модель порядка p :

$$g(\mathbf{w}, x) = [\hat{x}^{(1)}, \dots, \hat{x}^{(t)}],$$

где

$$\hat{x}^{(k)} = \begin{cases} x^{(k)}, & k = 1, \dots, p; \\ w_0 + \sum_{i=1}^p w_i x^{(k-i)}, & k = p+1, \dots, t. \end{cases}$$

Далее признаковое описание определяется аналогично случаю линейной регрессии (11).

3. Дискретное преобразование Фурье. Задан временной ряд

$$x = [x^{(0)}, \dots, x^{(t-1)}], \quad x^{(k)} \in \mathbb{C}, \quad k = 0, \dots, t-1.$$

Взяв в качестве аппроксимирующей модели обратное преобразование Фурье

$$g(\mathbf{w}, x) = [\hat{x}^{(0)}, \dots, \hat{x}^{(t-1)}],$$

где

$$\hat{x}^{(k)} = \frac{1}{t} \sum_{j=0}^{t-1} (w_{2j} + iw_{2j+1}) e^{(2\pi i/t)kj}, \quad k = 0, \dots, t-1, \quad (12)$$

получим, что признаковым описанием временного ряда x является прямое преобразование:

$$\mathbf{w}_g(x) = [w_0, \dots, w_{2t-1}], \quad (13)$$

где

$$w_{2k} + iw_{2k+1} = \sum_{j=0}^{t-1} x^{(j)} e^{-(2\pi i/t)kj}, \quad k = 0, \dots, t-1.$$

Переписывая (12) в матричном виде, заметим, что, как и в предыдущих случаях, параметры модели \mathbf{w} эквивалентно находятся при помощи линейной регрессии временного ряда на столбцы матрицы Фурье. Выбор лишь некоторых комплексных амплитуд соответствует регрессии временного ряда на соответствующие столбцы матрицы Фурье. Случай дискретного вейвлет-преобразования аналогичен.

Заметим, что в первых двух случаях используются билинейные аппроксимирующие моде-

ли $g(\mathbf{w}, x)$, а в третьем — линейная. Приведенные примеры демонстрируют большую общность построения пространства признаков при помощи моделей типа (9) и решения оптимизационной задачи (10). Вообще говоря, при $|X| \geq 2$ любая процедура построения признаковых описаний $\mathbf{f} : X \rightarrow \mathbb{R}^n$ задается эквивалентно решением оптимизационной задачи (10) при выборе соответствующей пары (g, ρ) .

5. Распределения признаков сегментов

Объединим идеи, изложенные в предыдущих разделах. Согласно аппроксимирующей модели (9) получим для каждого сегмента $s^{(k)} \in S(x) = \{s^{(1)}, \dots, s^{(p)}\}$ временного ряда x его признаковое описание $\mathbf{w}^{(k)} := \mathbf{w}_g(s^{(k)})$, решив оптимизационную задачу (10). Тогда всему набору сегментов $S(x)$ будет соответствовать выборка:

$$\mathbf{F} = (\mathbf{w}^{(1)}, \dots, \mathbf{w}^{(p)}). \quad (14)$$

Примем гипотезу простоты выборки (14).

Гипотеза 1. Выборка $\mathbf{F} = (\mathbf{f}^{(1)}, \dots, \mathbf{f}^{(p)})$ — простая, т. е. случайная, независимая и однородная, где $\mathbf{w}^{(k)} \sim P_0$.

Пусть имеется параметрическое семейство распределений $\{P_\theta\}_{\theta \in \Theta}$. Будем рассматривать вероятностную модель, в которой объект x зависит от случайного параметра θ .

Гипотеза 2. $p(x|\theta, y) = p(x|\theta)$.

Тогда

$$\begin{aligned} p(x, y) &= p(\mathbf{F}, y) = \\ &= \int_{\Theta} p(\mathbf{F}, \theta, y) d\theta = \int_{\Theta} p(\mathbf{F}|\theta) p(\theta, y) d\theta. \end{aligned}$$

При этом распределение $p(\theta, y)$ предлагается оценивать на этапе обучения, где признаковыми описаниями объектов x_i задачи классификации являются оценки параметров θ_i :

$$\hat{\theta}_i = T(x_i) = T(\mathbf{F}_i).$$

Получив оценку $\hat{p}(\theta, y)$, находим оценку плотности $\hat{p}(x, y)$:

$$\hat{p}(x, y) = \int_{\Theta} p(\mathbf{F}|\theta) \hat{p}(\theta, y) d\theta,$$

по которой строится байесовский классификатор.

В алгоритмической постановке задачи классификации получим $\hat{p}(y|\theta) = \delta(a(\theta), y)$ и

$$\hat{p}(\theta, y) = \delta(a(\theta), y)p(\theta).$$

Тогда

$$\begin{aligned} \hat{p}(x, y) &= \int_{\Theta} p(\mathbf{F}|\theta) \hat{p}(\theta, y) d\theta = \\ &= \int_{\Theta} p(\mathbf{F}|\theta) \delta(a(\theta), y) p(\theta) d\theta = \\ &= \int_{a^{-1}(y)} p(\mathbf{F}|\theta) p(\theta) d\theta = \int_{a^{-1}(y)} p(\theta|\mathbf{F}) p(\mathbf{F}) d\theta. \end{aligned}$$

Приближая распределение $p(\theta|\mathbf{F})$ вырожденным $\delta(\theta - T(\mathbf{F}))$, получим

$$\begin{aligned} \hat{p}(y|x) &= \int_{a^{-1}(y)} p(\theta|\mathbf{F}) d\theta = \int_{a^{-1}(y)} \delta(\theta - T(\mathbf{F})) d\theta = \\ &= \delta(a(T(\mathbf{F})), y). \end{aligned}$$

Таким образом, задача классификации временных рядов свелась к задаче классификации оценок параметров распределений семейства $\{P_{\theta}\}_{\theta \in \Theta}$.

В качестве оценок параметров θ предлагается брать оценки максимального правдоподобия:

$$\begin{aligned} \hat{\theta} = T(x) &= \arg \max_{\theta \in \Theta} \mathcal{L}(\theta | x) = \arg \max_{\theta \in \Theta} p(\mathbf{F}|\theta) = \\ &= \arg \max_{\theta \in \Theta} \prod_k p(\mathbf{w}^{(k)} | \theta). \end{aligned}$$

Заметим, что в частном случае тривиальной сегментации (8) и семейства вырожденных распределений оценка $\hat{\theta}$ является исходным признаковым описанием. Таким образом, предложенный подход к построению признакового описания временного ряда

$$\mathbf{f} : x \mapsto \hat{\theta}$$

является достаточно общим и при этом хорошо интерпретируется.

6 Алгоритм классификации

Для завершения построения классификатора временных рядов (3) построим многоклассовый классификатор b (2) по обучающей выборке $\{(\mathbf{f}(x), y) | (x, y) \in \mathcal{D}\}$.

Сведем задачу многоклассовой классификации к задачам бинарной классификации при помощи стратегий One-vs-All и One-vs-One.

В данной работе для решения задач бинарной классификации, где $Y = \{-1, +1\}$, берутся различные модификации SVM (support vector machine).

7 Вычислительный эксперимент

Вычислительный эксперимент проводился на данных для задачи классификации типов физической активности человека.

7.1 Датасет WISDM

Датасет (набор данных) WISDM [5] содержит показания акселерометра для 6 видов человеческой активности. Необработанные данные, представляющие собой последовательность размеченных показаний акселерометра (по тройке чисел на каждый отсчет времени с интервалом в 50 мс), были разбиты на временные ряды длиной по 200 отсчетов (10 с). Распределение полученных временных рядов по классам приведено в табл. 1.

Таблица 1 Распределение временных рядов по классам. Набор данных WISDM

Классы	Число объектов
1. Jogging (бежит)	1624
2. Walking (идет)	2087
3. Upstairs (поднимается)	549
4. Downstairs (спускается)	438
5. Sitting (сидит)	276
6. Standing (стоит)	231

7.1.1 Ручное выделение признаков

Выбор признаков. Каждая компонента временного ряда описывалась ее средним, стандартным отклонением, средним модулем отклонения от среднего, гистограммой с 10 областями равной ширины. Полученные признаки для каждой компоненты объединялись, и к ним добавлялся признак средней величины ускорения. Таким образом, каждый временной ряд описывался 40 признаками.

Классификатор. Задача многоклассовой классификации сводилась к задаче бинарной классификации при помощи подхода One-vs-One. В качестве бинарного классификатора использовался SVM с RBF (radial basis function) ядром и параметрами $C = 8,5$ и $\gamma = 0,12$.

Результаты. На диаграмме рис. 1 демонстрируется качество классификации при усреднении по $r = 50$ случайным разбиениям исходной выборки на тестовую и контрольную в пропорции 7 к 3.

Как видно из табл. 2, классы 2, 3 и 4 недостаточно хорошо отделяются друг от друга.

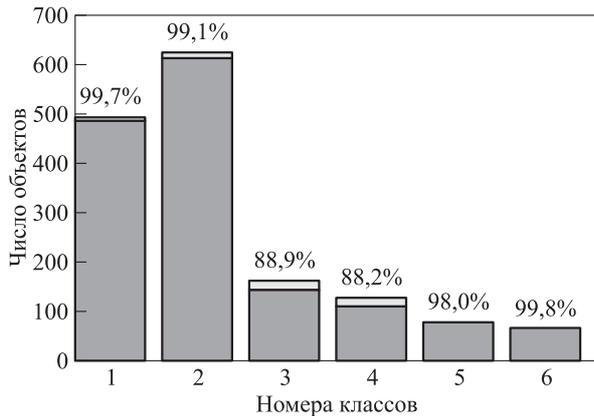


Рис. 1 Набор данных WISDM. Средняя точность 0,9726 — вычисляется по формуле (6). Средние точности классификации для каждого класса вычисляются по формуле (7)

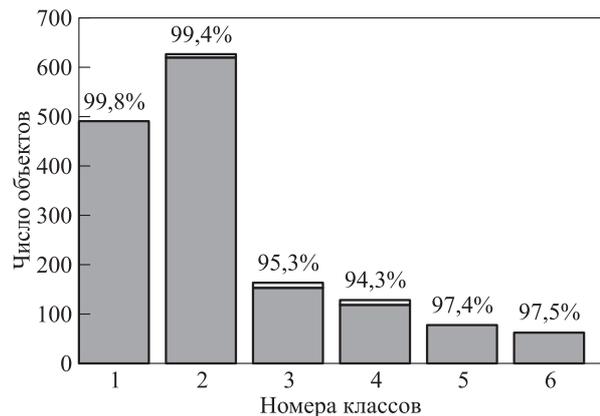


Рис. 2 Точность классификации для параметров модели авторегрессии в качестве признаков описаний

Таблица 2 Усредненная матрица неточностей. Ручное выделение признаков. Набор данных WISDM

Класс объекта	Предсказанный класс					
	1	2	3	4	5	6
1	1,00	0,00	0,00	0,00	0,00	0,00
2	0,00	0,99	0,01	0,00	0,00	0,00
3	0,03	0,04	0,89	0,04	0,00	0,00
4	0,02	0,05	0,05	0,88	0,00	0,00
5	0,01	0,00	0,00	0,00	0,98	0,00
6	0,00	0,00	0,00	0,00	0,00	1,00

Таблица 3 Усредненная матрица неточностей. Признаки, порожденные моделью авторегрессии. Набор данных WISDM

Класс объекта	Предсказанный класс					
	1	2	3	4	5	6
1	1,00	0,00	0,00	0,00	0,00	0,00
2	0,00	0,99	0,00	0,00	0,00	0,00
3	0,01	0,02	0,95	0,02	0,00	0,00
4	0,00	0,02	0,04	0,94	0,00	0,00
5	0,01	0,00	0,00	0,00	0,97	0,01
6	0,01	0,00	0,00	0,00	0,01	0,97

7.1.2 Модель авторегрессии

Признаковое описание. Во втором эксперименте в качестве признаков описаний временных рядов использовались все статистические функции, что брались в первом эксперименте, за исключением гистограммы, вместо которой использовалось 7 коэффициентов модели авторегрессии AR(6) (см. (2)). Таким образом, каждый временной ряд описывался 31 числом. Также проводилась предварительная нормализация признаков.

Классификатор. Задача многоклассовой классификации сводилась к задаче бинарной классификации при помощи подхода One-vs-All. В качестве бинарного классификатора использовалась SVM с RBF-ядром и параметрами $C = 8$ и $\gamma = 0,8$.

Результаты. На диаграмме рис. 2 и в табл. 3 показано качество классификации при усреднении по $r = 50$ случайным разбиениям исходной выборки на тестовую и контрольную в отношении 7 к 3.

Несмотря на неравномерное распределение объектов по классам, использование признакового

описания, порожденного моделью авторегрессии, позволяет значительно повысить качество классификации. Точность построенного классификатора минимальна для 4-го класса — Downstairs — и составляет 94,3%.

7.2 Датасет USC-HAD

Датасет USC-HAD [14] содержит показания акселерометра для 12 типов физической активности человека:

- 1) walk forward (идет вперед);
- 2) walk left (идет влево);
- 3) walk right (идет вправо);
- 4) go upstairs (подъем по лестнице);
- 5) go downstairs (спуск по лестнице);
- 6) run forward (бежит вперед);
- 7) jump up and down (делает прыжок);
- 8) sit and fidget (сидит);
- 9) stand (стоит);

- 10) sleep (спит);
- 11) elevator up (поднимается в лифте);
- 12) elevator down (спускается в лифте).

Выборка содержит примерно по 70 шестикомпонентных временных рядов для каждого класса, а средняя длина временного ряда — 3300. Частота записи измерений сенсора 100 Гц.

7.2.1 Модель авторегрессии и Фурье

Признаковое описание. Исходные временные ряды приводились к частоте 10 Гц при помощи осреднения.

В качестве признаков описаний преобразованных временных рядов брались статистические функции, описанные в п. 7.1.1, за исключением гистограммы. Также для каждой компоненты отдельно и для модуля результирующего ускорения и поворота добавлялось по 11 параметров авторегрессионной модели AR(10) (см. (2)). Затем проводилась нормализация признаков и добавлялись коэффициенты Фурье (13) с индексами 3–12. Таким образом, каждый 6-компонентный временной ряд описывался 128 признаками.

Классификатор. Задача многоклассовой классификации сводилась к задаче бинарной классификации при помощи подхода One-vs-One. В качестве бинарного классификатора использовалась SVM с RBF-ядром и параметрами $C = 10$ и $\gamma = 0,13$.

Результаты. На диаграмме рис. 3 показано качество классификации при усреднении по $r = 500$ случайным разбиениям исходной выборки на тестовую и контрольную в отношении 7 к 3.

Из табл. 4 видно, что использование коэффициентов Фурье значительно повысило качество классификации. Хуже всего класс 8 (sit and fidget) отделяется от класса 9 (stand). Точность классификации для него составляет 92,2%.

7.2.2 Классификация голосованием и классификация в пространстве распределений параметров

Рассмотрим алгоритм классификации в сочетании с процедурой сегментации временных рядов. В качестве процедуры сегментации $S(x)$ (см. (1)) будем использовать выделение сегментов фиксированной длины. Решим задачу классификации для первых 10 классов (за исключением «elevator up» и «elevator down», которые плохо отделяются друг от друга при малой длине сегментов) двумя алгоритмами.

В алгоритме голосования классификатор $b : \mathbb{R}^n \rightarrow Y$ обучается на новой обучающей выборке для сегментов исходных временных рядов

$$\mathcal{D}_S = \{(\mathbf{w}_g(s), y) : (x, y) \in \mathcal{D}, s \in S(x)\}.$$

Далее производится голосование $\hat{y} = \arg \max_y \sum_{s \in S(x)} 1 [b(\mathbf{w}_g(s)) = y]$.

Алгоритм классификации в пространстве гиперпараметров (распределений параметров аппрокси-

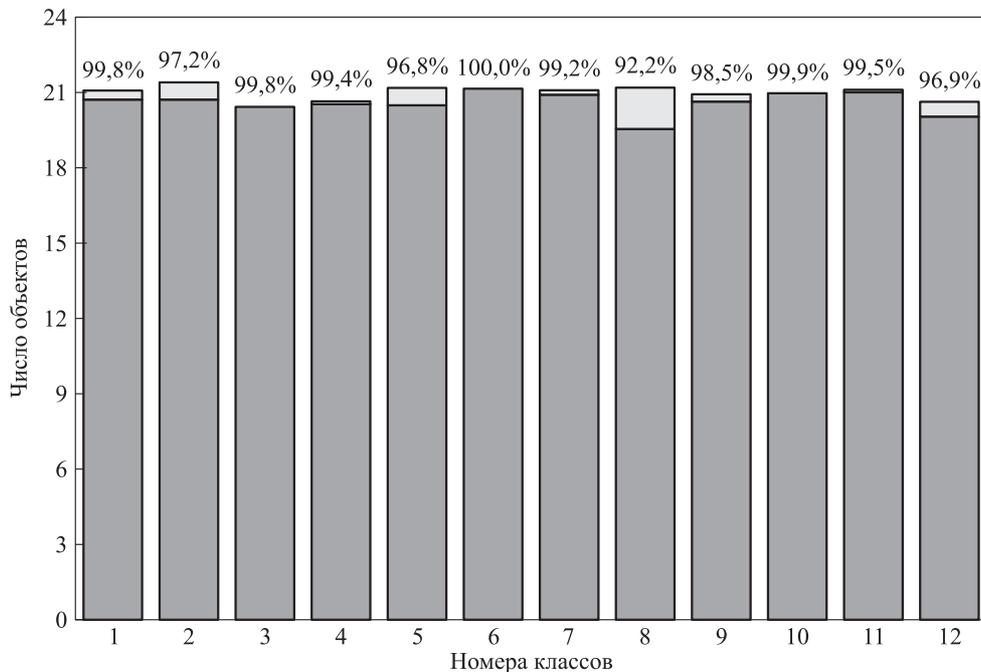


Рис. 3 Точность классификации для параметров модели авторегрессии в качестве признаков описаний

Таблица 4 Усредненная матрица неточностей. Признаки, порожденные моделью авторегрессии. Набор данных USC-HAD

Класс объекта	Предсказанный класс											
	1	2	3	4	5	6	7	8	9	10	11	12
1	0,99	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00
2	0,01	0,97	0,01	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00
3	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
4	0,00	0,00	0,00	0,99	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00
5	0,00	0,00	0,00	0,01	0,97	0,02	0,00	0,00	0,00	0,00	0,00	0,00
6	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	0,00
7	0,00	0,00	0,00	0,00	0,00	0,00	0,99	0,00	0,00	0,00	0,00	0,00
8	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,92	0,08	0,00	0,00	0,00
9	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,99	0,00	0,00	0,00
10	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00
11	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00
12	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,00	0,01	0,97

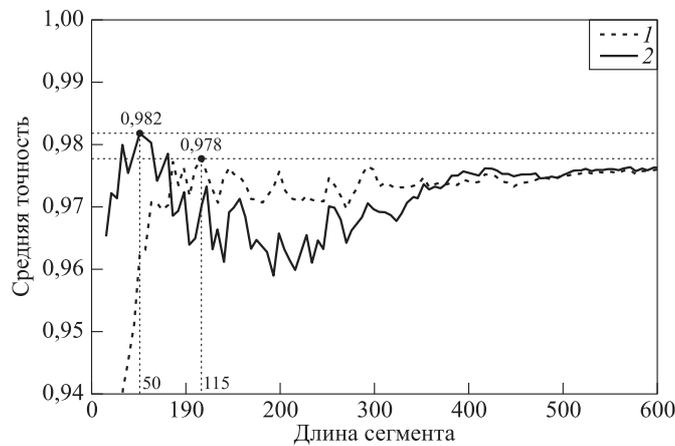


Рис. 4 Зависимость средней точности классификации от длины сегментов: 1 — голосование; 2 — гиперпараметры. Набор данных USC-HAD, первые 10 классов. Точность классификации вычисляется по формуле (6)

мирующих моделей был описан в разд. 5). В эксперименте использовалось семейство нормальных распределений с диагональной ковариационной матрицей.

Задача многоклассовой классификации решалась при помощи подхода One-vs-One бинарными классификаторами SVM с RBF-ядром и параметрами $C = 100$ и $\gamma = 0,017$.

На графике рис. 4 приведены результаты для средней точности решения задачи многоклассовой классификации обоими алгоритмами.

Из графика можно видеть, что оба алгоритма позволяют повысить качество классификации, причем алгоритм классификации в пространстве гиперпараметров при длине сегмента 50 достигает качества 98,2% и показывает результат выше, чем алгоритм голосования.

Объединим результаты из последних двух экспериментов. Будем обучать два классификатора. Пер-

вый классификатор a_1 — One-vs-One SVM с RBF-ядром и параметрами $C = 10$ и $\gamma = 0,13$ — будет разделять классы 11, 12 и первые десять классов для исходных временных рядов. Второй классификатор a_2 — One-vs-One SVM с RBF-ядром и параметрами $C = 100$ и $\gamma = 0,017$ — классификатор в пространстве гиперпараметров, описанный в предыдущем эксперименте.

Итоговый классификатор выглядит следующим образом:

$$a(x) = \begin{cases} a_1(x), & a_1(x) \in \{11, 12\}; \\ a_2(x) & \text{иначе.} \end{cases} \quad (15)$$

Результаты. На диаграмме рис. 5 и в табл. 5 демонстрируется качество классификации построенного классификатора (15) при усреднении по $r = 500$ случайным разбиениям исходной выборки на тестовую и контрольную в отношении 7 к 3.

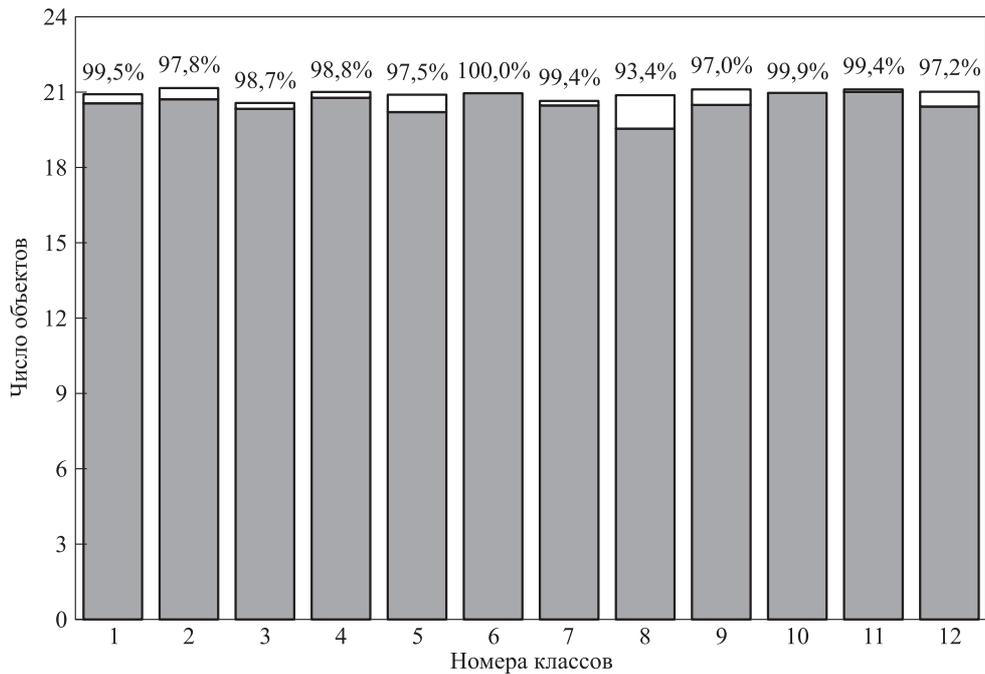


Рис. 5 Точность классификации для гиперпараметров в качестве признаков описаний. Набор данных USC-HAD

Таблица 5 Усредненная матрица неточностей. Признаки — гиперпараметры. Набор данных USC-HAD

Класс объекта	Предсказанный класс											
	1	2	3	4	5	6	7	8	9	10	11	12
1	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
2	0,01	0,98	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
3	0,00	0,00	0,99	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
4	0,00	0,00	0,00	0,99	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00
5	0,01	0,01	0,00	0,00	0,97	0,01	0,00	0,00	0,00	0,00	0,00	0,00
6	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	0,00
7	0,00	0,00	0,00	0,00	0,00	0,00	0,99	0,00	0,00	0,00	0,00	0,00
8	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,93	0,06	0,00	0,00	0,00
9	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,03	0,97	0,00	0,00	0,00
10	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00
11	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,99	0,00
12	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,00	0,01	0,97

8 Заключение

В работе показано, что метод признаков описания временного ряда оптимальными параметрами аппроксимирующих его моделей дает высокое качество решения задачи классификации. Предложенный метод вычислительно эффективен и не требователен к памяти вычислительного устройства.

В работе также предложен алгоритм классификации временных рядов в пространстве распределений параметров моделей, порождающих их сегменты. Он обобщает предыдущий метод клас-

сификации временных рядов и позволяет производить более тонкую настройку алгоритма классификации.

Литература

1. Geurts P., Wehenkel L. Segment and combine approach for non-parametric time-series classification // Knowledge discovery in databases: PKDD 2005. — Berlin—Heidelberg: Springer, 2005. P. 478—485.
2. Esling P., Agon C. Time-series data mining // ACM Comput. Surv., 2012. Vol. 45. No. 1. Article 12. P. 1—34.

3. Basil T., Lakshminarayan C. Automatic classification of heartbeats // 22nd European Signal Processing Conference Proceedings, 2014. P. 1542–1546.
4. Alomari M. H., Samaha A., AlKamha K. Automated classification of l/r hand movement eeg signals using advanced feature extraction and machine learning // Int. J. Adv. Comput. Sci. Appl., 2013. Vol. 4. No. 6. P. 207–212.
5. Kwapisz J. R., Weiss G. M., Moore S. A. Activity recognition using cell phone accelerometers // ACM SigKDD Explorations Newsletter, 2011. Vol. 12. No. 2. P. 74–82.
6. Gruber C., Coduro M., Sick B. Signature verification with dynamic rbf networks and time series motifs // 10th Workshop (International) on Frontiers in Handwriting Recognition. La Baule, 2006. P. 455–460.
7. Ding H., Trajcevski G., Scheuermann P., Wang X., Keogh E. Querying and mining of time series data: Experimental comparison of representations and distance measures // Proc. VLDB Endow, 2008. Vol. 1. No. 2. P. 1542–1552. doi:10.14778/1454159.1454226.
8. Jeong Y. S., Jeong M. K., Omitaomu O. A. Weighted dynamic time warping for time series classification // Pattern Recogn., 2011. Vol. 44. No. 9. P. 2231–2240. doi:10.1016/j.patcog.2010.09.022.
9. Nanopoulos A., Alcock R., Manolopoulos Y. Feature-based classification of time-series data // Int. J. Comput. Res., 2001. Vol. 10. P. 49–61.
10. Wiens J., Horvitz E., Gutttag J. V. Patient risk stratification for hospital-associated c. diff as a time-series classification task // Adv. Neur. Inform. Proc. Syst., 2012. Vol. 25. P. 467–475.
11. Mörchen F. Time series feature extraction for data mining using dwt and dft, 2003. Unpubl.
12. Kini B. V., Sekhar C. C. Large margin mixture of ar models for time series classification // Appl. Soft Comp., 2013. Vol. 13. No. 1. P. 361–371.
13. Кузнецов М. П., Ивкин Н. П. Алгоритм классификации временных рядов акселерометра по комбинированному признаковому описанию // Машинное обучение и анализ данных, 2015. Т. 1. № 11. С. 1471–1483.
14. Zhang M., Sawchuk A. A. USC-HAD: A daily activity dataset for ubiquitous activity recognition using wearable sensors // ACM Conference (International) on Ubiquitous Computing Workshop on Situation, Activity and Goal Awareness. — Pittsburgh, PA, USA, 2012.

Поступила в редакцию 10.05.16

FEATURE-BASED TIME-SERIES CLASSIFICATION

M. E. Karasikov^{1,2} and V. V. Strijov³

¹Moscow Institute of Physics and Technology, 9 Institutskiy Per., Dolgoprudny, Moscow Region 141700, Russian Federation

²Skolkovo Institute of Science and Technology, Skolkovo Innovation Center, Building 3, Moscow 143016, Russian Federation

³A. A. Dorodnicyn Computing Center, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

Abstract: The paper is devoted to the multiclass time series classification problem. The feature-based approach that uses meaningful and concise representations for feature space construction is applied. A time series is considered as a sequence of segments approximated by parametric models, and their parameters are used as time series features. This feature construction method inherits from the approximation model such unique properties as shift invariance. The authors propose an approach to solve the time series classification problem using distributions of parameters of the approximation model. The proposed approach is applied to the human activity classification problem. The computational experiments on real data demonstrate superiority of the proposed algorithm over baseline solutions.

Keywords: time series; multiclass classification; time series segmentation; hyperparameters of approximation model; autoregressive model; discrete Fourier transform

DOI: 10.14357/19922264160413

Acknowledgments

The work was supported by the Russian Foundation for Basic Research (project 16-37-00485).

References

1. Geurts, P., and L. Wehenkel. 2005. Segment and combine approach for non-parametric time-series classification. *Knowledge Discovery in Databases: PKDD 2005*. Berlin–Heidelberg: Springer. 478–485.
2. Esling, P., and C. Agon. 2012. Time-series data mining. *ACM Comput. Surv.* 45(1):12:1–12:34.
3. Basil, T., and C. Lakshminarayan. 2014. Automatic classification of heartbeats. *22nd European Signal Processing Conference Proceedings*. 1542–1546.
4. Alomari, M. H., A. Samaha, and K. AlKamha. 2013. Automated classification of l/r hand movement eeg signals using advanced feature extraction and machine learning. *Int. J. Adv. Comput. Sci. Appl.* 4(6):207–212.
5. Kwapisz, J. R., G. M. Weiss, and S. A. Moore. 2011. Activity recognition using cell phone accelerometers. *ACM SigKDD Explorations Newsletter* 12(2):74–82.
6. Gruber, C., M. Coduro, and B. Sick. 2006. Signature verification with dynamic rbf networks and time series motifs. *10th Workshop (International) on Frontiers in Handwriting Recognition*. La Baule. 455–460.
7. Ding, H., G. Trajcevski, P. Scheuermann, X. Wang, and E. Keogh. 2008. Querying and mining of time series data: Experimental comparison of representations and distance measures. *Proc. VLDB Endow* 1(2):1542–1552. doi: 10.14778/1454159.1454226.
8. Jeong, Y.S., M.K. Jeong, and O.A. Omitaomu. 2011. Weighted dynamic time warping for time series classification. *Pattern Recogn.* 44(9):2231–2240. doi: 10.1016/j.patcog.2010.09.022.
9. Nanopoulos, A., R. Alcock, and Y. Manolopoulos. 2001. Feature-based classification of time-series data. *Int. J. Comput. Res.* 10:49–61.
10. Wiens, J., E. Horvitz, and J. V. Guttag. 2012. Patient risk stratification for hospital-associated c. diff as a time-series classification task. *Adv. Neur. Inform. Proc. Syst.* 25:467–475.
11. Mörchen, F. 2003. Time series feature extraction for data mining using dwt and dft. Unpubl.
12. Kini, B. V., and C. C. Sekhar. 2013. Large margin mixture of ar models for time series classification. *Appl. Soft Comp.* 13(1):361–371.
13. Kuznetsov, M. P., and N. P. Ivkin. 2015. Algoritm klasifikatsii vremennykh ryadov akselerometra po kombinirovannomu priznakovomu opisaniyu [Time series classification algorithm using combined feature description]. *Mashinnoe obuchenie i analiz dannykh* [Machine Learning and Data Analysis] 1(11):1471–1483.
14. Zhang, M., and A. A. Sawchuk. 2012. USC-HAD: A daily activity dataset for ubiquitous activity recognition using wearable sensors. *ACM Conference (International) on Ubiquitous Computing Workshop on Situation, Activity and Goal Awareness*. Pittsburgh, PA.

Received May 10, 2016

Contributors

Karasikov Mikhail E. (b. 1992) — student, Moscow Institute of Physics and Technology, 9 Institutskiy Per., Dolgoprudny, Moscow Region 141700, Russian Federation; student, Skolkovo Institute of Science and Technology, Skolkovo Innovation Center, Building 3, Moscow 143016, Russian Federation; karasikov@phystech.edu

Strijov Vadim V. (b. 1967) — Doctor of Science in physics and mathematics, leading scientist, A. A. Dorodnicyn Computing Center, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; strijov@ccas.ru

БАЗА ДАННЫХ БЕЗЛИЧНЫХ ГЛАГОЛЬНЫХ КОНСТРУКЦИЙ РУССКОГО ЯЗЫКА*

Анна А. Зализняк¹, М. Г. Кружков²

Аннотация: Предлагается описание базы данных (БД) безличных глагольных конструкций (БГК), разработанной в целях информационной поддержки лингвистического исследования БГК русского языка в зеркале их перевода на иностранные языки. Показано, каким образом концепция построения надкорпусных БД (НБД) адаптирована под создание данного информационно-лингвистического ресурса. Переводные соответствия в БД БГК представлены в виде упорядоченных пар, элементами которых являются формальные описания соответствующих друг другу лексико-грамматических форм (ЛГФ) на языке оригинала и языке перевода. Также описана методика построения переводных соответствий в БД и рассмотрены проблемы, связанные с процедурой поиска безличных конструкций в электронных лингвистических корпусах, а также некоторые варианты их решения. Использование БД БГК и других НБД значительно расширяет возможности лингвистов, использующих корпусные методы как для моноязычного, так и для контрастивного анализа исследуемых языковых единиц, в том числе благодаря поисковым и статистическим функциям, интегрированным в эти БД.

Ключевые слова: компьютерная лингвистика; контрастивная лингвистика; информационные технологии; электронные корпуса текстов; надкорпусные базы данных; русский язык; безличные конструкции

DOI: 10.14357/19922264160414

1 Введение

База данных БГК русского языка и их переводных эквивалентов была создана в ходе реализации проекта «Контрастивное корпусное исследование глагольных конструкций русского языка: семантика, грамматика, идиоматика» с целью решения следующих задач:

- совершенствование метода унитарного контрастивного анализа, включающего построение моно- и полиэквиваленций и их последующий лингвистический и статистический анализ;
- разработка метода идентификации безличных конструкций;
- построение типологии конструкций русского языка, выражающих значение безличности.

Категория безличности является традиционным объектом русской грамматической теории; наиболее полно, последовательно и на корпусном материале она описана в статье [1]; тем не менее здесь еще остается ряд нерешенных вопросов. В рамках данного проекта рассматриваются только глагольные конструкции, составляющие лишь незначительную

часть всего многообразия русских безличных конструкций, представляющих собой, как известно, выдающуюся особенность русской грамматики [2, с. 413–430]. В данном исследовании принято расширенное понимание категории безличности: к классу безличных конструкций относятся также конструкции, которые в русской грамматике традиционно обозначаются как «неопределенно-личные» (с глаголом в форме 3 л. мн. ч.) и «обобщенно-личные» (с глаголом в форме 2 л. ед. ч., реже — 2 л. мн. ч.) (см. об этих конструкциях, в частности, [3]). Такое понимание является концептуальной основой для проектирования как поисковой подсистемы, так и в целом для создания БД русских БГК. Ее формирование представляет собой важный этап в развитии концепции НБД, а также на пути создания грамматики конструкций русского языка, опирающейся на корпусные данные.

В данном исследовании применяется метод *унитарного корпусного анализа*, предложенный и разработанный в ходе выполнения ряда проектов, суть которого состоит в том, что анализируемые языковые единицы рассматриваются «в зеркале перевода»: иноязычный текстовый эквивалент анализируемой единицы русского языка рассматривается как источник сведений о его семантике,

*Работа выполнена при поддержке РФФИ (проект 15-04-00507).

¹Институт языкознания Российской академии наук; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, anna.zalizniak@gmail.com

²Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, magnit75@yandex.ru

в том числе служит средством выявления имплицитных семантических компонентов. Таким образом, иноязычный текст в обоих направлениях перевода служит не объектом, а лишь инструментом анализа [4–6].

База данных БГК русского языка основана на концепции построения *надкорпусных баз данных*, описанной в работах [4–11]¹. В рамках этой концепции уже было создано три лингвистических БД: БД личных глагольных конструкций русского языка [7–9], БД лингвоспецифичных единиц русского языка [6, 10] и БД коннекторов русского языка [11]. В следующем разделе будет описана структура НБД и то, как она применяется для поиска и описания БГК русского языка и их переводных соответствий в БД БГК.

2 Структура надкорпусных баз данных и ее приложение к базе данных безличных глагольных конструкций

Надкорпусные базы данных используются для хранения информации о следующих объектах, выявляемых в параллельных корпусах²:

- языковые единицы некоторого языка, являющиеся предметом рассмотрения в рамках конкретного лингвистического исследования;
- соответствующие данным единицам фрагменты текста перевода на один или несколько языков;
- переводные соответствия, представляющие собой упорядоченные пары (= кортежи), объединяющие два вышеупомянутых объекта (на первой позиции — единица языка оригинала, на второй позиции — единица языка перевода).

В БД БГК эти три объекта выглядят следующим образом:

- (1) БГК русского языка;
- (2) соответствующие им конструкции в других языках (французском, немецком);
- (3) переводные соответствия, т. е. упорядоченные пары, объединяющие вышеупомянутые объекты.

Для этих объектов в БД БГК используется следующая терминология. Исследуемые конструкции русского языка, а также соответствующие им конструкции в других языках называются *лексико-грамматическими формами*³. Лексико-грамматическая форма перевода, соответствующая некоторой ЛГФ оригинала, называется ее *функционально эквивалентным фрагментом* (сокращенно ФЭФ⁴). Переводное соответствие, представляющее собой упорядоченную пару ⟨ЛГФ, ФЭФ⟩, называется *моноэквивалентией* (сокращенно МЭ)⁵. Такие переводные соответствия будем называть *прямыми*. Прямые соответствия позволяют описывать так называемые *модели перевода*.

В БД БГК также могут сохраняться *обратные* переводные соответствия, сформированные на основе переводов иноязычных текстов на русский язык. В этом случае глагольная конструкция русского языка, выявленная в переводе, находится на второй позиции пары в МЭ, а на первой — тот фрагмент иноязычного текста, который вызвал появление данной русской конструкции в переводе, т. е. послужил *стимулом перевода*⁶.

Нужно отметить, что в БД БГК и других НБД между множеством конструкций русского языка и множеством их ФЭФ в языках перевода существует важное различие. Конструкции русского языка являются объектом нашего исследования, поэтому класс таких конструкций изначально является закрытым.

Множество конструкций русского языка формируется на основе заранее заданных типов (например, в БД БГК в рассмотрение включаются лишь БГК). Что же касается их иноязычных соответствий, то они образуют открытый класс, поскольку заранее неизвестно, какие переводные эквиваленты исследуемых конструкций русского языка будут выявлены в процессе создания БД, и в этом и состоит одна из задач настоящего исследования: работа с параллельными текстами позволяет обнаружить соответствия, о которых до его начала исследователи могли не подозревать [15, с. 221; 16]. Соответственно, множество типов ЛГФ для языков перевода создается и расширяется в процессе построения МЭ по мере выявления новых видов переводных соответствий. При этом нередки случаи, когда в переводе для анализируемой русской конструкции вообще не находится никакого соответ-

¹ Термин «надкорпусная база данных» введен в работе [12].

² В настоящее время во всех НБД используются тексты параллельных подкорпусов Национального корпуса русского языка (НКРЯ, <http://www.ruscorpora.ru/>).

³ Термин был введен в [12].

⁴ Термин введен Д. О. Добровольским [13, 14].

⁵ Понятие *моноэквивалентии* было введено и определено в [7].

⁶ Подробнее о *моделях* и *стимулах* перевода см. [8, с. 102].

ствия или соответствие не может быть однозначно установлено¹.

Несмотря на вышеупомянутое различие, концепция НБД предполагает единую структуру описания ЛГФ на русском и других языках, в которой схемы описания конструкций на разных языках различаются только наборами используемых признаков.

Согласно концепции построения НБД исследователи не используют никакую заранее созданную классификацию рассматриваемых языковых единиц (в случае БД БГК — безличных глагольных форм), а ограничиваются системой аннотирования ЛГФ, предусматривающей присвоение каждой ЛГФ набора признаков из некоторого списка. Такой способ формирования БД БГК позволяет максимально ускорить процесс обработки параллельных текстов, отложив решение более сложных классификационных вопросов до этапа анализа исследователями уже размеченного массива данных.

Для разметки ЛГФ каждого языка (русский, французский, немецкий) используются свои наборы признаков, которые составляются исследователями перед началом работы и могут пополняться и видоизменяться в ходе строительства МЭ. Это в большей степени касается признаков ЛГФ иноязычных текстов, поскольку, как было сказано, в ходе работы прямые МЭ выявляют новые модели, а обратные — новые стимулы перевода. При этом НБД построена таким образом, что для добавления новых и модификации старых признаков не требуется корректировать структуру БД — идентификаторы и описания соответствующих признаков хранятся в таблицах реляционной БД и для внесения изменений достаточно добавить в соответствующую таблицу новые строки или изменить содержание уже имеющихся строк.

Для каждого языка признаки, использующиеся для аннотации ЛГФ, делятся на две группы: *базовые виды* и *дополнительные признаки*. Предполагается, что все множество исследуемых конструкций русского языка принадлежит к некоторому классу (в случае с БД БГК это безличные глагольные формы), который можно разбить на некоторое число непересекающихся подклассов, при этом принадлежность конструкции к определенному подклассу фиксируется с помощью признака, который называется *базовым видом*. Таким образом, каждая ЛГФ всегда должна иметь один и только один базовый вид.

Остальные признаки, использующиеся для аннотации ЛГФ, называются *дополнительными при-*

знаками. Каждой ЛГФ может быть присвоено ноль, один или более дополнительных признаков. Если в НБД фиксируется большое число дополнительных признаков, то они для удобства группируются в кластеры. При этом по ходу наполнения БД БГК базовые виды и дополнительные признаки могут добавляться, удаляться, объединяться, перемещаться из одного кластера в другой, а также могут создаваться новые кластеры дополнительных признаков (при этом иногда может возникать необходимость коррекции разметки для уже построенных ЛГФ).

Надо отметить, что выделение некоторой группы признаков в качестве набора базовых видов для заданного класса конструкций является в значительной степени условным. Этот выбор может определяться не концептуальными, а утилитарными соображениями, так как он ставит своей целью упрощение и ускорение процесса построения МЭ. Например, в БД БГК в качестве набора базовых видов русского языка можно было бы выделить некоторый набор типов безличных конструкций в соответствии с какой-либо существующей классификацией (см., например, [1]). Однако недостаток этого подхода состоит в том, что определение типа безличной конструкции часто является нетривиальной задачей. Более того, построение типологии безличных конструкций — это, наоборот, одна из задач, которая решается при помощи БД БГК. Поэтому в БД БГК в качестве базовых видов берутся лексические единицы — чаще всего это глагол в форме инфинитива, а такие свойства, как время, вид и лицо глагола, фигурируют в качестве дополнительных признаков; иногда в качестве базового вида используется глагол в определенной форме (например, в качестве самостоятельных ЛГФ рассматриваются единицы *кажется*, *может быть*); в качестве самостоятельных ЛГФ рассматриваются также фразеологические единицы с глагольной вершиной (например, *ничего не поделаешь*, *откуда ни возьмись* и т. п.). В других НБД базовые виды могут определяться грамматическими признаками².

3 Методика построения моноэквивалентий в базе данных безличных глагольных конструкций

Построение МЭ в БД БГК осуществляют пользователи, свободно владеющие обоими языками

¹ В таких случаях ФЭФ перевода получает, соответственно, пометы ZERO и UNKNOWN.

² Например, в БД личных глагольных конструкций базовые виды формируются комбинацией признаков «время», «вид» и «наклонение» [12–14].

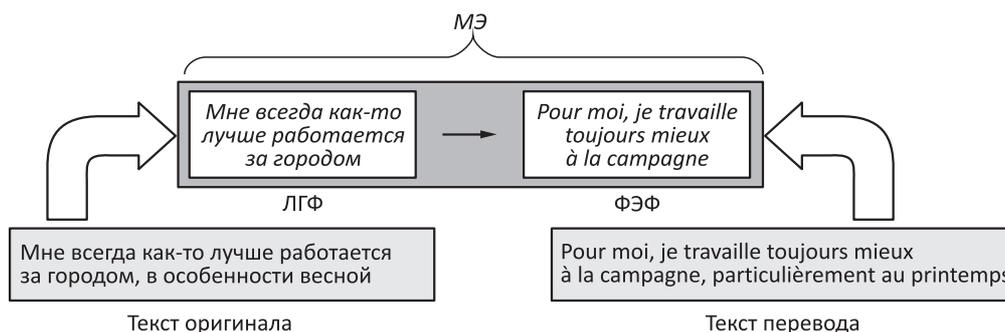


Рис. 1 Построение «скелета» МЭ (пример)

соответствующей языковой пары. Пользователей этой категории будем называть *строителями* МЭ. Система доступа к БД БГК является многопользовательской, благодаря чему построением МЭ в БД БГК могут одновременно заниматься несколько строителей.

Войдя в систему, строитель МЭ в соответствии с поставленной перед ним задачей задает нужное направление перевода (например, переводы с русского на французский), если нужно, выбирает для данного направления перевода конкретные произведения и их переводы, для которых ему предстоит строить МЭ. Затем строитель задает *первичный запрос* (см. разд. 4) на поиск безличных конструкций определенного типа в выбранной части параллельного корпуса. В результате выполнения первичного запроса система выдает строителю список выровненных пар фрагментов, в которых могут находиться искомые конструкции (обычно запрос специфицирует признаки одного или нескольких слов, входящих в текст русского фрагмента пары). Из-за имеющейся в русском языке лексической и грамматической омонимии, а также по ряду других причин (см. ниже) не все выданные по запросу пары в действительности содержат конструкции искомого вида. Строитель последовательно просматривает каждую из найденных пар, и если она действительно содержит искомую конструкцию, то для нее строится МЭ.

Сначала строитель создает «скелет» МЭ. Он выделяет в русском фрагменте данной пары *минимальный контекст* ЛГФ — совокупность элементов фразы, которая позволяет на базовом уровне интерпретировать употребление анализируемой русской ЛГФ в данной фразе. Затем он находит и выделяет в переводном фрагменте той же пары минимальный контекст ФЭФ для данной русской ЛГФ. Из выделенных слов в русском и переводном фрагментах создаются заготовки для русской ЛГФ и ее ФЭФ (переводной ЛГФ). Затем заготовка русской ЛГФ объединяется в пару с заготовкой ФЭФ перевода,

в результате чего создается неразмеченный «скелет» МЭ. Пример построения «скелета» МЭ схематично представлен на рис. 1.

Далее строитель размечает полученный «скелет» МЭ. В минимальных контекстах русской ЛГФ и переводного ФЭФ он выделяет так называемые «главные слова», т. е. слова, входящие в конструкцию: для русской ЛГФ это безличная глагольная форма с наиболее тесно связанными с ней словами (= элементами конструкции), а для ФЭФ перевода — функциональный эквивалент этой конструкции. Для примера из рис. 1 в русской части это будут слова *работается*, а также *мне*, *лучше*; во французской — *pour moi, je travaille, mieux*.

После этого русской ЛГФ и переводному ФЭФ присваиваются значения параметров «базовый вид» и «дополнительные признаки» (рис. 2). Если в списке уже имеющихся базовых видов отсутствует нужная единица, она добавляется в список; такая необходимость возникает, в особенности для ФЭФ перевода, поскольку, как упоминалось выше, по мере обработки параллельного корпуса могут выявляться новые модели и стимулы перевода.

Наконец, строитель может присвоить определенные признаки самой МЭ — из отдельного списка признаков, которые призваны фиксировать различные особенности переводного соответствия в целом и не могут быть отнесены по отдельности ни к оригиналу, ни к переводу. Одним из таких признаков является признак смены подлежащего при переводе (SubjCh). Пример размеченной МЭ схематично представлен на рис. 2.

Построенные и размеченные МЭ на следующем этапе проверяются экспертами. Эксперт может откорректировать состав слов, входящих в контекст ЛГФ и ФЭФ, и их «главные слова», изменить представленные признаки, дать оценку качества построенной МЭ, а также оставить текстовое замечание, что помогает поддерживать обратную связь между экспертами и строителями.

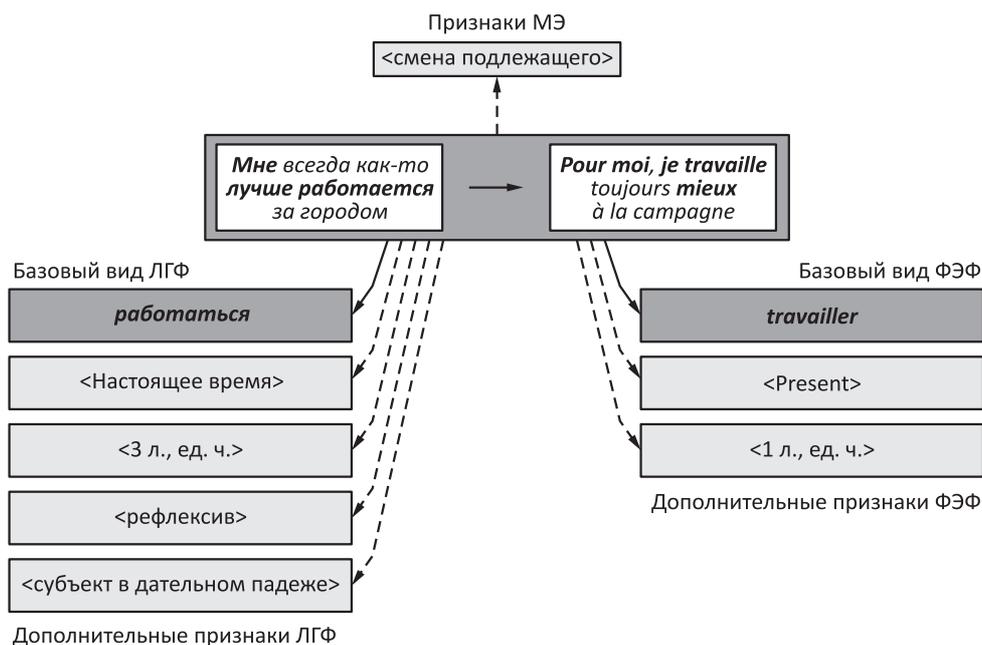


Рис. 2 Размеченная МЭ (пример). Выделены главные слова, указаны базовый вид и некоторые дополнительные признаки для ЛГФ и ФЭФ, указан признак МЭ

4 Построение первичных запросов

Чтобы облегчить поиск безличных конструкций, разработчики НБД создают *первичные поисковые запросы*, предоставляющие в распоряжение строителей множество пар предложений, в которых потенциально могут присутствовать БГК. Далее строители вручную выбирают те пары, где (в русской части) действительно имеется искомая конструкция, и осуществляют построение и аннотацию МЭ.

Первичные запросы базируются на морфологической разметке слов в параллельном подкорпусе НКРЯ (<http://www.ruscorpora.ru/corpora-morph.html>). В силу омонимичности многих русских словоформ в морфологической разметке НКРЯ допускается указание у одной и той же словоформы нескольких вариантов морфологического разбора, что приводит к появлению шума в результате выполнения первичных запросов.

Как уже говорилось, в данном исследовании принято расширенное понимание категории безличности — в том отношении, что сюда включаются также неопределенно-личные и обобщенно-

личные конструкции; свойством, объединяющим все типы безличных конструкций, считается отсутствие подлежащего, т.е. именной группы в номинативе. Однако при анализе предложения при помощи формальных процедур достаточно трудно определить отсутствие в нем подлежащего¹. С другой стороны, как было сказано, в рамках данного проекта рассмотрение ограничивается лишь *глагольными* конструкциями; при этом из рассмотрения на данном этапе исключаются конструкции с глаголами *быть, бывать, стать, становиться*, с предикативом (*сложно, вероятно, стыдно, холодно, надо, пора* и т. п.) и с инфинитивом (*тебе ходить, нечего сказать* и т. п.).

В частности, в БД БГК первичные запросы составлялись для поиска безличных форм, включающих глагол в форме:

- наст./буд. вр. 2-го л. ед. ч. (*сидишь тут целый день как дурак; тут за день так накувыркаешься*);
- наст./буд. вр. 3-го л. ед. ч. (*приходится согласиться; как вам не надоест*);
- наст. вр. 3-го л. мн. ч. (*с друзьями так не поступают*);

¹Ср. понятие «синтаксического нуля» в позиции подлежащего в [17]. Однако, как известно, отсутствие именной группы в номинативе — достаточно сложно идентифицируемый факт как с практической, так и с теоретической точки зрения (см. об этом, в частности, [18, 19]). Особую проблему составляет трактовка случаев отсутствия личного местоимения. Иначе говоря, современное описание грамматики русского языка не содержит набора признаков и критериев для формализованной идентификации отсутствия именной группы в номинативе.

- прош. вр. ср. р. ед. ч. (*оказалось, что он не виноват*);
- прош. вр. ср. р. мн. ч. (*к вам пришли*).

При этом в искомым конструкциях не должно быть подлежащего в именительном падеже, выраженного именной группой или местоимением, которое может согласоваться с данным глаголом.

Помимо уже упомянутой проблемы с шумом (когда в выдачу попадают слова, которые системой ошибочно принимаются за глаголы, например *опускал ноги с постели на пол*), при поиске безличных конструкций в БД БГК нередко могут возникать потери: безличные конструкции, присутствующие в корпусе, не попадают в выдачу. Это происходит, когда из результатов первичных запросов исключаются фрагменты, где фигурирует элемент, который может ошибочно интерпретироваться как подлежащее, согласующееся с искомым глаголом.

Вот несколько причин, которые могут обуславливать такую ошибочную интерпретацию слов в качестве подлежащего:

- некоторые слова, омонимичные существительным в именительном падеже, могут фигурировать в тексте в иной функции, например: *раз, уж, том, знать, чай, жила, стать, мол, берет, надел, постой* и т. д.:

Помилуй, чего тебе еще? от тебя и так уж несет розовой помадой. . .

В мыслях недостаточно последовательности, и, когда я излагаю их на бумаге, мне всякий раз кажется, что я утерял чутье к их органической связи.

Главным образом, я потому не поехал за границу, что вестей туда из России доходит мало, а знать хочется;

- многие местоимения (*то, что, что-то, все, всё* и т. д.), могут выступать как в роли подлежащего, так и в других функциях (союз, дискурсивное слово¹):

На это он заметил, что я еще слишком молода, что у меня еще в голове бродит.

Его всё тянет в ту сторону, где только и знают, что гуляют.

А то выходит по твоему рассказу, что он действительно родился!..

Ей всё хочется, чтобы все считали, что она покровительствует;

- найденное в непосредственном окружении глагола слово в именительном падеже может на самом деле относиться к другому глаголу:

Ну, будет другой редактор и даже, может быть, еще красноречивее прежнего.

Если как следует провентилировать этот вопрос, выходит, что я, в сущности, даже и не знал-то как следует покойника.

Фельдмаршал мой, кажется, говорит дело.

Здесь все дело, кажется, совершенно очевидно. . .

В совокупности эти и некоторые другие причины (например, возможность различного порядка следования подлежащего и сказуемого в русском предложении) могут приводить к заметным потерям, поэтому при построении первичных запросов разработчикам приходится искать способы их уменьшить. В частности:

- составляются списки «квазисуществительных», которые запрос не должен по умолчанию воспринимать как существительные (*раз, уж, том, знать* и т. д.);
- составляются списки местоимений, которые, в отличие от остальных местоимений, почти всегда должны интерпретироваться как потенциальные подлежащие (*который, которое, он, она, это* и т. д.);
- область поиска потенциального подлежащего ограничивается ближайшими к глаголу знаками препинания и т. д.

Поскольку, как было отмечено выше, не существует набора признаков и критериев для формализованной идентификации отсутствия именной группы в номинативе, при построении первичных запросов на поиск безличных форм разработчикам приходится искать приемлемый компромисс между уровнем шума и потерь, так как, если речь не идет о сплошном аннотировании текстов, уменьшение потерь почти всегда ведет к увеличению шума.

5 Поиск и статистика

После завершения аннотирования ЛГФ и МЭ в БД БГК становится достаточно легко находить в массиве данных те МЭ, которые удовлетворяют заданным критериям. Поисковая система БД БГК позволяет указывать при поиске МЭ следующие признаки:

- базовый вид ЛГФ оригинала;
- базовый вид ФЭФ перевода;
- дополнительные признаки ЛГФ;

¹ В некоторых случаях даже строителям МЭ бывает трудно без расширенного контекста определить, какую роль играет такое слово в конкретном примере, например: *Но теперь его вдруг что-то потянуло к людям.*

Таблица 1 Фрагмент результатов выполнения поискового запроса. Показано 4 из 22 найденных МЭ

Контекст ЛГФ оригинала	Вид и дополнительные признаки ЛГФ оригинала	Контекст ЛГФ перевода (ФЭФ)	Вид и дополнительные признаки ЛГФ перевода (ФЭФ)
Сестра теперь, впрочем, кажется , обеспечена. . .	кажется <Impers> <V-IPF> <Pres> <3sg> <Refl> <Parenth>	Ma soeur semble d'ailleurs désormais à l'abri du besoin. . .	sembler <Pr> <3sg>
кажется , доктор теперь уже лишнее	кажется <Impers> <V-IPF> <Pres> <3sg> <Refl> <Parenth>	le médecin semblait déjà inutile	sembler <Imparf> <3sg> <[V.] + {Adj.}>
Кажется , и печальми и радостями он управлял, как движением рук	кажется <Impers> <V-IPF> <Pres> <3sg> <Refl> <Parenth>	Il semble commander à ses joies et à ses tristesses comme il dirige les mouvements de ses bras	sembler <Pr> <3sg> <[V.] + {Inf.}>
все, казалось , лежит в торжественном покое.	кажется <Impers> <V-IPF> <Past> <Sg> <Refl> <Parenth>	Elle semblait silencieuse, solennelle	sembler <Imparf> <3sg> <[V.] + {Adj.}>

- дополнительные признаки ФЭФ;
- признаки МЭ;
- названия текстов, авторы текстов и переводов и т. д.

Все указанные признаки могут задаваться (или исключаться из результатов поиска) одновременно, при этом возвращаемый набор МЭ будет удовлетворять всем указанным требованиям. Например, в БД БГК пользователь может задать для поиска следующий набор признаков: (1) базовый вид ЛГФ оригинала — «*кажется*»; (2) дополнительный признак ЛГФ оригинала — <Parenth> (вводное слово); (3) базовый вид ФЭФ перевода — «*sembler*»; (4) признак МЭ — <SubjCh> (смена подлежащего). Такой запрос возвращает 22 результата (из общего массива в 2100 МЭ), 4 из которых показаны в табл. 1.

Данные БД БГК также можно использовать для получения статистики по моделям и стимулам перевода для различных анализируемых глагольных

конструкций на основе всего массива данных, содержащихся в БД БГК (или же на основе определенного подмножества этого массива). Разумеется, делать выводы на основе такой статистики следует лишь после лингвистической экспертизы представительного массива данных, поскольку результаты могут зависеть от состава анализируемого корпуса, выбранной исследователями схемы аннотации и т. д. Поэтому для пользователей БД предусмотрена возможность верифицировать полученную статистику: перейдя от количественных результатов непосредственно к МЭ, на основе которых они были сгенерированы, пользователь может оценить на качественном уровне зависимость значения указанных параметров от контекста, а также от жанра текста и стиля конкретного автора.

В качестве примера приводится таблица частотностей переводных соответствий (базовых видов французских ФЭФ) для русской ЛГФ «*хотеться*» (табл. 2). В БД БГК численные результаты (столбец

Таблица 2 Количественная статистика по вариантам перевода безличных конструкций с ЛГФ *хотеться*

хотеться 73		
ФЭФ французского языка	Количество МЭ	%
avoir envie	32	43,84
vouloir	24	32,88
selon ses desirs	2	2,74
UNKNOWN	2	2,74
préférer	2	2,74
brûler d'envie	2	2,74
avoir faim	2	2,74
aimer	1	1,37
tenir	1	1,37
chercher	1	1,37
avoir soif	1	1,37
avoir sommeil	1	1,37
désirer	1	1,37
envie	1	1,37

«Количество МЭ») оформлены в виде гиперссылок, по которым пользователи могут перейти на страницу, где находятся все МЭ, на основе которых были получены приведенные данные.

6 Заключение

В процессе создания БД БГК была разработана поисковая подсистема, которая обеспечивает выполнение широкого спектра запросов. Проведенные эксперименты продемонстрировали ее высокую эффективность с точки зрения решаемых лингвистических задач. Важно отметить, особую ценность для развития лингвистической теории полученного «шума» и обнаруженных «потерь». Их экспертный анализ позволит сформулировать более четкие критерии для их поиска в корпусах и БД и уточнить понятие безличной конструкции, что будет способствовать развитию грамматики конструкций русского языка.

База данных БГК стала уже четвертой лингвистической БД, созданной в рамках концепции построения НБД, что свидетельствует о жизнеспособности этой концепции в сфере разработки инструментов лингвистического сопоставительного анализа и формирования информационных ресурсов, не имеющих отечественных и зарубежных аналогов. Разработка БД БГК подтвердила, что концепция НБД и метод унидирекционального контрастного анализа применимы к исследованиям разноплановых языковых единиц и явлений (личные глагольные формы, лингвоспецифичные единицы, коннекторы, безличные глагольные формы, логико-семантические отношения в тексте).

Литература

1. Летуций А. Б. Безличность. Материалы для проекта корпусного описания русской грамматики. — М., 2011. <http://rusgram.ru/Безличность>.
2. Wierzbicka A. Semantics, culture, and cognition. Universal human concepts in culture-specific configurations. — New York — Oxford: Oxford University Press, 1992. 496 p.
3. Бульгина Т. В., Шмелев А. Д. Я, ты и другие в русском синтаксисе // Языковая концептуализация мира (на материале русской грамматики). — М.: Школа «Языки русской культуры», 1997. С. 335–352.
4. Бунтман Н. В., Зализняк Анна А., Зацман И. М., Кружков М. Г., Лошчилова Е. Ю., Сичинава Д. В. Информационные технологии корпусных исследований: принципы построения кросслингвистических баз данных // Информатика и её применения, 2014. Т. 8. Вып. 2. С. 98–110.
5. Зализняк Анна А., Зацман И. М., Инькова О. Ю., Кружков М. Г. Надкорпусные базы данных как лингвистический ресурс // Корпусная лингвистика-2015: Тр. 7-й Междунар. конф. — СПб.: СПбГУ, 2015. С. 211–218.
6. Зализняк Анна А. База данных межъязыковых эквивалентов как инструмент лингвистического анализа // Компьютерная лингвистика и интеллектуальные технологии: по мат-лам ежегодной Междунар. конф. «Диалог». — М.: РГГУ, 2016. Вып. 15(22). С. 763–775.
7. Loiseau S., Sitchinava D. V., Zalizniak A. A., Zatsman I. M. Information technologies for creating the database of equivalent verbal forms in the Russian–French multivariant parallel corpus // Информатика и её применения, 2013. Т. 7. Вып. 2. С. 100–109.
8. Zalizniak Anna A., Sitchinava D. V., Loiseau S., Krzhukov M., Zatsman I. M. Database of equivalent verbal forms in a Russian–French multivariant parallel corpus // 2013 Conference (International) on Artificial Intelligence. — Las Vegas, NV, USA: CSREA Press, 2013. Vol. 1. P. 101–107.
9. Krzhukov M. G., Buntman N. V., Loshchilova E. Ju., Sitchinava D. V., Zalizniak Anna A., Zatsman I. M. A database of Russian verbal forms and their French translation equivalents // Компьютерная лингвистика и интеллектуальные технологии: по мат-лам ежегодной Междунар. конф. «Диалог». — М.: РГГУ, 2014. Вып. 13(20). С. 284–296.
10. Зализняк Анна А. Лингвоспецифичные единицы русского языка в свете контрастного корпусного анализа // Компьютерная лингвистика и интеллектуальные технологии: по мат-лам ежегодной Междунар. конф. «Диалог». — М.: РГГУ, 2015. Вып. 14(21). С. 651–662.
11. Зацман И. М., Инькова О. Ю., Кружков М. Г., Попкова Н. А. Представление кроссязыковых знаний о коннекторах в надкорпусных базах данных // Информатика и её применения, 2016. Т. 10. Вып. 1. С. 106–118.

12. Крузжков М. Г. Информационные ресурсы контрастных лингвистических исследований: электронные корпуса текстов // Системы и средства информатики, 2015. Т. 25. № 2. С. 140–159.
13. Добровольский Д. О., Кретов А. А., Шаров С. А. Корпус параллельных текстов: архитектура и возможности использования // Национальный корпус русского языка: 2003–2005. — М.: Индрик, 2005. С. 263–296.
14. Добровольский Д. О., Кретов А. А., Шаров С. А. Корпус параллельных текстов // Научная и техническая информация. Сер. 2: Информационные процессы и системы, 2005. № 6. С. 16–27.
15. Stubbs M. Words and phrases. Corpus studies of lexical semantics. — Oxford: Blackwell, 2002. 287 p.
16. Zatsman I., Buntman N., Kruzchkov M., Nuriev V., Zalizniak Anna A. Conceptual framework for development of computer technology supporting cross-linguistic knowledge discovery // 15th European Conference on Knowledge Management Proceedings. — Reading, MA, USA: Academic Publishing International Ltd., 2014. P. 1063–1071.
17. Мельчук И. А. О синтаксическом нуле // Типология пассивных конструкций. Диатезы и залоги. — Л.: Наука, 1974. С. 343–361.
18. Guiraud-Weber M. L'effacement du sujet au nominatif dans l'énoncé en russe moderne // Revue des Études slaves, 1983. Vol. 55. No. 1. P. 79–86.
19. Гуро-Вебер М. Субъектные черты и проблема подлежащего в русском языке // Revue des Études slaves, 2002. Vol. 74. No. 2-3. P. 279–289.

Поступила в редакцию 13.10.16

DATABASE OF RUSSIAN IMPERSONAL VERBAL CONSTRUCTIONS

Anna A. Zalizniak^{1,2} and M. G. Kruzchkov²

¹Institute of Linguistics, Russian Academy of Sciences, 1-1 Bolshoy Kislovskiy pereulok, Moscow 125009, Russian Federation

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

Abstract: This article presents the Database of Russian Impersonal Verbal Constructions that has been developed to support a linguistic research of Russian impersonal verbal construction as mirrored in translations into other languages. This information resource was developed based on the concept of Supracorpora Databases (SCDBs). Translation correspondences in the Database of Russian Impersonal Verbal Constructions are presented as ordered pairs that combine formal descriptions of corresponding lexical-grammatical forms found in source and target texts of a parallel corpus. The paper also provides description of the methodology for creation of translation correspondences in the database. Some of the problems related to the task of finding Russian impersonal verbal constructions in corpora are considered and approaches to solving those problems are proposed. Thanks to integrated search and statistical functions, the Database of Russian Impersonal Verbal Constructions and other SCDBs significantly extend capabilities of linguistic experts using corpus-based methods to analyze specific linguistic items, both independently and in contrast with other languages.

Keywords: computer linguistics; contrastive linguistics; information technologies; electronic corpora; supracorpora databases; Russian language; impersonal constructions

DOI: 10.14357/19922264160414

Acknowledgments

This research was performed in the Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, with financial support of the Russian Foundation for Humanities (grant No. 15-04-00507).

References

1. Letuchiy, A. B. 2011. *Bezlichnost'. Materialy dlya proekta korpusnogo opisaniya russkoy grammatiki* [Impersonality. Materials for the Russian corpus-based grammar description project]. Moscow. Available at: <http://rusgram.ru/Безличность> (accessed October 12, 2016).
2. Wierzbicka, A. 1992. *Semantics, culture, and cognition. Universal human concepts in culture-specific configurations*. New York — Oxford: Oxford University Press. 496 p.
3. Bulygina, T. V., and A. D. Shmelev. 1997. Ya, ty i drugie v russkom sintaksise [Me, you, and others in Russian syntax]. *Yazykovaya kontseptualizatsiya mira (na materiale russkoy grammatiki)* [Language-based conceptualization

- of the world (based on Russian grammar)]. Moscow: Shkola “Yazyki russkoy kul’tury.” 335–352.
4. Buntman, N. V., Anna A. Zaliznyak, I. M. Zatsman, M. G. Kruzhkov, E. Yu. Loshchilova, and D. V. Sichinava. 2014. Informatsionnye tekhnologii korpusnykh issledovaniy: Printsipy postroeniya kross-lingvisticheskikh baz dannykh [Information technologies for corpus studies: Underpinnings for cross-linguistic database creation]. *Informatika i ee Primeneniya — Inform. Appl.* 8(2):98–110.
 5. Zaliznyak, Anna A., I. M. Zatsman, O. Yu. In’kova, and M. G. Kruzhkov. 2015. Nadkorpusnye bazy dannykh kak lingvisticheskiy resurs [Supracorpora databases as linguistic resource]. *7th Conference (International) on Corpus Linguistics Proceedings*. St. Petersburg: SPbGU. 211–218.
 6. Zaliznyak, Anna A. 2016. Baza dannykh mezh’yazykovykh ekvivalentsiy kak instrument lingvisticheskogo analiza [Database of cross-linguistic equivalences as a tool for linguistic analysis]. *Computer Linguistics and Intellectual Technologies: Conference (International) “Dialog” Proceedings*. Moscow: RGGU. 15(22):763–775.
 7. Loiseau, S., D. V. Sitchinava, Anna A. Zalizniak, and I. M. Zatsman. 2013. Information technologies for creating the database of equivalent verbal forms in the Russian–French multivariant parallel corpus. *Informatika i ee Primeneniya — Inform. Appl.* 7(2):100–109.
 8. Zalizniak, Anna A., D. V. Sitchinava, S. Loiseau, M. Kruzhkov, and I. M. Zatsman. 2013. Database of equivalent verbal forms in a Russian–French multivariant parallel corpus. *2013 Conference (International) on Artificial Intelligence*. Las Vegas, NV: CRSEA Press. 1:101–107.
 9. Kruzhkov, M. G., N. V. Buntman, E. Ju. Loshchilova, D. V. Sitchinava, Anna A. Zalizniak, and I. M. Zatsman. 2014. A database of Russian verbal forms and their French translation equivalents. *Computer Linguistics and Intellectual Technologies: Conference (International) “Dialog” Proceedings*. Moscow: RGGU. 13(20):284–296.
 10. Zalizniak, Anna A. 2015. Lingvospetsifichnye edinitsy russkogo yazyka v svete kontrastivnogo korpusnogo analiza [Lingvospecific units of Russian in the light of the contrast corpus analysis]. *Computer Linguistics and Intellectual Technologies: Conference (International) “Dialog” Proceedings*. Moscow: RGGU. 14(21):651–662.
 11. Zatsman, I. M., O. Yu. In’kova, M. G. Kruzhkov, and N. A. Popkova. 2016. Predstavlenie krossyazykovykh znaniy o konnektorakh v nadkorpusnykh bazakh dannykh [Representation of cross-lingual knowledge about connectors in supracorpora databases]. *Informatika i ee Primeneniya — Inform. Appl.* 10(1):106–118.
 12. Kruzhkov, M. G. 2015. Informatsionnye resursy kontrastivnykh lingvisticheskikh issledovaniy: Elektronnyye korpusa tekstov [Information resources for contrastive studies: Electronic text corpora]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 25(2):140–159.
 13. Dobrovol’skiy, D. O., A. A. Kretov, and S. A. Sharov. 2005. Korpus parallel’nykh tekstov: Arkhitektura i vozmozhnosti ispol’zovaniya [Corpus of parallel texts: Architecture and applications]. *Natsional’nyy korpus russkogo yazyka: 2003–2005* [Russian National Corpus: 2003–2005]. Moscow: Indrik. 263–296.
 14. Dobrovol’skiy, D. O., A. A. Kretov, and S. A. Sharov. 2005. Korpus parallel’nykh tekstov [Corpus of parallel texts]. *Nauchnaya i tekhnicheskaya informatsiya. Ser. 2 “Informatsionnye protsessy i sistemy”* [Scientific and technical information. Ser. 2 “Informational processes and systems”] 6:16–27.
 15. Stubbs, M. 2002. *Words and phrases: Corpus studies of lexical semantics*. — Oxford: Blackwell. 287 p.
 16. Zatsman, I., N. Buntman, M. Kruzhkov, V. Nuriev, and Anna A. Zalizniak. 2014. Conceptual framework for development of computer technology supporting cross-linguistic knowledge discovery. *15th European Conference on Knowledge Management Proceedings*. Reading, MA: Academic Publishing International. 1063–1071.
 17. Mel’chuk, I. A. 1974. O sintaksicheskom nule [On syntactic zero]. *Tipologiya passivnykh konstruksiy. Diately i zalogi* [Typology of passive constructions. Diatheses and voices]. Leningrad: Nauka. 343–361.
 18. Guiraud-Weber, M. 1983. L’effacement du sujet au nominatif dans l’énoncé en russe moderne [Diffusion of the subject in nominative case in modern Russian utterances]. *Revue des Études slaves* [J. Slavic Studies] 55(1):79–86.
 19. Guiraud-Weber, M. 2002. Sub’ektnye cherty i problema podlezhashchego v russkom yazyke [Subjective features and problem of the subject in Russian language]. *Revue des Études slaves* [J. Slavic Studies] 74(2-3):279–289.

Received October 13, 2016

Contributors

Zalizniak Anna A. (b. 1959) — Doctor of Science of philology, leading scientist, Institute of Linguistics, Russian Academy of Sciences, 1-1 Bolshoy Kislowskiy pereulok, Moscow, 125009, Russian Federation; 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; anna.zalizniak@gmail.com

Kruzhkov Mikhail G. (b. 1975) — leading programmer, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; magnit75@yandex.ru

Андреев Сергей Дмитриевич (р. 1984) — кандидат технических наук, доцент Российского университета дружбы народов; старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Белоусов Василий Владимирович (р. 1977) — кандидат технических наук, старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Бенинг Владимир Евгеньевич (р. 1954) — доктор физико-математических наук, профессор кафедры математической статистики факультета вычислительной математики и кибернетики МГУ им. М. В. Ломоносова; старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Гайдамака Юлия Васильевна (р. 1971) — кандидат физико-математических наук, доцент Российского университета дружбы народов; старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Горшенин Андрей Константинович (р. 1986) — кандидат физико-математических наук, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук; доцент кафедры вычислительной техники Института информационных технологий Московского технологического университета

Грушо Александр Александрович (р. 1946) — доктор физико-математических наук, профессор, заведующий лабораторией Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Данишевский Владислав Игоревич (р. 1992) — аспирант Балтийского федерального университета имени Иммануила Канта

Забезжайло Михаил Иванович (р. 1956) — кандидат физико-математических наук, доцент, заведующий лабораторией Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Зализняк Анна Андреевна (р. 1959) — доктор филологических наук, ведущий научный сотрудник Института языкознания Российской академии наук; ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Зацаринный Александр Алексеевич (р. 1951) — доктор технических наук, профессор, заместитель директора Федерального исследовательского центра «Информатика и управление» Российской академии наук

Зейфман Александр Израилевич (р. 1954) — доктор физико-математических наук, профессор, заведующий кафедрой прикладной математики Вологодского государственного университета; старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук; главный научный сотрудник ИСЭРТ РАН

Каданер Арсений Ильич (р. 1995) — студент механико-математического факультета Московского государственного университета им. М. В. Ломоносова; аналитик Сбербанка России

Карасиков Михаил Евгеньевич (р. 1992) — студент Московского физико-технического института; студент Сколковского института науки

Колесников Александр Васильевич (р. 1948) — доктор технических наук, профессор кафедры телекоммуникаций Балтийского федерального университета имени Иммануила Канта; старший научный сотрудник Калининградского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

Коновалов Михаил Григорьевич (р. 1950) — доктор технических наук, главный научный сотрудник Института проблем информатики Федерального ис-

следовательского центра «Информатика и управление» Российской академии наук

Королев Виктор Юрьевич (р. 1954) — доктор физико-математических наук, профессор, заведующий кафедрой математической статистики факультета вычислительной математики и кибернетики МГУ имени М. В. Ломоносова; ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Корчагин Александр Юрьевич (р. 1989) — младший научный сотрудник факультета вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова

Кружков Михаил Григорьевич (р. 1975) — ведущий программист Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Кудрявцев Алексей Андреевич (р. 1978) — кандидат физико-математических наук, доцент кафедры математической статистики факультета вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова

Листопад Сергей Викторович (р. 1984) — кандидат технических наук, старший научный сотрудник Калининградского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

Мельников Андрей Кимович (р. 1956) — кандидат технических наук, доцент, главный научный сотрудник НТЦ ЗАО «ИнформИнвестГрупп»

Разумчик Ростислав Валерьевич (р. 1984) — кандидат физико-математических наук, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук; доцент Российского университета дружбы народов

Ронжин Александр Федорович (р. 1954) — доктор физико-математических наук, профессор; главный научный сотрудник Института точной механики и вычислительной техники им. С. А. Лебедева Российской академии наук

Румовская София Борисовна (р. 1985) — программист I категории Калининградского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

Самуйлов Константин Евгеньевич (р. 1955) — доктор технических наук, профессор, заведующий кафедрой Российского университета дружбы на-

родов; ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Соколов Игорь Анатольевич (р. 1954) — академик (действительный член) Российской академии наук; доктор технических наук, директор Федерального исследовательского центра «Информатика и управление» Российской академии наук

Сопин Эдуард Сергеевич (р. 1987) — кандидат физико-математических наук, доцент Российского университета дружбы народов; старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Стрижов Вадим Викторович (р. 1967) — доктор физико-математических наук, ведущий научный сотрудник Вычислительного центра им. А. А. Дородницына 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

Титова Анастасия Игоревна (р. 1995) — студентка кафедры математической статистики факультета вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова

Хазеева Гелана Тагировна (р. 1993) — студентка факультета вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова

Черток Андрей Викторович (р. 1987) — младший научный сотрудник факультета вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова; директор по исследованиям и разработкам Сбербанка России

Шнурков Петр Викторович (р. 1953) — кандидат физико-математических наук, доцент Национального исследовательского университета «Высшая школа экономики»

Шоргин Сергей Яковлевич (р. 1952) — доктор физико-математических наук, профессор, заместитель директора Федерального исследовательского центра «Информатика и управление» Российской академии наук

АВТОРСКИЙ УКАЗАТЕЛЬ ЗА 2016 г.

	Вып.	Стр.
Агаларов М. Я. см. Агаларов Я. М.		
Агаларов Я. М., Агаларов М. Я., Шоргин В. С. Об оптимальном пороговом значении длины очереди в одной задаче максимизации дохода системы массового обслуживания типа $M/G/1$	2	70–79
Алексеевский Д. А. Применение контекстно-свободных грамматик для извлечения онтологии из текстов коротких описаний статей биологической тематики	1	119–128
Андреев С. Д. см. Гайдамака Ю. В.		
Андреев С. Д. см. Омётов А. Я.		
Архипов О. П., Архипов П. О., Сидоркин И. И. Вариант создания локальной системы координат для синхронизации изображений выбранных снимков	3	91–97
Архипов П. О. см. Архипов О. П.		
Белоусов В. В. см. Шнурков П. В.		
Белоусов В. В. см. Шнурков П. В.		
Бенинг В. Е. Вычисление асимптотического дефекта некоторых статистических процедур, основанных на выборках случайного объема	4	34–45
Борисов А. В., Босов А. В., Миллер Г. Б. Моделирование и мониторинг состояния VoIP-соединения	2	2–13
Босов А. В. см. Борисов А. В.		
Брюхов Д. О. см. Ступников С. А.		
Вольнова А. А. см. Калиниченко Л. А.		
Гайдамака Ю. В., Андреев С. Д., Сопин Э. С., Самуйлов К. Е., Шоргин С. Я. Анализ характеристик интерференции в модели взаимодействия устройств с учетом среды распространения сигнала	4	2–10
Гасилов А. В. см. Яковлев О. А.		
Гончаров А. В., Стрижов В. В. Метрическая классификация временных рядов со взвешенным выравниванием относительно центроидов классов	2	36–47
Гордов Е. П. см. Калиниченко Л. А.		
Горшенин А. К. Концепция онлайн-комплекса для стохастического моделирования реальных процессов	1	72–81
Горшенин А. К. см. Шнурков П. В.		
Горшенин А. К. см. Шнурков П. В.		
Грушо А. А., Грушо Н. А., Забейайло М. И., Тимонина Е. Е. Интеграция статистических и детерминистских методов анализа информационной безопасности	3	2–8
Грушо А. А., Забейайло М. И., Зацаринный А. А. Об одном способе сокращения вычислений при формировании замыканий Галуа	4	96–104
Грушо Н. А. см. Грушо А. А.		
Данишевский В. И. см. Колесников А. В.		
Забейайло М. И. см. Грушо А. А.		
Забейайло М. И. см. Грушо А. А.		
Зализняк Анна А., Кружков М. Г. База данных безличных глагольных конструкций русского языка	4	132–141
Засыпко В. В. см. Шнурков П. В.		
Захарова Т. В., Шестаков О. В. Анализ точности вейвлет-обработки аэродинамических картин обтекания	3	46–54
Зацаринный А. А., Сучков А. П. Системотехнические подходы к созданию системы поддержки принятия решений на основе ситуационного анализа	4	105–113
Зацаринный А. А. см. Грушо А. А.		

	Вып.	Стр.
Зацман И. М., Инькова О. Ю., Кружков М. Г., Попкова Н. А. Представление кросс-языковых знаний о коннекторах в надкорпусных базах данных	1	106–118
Зацман И. М. см. Минин В. А.		
Зейфман А. И. см. Королев В. Ю.		
Зейфман А. И. см. Королев В. Ю.		
Инькова О. Ю. см. Зацман И. М.		
Исаченко Р. В., Стрижов В. В. Метрическое обучение в задачах мультиклассовой классификации временных рядов	2	48–57
Каданер А. И. см. Черток А. В.		
Калиниченко Л. А., Вольнова А. А., Гордов Е. П., Киселева Н. Н., Ковалева Д. А., Малков О. Ю., Окладников И. Г., Подколodный Н. Л., Позаненко А. С., Пономарева Н. В., Ступников С. А., Фазлиев А. З. Проблемы доступа к данным в исследованиях с интенсивным использованием данных в России	1	2–22
Каллаос Н. К., Сейфуль-Мулюков Р. Б. Сложность и ее информационное содержание	1	129–139
Карасиков М. Е., Стрижов В. В. Классификация временных рядов в пространстве параметров порождающих моделей	4	121–131
Кириков И. А., Колесников А. В., Листопад С. В., Румовская С. Б. «Виртуальный консилиум» — инструментальная среда поддержки принятия сложных диагностических решений	3	81–90
Кириков И. А., Колесников А. В., Листопад С. В., Румовская С. Б. Мелкозернистые гибридные интеллектуальные системы. Часть 2: Двухнаправленная гибридизация	1	96–105
Киселева Н. Н. см. Калиниченко Л. А.		
Ковалева Д. А. см. Калиниченко Л. А.		
Ковалёв С. П. Применение метапрограммирования для повышения технологичности больших автоматизированных систем	1	56–66
Колесников А. В., Листопад С. В., Румовская С. Б., Данишевский В. И. Неформальная аксиоматическая теория ролевых визуальных моделей	4	114–120
Колесников А. В. см. Кириков И. А.		
Колесников А. В. см. Кириков И. А.		
Колин К. К. Гуманитарные аспекты проблемы информационной безопасности	3	111–121
Коновалов М. Г., Разумчик Р. В. О размещении заданий на двух серверах при неполном наблюдении	4	57–67
Корепанов Э. Р. см. Сеницын И. Н.		
Корепанов Э. Р. см. Сеницын И. Н.		
Королев В. Ю., Зейфман А. И., Корчагин А. Ю. Несимметричные распределения Линника как предельные законы для случайных сумм независимых случайных величин с конечными дисперсиями	4	21–33
Королев В. Ю., Корчагин А. Ю., Зейфман А. И. Теорема Пуассона для схемы испытаний Бернулли со случайной вероятностью успеха и дискретный аналог распределения Вейбулла	4	11–20
Корчагин А. Ю. см. Королев В. Ю.		
Корчагин А. Ю. см. Королев В. Ю.		
Кривенко М. П. Критерии значимости отбора признаков классификации	3	32–40
Кружков М. Г. см. Зализняк Анна А.		
Кружков М. Г. см. Зацман И. М.		
Кудрявцев А. А. Байесовские модели массового обслуживания и надежности: априорные распределения с компактным носителем	1	67–71
Кудрявцев А. А. Зависимые от коэффициента баланса характеристики в байесовских моделях с компактным носителем априорных распределений	3	77–80
Кудрявцев А. А., Палионная С. И. Байесовская рекуррентная модель роста надежности: параболическое распределение параметров	2	80–83

	Вып.	Стр.
Кудрявцев А. А., Титова А. И. Байесовские модели массового обслуживания и надежности: вырожденно-вейбулловский случай	4	68–71
Кучерявый Е. А. см. Омётов А. Я.		
Леонтьев Н. Д., Ушаков В. Г. Анализ системы обслуживания с входящим потоком авторегрессионного типа и относительным приоритетом	3	15–22
Листопад С. В. см. Кириков И. А.		
Листопад С. В. см. Кириков И. А.		
Листопад С. В. см. Колесников А. В.		
Малков О. Ю. см. Калиниченко Л. А.		
Марков А. С., Монахов М. М., Ульянов В. В. Разложения типа Корниша–Фишера для распределений статистик, построенных по выборкам случайного размера	2	84–91
Мейханаджян Л. А. Стационарные вероятности состояний в системе обслуживания конечной емкости с инверсионным порядком обслуживания и обобщенным вероятностным приоритетом	2	123–131
Мельников А. К., Ронжин А. Ф. Обобщенный статистический метод анализа текстов, основанный на расчете распределений вероятностей значений статистик	4	89–95
Миллер Г. Б. см. Борисов А. В.		
Минин В. А., Зацман И. М., Хавансков В. А., Шубников С. К. Интенсивность цитирования научных публикаций в изобретениях по информационно-компьютерным технологиям, патентуемых в России отечественными и зарубежными заявителями	2	107–122
Монахов М. М. см. Марков А. С.		
Наумов В. А., Самуйлов К. Е. О связи ресурсных систем массового обслуживания с сетями Эрланга	3	9–14
Окладников И. Г. см. Калиниченко Л. А.		
Омётов А. Я., Андреев С. Д., Тюрликов А. М., Кучерявый Е. А. Анализ производительности беспроводной системы агрегации данных с состязанием для современных сенсорных сетей	3	23–31
Палионная С. И. см. Кудрявцев А. А.		
Подколотный Н. Л. см. Калиниченко Л. А.		
Позаненко А. С. см. Калиниченко Л. А.		
Пономарева Н. В. см. Калиниченко Л. А.		
Попкова Н. А. см. Зацман И. М.		
Разумчик Р. В. см. Коновалов М. Г.		
Ронжин А. Ф. см. Мельников А. К.		
Румовская С. Б. см. Кириков И. А.		
Румовская С. Б. см. Кириков И. А.		
Румовская С. Б. см. Колесников А. В.		
Самуйлов К. Е. см. Гайдамака Ю. В.		
Самуйлов К. Е. см. Наумов В. А.		
Сейфуль-Мулюков Р. Б. см. Каллаос Н. К.		
Серебрянский С. М. см. Тырсин А. Н.		
Сидоркин И. И. см. Архипов О. П.		
Синицын В. И. см. Синицын И. Н.		
Синицын В. И. см. Синицын И. Н.		
Синицын И. Н. Аналитическое моделирование нормальных процессов в стохастических системах со сложными бесселевыми нелинейностями дробного порядка	3	55–65
Синицын И. Н. Ортогональные субоптимальные фильтры для нелинейных стохастических систем на многообразиях	1	34–44
Синицын И. Н., Корепанов Э. Р. Нормальные условно-оптимальные фильтры и экстраполяторы Пугачёва для стохастических систем, линейных относительно состояния	2	14–23

	Вып.	Стр.
Синицын И. Н., Синицын В. И. Аналитическое моделирование распределений в нелинейных стохастических системах на многообразиях методом эллипсоидальной аппроксимации	1	45–55
Синицын И. Н., Синицын В. И., Корепанов Э. Р. Эллипсоидальные субоптимальные фильтры для нелинейных стохастических систем на многообразиях	2	24–35
Скворцов Н. А. см. Ступников С. А.		
Соколов И. А. см. Черток А. В.		
Сопин Э. С. см. Гайдамака Ю. В.		
Стрижов В. В. см. Гончаров А. В.		
Стрижов В. В. см. Исаченко Р. В.		
Стрижов В. В. см. Карасиков М. Е.		
Ступников С. А., Брюхов Д. О., Скворцов Н. А. Анализ системного риска совместного кредитования над неоднородными коллекциями данных	1	23–33
Ступников С. А. см. Калиниченко Л. А.		
Сучков А. П. см. Зацаринный А. А.		
Тимонина Е. Е. см. Грушо А. А.		
Титова А. И. см. Кудрявцев А. А.		
Тырсин А. Н., Серебрянский С. М. Распознавание зависимостей на основе обратного отображения	2	58–64
Тюрликов А. М. см. Омётов А. Я.		
Ульянов В. В. см. Марков А. С.		
Ушаков В. Г. Система обслуживания с гиперэкспоненциальным входящим потоком и профилактиками прибора	2	92–97
Ушаков В. Г. см. Леонтьев Н. Д.		
Фазлиев А. З. см. Калиниченко Л. А.		
Федосеев А. А. К вопросу об уменьшении объема порций учебного материала при электронном обучении	3	105–110
Хавансков В. А. см. Минин В. А.		
Хазеева Г. Т. см. Черток А. В.		
Хохлов Ю. С. Многомерное дробное движение Леви и его приложения	2	98–106
Черток А. В., Каданер А. И., Хазеева Г. Т., Соколов И. А. Метод кумулятивных сумм для поиска смены режима в процессе Орнштейна–Уленбека на основе процесса Леви	4	46–56
Чичагов В. В. Асимптотические разложения средней абсолютной ошибки несмещенной оценки с равномерно минимальной дисперсией и оценки максимального правдоподобия в модели однопараметрического экспоненциального семейства решетчатых распределений	3	66–76
Шестаков О. В. Статистические свойства метода подавления шума, основанного на стабилизированной жесткой пороговой обработке	2	65–69
Шестаков О. В. Усиленный закон больших чисел для оценки риска в задаче реконструкции томографических изображений из проекций с коррелированным шумом	3	41–45
Шестаков О. В. см. Захарова Т. В.		
Шнурков П. В., Горшенин А. К., Белоусов В. В. Аналитическое решение задачи оптимального управления полумарковским процессом с конечным множеством состояний	4	72–88
Шнурков П. В., Засыпко В. В., Белоусов В. В., Горшенин А. К. Разработка алгоритма численного решения задачи оптимального управления инвестициями в закрытой динамической модели трехсекторной экономики	1	82–95
Шоргин В. С. см. Агаларов Я. М.		
Шоргин С. Я. см. Гайдамака Ю. В.		
Шубников С. К. см. Минин В. А.		
Яковлев О. А., Гасилов А. В. Ускоренный алгоритм стереосопоставления на основе геодезических вспомогательных коэффициентов	3	98–104

2016 AUTHOR INDEX

	Issue	Page
Agalarov M. Ya. see Agalarov Ya. M.		
Agalarov Ya. M., Agalarov M. Ya., and Shorgin V. S. About the optimal threshold of queue length in a particular problem of profit maximization in the $M/G/1$ queuing system	2	70–79
Alexeyevsky D. A. BioNLP ontology extraction from a restricted language corpus with context-free grammars	1	119–128
Andreev S. D. see Gaidamaka Yu. V.		
Andreev S. D. see Ometov A. Ya.		
Arkhipov O. P., Arkhipov P. O., and Sidorkin I. I. The option to create a local coordinate system for synchronization of selected images	3	91–97
Arkhipov P. O. see Arkhipov O. P.		
Belousov V. V. see Shnurkov P. V.		
Belousov V. V. see Shnurkov P. V.		
Bening V. E. Calculation of the asymptotic deficiency of some statistical procedures based on samples with random sizes	4	34–45
Borisov A. V., Bosov A. V., and Miller G. B. Modeling and monitoring of VoIP connection	2	2–13
Bosov A. V. see Borisov A. V.		
Briukhov D. O. see Stupnikov S. A.		
Callaos N. K. and Seyful-Mulyukov R. B. Complexity and its information content	1	129–139
Chertok A. V., Kadaner A. I., Khazeeva G. T., and Sokolov I. A. Regime switching detection for the Levy driven Ornstein–Uhlenbeck process using CUSUM methods	4	46–56
Chichagov V. V. Asymptotic expansions of mean absolute error of uniformly minimum variance unbiased and maximum likelihood estimators on the one-parameter exponential family model of lattice distributions	3	66–76
Danischevsky V. I. see Kolesnikov A. V.		
Fazliev A. Z. see Kalinichenko L. A.		
Fedoseev A. A. What is behind the concept of “knowledge in small packages”	3	105–110
Gaidamaka Yu. V., Andreev S. D., Sopin E. S., Samouylov K. E., and Shorgin S. Ya. Interference analysis of the device-to-device communications model with regard to a signal propagation environment	4	2–10
Gasilov A. V. see Yakovlev O. A.		
Goncharov A. V. and Strijov V. V. Metric time series classification using weighted dynamic warping relative to centroids of classes	2	36–47
Gordov E. P. see Kalinichenko L. A.		
Gorshenin A. K. Concept of online service for stochastic modeling of real processes	1	72–81
Gorshenin A. K. see Shnurkov P. V.		
Gorshenin A. K. see Shnurkov P. V.		
Grusho A. A., Grusho N. A., Zabezhailo M. I., and Timonina E. E. Integration of statistical and deterministic methods for analysis of information security	3	2–8
Grusho A. A., Zabezhailo M. I., and Zatsarinny A. A. On the advanced procedure to reduce calculation of Galois closures	4	96–104
Grusho N. A. see Grusho A. A.		
Havanskov V. A. see Minin V. A.		
Inkova O. Yu. see Zatsman I. M.		
Isachenko R. V. and Strijov V. V. Metric learning in multiclass time series classification problem	2	48–57

	Issue	Page
Kadaner A. I. see Chertok A. V.		
Kalinichenko L. A., Volnova A. A., Gordov E. P., Kiselyova N. N., Kovaleva D. A., Malkov O. Yu., Okladnikov I. G., Podkolodnyy N. L., Pozanenko A. S., Ponomareva N. V., Stupnikov S. A., and Fazliev A. Z. Data access challenges for data intensive research in Russia	1	2–22
Karasikov M. E. and Strijov V. V. Feature-based time-series classification	4	121–131
Khazeeva G. T. see Chertok A. V.		
Khokhlov Yu. S. Multivariate fractional Levy motion and its applications	2	98–106
Kirikov I. A., Kolesnikov A. V., Listopad S. V., and Rumovskaya S. B. Fine-grained hybrid intelligent systems. Part 2: Bidirectional hybridization	1	96–105
Kirikov I. A., Kolesnikov A. V., Listopad S. V., and Rumovskaya S. B. “Virtual council” — source environment supporting complex diagnostic decision making	3	81–90
Kiselyova N. N. see Kalinichenko L. A.		
Kolesnikov A. V., Listopad S. V., Rumovskaya S. B., and Danishevsky V. I. Informal axiomatic theory of the role visual models	4	114–120
Kolesnikov A. V. see Kirikov I. A.		
Kolesnikov A. V. see Kirikov I. A.		
Kolin K. K. Humanitarian aspects of information security	3	111–121
Konovalov M. G. and Razumchik R. V. Dispatching to two parallel nonobservable queues using only static information	4	57–67
Korchagin A. Yu. see Korolev V. Yu.		
Korchagin A. Yu. see Korolev V. Yu.		
Korepanov E. R. see Sinitsyn I. N.		
Korepanov E. R. see Sinitsyn I. N.		
Korolev V. Yu., Korchagin A. Yu., and Zeifman A. I. The Poisson theorem for Bernoulli trials with a random probability of success and a discrete analog of the Weibull distribution	4	11–20
Korolev V. Yu., Zeifman A. I., and Korchagin A. Yu. Asymmetric Linnik distributions as limit laws for random sums of independent random variables with finite variances	4	21–33
Koucheryavy E. A. see Ometov A. Ya.		
Kovaleva D. A. see Kalinichenko L. A.		
Kovalyov S. P. Metaprogramming to increase manufacturability of large-scale software-intensive systems	1	56–66
Krivenko M. P. Significance tests of feature selection for classification	3	32–40
Kruzhkov M. G. see Zalizniak Anna A.		
Kruzhkov M. G. see Zatsman I. M.		
Kudryavtsev A. A. Bayesian queueing and reliability models: <i>A priori</i> distributions with compact support	1	67–71
Kudryavtsev A. A. Characteristics dependent on the balance coefficient in Bayesian models with compact support of <i>a priori</i> distributions	3	77–80
Kudryavtsev A. A. and Palionnaia S. I. Bayesian recurrent model of reliability growth: Parabolic distribution of parameters	2	80–83
Kudryavtsev A. A. and Titova A. I. Bayesian queueing and reliability models: Degenerate-Weibull case	4	68–71
Leontyev N. D. and Ushakov V. G. Analysis of a queueing system with autoregressive arrivals and nonpreemptive priority	3	15–22
Listopad S. V. see Kirikov I. A.		
Listopad S. V. see Kirikov I. A.		
Listopad S. V. see Kolesnikov A. V.		
Malkov O. Yu. see Kalinichenko L. A.		
Markov A. S., Monakhov M. M., and Ulyanov V. V. Generalized Cornish–Fisher expansions for distributions of statistics based on samples of random size	2	84–91
Melnikov A. K. and Ronzhin A. F. Generalized statistical method of text analysis based on calculation of probability distributions of statistical values	4	89–95

	Issue	Page
Meykhanadzhyan L. A. Stationary characteristics of the finite capacity queueing system with inverse service order and generalized probabilistic priority	2	123–131
Miller G. B. see Borisov A. V.		
Minin V. A., Zatsman I. M., Havanskov V. A., and Shubnikov S. K. Intensity of citation of scientific publications in inventions on information and computer technologies patented in Russia by domestic and foreign applicants	2	107–122
Monakhov M. M. see Markov A. S.		
Naumov V. A. and Samouylov K. E. On relationship between queueing systems with resources and Erlang networks	3	9–14
Okladnikov I. G. see Kalinichenko L. A.		
Ometov A. Ya., Andreev S. D., Turlikov A. M., and Koucheryavy E. A. Performance analysis of a wireless data aggregation system with contention for contemporary sensor networks	3	23–31
Palionnaia S. I. see Kudryavtsev A. A.		
Podkolodnyy N. L. see Kalinichenko L. A.		
Ponomareva N. V. see Kalinichenko L. A.		
Popkova N. A. see Zatsman I. M.		
Pozanenko A. S. see Kalinichenko L. A.		
Razumchik R. V. see Konovalov M. G.		
Ronzhin A. F. see Melnikov A. K.		
Rumovskaya S. B. see Kirikov I. A.		
Rumovskaya S. B. see Kirikov I. A.		
Rumovskaya S. B. see Kolesnikov A. V.		
Samouylov K. E. see Gaidamaka Yu. V.		
Samouylov K. E. see Naumov V. A.		
Serebryanskii S. M. see Tyrsin A. N.		
Seyful-Mulyukov R. B. see Callaos N. K.		
Shestakov O. V. Statistical properties of the denoising method based on the stabilized hard thresholding	2	65–69
Shestakov O. V. The strong law of large numbers for the risk estimate in the problem of tomographic image reconstruction from projections with a correlated noise	3	41–45
Shestakov O. V. see Zakharova T. V.		
Shnurkov P. V., Gorshenin A. K., and Belousov V. V. Analytical solution of the optimal control task of a semi-Markov process with finite set of states	4	72–88
Shnurkov P. V., Zasytko V. V., Belousov V. V., and Gorshenin A. K. Development of the algorithm of numerical solution of the optimal investment control problem in the closed dynamical model of three-sector economy	1	82–95
Shorgin S. Ya. see Gaidamaka Yu. V.		
Shorgin V. S. see Agalarov Ya. M.		
Shubnikov S. K. see Minin V. A.		
Sidorkin I. I. see Arkhipov O. P.		
Sinitsyn I. N. Analytical modeling of processes in stochastic systems with complex fractional order Bessel nonlinearities	3	55–65
Sinitsyn I. N. Orthogonal suboptimal filters for nonlinear stochastic systems on manifolds	1	34–44
Sinitsyn I. N. and Korepanov E. R. Normal Pugachev conditionally-optimal filters and extrapolators for state linear stochastic systems	2	14–23
Sinitsyn I. N. and Sinitsyn V. I. Analytical modeling of distributions in stochastic systems on manifolds based on ellipsoidal approximation	1	45–55
Sinitsyn I. N., Sinitsyn V. I., and Korepanov E. R. Ellipsoidal suboptimal filters for nonlinear stochastic systems on manifolds	2	24–35
Sinitsyn V. I. see Sinitsyn I. N.		
Sinitsyn V. I. see Sinitsyn I. N.		
Skvortsov N. A. see Stupnikov S. A.		
Sokolov I. A. see Chertok A. V.		

	Issue	Page
Sopin E. S. see Gaidamaka Yu. V.		
Strijov V. V. see Goncharov A. V.		
Strijov V. V. see Isachenko R. V.		
Strijov V. V. see Karasikov M. E.		
Stupnikov S. A., Briukhov D. O., and Skvortsov N. A. Co-lending systemic risk analysis over heterogeneous data collections	1	23–33
Stupnikov S. A. see Kalinichenko L. A.		
Suchkov A. P. see Zatsarinny A. A.		
Timonina E. E. see Grusho A. A.		
Titova A. I. see Kudryavtsev A. A.		
Turlikov A. M. see Ometov A. Ya.		
Tyrsin A. N. and Serebryanskii S. M. Recognition of dependences on the basis of inverse mapping	2	58–64
Ulyanov V. V. see Markov A. S.		
Ushakov V. G. Queuing system with working vacations and hyperexponential input stream	2	92–97
Ushakov V. G. see Leontyev N. D.		
Volnova A. A. see Kalinichenko L. A.		
Yakovlev O. A. and Gasilov A. V. Speeded-up stereo matching using geodesic support weights	3	98–104
Zabekhailo M. I. see Grusho A. A.		
Zabekhailo M. I. see Grusho A. A.		
Zakharova T. V. and Shestakov O. V. Precision analysis of wavelet processing of aerodynamic flow patterns	3	46–54
Zalizniak Anna A. and Kruzhkov M. G. Database of Russian impersonal verbal constructions	4	132–141
Zasytko V. V. see Shnurkov P. V.		
Zatsarinny A. A. and Suchkov A. P. Systems engineering approaches to the establishment of a system for decision support based on situational analysis	4	105–113
Zatsarinny A. A. see Grusho A. A.		
Zatsman I. M., Inkova O. Yu., Kruzhkov M. G., and Popkova N. A. Representation of cross-lingual knowledge about connectors in supracorpora databases	1	106–118
Zatsman I. M. see Minin V. A.		
Zeifman A. I. see Korolev V. Yu.		
Zeifman A. I. see Korolev V. Yu.		

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 \LaTeX (.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 15 страниц указанного формата. При превышении указанного объема редколлегия вправе потребовать от автора сокращения объема рукописи.

Сокращения слов, помимо стандартных, не допускаются. Допускается минимальное количество аббревиатур.

Все страницы рукописи нумеруются.

Шаблоны примеров оформления представлены в Интернете: <http://www.ipiran.ru/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_rus/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://translit.ru/> бесплатной программой транслитерации русского текста в латиницу, при этом в закладке «варианты. . . » следует выбрать опцию BGN.

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. *Rus. J. Electrochem.* 44(8):926–930. doi:10.1134/S10231935080077.

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 resursosberegayushchie tekhnologii nedropol'zovaniya i povysheniya neftegazootdachi"* [6th Symposium (International) "New Energy Saving Subsoil Technologies and the Increasing of the Oil and Gas Impact" Proceedings]. Moscow. 267–272.

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. — N.Y.: Wiley, 1974. 521 p.)

2. В англоязычной части:

Timoshenko, S. P., D. H. Young, and W. Weaver. 1974. *Vibration problems in engineering*. 4th ed. N.Y.: Wiley. 521 p.

Описание неопубликованного документа:

Latypov, A. R., M. M. Khasanov, and V. A. Baikov. 2004. Geology and production (NGT GiD). Certificate on official registration of the computer program No. 2004611198. (In Russian, unpubl.)

Описание интернет-ресурса:

Pravila tsitirovaniya istochnikov [Rules for the citing of sources]. Available at: <http://www.scribd.com/doc/1034528/> (accessed February 7, 2011).

Описание диссертации или автореферата диссертации:

Semenov, V. I. 2003. *Matematicheskoe modelirovanie plazmy v sisteme kompaktnyy tor* [Mathematical modeling of the plasma in the compact torus]. D.Sc. Diss. Moscow. 272 p.

Kozhunova, O. S. 2009. *Tekhnologiya razrabotki semanticheskogo slovarya informatsionnogo monitoringa* [Technology of development of semantic dictionary of information monitoring system]. PhD Thesis. Moscow: IPI RAN. 23 p.

Описание ГОСТа:

GOST 8.586.5-2005. 2007. *Metodika vypolneniya izmereniy. Izmerenie 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: rust@ipiran.ru (Сейфуль-Мулюков Рустем Бадриевич)

<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 \LaTeX (.tex) formats, and additionally in the .pdf format. All documents may be sent by e-mail or provided on a CD or diskette. A hard copy submission is not necessary.
7. The recommended typesetting instructions for manuscript.

Pages parameters: format A4, portrait orientation, document margins (cm): left — 2.5, right — 1.5, above — 2.0, below — 2.0, footer 1.3.

Text: font — Times New Roman, font size — 14, paragraph indent — 0.5, line spacing — 1.5, justified alignment.

The recommended manuscript size: not more than 15 pages of the specified format. If the specified size exceeded, the editorial board is entitled to require the author to reduce the manuscript.

Use only standard abbreviations. Avoid abbreviations in the title and abstract. The full term for which an abbreviation stands should precede its first use in the text unless it is a standard unit of measurement.

All pages of the manuscript should be numbered.

The templates for the manuscript typesetting are presented on site: <http://www.ipiran.ru/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.ru>, option BGN).
Please take into account the following examples of Russian references appearance:
- Article in journal:**
Zhang, Z., and D. Zhu. 2008. Experimental research on the localized electrochemical micromachining. *Rus. J. Electrochem.* 44(8):926–930. doi:10.1134/S1023193508080077.
- Journal article in electronic format:**
Swaminathan, V., E. Lepkoswka-White, and B. P. Rao. 1999. Browsers or buyers in cyberspace? An investigation of electronic factors influencing electronic exchange. *JCMC* 5(2). Available at: <http://www.ascusc.org/jcmc/vol5/issue2/> (accessed April 28, 2011).
- 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 resursosberegayushchie tekhnologii nedropol’zovaniya i povysheniya neftegazoidachi” [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 slovary informatsionnogo monitoringa [Technology of development of semantic dictionary of information monitoring system]. PhD Thesis. Moscow: IPI RAN. 23 p.
- State standards and patents:**
GOST 8.586.5-2005. 2007. Metodika vypolneniya izmereniy. Izmerenie raskhoda i kolichestva zhidkostey i gazov s pomoshch’yu standartnykh suzhayushchikh ustroystv [Method of measurement. Measurement of flow rate and volume of liquids and gases by means of orifice devices]. M.: Standardinform Publs. 10 p.
Bolshakov, M. V., A. V. Kulakov, A. N. Lavrenov, and M. V. Palkin. 2006. Sposob orientirovaniya po krenu letatel’nogo apparata s opticheskoy 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: rust@ipiran.ru (to Prof. Rustem Seyful-Mulyukov)
<http://www.ipiran.ru/english/journal.asp>