

Информатика и её применения

Том 10 Выпуск 3 Год 2016

СОДЕРЖАНИЕ

Интеграция статистических и детерминистских методов анализа информационной безопасности А. А. Грушо, Н. А. Грушо, М. И. Забежайло, Е. Е. Тимонина	2
О связи ресурсных систем массового обслуживания с сетями Эрланга В. А. Наумов, К. Е. Самуйлов	9
Анализ системы обслуживания с входящим потоком авторегрессионного типа и относительным приоритетом Н. Д. Леонтьев, В. Г. Ушаков	15
Анализ производительности беспроводной системы агрегации данных с состязанием для современных сенсорных сетей А. Я. Омётов, С. Д. Андреев, А. М. Тюрликов, Е. А. Кучерявый	23
Критерии значимости отбора признаков классификации М. П. Кривенко	32
Усиленный закон больших чисел для оценки риска в задаче реконструкции томографических изображений из проекций с коррелированным шумом О. В. Шестаков	41
Анализ точности вейвлет-обработки аэродинамических картин обтекания Т. В. Захарова, О. В. Шестаков	46
Аналитическое моделирование нормальных процессов в стохастических системах со сложными бесселевыми нелинейностями дробного порядка И. Н. Сеницын	55
Асимптотические разложения средней абсолютной ошибки несмещенной оценки с равномерно минимальной дисперсией и оценки максимального правдоподобия в модели однопараметрического экспоненциального семейства решетчатых распределений В. В. Чичагов	66
Зависимые от коэффициента баланса характеристики в байесовских моделях с компактным носителем априорных распределений А. А. Кудрявцев	77
«Виртуальный консилиум» — инструментальная среда поддержки принятия сложных диагностических решений И. А. Кириков, А. В. Колесников, С. В. Листопад, С. Б. Румовская	81
Вариант создания локальной системы координат для синхронизации изображений выбранных снимков О. П. Архипов, П. О. Архипов, И. И. Сидоркин	91
Ускоренный алгоритм стереосопоставления на основе геодезических вспомогательных коэффициентов О. А. Яковлев, А. В. Гасилов	98
К вопросу об уменьшении объема порций учебного материала при электронном обучении А. А. Федосеев	105
Гуманитарные аспекты проблемы информационной безопасности К. К. Колин	111
Об авторах	122
Правила подготовки рукописей	124
Requirements for manuscripts	127

Технический редактор *Л. Кокушкина* Художественный редактор *М. Седакова*
Сдано в набор 11.06.16. Подписано в печать 14.09.16. Формат 60 x 84 / 8
Бумага офсетная. Печать цифровая. Усл.-печ. л. 14,88. Уч.-изд. л. 12,0. Тираж 100 экз.
Заказ № 940

Издательство «ТОРУС ПРЕСС», Москва 121614, ул. Крылатская, 29-1-43
Отпечатано в НИПКЦ «Восход-А» с готовых файлов
Москва 109052, ул. Смирновская, д. 25, стр. 3

ИНТЕГРАЦИЯ СТАТИСТИЧЕСКИХ И ДЕТЕРМИНИСТСКИХ МЕТОДОВ АНАЛИЗА ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ*

А. А. Грушо¹, Н. А. Грушо², М. И. Забейло³, Е. Е. Тимонина⁴

Аннотация: Статья посвящена разработке методов автоматического анализа и управления механизмами информационной безопасности в облачных вычислительных средах. Рассмотренные подходы основаны на синтезе вероятностно-статистических и детерминистских методов анализа ситуаций, встречающихся в задачах информационной безопасности. Статистический анализ позволяет сформировать множество объектов для детерминистского (логического) анализа. Поскольку детерминистские методы требуют больших объемов вычислений, предварительная статистическая обработка позволяет выделить для детерминистского (логического) анализа доступные для вычисления объемы данных. В работе детерминистские методы представлены аналогами поиска причинно-следственных связей. Применение эвристик и правдоподобных рассуждений может породить недостоверные выводы, которые связаны со случайным характером исходных данных, поэтому проводится анализ возможностей случайного порождения заключений детерминистского характера. Предложенные методы анализа ориентированы на двухуровневую архитектуру построения информационной безопасности в облачных вычислительных средах. В этой архитектуре автоматический интеллектуальный анализ данных порождает на верхнем уровне быструю реакцию для разрешения конфликтов в вычислительных процессах или выявления функционирования вредоносного кода.

Ключевые слова: облачные вычислительные среды; информационная безопасность; вероятностно-статистические и детерминистские (логические) методы анализа; эвристические алгоритмы; взаимное влияние данных

DOI: 10.14357/19922264160301

1 Введение

Изложим идеологию взаимного дополнения статистических и логических методов. Статистическая обработка данных основана на использовании технологии усреднения (сглаживания). Тогда значимые отклонения от средних сглаженных данных порождают исходные данные для получения новой информации [1].

Эта идея основана на том, что значимые отклонения могут иметь закономерную структуру, которая отсутствует в основной массе данных, использованных для сглаживания. В качестве примера таких задач можно указать способы определения и выявления признаков атак с помощью анализа логов [2, 3].

В работе [4] предлагается искать локальные соотношения, которым удовлетворяют случайные последовательности.

Таким образом статистические методы позволяют выделить множество фрагментов данных, отличающихся от средних сглаженных данных. Таких фрагментов немного, поэтому далее использовать статистические методы анализа неэффективно.

Вместе с тем эти фрагменты можно рассматривать как множество объектов для применения логических методов анализа [5, 6]. Эти методы позволяют строить эвристические закономерности, анализировать причинно-следственные связи, автоматически порождать гипотезы о структурных зависимостях в данных [6].

Логические методы, использующие эвристику и правдоподобные рассуждения, при всей строгости логического анализа не гарантируют достоверности выводов. Поэтому целесообразно сравнивать полученные выводы с результатами оценок, полученных вероятностными методами, в предположе-

* Работа поддержана РФФИ (проекты 15-29-07981 и 15-07-02053).

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

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

³Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, m.zabehailo@yandex.ru

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

нии стохастического характера результатов логической обработки данных. Это особенно важно, когда методы интеллектуального анализа данных применяются многократно на больших объемах данных. В таких многократных случаях требуется автоматическая интеллектуальная обработка данных, так как объем необходимых вычислений существенно превышает человеческие возможности.

В продолжение идей ДСМ (Джон Стюарт Милль) метода [6] в работе приведена «смягченная» модель анализа причинно-следственных связей в данных и простейшая стохастическая модель ее валидации.

2 Влияние характеристик процессов на их целевые свойства

Рассмотрим $r + 1$ процессов ξ_0, \dots, ξ_r , $r \geq 0$. Пусть ξ_0 — целевой процесс, его состояния содержат интересующее нас состояние (свойство) p . Задача состоит в выявлении влияния остальных процессов ξ_1, \dots, ξ_r на появление свойства p в процессе ξ_0 . В данной работе понятие влияния отличается от известных функций влияния [7]. Пусть время дискретно, что позволяет точно синхронизировать все процессы. Пусть $t_1, t_2, \dots, t_n, \dots$ — моменты появления свойства p в процессе ξ_0 . Для простоты исследуем влияние процесса ξ_1 на появление свойства p . Для этого рассмотрим фрагменты процесса ξ_1 в промежутках времени $t_1, t_2, \dots, t_n, \dots$. Будем считать, что процесс ξ_1 принимает значения в алфавите $A = \{a_1, \dots, a_m\}$.

В усложненной схеме поиска влияния процесса ξ_1 на появление свойства p следует рассматривать подпоследовательности длин, больших либо равных l , для выделенных фрагментов. Параметр l характеризует степень положительного влияния ξ_1 на появление свойства p . Наличие достаточно большого количества фрагментов, содержащих одинаковые подпоследовательности, можно рассматривать как довод в пользу влияния процесса ξ_1 на появление свойства p процесса ξ_0 .

В более простой схеме следует рассматривать множества элементов в каждом фрагменте и выбирать такие фрагменты для определения влияния, в которых число элементов в пересечении множеств этих фрагментов больше либо равно l .

Выбор между упрощенной или усложненной схемами связан со сложностью вычислений и/или со схемами расчета вероятностей ошибочного (случайного) решения о наличии влияния ξ_1 на появление свойства p .

Рассмотрим упрощенную схему и обозначим через $T_0 < T_1$ два пороговых значения времени. Пусть $\Phi^+(i)$ — множество различных элементов, расположенных на расстоянии, не превосходящем порог T_0 , от p в i -м фрагменте. Поскольку на появление свойства p могут влиять различные множества элементов из ξ_1 , то выбираем подпоследовательность из k фрагментов i_1, i_2, \dots, i_k , $k \geq 2$, таких что

$$\left| \bigcap_{i \in \{i_1, \dots, i_k\}} \Phi^+(i) \right| \geq l.$$

Замечание 1. В случае исследования влияния r процессов на появление свойства p надо брать пересечение множеств $\Phi_j^+(i)$, $j = 1, \dots, r$.

В данной схеме необходимо оценить возможность ошибочного решения о влиянии процесса ξ_1 на появление свойства p . Рассмотрим матрицу $K^+ = \|k_{ij}^+\|$, столбцы которой пронумерованы числами от 1 до m , а строки пронумерованы номерами фрагментов процесса ξ_1 , определяемыми моментами времени $t_1, t_2, \dots, t_n, \dots$, элементы k_{ij}^+ матрицы K^+ равны 1 или 0, при этом $k_{ij}^+ = 1$, если элемент $a_j \in \Phi^+(i)$. Матрица K_n^+ получается из матрицы K^+ ограничением на первые n строк.

Число единиц в каждой строке матрицы K^+ зависит от длин фрагментов и вероятностей появления букв алфавита A . В простейшей модели ошибочного влияния будем считать, что элементы матрицы K^+ появляются независимо друг от друга с одинаковой вероятностью q появления единицы. Вероятность появления заданного подмножества мощности l в данной строке равна q^l . Вероятность P_s^+ того, что данное подмножество встретится в не менее чем s строках матрицы K_n^+ , равна

$$P_s^+ = \sum_{v=s}^n \binom{n}{v} q^{lv} (1 - q^l)^{n-v}.$$

Используя теорему Муавра–Лапласа [8], получаем оценку s , при которой вероятность случайного влияния данного подмножества стремится к 0 при $n \rightarrow \infty$:

$$s \geq q^l n + \ln n \sqrt{nq^l(1 - q^l)}.$$

В таблице представлены численные расчеты значений s для различных значений q , l и n .

Используя неравенство Маркова [9], получим, что вероятность случайного появления какого-либо множества, содержащего число элементов, большее или равное l , и влияющего на появление свойства p в процессе ξ_1 , стремится к нулю.

Выше рассматривалось положительное влияние на появление свойства p . Однако согласно [6] следует рассматривать также и негативное влияние. Множество $\Phi^-(i)$ негативно влияет на появление

Оценка s , при которой вероятность случайного влияния подмножества мощности l стремится к 0

Вероятность появления единицы $q = 0,5$			Вероятность появления единицы $q = 0,9$		
n	l	s	n	L	s
100	3	19	100	3	79
400	5	20	400	5	245
10000	10	21	10000	10	3500

свойства p , если для каждого состояния $a_j \in \Phi^-(i)$ расстояние от a_j до p больше порога T_1 . Тогда негативное влияние на появление свойства p существует, если существуют фрагменты i_1, i_2, \dots, i_k , $k \geq 2$, такие что $\left| \bigcap_{i \in \{i_1, \dots, i_k\}} \Phi^-(i) \right| \geq l$. Параметр l характеризует степень негативного влияния ξ_1 на появление свойства p .

Для ускорения поиска негативного влияния рассмотрим расстояние $d(p, p')$ между свойствами p и p' . Выберем те свойства p' , для которых $d(p, p') > T_1$. Для таких p' найдем положительное влияние $\Phi^+(p')$. Тогда расстояние $d(p, \Phi^+(p')) \geq T_1$, так как $\Phi^+(p')$ предшествует появлению свойства p' .

Влияние $\Phi^{(0)}(i)$ — это нулевое влияние, если существуют фрагменты i_1, i_2, \dots, i_k , $k \geq 2$, такие что $d\left(\bigcap_{i \in \{i_1, \dots, i_k\}} \Phi^{(0)}(i), p\right) < T_0$, и существуют фрагменты j_1, j_2, \dots, j_v , $v \geq 2$, такие что $d\left(\bigcap_{i \in \{j_1, \dots, j_v\}} \Phi^{(0)}(i), p\right) \geq T_1$. Возможность случайного появления негативного влияния $\Phi^-(i)$ или нулевого влияния $\Phi^{(0)}(i)$ оценивается аналогично тому, как это было сделано для положительного влияния $\Phi^+(i)$.

3 Графы влияния

Рассмотренную в предыдущем разделе схему влияния можно обобщить с помощью графов влияния. Вершинами графа влияния являются свойства объектов влияния (события, участие в порождении информационного потока, вызов функции компьютерной системы, элемент решения некоторой задачи, фрагмент вычислительного процесса и т.д.), а дуги графа отражают влияние одних свойств на другие. Каждая вершина состоит из двух частей «+» и «-». Дуга графа из вершины i в часть «+» вершины j означает положительное влияние i на j , т.е. участие в порождении свойства j . Дуга графа из вершины i в часть «-» вершины j означает отрицательное влияние i на j , т.е. препятствует

появлению свойства j . Если из одной вершины i идут две дуги в «+» и «-» вершины j , то этот случай обозначим «0», что значит противоречие (неправильное понимание влияния). Если дуги из i в j нет, то этот случай будем обозначать через «τ», что означает отсутствие влияния или отсутствие знания о влиянии.

Дуги в графах влияния можно маркировать степенью влияния или силой влияния. Степень влияния ранжируется между максимальным влиянием, что является причиной свойства или отсутствия свойства, и минимальным влиянием, что означает невлияние на появление данного свойства.

В связи с графами влияния возникают следующие задачи.

- 1. Порождение графов влияния.** В разд. 2 рассмотрен случай выявления влияний «+», «-», «0» и структуры этих влияний в виде свойства процесса ξ_1 .
- 2. Преобразование графов влияния.** Ясно, что транзитивное замыкание двух последовательных дуг из x в y и из y в z означает возможность влияния x на z , но более слабое, чем влияние x на y и y на z . Таким образом, введение дополнительной вершины позволяет снижать влияние одного свойства на другое. Этот вывод использовался в работе [10]. Наоборот, возможность агрегирования всех вершин, имеющих влияние на данную, так, что из агрегированного множества не выходит других дуг, кроме как в данную вершину, означает усиление влияния до причины (в смысле Д. С. Милля [6]).
- 3. Пути использования графов влияния в задачах обеспечения информационной безопасности.**

4 Примеры описания влияний

4.1 Однородная простая цепь Маркова

Введенные понятия проще всего объяснить на примере простой однородной цепи Маркова на множестве состояний $A = \{a_1, \dots, a_m\}$ с матрицей переходных вероятностей P . Если в матрице переходных вероятностей P все ненулевые вероятности заменить на 1, то получится часть матрицы смежности графа влияний на переходы из состояния в состояние, описывающего влияние «+» для $T_0 = 1$.

Пусть из вершины x в вершину y ведет единственная дуга, других дуг из вершины x нет и нет других дуг в вершину y . Тогда состояние x является причиной состояния y (рис. 1).



Рис. 1 Состояние x является причиной состояния y

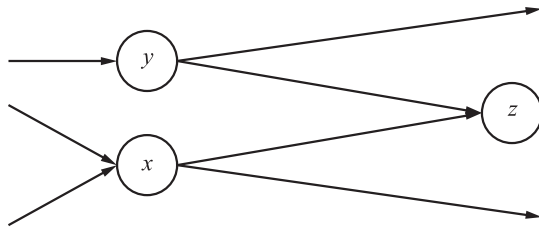


Рис. 2 x и y влияют на «+»-свойства z

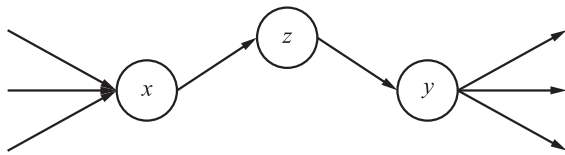


Рис. 3 Влияние «-»-свойства x на свойство y

Если из вершин x и y ведут дуги в вершину z , то x влияет на «+» вершины z и y влияет на «+» вершины z (рис. 2).

Если из вершины x нет дуги в вершину y , то это можно рассматривать как влияние «-»-свойства x на свойство y для $T_1 = 2$ (рис. 3).

В случае заданной однородной цепи Маркова невозможно появление «0» и «т».

Для однородной цепи Маркова с одним ациклическим эргодическим классом транзитивное замыкание графа влияний на определенное число шагов позволяет обосновать влияние «+» каждого свойства x на каждое свойство y . При этом для каждого шага замыкания дуги в части «-»-свойств строятся отдельно от предыдущих и последующих шагов итераций.

4.2 Однородная сложная цепь Маркова

Для однородной сложной цепи Маркова глубины 2 переходы определяются графом, вершинами которого являются пары из множества A^2 , причем допустимы только переходы вида (x, y) в (y, z) . В таком графе влияние «+» на свойство z будет определяться ориентированными цепочками длины 2.

При переходе на влияние множеств может возникнуть влияние «0».

5 Использование графов влияния в задачах обеспечения информационной безопасности

5.1 Идентификация безопасного обращения к ресурсу

Одним из признаков функционирования вредоносного кода в программной среде является обращение к набору библиотек, функции которых позволяют реализовать вредоносное воздействие (критические функции). Это не означает, что обращение к этим библиотекам недопустимо для легальных программ. При этом даже в легальном вычислительном процессе обращение к критическим функциям не всегда является необходимым.

Классификация программ по возможностям обращения к критическим функциям является сложной задачей. Однако можно выделить признаки (влияние), позволяющие идентифицировать возможность обращения к критическим функциям. Появление таких признаков в легальном вычислительном процессе позволяет сделать вывод о том, что обращение к критическим функциям будет легальным. При отсутствии указанных признаков обращение к критическим функциям может означать функционирование враждебного кода. Тогда такое обращение следует рассматривать как событие безопасности и принимать меры по предотвращению враждебного воздействия.

Исследование влияния на возможность обращения к критическим функциям можно проводить предварительно, а в ходе вычислительного процесса использовать сигнатуры влияния в легальных вычислительных процессах.

5.2 Разрешение конфликтов в вычислительных процессах с помощью влияния

Конфликты в вычислительных процессах могут возникать по различным причинам. Например, выполнение процесса блокируется нарушением какого-либо требования политики безопасности одного из участников облачных вычислений. Возникает задача «отката» вычислительного процесса и построения новой траектории процесса, не позволяющей снова выйти на конфликт.

Рассмотрим эвристический алгоритм решения этой задачи с использованием влияний. Пусть p является причиной конфликта, F — влияние на появление p . Задача состоит в том, чтобы найти новые, близкие к F траектории, не ведущие к p . Поскольку F описывает влияние на появление p , то возможны случаи появления F без последующе-

го появления p . Обозначим через X множество состояний процесса, которые следуют за F , но не совпадают с p .

Рассмотрим влияние на каждый из элементов множества X . И пусть F' принадлежит этому множеству, т.е. F' описывает влияние на появление какого-то элемента из X , который не совпадает с p . Множество F' не совпадает с F , но является влиянием для свойства, которому может предшествовать F . Таким образом, F' может рассматриваться как признак альтернативной траектории вычислительного процесса, не приводящей к конфликту.

Возможны дополнительные ограничения на альтернативные траектории вычислительного процесса, а именно: начало альтернативной траектории должно совпадать с точкой «отката» первоначальной траектории, приведшей к конфликту. В отброшенном фрагменте первоначальной траектории вычислительного процесса могли быть выполнены необходимые функции, не входившие в множество влияния F на появление конфликта. Эти выполненные функции можно оставить в новом варианте траектории вычислительного процесса. Таким образом, влияние позволяет построить эвристический алгоритм разрешения конфликтов в вычислительном процессе, отличный от полного перебора.

6 Заключение

В работе рассматриваются пути совместного использования статистических и логических методов анализа больших стохастических данных.

Статистический анализ позволяет сформировать множество объектов для детерминистского (логического) анализа. Поскольку детерминистские методы требуют больших объемов вычислений, то предварительная статистическая обработка позволяет выделить доступные для вычисления объемы данных для детерминистского (логического) анализа.

В работе детерминистские методы представлены простейшим аналитическим поиском причинно-следственных связей [6]. Применение эвристических и правдоподобных рассуждений может породить недостоверные выводы, которые связаны со случайным характером исходных данных, поэтому требуется проверка возможности получения подобных выводов случайно в простейших вероятностных моделях. В работе предложенные методы анализа ориентированы на построение системы анализа информационной безопасности в облачных вычислительных средах. Они позволяют построить механизмы разрешения конфликтов в вы-

числительных процессах и идентифицировать легальное использование критических ресурсов.

Рассмотренные методы интеллектуального анализа данных предполагают двухуровневую архитектуру решения проблем информационной безопасности, при которой сложные задачи решаются в режиме офлайн и реализуются в виде скоростных алгоритмов принятия решений на основе сигнатур.

Литература

1. Тьюки Дж. Анализ результатов наблюдений. Разведочный анализ / Пер. с англ. — М.: Мир, 1981. 696 с. (Tukey J. W. Exploratory data analysis. — Addison-Wesley Pub. Co, 1977. 688 p.)
2. Норткат С., Купер М., Фирноу М., Фредерик К. Анализ типовых нарушений безопасности в сетях / Пер. с англ. — М.: Вильямс, 2001. 464 с. (Norcutt S., Cooper M., Fearnow M., Frederik K. Intrusion signatures and analysis. — New Readers Pub., 2001. 408 p.)
3. Грушо А., Забежайло М., Зацаринный А. Контроль и управление информационными потоками в облачной среде // Информатика и её применения, 2015. Т. 9. Вып. 4. С. 95–101.
4. Бендат Дж., Пирсол А. Прикладной анализ случайных данных / Пер. с англ. — М.: Мир, 1989. 540 с. (Bendat J. S., Piersol A. G. Random data: Analysis and measurement procedures. — 2nd ed. — New York, NY, USA: John Wiley and Sons, 1986. 566 p.)
5. Журавлев Ю. И. Корректные алгебры над множеством некорректных (эвристических) алгоритмов // Кибернетика, 1977. Ч. I. № 4. С. 5–17; Ч. II. № 6. С. 21–27; Кибернетика, 1978. Ч. III. № 2. С. 35–43.
6. Панкратова Е. С., Финн В. К. Автоматическое порождение гипотез / Под общ. ред. В. К. Финна. — М.: Либроком, 2009. 528 с.
7. Хампель Ф., Рончетти Э., Рауссеув П., Штаэль В. Робастность в статистике. Подход на основе функций влияния / Пер. с англ. — М.: Мир, 1989. 519 с. (Hampel F. R., Ronchetti E. M., Rousseeuw P. J., Stahel W. A. Robust statistics. The approach based on influence functions. — New York, NY, USA: John Willey and Sons, 1986. 509 p.)
8. Феллер В. Введение в теорию вероятностей и ее приложения / Пер. с англ. — М.: Мир, 1967. Т. 1. 499 с. (Feller W. An introduction to probability theory and its applications. — 2nd ed. — New York, NY, USA: John Wiley and Sons, 1950. Vol. 1. 520 p.)
9. Лозэ М. Теория вероятностей / Пер. с англ. — М.: ИЛ, 1962. 720 с. (Loève M. Probability theory. — Princeton, NJ, USA: D Van Nostrand, 1955. 701 p.)
10. Грушо А. А., Грушо Н. А., Тимонина Е. Е., Шоргин С. Я. Безопасные архитектуры распределенных систем // Системы и средства информатики, 2014. Т. 24. № 3. С. 18–31.

Поступила в редакцию 27.06.16

INTEGRATION OF STATISTICAL AND DETERMINISTIC METHODS FOR ANALYSIS OF INFORMATION SECURITY

A. A. Grusho, N. A. Grusho, M. I. Zabezhailo, and E. E. Timonina

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

Abstract: The paper is devoted to the methods of automatic analysis of information security and control mechanisms in cloudy computing environments. The considered approaches are based on synthesis of probabilistic and statistical and deterministic methods. The statistical analysis allows creating a set of objects for the deterministic (logical) analysis. As deterministic methods demand large volumes of calculations, preliminary statistical processing allows to reduce volumes of data for the deterministic (logical) analysis. In the paper, deterministic methods are presented by analogs of search of causal relationships. Application of heuristic and plausible reasonings can generate doubtful conclusions which are connected with random character of source data. Therefore, the analysis of random generation of deterministic conclusions is considered. The suggested methods of the analysis are focused on two-level architecture of information security system in cloudy computing environments. In this architecture, a slow automatic data mining generates at the top level fast reaction for resolution of conflicts in computing processes or identification of malicious code functioning.

Keywords: cloudy computing environments; information security; probabilistic and statistical and deterministic (logical) methods of the analysis; heuristic algorithms; mutual influence of data

DOI: 10.14357/19922264160301

Acknowledgments

The paper was supported by the Russian Foundation for Basic Research (projects 15-29-07981 and 15-07-02053).

References

1. Tukey, J. W. 1977. *Exploratory data analysis*. Addison-Wesley Pub. Co. 688 p.
2. Nortcutt, S., M. Cooper, M. Fearnow, and K. Frederik. 2001. *Intrusion signatures and analysis*. New Readers Pub. 408 p.
3. Grusho, A., M. Zabezhaylo, and A. Zatsarinny. 2015. Kontrol' i upravlenie informatsionnymi potokami v oblachnoy srede [Control and management of information streams in the cloudy environment]. *Informatika i ee Primeneniya — Inform. Appl.* 9(4):95–101.
4. Bendat, J. S., and A. G. Piersol. 1986. *Random data: Analysis and measurement procedures*. 2nd ed. New York, NY: John Wiley and Sons. 566 p.
5. Zhuravlev, Yu. I. 1977–1978. Korrektnye algebrы nad mnozhestvom nekorrektnykh (evristicheskikh) algoritmov. Ch. I (1977); Ch. II (1977); Ch. III (1978) [Correct algebras over a set of incorrect (heuristic) algorithms. Part I (1977), Part II (1977), Part III (1978)]. *Kibernetika [Cybernetics]* I(4):5–17; II(6):21–27; III(2):35–43.
6. Pankratova, E. S., and V. K. Finn. 2009. *Avtomatischeskoe porozhdenie gipotez [Automatic hypothesis generation in intelligent systems]*. Ed. V. K. Finn. Moscow: Librokom. 528 p.
7. Hampel F. R., E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel. 1986. *Robust statistics. The approach based on influence functions*. John Wiley and Sons, Inc. 509 p.
8. Feller, W. 1950. *An introduction to probability theory and its applications*. 2d ed. New York, NY: John Wiley and Sons, Inc. Vol. 1. 520 p.
9. Loève, M. 1955. *Probability theory*. Princeton, NJ: D Van Nostrand. 701 p.
10. Grusho, A., N. Grusho, E. Timonina, and S. Shorgin. 2014. Bezopasnye arkhitektury raspredelennykh sistem [Secure architecture of the distributed systems]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 24(3):18–31.

Received June 27, 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

Grusho Nick A. (b. 1982) — Candidate of Science (PhD) in physics and mathematics, senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Sciences and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; info@itake.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

Timonina Elena E. (b. 1952) — Doctor of Science in technology, professor, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Sciences and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; eltimon@yandex.ru

О СВЯЗИ РЕСУРСНЫХ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ С СЕТЯМИ ЭРЛАНГА*

В. А. Наумов¹, К. Е. Самуйлов²

Аннотация: Рассматривается модель многолинейной системы массового обслуживания (СМО) с потерями, вызванными нехваткой ресурсов, необходимых для обслуживания заявок. Принятая на обслуживание заявка занимает случайные объемы ресурсов нескольких типов с заданными функциями распределения. Случайные векторы, описывающие требования заявок к ресурсам, не зависят от процессов поступления и обслуживания заявок, независимы в совокупности и одинаково распределены. Интерес, как и в задаче Эрланга, представляет вычисление вероятности потери поступающей заявки из-за нехватки ресурсов. Показана связь между мультисервисными сетями Эрланга и ресурсными СМО, что позволяет решать задачу вычисления вероятности потерь в ресурсной СМО с помощью известных методов, разработанных для мультисервисных сетей.

Ключевые слова: мультисервисная сеть; сеть Эрланга; система массового обслуживания; ресурсная СМО; случайный объем ресурсов; вероятность потерь; решетчатая функция

DOI: 10.14357/19922264160302

1 Введение

Рассмотрим многолинейные СМО с потерями, разнотипными ресурсами и пуассоновским входящим потоком, которые функционируют следующим образом. Поступившая заявка теряется, если в момент поступления количество какого-либо требуемого ей ресурса превышает количество свободного ресурса этого типа либо если число обслуживаемых заявок достигло максимума. В момент начала обслуживания заявки суммарный объем свободного ресурса каждого типа уменьшается на величину ресурса, выделенного этой заявке. В момент окончания обслуживания заявки суммарный объем свободного ресурса каждого типа увеличивается на величину ресурса, выделенного этой заявке при поступлении.

Ресурсные СМО с пуассоновским входящим потоком исследуются давно. В [1] получены стационарные распределения числа заявок в системе и объема занятого ресурса для СМО с экспоненциальной функцией распределения длительности обслуживания и произвольной функцией распределения объемов ресурса. Эти результаты обобщены в [2] на СМО с произвольной функцией распределения длительностей обслуживания и в [3] на системы с множественными ресурсами. Дальнейшие обобщения, рассмотренные в [4, 5], включают системы, у которых время обслуживания заявки и объемы выделенных ей ресурсов являются зави-

симыми случайными величинами. В [4] получено стационарное распределение и вероятность потери для СМО с произвольной совместной функцией распределения длительности обслуживания и объема единственного ресурса. Эти результаты обобщены в [5] на СМО, в которых каждая заявка характеризуется тремя зависимыми случайными признаками: числом приборов, необходимых для обслуживания, объемом ресурса и временем обслуживания. Применение модели ресурсных СМО к анализу вероятностных характеристик беспроводных гетерогенных сетей 5-го поколения было предложено в [6].

Хорошо изучены мультисервисные сети Эрланга [7, 8] с соединениями нескольких типов, в которых каждому соединению в каждом звене сети выделяется определенное число каналов, т.е. ресурсов сети, удерживаемое до завершения соединения.

В настоящей работе исследуется связь между мультисервисными сетями Эрланга и ресурсными СМО с арифметическими функциями распределения объемов требуемых ресурсов, которыми сколь угодно точно можно аппроксимировать любые функции распределения объемов ресурсов. Целью работы является исследование приближенного подхода к вычислению вероятностных характеристик ресурсных СМО. Для краткости будем опускать слово «мультисервисные» в названии сетей Эрланга.

*Работа выполнена при частичной финансовой поддержке РФФИ (проекты 16-07-00766, 15-07-03051 и 15-07-03608).

¹Исследовательский институт инноваций, г. Хельсинки, Финляндия, valeriy.naumov@pfi.fi

²Российский университет дружбы народов; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, ksam@sci.pfu.edu.ru

2 Сети Эрланга

Рассмотрим сеть массового обслуживания с потерями, состоящую из некоторого числа узлов, соединенных звеньями.

Пусть общее число звеньев сети равно M , емкость m -го звена равна N_m , $\mathbf{N} = (N_1, N_2, \dots, N_M)$ и $\mathbf{N}(\mathbf{n}) = \{\mathbf{i} \in \mathbf{N}^M | \mathbf{0} \leq \mathbf{i} \leq \mathbf{n}\}$, где \mathbf{N} — множество неотрицательных целых чисел. Между узлами сети могут быть установлены соединения L различных классов, каждый из которых однозначно характеризуется своими требованиями к емкости звеньев сети. Требование к числу каналов соединений l -го класса задается вектором $\mathbf{n}_l = (n_{l1}, n_{l2}, \dots, n_{lM})$, где n_{lj} есть число каналов, занимаемых соединением на j -м звене сети.

Предположим, что запросы на установление в сети соединения l -го класса образуют пуассоновский поток интенсивности λ_l , причем средняя продолжительность соединений l -го класса равна $b_l < \infty$. Если при поступлении запроса на установление нового соединения в сети недостаточно свободных каналов или уже установлено максимально возможное число K соединений, происходит блокировка запроса.

Обозначим $v_l(t)$ число соединений l -го класса, установленных в сети в момент t , $\mathbf{v}(t) = (v_1(t), v_2(t), \dots, v_L(t))$, и \mathbf{K} — пространство состояний процесса $\mathbf{v}(t)$, представляющее собой множество неотрицательных целочисленных векторов $\mathbf{k} = (k_1, k_2, \dots, k_L)$, удовлетворяющих неравенствам $k_1 + k_2 + \dots + k_L \leq K$ и $k_1 \mathbf{n}_1 + k_2 \mathbf{n}_2 + \dots + k_L \mathbf{n}_L \leq \mathbf{N}$.

Стационарное распределение $\phi(\mathbf{k}) = \lim_{t \rightarrow \infty} P\{\mathbf{v}(t) = \mathbf{k}\}$ процесса $\mathbf{v}(t)$ зависит от функций распределения продолжительности соединений лишь посредством средних значений и дается следующей формулой [7]:

$$\left. \begin{aligned} \phi(\mathbf{0}) &= \sum_{\mathbf{k} \in \mathbf{K}} \frac{(\lambda_1 b_1)^{k_1}}{k_1!} \dots \frac{(\lambda_L b_L)^{k_L}}{k_L!}; \\ \phi(\mathbf{k}) &= \phi(\mathbf{0}) \frac{(\lambda_1 b_1)^{k_1}}{k_1!} \dots \frac{(\lambda_L b_L)^{k_L}}{k_L!}, \quad \mathbf{k} \in \mathbf{K}. \end{aligned} \right\} (1)$$

Пусть $w_m(t)$ — число занятых в момент t каналов m -го звена сети и $\mathbf{w}(t) = (w_1(t), w_2(t), \dots, w_M(t))$. Зная распределение вероятностей $\phi(\mathbf{k})$, легко найти совместное распределение числа установленных в сети соединений и занятых ими каналов:

$$\begin{aligned} \psi_k(\mathbf{i}) &= \\ &= \lim_{t \rightarrow \infty} P\{v_1(t) + v_2(t) + \dots + v_L(t) = k, \mathbf{w}(t) = \mathbf{i}\} = \end{aligned}$$

$$\begin{aligned} &= \sum_{\substack{\mathbf{k} \in \mathbf{K} \\ k_1 + \dots + k_L = k \\ k_1 \mathbf{n}_1 + \dots + k_L \mathbf{n}_L = \mathbf{i}}} \phi(\mathbf{k}) = \phi(\mathbf{0}) \sum_{\substack{\mathbf{k} \in \mathbf{K} \\ k_1 + \dots + k_L = k \\ k_1 \mathbf{n}_1 + \dots + k_L \mathbf{n}_L = \mathbf{i}}} \frac{(\lambda_1 b_1)^{k_1}}{k_1!} \dots \frac{(\lambda_L b_L)^{k_L}}{k_L!} = \\ &= \phi(\mathbf{0}) \frac{\rho^k}{k!} \sum_{\substack{\mathbf{k} \in \mathbf{K} \\ k_1 + \dots + k_L = k \\ k_1 \mathbf{n}_1 + \dots + k_L \mathbf{n}_L = \mathbf{i}}} \frac{k!}{k_1! \dots k_L!} \left(\frac{\lambda_1 b_1}{\rho}\right)^{k_1} \dots \left(\frac{\lambda_L b_L}{\rho}\right)^{k_L}, \\ &\quad \mathbf{i} \in \mathbf{N}(\mathbf{N}), \quad k = 0, 1, \dots, K, \quad (2) \end{aligned}$$

где $\rho = \lambda_1 b_1 + \lambda_2 b_2 + \dots + \lambda_L b_L$. Формула (2) допускает простую вероятностную интерпретацию. Рассмотрим распределение вероятностей $\pi(\mathbf{i})$, $\mathbf{i} \in \mathbf{N}(\mathbf{N})$, случайного целочисленного вектора, с положительной вероятностью принимающего лишь значения из подмножества $\mathbf{L} = \{\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_L\}$ множества $\mathbf{N}(\mathbf{N})$, причем значение \mathbf{n}_l принимается с вероятностью $\lambda_l b_l / \rho$, т. е.

$$\pi(\mathbf{i}) = \begin{cases} \frac{\lambda_l b_l}{\rho}, & \mathbf{i} = \mathbf{n}_l; \\ 0, & \mathbf{i} \neq \mathbf{n}_l. \end{cases} \quad (3)$$

С учетом очевидного равенства $\varphi(\mathbf{0}) = \psi_0(\mathbf{0})$ формулу (2) можно записать следующим образом:

$$\psi_k(\mathbf{i}) = \psi_0(\mathbf{0}) \pi^{(k)}(\mathbf{i}) \frac{\rho^k}{k!}, \quad \mathbf{i} \in \mathbf{N}(\mathbf{N}), \quad (4)$$

где $\pi^{(k)}(\mathbf{i})$ есть k -кратная свертка распределения вероятностей (3).

Положим $C(\mathbf{r}) = 0$, если вектор \mathbf{r} не является неотрицательным, а для неотрицательных векторов \mathbf{r} определим величины $C(\mathbf{r})$ следующим образом:

$$C(\mathbf{r}) = 1 + \sum_{k=1}^K \frac{\rho^k}{k!} \sum_{\mathbf{i} \in \mathbf{N}(\mathbf{r})} \pi^{(k)}(\mathbf{i}). \quad (5)$$

Эти величины играют роль нормировочных констант для распределений вероятностей (1) и (2), поскольку справедливо равенство:

$$\phi(\mathbf{0}) = \psi_0(\mathbf{0}) = C(\mathbf{N})^{-1}. \quad (6)$$

Условная вероятность блокировки запроса на установление соединения при условии, что запрашивается соединение l -го класса, дается следующей формулой [7]:

$$B_l = 1 - \frac{C(\mathbf{N} - \mathbf{n}_l)}{C(\mathbf{N})}.$$

Отсюда вытекает выражение для безусловной вероятности блокировки запроса на установление соединения:

$$B = 1 - \frac{1}{C(\mathbf{N})} \sum_{l=1}^L \frac{\lambda_l}{\lambda} C(\mathbf{N} - \mathbf{n}_l), \quad (7)$$

где $\lambda = \lambda_1 + \lambda_2 + \dots + \lambda_L$ есть интенсивность суммарного потока поступающих запросов.

3 Ресурсная система массового обслуживания

Рассмотрим ресурсную СМО с ресурсами M типов, на которую поступает пуассоновский поток заявок с параметром λ . Обозначим R_m общий объем ресурса типа m и $\mathbf{R} = (R_1, R_2, \dots, R_M)$. Поступившая j -я заявка характеризуется длительностью обслуживания s_j и вектором объемов необходимых ей ресурсов $\mathbf{r}_j = (r_{j1}, r_{j2}, \dots, r_{jM})$. Случайные векторы (s_j, \mathbf{r}_j) , $j = 1, 2, \dots$, не зависят от моментов поступления заявок, независимы в совокупности и имеют одинаковую совместную функцию распределения $H(t, \mathbf{x}) = P\{s_j \leq t, \mathbf{r}_j \leq \mathbf{x}\}$. Обозначим через $F(\mathbf{x}) = P\{\mathbf{r}_j \leq \mathbf{x}\}$ функцию распределения объемов требуемых заявке ресурсов, $B(t) = P\{s_j \leq t\}$ — функцию распределения длительности обслуживания, $b < \infty$ — среднюю длительность обслуживания и $\rho = \lambda b$. Для простоты будем считать, что $F(\mathbf{R}) = 1$, т.е. требование заявкой любого ресурса не превосходит его общего объема.

Состояние рассматриваемой системы в момент t можно описать случайным процессом $X(t) = (\xi(t), \gamma(t))$. Здесь $\xi(t)$ — число заявок в системе и $\gamma(t) = (\gamma_1(t), \dots, \gamma_M(t))$ — вектор объемов занимаемых ими ресурсов. При условии конечности среднего времени обслуживания b стационарное распределение процесса $X(t)$

$$\left. \begin{aligned} p_0 &= \lim_{t \rightarrow \infty} P\{\xi(t) = 0\}; \\ P_k(\mathbf{x}) &= \lim_{t \rightarrow \infty} P\{\xi(t) = k; \gamma(t) \leq \mathbf{x}\}, \\ &0 \leq \mathbf{x} \leq \mathbf{R}, \quad k = 0, 1, \dots, K, \end{aligned} \right\} \quad (8)$$

имеет следующий вид:

$$\left. \begin{aligned} p_0 &= \left(1 + \sum_{k=1}^K G^{(k)}(\mathbf{R}) \frac{\rho^k}{k!}\right)^{-1}; \\ P_k(\mathbf{x}) &= p_0 G^{(k)}(\mathbf{x}) \frac{\rho^k}{k!}, \quad k = 1, 2, \dots, K. \end{aligned} \right\} \quad (9)$$

Здесь K — максимальное число заявок в системе и $G^{(k)}(\mathbf{x})$ — k -кратная свертка функции распределения

$$G(\mathbf{x}) = \frac{1}{b} \int_0^\infty \int_{\mathbf{y} \leq \mathbf{x}} t H(dt, d\mathbf{y}). \quad (10)$$

Из формул (8), в частности, вытекает следующее выражение для вероятности потери заявки в ресурсной СМО:

$$B = 1 - p_0 \left(1 + \sum_{k=1}^{K-1} \left(G^{(k)} * F\right)(\mathbf{R}) \frac{\rho^k}{k!}\right),$$

где $(G^{(k)} * F)(\mathbf{x})$ есть свертка функций распределения $G^{(k)}(\mathbf{x})$ и $F(\mathbf{x})$.

Справедливость формул (9) можно установить путем очевидного обобщения на произвольное число ресурсов результатов работы [5]. Ниже будет показано, как для вычисления стационарных характеристик ресурсных СМО с множественными дискретными ресурсами [6] можно использовать сети Эрланга. Попутно будут доказаны формулы (9) для решетчатых функций распределения объемов ресурсов $F(\mathbf{x})$.

Определение (10) функции распределения $G(\mathbf{x})$ станет понятнее, если ввести условное среднее время обслуживания заявки $b(\mathbf{x})$ при условии, что вектор объемов необходимых ей ресурсов равен \mathbf{x} . Это условное среднее время обслуживания можно определить как функцию, которая при любом векторе \mathbf{x} удовлетворяет следующему уравнению [9]:

$$\int_{\mathbf{y} \leq \mathbf{x}} \int_0^\infty t H(dt, d\mathbf{y}) = \int_{\mathbf{y} \leq \mathbf{x}} b(\mathbf{y}) D(d\mathbf{y}). \quad (11)$$

Используя равенство (11), выражение (10) для функции $G(\mathbf{x})$ можно переписать следующим образом:

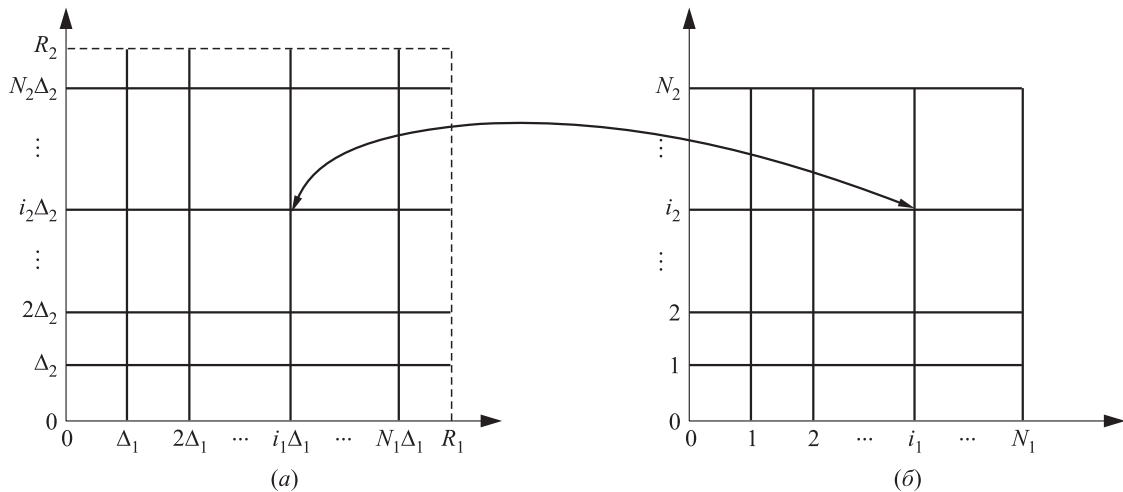
$$G(\mathbf{x}) = \frac{1}{b} \sum_{\mathbf{y} \leq \mathbf{x}} b(\mathbf{y}) F(d\mathbf{y}). \quad (12)$$

4 Связь ресурсных систем массового обслуживания с сетями Эрланга

В дальнейшем будем считать, что функция распределения требуемых объемов ресурсов $F(\mathbf{x})$ является решетчатой с некоторыми координатными шагами $\Delta_1, \Delta_2, \dots, \Delta_M > 0$ и положим $\mathbf{N} = (N_1, N_2, \dots, N_M)$, где N_m — целая часть числа R_m / Δ_m . В этом случае векторы требуемых объемов ресурсов \mathbf{r}_j с положительной вероятностью могут принимать лишь значения вида $\mathbf{Dk} = (k_1 \Delta_1, k_2 \Delta_2, \dots, k_M \Delta_M)$ где $\mathbf{k} = (k_1, k_2, \dots, k_M)$ — целочисленный вектор, а \mathbf{D} — диагональная матрица с элементами Δ_i на диагонали. Поэтому вместо вероятностей $P_k(\mathbf{x})$ удобнее рассматривать дискретное распределение вероятностей

$$p_k(\mathbf{i}) = \lim_{t \rightarrow \infty} P\{\xi(t) = k; \gamma(t) = \mathbf{Di}\}, \quad \mathbf{i} \in \mathbf{N}(\mathbf{N}), \quad k = 1, 2, \dots, K, \quad (13)$$

зная которое, легко вычислить вероятности (8):



Состояния процесса $\gamma(t)$, объема ресурсов, занятых в ресурсной СМО (а) и состояния процесса $w(t)$, числа занятых каналов в сети Келли (б)

$$\left. \begin{aligned} p_0 &= p_0(\mathbf{0}); \\ P_k(\mathbf{x}) &= \sum_{\substack{\mathbf{i} \in \mathbf{N}(\mathbf{N}) \\ \mathbf{D}\mathbf{i} \leq \mathbf{x}}} p_k(\mathbf{i}), \\ \mathbf{0} \leq \mathbf{x} \leq \mathbf{R}, \quad k &= 1, 2, \dots, K. \end{aligned} \right\} \quad (14)$$

Обозначим $f(\mathbf{i})$ вероятность того, что j -й заявке требуется вектор объемов ресурсов $\mathbf{r}_j = \mathbf{D}\mathbf{i}$, и $b(\mathbf{i})$ — условную среднюю длительность обслуживания j -й заявки при условии, что $\mathbf{r}_j = \mathbf{D}\mathbf{i}$. Перенумеруем все элементы множества $L = \{\mathbf{i} \in \mathbf{N}(\mathbf{N}) | f(\mathbf{i}) > 0\}$ и обозначим $\mathbf{n}_l = (n_{l1}, n_{l2}, \dots, n_{lM})$ его l -й элемент. В этих обозначениях функции распределения $F(\mathbf{x})$ и $G(\mathbf{x})$ могут быть записаны следующим образом:

$$F(\mathbf{x}) = \sum_{\substack{\mathbf{i} \in L \\ \mathbf{D}\mathbf{i} \leq \mathbf{x}}} f(\mathbf{i}); \quad G(\mathbf{x}) = \frac{1}{b} \sum_{\substack{\mathbf{i} \in L \\ \mathbf{D}\mathbf{i} \leq \mathbf{x}}} f(\mathbf{i})b(\mathbf{i}). \quad (15)$$

Таким образом, $G(\mathbf{x})$ является функцией распределения некоторого случайного вектора, принимающего значение $\mathbf{D}\mathbf{i}$ с вероятностью

$$g(\mathbf{i}) = \frac{1}{b} f(\mathbf{i})b(\mathbf{i}), \quad \mathbf{i} \in L. \quad (16)$$

Введем в рассмотрение вспомогательную сеть Эрланга, поведение которой во времени синхронизировано с поведением исходной ресурсной СМО. В момент поступления в СМО заявки, требующей вектор ресурсов $\mathbf{D}\mathbf{n}_l$ и обслуживание в течение времени τ , во вспомогательную сеть Эрланга поступает запрос на установление соединения класса l продолжительностью τ , а вектор \mathbf{n}_l задает требование к числу каналов этого соединения. Моменты поступления заявок в СМО являются моментами установления соединений в сети, а моменты ухода

заявок из СМО являются моментами разъединения соединений, и только они. Кроме того, если поступившая в СМО заявка теряется, то запрос на установление соответствующего соединения в сети также теряется. Ясно, что запросы на установление во вспомогательной сети соединений l -го класса образуют пуассоновский поток интенсивности $\lambda_l = \lambda f(\mathbf{n}_l)$, средняя продолжительность соединений l -го класса равна $b_l = b(\mathbf{n}_l)$, а распределение вероятностей (3) совпадает с распределением вероятностей (16), т. е. $\pi(\mathbf{i}) = g(\mathbf{i}), \mathbf{i} \in L$.

Нетрудно видеть, что между процессами $\xi(t)$ и $\gamma(t) = (\gamma_1(t), \dots, \gamma_M(t))$, описывающими поведение ресурсной СМО, и процессами $\mathbf{v}(t) = (v_1(t), v_2(t), \dots, v_L(t))$ и $\mathbf{w}(t) = (w_1(t), w_2(t), \dots, w_M(t))$, описывающими вспомогательную сеть Эрланга, существует простая связь, проиллюстрированная на рисунке:

$$\xi(t) = v_1(t) + v_2(t) + \dots + v_L(t), \quad \gamma(t) = \mathbf{D}\mathbf{w}(t).$$

Поэтому для распределений вероятностей (2) и (13) имеет место равенство:

$$p_k(\mathbf{i}) = \psi_k(\mathbf{i}), \quad \mathbf{i} \in \mathbf{N}(\mathbf{N}), \quad k = 0, 1, \dots, K,$$

и распределение вероятностей (2) можно записать в следующем виде:

$$p_k(\mathbf{i}) = p_0(\mathbf{0})g^{(k)}(\mathbf{i})\frac{\rho^k}{k!}, \quad \mathbf{i} \in \mathbf{N}(\mathbf{N}), \quad k = 0, 1, \dots, K,$$

где $g^{(k)}(\mathbf{i})$ есть k -кратная свертка распределения вероятностей (16). Используя равенства (6) и (7), вероятности простоя и потери заявки можно выразить через нормировочные константы (5):

$$p_0 = C(\mathbf{N})^{-1}; \quad B = 1 - \frac{1}{C(\mathbf{N})} \sum_{\mathbf{i} \in \mathbf{N}(\mathbf{N})} C(\mathbf{i})f(\mathbf{N} - \mathbf{i}).$$

Кроме того, из формул (4), (5), (14) и (15) вытекают следующие выражения для стационарного распределения числа заявок в системе и объемов занимаемых ими ресурсов:

$$\begin{aligned} p_0^{-1} &= 1 + \sum_{k=1}^K \frac{\rho^k}{k!} \sum_{\mathbf{i} \in \mathbf{N}(\mathbf{N})} g^{(k)}(\mathbf{i}) = \\ &= 1 + \sum_{k=1}^K \frac{\rho^k}{k!} \sum_{\substack{\mathbf{i} \in \mathbf{N}(\mathbf{N}) \\ \mathbf{D}\mathbf{i} \leq \mathbf{R}}} g^{(k)}(\mathbf{i}) = 1 + \sum_{k=1}^K G^{(k)}(\mathbf{R}) \frac{\rho^k}{k!}; \end{aligned}$$

$$P_k(\mathbf{x}) = p_0 \sum_{\substack{\mathbf{i} \in \mathbf{N}(\mathbf{N}) \\ \mathbf{D}\mathbf{i} \leq \mathbf{x}}} g^{(k)}(\mathbf{i}) \frac{\rho^k}{k!} = p_0 G^{(k)}(\mathbf{x}) \frac{\rho^k}{k!},$$

$$k = 1, 2, \dots, K.$$

Таким образом, для решетчатых функций распределения требуемых объемов ресурсов $F(\mathbf{x})$ доказаны формулы (9).

5 Заключение

В работе показано, что каждой СМО с арифметической функцией распределения множественных ресурсов соответствует некоторая вспомогательная сеть Эрланга. Поскольку стационарные распределения случайных процессов, описывающих ресурсную СМО и соответствующую ей сеть Эрланга, связаны простыми равенствами, для анализа ресурсных СМО имеется принципиальная возможность применения известных алгоритмов для анализа сетей Эрланга.

Соответствующие ресурсным СМО сети Эрланга имеют одну особенность, не характерную для типичных сетей Эрланга — число классов соединений вспомогательной сетей Эрланга может быть очень большим. Так, если $f(\mathbf{i}) > 0$ для всех $\mathbf{i} \in \mathbf{N}(\mathbf{N})$, то множество классов соединений сетей Эрланга состоит из всех целочисленных векторов

$\mathbf{i} = (i_1, i_2, \dots, i_M)$ в интервале $\mathbf{0} \leq \mathbf{i} \leq \mathbf{N}$ и число классов соединений вспомогательной сети равно произведению $(N_1 + 1)(N_2 + 1) \cdots (N_M + 1)$. Поэтому необходимы дополнительные исследования для того, чтобы выяснить, какие точные и приближенные методы анализа сетей Эрланга применимы к ресурсным СМО.

Литература

1. Ромм Э. Л., Скитович В. В. Об одном обобщении задачи Эрланга // Автоматика и телемеханика, 1971. № 6. С. 164–167.
2. Тихоненко О. М. Определение характеристик систем обслуживания с ограниченной памятью // Автоматика и телемеханика, 1997. № 6. С. 105–110.
3. Наумов В. А., Самуйлов К. Е., Самуйлов А. К. О суммарном объеме ресурсов, занимаемых обслуживаемыми заявками // Автоматика и телемеханика, 2016. № 8. С. 105–110, 125–135.
4. Тихоненко О. М., Климович К. Г. Анализ систем обслуживания требований случайной длины при ограниченном суммарном объеме // Проблемы передачи информации, 2001. Т. 37. Вып. 1. С. 78–88.
5. Тихоненко О. М. Обобщенная задача Эрланга для систем обслуживания с ограниченным суммарным объемом // Проблемы передачи информации, 2005. Т. 41. Вып. 3. С. 64–75.
6. Naumov V., Samouylov K., Yarkina N., Sopin E., Andreev S., Samuylov A. LTE performance analysis using queuing systems with finite resources and random requirements // 7th Congress (International) on Ultra Modern Telecommunications and Control Systems ICUMT-2015 Proceedings. — Piscataway, NJ, USA: IEEE, 2015. P. 100–103.
7. Kelly F. P. Loss networks // Ann. App. Probab., 1991. Vol. 1. No. 3. P. 319–378.
8. Наумов В. А., Самуйлов К. Е., Гайдамака Ю. В. Мультипликативные решения конечных цепей Маркова. — М.: РУДН, 2015. 159 с.
9. Гихман И. И., Скороход А. В. Теория случайных процессов. — М.: Наука, 1971. Т. 1. 664 с.

Поступила в редакцию 29.07.16

ON RELATIONSHIP BETWEEN QUEUING SYSTEMS WITH RESOURCES AND ERLANG NETWORKS

V. A. Naumov¹ and K. E. Samouylov^{2,3}

¹Service Innovation Research Institute, 30 D Lönnrotinkatu, Helsinki 00180, Finland

²Peoples' Friendship University of Russia, 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation

³Institute of Informatics Problems, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

Abstract: The paper considers a model of a multiserver queuing system (QS) with losses caused by the lack of resources required to service customers. During its service, each customer occupies a particular amount of resources of several types. Random vectors, describing the requirements of customers to resources, do not depend on the arrival process and service times and are mutually independent and identically distributed with the general cumulative distribution function. Like in the Erlang problem, the task is to calculate the probability of losses of an arriving customer caused by the lack of resources. The paper shows the relationship between multiservice loss networks and queuing systems with resources, which makes it possible to solve the problem of calculating the loss probability in the queuing systems with resources using known methods developed for multiservice loss networks.

Keywords: multiservice network; Erlang network; queuing system; queuing system with resources; random amount of resources; loss probability; arithmetic probability distribution

DOI: 10.14357/19922264160302

Acknowledgments

The work was partly supported by the Russian Foundation for Basic Research (projects 16-07-00766, 15-07-03051, and 15-07-03608).

References

1. Romm, E., and V. Skitovitch. 1971. On certain generalization of problem of Erlang. *Automation Remote Control* 32(6):1000–1003.
2. Tikhonenko, O. 1997. The determination of service characteristics under limited memory. *Automation Remote Control* 58(6):969–973.
3. Naumov, V., K. Samouylov, and A. Samuylov. 2016. On the total amount of resources occupied by serviced customers. *Automation Remote Control* 77(8): 1419–1427.
4. Tikhonenko, O., and K. Klimovitch. 2001. Analysis of queuing systems for random-length arrivals with limited cumulative volume. *Problems Information Transmission* 37(1):70–79.
5. Tikhonenko, O. 2005. Generalized Erlang problem for queueing systems with bounded total size. *Problems Information Transmission* 41(3):243–253.
6. Naumov, V., K. Samouylov, N. Yarkina, E. Sopin, S. Andreev, and A. Samuylov. 2015. LTE performance analysis using queuing systems with finite resources and random requirements. *7th Congress (International) on Ultra Modern Telecommunications and Control Systems ICUMT-2015 Proceedings*. Piscataway, NJ: IEEE. 100–103.
7. Kelly, F. P. 1991. Loss networks. *Ann. App. Probab.* 1(3):319–378.
8. Naumov, V., Yu. Gaidamaka, and K. Samouylov. 2015. *Mul'tiplikativnye resheniya konechnykh tsepey Markova* [Product form solutions for finite Markov chains]. Moscow: RUDN. 159 p.
9. Gihman, I., and A. Skorohod. 1971. *The theory of stochastic processes*. New York – Heidelberg – Berlin: Springer-Verlag. Vol. I, 1974, 574 p.

Received July 29, 2016

Contributors

Naumov Valeriy A. (b. 1950) — Candidate of Science (PhD) in physics and mathematics, Research Director, Service Innovation Research Institute, 30 D Lönnrotinkatu, Helsinki 00180, Finland; valeriy.naumov@pfu.fi

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; 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

АНАЛИЗ СИСТЕМЫ ОБСЛУЖИВАНИЯ С ВХОДЯЩИМ ПОТОКОМ АВТОРЕГРЕССИОННОГО ТИПА И ОТНОСИТЕЛЬНЫМ ПРИОРИТЕТОМ*

Н. Д. Леонтьев¹, В. Г. Ушаков²

Аннотация: Рассматривается одноканальная система массового обслуживания с неограниченным числом мест для ожидания, в которую поступают два потока требований: первый поток — пуассоновский, а второй — неординарный пуассоновский (т. е. пуассоновский поток групп требований). Требования из первого потока имеют относительный приоритет перед требованиями второго потока. Особенностью системы является авторегрессионная зависимость размеров групп требований второго потока: размер n -й поступившей в систему группы требований либо с некоторой фиксированной вероятностью равен размеру $(n - 1)$ -й поступившей в систему группы требований, либо с дополнительной вероятностью является независимой от него случайной величиной. Длительности обслуживания требований каждого потока являются независимыми случайными величинами с произвольным распределением. Найдена производящая функция совместного распределения числа требований каждого потока в системе в произвольный момент времени.

Ключевые слова: теория массового обслуживания; нестационарный режим; системы с групповым поступлением требований; относительный приоритет

DOI: 10.14357/19922264160303

1 Введение

При моделировании передачи данных в телекоммуникационных сетях важно учитывать природу потоков информации и характеристики потоков в зависимости от приложений. В простейших моделях предполагается, что все пакеты информации имеют фиксированную конечную длину, а размеры пакетов независимы. Однако в ряде случаев необходимо рассматривать более сложные конструкции входящего потока, которые позволяют учитывать неоднородную и коррелированную природу потоков данных. В данной работе рассматривается система массового обслуживания типа $M|G|1$ с двумя потоками требований, один из которых является приоритетным, а другой — коррелированным потоком групп требований.

Настоящая работа обобщает результаты статьи [1] на случай двух потоков, требования одного из которых имеют относительный приоритет.

2 Описание системы

Рассматривается одноканальная система обслуживания, в которую поступают два потока требований:

- первый поток — пуассоновский с интенсивностью a_1 ,
- второй поток является пуассоновским потоком групп требований.

Интенсивность поступления групп требований равна a_2 . Группа состоит из случайного числа требований и содержит k требований с вероятностью h_k , $k = 1, \dots, M$. Между размерами двух последовательно поступающих групп требований имеется следующая зависимость: размер n -й поступающей группы требований либо с вероятностью $0 \leq p < 1$ равен размеру $(n - 1)$ -й группы, либо с вероятностью $1 - p$ является независимой от него случайной величиной. Требования из первого потока имеют относительный приоритет по отношению к требованиям из второго потока. Иными словами, прерывание уже начатого обслуживания не допускается и требования из второго потока могут поступать на обслуживание только при отсутствии в очереди требований из первого потока. Будем считать, что число мест для ожидания не ограничено, а длительность обслуживания требований i -го потока имеет функцию распределения $B_i(x)$ с плотностью $b_i(x)$ и преобразованием Лапласа $\beta_i(s)$.

* Работа выполнена при финансовой поддержке РФФИ (проект 15-07-02354).

¹ Факультет вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова, ndleontyev@gmail.com

² Факультет вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова; Институт проблем информатики ФИЦ ИУ РАН, vgushakov@mail.ru

Определим следующие случайные процессы:

$I(t)$ — номер потока, требование из которого обслуживается в момент t ;

$L_i(t)$ — число требований i -го потока в системе в момент времени t ;

$X(t)$ — время, прошедшее с начала обслуживания требования, находящегося на обслуживании в момент t ;

$N(t)$ — размер последней поступившей в систему до момента t группы требований 2-го потока.

Положим

$$P_i(n_1, n_2, k, x, t) = \frac{\partial}{\partial x} \mathbf{P}(I(t) = i, L_1(t) = n_1, L_2(t) = n_2, N(t) = k, X(t) < x);$$

$$P(0, k, t) = \mathbf{P}(L_1(t) = 0, L_2(t) = 0, N(t) = k).$$

Обозначим

$$\pi_1(z_1, z_2, k, x, s) = \sum_{n_1=1}^{\infty} z_1^{n_1} \sum_{n_2=0}^{\infty} z_2^{n_2} \int_0^{\infty} e^{-st} P_1(n_1, n_2, k, x, t) dt;$$

$$\pi_2(z_1, z_2, k, x, s) = \sum_{n_1=0}^{\infty} z_1^{n_1} \sum_{n_2=1}^{\infty} z_2^{n_2} \int_0^{\infty} e^{-st} P_2(n_1, n_2, k, x, t) dt;$$

$$\pi_0(k, s) = \int_0^{\infty} e^{-st} P(0, k, t) dt$$

при $|z_1| \leq 1, |z_2| \leq 1, \operatorname{Re}(s) > 0$.

3 Основные результаты

Соотношения для определения $\pi_i(z_1, z_2, k, x, s)$ содержатся в следующей основной теореме.

Теорема. *Функции $\pi_i(z_1, z_2, k, x, s)$ при $|z_1| < 1, |z_2| < 1, k = 1, \dots, M, x \geq 0, \operatorname{Re}(s) > 0$ определяются по следующим формулам:*

$$\pi_i(z_1, z_2, k, x, s) = (1 - B_i(x)) \sum_{n=1}^M C_{i,n}(z_1, z_2, s) \times$$

$$\times \frac{(1-p)h_k a_2 z_2^k}{\tilde{\lambda}_n(z_2) - p a_2 z_2^k} \exp\left(-\left(s + a_1 - a_1 z_1 + a_2 - \tilde{\lambda}_n(z_2)\right)x\right),$$

где

$$C_{1,n}(z_1, z_2, s) = \frac{1}{1 - z_1^{-1} \beta_1 (s + a_1 - a_1 z_1 + a_2 - \tilde{\lambda}_n(z_2))} \times$$

$$\times \frac{\prod_{i=1}^M (\tilde{\lambda}_n(z_2) - p a_2 z_2^i)}{\prod_{\substack{j=1 \\ j \neq n}}^M (\tilde{\lambda}_n(z_2) - \tilde{\lambda}_j(z_2))} \sum_{m=1}^M \frac{1}{\tilde{\lambda}_n(z_2) - p a_2 z_2^m} \times$$

$$\times \left(b_m(z_1, z_2, s) - \frac{1 - z_2^{-1} \beta_2 (s + a_1 - a_1 z_1 + a_2 - \tilde{\lambda}_n(z_2))}{1 - z_2^{-1} \beta_2 (s + a_1 - a_1 z_{1,n}(z_2, s) + a_2 - \tilde{\lambda}_n(z_2))} \times \right.$$

$$\left. \times b_m(z_{1,n}(z_2, s), z_2, s) \right);$$

$$C_{2,n}(z_1, z_2, s) = \frac{1}{1 - z_2^{-1} \beta_2 (s + a_1 - a_1 z_{1,n}(z_2, s) + a_2 - \tilde{\lambda}_n(z_2))} \times$$

$$\times \frac{\prod_{i=1}^M (\tilde{\lambda}_n(z_2) - p a_2 z_2^i)}{\prod_{\substack{j=1 \\ j \neq n}}^M (\tilde{\lambda}_n(z_2) - \tilde{\lambda}_j(z_2))} \sum_{m=1}^M \frac{b_m(z_{1,n}(z_2, s), z_2, s)}{\tilde{\lambda}_n(z_2) - p a_2 z_2^m};$$

$$b_m(z_1, z_2, s) = -(s + a_1 - a_1 z_1 + a_2) \pi_0(m, s) + h_m +$$

$$+ \left[p \pi_0(m, s) a_2 z_2^m + (1-p) \sum_{n=1}^M \pi_0(n, s) h_m a_2 z_2^m \right];$$

$\tilde{\lambda}_1(z_2), \dots, \tilde{\lambda}_M(z_2)$ определяются из уравнения

$$\prod_{i=1}^M (p a_2 z_2^i - \tilde{\lambda}) +$$

$$+ \sum_{i=1}^M (1-p) h_i a_2 z_2^i \prod_{\substack{j=1 \\ j \neq i}}^M (p a_2 z_2^j - \tilde{\lambda}) = 0,$$

а $\pi_0(1, s), \dots, \pi_0(M, s)$ — из системы

$$\sum_{m=1}^M \prod_{\substack{j=1 \\ j \neq m}}^M (\tilde{\lambda}_n - p a_2 z_{2,n}^j) \left(- (s + a_1 - a_1 z_{1,n} + \right.$$

$$\left. + a_2) \pi_0(m, s) + h_m + \tilde{\lambda}_n(z_{2,n}) \pi_0(m, s) \right) = 0,$$

¹ В случае, когда система свободна, можно для определенности положить $I(t) = 0$ и $X(t) = 0$.

в которой $z_{1,n} = z_{1,n}(z_{2,n}, s)$ и $z_{1,n}(z_2, s)$ — решение функционального уравнения

$$z_{1,n} = \beta_1 \left(s + a_1 - a_1 z_{1,n} + a_2 - \tilde{\lambda}_n(z_2) \right),$$

а $z_{2,n} = z_{2,n}(s)$ — решение функционального уравнения

$$z_{2,n} = \beta_2 \left(s + a_1 - a_1 z_{1,n}(z_{2,n}, s) + a_2 - \tilde{\lambda}_n(z_{2,n}) \right).$$

Доказательство. Функции $P_i(n_1, n_2, k, x, t)$ и $P(0, k, t)$ удовлетворяют соотношениям:

$$\begin{aligned} P_1(n_1, n_2, k, x + \Delta, t + \Delta) &= \\ &= P_1(n_1, n_2, k, x, t) [1 - (a_1 + a_2 + \eta_1(x))\Delta] + \\ &+ \mathbf{1}_{\{n_1 > 1\}} P_1(n_1 - 1, n_2, k, x, t) a_1 \Delta + \\ &+ \mathbf{1}_{\{n_2 \geq k\}} \left[p P_1(n_1, n_2 - k, k, x, t) a_2 \Delta + \right. \\ &\left. + (1 - p) \sum_{m=1}^M P_1(n_1, n_2 - k, m, x, t) h_k a_2 \Delta \right]; \quad (1) \end{aligned}$$

$$\begin{aligned} P_2(n_1, n_2, k, x + \Delta, t + \Delta) &= \\ &= P_2(n_1, n_2, k, x, t) [1 - (a_1 + a_2 + \eta_2(x))\Delta] + \\ &+ \mathbf{1}_{\{n_1 \geq 1\}} P_2(n_1 - 1, n_2, k, x, t) a_1 \Delta + \\ &+ \mathbf{1}_{\{n_2 > k\}} \left[p P_2(n_1, n_2 - k, k, x, t) a_2 \Delta + \right. \\ &\left. + (1 - p) \sum_{m=1}^M P_2(n_1, n_2 - k, m, x, t) h_k a_2 \Delta \right]; \quad (2) \end{aligned}$$

$$\begin{aligned} P(0, k, t + \Delta) &= P(0, k, t) [1 - (a_1 + a_2) \Delta] + \\ &+ \int_0^\infty P_1(1, 0, k, x, t) \eta_1(x) dx \Delta + \\ &+ \int_0^\infty P_2(0, 1, k, x, t) \eta_2(x) dx \Delta; \quad (3) \end{aligned}$$

$$\begin{aligned} \int_0^\Delta P_1(n_1, n_2, k, u, t + \Delta) du &= \\ &= \int_0^\infty P_1(n_1 + 1, n_2, k, x, t) \eta_1(x) dx \Delta + \\ &+ \int_0^\infty P_2(n_1, n_2 + 1, k, x, t) \eta_2(x) dx \Delta + \\ &+ \delta_{n_1, 1} \delta_{n_2, 0} P(0, k, t) a_1 \Delta; \quad (4) \end{aligned}$$

$$\begin{aligned} \int_0^\Delta P_2(0, n_2, k, u, t + \Delta) du &= \\ &= \int_0^\infty P_1(1, n_2, k, x, t) \eta_1(x) dx \Delta + \\ &+ \int_0^\infty P_2(0, n_2 + 1, k, x, t) \eta_2(x) dx \Delta + \\ &+ \delta_{n_2, k} \left[p P(0, k, t) a_2 \Delta + \right. \\ &\left. + (1 - p) \sum_{m=1}^M P(0, m, t) h_k a_2 \Delta \right], \quad (5) \end{aligned}$$

где $\eta_i(x) = b_i(x)/(1 - B_i(x))$.

Переходя к пределу при $\Delta \rightarrow 0$ в (1)–(5), имеем:

$$\begin{aligned} \frac{\partial P_1(n_1, n_2, k, x, t)}{\partial t} + \frac{\partial P_1(n_1, n_2, k, x, t)}{\partial x} &= \\ &= -(a_1 + a_2 + \eta_1(x)) P_1(n_1, n_2, k, x, t) + \\ &+ \mathbf{1}_{\{n_1 > 1\}} P_1(n_1 - 1, n_2, k, x, t) a_1 + \\ &+ \mathbf{1}_{\{n_2 \geq k\}} \left[p P_1(n_1, n_2 - k, k, x, t) a_2 + \right. \\ &\left. + (1 - p) \sum_{m=1}^M P_1(n_1, n_2 - k, m, x, t) h_k a_2 \right]; \quad (6) \end{aligned}$$

$$\begin{aligned} \frac{\partial P_2(n_1, n_2, k, x, t)}{\partial t} + \frac{\partial P_2(n_1, n_2, k, x, t)}{\partial x} &= \\ &= -(a_1 + a_2 + \eta_2(x)) P_2(n_1, n_2, k, x, t) + \\ &+ \mathbf{1}_{\{n_1 \geq 1\}} P_2(n_1 - 1, n_2, k, x, t) a_1 + \\ &+ \mathbf{1}_{\{n_2 > k\}} \left[p P_2(n_1, n_2 - k, k, x, t) a_2 + \right. \\ &\left. + (1 - p) \sum_{m=1}^M P_2(n_1, n_2 - k, m, x, t) h_k a_2 \right]; \quad (7) \end{aligned}$$

$$\begin{aligned} \frac{\partial P(0, k, t)}{\partial t} &= -(a_1 + a_2) P(0, k, t) + \\ &+ \int_0^\infty P_1(1, 0, k, x, t) \eta_1(x) dx + \\ &+ \int_0^\infty P_2(0, 1, k, x, t) \eta_2(x) dx; \quad (8) \end{aligned}$$

$$\begin{aligned}
 P_1(n_1, n_2, k, 0, t) &= \\
 &= \int_0^\infty P_1(n_1 + 1, n_2, k, x, t) \eta_1(x) dx + \\
 &+ \int_0^\infty P_2(n_1, n_2 + 1, k, x, t) \eta_2(x) dx + \\
 &\quad + \delta_{n_1, 1} \delta_{n_2, 0} P(0, k, t) a_1; \quad (9)
 \end{aligned}$$

$$P_2(n_1, n_2, k, 0, t) = 0, \quad n_1 > 0; \quad (10)$$

$$\begin{aligned}
 P_2(0, n_2, k, 0, t) &= \int_0^\infty P_1(1, n_2, k, x, t) \eta_1(x) dx + \\
 &+ \int_0^\infty P_2(0, n_2 + 1, k, x, t) \eta_2(x) dx + \\
 &+ \delta_{n_2, k} \left[p P(0, k, t) a_2 + \right. \\
 &\quad \left. + (1 - p) \sum_{m=1}^M P(0, m, t) h_k a_2 \right]. \quad (11)
 \end{aligned}$$

Переходя в уравнениях (6)–(11) к производящим функциям и преобразованиям Лапласа по t , получим:

$$\begin{aligned}
 \frac{\partial \pi_i(z_1, z_2, k, x, s)}{\partial x} &= \\
 &= -(s + a_1 - a_1 z_1 + a_2 + \eta_i(x)) \pi_i(z_1, z_2, k, x, s) + \\
 &+ \left[p \pi_i(z_1, z_2, k, x, s) a_2 z_2^k + \right. \\
 &\quad \left. + (1 - p) \sum_{m=1}^M \pi_i(z_1, z_2, m, x, s) h_k a_2 z_2^k \right]; \\
 (s + a_1 + a_2) \pi_0(k, s) - h_k &= \\
 &= \int_0^\infty \int_0^\infty P_1(1, 0, k, x, t) \eta_1(x) dx e^{-st} dt + \\
 &+ \int_0^\infty \int_0^\infty P_2(0, 1, k, x, t) \eta_2(x) dx e^{-st} dt; \quad (12)
 \end{aligned}$$

$$\begin{aligned}
 \pi_1(z_1, z_2, k, 0, s) &= z_1^{-1} \int_0^\infty \pi_1(z_1, z_2, k, x, s) \eta_1(x) dx - \\
 &- \sum_{n_2=0}^\infty z_2^{n_2} \int_0^\infty \int_0^\infty P_1(1, n_2, k, x, t) \eta_1(x) dx e^{-st} dt + \\
 &+ z_2^{-1} \int_0^\infty \pi_2(z_1, z_2, k, x, s) \eta_2(x) dx -
 \end{aligned}$$

$$\begin{aligned}
 &- z_2^{-1} \sum_{n_2=1}^\infty z_2^{n_2} \int_0^\infty \int_0^\infty P_2(0, n_2, k, x, t) \eta_2(x) dx e^{-st} dt + \\
 &\quad + \pi_0(k, s) a z_1; \quad (13)
 \end{aligned}$$

$$\begin{aligned}
 \sum_{n_1=1}^\infty z_1^{n_1} \sum_{n_2=1}^\infty z_2^{n_2} \int_0^\infty P_2(n_1, n_2, k, 0, t) e^{-st} dt &= 0; \\
 &\quad (14)
 \end{aligned}$$

$$\begin{aligned}
 \sum_{n_2=1}^\infty z_2^{n_2} \int_0^\infty P_2(0, n_2, k, 0, t) e^{-st} dt &= \\
 &= \sum_{n_2=1}^\infty z_2^{n_2} \int_0^\infty \int_0^\infty P_1(1, n_2, k, x, t) \eta_1(x) dx e^{-st} dt + \\
 &+ z_2^{-1} \sum_{n_2=1}^\infty z_2^{n_2} \int_0^\infty \int_0^\infty P_2(0, n_2, k, x, t) \eta_2(x) dx e^{-st} dt - \\
 &- \int_0^\infty \int_0^\infty P_2(0, 1, k, x, t) \eta_2(x) dx e^{-st} dt + \\
 &+ \left[p \pi_0(k, s) a_2 z_2^k + (1 - p) \sum_{m=1}^M \pi_0(m, s) h_k a_2 z_2^k \right]. \quad (15)
 \end{aligned}$$

Суммируя уравнения (12)–(15), приходим к системе:

$$\begin{aligned}
 \frac{\partial \pi_i(z_1, z_2, k, x, s)}{\partial x} &= \\
 &= -(s + a_1 - a_1 z_1 + a_2 + \eta_i(x)) \pi_i(z_1, z_2, k, x, s) + \\
 &+ \left[p \pi_i(z_1, z_2, k, x, s) a_2 z_2^k + \right. \\
 &\quad \left. + (1 - p) \sum_{m=1}^M \pi_i(z_1, z_2, m, x, s) h_k a_2 z_2^k \right]; \quad (16)
 \end{aligned}$$

$$\begin{aligned}
 \pi_1(z_1, z_2, k, 0, s) + \pi_2(z_1, z_2, k, 0, s) &= \\
 &= z_1^{-1} \int_0^\infty \pi_1(z_1, z_2, k, x, s) \eta_1(x) dx + \\
 &+ z_2^{-1} \int_0^\infty \pi_2(z_1, z_2, k, x, s) \eta_2(x) dx - \\
 &- (s + a_1 - a_1 z_1 + a_2) \pi_0(k, s) + h_k + \\
 &+ \left[p \pi_0(k, s) a_2 z_2^k + (1 - p) \sum_{m=1}^M \pi_0(m, s) h_k a_2 z_2^k \right]. \quad (17)
 \end{aligned}$$

Обозначим

$$\pi_i(z_1, z_2, k, x, s) = (1 - B_i(x)) \tilde{\pi}_i(z_1, z_2, k, x, s).$$

В новых обозначениях (16) примет вид:

$$\begin{aligned} \frac{\partial \tilde{\pi}_i(z_1, z_2, k, x, s)}{\partial x} = & - (s + a_1 - a_1 z_1 + a_2) \tilde{\pi}_i(z_1, z_2, k, x, s) + \\ & + \left[\sum_{m=1}^M p \tilde{\pi}_i(z_1, z_2, k, x, s) a_2 z_2^k + \right. \\ & \left. + (1 - p) \sum_{m=1}^M \tilde{\pi}_i(z_1, z_2, m, x, s) h_m a_2 z_2^k \right]. \end{aligned} \quad (18)$$

Это линейная система дифференциальных уравнений первого порядка с постоянными коэффициентами, решение которой можно записать в виде:

$$\begin{aligned} \tilde{\pi}_i(z_1, z_2, k, x, s) = & \sum_{n=1}^M C_{i,n}(z_1, z_2, s) u_{kn}(z_2) \exp \left(- \left(s + a_1 - \right. \right. \\ & \left. \left. - a_1 z_1 + a_2 - \tilde{\lambda}_n(z_2) \right) x \right), \end{aligned} \quad (19)$$

где $\tilde{\lambda}_k(z_2) = \lambda_k(z_1, z_2, s) + (s + a_1 - a_1 z_1 + a_2)$, $k = 1, \dots, M$; $\lambda_1(z_1, z_2, s), \dots, \lambda_M(z_1, z_2, s)$ — собственные значения матрицы системы: $u_1(z_2) = (u_{11}(z_2), \dots, u_{M1}(z_2))^T, \dots, u_M(z_2) = (u_{1M}(z_2), \dots, u_{MM}(z_2))^T$ — соответствующие собственные векторы. Заметим, что матрицы систем дифференциальных уравнений для $\tilde{\pi}_1(z_1, z_2, k, x, s)$ и $\tilde{\pi}_2(z_1, z_2, k, x, s)$ одинаковы, а следовательно, собственные значения и собственные векторы в записи решений совпадают.

Функции $\tilde{\lambda}_1(z_2), \dots, \tilde{\lambda}_M(z_2)$ являются решениями характеристического уравнения:

$$\prod_{i=1}^M (p a_2 z_2^i - \tilde{\lambda}) + \sum_{i=1}^M (1-p) h_i a_2 z_2^i \prod_{\substack{j=1 \\ j \neq i}}^M (p a_2 z_2^j - \tilde{\lambda}) = 0. \quad (20)$$

Подставляя (19) в (18), находим:

$$u_{mn}(z_2) = \frac{(1-p) h_m a_2 z_2^m}{\tilde{\lambda}_n(z_2) - p a_2 z_2^m}. \quad (21)$$

Подставив (19) в (17), получим:

$$\begin{aligned} \sum_{n=1}^M C_{1,n}(z_1, z_2, s) u_{mn}(z_2) + \\ + \sum_{n=1}^M C_{2,n}(z_1, z_2, s) u_{mn}(z_2) = \end{aligned}$$

$$\begin{aligned} = z_1^{-1} \sum_{n=1}^M C_{1,n}(z_1, z_2, s) u_{mn}(z_2) \times \\ \times \beta_1 \left(s + a_1 - a_1 z_1 + a_2 - \tilde{\lambda}_n(z_2) \right) + \\ + z_2^{-1} \sum_{n=1}^M C_{2,n}(z_1, z_2, s) u_{mn}(z_2) \times \\ \times \beta_2 \left(s + a_1 - a_1 z_1 + a_2 - \tilde{\lambda}_n(z_2) \right) - \\ - (s + a_1 - a_1 z_1 + a_2) \pi_0(m, s) + h_m + \\ + \left[p \pi_0(m, s) a_2 z_2^m + (1-p) \sum_{n=1}^M \pi_0(n, s) h_n a_2 z_2^m \right]. \end{aligned}$$

Перепишем это уравнение в виде:

$$\begin{aligned} \sum_{n=1}^M \left[C_{1,n}(z_1, z_2, s) \left(1 - \right. \right. \\ \left. \left. - z_1^{-1} \beta_1 \left(s + a_1 - a_1 z_1 + a_2 - \tilde{\lambda}_n(z_2) \right) \right) + \right. \\ \left. + C_{2,n}(z_1, z_2, s) \left(1 - \right. \right. \\ \left. \left. - z_2^{-1} \beta_2 \left(s + a_1 - a_1 z_1 + a_2 - \tilde{\lambda}_n(z_2) \right) \right) \right] u_{mn}(z_2) = \\ = b_m(z_1, z_2, s), \end{aligned} \quad (22)$$

где

$$\begin{aligned} b_m(z_1, z_2, s) = & - (s + a_1 - a_1 z_1 + a_2) \pi_0(m, s) + h_m + \\ & + \left[p \pi_0(m, s) a_2 z_2^m + (1-p) \sum_{n=1}^M \pi_0(n, s) h_n a_2 z_2^m \right]. \end{aligned}$$

Подставим (21) в (22) и поделим обе части полученного уравнения на $(1-p) h_m a_2 z_2^m$:

$$\begin{aligned} \sum_{n=1}^M \left[C_{1,n}(z_1, z_2, s) \left(1 - \right. \right. \\ \left. \left. - z_1^{-1} \beta_1 \left(s + a_1 - a_1 z_1 + a_2 - \tilde{\lambda}_n(z_2) \right) \right) + \right. \\ \left. + C_{2,n}(z_1, z_2, s) \left(1 - \right. \right. \\ \left. \left. - z_2^{-1} \beta_2 \left(s + a_1 - a_1 z_1 + a_2 - \tilde{\lambda}_n(z_2) \right) \right) \right] \times \\ \times \frac{1}{\tilde{\lambda}_n(z_2) - p a_2 z_2^m} = \frac{b_m(z_1, z_2, s)}{(1-p) h_m a_2 z_2^m}. \end{aligned}$$

Это система линейных алгебраических уравнений с матрицей Коши. Ее решение записывается в виде¹:

¹Про обращение матриц Коши см. [2].

$$\begin{aligned}
 & C_{1,n}(z_1, z_2, s) \left(1 - \right. \\
 & \quad \left. - z_1^{-1} \beta_1 \left(s + a_1 - a_1 z_1 + a_2 - \tilde{\lambda}_n(z_2) \right) \right) + \\
 & \quad + C_{2,n}(z_1, z_2, s) \left(1 - \right. \\
 & \quad \left. - z_2^{-1} \beta_2 \left(s + a_1 - a_1 z_1 + a_2 - \tilde{\lambda}_n(z_2) \right) \right) = \\
 & = \frac{\prod_{i=1}^M (\tilde{\lambda}_n(z_2) - pa_2 z_2^i)}{\prod_{\substack{j=1 \\ j \neq n}}^M (\tilde{\lambda}_n(z_2) - \tilde{\lambda}_j(z_2))} \sum_{m=1}^M \frac{\prod_{\substack{i=1 \\ i \neq n}}^M (pa_2 z_2^m - \tilde{\lambda}_i(z_2))}{\prod_{\substack{j=1 \\ j \neq m}}^M (pa_2 z_2^m - pa_2 z_2^j)} \times \\
 & \quad \times \frac{b_m(z_1, z_2, s)}{(1-p)h_m a_2 z_2^m}.
 \end{aligned}$$

Далее, поскольку функции $\tilde{\lambda}_m(z_2)$, $m = 1, \dots, M$, являются решениями уравнения (20), можно записать:

$$\begin{aligned}
 & \prod_{i=1}^M (pa_2 z_2^i - \tilde{\lambda}) + \sum_{i=1}^M (1-p)h_i a_2 z_2^i \prod_{\substack{j=1 \\ j \neq i}}^M (pa_2 z_2^j - \tilde{\lambda}) = \\
 & = \prod_{j=1}^M (\tilde{\lambda}_j(z_2) - \tilde{\lambda}). \quad (23)
 \end{aligned}$$

Подставляя в (23) $\tilde{\lambda} = pa_2 z_2^m$, получим:

$$\begin{aligned}
 & (1-p)h_m a_2 z_2^m \prod_{\substack{j=1 \\ j \neq m}}^M (pa_2 z_2^j - pa_2 z_2^m) = \\
 & = \prod_{j=1}^M (\tilde{\lambda}_j(z_2) - pa_2 z_2^m).
 \end{aligned}$$

Отсюда

$$\begin{aligned}
 & C_{1,n}(z_1, z_2, s) \left(1 - \right. \\
 & \quad \left. - z_1^{-1} \beta_1 \left(s + a_1 - a_1 z_1 + a_2 - \tilde{\lambda}_n(z_2) \right) \right) + \\
 & \quad + C_{2,n}(z_1, z_2, s) \left(1 - \right. \\
 & \quad \left. - z_2^{-1} \beta_2 \left(s + a_1 - a_1 z_1 + a_2 - \tilde{\lambda}_n(z_2) \right) \right) = \\
 & = \frac{\prod_{i=1}^M (\tilde{\lambda}_n(z_2) - pa_2 z_2^i)}{\prod_{\substack{j=1 \\ j \neq n}}^M (\tilde{\lambda}_n(z_2) - \tilde{\lambda}_j(z_2))} \sum_{m=1}^M \frac{b_m(z_1, z_2, s)}{\tilde{\lambda}_n(z_2) - pa_2 z_2^m}. \quad (24)
 \end{aligned}$$

Рассмотрим уравнение:

$$z_1 = \beta_1 \left(s + a_1 - a_1 z_1 + a_2 - \tilde{\lambda}_n(z_2) \right). \quad (25)$$

Обе части уравнения являются аналитическими в области $|z_1| \leq 1$ функциями. Имеем:

$$\begin{aligned}
 & \left| \beta_1 \left(s + a_1 - a_1 z_1 + a_2 - \tilde{\lambda}_n(z_2) \right) \right| \leq \\
 & \leq \beta_1 \left(\operatorname{Re} \left(s + a_1 - a_1 z_1 + a_2 - \tilde{\lambda}_n(z_2) \right) \right) \leq \\
 & \leq \beta_1 (\operatorname{Re}(s)) < 1 = |z_1|
 \end{aligned}$$

при $|z_1| = 1$. В силу теоремы Руше отсюда следует, что функциональное уравнение (25) имеет единственное решение $z_1 = z_{1,n}(z_2, s)$, причем функция $z_{1,n}(z_2, s)$ является аналитической в области $|z_2| \leq 1 \times \operatorname{Re}(s) > 0$.

Подставляя $z_1 = z_{1,n}(z_2, s)$ в уравнение (24), получим:

$$\begin{aligned}
 & C_{2,n}(z_{1,n}(z_2, s), z_2, s) \left(1 - \right. \\
 & \quad \left. - z_2^{-1} \beta_2 \left(s + a_1 - a_1 z_{1,n}(z_2, s) + a_2 - \tilde{\lambda}_n(z_2) \right) \right) = \\
 & = \frac{\prod_{i=1}^M (\tilde{\lambda}_n(z_2) - pa_2 z_2^i)}{\prod_{\substack{j=1 \\ j \neq n}}^M (\tilde{\lambda}_n(z_2) - \tilde{\lambda}_j(z_2))} \times \\
 & \quad \times \sum_{m=1}^M \frac{b_m(z_{1,n}(z_2, s), z_2, s)}{\tilde{\lambda}_n(z_2) - pa_2 z_2^m}, \quad (26)
 \end{aligned}$$

откуда

$$\begin{aligned}
 & C_{2,n}(z_{1,n}(z_2, s), z_2, s) = \\
 & = \frac{1}{1 - z_2^{-1} \beta_2 (s + a_1 - a_1 z_{1,n}(z_2, s) + a_2 - \tilde{\lambda}_n(z_2))} \times \\
 & \quad \times \frac{\prod_{i=1}^M (\tilde{\lambda}_n(z_2) - pa_2 z_2^i)}{\prod_{\substack{j=1 \\ j \neq n}}^M (\tilde{\lambda}_n(z_2) - \tilde{\lambda}_j(z_2))} \sum_{m=1}^M \frac{b_m(z_{1,n}(z_2, s), z_2, s)}{\tilde{\lambda}_n(z_2) - pa_2 z_2^m}.
 \end{aligned}$$

Заметим, что $\tilde{\pi}_2(z_1, z_2, k, x, s)$, а значит, и $C_{2,n}(z_1, z_2, s)$, не зависит от z_1 . Из этого факта вытекает, что можно записать:

$$\begin{aligned}
 & C_{2,n}(z_1, z_2, s) = \\
 & = \frac{1}{1 - z_2^{-1} \beta_2 (s + a_1 - a_1 z_{1,n}(z_2, s) + a_2 - \tilde{\lambda}_n(z_2))} \times \\
 & \quad \times \frac{\prod_{i=1}^M (\tilde{\lambda}_n(z_2) - pa_2 z_2^i)}{\prod_{\substack{j=1 \\ j \neq n}}^M (\tilde{\lambda}_n(z_2) - \tilde{\lambda}_j(z_2))} \sum_{m=1}^M \frac{b_m(z_{1,n}(z_2, s), z_2, s)}{\tilde{\lambda}_n(z_2) - pa_2 z_2^m}. \quad (27)
 \end{aligned}$$

Подставляя (27) в (24), получим

$$C_{1,n}(z_1, z_2, s) \left(1 - z_1^{-1} \beta_1 (s + a_1 - a_1 z_1 + a_2 - \tilde{\lambda}_n(z_2)) \right) = \frac{\prod_{i=1}^M (\tilde{\lambda}_n(z_2) - pa_2 z_2^i)}{\prod_{\substack{j=1 \\ j \neq n}}^M (\tilde{\lambda}_n(z_2) - \tilde{\lambda}_j(z_2))} \sum_{m=1}^M \frac{1}{\tilde{\lambda}_n(z_2) - pa_2 z_2^m} \times (b_m(z_1, z_2, s) - \frac{1 - z_2^{-1} \beta_2 (s + a_1 - a_1 z_1 + a_2 - \tilde{\lambda}_n(z_2))}{1 - z_2^{-1} \beta_2 (s + a_1 - a_1 z_{1,n}(z_2, s) + a_2 - \tilde{\lambda}_n(z_2))} \times b_m(z_{1,n}(z_2, s), z_2, s)),$$

т. е.

$$C_{1,n}(z_1, z_2, s) = \frac{1}{1 - z_1^{-1} \beta_1 (s + a_1 - a_1 z_1 + a_2 - \tilde{\lambda}_n(z_2))} \times \frac{\prod_{i=1}^M (\tilde{\lambda}_n(z_2) - pa_2 z_2^i)}{\prod_{\substack{j=1 \\ j \neq n}}^M (\tilde{\lambda}_n(z_2) - \tilde{\lambda}_j(z_2))} \sum_{m=1}^M \frac{1}{\tilde{\lambda}_n(z_2) - pa_2 z_2^m} \times (b_m(z_1, z_2, s) - \frac{1 - z_2^{-1} \beta_2 (s + a_1 - a_1 z_1 + a_2 - \tilde{\lambda}_n(z_2))}{1 - z_2^{-1} \beta_2 (s + a_1 - a_1 z_{1,n}(z_2, s) + a_2 - \tilde{\lambda}_n(z_2))} \times b_m(z_{1,n}(z_2, s), z_2, s)).$$

Остается найти $\pi_0(m, s)$, $m = 1, \dots, M$. Рассмотрим уравнение:

$$z_2 = \beta_2 (s + a_1 - a_1 z_{1,n}(z_2, s) + a_2 - \tilde{\lambda}_n(z_2)). \quad (28)$$

Обе части уравнения являются аналитическими в области $|z_2| \leq 1$ функциями. Имеем:

$$\left| \beta_2 (s + a_1 - a_1 z_{1,n}(z_2, s) + a_2 - \tilde{\lambda}_n(z_2)) \right| \leq \beta_2 \left(\operatorname{Re} (s + a_1 - a_1 z_{1,n}(z_2, s) + a_2 - \tilde{\lambda}_n(z_2)) \right) \leq \beta_2 (\operatorname{Re}(s)) < 1 = |z_2|$$

при $|z_2| = 1$. В силу теоремы Руше отсюда следует, что функциональное уравнение (28) имеет единственное решение $z_2 = z_{2,n}(s)$, причем функция $z_{2,n}(s)$ является аналитической в области $\operatorname{Re}(s) > 0$.

¹ В дальнейшем будем для краткости писать $z_{1,n}$ вместо $z_{1,n}(z_{2,n}(s), s)$, $z_{2,n}$ вместо $z_{2,n}(s)$ и $\tilde{\lambda}_n$ вместо $\tilde{\lambda}_n(z_{2,n})$.

Подставляя $z_2 = z_{2,n}(s)$ в уравнение (26), приходим после ряда преобразований к уравнению¹:

$$\sum_{m=1}^M \prod_{\substack{j=1 \\ j \neq m}}^M (\tilde{\lambda}_n - pa_2 z_{2,n}^j) b_m(z_{1,n}, z_{2,n}, s) = 0. \quad (29)$$

Вспомним, что

$$b_m(z_1, z_2, s) = -(s + a_1 - a_1 z_1 + a_2) \pi_0(m, s) + h_m + \left[p \pi_0(m, s) a_2 z_2^m + (1 - p) \sum_{n=1}^M \pi_0(n, s) h_n a_2 z_2^m \right].$$

С учетом уравнения (20) будем иметь:

$$\sum_{m=1}^M \prod_{\substack{j=1 \\ j \neq m}}^M (\tilde{\lambda}_n - pa_2 z_{2,n}^j) \left[p \pi_0(m, s) a_2 z_{2,n}^m + (1 - p) \sum_{k=1}^M \pi_0(k, s) h_k a_2 z_{2,n}^m \right] = \sum_{m=1}^M \prod_{\substack{j=1 \\ j \neq m}}^M (\tilde{\lambda}_n - pa_2 z_{2,n}^j) \times \left[(pa_2 z_{2,n}^m - \tilde{\lambda}_n) \pi_0(m, s) + \tilde{\lambda}_n \pi_0(m, s) + (1 - p) \sum_{k=1}^M \pi_0(k, s) h_k a_2 z_{2,n}^m \right] = - \prod_{j=1}^M (\tilde{\lambda}_n - pa_2 z_{2,n}^j) \sum_{m=1}^M \pi_0(m, s) + \sum_{m=1}^M \prod_{\substack{j=1 \\ j \neq m}}^M (\tilde{\lambda}_n - pa_2 z_{2,n}^j) \tilde{\lambda}_n \pi_0(m, s) + \prod_{j=1}^M (\tilde{\lambda}_n - pa_2 z_{2,n}^j) \sum_{m=1}^M \pi_0(m, s) = \sum_{m=1}^M \prod_{\substack{j=1 \\ j \neq m}}^M (\tilde{\lambda}_n - pa_2 z_{2,n}^j) \tilde{\lambda}_n \pi_0(m, s).$$

Возвращаясь к уравнению (29), получим:

$$\sum_{m=1}^M \prod_{\substack{j=1 \\ j \neq m}}^M (\tilde{\lambda}_n - pa_2 z_{2,n}^j) \left(-(s + a_1 - a_1 z_{1,n} + a_2) \pi_0(m, s) + h_m + \tilde{\lambda}_n \pi_0(m, s) \right) = 0.$$

Там самым доказательство теоремы завершено.

Литература

1. Леонтьев Н. Д., Ушаков В. Г. Анализ системы обслуживания с входящим потоком авторегрессионного типа // Информатика и её применения, 2014. Т. 8. Вып. 3. С. 39–44.
2. Schechter S. On the inversion of certain matrices // Math. Tab. Aids Comput., 1959. Vol. 13. No. 66. P. 73–77.

Поступила в редакцию 11.05.16

ANALYSIS OF A QUEUEING SYSTEM WITH AUTOREGRESSIVE ARRIVALS AND NONPREEMPTIVE PRIORITY

N. D. Leontyev¹ and V. G. Ushakov^{1,2}

¹Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, Moscow 119991, GSP-1, 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: The paper studies a single server queueing system with infinite capacity and with two arrival streams, one of which is Poisson and the other is batch Poisson. The customers from the first stream have nonpreemptive priority over the customers from the second. A feature of the system under study is autoregressive dependence of the sizes of the batches from the second arrival stream: the size of the n th batch is equal to the size of the $(n - 1)$ st batch with a fixed probability and is an independent random variable with complementary probability. Service times of the customers from each stream are supposed to be independent random variables with specified distributions. The main object of the study is the number of the customers from each stream in the system at an arbitrary moment. The relations derived make it possible to find Laplace transform in time of probability generating function of the transient queue length and also a number of additional characteristics.

Keywords: queueing theory; transient behavior; batch arrivals; nonpreemptive priority

DOI: 10.14357/19922264160303

Acknowledgments

This work was supported by the Russian Foundation for Basic Research (project No. 15-07-02354).

References

1. Leontyev, N. D., and V. G. Ushakov. 2014. Analiz sistemy obsluzhivaniya s vkhodyashchim potokom avtoregressivnogo tipa [Analysis of queueing system with autoregressive arrivals]. *Informatika i ee Primeneniya — Inform. Appl.* 8(3):39–44.
2. Schechter, S. 1959. On the inversion of certain matrices. *Math. Tab. Aids Comput.* 13(66):73–77.

Received May 11, 2016

Contributors

Leontyev Nikolai D. (b. 1988) — PhD student, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, Moscow 119991, GSP-1, Russian Federation; ndleontyev@gmail.com

Ushakov Vladimir G. (b. 1952) — Doctor of Science in physics and mathematics, professor, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, Moscow 119991, GSP-1, 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; vgushakov@mail.ru

АНАЛИЗ ПРОИЗВОДИТЕЛЬНОСТИ БЕСПРОВОДНОЙ СИСТЕМЫ АГРЕГАЦИИ ДАННЫХ С СОСТЯЗАНИЕМ ДЛЯ СОВРЕМЕННЫХ СЕНСОРНЫХ СЕТЕЙ*

А. Я. Омётов¹, С. Д. Андреев², А. М. Тюрликов³, Е. А. Кучерявый⁴

Аннотация: Рассматривается беспроводная система связи, учитывающая особенности современных сенсорных сетей, в которых устройства передают свои данные на множество промежуточных агрегирующих узлов, имеющих подключение к сети Интернет по технологии IEEE 802.11-2014 (WiFi). Предполагается, что агрегатор осуществляет пересылку данных от многих сенсоров, участвуя при этом в состязании за общий канал связи с другими агрегаторами. Предлагается аналитическая модель такого состязания, учитывающая специфику алгоритма разрешения коллизий, характеристики протокола доступа к каналу, а также возможность потери данных на узле агрегации. Полученные аналитические результаты сопоставляются с данными имитационного моделирования, и вычисляется максимальное количество поддерживаемых системой связи сенсоров.

Ключевые слова: Интернет вещей; беспроводные сенсорные сети; регенеративный анализ; БЛВС; стандарт IEEE 802.11-2014

DOI: 10.14357/19922264160304

1 Введение

В последнее время все более усиливается влияние беспроводных технологий на современное общество, что, в свою очередь, предвещает рост научного интереса к данной тематике в ближайшие годы и влечет за собой потенциальную возможность установления беспроводного соединения в любом месте и в любое время [1]. Данная возможность является привлекательной для внедрения концепции «Интернета вещей» (Internet of Things, IoT) [2]. При интеграции IoT многие устройства могут быть оборудованы сенсорами и расширяющими модулями, с помощью которых появляется возможность обрабатывать и передавать информацию без вмешательства человека. Данные беспроводные технологии открывают дорогу для широкого спектра сервисов, начиная с удаленного наблюдения и заканчивая здравоохранением.

Основной целью текущих исследований является разработка системы связи, использующей подходящие IoT-технологии для реализации подключения разнородных сенсоров.

Исторически беспроводные сети разрабатывались для использования людьми, а для использования их машинами необходима значительная опти-

мизация, что определяет необходимость разработки или улучшения технологий связи для поддержки большого числа устройств [3]. Основными требованиями к таким технологиям остаются низкая сложность обработки данных, дешевизна в производстве и высокая энергоэффективность.

Беспроводные локальные сети на основе стандарта IEEE 802.11 (WiFi) являются одним из самых распространенных технических решений для организации беспроводного доступа в домах и на предприятиях. Благодаря их высокой пропускной способности, относительно низкой стоимости и повсеместной распространенности использование WiFi для сценариев IoT является все более привлекательным.

В данной работе производится исследование принципиальной возможности и эффективности использования современной технологии WiFi с учетом ее технических характеристик для типовых сценариев IoT. В частности, предлагается аналитическая модель для учета особенностей работы технологии WiFi (состязание между агрегаторами, особенности протокола доступа, режимы передачи данных и т.д.), основанная на теории регенерирующих процессов. Аналитические результаты

* Исследование выполнено при частичной финансовой поддержке РФФИ (проект 15-07-03051), а также Фонда содействия развитию малых форм предприятий в научно-технической сфере в рамках программы «УМНИК» по договору № 8268ГУ2015 от 02.12.2015.

¹ Санкт-Петербургский государственный университет телекоммуникаций им. М. А. Бонч-Бруевича, alexander.ometov@gmail.com

² Российский университет дружбы народов, serge.andreev@gmail.com

³ Санкт-Петербургский государственный университет аэрокосмического приборостроения, turlikov@vu.spb.ru

⁴ Национальный исследовательский университет «Высшая школа экономики», ykoucheryavy@hse.ru

сопоставляются с данными, полученными имитационным моделированием, и делается вывод о наибольшем возможном количестве сенсоров, поддерживаемых такой системой связи.

2 Модель системы и анализ

2.1 Описание сценария и протокола

В данной работе рассматривается изолированный сегмент (кластер) беспроводной сети для IoT-приложений со статичным размещением M агрегирующих узлов, в котором отсутствуют «скрытые» станции. Топология данной сети представлена на рис. 1.

Агрегаторы оборудованы двумя модулями беспроводной связи: WiFi и ZigBee. Сенсоры передают собранную ими информацию на агрегаторы посредством ZigBee, и далее поток данных перенаправляется в канал WiFi, соединяющий агрегаторы с сетью Интернет. Агрегаторы взаимодействуют в нелицензированном частотном диапазоне WiFi и используют протокол случайного множественного доступа для передачи накопленных данных в общий канал связи. При передаче более чем от одного агрегатора одновременно возникает наложение таких передач в канале — коллизия. В случае если передавало только одно устройство, передача считается успешной в предположении, что в канале отсутствует шум. Третье возможное событие — пустой слот — происходит в случае, если ни один из агрегаторов не осуществлял передачу. В то же время известные решения для сотовых сетей в данной работе рассматриваться не будут [4].

В данной работе рассматривается система среднего/большого производства, где в каждом помещении находится один агрегатор, обслуживающий не более 20 узлов. Интерференция и конфликты в канале сенсор—агрегатор для данного исследования не являются критическими, так как алгоритмы беспроводной связи ближнего радиуса действия не подвержены интерференции от кластеров из соседних помещений, в то время как агрегаторы могут конфликтовать в связи с более широким радиусом действия беспроводного покрытия.

Таким образом, в работе предполагается худший случай насыщенного трафика между всеми агрегирующими узлами [5]. Иными словами, на уровне управления доступом к среде можно наблюдать целиком заполненный буфер исходящих сообщений. Более высокие уровни модели связи в данной работе не рассматриваются ввиду предположения о простоте сенсоров. Данное допущение дает возможность оценивать наихудший сценарий загрузки сети и позволяет производить оценку с точки зрения пропускной способности насыщения S .

Согласно спецификации стандарта IEEE 802.11, процесс доступа агрегаторов к общему каналу связи основан на алгоритме двоичной экспоненциальной отсрочки (ДЭО) и состоит в выборе случайного интервала отсрочки перед передачей из некоторого окна отсрочки (CW). В процессе работы алгоритма ДЭО в случае передачи одновременно двумя или более агрегаторами происходит коллизия. Она отслеживается на стороне получателя (точки доступа), а каждому из передающих абонентов предоставляется возможность повторной передачи, если это позволит сделать счетчик повторных передач (RC).

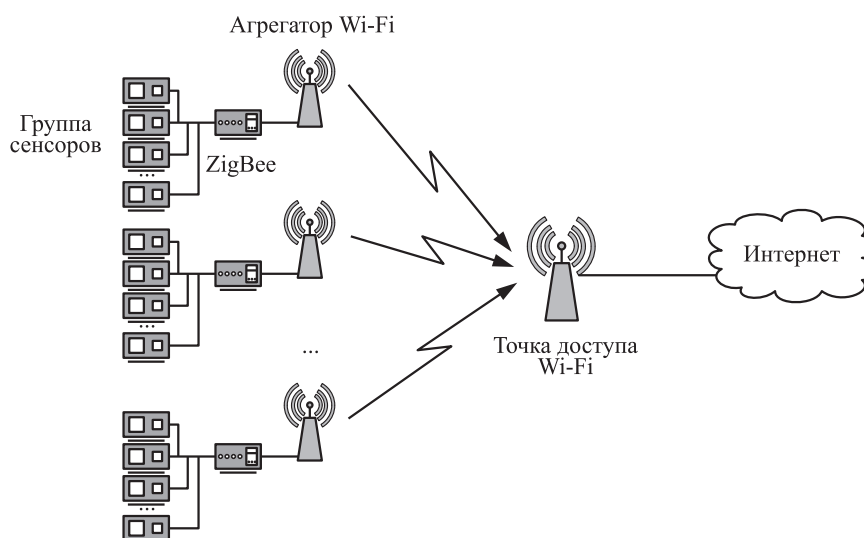


Рис. 1 Предполагаемая топология беспроводной сети

Основные обозначения, использованные в работе

Параметр	Обозначение
Максимальная длительность доступа	$T_{\text{ТХОП}}$
Количество агрегаторов	M
Текущее окно отсрочки	W_i, CW
Начальное окно отсрочки	W_0
Длительность пустого слота	Σ
Длительность AIFS	T_{AIFS}
Длительность SIFS	T_{SIFS}
Длительность BA	T_{BA}
Длительность RTS	T_{RTS}
Длительность CTS	T_{CTS}
Длительность CFE	T_{CFE}
Средняя длительность блока данных	$T_{E[P]}$
Длительность передачи преамбулы P	T_P

В данном случае значение окна отсрочки CW будет удвоено ($W_{i+1} = 2W_i$) с целью уменьшения вероятности повторной коллизии. В то же время будет уменьшен счетчик повторных передач RC . Возможный рост CW ограничен максимальным значением ($\text{CW}_{\text{max}} = 2^m W_0$), где m определяется как «степень» отсрочки. Основные используемые в работе сокращения представлены в таблице.

Данная схема передачи может использовать два альтернативных механизма доступа, подробно описанных в спецификации IEEE 802.11. При использовании механизма базового доступа в канал (Basic) пакет данных (либо агрегированная группа пакетов с единой преамбулой P) передается незамедлительно после ожидания регуляционного межкадрового интервала (AIFS) и случайного времени отсрочки (ВОТ). Информация об успешной передаче пакета/блока данных передается в блоковом подтверждении (BA). Механизм доступа «запрос на отправку/разрешение отправки» (RTS/CTS) использует алгоритм «двукратного рукопожатия» при передаче сообщения. Иными словами, канал резервируется за определенным агрегатором на время его предполагаемой передачи.

В результате агрегации пакетов данных на физическом уровне, на высоких скоростях передачи влияние служебной информации на пропускную способность становится менее значительно. Однако длительность нового пакета данных после агрегирования не должна превышать интервала захвата среды передачи (ТХОП), включая необходимые межкадровые интервалы (SIFS), блоковое подтверждение (BA) и (опционально) RTS/CTS. При необходимости агрегатор может также преждевременно закончить передачу, в случае если длительность передачи пакета данных оказалась меньше ТХОП, отправив сообщение об освобождении канала (CFE).

Согласно стандарту после первого интервала AIFS значение счетчика отсрочки (BC) выбирается как равномерно распределенная величина W_i в промежутке между 0 и $W_0 - 1$, где W_i является окном отсрочки CW . После каждого пустого слота BC уменьшается на единицу. Как только BC достигает нуля, выбранный агрегатор пытается передать. При возникновении коллизии происходит повторная передача, в случае если счетчик повторных передач RC не равен нулю. В случае повторной передачи $W_i = 2W_{i-1}$. Рост CW также ограничен максимальным значением W_{max} , но агрегатор может продолжать передавать повторно, пока RC не достигнет нуля. В момент, когда пакет успешно передан или принято решение об отказе от передачи, CW устанавливается в некоторое начальное значение W_0 . Эквивалентно $\text{CW}_{\text{max}} = 2^m W_0$, где m — степень отсрочки.

Далее представлен анализ работы системы доступа к каналу, функционирующей на основе алгоритма ДЭО с потерями согласно описанию выше. В работе [6] предложено рассматривать два параметра функционирования системы на основании циклов регенерации: вероятность передачи p_t пользователем в конкретный временной слот и вероятность ошибки p_c в случае осуществленной передачи. Данные параметры предполагаются неизменными на всем протяжении работы насыщенной системы. Вследствие данного допущения весь кластер может быть рассмотрен с точки зрения одного «меченого» абонента этой системы. Влияние всех прочих факторов учитывается в значении вероятности ошибки p_c . Необходимо отметить, что данная замена допустима только в случае справедливой системы, т. е. когда все пользователи получают приблизительно равную долю времени доступа к каналу [7].

2.2 Общие понятия для системы без потерь

Рассмотрим модель, представленную на рис. 2, основываясь на концепции *циклов регенерации*.

Данная модель рассматривает равные временные слоты, где начало передачи пакета совпадает с началом слота. Каждая передача занимает в точности один слот. Подобное упрощение типично для моделирования протоколов случайного множественного доступа [8] и позволяет легко масштабировать модель согласно требуемым временным характеристикам конкретного протокола, что будет показано в следующем подразделе. Отметим, что меченый агрегатор имеет следующую вероятность конфликта в произвольно взятом слоте:

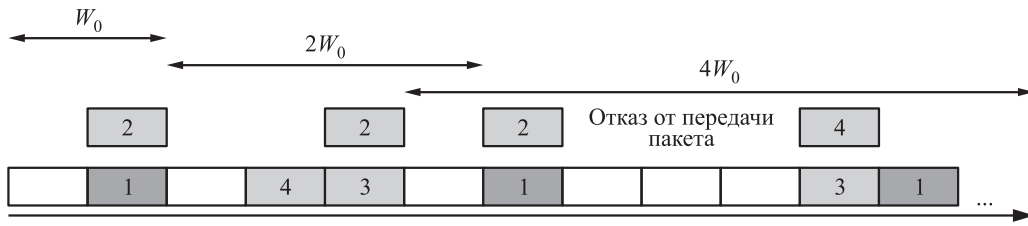


Рис. 2 Упрощенная модель алгоритма ДЭО

$$p_c = 1 - (1 - p_t)^{M-1}. \quad (1)$$

Важно также обратить внимание, что вероятность передачи в канал для данного агрегатора может быть рассчитана как отношение числа попыток передачи на пакет $B^{(i)}$ в течение некоторого цикла регенерации к длительности данного цикла в слотах $D^{(i)}$:

$$p_t = \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n B^{(i)}}{\sum_{i=1}^n D^{(i)}} = \frac{E[B]}{E[D]}. \quad (2)$$

Предполагая, что система находится в насыщении и не имеет потерь, можно с легкостью получить $E[B]$:

$$\begin{aligned} E[B] &= \sum_{i=1}^{\infty} \Pr\{B = i\} = \\ &= (1 - p_c) \sum_{i=1}^{\infty} i p_c^{i-1} = \frac{1}{1 - p_c}. \end{aligned} \quad (3)$$

Выражение для $E[D]$ может быть получено аналогичным способом [9]:

$$\begin{aligned} E[D] &= \sum_{i=1}^{\infty} D^{(i)} \Pr\{D = i\} = \\ &= (1 - p_c) \sum_{i=1}^{\infty} D^{(i)} p_c^{i-1}, \end{aligned} \quad (4)$$

где $D^{(i)}$ — длина цикла регенерации при условии, что было произведено i попыток передачи.

Основываясь на зависимости i и m , можно также получить следующие правила вычисления величины $D^{(i)}$:

$$D^{(i)} = \begin{cases} 2^{i-1} W_0 - \frac{W_0 - 1}{2}, & \text{если } 1 \leq i \leq m + 1; \\ 2^{m-1} W_0 (i - m + 1) - \frac{W_0 - i}{2}, & \text{если } i > m + 1. \end{cases} \quad (5)$$

Подставив (5) в (4), после преобразования получим:

$$\begin{aligned} E[D] &= \\ &= \frac{(1 - 2p_c)(W_0 + 1) + p_c W_0 (1 - (2p_c)^m)}{2(1 - 2p_c)(1 - p_c)}. \end{aligned} \quad (6)$$

Далее, подставляя (3) и (6) в (2), имеем:

$$p_t = \frac{2(1 - 2p_c)}{(1 - 2p_c)(W_0 + 1) + p_c W_0 (1 - (2p_c)^m)}. \quad (7)$$

Следует отметить, что аналогичные результаты были получены в известной работе [10], где аналитические расчеты были основаны на двумерных цепях Маркова, которые сложно масштабировать на случай дополнительных параметров системы. Здесь же использован иной математический подход, являющийся более простым, но не менее эффективным. Выражения (1) и (7) составляют систему двух нелинейных уравнений с неизвестными p_c и p_t , которую можно решить численно.

2.3 Система с потерями

В данном подразделе представлен анализ исходной системы, работающей с потерями, т.е. с определенным максимальным числом повторных передач K для отдельно взятого пакета. Для вычисления p_t в данной системе определяем $E[B]$ и $E[D]$ согласно (2), но выражение для расчета среднего числа попыток передачи в цикле $E[B]$ необходимо модифицировать следующим образом:

$$\begin{aligned} E[B] &= \sum_{i=1}^{K+1} i \Pr\{B = i\} = \\ &= (1 - p_c) \sum_{i=1}^{K+1} i p_c^{i-1} (K + 1) p_c^{K+1} = \frac{1 - p_c^{K+1}}{1 - p_c}. \end{aligned} \quad (8)$$

Далее вычисляем среднюю длительность цикла регенерации как

$$\begin{aligned} E[D] &= \sum_{i=1}^{K+1} D^{(i)} \Pr\{B = i\} = \\ &= (1 - p_c) \sum_{i=1}^{K+1} D^{(i)} p_c^{i-1} + D(K + 1) p_c^{K+1}. \end{aligned}$$

Возможны две ситуации, зависящие от соотношения m и K : когда $K \leq m$ и $K > m$. В первом случае

$$E[D'] = (1-p_c) \left[\sum_{i=1}^{K+1} \left(2^{i-1}W_0 - \frac{W_0-2}{2} \right) p_c^{i-1} \right] + p_c^{K-1} \left(2^K W_0 - \frac{W_0-(K+1)}{2} \right). \quad (9)$$

Соответственно, вероятность выхода в канал p'_t может быть получена при расчете выражения (2) в результате подстановки $E[B]$ из (8) и $E[D]$ (в данном случае $E[D']$) из (9):

$$p'_t = 2(1-2p_c)(1-2p_c^{K+1}) / \left[(1-2p_c) \left(1 - (2p_c)^{K+1} \right) + W_0(1-p_c) \left(1 - (2p_c)^{K+1} \right) \right]. \quad (10)$$

Чтобы рассчитать второй случай (когда $K > m$), необходимо получить соответствующее значение $E[D'']$:

$$E[D''] = (1-p_c) \left[\sum_{i=1}^{m+1} \left(2^{i-1}W_0 - \frac{W_0-i}{2} \right) p_c^{i-1} + \sum_{i=1}^{m+1} \left(2^{i-1}W_0(i-m+1) - \frac{W_0-i}{2} \right) p_c^{i-1} \right] + p_c^{K-1} \left(2^K W_0 - \frac{W_0-(K+1)}{2} \right). \quad (11)$$

Аналогично можно получить вероятность передачи p''_t , подставляя $E[B]$ из (8) и $E[D]$ (в данном случае $E[D'']$) из (11) в (2):

$$p''_t = 2(1-2p_c)(1-2p_c^{K+1}) / \left[(1-2p_c) \times \left(W_0 \left(1 - (2p_c)^{K+1} \right) \right) + \left(1 - (2p_c)^{K+1} \right) + W_0 p_c \left(1 - (2p_c)^m \right) \right]. \quad (12)$$

Итак, вероятность выхода в канал p_t в системе с потерями может быть получена двумя способами: как p'_t из (10) или как p''_t из (12), в зависимости от соотношения K и m . Решая нелинейную систему уравнений (8), получаем итоговое значение p_t .

Также в данной работе были использованы фактические длительности служебных сообщений WiFi для вычисления пропускной способности насыщения, которая может быть воспроизведена по аналогии с [8, 10, 11] и множеством других работ. Основное отличие от вышеприведенной модели заключается в том, что используются разные длительности слотов для различных событий в канале, которые соответствуют текущей спецификации

стандарта IEEE 802.11-2014. Далее рассматривается работа механизма доступа «RTS/CTS». Размеры слотов разной длительности следующие: σ соответствует длительности пустого слота, T_S — длительности успешной передачи, а T_C — длительности коллизии. Длительности успеха и коллизии могут быть рассчитаны согласно [7] как

$$T_S = T_{RTS} + T_{SIFS} + T_{CTS} + T_{SIFS} + T_P + T_{E[P]} + T_{SIFS} + T_{BA} + T_{AIFS}, \\ T_C = T_{RTS} + T_{AIFS}.$$

В итоге можно получить пропускную способность насыщения S :

$$S = \frac{P_t P_S E[P]}{(1-P_t)\sigma + P_t P_S T_S + P_t(1-P_S)T_C},$$

где $P_t = 1 - (1-p_t)^M$ — вероятность того, что в системе данные передавал только один агрегатор, $P_S = M p_t (1-p_t)^{M-1} p_t^{-1}$ — вероятность успешной передачи (при условии, что передавал один агрегатор).

В системе с потерями можно также выписать вероятность того, что агрегатор отказывается от передачи пакета после K неуспешных повторных передач, вызванных коллизиями: $P_d = p_c^{K-1}$.

3 Результаты и выводы

В данном разделе представлены результаты моделирования для современной версии протокола IEEE 802.11. Также рассмотрено их сопоставление с аналитическими результатами, полученными выше. Реализованная система имитационного моделирования является гибким программным инструментом, включающим различные сценарии взаимодействия WiFi-агрегаторов, а также набор необходимых механизмов для управления доступом к среде. Система моделирования была откалибрована по результатам, представленным в работе [10] (зависимость пропускной способности S от количества абонентов M в канале). Для этого был оцифрован соответствующий график и произведено наложение полученных результатов на воспроизведенные данные. Результаты можно наблюдать на рис. 3, а.

Следует отметить, что результаты работы [10] и проведенного в данной работе моделирования совпадают в аналогичных условиях и с максимальной скоростью передачи $T = 1$ Мбит/с (система без потерь).

Далее рассмотрим сценарий, в котором M агрегаторов в кластере используют реалистичные длительности слотов согласно текущей спецификации

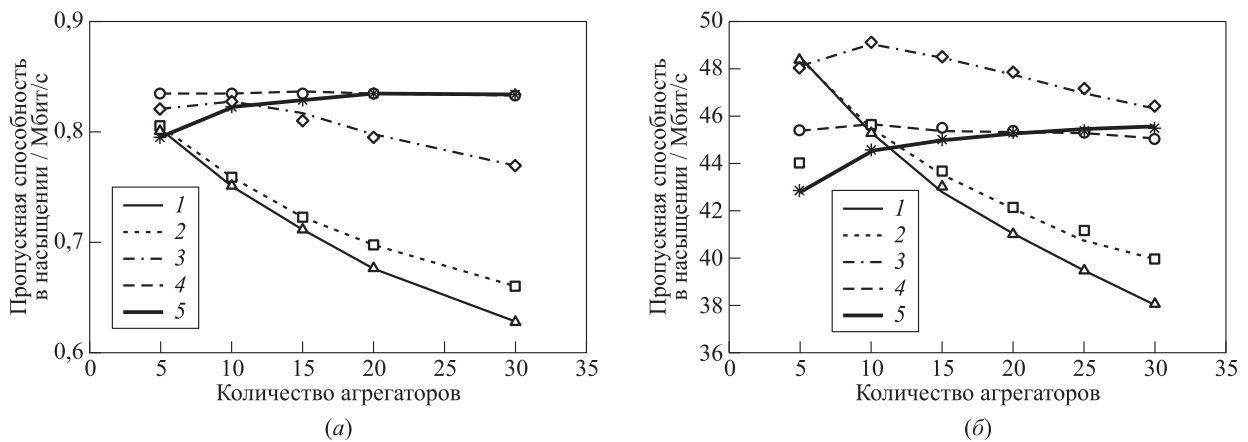


Рис. 3 Пропускная способность насыщения при $T = 1$ (а) и 65 Мбит/с (б) и $K \rightarrow \infty$: 1 — Basic, $W_0 = 32, m = 3$; 2 — Basic, $W_0 = 32, m = 5$; 3 — Basic, $W_0 = 128, m = 3$; 4 — RTS/CTS, $W_0 = 32, m = 3$; 5 — RTS/CTS, $W_0 = 128, m = 3$

IEEE 802.11-2014. Для этого установим более высокую скорость передачи $T = 65$ Мбит/с и сопоставим анализ (кривые) с результатами моделирования (символы) на рис. 3, б.

Рассмотрение системы, работающей без потерь (в которой количество попыток повторной передачи не ограничено), воссоздает ситуацию, когда M агрегаторов отправляют сообщения к точке доступа WiFi в насыщенном режиме. С помощью предлагаемого подхода получена максимально достижимая пропускная способность системы. В целом зависимость на рис. 3, б аналогична результатам, приведенным на рис. 3, а, но значение достижимой пропускной способности в случае $T = 65$ Мбит/с значительно выше.

В заключение рассмотрим систему предложенной топологии (см. рис. 1) с различным числом агрегаторов (5–16) и определим максимально возможное число сенсоров, которые могут быть обслужены такой системой.

Результаты на рис. 4 получены при условии, что каждый сенсор передает пакеты на агрегатор со средней скоростью 256 бит/с, а максимальные пропускные способности в восходящем канале (к точке доступа WiFi) взяты согласно результатам, представленным на рис. 3, б. В работе [12] представлена статистика реальной плотности размещения сенсоров в городских условиях, что для типового радиуса покрытия точки доступа WiFi (до 300 м) дает около 1000 устройств на агрегатор. При этом

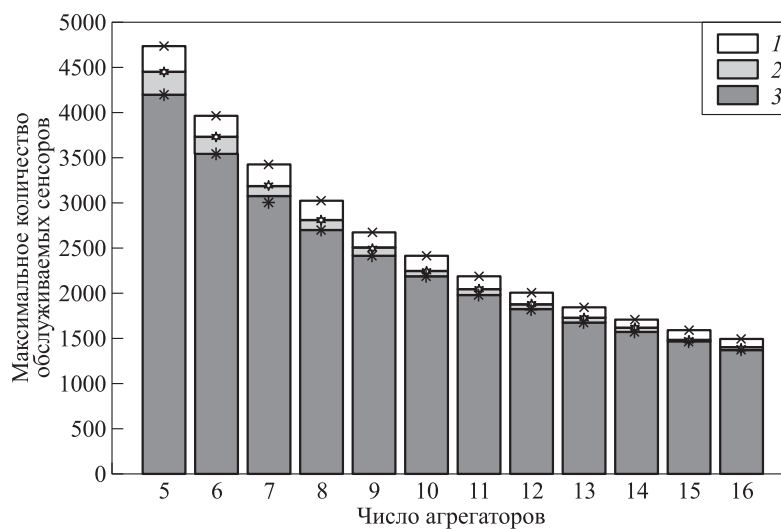


Рис. 4 Максимальное число сенсоров с агрегацией в канале WiFi: 1 — Basic, $W_0 = 128, m = 3$; 2 — RTS/CTS, $W_0 = 32, m = 3$; 3 — RTS/CTS, $W_0 = 128, m = 5$

на рис. 4 видно, что при использовании предложенной в текущем исследовании топологии сети, использующей WiFi-агрегацию, удается достичь расчетного числа поддерживаемых сенсоров даже при достаточно большом количестве агрегирующих устройств.

4 Заключение

В данной работе проведен анализ системы передачи данных от сенсорных устройств с их промежуточной агрегацией и пересылкой по технологии WiFi (IEEE 802.11-2014). Предполагается, что соответствующая топология станет типовой для многих IoT-приложений с большим количеством сенсоров, и рассматривается этап состязания между WiFi-агрегаторами за выход в беспроводной канал связи.

Предлагается аналитическая модель, построенная на основе теории регенерирующих процессов и учитывающая основные особенности работы протокола доступа к каналу, а также алгоритма разрешения коллизий. В частности, предполагается, что данные на агрегаторе могут быть потеряны, если они превысили определенное количество попыток повторной передачи. Значение пропускной способности насыщения, полученное в рамках предложенной модели, сопоставляется с данными имитационного моделирования. Делается вывод об их совпадении и обосновывается максимальное число сенсоров, которые могут быть обслужены системой с рассматриваемой топологией. Полученные значения существенно превышают ожидаемое число сенсоров в городских условиях даже при достаточно большом количестве агрегаторов, что подтверждает целесообразность использования технологии WiFi в беспроводных системах агрегации данных с большим числом сенсоров.

Литература

1. *Ahmadian A., Galinina O. S., Gudkova I. A., Andreev S. D., Shorgin S. Ya., Samouylov K. E.* On capturing spatial di-

versity of joint M2M/H2H dynamic uplink transmissions in 3GPP LTE cellular system // Next Generation Teletraffic and Wired/Wireless Advanced Networking Conference (International) Proceedings. — Lecture notes in computer science ser. — St. Petersburg, Russia, 2015. Vol. 9247. P. 407–421.

2. *Кучерявый А. Е.* Самоорганизующиеся сети и новые услуги // Электросвязь, 2009. Вып. 1. С. 19–23.
3. *Восков Л. С.* Беспроводные сенсорные сети и прикладные проекты // Автоматизация и IT в энергетике, 2009. № 2-3. С. 44–49.
4. *Косинов М. И., Шорин О. А.* Повышение емкости сетевой системы связи при использовании зон перекрытия // Электросвязь, 2003. Вып. 1. С. 18–20.
5. *Гайдамака Ю. В., Печинкин А. В., Разумчик Р. В., Самуйлов А. К., Самуйлов К. Е., Соколов И. А., Соппин Э. С., Шоргин С. Я.* Распределение времени выхода из множества состояний перегрузки в системе $M|M|1|(L, H)|(H, R)$ с гистерезисным управлением нагрузкой // Информатика и её применения, 2013. Т. 7. Вып. 4. С. 20–33.
6. *Bianchi G.* Performance analysis of the IEEE 802.11 distributed coordination function // IEEE J. Sel. Area. Comm., 2000. Vol. 18. No. 3. P. 535–547.
7. *Skordoulis D., Ni Q., Chen H. H., Stephens A. P., Liu C., Jamalipour A.* IEEE 802.11n MAC frame aggregation mechanisms for next-generation high-throughput WLANs // IEEE Wirel. Commun., 2008. Vol. 15. No. 1. P. 40–47.
8. *Sharma G., Ganesh A., Key P., Needham R.* Performance analysis of contention based medium access control protocols // IEEE Trans. Inform. Theory, 2009. Vol. 55. No. 4. P. 1665–1682.
9. *Malone D., Duffy K., Leith D.* Modeling the 802.11 distributed coordination function in non-saturated heterogeneous conditions // IEEE/ACM Trans. Networks, 2007. Vol. 15. No. 1. P. 159–172.
10. *Bordenave C., McDonald D., Proutire A.* Random multi-access algorithms — a mean field analysis // Rapport de Recherche, 2005. Vol. 5632. P. 1–12.
11. *Andreev S., Koucheryavy Y., Sousa L.* Calculation of transmission probability in heterogeneous ad hoc networks // IEEE Baltic Congress on Future Internet and Communications (BCFIC) Proceedings, 2011. P. 75–82.
12. *Ortiz S.* IEEE 802.11n: The road ahead // IEEE Computer, 2009. Vol. 42. No. 7. P. 13–15.

Поступила в редакцию 06.04.16

PERFORMANCE ANALYSIS OF A WIRELESS DATA AGGREGATION SYSTEM WITH CONTENTION FOR CONTEMPORARY SENSOR NETWORKS

A. Ya. Ometov¹, S. D. Andreev², A. M. Turlikov³, and E. A. Koucheryavy⁴

¹Saint-Petersburg State University of Telecommunications, 22B Pr. Bolshhevikov, St. Petersburg 193232, Russian Federation

²Peoples' Friendship University of Russia, 3 Ordzhonikidze Str., Moscow 115419, Russian Federation

³State University of Aerospace Instrumentation, 67 Bolshaya Morskaya Str., St. Petersburg 190000, Russian Federation

⁴National Research University Higher School of Economics, 30 Myasnitskaya Str., Moscow 101000, Russian Federation

Abstract: The paper considers a wireless communication system with a number of sensing devices that transmit their data to multiple aggregating nodes connected to Internet via IEEE 802.11-2012 (WiFi) technology. It is assumed that an aggregator retransmits data from many sensors by competing with other aggregators for the shared channel. The paper proposes an analytical model taking into account the features of the collision resolution algorithm, the properties of the channel access protocol, as well as the possibility to discard data at the aggregator. The obtained analytical results are compared with the simulation data, and the maximum number of supported sensors in the communication system is estimated.

Keywords: Internet of Things; wireless sensor networks; saturated system; regenerative analysis; WLAN; IEEE 802.11-2014 standard

DOI: 10.14357/19922264160304

Acknowledgments

This work is supported by the Russian Foundation for Basic Research (project No. 15-07-03051) and the Foundation for Assistance to Small Innovative Enterprises (FASIE) within the program “UMNIK” under grant 8268GU2015 (02.12.2015).

References

1. Ahmadian, A. M., O. S. Galinina, I. A. Gudkova, S. D. Andreev, S. Ya. Shorgin, and K. E. Samouylov. 2015. On capturing spatial diversity of joint M2M/H2H dynamic uplink transmissions in 3GPP LTE cellular system. *International Next Generation Teletraffic and Wired/Wireless Advanced Networking Conference Proceedings*. Lecture notes in computer science ser. St. Petersburg, Russia. 9247:407–421.
2. Koucheryavy, A. E. 2009. Samoorganizuyushchiesya seti i novye uslugi [Ad hoc networks and new services]. *Elektrosvyaz'* [Telecommunications] 1:19–23.
3. Voskov, L. S. 2009. Besprovodnye sensornye seti i prikladnye proekty [Wireless networks and application projects]. *Avtomatizatsiya i IT v energetike* [IT automatization in Energetics] 2-3:44–49.
4. Kosinov, M. I., and O. A. Shorin. 2003. Povyshenie emkosti sotovoy sistemy svyazi pri ispol'zovanii zon perekrytiya [Increasing cellular network capacity utilizing the overlapping zones]. *Elektrosvyaz'* [Telecommunications] 1:18–20.
5. Gaydamaka, Yu. V., A. V. Pechinkin, R. V. Razumchik, A. K. Samuylov, K. E. Samuylov, I. A. Sokolov, E. S. Sopin, and S. Ya. Shorgin. 2013. Raspredelenie vremeni vykhoda iz mnozhestva sostoyaniy peregruzki v sisteme $M|M|1|(\langle L, H \rangle)|(\langle H, R \rangle)$ s gistereznym upravleniem nagruzkoy [The distribution of the return time from the set of overload states to the set of normal load states in a system $M|M|1|(\langle L, H \rangle)|(\langle H, R \rangle)$ with hysteretic load control]. *Informatika i ee Primeneniya — Inform. Appl.* 7(4): 20–33.
6. Bianchi, G. 2000. Performance analysis of the IEEE 802.11 distributed coordination function. *IEEE J. Sel. Area. Comm.* 18(3):535–547.
7. Skordoulis, D., Q. Ni, H. H. Chen, A. P. Stephens, C. Liu, and A. Jamalipour. 2008. IEEE 802.11 n MAC frame aggregation mechanisms for next-generation high-throughput WLANs. *IEEE Wirel. Commun.* 15(1):40–47.
8. Sharma, G., A. Ganesh, P. Key, and R. Needham. 2009. Performance analysis of contention based medium access control protocols. *IEEE Trans. Inform. Theory* 55(4):1665–1682.

9. Malone, D., K. Duffy, and D. Leith. 2007. Modeling the 802.11 distributed coordination function in non-saturated heterogeneous conditions. *IEEE/ACM Trans. Networks* 15(1):159–172.
10. Bordenave, C., D. McDonald, and A. Proutire. 2005. Random multi-access algorithms — a mean field analysis. *Rapport de Recherche* 5632:1–12.
11. Andreev, S., Y. Koucheryavy, and L. Sousa. 2011. Calculation of transmission probability in heterogeneous ad hoc networks. *IEEE Baltic Congress on Future Internet and Communications (BCFIC) Proceedings*. 75–82.
12. Ortiz, S. 2009. IEEE 802.11n: The road ahead. *IEEE Computer* 42(7):13–15.

Received April 06, 2016

Contributors

Ometov Aleksandr Ya. (b. 1991) — PhD student, St. Petersburg State University of Telecommunications, 22B Bolshevnikov Pr., St. Petersburg 193232, Russian Federation; alexander.ometov@gmail.com

Andreev Sergey D. (b. 1984) — Candidate of Sciences, PhD; associate professor, Peoples' Friendship University of Russia, 3 Ordzhonikidze Str., Moscow 115419, Russian Federation; serge.andeev@gmail.com

Turlikov Andrey M. (b. 1957) — Doctor of Sciences, professor; Head of Department, St. Petersburg State University of Aerospace Instrumentation, 67 Bolshaya Morskaya Str., St. Petersburg 190000, Russian Federation; turlikov@vu.spb.ru

Koucheryavy Evgeni A. (b. 1974) — Candidate of Sciences, PhD; professor, National Research University Higher School of Economics, 20 Myasnitskaya Str., Moscow 101000, Russian Federation; ykoucheryavy@hse.ru

КРИТЕРИИ ЗНАЧИМОСТИ ОТБОРА ПРИЗНАКОВ КЛАССИФИКАЦИИ

М. П. Кривенко¹

Аннотация: Рассмотрена задача отбора признаков классификации и вопросы оценивания качества получаемых решений. Среди различных методов отбора признаков внимание обращено на последовательные процедуры; мерой качества классификации выбрана вероятность правильной классификации. Для оценивания этой характеристики предложено использовать метод перепроверки и бутстреп-метод, а для исследования ряда выборочных значений — сравнительный анализ доверительных интервалов и критерии однородности биномиальных пропорций. При построении байесовского классификатора для данных применялась модель смеси нормальных распределений; ее параметры оценивались с помощью EM (expectation–maximization) алгоритма. В качестве эксперимента рассмотрена задача обоснованного выбора признаков классификации при прогнозировании типа мочевых камней в урологии. Показано, что сокращать совокупность анализируемых показателей можно не только без потери качества принимаемых решений, но и с повышением вероятности правильного прогноза типа камня.

Ключевые слова: селекция признаков; последовательная селекция вперед и назад; байесовская классификация данных; проверка однородности двоичных последовательностей; прогноз типа камня в урологии

DOI: 10.14357/19922264160305

1 Введение

Сокращение числа переменных (признаков) может привести к повышению эффективности классификации данных и к более глубокому пониманию их природы. Соответствующие постановки задач и применяемые методы востребованы на практике в весьма разнообразных областях распознавания образов и машинного обучения: категоризации текстов, дистанционном сканировании, обнаружении наркотиков, маркетинге, обработке речи, распознавании рукописных символов, медицине и т. д. Особый интерес вызывают задачи, когда данные оказываются высокоразмерными, а объем обучающей выборки относительно мал. Получить предварительное представление о соответствующей проблеме можно из разд. 10 [1], примерами обстоятельных обзоров могут служить работы [2, 3].

Причины, по которым приходится корректировать совокупность анализируемых признаков, могут заключаться в следующем:

- повышение эффективности обучения и применения классификатора путем сокращения вычислительных затрат (уменьшение времени обработки и освобождение памяти от хранения ненужных атрибутов);
- снижение стоимости последующего сбора данных за счет измерения только тех переменных, которые имеют отношение к распознаванию;

- создание предпосылок для повышения качества классификатора;
- упрощение описания классификатора благодаря более ясному пониманию природы данных и структуры модели;
- предоставление возможности специалистам в предметной области прояснить суть протекающих процессов.

2 Задачи и методы селекции признаков

Сокращение числа признаков может осуществляться либо путем отбора (селекции, выбора подмножества исходных переменных), либо путем извлечения (формирования, определения линейного или нелинейного преобразования совокупности исходных переменных для получения меньшего набора). Далее речь пойдет только о методах отбора, так как упрощение модели данных и повышение качества классификации ставятся в данной работе во главу угла.

Постановка задачи построения методов селекции подразумевает оптимизацию некоторой целевой функции. В зависимости от того, как связаны задачи предварительной обработки состава анализируемых признаков и задачи классификации дан-

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

ных, обычно выделяются две основные категории методов:

- (1) не связанные со свойствами классификатора, их обычно называют методами фильтрации;
- (2) учитывающие характеристики последующей классификации данных, к ним относятся настраиваемые классификатор методы (НКл-методы) и формирующие классификатор методы (ФКл-методы).

Понятно, что от второй группы методов следует ждать лучшие по сравнению с методами фильтрации результаты. Настраиваемые классификатор методы (*wrapper methods*) построены так, что подмножества признаков оцениваются с учетом прогностической точности алгоритма классификации, т.е. подход зависит от классификатора. Вычислительно они более требовательны, чем методы фильтрации. Формирующие классификатор методы (*embedded methods*) отличаются тем, что поиск оптимального набора признаков встроен в структуру классификатора, а не реализуется отдельно от построения классификатора. Подход зависит от классификатора, и его можно рассматривать как поиск решения в объединенном пространстве признаков и моделей классификатора.

Обычно многие признаки оказываются неинформативными (не способствующими классификации) или избыточными (дублирующими, существенно связанными между собой признаками). Классификация может стать более продуктивной и эффективной только при использовании релевантных и неизбыточных признаков [4].

В процессе итерационного отбора признаков есть два ключевых шага — оценивание и генерирование подмножества признаков, суть которых состоит в следующем:

- оценивание определяет качество некоторого подмножества признаков; НКл- и ФКл-методы, как правило, используют меры, основанные на производительности классификатора; методы фильтрации — меры, основанные на свойствах данных;
- генерирование является средством создания некоторого подмножества признаков; процедура может основываться на простом ранжировании отдельных признаков или заключаться в коррекции текущего состава подмножества путем добавления или удаления элементов.

Для того чтобы выбрать подходящий набор признаков, необходимо средство измерения способности признака внести свой вклад в отделимость классов либо индивидуально, либо в контексте

других уже выбранных признаков, т.е. необходимы средства измерения значимости и избыточности. Меры, учитывающие правила классификации, подразумевают использование выделенного набора признаков при формировании и применении классификатора, т.е. отличающиеся наборы признаков могут давать различные классификаторы. Широко используемым примером подобных мер может служить вероятность ошибок классификатора и оценка ее значения с помощью частоты появления ошибок классификатора, обученного с помощью выбранного набора признаков.

Существуют три основные категории алгоритмов поиска для выбора подмножества признаков: гарантированный, последовательный и случайный [2].

Гарантированный поиск обеспечивает нахождение оптимального (в смысле заданного критерия) подмножества признаков. Конечно, полный (исчерпывающий) поиск является гарантированным, но стратегия поиска необязательно должна быть исчерпывающей, чтобы стать гарантированной (например, метод ветвей и границ гарантирует получение решения).

Последовательный поиск: признаки добавляются или удаляются последовательно (последовательный вперед или назад отбор). Такие методы не являются оптимальными, но они просты в реализации и быстро дают результат.

Случайный поиск подразумевает подходы, которые используют случайные механизмы при реализации указанных выше способов. Встраивание случайности в методы отбора признаков может быть полезным в следующих случаях: набор возможных подмножеств признаков велик и не доступен для обработки; детерминированные алгоритмы оказываются подвержены попаданию в ловушку локальных экстремумов критерия отбора признаков; когда выгоды от получения хорошего решения значительно перевешивают возникающие затраты, т.е. стоит потратить время на привлечение случайного механизма, чтобы перепроверить полученные результаты. Варианты случайных процедур отбора признаков приведены в [5].

По ряду причин обычно приходится отказываться от гарантированного поиска в пользу субоптимальных методов (последовательный и случайный поиск): высокие вычислительные сложности получения оптимальных решений, не всегда выполняются условия того, чтобы неисчерпывающий поиск становился гарантированным (например, условие монотонности для метода ветвей и границ).

Последовательный поиск включает в первую очередь последовательный отбор вперед (*sequential forward selection* — SFS) и последовательный отбор

назад (sequential backward selection — SBS), а также различные обобщения этих двух методов.

Метод SFS итерационно добавляет по одному новые признаки к уже сформированному набору, руководствуясь качеством получаемого решения. Основным недостатком этого метода является то, что он не включает в себя механизм для удаления ранее добавленных признаков из уже сформированного на очередном шаге набора признаков, хотя добавления могут сделать ненужными признаки, ранее уже вошедшие в формируемый набор. Метод SBS отличается от SFS тем, что признаки не добавляются, а удаляются.

В данной работе в качестве классификатора рассматривается байесовская решающая процедура, причем модель данных — смесь нормальных распределений. В качестве показателя эффективности классификации выступает вероятность правильной классификации P_c . Оценка этого показателя есть частота правильной классификации при многократном применении классификатора. Из-за высокой сложности постановок реальных задач (многомерность данных, применение итерационного EM-алгоритма для оценивания параметров смеси, ограниченность объема обучающих данных) при оценивании показателя приходится прибегать к методам управления обработкой выборки.

Для селекции признаков предлагается применять последовательный метод с настройкой классификатора для каждого апробируемого набора признаков. При этом использовался как метод SBS, так и SFS. Выбор именно селекции и последовательных методов отбора объясняется рядом причин:

- простота реализации,
- на начальном, поисковом, этапе исследований порождаются информативные наборы признаков, обеспечивающие возможность интерпретации полученных решений, в частности с позиций качества классификатора появляется возможность высказать предположения об эффективной размерности пространства признаков, выделить наименее и наиболее информативные признаки, увериться или усомниться относительно сложившихся в предметной области моделей данных.

3 Критерии значимости

Последовательные методы отбора включают генерирование варианта набора признаков и оценивание для него вероятности правильной классификации. Таким образом, при использовании определенного метода управления обработкой выборки получаем бинарную последовательность —

последовательность испытаний с двумя исходами («успех», когда классификатор не ошибся, «неуспех» в противном случае). При применении бутстреп-метода речь идет о последовательности испытаний Бернулли, тогда оценка вероятности правильной классификации есть случайная величина, имеющая биномиальное распределение. Если же привлекается метод перепроверки, то использование биномиального распределения даст некоторое приближение для распределения оценки вероятности правильной классификации, позволяющее получить первое представление о реальном качестве анализируемого набора признаков. Для обоих методов в ходе их реализации возникают последовательности наблюдаемых значений, которые могут стать источником непараметрических оценок для требуемых величин.

Дадим постановки возможных задач анализа, возникающих в связи с анализом результатов отбора признаков. Имеется k независимых биномиальных популяций X_1, \dots, X_k , распределения которых

$$X_i \sim \text{Binomial}(n_i, \pi_i), \quad i = 1, \dots, k,$$

где n_1, \dots, n_k известны и, возможно, различны, а вероятности успеха $0 \leq \pi_1, \dots, \pi_k \leq 1$ неизвестны. Со значением i связаны следующие действия: выбор определенного набора признаков, построение соответствующего классификатора на основе смеси распределений, оценивание по n_i наблюдаемым значениям вероятности правильной классификации π_i .

Для определенного значения i оценка максимального правдоподобия вероятности успеха есть $\hat{\pi}_i = X_i/n_i$. При больших значениях n_i и условиях, что ни π_i , ни $1 - \pi_i$ не являются малыми величинами, $\hat{\pi}_i$ имеет приблизительно нормальное распределение. Более практичной в случае сравнительного анализа результатов отбора признаков оказывается интервальная оценка для π_i . Используемый для этих целей доверительный интервал может быть двухсторонний или односторонний, а также принимать различную форму [6]. Наиболее распространенным при построении доверительного интервала является использование нормальной аппроксимации. При этом в случае одностороннего интервала из-за асимметрии истинного распределения ошибка может оказаться существенной, в [7] дается простое и эффективное решение этой проблемы.

Пусть классификатор оперирует с M классами, вероятности появления которых равны p_1, \dots, p_M . На практике излишней оказывается проверка того, отличается ли построенная классификация от

действий «наугад», т. е. отнесения некоторого объекта к определенному классу случайным образом и только в соответствии со значениями p_1, \dots, p_M . Дело в том, что классификация «наугад» дает $P_c = \sum_{j=1}^M p_j^2$; эта величина не меньше и может приближаться к 1 при возрастании разброса значений p_1, \dots, p_M . Таким образом, наблюдаемое кажущееся большим значение P_c может и не говорить о достоинствах принятого классификатора. Простым способом контроля значимости результатов, получаемых в ходе селекции признаков, является графический анализ зависимости односторонних доверительных интервалов от размерности признакового пространства с указанием уровня, отвечающего классификации «наугад»: выход нижней границы доверительного интервала за этот уровень говорит о том, что соответствующий вариант набора признаков вообще не информативен.

Более общий подход в анализе совокупности результатов отбора признаков заключается в построении и использовании тестов однородности биномиальных пропорций. Нулевая гипотеза об однородности проверяется против альтернативы весьма общего вида:

$$H_0 : \pi_1 = \dots = \pi_k = \pi_0$$

против

$$H_1 : \pi_i \neq \pi_j$$

для некоторых $i \neq j$, где π_0 обычно неизвестно.

Достаточно много работ посвящено разработке подобных тестов, особенно для $k = 2$, что соответствует случаю 2×2 таблицы сопряженности (например, в [8] рассматриваются 22 различных теста). При анализе результатов селекции признаков речь идет о сравнении двух определенных наборов признаков с возможностью ответить на вопрос, какой из них более информативен. Интересно, что в этом случае существует исчерпывающее решение соответствующей проблемы — равномерно наиболее мощный несмещенный (РНМН) критерий, основанный на статистике числа успехов в одной популяции при условии конкретного суммарного числа успехов в обеих популяциях [9, разд. 4.5]. Но сказать, что этот критерий широко известен и активно используется, нельзя. Причина проста: он основывается на гипергеометрическом распределении и трудоемок в применении.

Автор данной статьи занимался вопросами точных вычислений для гипергеометрического распределения и его аппроксимации с помощью биномиального, пуассоновского, нормального и бета-распределений [10], что позволило обеспечить корректное использование РНМН-критерия для сравнения двух биномиальных совокупностей.

В более общем случае $k > 2$ имеется также достаточно много критериев, построенных на различных принципах и способах практической реализации (см., например, [11, 12]); в первую очередь речь идет о следующих критериях: точные; стандартные Пирсона (Pearson) и Вилкса (Wilks); Потхоффа–Витингхилла (Potthoff–Whittinghill); Сю (Xu); Пауля и Денга (Paul and Deng), а также модификации отдельных из них. Сравнительный анализ основных процедур из приведенного перечня был проведен в [12] методом моделирования с учетом ситуаций, когда присутствуют так называемые разреженные данные (sparse data): некоторые n_i малы или некоторые из π_i близки к 0 или 1. Этот анализ позволил сформулировать следующие выводы:

- обнаружены случаи разреженных данных, для которых стандартные тесты, тест Потхоффа–Витингхилла и его модификация, тест Пауля и Денга выполняются неадекватно;
- тесты точные и Сю обладают адекватными характеристиками в любых рассмотренных условиях моделирования, при этом точные методы в целом схожи между собой, но некоторые различия возникли между точными тестами и тестом Сю.

Данные выводы основаны на результатах моделирования, поэтому всегда остается место для сомнений. Его можно рассеять только при эмпирическом сравнительном анализе критериев в рамках конкретных ограничений.

4 Эксперименты

Совместно со специалистами НИИ урологии и интервенционной радиологии им. Н. А. Лопаткина — филиал ФГБУ «НМИРЦ» Минздрава России С. А. Головановым и А. В. Сивковым проводились исследования возможности прогнозировать химический состав мочевых камней у пациентов с уролитиазом по метаболическим показателям мочи и сывортки крови.

Задача прогнозирования состава камня по набору показателей была сформулирована как задача обучаемой классификации типов камней, распадающаяся на следующие отдельные элементы.

1. Предполагается, что исследователем задана классификация типов камней по составу. Она может включать, в принципе, произвольное число классов M и должна быть четкой, т. е. любой камень по составу должен соответствовать только одному классу.

2. Для отдельных классов принимается вероятностное описание входящих в него наборов показателей. Например, если u — вектор значений всех показателей, то считается известным распределение $f_i(u)$ для каждого j -го класса. Кроме этого, предполагаются заданными вероятности p_j появления классов, $j = 1, \dots, M$. В качестве классификатора рассматривался байесовский классификатор с единичной функцией потерь.
3. Для реализации описанной схемы вместо p_j и $f_j(u)$ подставляются их оценки p_j^* и $f_j^*(u)$. Для p_j^* это не что иное, как преваленс (prevalence) — доля субъектов в популяции, которые имеют камень из i -го класса. В качестве $f_j^*(u)$ предлагается применять модель смеси нормальных распределений и EM-алгоритм для оценивания параметров этой модели.

Для решения задачи прогнозирования исходные данные о составе мочевых камней и показателях состояния пациентов представлялись как таблица «объект—признак», где объекты — пациенты, а признаки включали группу признаков, характеризующих химический состав камня пациента, а также группу метаболических и антропологических признаков пациента.

В группу признаков, характеризующих химический состав камня пациента, входили такие минеральные компоненты мочевых камней, как вееллит (WH), веделлит (WD), мочева кислота безводная (UA), мочева кислота дигидрат (UADH), аммония урат (AMUR), даллит, или карбонатапатит (Dh), брушит (BRU), струвит (STRU), цистин (CYS). Здесь и далее в подобных ситуациях в скобках приводятся общепринятые обозначения показателей.

В данной работе рассматривалась классификация типов камней по составу, описанная в табл. 1. Задание порога в 50% отражает привычное представление о классификации на основе доминирующего значения того или иного компонента.

В группе метаболических признаков пациента были представлены биохимические показатели сыворотки крови — общий кальций (Ca_ser), мочева

кислота (UA_ser), фосфаты (P_ser); биохимические показатели суточной экскреции с мочой кальция (Ca_ur), мочевои кислоты (UA_ur), фосфатов (P_ur).

Учитывая важное влияние концентраций этих веществ в конечной моче на степень перенасыщенности мочи и, следовательно, на ее литогенный потенциал, рассчитывали также концентрационные показатели (в мМол/л), такие как концентрация в моче кальция (CaUrC), мочевои кислоты (UaUrC), фосфатов (PhUrC). Помимо этого учитывали такие физико-химические показатели мочи, как удельный вес (Spec_Grav), pH мочи (pH), суточный объем мочи (Diuresis) и антропологические показатели пациента — рост (Height) и вес (Weight). Таким образом, в задаче обучаемой классификации использовались 14 показателей: Ca_ser, UA_ser, P_ser, Ca_ur, UA_ur, P_ur, CaUrC, UaUrC, PhUrC, Spec_Grav, pH, Diuresis, Height, Weight. Полученный классификатор использовался для прогноза типа камня по представляемым данным о показателях. Для того чтобы охарактеризовать качество классификатора, использовалась вероятность правильной классификации. Чем выше значение вероятности правильной классификации, тем выше точность прогноза. Анализ полученных результатов позволил сделать вывод, что прогноз возможен, но надо быть готовым к не очень высоким результатам при увеличении числа классов (в частности, от $M = 2$ к 4). Источником повышения эффективности прогноза может стать уточнение набора показателей, на основе которых строится классификация.

Оценивание вероятности правильной классификации можно проводить одним из следующих способов:

- повторно используя обучающую выборку как для построения классификатора, так и для получения P_c^* (Т-оценка, от слова thrifty);
- привлекая метод перепроверки (CV-оценка, от термина cross-validation);
- используя бутстреп-метод (В-оценка, от термина bootstrap).

В данной работе применялся классический вариант метода перепроверки: исключим из исходной

Таблица 1 Классификация типов камней по составу

Названия типа классификации и камней	Число классов	Правила классификации: номер класса, условие
Смешанная общего вида: оксалатные, уратные, фосфатные, пр.	4	1, если WH + WD > 50% 2, если UA + UADH + Amur > 50% 3, если Dh + BRU + STRU > 50% 4 иначе

совокупности данных i -й объект; для оставшегося набора данных построим байесовский классификатор и применим его к исключенному объекту; далее, вспомнив номер класса, к которому в действительности принадлежит i -й объект, сравним его с тем, который дала байесовская процедура, и получим ответ, совершила она ошибку или нет; подобные действия повторим для каждого значения i . В результате будет получена CV-оценка для P_c . Обычно считается, что она является более качественной, чем Т-оценка, обладающая «завышенным оптимизмом» (подгонка модели байесовского классификатора и оценка его качества происходит по одной и той же обучающей выборке).

Описание В-оценки начнем с модели байесовского классификатора. Предположим, что распределение данных, относящихся к j -му классу, есть смесь нормальных распределений

$$f_j(u) = \sum_{i=1}^k q_{ji} N(\mu_{ji}, \Sigma_{ji}) .$$

Тогда плотность распределения данных, подвергавшихся байесовской классификации, представима в виде

$$g(u) = \sum_{j=1}^M p_j \sum_{i=1}^k q_{ji} N(u, \mu_{ji}, \Sigma_{ji}) .$$

Но реально $g(u)$ не задана, а имеется лишь обучающая выборка, с помощью которой можно получить оценку

$$g^*(u) = \sum_{j=1}^M p_j^* \sum_{i=1}^k q_{ji}^* N(u, \mu_{ji}^*, \Sigma_{ji}^*) .$$

Таким образом, становятся заданными все элементы теперь уже эмпирического байесовского классификатора, с помощью которого любое x можно отнести к некоторому классу. Качество классификатора характеризуется с помощью вероятности правильной классификации P_c , которую посчитать аналитически затруднительно, а потому она бу-

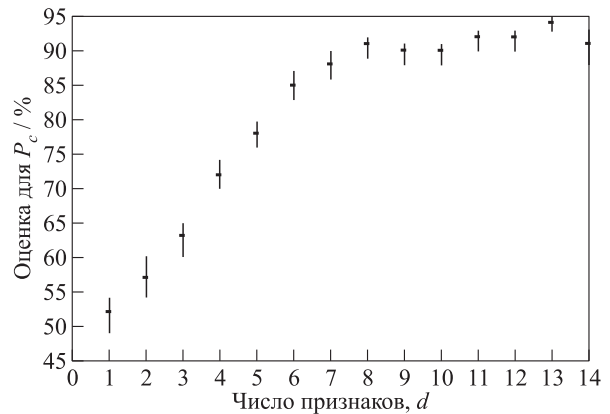


Рис. 1 Зависимость эффективности селекции от числа признаков d в случае SBS и В-оценки величины P_c (для отдельного d приведено значение оценки — горизонтальный штрих и 90%-ный доверительный интервал — вертикальный отрезок)

дет оцениваться с помощью моделирования обучающей выборки из $g^*(u)$ — так называемой бутстреп-выборки x^B . В результате получаем алгоритм бутстреп-оценивания P_c , включающий следующие шаги: оценивание по обучающей выборке параметров смеси $g^*(u)$; формирование бутстреп-выборки x^B из $g^*(u)$; классификация с помощью $g^*(u)$ данных из x^B и подсчет числа случаев правильной классификации. Такая процедура позволяет для любого набора признаков дать оценку вероятности правильной классификации. Теперь для различных наборов признаков на основе биномиального распределения можно строить необходимые доверительные интервалы, проверять гипотезу об однородности результатов анализа классификаторов для ряда наборов признаков, сравнивать отдельные варианты наборов. Результаты отбора признаков с помощью метода SBS и В-оценки для P_c представлены на рис. 1.

Результаты исследования наиболее интересных наборов признаков отражены в табл. 2. Для критерия Сю действенность предложенной им аппрок-

Таблица 2 Выбор эффективной размерности признаков

№	Набор значений d	Критерий значимости	Критический уровень значимости	Принятие нулевой гипотезы об однородности при 5%-ном уровне значимости
1	8, 9, 10, 11, 12, 13, 14	Сю	4,8%	Отвергается
2	8, 9, 10, 11, 12, 14	Сю	61,1%	Принимается
3	7, 8, 9, 10, 11, 12, 13, 14	Сю	0,0%	Отвергается
4	13, 14	РНМН	1,2%	Отвергается
5	8, 13	РНМН	1,0%	Отвергается

симации проверялась с помощью моделирования и полностью подтвердилась. Анализ табл. 2 позволяет сделать следующие практически важные выводы:

- значения размерности пространства признаков от 8 до 14 с точки зрения эффективности классификации практически эквивалентны (строки 1 и 2 табл. 2);
- снижение размерности до 7 дает значимые потери в эффективности классификации (строка 3 табл. 2);
- значение размерности 13 заслуживает пристального внимания, так как может привести к повышению эффективности классификации по сравнению с обычным использованием всех показателей (строка 4 табл. 2) или по сравнению с применением наиболее «экономичного» варианта $d = 8$ (строка 5 табл. 2).

Практически те же результаты дал и метод SFS.

Данная схема бутстреп-анализа позволяет ответить на вопрос, на что можно рассчитывать с точки зрения качества классификации при использовании байесовского классификатора на основе смеси нормальных распределений и данных, распределенных действительно как смесь нормальных распределений.

Метод перепроверки из-за зависимости отдельных экспериментов по оценке качества классификатора не позволяет полностью формально сопоставлять их результаты, но по сравнению с результатами бутстреп-анализа он дает возможность получить представление об адекватности принятой модели данных (рис. 2).

Из рис. 2 видно, что метод перепроверки дает существенно более осторожные результаты, хотя из

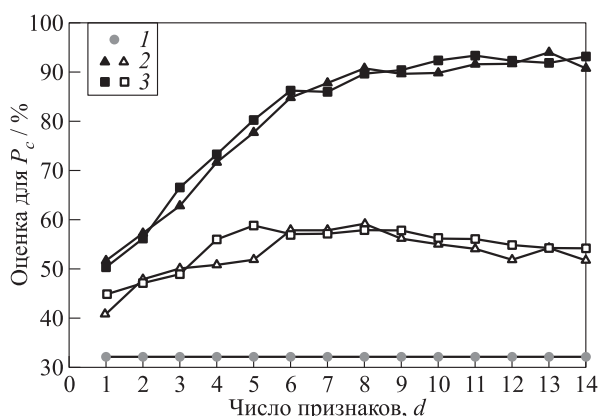


Рис. 2 Зависимость эффективности селекции от числа признаков d при классификации «наугад» (1) и с использованием байесовских классификаторов (2 — SBS; 3 — SFS; черные значки — В-оценка для P_c ; пустые значки — CV-оценка для P_c)

них опять же следует целесообразность применения для прогнозирования состава камня не всей, а лишь части совокупности показателей.

Снижение эффективности по методу перепроверки по сравнению с бутстреп-методом можно частично объяснить примененным вариантом CV-оценки и ограниченным объемом исходных данных. Для иллюстрации этого анализировалась зависимость эффективности классификации от объема обучающей выборки. Оказалось, что в рассматриваемом случае CV-оценка является нижней границей для истинного значения, а T- и B-оценки, близкие между собой, служат верхними границами. Их значения с ростом размера исходной выборки сходятся к истинной величине вероятности правильной классификации, но достигают этого значения при значительных размерах исходной выборки (более 2000 наблюдаемых значений).

5 Заключение

Существуют различные подходы к решению задачи отбора признаков, еще более богатым оказывается набор соответствующих методов. В прикладных областях к этому многообразию добавляется фактическая неоднозначность получающегося решения. Оно в принципе может оказаться формально единственным, но фактически значимо отличным от множества других. По этой причине возрастает роль критериев значимости полученных решений. Для нужд задачи классификации данных предлагается использовать такую интегральную характеристику эффективности, как вероятность правильной классификации. Для ее оценивания необходимо прибегать к анализу последовательности испытаний Бернулли, при этом полностью корректные результаты будут получены при использовании бутстреп-метода обработки исходной выборки.

Проведенные эксперименты в задаче прогнозирования химического состава мочевых камней создают предпосылки для повышения качества получаемых решений (сокращение перечня проводимых анализов может привести к росту точности прогноза), дают толчок специалистам в предметной области для прояснения сути протекающих процессов.

Литература

1. *Webb A. R., Copsey K. D. Statistical pattern recognition. — 3rd ed. — Chichester, U.K.: John Wiley & Sons, 2011. 616 p.*

2. Liu H., Yu L. Toward integrating feature selection algorithms for classification and clustering // *IEEE Trans. Knowl. Data Eng.*, 2005. Vol. 17. P. 491–502.
3. Saeyns Y., Inza I., Larrannaga P. A review of feature selection techniques in bioinformatics // *Bioinformatics*, 2007. Vol. 23. P. 2507–2517.
4. Yu L., Liu H. Efficient feature selection via analysis of relevance and redundancy // *J. Machine Learning Res.*, 2004. Vol. 5. P. 1205–1224.
5. Stracuzzi D. J. Randomized feature selection // *Computational methods of feature selection*. — Boca Raton, FL, USA: Chapman and Hall/CRC, 2007. P. 41–62.
6. Dasgupta A., Zhang T. Binomial and multinomial parameters, inference on // *Encyclopedia of statistical sciences*. — New York, NY, USA: John Wiley & Sons, 2006. P. 501–519.
7. Hall P. Improving the normal approximation when constructing one-sided confidence intervals for binomial or Poisson parameters // *Biometrika*, 1982. Vol. 69. P. 647–652.
8. Upton G. J. G. A comparison of alternative tests for the 2×2 comparative trial // *J. Roy. Stat. Soc. A*, 1982. Vol. 145. P. 86–105.
9. Lehmann E. L., Romano J. P. Testing statistical hypotheses. — 3rd ed. — New York, NY, USA: Springer, 2005. 784 p.
10. Кривенко М. П. Задачи выборочного контроля при досмотре лиц, багажа и транспорта // *Обзорные прикладной и промышленной математики*, 2011. Vol. 18. P. 125–126.
11. Potthoff R. F. Homogeneity, Potthoff–Whittighill tests of // *Encyclopedia of statistical sciences*. — New York, NY, USA: John Wiley & Sons, 2006. P. 3217–3220.
12. Klein M., Linton P. On a comparison of tests of homogeneity of binomial proportions. — Washington: Center for Statistical Research & Methodology Research and Methodology Directorate U.S. Census Bureau, 2013. <https://www.census.gov/srd/papers/pdf/rrs2013-03.pdf>.

Поступила в редакцию 14.06.16

SIGNIFICANCE TESTS OF FEATURE SELECTION FOR CLASSIFICATION

M. P. Krivenko

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

Abstract: The paper considers the problem of feature selection for classification and issues related to the assessment of the quality of the solutions. Among the different methods of feature selection, attention is paid to sequential procedures; the probability of the correct classification is used to measure the quality of the classification. To evaluate this indicator, it is proposed to use cross-validation and the bootstrap method. At the same time, to investigate the set of sample values of probability of the correct classification, it is suggested to use comparative analysis of confidence intervals and the test for homogeneity of binomial proportions. While constructing Bayesian classifier as the data model mixture of normal distributions is adopted, the model parameters are estimated by the expectation–maximization algorithm. As an experiment, the paper considers the problem of well-thoughtout choice of classification characteristics when predicting the type of urinary stones in urology. It is demonstrated that the set of used features can be reduced not only without losing the quality of decisions, but also with increase of probability of correct prediction of the stone type.

Keywords: feature selection; sequential forward and backward selections; Bayes classification; test of homogeneity of binomial proportions; prediction of stone types in urology

DOI: 10.14357/19922264160305

References

1. Webb, A. R., and K. D. Copsey. 2011. *Statistical pattern recognition*. 3rd ed. Chichester, U.K.: John Wiley & Sons. 616 p.
2. Liu, H., and L. Yu. 2005. Toward integrating feature selection algorithms for classification and clustering. *IEEE Trans. Knowl. Data Eng.* 17:491–502.
3. Saeyns, Y., I. Inza, and P. Larrannaga. 2007. A review of feature selection techniques in bioinformatics. *Bioinformatics* 23:2507–2517.
4. Yu, L., and H. Liu. 2004. Efficient feature selection via analysis of relevance and redundancy. *J. Machine Learning Res.* 5:1205–1224.
5. Stracuzzi, D. J. 2007. Randomized feature selection. *Computational methods of feature selection*. Boca Raton, FL: Chapman and Hall/CRC. 41–62.
6. Dasgupta, A., and T. Zhang. 2006. Binomial and multinomial parameters, inference on. *Encyclopedia of statistical sciences*. New York, NY: John Wiley & Sons. 501–519.

7. Hall, P. 1982. Improving the normal approximation when constructing one-sided confidence intervals for binomial or Poisson parameters. *Biometrika* 69:647–652.
8. Upton, G. J. G. 1982. A comparison of alternative tests for the 2×2 comparative trial. *J. Roy. Stat. Soc. A* 145:86–105.
9. Lehmann, E. L., and J. P. Romano. 2005. *Testing statistical hypotheses*. 3rd ed. New York, NY: Springer. 784 p.
10. Krivenko, M. P. 2011. Zadachi vyborochnogo kontrolya pri dosmotre lits, bagazha i transporta [Tasks of sampling during the inspection of individuals, baggage and transport]. *Obozrenie prikladnoy i promyshlennoy matematiki* [Review of applied and industrial mathematics] 18:125–126.
11. Potthoff, R. F. 2006. Homogeneity, Potthoff–Whittighill tests of. *Encyclopedia of statistical sciences*. New York, NY: John Wiley & Sons. 3217–3220.
12. Klein, M., and P. Linton. 2013. On a comparison of tests of homogeneity of binomial proportions. Center for Statistical Research & Methodology Research and Methodology Directorate U.S. Census Bureau Washington. Available at: <https://www.census.gov/srd/papers/pdf/rrs2013-03.pdf> (accessed April 25, 2016).

Received June 14, 2016

Contributors

Krivenko Mikhail P. (b. 1946) — Doctor of Science in technology, professor, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; mkrivenko@ipiran.ru

УСИЛЕННЫЙ ЗАКОН БОЛЬШИХ ЧИСЕЛ ДЛЯ ОЦЕНКИ РИСКА В ЗАДАЧЕ РЕКОНСТРУКЦИИ ТОМОГРАФИЧЕСКИХ ИЗОБРАЖЕНИЙ ИЗ ПРОЕКЦИЙ С КОРРЕЛИРОВАННЫМ ШУМОМ*

О. В. Шестаков¹

Аннотация: Методы вейвлет-анализа, основанные на процедуре пороговой обработки коэффициентов разложения проекций, широко используются при решении задач реконструкции томографических изображений, возникающих в медицине, биологии, астрономии и других областях. Их привлекательность заключается, во-первых, в скорости алгоритмов, а во-вторых, в возможности реконструкции локальных участков изображения по неполным проекционным данным, что имеет ключевое значение, например, для медицинских приложений, где пациента нежелательно подвергать лишней дозе облучения. Анализ погрешностей этих методов представляет собой важную практическую задачу, поскольку позволяет оценить качество как самих методов, так и используемого оборудования. В работе рассматривается метод вейвлет-вейглет-разложения при реконструкции томографических изображений в модели с коррелированным аддитивным шумом. Доказывается, что для несмещенной оценки среднеквадратичного риска при пороговой обработке коэффициентов вейвлет-вейглет-разложения функции изображения выполняется усиленный закон больших чисел, т. е. эта оценка является сильно состоятельной.

Ключевые слова: вейвлеты; пороговая обработка; оценка среднеквадратичного риска; преобразование Радона

DOI: 10.14357/19922264160306

1 Введение

Методы реконструкции изображений поглощающих, излучающих и отражающих объектов и сред получили широкое распространение в самых разнообразных областях, включая медицину, биологию, физику плазмы, газовую динамику, геофизику, астрономию и радиолокацию. В частности, эти методы революционизировали медицинскую диагностику, поскольку позволили получать изображения внутренних органов человека без оперативного вмешательства. При построении и анализе статистической модели данных, получаемых в ходе томографического эксперимента, необходимо учитывать погрешности, которые возникают в реальных экспериментах из-за дискретизации исходной информации, несовершенства оборудования, характеристик используемого излучения, наличия фонового излучения и внутренних шумов приемника и т. д. В данной работе рассматривается модель стационарного коррелированного шума и исследуются свойства оценки риска в методе пороговой обработки коэффициентов вейвлет-вейглет-разложения функции томографического изобра-

жения. Такая оценка позволяет анализировать качество полученного изображения на основе только наблюдаемых данных, что очень важно, например, в медицинских приложениях. В работах [1, 2] показано, что при определенных условиях данная оценка является состоятельной и асимптотически нормальной. В данной работе доказывается свойство сильной состоятельности.

2 Преобразование Радона и его обращения с помощью метода вейвлет-вейглет-разложения

Преобразование Радона на плоскости определяется как набор интегралов от функции f по всевозможным прямым:

$$\text{Rf}(s, \theta) = \int_{L_{s, \theta}} f(x, y) dl,$$

где

$$L_{s, \theta} = \{(x, y) : x \cos \theta + y \sin \theta - s = 0\}.$$

*Работа выполнена при частичной финансовой поддержке РФФИ (проект 15-07-02652).

¹Московский государственный университет им. М. В. Ломоносова, кафедра математической статистики факультета вычислительной математики и кибернетики; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, oshestakov@cs.msu.su

Задача реконструкции томографического изображения состоит в обращении преобразования R . Эту задачу можно решать разными способами. Воспользуемся методом вейвлет-вейглет-разложения, предложенным в работе [3], который позволяет учитывать различные локальные особенности изображений.

Пусть задана масштабирующая функция $\phi(x)$ и вейвлет-функция $\psi(x)$. Определим функции

$$\left. \begin{aligned} \psi_{j,k_1,k_2}^{[1]}(x,y) &= 2^j \phi(2^j x - k_1) \psi(2^j y - k_2); \\ \psi_{j,k_1,k_2}^{[2]}(x,y) &= 2^j \psi(2^j x - k_1) \phi(2^j y - k_2); \\ \psi_{j,k_1,k_2}^{[3]}(x,y) &= 2^j \psi(2^j x - k_1) \psi(2^j y - k_2), \end{aligned} \right\} (1)$$

где $j, k_1, k_2 \in \mathbb{Z}$. Семейство $\{\psi_{j,k_1,k_2}^{[\lambda]}\}_{\lambda,j,k_1,k_2}$ образует ортонормированный базис в $L^2(\mathbb{R}^2)$. Индекс j в (1) называется масштабом, а индексы k_1 и k_2 — сдвигами. Функция ψ должна удовлетворять определенным требованиям, однако ее можно выбрать таким образом, чтобы она обладала некоторыми полезными свойствами, например была дифференцируемой нужное число раз и имела заданное число M нулевых моментов, т. е.

$$\int_{-\infty}^{\infty} x^k \psi(x) dx = 0, \quad k = 0, \dots, M-1.$$

В данной работе предполагается, что используются вейвлеты Мейера [4], преобразование Фурье которых обладает достаточным количеством непрерывных производных.

Вейвлет-разложение функции f имеет вид:

$$f = \sum_{\lambda,j,k_1,k_2} \langle f, \psi_{j,k_1,k_2}^{[\lambda]} \rangle \psi_{j,k_1,k_2}^{[\lambda]}. \quad (2)$$

Идея метода вейвлет-вейглет-разложения заключается в том, чтобы заменить скалярные произведения в (2) на величины, выражающиеся через Rf , а не через f . Определим последовательность функций

$$\xi_{j,k_1,k_2}^{[\lambda]} = \frac{1}{4\pi} I^{-1} R \psi_{j,k_1,k_2}^{[\lambda]},$$

где I^p — потенциал Рисса, определяемый через преобразование Фурье по формуле $\widehat{I^p g}(\omega) = |\omega|^{-p} \widehat{g}(\omega)$. В [5] показано, что последовательность функций $\{2^{-j/2} \xi_{j,k_1,k_2}^{[\lambda]}\}_{\lambda,j,k_1,k_2}$ образует устойчивый базис. При этом выполнено $\langle f, \psi_{j,k_1,k_2}^{[\lambda]} \rangle = \langle Rf, \xi_{j,k_1,k_2}^{[\lambda]} \rangle$. Функции $2^{-j/2} \xi_{j,k_1,k_2}^{[\lambda]}$ получили название «вейглеты» за свою схожесть с вейвлетами. Таким образом, формула обращения в методе вейвлет-вейглет-разложения выглядит следующим образом:

$$f = \sum_{\lambda,j,k_1,k_2} \langle Rf, \xi_{j,k_1,k_2}^{[\lambda]} \rangle \psi_{j,k_1,k_2}^{[\lambda]}.$$

3 Модель данных и оценка риска

На практике функция f обычно задана в дискретных отсчетах на круге. Без ограничения общности будем считать, что это круг единичного радиуса с центром в начале координат. Проекционные данные в этом случае измеряются при $(s, \theta) \in [-1, 1] \times [0, \pi]$. Пусть $\{e_{i,j}, i, j \in \mathbb{Z}\}$ — стационарный гауссовский процесс с ковариационной последовательностью $r_{k,p} = \mathbf{cov}(e_{i,j}, e_{i+k,j+p})$. Модель проекционных данных с шумом выглядит следующим образом:

$$Y_{i,j} = Rf_{i,j} + e_{i,j} \quad i = 1, \dots, 2^J, \quad j = 1, \dots, 2^J,$$

где 2^J — число отсчетов.

Структура ковариации шума для преобразования Радона должна отражать типичную ситуацию: на практике проекции для разных углов регистрируются независимо друг от друга. В рассматриваемой модели ошибок предполагаются независимые наблюдения в случае разных углов и стационарный гауссовский шум с нулевым математическим ожиданием, конечной дисперсией и ковариационной последовательностью $r_\delta \sim A\delta^{-\alpha}$ ($0 < \alpha < 1$) для одинаковых углов (так называемая долгосрочная зависимость).

При переходе к дискретному вейвлет-вейглет-преобразованию получается следующая модель дискретных коэффициентов [2]:

$$X_{j,k_1,k_2}^{[\lambda]} = \mu_{j,k_1,k_2}^{[\lambda]} + 2^{(1-\alpha)J/2} e_{j,k_1,k_2}^{[\lambda]}, \quad (3)$$

где $\mu_{j,k_1,k_2}^{[\lambda]} = 2^J \langle Rf, \xi_{j,k_1,k_2}^{[\lambda]} \rangle$, $e_{j,k_1,k_2}^{[\lambda]} = \int \xi_{j,k_1,k_2}^{[\lambda]} d\mathbf{B}'_H$, а $\mathbf{B}'_H(s, \theta)$ — случайная функция, которая для каждого фиксированного угла θ представляет собой дробное броуновское движение $\mathbf{B}_H(s)$ с $H = 1 - \alpha/2$ и имеет некоррелированные приращения по θ . Дисперсии коэффициентов модели имеют вид:

$$\sigma_{\lambda,j}^2 = C_{\lambda,\alpha} 2^{(1-\alpha)J} 2^{j\alpha},$$

где константы $C_{\lambda,\alpha}$ ($\lambda = 1, 2, 3$) зависят от параметров A, α и выбранного вейвлет-базиса.

Для подавления шума используется процедура пороговой обработки. Ее смысл заключается в удалении достаточно маленьких коэффициентов вейвлет-вейглет-разложения, которые считаются шумом. В данной работе рассматривается так называемая мягкая пороговая обработка с порогом $T_{\lambda,j}$, зависящим от масштаба j . К каждому коэффициенту применяется функция $\rho_{T_{\lambda,j}}(x) = \mathbf{sgn}(x) (|x| - T_{\lambda,j})_+$, т. е. коэффициенты, которые по модулю меньше порога $T_{\lambda,j}$, обнуляются,

а абсолютные величины остальных коэффициентов уменьшаются на величину порога. Погрешность (или риск) мягкой пороговой обработки определяется следующим образом:

$$R_J(f) = \sum_{j=0}^{J-1} \sum_{k_1=0}^{2^j-1} \sum_{k_2=0}^{2^j-1} \sum_{\lambda=1}^3 \mathbb{E} \left(\mu_{j,k_1,k_2}^{[\lambda]} - \rho_{T_{\lambda,j}}(X_{j,k_1,k_2}^{[\lambda]}) \right)^2. \quad (4)$$

В [6] предложено использовать порог $T_{\lambda,j} = \sqrt{2 \ln 2^{2j}} \sigma_{\lambda,j}$, названный универсальным. В дальнейшем будет использоваться именно такой вид порога. В выражении (4) присутствуют неизвестные величины $\mu_{j,k_1,k_2}^{[\lambda]}$, поэтому вычислить значение $R_J(f)$ нельзя. Однако его можно оценить. В качестве оценки риска используется следующая величина [7]:

$$\widehat{R}_J(f) = \sum_{j=0}^{J-1} \sum_{k_1=0}^{2^j-1} \sum_{k_2=0}^{2^j-1} \sum_{\lambda=1}^3 F[(X_{j,k_1,k_2}^{[\lambda]})^2, T_{\lambda,j}], \quad (5)$$

где $F(x, T) = (x - \sigma^2) \mathbf{1}(|x| \leq T) + (\sigma^2 + T^2) \mathbf{1}(|x| > T)$. Величина $\widehat{R}_J(f)$ является несмещенной оценкой для $R_J(f)$ [4].

4 Сильная состоятельность оценки риска

В работе [1] исследовались асимптотические свойства оценки (5) в модели с независимым шумом. Было показано, что при определенных условиях гладкости эта оценка является состоятельной и асимптотически нормальной. В модели с коррелированным шумом аналогичные результаты получены в работе [2]. В данной работе доказывается сильная состоятельность оценки (5) в модели (3).

Приведем вспомогательное утверждение [8], в котором оценивается вероятность отклонения суммы ограниченных слабовзависимых случайных величин от ее математического ожидания.

Лемма. Пусть $\{X_i, i \in Z\}$ — последовательность случайных величин таких, что $\mathbb{E}X_i = 0$ и $|X_i| \leq b$ п.в. для всех $i \in Z$, где $b > 0$ — некоторая константа. Тогда для любого $q \in [1, n/2]$ и любого $\varepsilon > 0$

$$\mathbb{P} \left(\left| \sum_{i=1}^n X_i \right| > n\varepsilon \right) \leq 4 \exp \left\{ -\frac{\varepsilon^2}{8b^2} q \right\} + 22 \left(1 + \frac{4b}{\varepsilon} \right)^{1/2} q \alpha \left(\left\lceil \frac{n}{2q} \right\rceil \right), \quad (6)$$

где $\alpha(k)$ — коэффициент α -перемешивания последовательности $\{X_i, i \in Z\}$.

Докажем теперь сильную состоятельность оценки (5).

Теорема. Пусть $f \in L^2(\mathbb{R}^2)$ задана в круге единичного радиуса с центром в начале координат. Тогда имеет место сходимость

$$\frac{\widehat{R}_J(f) - R_J(f)}{2^{\lambda J}} \rightarrow 0 \text{ п.в. при } J \rightarrow \infty \quad (7)$$

для любого $\lambda > 2$.

Доказательство: Обозначим

$$f_{j,k_1,k_2}^{[\lambda]} = F \left[\left(X_{j,k_1,k_2}^{[\lambda]} \right)^2, T_{\lambda,j} \right] - \mathbb{E} F \left[\left(X_{j,k_1,k_2}^{[\lambda]} \right)^2, T_{\lambda,j} \right].$$

Выберем произвольное $0 < p < 1$ и разобьем числитель в (7) на две суммы:

$$\widehat{R}_J(f) - R_J(f) = \widehat{R}_1 + \widehat{R}_2,$$

где

$$\widehat{R}_1 = \sum_{j=0}^{[pJ]-1} \sum_{k_1=0}^{2^j-1} \sum_{k_2=0}^{2^j-1} \sum_{\lambda=1}^3 f_{j,k_1,k_2}^{[\lambda]};$$

$$\widehat{R}_2 = \sum_{j=[pJ]}^{J-1} \sum_{k_1=0}^{2^j-1} \sum_{k_2=0}^{2^j-1} \sum_{\lambda=1}^3 f_{j,k_1,k_2}^{[\lambda]}.$$

Сначала рассмотрим \widehat{R}_2 . Для произвольного $\delta > 0$ имеем:

$$p_J = \mathbb{P} \left(\left| \widehat{R}_2 \right| > \delta 2^{\lambda J} \right) \leq \sum_{j=[pJ]}^{J-1} \mathbb{P} \left(\left| \sum_{k_1=0}^{2^j-1} \sum_{k_2=0}^{2^j-1} \sum_{\lambda=1}^3 f_{j,k_1,k_2}^{[\lambda]} \right| > \delta J^{-1} 2^{\lambda J} \right). \quad (8)$$

Величины $F[(X_{j,k_1,k_2}^{[\lambda]})^2, T_{\lambda,j}]$ ограничены:

$$-\sigma_{\lambda,j}^2 \leq F \left[\left(X_{j,k_1,k_2}^{[\lambda]} \right)^2, T_{\lambda,j} \right] \leq \sigma_{\lambda,j}^2 + T_{\lambda,j}^2.$$

Повторяя рассуждения работы [9], можно показать, что в силу свойств вейвлетов Мейера при каждом j слагаемые в сумме под вероятностью в (8) удовлетворяют свойству ρ -перемешивания с коэффициентом $\rho(k) \leq Ck^{-M}$, где $C > 0$ — некоторая константа, а M зависит от выбранного вейвлета Мейера и может быть сделано достаточно большим.

Для коэффициентов α -перемешивания и ρ -перемешивания справедливо неравенство $4\alpha(k) \leq \rho(k)$ [10]. Применяя неравенство (6) с $q = 2^{2\theta j}$

($\theta < 1$) для каждого j в сумме (8) и выбирая такой вейвлет Мейера, чтобы M было достаточно большим, получаем:

$$p_J \leq c_1 J \times \max_{[pJ] \leq j \leq J-1} \left\{ \exp \left[-c_2 J^{-3} 2^{2(\lambda-1+\alpha)J + (2\theta-2\alpha-4)j} \right] \right\} + o_J. \quad (9)$$

Здесь c_1 и c_2 — некоторые положительные константы, а o_J убывает по J быстрее, чем $2^{-M_0 p J}$, где M_0 — некоторое положительное число, зависящее от M .

Имеем $2\theta - 2\alpha - 4 < 0$, и при $j = = J$ правая часть (9) не превосходит $c_1 J \exp \left[-c_2 J^{-3} 2^{(2\lambda-6+2\theta)J} \right]$. Поскольку $\theta < 1$ можно выбрать произвольно, для того чтобы выполнялось неравенство $2\lambda - 6 + 2\theta > 0$, достаточно потребовать $\lambda > 2$. При таком выборе λ

$$\sum_{J=1}^{\infty} p_J < \infty$$

и в силу леммы Бореля—Кантелли для любого $\delta > 0$ событие $\left\{ \left| \widehat{R}_2 \right| > \delta 2^{\lambda J} \right\}$ осуществляется лишь конечное число раз. Следовательно, $\widehat{R}_2 2^{-\lambda J} \rightarrow 0$ п. в.

Рассмотрим теперь \widehat{R}_1 . Общее число слагаемых в \widehat{R}_1 не превосходит $2^{2[pJ]}$. Число слагаемых при каждом j равно $3 \cdot 2^{2j}$, и при фиксированном j каждое слагаемое не превосходит по модулю $B J \cdot 2^{J(1-\alpha)+\alpha j}$, где $B > 0$ — некоторая константа. Следовательно, $\left| \widehat{R}_1 \right| \leq B_1 J \cdot 2^{J(1-\alpha+(\alpha+2)p)}$ п. в., где $B_1 > 0$ — некоторая константа. При любых $0 < \alpha < < 1$ и $\lambda > 2$ можно выбрать такое $0 < p < 1$, что $\lambda - 1 + \alpha - (\alpha + 2)p > 0$. Следовательно, $\widehat{R}_1 \cdot 2^{-\lambda J} \rightarrow 0$ п. в. Теорема доказана.

Литература

1. Маркин А. В., Шестаков О. В. Асимптотики оценки риска при пороговой обработке вейвлет-вейвлет коэффициентов в задаче томографии // Информатика и её применения, 2010. Т. 4. Вып. 2. С. 36–45.
2. Ерошенко А. А., Шестаков О. В. Асимптотические свойства оценки риска в задаче восстановления изображения с коррелированным шумом при обращении преобразования Радона // Информатика и её применения, 2014. Т. 8. Вып. 4. С. 32–40.
3. Donoho D. Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition // Appl. Comput. Harmonic Anal., 1995. Vol. 2. P. 101–126.
4. Mallat S. A wavelet tour of signal processing. — Academic Press, 1999. 857 p.
5. Lee N. Wavelet-vaguelette decompositions and homogeneous equations. Purdue University, 1997. PhD Thesis. 93 p.
6. Kolaczyk E. D. Wavelet methods for the inversion of certain homogeneous linear operators in the presence of noisy data. Stanford University, 1994. PhD Thesis. 163 p.
7. Donoho D., Johnstone I. M. Adapting to unknown smoothness via wavelet shrinkage // J. Amer. Stat. Assoc., 1995. Vol. 90. P. 1200–1224.
8. Bosq D. Nonparametric statistics for stochastic processes: Estimation and prediction. — New York, NY, USA: Springer-Verlag, 1996. 169 p.
9. Johnstone I. M. Wavelet shrinkage for correlated data and inverse problems: Adaptivity results // Statistica Sinica, 1999. Vol. 9. No. 1. P. 51–83.
10. Bradley R. C. Basic properties of strong mixing conditions. A survey and some open questions // Probab. Surveys, 2005. Vol. 2. P. 107–144.

Поступила в редакцию 01.07.16

THE STRONG LAW OF LARGE NUMBERS FOR THE RISK ESTIMATE IN THE PROBLEM OF TOMOGRAPHIC IMAGE RECONSTRUCTION FROM PROJECTIONS WITH A CORRELATED NOISE

O. V. Shestakov^{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: Methods of wavelet analysis based on thresholding of coefficients of the projection decomposition are widely used for solving the problems of tomographic image reconstruction in medicine, biology, astronomy, and other areas. These methods are easily implemented through fast algorithms; so, they are very appealing in practical

situations. Besides, they allow the reconstruction of local parts of the images using incomplete projection data, which is essential, for example, for medical applications, where it is not desirable to expose the patient to the redundant radiation dose. Wavelet thresholding risk analysis is an important practical task, because it allows determining the quality of the techniques themselves and of the equipment which is being used. The present paper considers the problem of estimating the function by inverting the Radon transform in the model of data with correlated noise. The paper considers the wavelet-vaguelette decomposition method of reconstructing tomographic images in the model with a correlated noise. It is proven that the unbiased mean squared error risk estimate for thresholding wavelet-vaguelette coefficients of the image function satisfies the strong law of large numbers, i. e., it is a strongly consistent estimate.

Keywords: wavelets; thresholding; MSE risk estimate; Radon transform

DOI: 10.14357/19922264160306

Acknowledgments

The work was partly supported by the Russian Foundation for Basic Research (project 15-07-02652).

References

1. Markin, A. V., and O. V. Shestakov. 2010. Asimptotiki otsenki riska pri porogovoy obrabotke veyvlet-veyglet ko-effitsientov v zadache tomografii [The asymptotics of risk estimate for wavelet-vaguelette coefficients' thresholding in the problems of tomography]. *Informatika i ee Primeneniya — Inform. Appl.* 4(2):36–45.
2. Eroshenko, A. A., and O. V. Shestakov. 2014. Asimptoticheskie svoystva otsenki riska v zadache vosstanovleniya izobrazheniya s korrelirovannym shumom pri obrashchenii preobrazovaniya Radona [Asymptotic properties of risk estimate in the problem of reconstructing images with correlated noise by inverting the Radon transform]. *Informatika i ee Primeneniya — Inform. Appl.* 8(4):32–40.
3. Donoho, D. 1995. Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition. *Appl. Comput. Harmonic Anal.* 2:101–126.
4. Mallat, S. 1999. *A wavelet tour of signal processing*. New York, NY: Academic Press. 857 p.
5. Lee, N. 1997. Wavelet-vaguelette decompositions and homogeneous equations. West Lafayette: Purdue University. PhD Thesis. 93 p.
6. Kolaczyk, E. D. 1994. Wavelet methods for the inversion of certain homogeneous linear operators in the presence of noisy data. Stanford: Stanford University. PhD Thesis. 163 p.
7. Donoho, D., and I. M. Johnstone. 1995. Adapting to unknown smoothness via wavelet shrinkage. *J. Amer. Stat. Assoc.* 90:1200–1224.
8. Bosq, D. 1996. *Nonparametric statistics for stochastic processes: Estimation and prediction*. New York, NY: Springer-Verlag. 169 p.
9. Johnstone, I. M. 1999. Wavelet shrinkage for correlated data and inverse problems: Adaptivity results. *Statistica Sinica* 9(1):51–83.
10. Bradley, R. C. 2005. Basic properties of strong mixing conditions. A survey and some open questions. *Probab. Surveys* 2:107–144.

Received July 01, 2016

Contributor

Shestakov Oleg V. (b. 1976) — Doctor of Science in physics and mathematics, associate professor, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; oshestakov@cs.msu.su

АНАЛИЗ ТОЧНОСТИ ВЕЙВЛЕТ-ОБРАБОТКИ АЭРОДИНАМИЧЕСКИХ КАРТИН ОБТЕКАНИЯ

Т. В. Захарова¹, О. В. Шестаков²

Аннотация: Предлагается новый метод обработки зашумленных аэродинамических картин обтекания, основанный на технике вейвлет-анализа. Вейвлет-методы подавления шума, основанные на процедуре пороговой обработки, широко используются при анализе сигналов и изображений. Их привлекательность заключается, во-первых, в скорости алгоритмов построения оценок, а во-вторых, в возможности лучшей, чем линейные методы, адаптации к функциям, имеющим на разных участках различную степень регулярности. Анализ погрешностей этих методов представляет собой важную практическую задачу, поскольку он позволяет оценить качество как самих методов, так и используемого оборудования. Проведена верификация метода на основе сравнительного анализа с результатами обработки ранее разработанным дискриминантным методом. Рассчитанная оценка погрешности обработки при этом согласуется с теоретическими результатами для этой оценки.

Ключевые слова: вейвлет-анализ; пороговая обработка; несмещенная оценка риска; аэродинамический поток

DOI: 10.14357/19922264160307

1 Введение

Обработка результатов экспериментов в аэродинамической трубе требует исключительной точности. Результаты таких испытаний влияют на кинематические возможности проектируемой техники, экономичность и безопасность ее использования. Повысить точность можно путем усложнения экспериментальной установки либо использованием специализированной обработки результатов эксперимента. Первый путь, как правило, весьма дорог и сложен: для аэродинамических экспериментов это означает улучшение системы фильтрации воздуха и поддержания постоянства его температуры, влажности и т. д., усложнение оптических систем для снятия характеристик обтекания. Второй требует разработки и верификации методики оценки истинных значений по имеющимся данным — зачастую это сделать проще и дешевле; таким образом, разработка специальных алгоритмов обработки данных весьма актуальна в данной области.

В рамках данной работы непосредственные физические характеристики газовой струи заменены условным эквивалентом — цветом на теневой картине обтекания. Задача состоит в очистке данного сигнала от шума, возникающего вследствие нали-

чия в газовой струе и на оптике аппаратуры фото/видеосъемки пыли и т. п. [1, 2]. Для решения данной задачи в работе используются методы вейвлет-анализа, которые прекрасно зарекомендовали себя при анализе и обработке нестационарных сигналов и изображений [3, 4].

2 Вейвлет-методы обработки изображений

2.1 Пороговая обработка вейвлет-коэффициентов изображения

Большинство алгоритмов подавления шума, использующие методы вейвлет-анализа, действуют по принципу «вейвлет-преобразование — обработка вейвлет-коэффициентов — обратное вейвлет-преобразование». В данной работе рассматривается один из способов обработки вейвлет-коэффициентов — так называемая мягкая пороговая обработка.

Пусть $N = 2^J$ для некоторого $J > 0$; $f(i, j)$, $0 \leq i, j < N$, — матрица изображения; $\{W(i, j)\}_{0 \leq i, j < N}$ — помехи (шум), являющиеся реализацией некоторого случайного процесса. В данной работе предполагается, что $\{W(i, j)\}_{0 \leq i, j < N}$ —

¹Московский государственный университет им. М. В. Ломоносова, факультет вычислительной математики и кибернетики; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, lsa@cs.msu.ru

²Московский государственный университет им. М. В. Ломоносова, кафедра математической статистики факультета вычислительной математики и кибернетики; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, oshestakov@cs.msu.ru

независимые одинаково распределенные нормальные случайные величины с нулевым средним и дисперсией σ^2 . Искаженный сигнал определяется формулой:

$$X(i, j) = f(i, j) + W(i, j), \quad 0 \leq i, j < N.$$

После применения дискретного вейвлет-преобразования $Y = W(X)$ [4] получается следующая модель эмпирических вейвлет-коэффициентов:

$$Y^{[\lambda]}(j, k, l) = f_W^{[\lambda]}(j, k, l) + Z^{[\lambda]}(j, k, l), \\ \lambda = 1, 2, 3, \quad 0 \leq j < J, \quad 0 \leq k, l < 2^j,$$

где $f_W(j, k, l)$ — вейвлет-коэффициенты «чистого» изображения, а шумовые коэффициенты $Z^{[\lambda]}(j, k, l)$ в силу ортогональности вейвлет-преобразования имеют такую же статистическую структуру, как $W(i, j)$. Индекс j называется масштабом и отвечает за «размер» вейвлет-коэффициента (размер участка изображения, за который отвечает данный коэффициент), индексы k и l — сдвиги, отвечающие за «положение» вейвлет-коэффициента (расположение участка изображения, за который отвечает данный коэффициент), а индекс λ описывает тип вейвлет-коэффициента: при $\lambda = 1$ вейвлет-коэффициенты называются вертикальными, при $\lambda = 2$ — горизонтальными, а при $\lambda = 3$ — диагональными [3, 5, 6].

Оценки вейвлет-коэффициентов изображения $\hat{f}_W(j, k, l)$ вычисляются покомпонентно, т. е.

$$\hat{f}_W^{[\lambda]}(j, k, l) = \rho_T(Y^{[\lambda]}(j, k, l)),$$

где $\rho_T(x)$ — функция мягкой пороговой обработки с порогом $T > 0$, задаваемая формулой:

$$\rho_T(x) = \begin{cases} x - T, & \text{если } x > T; \\ x + T, & \text{если } x < -T; \\ 0, & \text{если } |x| \leq T. \end{cases}$$

Таким образом, в результате пороговой обработки обнуляются те вейвлет-коэффициенты, абсолютная величина которых не превосходит порога, а абсолютная величина остальных вейвлет-коэффициентов уменьшается на величину порога.

Риск (среднеквадратичная погрешность) мягкой пороговой обработки определяется формулой:

$$r_f = \sum_{\lambda=1}^3 \sum_{j=0}^{J-1} \sum_{i,j=0}^{2^j-1} \mathbb{E} \left| \hat{f}_W^{[\lambda]}(j, k, l) - f_W^{[\lambda]}(j, k, l) \right|^2. \quad (1)$$

Выражение (1) содержит неизвестные величины «чистых» вейвлет-коэффициентов $f_W^{[\lambda]}(j, k, l)$, поэтому на практике вычислить его нельзя. Однако

его можно оценить, используя только известные эмпирические вейвлет-коэффициенты $Y^{[\lambda]}(j, k, l)$. В каждом слагаемом если $|Y^{[\lambda]}(j, k, l)| > T$, то вклад этого слагаемого в риск составляет $\sigma^2 + T^2$, а если $|Y^{[\lambda]}(j, k, l)| \leq T$, то вклад составляет $[f_W^{[\lambda]}(j, k, l)]^2$. Поскольку $\mathbb{E}[Y^{[\lambda]}(j, k, l)]^2 = \sigma^2 + [f_W^{[\lambda]}(j, k, l)]^2$, величину $[f_W^{[\lambda]}(j, k, l)]^2$ можно оценить разностью $[Y^{[\lambda]}(j, k, l)]^2 - \sigma^2$.

Таким образом, в качестве оценки риска можно использовать следующую величину:

$$\hat{r}_f = \sum_{\lambda=1}^3 \sum_{j=0}^{J-1} \sum_{i,j=0}^{2^j-1} F[[Y^{[\lambda]}(j, k, l)]^2], \quad (2)$$

где $F[x] = (x - \sigma^2)\mathbf{1}(|x| \leq T^2) + (\sigma^2 + T^2)\mathbf{1}(|x| > T^2)$. Донохо и Джонстон показали [7, 8], что при мягкой пороговой обработке для оценки риска (2) справедливо следующее утверждение.

Лемма 1. $\mathbb{E}\hat{r}_f = r_f$, т. е. \hat{r}_f является несмещенной оценкой для r_f .

Исследуем теперь вопрос выбора порога.

2.2 Предпосылки выбора порога вейвлет-обработки

Лемма 2. Пусть z_1, z_2, \dots, z_n — независимые случайные величины, имеющие нормальное распределение с нулевым средним и дисперсией σ^2 ;

$$A_n = \left\{ \max_{1 \leq i \leq n} |z_i| > \sigma\sqrt{2 \ln n} \right\};$$

$$B_n(t) = \left\{ \max_{1 \leq i \leq n} |z_i| > \sigma t + \sigma\sqrt{2 \ln n} \right\}.$$

Тогда

$$P(A_n) \rightarrow 0 \text{ при } n \rightarrow \infty; \quad P(B_n(t)) < e^{-t^2/2}.$$

Порог $T = \sigma\sqrt{2 \ln n}$ согласно лемме 2 с большой вероятностью удаляет основной шум, а оставшийся шум будет незначительным, так как вероятность B_n экспоненциально убывает.

На практике часто приходится оценивать значение σ . Обычно в качестве такой оценки используется величина

$$\hat{\sigma} = \frac{1}{C_{3/4}} \text{MAD}, \quad (3)$$

где MAD (Median Absolute Deviation) — выборочное медианное абсолютное отклонение, построенное по вейвлет-коэффициентам при $j = J - 1$ (считается, что на этом масштабе коэффициенты содержат

только шум [6]), а $C_{3/4}$ — 3/4-квантиль стандартного нормального распределения.

В силу приведенных рассуждений в данной работе используется порог

$$T_U = \hat{\sigma} \sqrt{2 \ln n}, \quad n = N^2,$$

который называется «универсальным».

В работах [9, 10] доказаны следующие утверждения, позволяющие статистически оценить погрешность метода пороговой обработки, используя только известные эмпирические вейвлет-коэффициенты.

Теорема 1. Пусть функция, описывающая изображение, является кусочно-регулярной по Липшицу с показателем $\gamma > 0$. Тогда при использовании мягкой пороговой обработки с порогом T_U

$$\frac{\hat{r}_f - r_f}{N^2} \rightarrow 0 \text{ при } N \rightarrow \infty. \quad (4)$$

Теорема 2. Пусть функция, описывающая изображение, является кусочно-регулярной по Липшицу с показателем $\gamma > 1/2$. Тогда при использовании мягкой пороговой обработки с порогом T_U

$$P\left(\frac{\hat{r}_f - r_f}{\sqrt{2}\hat{\sigma}^2 N} < x\right) \Rightarrow \Phi_{\Upsilon}(x) \text{ при } N \rightarrow \infty. \quad (5)$$

Здесь $\Phi_{\Upsilon}(x)$ — функция распределения нормального закона с нулевым средним и дисперсией $\Upsilon^2 = [2C_{3/4}\phi(C_{3/4})]^{-2} - 1$, а $\phi(x)$ — плотность стандартного нормального распределения.

Теорема 1 говорит о том, что \hat{r}_f является состоятельной оценкой, а теорема 2 дает возможность строить асимптотические доверительные интервалы для величины r_f и оценивать отклонение \hat{r}_f от r_f .

Рассмотрим применение этих результатов на примере обработки аэродинамических изображений.

3 Обработка цветных теневых картин аэродинамического эксперимента

3.1 Постановка задачи

В аэродинамической трубе установлено тело (использовались данные, полученные с цилиндра, ось которого была сначала установлена перпендикулярно набегающему потоку и параллельно оптической оси регистрирующей аппаратуры, а затем —

перпендикулярно набегающему потоку и оптической оси регистрирующей аппаратуры).

Исследуется структура обтекающей его газовой струи. При помощи теневого прибора производится снятие характеристик обтекания, при этом на выходе получается так называемая цветная теневая картина обтекания, т. е. изображение, на котором соответствующими цветами фиксируется плотность потока [1, 2, 11, 12].

Дешифровка цветной теневой картины в значительной мере осложняется влиянием зашумления. Оно обуславливается многими факторами, например наличием частиц пыли в набегающем потоке и на оптических элементах фотоаппаратуры, при помощи которой ведется съемка теневой картины, т. е. как собственно возникающими в исследуемой газовой струе (внутренними), так и вносимыми измерительной установкой (внешними). Естественным в этом случае можно считать предположение о том, что их влияние на результат измерения описывается совокупностью независимых одинаково распределенных нормальных случайных величин. Вообще говоря, следует заметить, что цветорегистрирующая аппаратура может зашумлять цвет из разных частей спектра по-разному, в этом случае нужно обладать соответствующими данными о зависимости точности работы установки от входного сигнала. При отсутствии указанных сведений будем считать, что систематической ошибки измерительная аппаратура не создает, а шум, вносимый ею в данные, от самого сигнала не зависит.

В изначальной постановке необходимо создать методику оценки плотности по цветной теневой картине потока в аэродинамической трубе. Алгоритм должен быть устойчив как к внутренним, так и к внешним шумам и с достаточной точностью оценивать численные характеристики обтекания.

В такой постановке задача несколько лет назад была успешно решена при помощи так называемого дискриминантного метода [11–14], он будет кратко описан ниже.

В данной работе исследуется вопрос о возможности реализовать процедуру обработки цветной теневой картины методами вейвлет-анализа. При этом, ввиду отсутствия прямого (без использования результатов работы дискриминантного метода) алгоритма дешифровки цветной теневой картины, исследуется вопрос восстановления цветности изображения, а не собственно плотности потока.

Особое внимание уделено точности обработки области перед телом, установленным в аэродинамической трубе, так как дискриминантный метод оптимизирован для анализа этой области.

3.2 Решение задачи при помощи дискриминантного метода

Входная цветная теневая картина представляет собой графический файл в формате bmp. Размер изображения 512×512 пикселей. Каждый пиксель изображения — элемент цветового пространства (R, G, B) . Отобразим его на единичную сферу, т. е. осуществим следующее преобразование:

$$(r, g, b) \rightarrow \left(\frac{r}{\sqrt{r^2 + g^2 + b^2}}, \frac{g}{\sqrt{r^2 + g^2 + b^2}}, \frac{b}{\sqrt{r^2 + g^2 + b^2}} \right) = (\cos \alpha, \cos \beta, \cos \gamma). \quad (6)$$

Всюду далее будем называть цветовым пространством единичную сферу, на которой расположены векторы $(\cos \alpha, \cos \beta, \cos \gamma)$ — элементы пространства (R, G, B) после преобразования (6).

Все цветовое пространство разобьем на так называемые эталонные классы: произведем разбиение полосы спектра, зафиксированной измерительным прибором, на прямоугольные сегменты, для каждого из них в терминах косинусов (из приведенного выше преобразования) вычислим дисперсию $\{D_{i1}, D_{i2}, D_{i3}\}$ и выборочное среднее $\{z_{i1}, z_{i2}, z_{i3}\}$ (так называемый центр) класса с номером i . Фактически эти характеристики параметризуют зависимость погрешности определения прибором цвета в разных частях спектра. Каждый класс эквивалентности характеризуется также числом X_i — индексом цветности.

Введем дискриминантную функцию

$$d_i(v) = \sqrt{a_{i1}(v_1 - z_{i1})^2 + a_{i2}(v_2 - z_{i2})^2 + a_{i3}(v_3 - z_{i3})^2},$$

где $v = (v_1, v_2, v_3)$ — произвольный вектор цветового пространства;

$$a_{ik} = \frac{\min_{j \neq i}(z_{ik} - z_{jk})}{D_{ik}}, \quad k = 1, 2, 3;$$

D_{ik} — выборочная дисперсия v_k на i -м эталонном классе.

При помощи функций $d_i(v)$ по сути производится вычисление расстояний ρ от центров эталонных классов до точек цветового пространства v . Теоретически точка v находится в i -м эталонном классе, если $d_i(v) = \min_j d_j(v)$.

Для каждого пикселя анализируемого изображения вычислим два наиболее близких класса (для простоты проиндексируем их: 1 — ближайший и 2 —

второй в порядке удаления) и на основе расстояний ρ_1 и ρ_2 до них и индексов цветности соответственно X_1 и X_2 будем вычислять цветность X вектора v по формуле

$$X = X_1 + \frac{\rho_1}{\rho_1 + \rho_2} |X_1 - X_2|.$$

Далее производится восстановление плотности и цветной теневой картины на основе сформированных индексов цветности, сведений об эталонных классах и параметров измерительной установки.

Дискриминантный метод показал хорошую устойчивость как к внешним, так и к внутренним шумам и позволил с необходимой точностью восстановить истинные значения плотности потока и, соответственно, цвета на теневой картине [12–14].

Нужно отметить тот факт, что константы настройки дискриминантного метода установлены таким образом, чтобы с максимальной точностью оценивать плотность потока в отдельном участке спектра, соответствующем области набегающего потока прямо перед столкновением его с телом.

3.3 Решение задачи средствами вейвлет-анализа

Как уже отмечалась, общая концепция любого алгоритма обработки изображений, основанного на вейвлет-преобразовании, заключается в следующем:

- (1) преобразование (декомпозиция) изображения;
- (2) обработка массива вейвлет-коэффициентов;
- (3) реконструкция изображения.

Декомпозиция и реконструкция — взаимосвязанные процессы, так как должны проводиться с использованием наперед заданного вида вейвлета и глубины разложения. Таким образом, сразу появляются два параметра алгоритма: используемый вейвлет и глубина декомпозиции. По результатам тестов с использованием различных типов вейвлетов наилучший результат был получен для так называемого обратного биортогонального семейства. Глубина разложения сказывается на размере деталей, которые все еще можно считать шумовыми, этот параметр подбирался исключительно экспериментально: заранее можно было лишь сказать, что, исходя из информационного смысла вейвлет-коэффициентов, задействовать слишком большую глубину разложения не имело смысла.

Отметим тот факт, что отсутствуют какие-либо априорные сведения о степени зашумления изображения. Поскольку в качестве способа обработки массива вейвлет-коэффициентов заявлена пороговая обработка, то в силу леммы 2 получаем, что достаточно оценить дисперсию шума (порог выбран универсальным).

Цветная теневая картина — изображение, где каждый пиксель кодируется триплетом (R, G, B) . Предлагаемый алгоритм вейвлет-обработки оценивает пороги для каждой цветовой компоненты в отдельности, тем самым отчасти учитывается тот факт, что зашумление может по-разному проявляться в разных частях спектра. Эти принципы вполне соответствуют специфике обрабатываемого сигнала: в ходе специальной проверки были зафиксированы отличия в зашумлении для разных цветовых компонент.

Одной из главных характеристик сигнала является его энергия. В терминах изображения она вычисляется как

$$\|C_0\| = \sum_{i,j=1}^N (r_{i,j}^2 + g_{i,j}^2 + b_{i,j}^2), \quad (7)$$

где $(r_{i,j}, g_{i,j}, b_{i,j})$ — пиксель изображения.

Результат сравнения энергий, заключающихся в аппроксимации и детализации изображения при глубине разложения не более 4, показал, что значительная доля сигнала заключена в аппроксимирующих коэффициентах, поэтому малые изменения, вносимые в них, могут серьезно повлиять на результат обработки. Вследствие этого было принято решение о невнесении каких-либо изменений в коэффициенты аппроксимации. Этот вывод вполне согласуется с физическим смыслом: низкочастотный вейвлет-фильтр выделяет главные особенности сигнала, по сути, создавая сглаженный и уменьшенный его вариант, изменять его — подвергаться значительному риску потери информации о сигнале. Коэффициенты детализации же как раз характеризуют отличия аппроксимации от оригинала, большие по величине вряд ли соответствуют шуму (в терминах постановки задачи), в то время как малые (они и отвечают зашумлению) можно приравнять к 0, что как раз соответствует пороговой обработке сигнала.

Дополнительное исследование вейвлет-коэффициентов показало, что в качестве оценки σ можно взять величину $\hat{\sigma}$, определяемую формулой (3) для каждого набора коэффициентов детализации.

Описанный алгоритм вейвлет-обработки показал высокую точность восстановления цветности теневых аэродинамических картин, что подтверждается проведенным сравнительным анализом

с теневыми картинками, обработанными дискриминантным методом [11–14], для верификации которого использовалась система оптических клиньев.

3.4 Примеры обработки изображений с использованием дискриминантного и вейвлет-методов

В данном разделе в качестве примеров обработки изображений приводятся результаты обработки картин обтекания цилиндра. Так как дискриминантный метод оптимизирован (т.е. демонстрирует наивысшую точность результатов) для обработки зоны цветового спектра соответствующей плотности потока перед исследуемым телом, то особый акцент будет сделан на сравнении получаемых результатов именно в этой области изображения.

3.4.1 Исходные данные и сравнение результатов обработки

Для данной публикации все приводимые ниже рисунки конвертированы в формат «Градации серого». Исходные цифровые цветные изображения в формате RGB, взятые для анализа и в дальнейшем обработанные обоими методами, приведены в приложении [15].

На рис. 1, *a* четко заметно возмущение, возникающее в поле плотности набегающего потока перед телом. Уточнение значений плотности газа в данной области — основная цель создания дискриминантного метода. На рис. 1, *б* данная особенность не просматривается.

Поскольку особое значение имеет обработка области перед телом, приведем также соответствующим образом кадрированное изображение, полученное на основе первого (рис. 2).

Введем обозначения, которые будут использоваться далее:

- r, g, b — значения соответствующих цветовых компонент пикселя из диапазона $0 \dots 255$;
- пара нижних индексов для цветовых компонент — координаты пикселя в изображении от верхнего левого угла;
- изображение имеет размер $N \times N$ пикселей ($N = 512$);
- верхний индекс `discr` свидетельствует о принадлежности значения к результатам обработки дискриминантным методом, `wave` — вейвлет-алгоритмом;
- под элементом изображения понимается некоторая цветовая компонента пикселя.

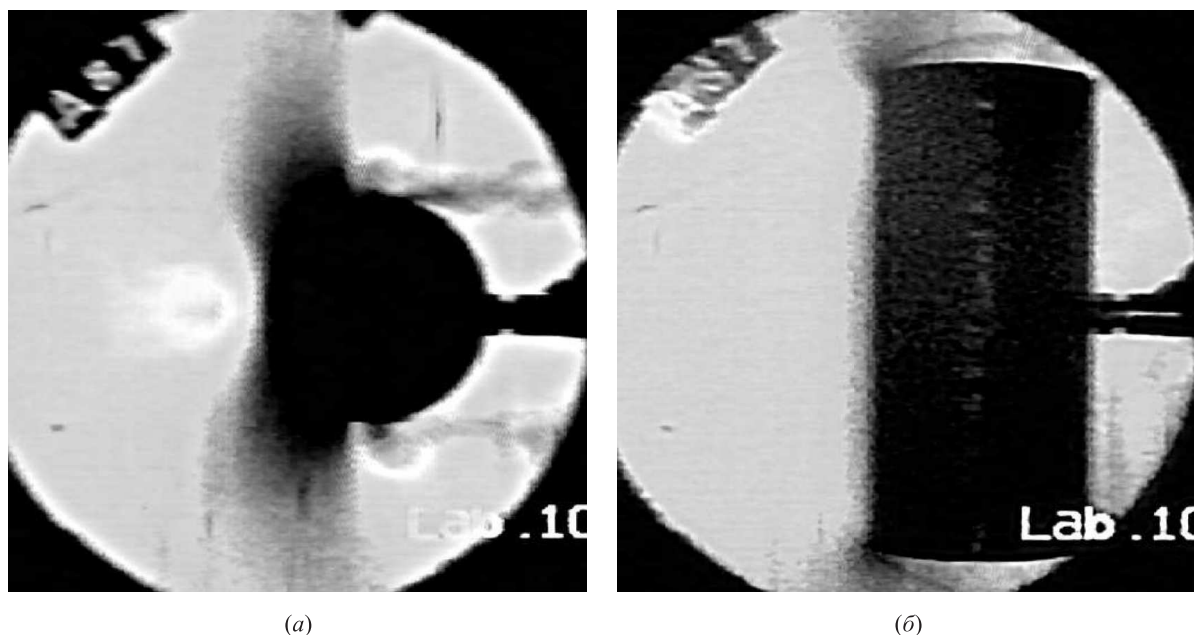


Рис. 1 Изображения 1 (а) и 2 (б) до обработки



Рис. 2 Кадрованное изображение 1 до обработки

За критерий сравнения работы дискриминантного метода и вейвлет-обработки выберем число элементов изображения, различающихся не более чем на 2 единицы, и обозначим его через δ . Тогда

$$\delta = \sum_{i,j=1,\dots,N} (\mathbb{I}\{|r_{i,j}^{\text{discr}} - r_{i,j}^{\text{wave}}| \leq 2\} + \mathbb{I}\{|g_{i,j}^{\text{discr}} - g_{i,j}^{\text{wave}}| \leq 2\} + \mathbb{I}\{|b_{i,j}^{\text{discr}} - b_{i,j}^{\text{wave}}| \leq 2\}).$$

3.4.2 Результаты обработки вейвлет-методом

Результаты обработки исходных теневых картин вейвлет-методом представлены на рис. 3 и 4.

Расчет значения критерия δ показал, что около 80% элементов изображений обрабатываются одинаково обоими рассматриваемыми методами.

Значения нормированных оценок среднеквадратичной погрешности для изображения 1 равны 24,7, 30,1 и 41,6 для цветовых компонент r , g и b соответственно. Эти же значения для изображения 2 равны 23,9, 29,7 и 40,9. На основе теоремы 2 можно сделать вывод, что значения нормированных оценок среднеквадратичной погрешности, которые вычисляются с использованием только известных эмпирических вейвлет-коэффициентов, с вероятностью 0,95 отличаются от истинных среднеквадратичных погрешностей не более чем на 1,1, 1,34 и 1,99 для изображения 1 и не более чем на 1,05, 1,32 и 1,9 для изображения 2.

Таким образом, вейвлет-метод, несмотря на кардинально другой подход к способу оценивания зашумления, дает результаты, схожие с теми, что получаются при обработке изображений дискриминантным методом. Наличие параметризации вейвлет-алгоритма позволяет применять его к изображениям с различной степенью и характером зашумления. Например, в случае увеличения разрешения входного сигнала необходимо будет изменить глубину разложения, чтобы величина деталей, среди которых алгоритм отсеивает

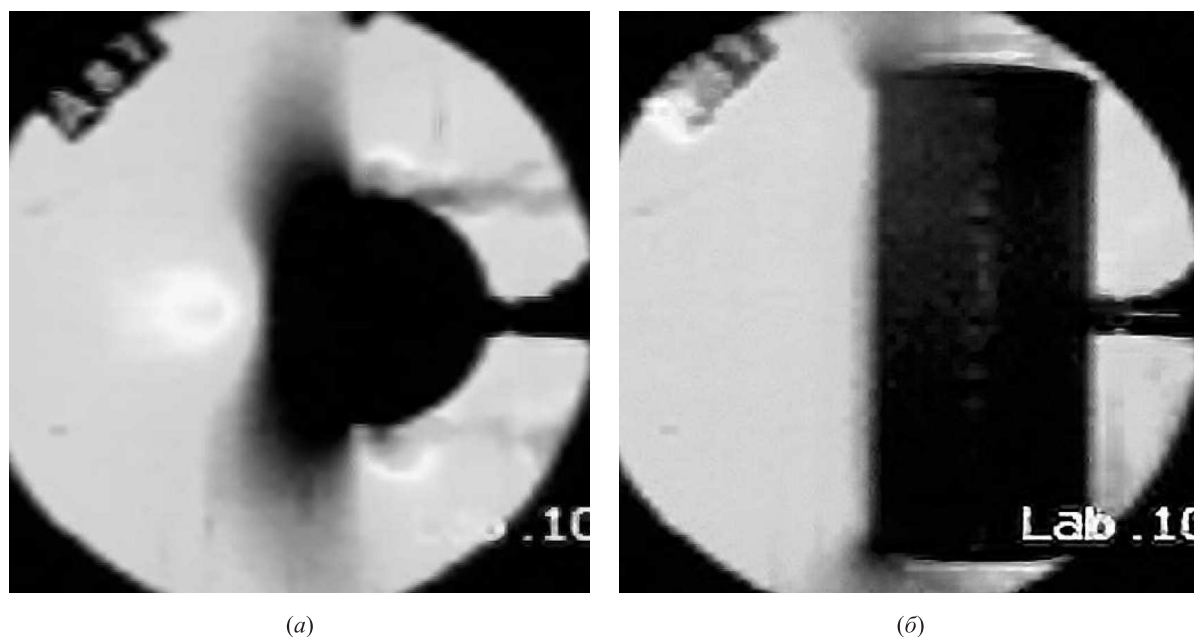


Рис. 3 Изображения 1 (а) и 2 (б) после вейвлет-обработки с глубиной декомпозиции 3 и мягким порогом



Рис. 4 Изображение 1 после вейвлет-обработки с глубиной декомпозиции 3 и мягким порогом и последующего кадрирования

шумовые, соответствующим образом перемасштабировалась. Выбранный в качестве оптимального для тестовой группы изображений вейвлет из обратного биортогонального семейства безусловно не является гарантированно лучшим для обработки

всех теневых картин, он показывал наилучшие результаты в рамках имевшегося набора изображений. Вместе с тем значения параметров гладкости исходного и итогового сигнала вполне соответствуют ожидаемым значениям с точки зрения физического смысла цветной теневой картины и характеристик тестовых изображений: вследствие зашумленности исходный сигнал должен обладать минимальной гладкостью, в то время как генерируемый — значительно большей.

4 Выводы

В ходе решения данной задачи был разработан алгоритм, основанный на средствах, предлагаемых вейвлет-анализом. Для верификации результатов было проведено сравнение с дискриминантным методом обработки аэродинамических картин.

В ходе тестирования была показана возможность настройки вейвлет-алгоритма для обработки имевшейся группы цветных теневых картин с достаточной точностью относительно результатов обработки дискриминантным методом.

Был проведен расчет нормированной оценки среднеквадратичной погрешности по экспериментальным данным для каждого цветового канала. Полученные эмпирические значения погрешности хорошо согласуются с теоретическими значениями, рассчитанными на основе свойств ее предельного распределения.

Литература

1. Холдер Д., Норт Р. Теневые методы в аэродинамике / Пер с англ. — М.: Мир, 1966. 180 с. (Holder D. W., North R. J. Schlieren methods. — National Physics Laboratory. Notes on applied science No. 31. — London: H.M.S.O., 1963. 106 p.)
2. Краснов Н.Ф. Аэродинамика. Т. 2: Методы аэродинамического расчета. — 4-е изд. — М.: Высшая школа, 2010. 416 с.
3. Смоленцев Н. К. Основы теории вейвлетов. Вейвлеты в Matlab. — М.: ДМК Пресс, 2005. 157 с.
4. Захарова Т. В., Шестаков О. В. Вейвлет-анализ и его приложения. — 2-е изд., перераб. и доп. — М.: ИНФРА-М, 2012. 157 с.
5. Bogges A., Narkovich F. A. A first course in wavelets with Fourier analysis. — Prentice Hall, 2001. 293 p.
6. Малла С. Вейвлеты в обработке сигналов / Пер. с англ. — М.: Мир, 2005. 671 с. (Mallat S. A wavelet tour of signal processing. — 2nd ed. — Elsevier, 1999. 661 p.)
7. Donoho D., Johnstone I. Ideal spatial adaptation via wavelet shrinkage // *Biometrika*, 1994. Vol. 81. P. 425–455.
8. Donoho D., Johnstone I. Adapting to unknown smoothness via wavelet shrinkage // *J. Am. Statist. Assoc.*, 1995. Vol. 90. P. 1200–1224.
9. Маркин А. В. Предельное распределение оценки риска при пороговой обработке вейвлет-коэффициентов // *Информатика и её применения*, 2009. Т. 3. Вып. 4. С. 57–63.
10. Маркин А. В., Шестаков О. В. О состоятельности оценки риска при пороговой обработке вейвлет-коэффициентов // *Вестн. Моск. ун-та. Сер. 15. Вычисл. матем. и киберн.*, 2010. Вып. 1. С. 26–34.
11. Zakharova T. V., Berezentsev M. V. About a method of supervision classification in the decision of aerodynamic problems // 24th Seminar (International) on Stability Problems for Stochastic Models Transactions. — Riga: ТТИ, 2004. P. 353–356.
12. Захарова Т. В. Метод распознавания для восстановления изображений цветных теневых картин // *Обозрение прикладной и промышленной математики*, 2005. Т. 12. Вып. 4. С. 967–968.
13. Захарова Т. В., Шагуров Э. А. Определение плотности аэродинамического потока обтекания методом цветовой фильтрации // *Математическое моделирование*, 2013. Т. 25. Вып. 12. С. 103–109.
14. Захарова Т. В., Шагуров Э. А. Оптимизация метода цветовой фильтрации для решения задач аэродинамики // *Обозрение прикладной и промышленной математики*, 2013. Т. 20. Вып. 4. С. 545–548.
15. Захарова Т. В., Шестаков О. В. Приложение к статье «Анализ точности вейвлет-обработки аэродинамических картин обтекания». <http://www.ipiran.ru/publications/pictures.pdf>.

Поступила в редакцию 07.05.16

PRECISION ANALYSIS OF WAVELET PROCESSING OF AERODYNAMIC FLOW PATTERNS

T. V. Zakharova^{1,2} and O. V. Shestakov^{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: This paper is devoted to a new method of aerodynamic flow pattern processing based on the wavelet analysis. Wavelet thresholding techniques are widely used in signal and image processing. These methods are easily implemented through fast algorithms; so, they are very appealing in practical situations. Besides, they adapt to function classes with different amounts of smoothness in different locations more effectively than the usual linear methods. Wavelet thresholding risk analysis is an important practical task, because it allows determining the quality of the techniques themselves and the equipment which is being used. Comparative analysis using the discriminant method was carried out to verify the new method. The empirical estimated error of processing is consistent with theoretical results for this estimate.

Keywords: wavelet analysis; thresholding; unbiased risk estimate; aerodynamic flow

DOI: 10.14357/19922264160307

References

1. Holder, D. W., and R. J. North. 1963. *Schlieren methods*. National Physics Laboratory. Notes on applied science No. 31. London: H.M.S.O. 106 p.
2. Krasnov, N. F. 2010. *Aerodinamika. T. 2: Metody aerodinamicheskogo rascheta* [Aerodynamics. Vol. 2: The aerodynamic calculation methods]. 4th ed. Moscow: Vysshaya Shkola. 416 p.
3. Smolentsev, N. K. 2008. *Osnovy teorii veyvletov. Veyvlety v Matlab* [Foundations of the theory of wavelets. Wavelets in Matlab]. Moscow: DMK Press. 448 p.
4. Zakharova, T. V., and O. V. Shestakov. 2012. *Veyvlet-analiz i ego prilozheniya* [Wavelet analysis and its applications]. 2nd ed. Moscow: INFRA-M. 157 p.
5. Boggess, A., and F. A. Narkovich. 2001. *A first course in wavelets with Fourier analysis*. Prentice Hall. 293 p.
6. Mallat, S. 1999. *Wavelet tour of signal processing*. Elsevier. 661 p.
7. Donoho, D., and I. Johnstone. 1994. Ideal spatial adaptation via wavelet shrinkage. *Biometrika* 81:425–455.
8. Donoho, D., and I. Johnstone. 1995. Adapting to unknown smoothness via wavelet shrinkage. *J. Am. Statist. Assoc.* 90:1200–1224.
9. Markin, A. V. 2009. Predel'noe raspredelenie otsenki riska pri porogovoy obrabotke veyvlet-koeffitsientov [Limit distribution of risk estimate of wavelet coefficient thresholding]. *Informatika i ee Primeneniya — Inform. Appl.* 3(4):57–63.
10. Markin, A. V., and O. V. Shestakov. 2010. O sostoyatel'nosti otsenki riska pri porogovoy obrabotke veyvlet-koeffitsientov [Consistency of risk estimation with thresholding of wavelet coefficients]. *Vestn. Mosk. un-ta. Ser. 15. Vychisl. matem. i kibern.* [Moscow University Computational Mathematics and Cybernetics] 1:26–34.
11. Zakharova, T. V., and M. V. Berezentsev. 2004. About a method of supervision classification in the decision of aerodynamic problems. *24th Seminar (International) on Stability Problems for Stochastic Models Transactions*. Riga: TTI. 353–356.
12. Zakharova, T. V. 2005. Metod raspoznavaniya dlya vosstanovleniya izobrazheniy tsvetnykh tenevykh kartin [Recognition method for recovery of nonferrous shadow paintings images]. *Obozrenie Prikladnoy i Promyshlennoy Matematiki* [Review of Applied and Industrial Mathematics] 12(4):967–968.
13. Zakharova, T. V., and E. A. Shagirov. 2013. Opredelenie plotnosti aerodinamicheskogo potoka obtekaniya metodom tsvetovoy fil'tratsii [Density determination of aerodynamic flow color flow filtration method]. *Matematicheskoe modelirovanie* [Mathematical Modeling] 25(12):103–109.
14. Zakharova, T. V., and E. A. Shagirov. 2013. Optimizatsiya metoda tsvetovoy fil'tratsii dlya resheniya zadach aerodinamiki [Optimization of the color filter method for solving the problems of aerodynamics]. *Obozrenie Prikladnoy i Promyshlennoy Matematiki* [Review of Applied and Industrial Mathematics] 20(4):545–548.
15. Zakharova, T. V., and O. V. Shestakov. Appendix to the paper “Precision analysis of wavelet processing of aerodynamic flow patterns.” Available at: <http://www.ipiran.ru/publications/pictures.pdf> (accessed August 29, 2016).

Received May 07, 2016

Contributors

Zakharova Tatiana V. (b. 1962) — Candidate of Science (PhD) in physics and mathematics; senior lecturer, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; lsa@cs.msu.ru

Shestakov Oleg V. (b. 1976) — Doctor of Science in physics and mathematics, associate professor, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; oshestakov@cs.msu.ru

АНАЛИТИЧЕСКОЕ МОДЕЛИРОВАНИЕ НОРМАЛЬНЫХ ПРОЦЕССОВ В СТОХАСТИЧЕСКИХ СИСТЕМАХ СО СЛОЖНЫМИ БЕССЕЛЕВЫМИ НЕЛИНЕЙНОСТЯМИ ДРОБНОГО ПОРЯДКА*

И. Н. Сеницын¹

Аннотация: Рассматриваются методы аналитического моделирования (МAM) нормальных (гауссовских) процессов в гауссовских и негауссовских стохастических системах (СтС) со сложными бесселевыми нелинейностями (БН) дробного порядка (БНДП) (сферическими, модифицированными сферическими и описываемыми функциями Эйри). Приведены необходимые сведения из теории бесселевых функций дробного порядка. Даны формулы для коэффициентов статистической линеаризации БНДП. Особое внимание уделено алгоритмам вычисления этих коэффициентов, основанным на степенных разложениях. Приведены алгоритмы МAM нормальных процессов в СтС с БНДП на основе метода нормальной аппроксимации (МНА) и метода статистической линеаризации (МСЛ). Разработаны методы вычисления типовых интегралов для БНДП. Рассмотрен ряд тестовых примеров. Сделаны основные выводы и некоторые обобщения.

Ключевые слова: бесселева нелинейность дробного порядка (БНДП); метод аналитического моделирования (МAM); метод нормальной аппроксимации (МНА); метод статистической линеаризации (МСЛ); модифицированные сферические бесселевы нелинейности; нелинейности Эйри; нормальный (гауссовский) стохастический процесс; сферические бесселевы нелинейности

DOI: 10.14357/19922264160308

1 Введение

В [1] рассматриваются МAM процессов в динамических системах со сложными цилиндрическими БН целого порядка при гармонических и стохастических, узкополосных и широкополосных возмущениях. Даны необходимые сведения из теории бесселевых функций и сложных БН (СБН). Приведено методическое и алгоритмическое обеспечение МAM на основе МСЛ и МНА для стохастических широкополосных процессов типа белого шума. Рассмотрены особенности МAM для гармонических и узкополосных стохастических процессов. В приложении приведены формулы для коэффициентов МСЛ для типовых бесселевых нелинейностей целого порядка. В качестве тестовых примеров рассмотрены задачи МAM процессов в одномерных системах с аддитивными и мультипликативными белыми шумами. Особое внимание уделено процессам в бесселевом осцилляторе в условиях различных возмущений. Заключение содержит основные выводы и некоторые обобщения.

Рассмотрены особенности МAM нормальных процессов на основе МНА (МСЛ) в гауссовских

и негауссовских СтС со сложными БНДП. В разд. 2, 3 и приложениях П1 и П2 даны необходимые сведения из теории действительных бесселевых функций и сложных БНДП. Разделы 2–7 посвящены степенным алгоритмам вычисления интегралов МНА (МСЛ). В разд. 8 приведены алгоритмы МAM нормальных процессов в СтС с БНДП, а в приложении П3 — тестовые примеры. Заключение содержит основные выводы и некоторые обобщения.

2 Функции Бесселя дробного порядка

Как известно [2–4], наиболее распространенными функциями Бесселя (ФБ) дробного порядка являются сферические и модифицированные сферические ФБ, а также функции Эйри.

Сферические ФБ являются частными решениями обыкновенного линейного дифференциального уравнения второго порядка следующего вида:

$$z^2 \frac{d^2 w}{dz^2} + 2z \frac{dw}{dz} + [z^2 - \nu(\nu + 1)] w = 0. \quad (1)$$

*Работа выполнена при поддержке ОНИТ РАН (проект 0063-2015-0017 П.3).

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

Сферические ФБ первого и второго рода являются действительными функциями и определяются следующими формулами:

$$j_n(z) = \sqrt{\frac{\pi}{2z}} J_{n+1/2}(z);$$

$$y_n(z) = \sqrt{\frac{\pi}{2z}} Y_{n+1/2}(z).$$

Здесь введены следующие обозначения (см. приложение П1): $J_\nu(z)$ — ФБ первого рода; $Y_\nu(z)$ — ФБ второго рода (называемые также функциями Вебера или Неймана, причем $Y_\nu(z) \equiv N_\nu(z)$).

Модифицированные сферические ФБ первого и второго рода удовлетворяют уравнению (1) при замене z на iz :

$$z^2 \frac{d^2 w}{dz^2} + 2z \frac{dw}{dz} - [z^2 + \nu(\nu + 1)] w = 0$$

и при $\nu = \pm(n \pm 1/2)$ определяются следующими формулами:

$$i_n(z) = \sqrt{\frac{\pi}{2z}} I_{n+1/2}(z);$$

$$\tilde{i}_n(z) = \sqrt{\frac{\pi}{2z}} I_{-n-1/2}(z).$$

Здесь $I_n(z)$ — модифицированная цилиндрическая ФБ (см. приложение П1).

Функции Эйри $Ai(z)$ и $Bi(z)$ удовлетворяют дифференциальному уравнению вида

$$\frac{d^2 w}{dz^2} - zw = 0$$

и определяются следующими формулами:

$$\left. \begin{aligned} Ai(z) &= c_1 f(z) - c_2 g(z); \\ Bi(z) &= \sqrt{3} [c_1 f(z) + c_2 g(z)]. \end{aligned} \right\} \quad (2)$$

Здесь

$$f(z) = \sum_{l=0}^{\infty} 3^l \left(\frac{1}{3}\right)_l \frac{z^{3l}}{(3l)!};$$

$$g(z) = \sum_{l=0}^{\infty} 3^l \left(\frac{2}{3}\right)_l \frac{z^{3l+1}}{(3l+1)!},$$

где

$$c_1 = Ai(0) = \frac{Bi(0)}{3} = \frac{3^{-2/3}}{\Gamma(2/3)} \approx 0,35502;$$

$$c_2 = -Ai'(0) = \frac{Bi'(0)}{\sqrt{3}} = \frac{3^{-1/3}}{\Gamma(1/3)} \approx 0,25881;$$

$$\left(a + \frac{1}{3}\right)_0 = 1;$$

$$3^l \left(a + \frac{1}{3}\right)_l = (3a + 1)(3a + 4) \cdots (3a + 3l - 2).$$

В приложении П2 приведены необходимые свойства и степенные представления ФБ дробно-порядка.

3 Сложные бесселевы нелинейности

Рассмотрим скалярное безынерционное нелинейное детерминированное преобразование вида

$$Z = \varphi(Y). \quad (3)$$

Здесь под φ понимается ФБ. Нелинейности, описываемые (3), следуя [1], будем называть бесселевыми нелинейностями.

Примерами СБН [1], получаемых посредством сумм типовых БН, могут служить следующие:

$$\varphi^{\text{СБН}}(Y, t) = \sum_{r=1}^n l_{rt} \varphi_r^{\text{БН}}(Y);$$

$$\varphi^{\text{СБН}}(Y, t) = \sum_{r=1}^n l_{rt} \varphi_r^{\text{БН}}(Y) \varphi_r^{\text{АН}}(Y),$$

а также дробно-рациональные представления:

$$\varphi^{\text{СБН}}(Y, t) = \frac{\sum_{r=1}^{n'} l'_{rt} \varphi_r^{\text{БН}}(Y)}{\sum_{r=1}^{n''} l''_{rt} \varphi_r^{\text{БН}}(Y)};$$

$$\varphi^{\text{СБН}}(Y, t) = \frac{\sum_{r=1}^{n'} l'_{rt} \varphi_r^{\text{БН}}(Y) \varphi_r^{\text{АН}}(Y)}{\sum_{r=1}^{n''} l''_{rt} \varphi_r^{\text{БН}}(Y) \varphi_r^{\text{АН}}(Y)},$$

где $\varphi_r^{\text{БН}}(Y)$, $\varphi_r^{\text{АН}}(Y)$ и $\varphi_r^{\text{АН}}(Y)$ — типовые бесселевы функции; l_{rt} , l'_{rt} и l''_{rt} — коэффициенты, зависящие от времени t ; $\varphi_r^{\text{АН}}(Y)$, $\varphi_r^{\text{АН}}(Y)$ и $\varphi_r^{\text{АН}}(Y)$ — алгебраические нелинейности (многочлены, степенные, иррациональные, дробно-рациональные и другие функции).

Другими примерами СБН являются нелинейности, получаемые путем соответствующего преобразования аргумента:

$$\begin{aligned} \varphi^{\text{СБН}}(Y, t) &= \varphi^{\text{АН}}(\psi^{\text{БН}}(Y, t), t); \\ \varphi^{\text{СБН}}(Y, t) &= \varphi^{\text{БН}}(\psi^{\text{АН}}(Y, t), t), \end{aligned}$$

где $\varphi^{\text{АН}}$, $\psi^{\text{АН}}$, $\varphi^{\text{БН}}(Y, t)$ и $\psi^{\text{БН}}(Y, t)$ — типовые алгебраические и бесселевы трансцендентные нелинейности.

В качестве примеров скалярных СБН векторного аргумента $Y = [Y_1, \dots, Y_p]^T$ рассмотрим следующие:

$$\begin{aligned} \varphi^{\text{СБН}}(Y, t) &= \sum_{r=1}^n \prod_{h=1}^H l_{rh,t} \varphi_{rh}^{\text{БН}}(Y_h); \\ \varphi^{\text{СБН}}(Y, t) &= \frac{\sum_{r=1}^{n'} \prod_{h=1}^{H'} l'_{rh,t} \varphi'_{rh}{}^{\text{БН}}(Y_h)}{\sum_{r=1}^{n''} \prod_{h=1}^{H''} l''_{rh,t} \varphi''_{rh}{}^{\text{БН}}(Y_h)}. \end{aligned}$$

В случае векторных и матричных СБН приведенные формулы имеют место для соответствующих компонент.

4 Статистическая линейаризация бесселевых нелинейностей дробного порядка

Рассмотрим статистическую линейаризацию по Казакову [5–7] бесселевой нелинейности (3) (индекс «СБН» для краткости опускается) при несимметричном ($m_y \neq 0$) гауссовском стохастическом входном сигнале Y_t :

$$Y_t = Y(t) = m_y + Y_t^0,$$

где m_y — математическое ожидание, а D_y — дисперсия, $Y_t^0 = Y^0(t) = Y(t) - m_y$. В соответствии с МСЛ зависимость аппроксимируется следующим выражением:

$$Z_t = \varphi_0(m_y, D_y) + k_1(m_y, D_y) Y_t^0.$$

Здесь φ_0 и k_1 — коэффициенты статистической линейаризации, зависящие от m_y и D_y и определяемые по формулам:

$$\begin{aligned} \varphi_0 &= \varphi_0(m_y, D_y) = \\ &= \frac{1}{\sqrt{2\pi D_y}} \int_{-\infty}^{\infty} \varphi(\eta) e^{-(\eta - m_y)^2 / (2D_y)} d\eta; \end{aligned}$$

$$\begin{aligned} k_1 &= k_1(m_y, D_y) = \\ &= \frac{1}{\sqrt{2\pi D_y}} \int_{-\infty}^{\infty} (\eta - m_y) \varphi(\eta) e^{-(\eta - m_y)^2 / (2D_y)} d\eta = \\ &= \frac{\partial \varphi_0(m_y, D_y)}{\partial m_y}. \end{aligned}$$

Для нечетных функций φ имеем:

$$\begin{aligned} \varphi_0(m_y, D_y) &= k_0(m_y, D_y) m_y; \\ k_0(m_y, D_y) &= \frac{1}{m_y \sqrt{2\pi D_y}} \int_{-\infty}^{\infty} \varphi(\eta) e^{-(\eta - m_y)^2 / (2D_y)} d\eta. \end{aligned}$$

В зависимости от аналитической формы представления СБН различают следующие методы вычисления коэффициентов статистической линейаризации [4]:

- выражения через элементарные функции;
- представления по формулам Рэлея;
- степенные разложения;
- многочленные приближения;
- дробно-рациональные приближения;
- разложения в цепные дроби;
- асимптотические формулы и приближения;
- рекуррентные соотношения.

В статье ограничимся разработкой алгоритмов расчетов, основанных на степенных разложениях.

5 Коэффициенты статистической линейаризации сферических бесселевых нелинейностей $j_n(z)$ и модифицированных сферических нелинейностей $i_n(z)$ первого рода

Используя степенные представления функций $i_n(z)$ и $j_n(z)$ [2–4]

$$\begin{aligned} j_n(z) &= \sum_{k=0}^{\infty} \frac{(-1)^k z^{2k+n}}{(2k)!! (2k+2n+1)!!}; \\ i_n(z) &= \sum_{k=0}^{\infty} \frac{z^{2k+n}}{(2k)!! (2k+2n+1)!!} \end{aligned}$$

и табличные интегралы

$$\int_0^\infty t^{2m+1} e^{-ct^2/2} dt = \frac{m!}{2^m c^{m+1}};$$

$$\int_0^\infty t^{2m} e^{-ct^2/2} dt = \frac{(2m-1)!!}{2c^m} \sqrt{\frac{2\pi}{c}},$$

придем к следующим степенным алгоритмам вычисления коэффициентов статистической линеаризации $\varphi_0^{j_n}$, $k_1^{j_n}$ и $\varphi_0^{i_n}$, $k_1^{i_n}$:

$$\varphi_0^{j_n}(m_y, D_y) = \sum_{k=0}^\infty \frac{(-1)^k}{(2k)!!(2k+2n+1)!!} \varphi_0^{z^{2k+n}}(m_y, D_y); \quad (4)$$

$$\varphi_0^{i_n}(m_y, D_y) = \sum_{k=0}^\infty \frac{1}{(2k)!!(2k+2n+1)!!} \varphi_0^{z^{2k+n}}(m_y, D_y); \quad (5)$$

$$\left. \begin{aligned} k_1^{j_n}(m_y, D_y) &= \varphi_0^{z^{j_n}}(m_y, D_y) - \\ &\quad - m_y \varphi_0^{j_n}(m_y, D_y) \quad (m_y \neq 0); \\ k_1^{j_n}(0, D_y) &= \varphi_0^{z^{j_n}}(0, D_y) \quad (m_y = 0); \end{aligned} \right\} (6)$$

$$\left. \begin{aligned} k_1^{i_n}(m_y, D_y) &= \varphi_0^{z^{i_n}}(m_y, D_y) - \\ &\quad - m_y \varphi_0^{i_n}(m_y, D_y) \quad (m_y \neq 0); \\ k_1^{i_n}(0, D_y) &= \varphi_0^{z^{i_n}}(0, D_y) \quad (m_y = 0). \end{aligned} \right\} (7)$$

Здесь введены обозначения:

$$\varphi_0^{z^h}(m_y, D_y) = \frac{\exp(-m_y^2/(2D_y))}{\sqrt{2\pi D_y}} \times \sum_{l=0}^\infty \frac{[1 + (-1)^{h+l}]}{l!} A^{h+l}(D_y) \left(\frac{m_y}{D_y}\right)^l$$

$$\left. \begin{aligned} &\quad (h = 1, 2, \dots); \\ A^{2h}(D_y) &= \frac{(2h-1)!!}{2} D_y^h \sqrt{2\pi D_y}; \\ A^{2h+1}(D_y) &= \frac{h!}{2^h} D_y^{h+1} \quad (h = 1, 2, \dots). \end{aligned} \right\} (8)$$

Таким образом, при $m_y > 0$, $D_y > 0$ в основе степенного алгоритма вычисления коэффициентов статистической линеаризации сферической БН $j_n(z)$ лежат формулы (4) и (6) при условиях (8) и (9) (теорема 1).

В основе теоремы 2 для модифицированных сферических нелинейностей $i_n(z)$ лежат формулы (5) и (7) при условиях (8) и (9) и $m_y > 0$, $D_y > 0$.

6 Коэффициенты статистической линеаризации сферических бесселевых нелинейностей $y_n(z)$ и модифицированных сферических нелинейностей $\tilde{i}_n(z)$ второго рода

Используя степенные представления функций $y_n(z)$ [2-4]:

$$y_n(z) = - \sum_{k=0}^n \frac{(2n-2k-1)!!}{(2k)!!} z^{2k+n} - \sum_{k=n}^\infty \frac{(-1)^k}{((2k)!!(2k-2n+1)!!)} z^{2k+n},$$

получим следующий алгоритм вычисления коэффициентов статистической линеаризации $\varphi_0^{y_n}(m_y, D_y)$ и $k^{y_n}(m_y, D_y)$:

$$\varphi_0^{y_n}(m_y, D_y) = - \sum_{k=0}^\infty \frac{(2n-2k-1)!!}{(2k)!!} \varphi_0^{z^{2k+n}}(m_y, D_y) - \sum_{k=n}^\infty \frac{(-1)^k}{(2k)!!(2k-2n+1)!!} \varphi_0^{z^{2k+n}}(m_y, D_y); \quad (10)$$

$$\left. \begin{aligned} k_1^{y_n}(m_y, D_y) &= \varphi_0^{z^{y_n}}(m_y, D_y) - \\ &\quad - m_y \varphi_0^{y_n}(m_y, D_y) = \frac{\partial \varphi_0^{y_n}}{\partial m_y}(m_y, D_y) \quad (m_y \neq 0); \\ k_1^{y_n}(0, D_y) &= \varphi_0^{z^{y_n}}(0, D_y) \quad (m_y = 0). \end{aligned} \right\} (11)$$

Здесь $\varphi_0^{z^h}(m_y, D_y)$ определены (8) и (9).

Аналогично имеем для $\tilde{i}_n(z)$:

$$\tilde{i}_n(z) = \sum_{k=0}^{n-1} \frac{(-1)^{n+k}(2n-2k-1)!!}{(2k)!!} z^{2k-n-1} + \sum_{k=n}^\infty \frac{(2k+1-2n)!!}{(2k)!!} z^{2k-n-1};$$

$$\varphi_0^{\tilde{i}_n}(m_y, D_y) = \sum_{k=0}^{n-1} \frac{(-1)^{n+k}(2n-2k-1)!!}{(2k)!!} \varphi_0^{z^{2k-n-1}}(m_y, D_y) + \sum_{k=n}^\infty \frac{(2k+1-2n)!!}{(2k)!!} \varphi_0^{z^{2k-n-1}}(m_y, D_y); \quad (12)$$

$$\left. \begin{aligned} k_1^{\tilde{i}n} (m_y, D_y) &= \varphi_0^{\tilde{z}i_n} (m_y, D_y) - \\ &- m_y \varphi_0^{\tilde{i}n} (m_y, D_y) \quad (m_y \neq 0); \\ k_1^{\tilde{i}n} (0, D_y) &= \varphi_0^{\tilde{z}i_n} (0, D_y) \quad (m_y = 0). \end{aligned} \right\} \quad (13)$$

Здесь $\varphi_0^{\tilde{z}i_n} (m_y, D_y)$ определены (8) и (9).

В основе **теоремы 3** для алгоритма расчета коэффициентов статистической линеаризации сферической БН $y_n(z)$ лежат формулы (10) и (11) при условиях (8) и (9) и $m_y > 0, D_y > 0$.

Алгоритм расчета коэффициентов статистической линеаризации модифицированной сферической БН $\tilde{i}_n(z)$ определяется формулами (12) и (13) при условиях (8) и (9) и $m_y > 0, D_y > 0$ (**теорема 4**).

7 Коэффициенты статистической линеаризации нелинейностей Эйри

Применяя степенные разложения функций $f(z)$ и $g(z)$ в (2), получим следующие алгоритмы вычисления коэффициентов статистической линеаризации:

$$\varphi_0^{Ai} (m_y, D_y) = c_1 \varphi_0^f (m_y, D_y) - c_2 \varphi_0^g (m_y, D_y); \quad (14)$$

$$\varphi_0^{Bi} (m_y, D_y) = \sqrt{3} \left[c_1 \varphi_0^f (m_y, D_y) + c_2 \varphi_0^g (m_y, D_y) \right]; \quad (15)$$

$$\left. \begin{aligned} k_1^{Ai} (m_y, D_y) &= \varphi_0^{zAi} (m_y, D_y) - \\ &- m_y \varphi_0^{Ai} (m_y, D_y) = \frac{\partial \varphi_0^{Ai}}{\partial m_y} (m_y, D_y) \quad (m_y \neq 0); \\ k_1^{Ai} (0, D_y) &= k_1^{zAi} (0, D_y); \end{aligned} \right\} \quad (16)$$

$$\left. \begin{aligned} k_1^{Bi} (m_y, D_y) &= \varphi_0^{zBi} (m_y, D_y) - \\ &- m_y \varphi_0^{Bi} (m_y, D_y) = \frac{\partial \varphi_0^{Bi}}{\partial m_y} (m_y, D_y) \quad (m_y \neq 0); \\ k_1^{Bi} (0, D_y) &= k_1^{zBi} (0, D_y). \end{aligned} \right\} \quad (17)$$

Здесь

$$\left. \begin{aligned} \varphi_0^f (m_y, D_y) &= \sum_{l=0}^{\infty} 3^l \left(\frac{1}{3} \right)_l \frac{\varphi_0^{3l} (m_y, D_y)}{(3l)!}; \\ \varphi_0^g (m_y, D_y) &= \sum_{l=0}^{\infty} 3^l \left(\frac{2}{3} \right)_l \frac{\varphi_0^{3l+1} (m_y, D_y)}{(3l+1)!}. \end{aligned} \right\} \quad (18)$$

Таким образом, в основе алгоритма расчета коэффициентов статистической линеаризации функции $Ai(z)$ лежат формулы (14), (16) и (18) при условиях (8) и (9) и $m_y > 0, D_y > 0$ (**теорема 5**).

Алгоритм расчета функции $Bi(z)$ определяется формулами (15) и (17) при условиях (8) и (9) и $m_y > 0, D_y > 0$ (**теорема 6**).

8 Аналитическое моделирование нормальных процессов в стохастических системах со сложными бесселевыми нелинейностями дробного порядка

Как известно [1, 5, 6], уравнения конечномерных непрерывных нелинейных систем со стохастическими возмущениями путем расширения вектора состояния СтС могут быть записаны в виде следующего векторного стохастического дифференциального уравнения Ито:

$$dY_t = a(Y_t, t) dt + b(Y_t, t) dW_0 + \int_{R_0} c(Y_t, t, v) P^0(dt, dv), \quad Y(t_0) = Y_0. \quad (19)$$

Здесь Y_t — $(p \times 1)$ -мерный вектор состояния, $Y_t \in \Delta_y$ (Δ_y — многообразие состояний); $a = a(Y_t, t)$ и $b = b(Y_t, t)$ — известные $(p \times 1)$ -мерная и $(p \times m)$ -мерная функции Y_t и t ; $W_0 = W_0(t) - (r \times 1)$ -мерный винеровский стохастический процесс (СтП) интенсивности $\nu_0 = \nu_0(t)$; $c(Y_t, t, v)$ — $(p \times 1)$ -мерная функция Y_t, t и вспомогательного $(q \times 1)$ -мерного параметра v ; $\int_{\Delta} dP^0(t, A)$ — центрированная пуассоновская мера, определяемая

$$\int_{\Delta} dP^0(t, A) = \int_{\Delta} dP(t, A) = \int_{\Delta} \nu_P(t, A) dt.$$

При этом принято: \int_{Δ} — число скачков пуассоновского СтП в интервале времени $\Delta = (t_1, t_2]$; $\nu_P(t, A)$ — интенсивность пуассоновского СтП $P(t, A)$; A — некоторое борелевское множество пространства R_0^q с выколотым началом. Начальное значение Y_0 представляет собой случайную величину,

не зависящую от приращений $W_0(t)$ и $P(t, A)$ на интервалах времени, следующих за $t_0, t_0 \leq t_1 \leq t_2$, для любого множества A . Элементы векторно-матричных функций $a(Y_t, t)$, $b(Y_t, t)$ и $c(Y_t, t, v)$ являются СБН дробного порядка.

Для аддитивных гауссовских (нормальных) и обобщенных пуассоновских возмущений уравнение (19) принимает вид [1, 5, 6]:

$$\dot{Y} = a(Y_t, t) + b_0(t)V; \quad V = \dot{W}; \quad Y(t_0) = Y_0. \quad (20)$$

Здесь W — СтП с независимыми приращениями, представляющий собой смесь нормального и обобщенного пуассоновского СтП.

Если предположить существование конечных вероятностных моментов второго порядка для моментов времени t_1 и t_2 , то уравнения МНА примут следующий вид [1, 5, 6]:

– для характеристических функций:

$$\left. \begin{aligned} g_1^N(\lambda; t) &= \exp \left[i\lambda^T m_t - \frac{1}{2} \lambda^T K_t \lambda \right]; \\ g_{t_1, t_2}^N(\lambda_1, \lambda_2; t_1, t_2) &= \\ &= \exp \left[i\bar{\lambda}^T \bar{m}_2 - \frac{1}{2} \bar{\lambda}^T \bar{K}_2 \bar{\lambda} \right], \end{aligned} \right\} \quad (21)$$

где

$$\bar{\lambda} = [\lambda_1^T \lambda_2^T]^T; \quad \bar{m}_2 = [m_{t_1}^T m_{t_2}^T]^T; \\ \bar{K}_2 = \begin{bmatrix} K(t_1, t_1) & K(t_1, t_2) \\ K(t_2, t_1) & K(t_2, t_2) \end{bmatrix};$$

– для математических ожиданий m_t , ковариационной матрицы K_t и матрицы ковариационных функций $K(t_1, t_2)$:

$$\left. \begin{aligned} \dot{m}_t &= a_1(m_t, K_t, t), \quad m_0 = m(t_0); \\ \dot{K}_t &= a_2(m_t, K_t, t), \quad K_0 = K(t_0); \\ \frac{\partial K(t_1, t_2)}{\partial t_2} &= K(t_1, t_2) a_{21}(m_{t_2}, K_{t_2}, t_2)^T, \\ &K(t_1, t_1) = K_{t_1}. \end{aligned} \right\} \quad (22)$$

Здесь приняты следующие обозначения:

$$\begin{aligned} m_t &= M_{\Delta_y}^N [Y_t]; \quad Y_t^0 = Y_t - m_t; \\ K_t &= M_{\Delta_y}^N [Y_t^0 Y_t^{0T}]; \quad K(t_1, t_2) = M_{\Delta_y}^N [Y_{t_1}^0 Y_{t_2}^{0T}]; \\ a_1 &= a_1(m_t, K_t, t) = M_{\Delta_y}^N [a(Y_t, t)]; \\ a_2 &= a_2(m_t, K_t, t) = \\ &= a_{21}(m_t, K_t, t) + a_{21}(m_t, K_t, t)^T + a_{22}(m_t, K_t, t); \\ a_{21} &= a_{21}(m_t, K_t, t) = M_{\Delta_y}^N [a(Y_t, t) Y_t^{0T}]; \end{aligned}$$

$$a_{22} = a_{22}(m_t, K_t, t) = M_{\Delta_y}^N [\bar{\sigma}(Y_t, t)];$$

$$\sigma(Y_t, t) = b(Y_t, t) \nu_0(t) b(Y_t, t)^T;$$

$$\bar{\sigma}(Y_t, t) =$$

$$= \sigma(Y_t, t) + \int_{R_0^d} c(Y_t, t, v) c(Y_t, t, v)^T \nu_P(t, dv), \quad (23)$$

где $M_{\Delta_y}^N$ — символ вычисления математического ожидания для нормальных распределений (21) на гладком многообразии Δ_y .

Для стационарных СтС нормальные стационарные СтП — если они существуют, то $m_t = m^*$, $K_t = K^*$, $K(t_1, t_2) = k(\tau)$ ($\tau = t_1 - t_2$) — определяются уравнениями [1, 5, 6]:

$$\left. \begin{aligned} a_1(m^*, K^*) &= 0; \quad a_2(m^*, K^*) = 0; \\ \dot{k}_\tau(\tau) &= a_{21}(m^*, K^*) K^{*-1} k(\tau); \\ k(0) &= K^* \quad (\forall \tau > 0); \\ k(\tau) &= k(-\tau)^T \quad (\forall \tau < 0). \end{aligned} \right\} \quad (24)$$

При этом необходимо, чтобы матрица $a_{21}(m^*, K^*) = a_{21}^*$ была асимптотически устойчивой.

Уравнения МНА в случае СтС (20) переходят в уравнения МСЛ [1, 5, 6]:

$$\left. \begin{aligned} \dot{m}_t &= a_1(m_t, K_t, t), \quad m_0 = m(t_0); \\ \dot{K}_t &= k_1^a(m_t, K_t, t) K_t + K_t k_1^a(m_t, K_t, t)^T + \\ &+ \sigma_0(t), \quad K_0 = K(t_0); \\ \frac{\partial K(t_1, t_2)}{\partial t_2} &= K(t_1, t_2) K_{t_2} k_1^a(m_{t_2}, K_{t_2}, t_2)^T, \\ &K(t_1, t_2) = K_{t_1}, \end{aligned} \right\} \quad (25)$$

где

$$a(Y_t, t) = a_0(m_t, K_t) + k_1^a(m_t, K_t) Y_t^0;$$

$$k_1^a(m_t, K_t, t) = \left[\left(\frac{\partial}{\partial m_t} \right) a_0(m_t, K_t, t)^T \right]^T;$$

$$b(Y_t, t) = b_0(t); \quad \sigma(Y_t, t) = b_0(t) \nu(t) b_0(t)^T = \sigma_0(t).$$

Для стационарных СтС (20) при условии асимптотической устойчивости матрицы $k_1^a(m^*, K^*)$ в основе МСЛ лежат уравнения (24), записанные в виде:

$$\left. \begin{aligned} a_0(m^*, K^*) &= 0; \\ k_1^a(m^*, K^*) K^* + K^* k_1^a(m^*, K^*)^T + \bar{\sigma}_0 &= 0; \\ \dot{k}_\tau(\tau) &= k_1^a(m^*, K^*) k(\tau); \\ k(0) &= K^* \quad (\forall \tau > 0); \\ k(\tau) &= k(-\tau)^T \quad (\forall \tau < 0). \end{aligned} \right\} \quad (26)$$

Теорема 7. Если существуют интегралы (23), то уравнения (21) и (22) лежат в основе нестационарных алгоритмов МАМ для негауссовских СтС (19), а уравнения (25) — для негауссовских СтС (20).

Теорема 8. Если СтС (19) и (20) стационарны и существует стационарный нормальный процесс и матрица a_{21}^* асимптотически устойчива, то уравнения (24) и (26) лежат в основе стационарных алгоритмов МАМ.

Для гауссовских СтС теоремы 7 и 8 упрощаются, если принять $c(Y_t, t, v) \equiv 0$ в (19) и $V = V_0, \nu^V = v^{V_0}$ в (20).

Для алгоритмизации МНА необходимо уметь вычислять следующие интегралы:

$$\left. \begin{aligned} I_0^a &= I_0^a(m_t, K_t, t) = a_1(m_t, K_t, t) = \\ &= M_{\Delta_y}^N [a(Y_t, t)]; \\ I_1^a &= I_1^a(m_t, K_t, t) = a_{21}(m_t, K_t, t) = \\ &= M_{\Delta_y}^N [a(Y_t, t) Y_t^{0T}]; \\ I_0^{\bar{\sigma}} &= I_0^{\bar{\sigma}}(m_t, K_t, t) = a_{22}(m_t, K_t, t) = \\ &= M_N [\bar{\sigma}(Y_t, t)], \end{aligned} \right\} (27)$$

а для МСЛ достаточно вычислить первый интеграл в (27), причем интеграл I_1^a вычисляется по формуле [1, 5, 6]

$$k_1^a = k_1^a(m_t, K_t, t) = \left[\left(\frac{\partial}{\partial m_t} \right) I_0^a(m_t, K_t, t) \right]^T.$$

Важно иметь в виду, что уравнения МНА (МСЛ) содержат интегралы I_0^a, I_1^a и $I_0^{\bar{\sigma}}$ в виде соответствующих коэффициентов. Поэтому процедура вычисления интегралов должна быть согласована с методом численного решения обыкновенных дифференциальных уравнений для m_t, K_t и $K(t_1, t_2)$. Эти коэффициенты допускают дифференцирование по m_t и K_t , так как под интегралом стоит сглаживающая нормальная плотность.

В [8] изложены алгоритмы дискретного аналитического и статистического моделирования типовых распределений (в том числе нормальных) в нелинейных СтС на многообразиях. Алгоритмы дискретного аналитического и статистического моделирования для СтС с СБН, а также смешанные алгоритмы различной степени точности относительно шага интегрирования также представлены в [8].

В приложении ПЗ приведены тестовые примеры.

9 Заключение

На основе МНА (МСЛ) разработано методическое и алгоритмическое (степенное) обеспечение аналитического моделирования нормальных процессов в гауссовских и негауссовских СтС с БНДП (сферическими, модифицированными сферическими и нелинейностями Эйри).

Алгоритмы положены в основу разрабатываемого инструментального программного обеспечения для решения задач надежности и безопасности систем и средств информатики и управления.

Алгоритмы допускают обобщение на случай аппроксимации БНДП многочленными, дробно-рациональными и другими представлениями.

Теоретический и практический интерес представляют символьные алгоритмы.

Приложения

П1. Функции Бесселя первого и второго рода [2–4]

$$J_\nu(z) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k! \Gamma(k + \nu + 1)} \left(\frac{z}{2} \right)^{2k + \nu} \quad (\nu \neq \pm n);$$

$$Y_\nu(z) = \frac{1}{\sin \nu \pi} [J_\nu(z) \cos \nu \pi - J_{-\nu}(z)] \quad (\nu \neq \pm n);$$

$$Y_n(z) = \frac{2}{\pi} J_n(z) \ln \frac{z}{2} - \frac{1}{\pi} \sum_{k=0}^{n-1} \frac{(n-k-1)!}{k!} \left(\frac{z}{2} \right)^{2k-n} - \frac{1}{\pi} \sum_{k=0}^{\infty} \frac{(-1)^k}{(n+k)! k!} [\psi(n+k+1) + \psi(k+1)] \left(\frac{z}{2} \right)^{2k+n},$$

где $\psi(l) = \Gamma'(l)/\Gamma(l) = (d/dl) \ln \Gamma(l)$;

$$J_{-n}(z) = (-1)^n J_n(z), \quad Y_{-n}(z) = (-1)^n Y_n(z);$$

$$I_\nu(z) = \sum_{k=0}^{\infty} \frac{1}{k! \Gamma(k + \nu + 1)} \left(\frac{z}{2} \right)^{2k + \nu};$$

$$K_\nu(z) = \frac{\pi}{2 \sin \nu \pi} [I_{-\nu}(z) - I_\nu(z)] \quad (\nu \neq \pm n);$$

$$\begin{aligned} K_n(z) &= \\ &= (-1)^{n+1} I_n(z) \ln \frac{z}{2} + \frac{1}{2} \sum_{k=0}^{n-1} \frac{(n-k-1)!}{k!} \left(\frac{z}{2} \right)^{2k-n} + \\ &+ \frac{(-1)^n}{2} \sum_{k=0}^{\infty} \frac{\psi(k+n+1) + \psi(k+1)}{(n+k)! k!} \left(\frac{z}{2} \right)^{n+2k}; \end{aligned}$$

$$\begin{aligned}
J_{\pm n \pm (1/2)}(z) &= \\
&= (\mp)^n \sqrt{\frac{2}{\pi}} z^{n+(1/2)} \left(\frac{1}{z} \frac{d}{dz} \right)^n \left(\frac{1}{z} \left\{ \frac{\sin}{z} \cos z \right\} \right) = \\
&= \sqrt{\frac{2}{\pi z}} \left[\left\{ \begin{array}{l} \sin \left(z - \frac{n\pi}{2} \right) \\ \cos \left(z + \frac{n\pi}{2} \right) \end{array} \right\} \sum_{k=0}^{[n/2]} \frac{(-1)^k (n+2k)!}{(2k)!(n-2k)!(2z)^{2k}} \pm \right. \\
&\quad \left. \pm \left\{ \begin{array}{l} \cos \left(z - \frac{n\pi}{2} \right) \\ \sin \left(z + \frac{n\pi}{2} \right) \end{array} \right\} \times \right. \\
&\quad \left. \times \sum_{k=0}^{[(n-1)/2]} \frac{(-1)^k (n+2k+1)!}{(2k+1)!(n-2k-1)!(2z)^{2k+1}} \right];
\end{aligned}$$

$$J_{\pm(1/2)}(z) = \sqrt{\frac{2}{\pi z}} \left\{ \begin{array}{l} \sin z \\ \cos z \end{array} \right\};$$

$$J_{\pm(3/2)}(z) = \sqrt{\frac{2}{\pi z}} \left[\pm \frac{1}{z} \left\{ \begin{array}{l} \sin z \\ \cos z \end{array} \right\} - \left\{ \begin{array}{l} \cos z \\ \sin z \end{array} \right\} \right];$$

$$J_{\pm(5/2)}(z) = \sqrt{\frac{2}{\pi z}} \left[\left(\frac{3}{z^2} - 1 \right) \left\{ \begin{array}{l} \sin z \\ \cos z \end{array} \right\} \mp \frac{3}{z} \left\{ \begin{array}{l} \cos z \\ \sin z \end{array} \right\} \right];$$

$$Y_{\pm n \pm 1/2}(z) = \mp (-1)^n J_{\mp n \mp 1/2}(z);$$

$$Y_{\pm 1/2}(z) = \mp \sqrt{\frac{2}{\pi z}} \left\{ \begin{array}{l} \cos z \\ \sin z \end{array} \right\};$$

$$I_{-n}(z) = I_n(z); \quad I_n(z) = (-1)^n I_n(z);$$

$$\begin{aligned}
I_{\pm n \pm 1/2}(z) &= z^n \left(\frac{1}{z} \frac{d}{dz} \right)^n \left[\frac{1}{z} \left\{ \begin{array}{l} \text{sh } z \\ \text{ch } z \end{array} \right\} \right] = \\
&= \frac{1}{\sqrt{2\pi z}} \left[e^z \sum_{k=0}^n \frac{(-1)^k (n+k)!}{k!(n-k)!(2z)^k} \pm \right. \\
&\quad \left. + \pm (-1)^{n+1} e^{-z} \sum_{k=0}^n \frac{(n+k)!}{k!(n-k)!(2z)^k} \right];
\end{aligned}$$

$$I_{\pm 1/2}(z) = \sqrt{\frac{2}{\pi z}} \left\{ \begin{array}{l} \text{sh } z \\ \text{ch } z \end{array} \right\};$$

$$I_{\pm 3/2}(z) = \sqrt{\frac{2}{\pi z}} \left[\left\{ \begin{array}{l} \text{ch } z \\ \text{sh } z \end{array} \right\} \mp \frac{1}{z} \left\{ \begin{array}{l} \text{sh } z \\ \text{ch } z \end{array} \right\} \right];$$

$$I_{\pm 5/2}(z) = \sqrt{\frac{2}{\pi z}} \left[\left(\frac{3}{2z^2} + 1 \right) \left\{ \begin{array}{l} \text{sh } z \\ \text{ch } z \end{array} \right\} - \frac{3}{z} \left\{ \begin{array}{l} \text{ch } z \\ \text{sh } z \end{array} \right\} \right];$$

$$K_{n+1/2}(z) = \sqrt{\frac{\pi}{2z}} e^{-z} \sum_{k=0}^n \frac{(n+k)!}{k!(n-k)!(2z)^k};$$

$$K_{\pm 1/2}(z) = \sqrt{\frac{\pi}{2z}} e^{-z}; \quad K_{\pm 3/2}(z) = \frac{\pi}{2z} \left(1 + \frac{1}{z} \right) e^{-z}.$$

П2. Некоторые свойства функции Бесселя дробного порядка

Формула Рэлея:

$$j_n(z) = z^n \left(-\frac{1}{z} \frac{d}{dz} \right)^n \frac{\sin z}{z};$$

$$y_n(z) = -z^n \left(-\frac{1}{z} \frac{d}{dz} \right)^n \frac{\cos^2 z}{z};$$

$$i_n(z) = z^n \left(\frac{1}{z} \frac{d}{dz} \right)^n \frac{\text{sh } z}{z};$$

$$\tilde{i}_n(z) = z^n \left(\frac{1}{z} \frac{d}{dz} \right)^n \frac{\text{ch } z}{z}.$$

Производные w_n : $j_n, i_n, y_n, \tilde{i}_n$:

$$\left(\frac{1}{z} \frac{d}{dz} \right)^m [z^{n+1} w_n(z)] = z^{n-m+1} w_{n-m}(z);$$

$$\left(\frac{1}{z} \frac{d}{dz} \right)^m [z^{-n} w_n(z)] = (\pm)^m z^{-n-m} w_{n+m}(z)$$

$$(n = 0, \pm 1, \pm 2, \dots, m = 1, 2, \dots).$$

Знак минус относится к j_n и y_n , а плюс и минус — к i_n .

Основные соотношения:

$$y_n(z) = (-1)^{n+1} j_{-n-1}(z) \quad (n = 0, \pm 1, \pm 2, \dots);$$

$$\begin{aligned}
\text{Ai}(z) &= \frac{1}{3} \sqrt{z} [I_{-1/3}(b) - I_{1/3}(b)] = \\
&= \frac{1}{\pi} \sqrt{\frac{z}{3}} K_{1/3}(b) \quad \left(b = \frac{2}{3} z^{3/2} \right);
\end{aligned}$$

$$\text{Ai}(-z) = \frac{1}{3} \sqrt{z} [J_{1/3}(b) + J_{-1/3}(b)];$$

$$\text{Bi}(z) = \sqrt{\frac{z}{3}} [J_{-1/3}(b) + J_{1/3}(b)];$$

$$\text{Bi}(-z) = \sqrt{\frac{z}{3}} [J_{-1/3}(b) - J_{1/3}(b)];$$

$$j_0(z) = \frac{\sin z}{z}; \quad j_1(z) = \frac{\sin z}{z^2} - \frac{\cos z}{z};$$

$$j_2(z) = \left(\frac{3}{z^3} - \frac{1}{z} \right) \sin z - \frac{3}{z^2} \cos z;$$

$$y_0(z) = -j_{-1}(z) = -\frac{\cos z}{z};$$

$$y_1(z) = j_{-2}(z) = -\frac{\cos z}{z^2} - \frac{\sin z}{z};$$

$$y_2(z) = -j_{-3}(z) = \left(-\frac{3}{z^3} + \frac{1}{z} \right) \cos z - \frac{3}{z^2} \sin z;$$

$$i_0(z) = \frac{\text{sh } z}{z}; \quad i_1(z) = -\frac{\text{sh } z}{z^2} + \frac{\text{ch } z}{z};$$

$$i_2(z) = \left(\frac{3}{z^3} + \frac{1}{z} \right) \text{sh } z - \frac{3}{z^2} \text{ch } z;$$

$$\tilde{i}_0(z) = \frac{\text{ch } z}{z}; \quad \tilde{i}_1(z) = \frac{\text{sh } z}{z} - \frac{\text{ch } z}{z};$$

$$\tilde{i}_2(z) = -\frac{3}{z^2} \text{sh } z + \left(\frac{3}{z^3} + \frac{1}{z} \right) \text{ch } z.$$

Рекуррентные соотношения $w_n(z) : j_n(z), y_n(z)$:

$$w_{n-1}(z) + w_{n+1}(z) = \frac{2n+1}{z} w_n(z);$$

$$nw_{n-1}(z) - (n+1)w_{n+1}(z) = (2n+1) \frac{d}{dz} w_n(z);$$

$$\frac{n+1}{z} w_n(z) + \frac{d}{dz} w_n(z) = w_{n-1}(z);$$

$$\frac{n}{z} w_{n-1}(z) - \frac{d}{dz} w_n(z) = w_{n+1}(z);$$

$$i_{n-1}(z) - i_{n+1}(z) = \frac{(2n+1)}{z} w_n;$$

$$ni_{n-1}(z) + (n+1)i_{n+1}(z) = (2n+1) \frac{d}{dz} i_n(z);$$

$$\frac{n+1}{z} i_n(z) + \frac{d}{dz} i_n(z) = i_{n-1}(z);$$

$$-\frac{n}{z} i_n(z) + \frac{d}{dz} i_n(z) = i_{n+1}(z).$$

Точность вычислений может быть проверена как с помощью тождеств

$$j_n(z)y_{n-1}(z) - j_{n-1}(z)y_n(z) = z^{-2};$$

$$y_{n+1}(z)y_{n-1}(z) - j_{n-1}(z)y_{n+1}(z) = (2n+1)z^{-2};$$

$$i_n(z)\tilde{i}_n(z) - i_{n-1}(z)\tilde{i}_{n-1}(z) = (-1)^{n+1}z^{-2},$$

так и нахождением нулей [3–6].

П3. Тестовые примеры

П3.1. При $m_y = 0$ для сложных нелинейностей вида

$$\varphi(Y) = J_{1/2}(Y)1(Y) = 1(Y)\sqrt{\frac{2}{\pi|Y|}} \sin Y;$$

$$\varphi(Y) = J_{-1/2}(Y)1(Y) = 1(Y)\sqrt{\frac{2}{\pi|Y|}} \cos Y,$$

используя следующие известные табличные интегралы [3–6]:

$$\int_0^\infty x^{\mu-1} e^{-\beta x^2} \sin ax \, dx = \frac{ae^{-a^2/4\beta}}{2^{\beta(\mu+1)/2}} \Gamma\left(\frac{1+\mu}{2}\right) {}_1F_1\left(\frac{\mu}{2}; \frac{3}{2}; -\frac{a^2}{4\beta}\right) \quad (\text{Re } \beta > 0, \text{ Re } \mu > -1);$$

$$\int_0^\infty x^{\mu-1} e^{-\beta x^2} \cos ax \, dx = \frac{\Gamma(\mu/2)}{2\beta^{\mu/2}} {}_1F_1\left(\frac{\mu}{2}; \frac{1}{2}; -\frac{a^2}{4\beta}\right) \quad (\text{Re } \beta > 0, \text{ Re } \mu > 0),$$

при $\mu = 1/2$ получаем точные выражения для $\varphi_0(0, D_y)$:

$$\varphi_0(0, D_y) = \frac{\Gamma(3/4)}{(2D_y)^{1/4}} e^{-D_y/2} {}_1F_1\left(\frac{1}{4}; \frac{3}{2}; -\frac{D_y}{2}\right);$$

$$\varphi_0(0, D_y) = \frac{1}{2^{5/4} D_y^{1/4}} {}_1F_1\left(\frac{1}{4}; \frac{3}{2}; -\frac{D_y}{2}\right).$$

Здесь ${}_1F_1(a; b; \xi)$ – вырожденная гипергеометрическая функция [3–6], допускающая представление:

$${}_1F_1(a; b; \xi) = 1 + \frac{a}{b} \frac{\xi}{1!} + \frac{a(a+1)}{b(b+1)} \frac{\xi^2}{2!} + \frac{a(a+1)(a+2)}{b(b+1)(b+2)} \frac{\xi^3}{3!} + \dots$$

При $m_y \neq 0$ и для $\mu = 1/2$, если принять во внимание следующие табличные интегралы [3–6]:

$$\int_0^\infty x^{\mu-1} e^{-\gamma x - \beta x^2} \sin ax \, dx = -\frac{i}{2(2\beta)^{\mu/2}} \exp\left(\frac{\gamma^2 - a^2}{8\beta}\right) \Gamma(\mu) \times \left\{ \exp\left(-\frac{ia\gamma}{4\beta}\right) D_{-\mu}\left(\frac{\gamma - ia}{\sqrt{2\beta}}\right) - \exp\left(\frac{ia\gamma}{4\beta}\right) D_{-\mu}\left(\frac{\gamma + ia}{\sqrt{2\beta}}\right) \right\}$$

$$(\gamma^2 = -1, \text{ Re } \mu > -1, \text{ Re } \beta > 0, a > 0);$$

$$\int_0^\infty x^{\mu-1} e^{-\gamma x - \beta x^2} \cos ax \, dx = \frac{1}{2(2\beta)^{\mu/2}} \exp\left(\frac{\gamma^2 - a^2}{8\beta}\right) \Gamma(\mu) \times \left\{ \exp\left(-\frac{ia\gamma}{4\beta}\right) D_{-\mu}\left(\frac{\gamma^2 - ia}{\sqrt{2\beta}}\right) + \exp\left(\frac{ia\gamma}{4\beta}\right) D_{-\mu}\left(\frac{\gamma^2 + ia}{\sqrt{2\beta}}\right) \right\} \quad (\text{Re } \mu > 0, \text{ Re } \beta > 0, a > 0),$$

имеем следующие точные выражения:

$$\varphi_0(m_y, D_y) = -\frac{i(2\beta)^{1/4}}{2^{3/2}} \exp\left(-\frac{m_y^2}{2D_y}\right) \times \left[\exp\left(-\frac{i\gamma}{4\beta}\right) D_{-1/2}\left(\frac{\gamma - i}{\sqrt{2\beta}}\right) - \exp\left(\frac{i\gamma}{4\beta}\right) D_{-1/2}\left(\frac{\gamma + i}{\sqrt{2\beta}}\right) \right];$$

$$\varphi_0(m_y, D_y) = \frac{(2\beta)^{1/4}}{2^{3/2}} \exp\left(-\frac{m_y^2}{2D_y}\right) \times \left[\exp\left(-\frac{i\gamma}{4\beta}\right) D_{-1/2}\left(\frac{\gamma - i}{\sqrt{2\beta}}\right) + \exp\left(\frac{i\gamma}{4\beta}\right) D_{-1/2}\left(\frac{\gamma + i}{\sqrt{2\beta}}\right) \right].$$

Здесь $D_\mu(\cdot)$ — функция параболического цилиндра [2–4];

$$\gamma = \gamma(m_y, D_y) = -\frac{m_y}{D_y}; \quad \beta = \beta(D_y) = \frac{1}{2D_y}.$$

ПЗ.2. Для одномерной бесселевой системы с аддитивным белым шумом

$$\dot{Y}_t = \alpha + \beta w_n(cY_t) + \gamma V$$

(α, β, γ и c — постоянные; V — гауссовский белый шум интенсивности ν , $w_n: j_n, i_n, y_n, i_n, Ai, Vi$) алгоритмы МАМ на основе МСЛ (теоремы 7 и 8) имеют следующий вид:

$$\left. \begin{aligned} \dot{m}_y &= \alpha + \beta \varphi_0^{w_n}(m_y, D_y); \\ \dot{D}_y &= 2\beta k_1^{w_n}(m_y, D_y) + \sigma \quad (\sigma = \gamma^2 \nu); \end{aligned} \right\} \quad (28)$$

$$\left. \begin{aligned} \alpha^* + \beta^* \varphi_0^{*w_n}(m^*, D^*) &= 0; \\ 2\beta^* k_1^{*w_n}(m^*, D^*) D^* + \sigma^* &= 0 \quad (\beta > 0). \end{aligned} \right\} \quad (29)$$

Коэффициенты $\varphi_0^{w_n}$ и $\varphi_0^{*w_n}$ в (28) и (29) определяются согласно алгоритмам теорем 1–6.

ПЗ.3. Для одномерной бесселевой системы с мультипликативным гауссовским белым шумом

$$\dot{Y}_t = \alpha + \beta Y_t + \gamma w_n(cY_t) V$$

алгоритм МАМ согласно МНА определяется следующими уравнениями:

$$\dot{m}_y = \alpha + \beta m_y; \quad \dot{D}_y = 2\beta D_y + \sigma(m_y, D_y);$$

$$\alpha^* + \beta^* m^* = 0; \quad 2\beta^* D^* + \sigma(m^*, D^*) = 0 \quad (\beta < 0).$$

Здесь

$$\begin{aligned} \sigma(m_y, D_y) &= \\ &= \frac{\gamma^2 \nu}{\sqrt{2\pi D_y}} \int_{-\infty}^{\infty} w_n^2(c\eta) e^{-(\eta - m_y)^2 / (2D_y)} d\eta. \end{aligned} \quad (30)$$

Интеграл (30) вычисляются на основе теорем 1–6, если принять во внимание известную формулу для квадрата степенного ряда [3]:

$$\left(\sum_{h=0}^{\infty} a_h z^h \right)^2 = \sum_{h=0}^{\infty} b_h z^h,$$

$$b_0 = a_0, \quad b_h = \frac{1}{h a_0} \sum_{\rho=1}^h (3\rho - h) a_\rho b_{h-\rho}.$$

Литература

1. Синицын И. Н. Аналитическое моделирование процессов в динамических системах с цилиндрическими бесселевыми нелинейностями // Информатика и её применения, 2015. Т. 9. Вып. 4. С. 39–49.
2. Градштейн И. С., Рыжик И. М. Таблицы интегралов, сумм, рядов и произведений. — М.: ГИФМЛ, 1963. 1100 с.
3. Справочник по специальным функциям / Под ред. М. Абрамовича, И. Стигана. — М.: Наука, 1979. 832 с.
4. Попов Б. А., Теслер Г. С. Вычисление функций на ЭВМ: Справочник. — Киев: Наукова Думка, 1984. 599 с.
5. Пугачёв В. С., Синицын И. Н. Стохастические дифференциальные системы. Анализ и фильтрация. — М.: Наука, 1990. 632 с.
6. Пугачёв В. С., Синицын И. Н. Теория стохастических систем. — М.: Логос, 2000; 2004. 1000 с.
7. Синицын И. Н., Синицын В. И. Лекции по нормальной и эллипсоидальной аппроксимации распределений в стохастических системах. — М.: ТОРУС ПРЕСС, 2013. 488 с.
8. Синицын И. Н. Параметрическое статистическое и аналитическое моделирование распределений в нелинейных стохастических системах на многообразиях // Информатика и её применения, 2013. Т. 7. Вып. 2. С. 4–16.

Поступила в редакцию 22.03.16

ANALYTICAL MODELING OF PROCESSES IN STOCHASTIC SYSTEMS WITH COMPLEX FRACTIONAL ORDER BESSEL NONLINEARITIES

I. N. Sinitsyn

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

Abstract: Methods of analytical modeling for normal (Gaussian) processes in Gaussian and non-Gaussian stochastic systems with complex fractional order Bessel nonlinearities (spherical, modified spherical, and Airy) are developed. Necessary information about Bessel fractional order functions is given. Coefficients of statistical

linearization for typical fractional order Bessel nonlinearities are presented. Special attention is paid to the series algorithms. Analytical modeling algorithms have been developed for nonstationary and stationary normal processes. Test examples are presented. Main conclusions and generalizations are mentioned.

Keywords: Airy nonlinearity; Bessel function of fractional order; Bessel nonlinearity; method of analytical modeling; method of normal approximation; method of statistical linearization; modified spherical Bessel function; normal (Gaussian) process; spherical Bessel function; stochastic process

DOI: 10.14357/19922264160308

Acknowledgments

The work was supported by the Department for Nanotechnologies and Information Technologies (ONIT) of the Russian Academy of Sciences (project 0063-2015-0017.III.3).

References

1. Sinitsyn, I. N. 2015. Analiticheskoe modelirovanie protsessov v dinamicheskikh sistemakh s tsilindricheskimi besselevymi nelineynostyami [Analytical modeling of processes in dynamical systems with cylindrical Bessel nonlinearities]. *Informatika i ee Primeneniya — Inform. Appl.* 9(4):39–49.
2. Gradshteyn, I. S., and I. M. Ryzhik. 1963. *Tablitsy integralov, summ, ryadov i proizvedeniy* [Tables of integrals, sums, series and products]. Moscow: GIFML. 1100 p.
3. Abramovich, M., and I. Stigan, eds. 1979. *Spravochnik po spetsial'nym funktsiyam* [Handbook of mathematical functions]. Moscow: Nauka. 832 p.
4. Popov, B. A., and G. S. Tesler. 1984. *Vychislenie funktsiy na EVM: Spravochnik* [Calculation of functions on the computer: Handbook]. Kiev: Naukova Dumka. 599 p.
5. Pugachev, V. S., and I. N. Sinitsyn. 1987. *Stochastic differential systems. Analysis and filtering*. Chichester – New York, NY: John Wiley. 549 p.
6. Pugachev, V. S., and I. N. Sinitsyn. 2000, 2004. *Teoriya stokhasticheskikh sistem* [Stochastic systems. Theory and applications]. Moscow: Logos. 1000 p.
7. Sinitsyn, I. N., and V. I. Sinitsyn. 2013. *Lektsii po normal'noy i ellipsoidal'noy approksimatsii raspredeleniy v stokhasticheskikh sistemakh* [Lectures on normal and ellipsoidal approximation of distributions in stochastic systems]. Moscow: TORUS PRESS. 488 p.
8. Sinitsyn, I. N. 2013. Parametricheskoe statisticheskoe i analiticheskoe modelirovanie raspredeleniy v nelineynykh stokhasticheskikh sistemakh na mnogoobraznykh [Parametrical statistical and analytical modeling of distributions in nonlinear stochastic systems on manifolds]. *Informatika i ee Primeneniya — Inform. Appl.* 7(2):4–16.

Received March 22, 2016

Contributor

Sinitsyn Igor N. (b. 1940) — Doctor of Science in technology, professor, Honored scientist of RF, Head of Department, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; sinitsin@dol.ru

АСИМПТОТИЧЕСКИЕ РАЗЛОЖЕНИЯ СРЕДНЕЙ АБСОЛЮТНОЙ ОШИБКИ НЕСМЕЩЕННОЙ ОЦЕНКИ С РАВНОМЕРНО МИНИМАЛЬНОЙ ДИСПЕРСИЕЙ И ОЦЕНКИ МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ В МОДЕЛИ ОДНОПАРАМЕТРИЧЕСКОГО ЭКСПОНЕНЦИАЛЬНОГО СЕМЕЙСТВА РЕШЕТЧАТЫХ РАСПРЕДЕЛЕНИЙ*

В. В. Чичагов¹

Аннотация: Рассмотрена модель повторной выборки фиксированного объема n из решетчатого распределения, принадлежащего естественному однопараметрическому экспоненциальному семейству. При неограниченном возрастании n найдены асимптотические разложения средних абсолютных ошибок несмещенной оценки с равномерно минимальной дисперсией (НОРМД) и оценки максимального правдоподобия (ОМП) заданной параметрической функции $G[a]$. Отдельно исследован случай, когда $G'[a] = 0$, но $G''[a] \neq 0$. В случае распределения Пуассона для двух параметрических функций проведена оценка относительной погрешности вычисления разности средних абсолютных ошибок НОРМД и ОМП с помощью полученных асимптотических разложений. Установлено, что асимптотические результаты при достаточно большом объеме выборки позволяют сравнивать НОРМД и ОМП с помощью такого показателя качества оценок, как средняя абсолютная ошибка.

Ключевые слова: экспоненциальное семейство; решетчатое распределение; несмещенная оценка; оценка максимального правдоподобия; асимптотическое разложение

DOI: 10.14357/19922264160309

1 Введение

Несмещенные оценки с равномерно минимальной дисперсией и оценки максимального правдоподобия играют существенную роль в современных статистических исследованиях. Выбор лучшей из них можно сделать, используя одну из мер близости статистической оценки к истинному значению оцениваемой функции. В данной работе в качестве такой меры близости используется средняя абсолютная ошибка статистической оценки.

Определение. Средней абсолютной ошибкой статистической оценки $T[\mathbf{X}]$ при оценивании значения параметрической функции $G[a]$ будем называть функционал

$$\text{MAE}[T[\mathbf{X}], G[a]] = \mathbf{E}|T[\mathbf{X}] - G[a]|, \quad (1)$$

где $\mathbf{X} = (X_1, X_2, \dots, X_n)$ — выборка, по которой строится оценка $T[\mathbf{X}]$.

Этот показатель качества статистической оценки является более естественной мерой ее точности, но в меньшей степени поддается математическому анализу по сравнению с такой распространенной мерой точности, как среднеквадратическая ошибка статистической оценки. Получить явные форму-

лы для вычисления функционала (1) удастся лишь в редких случаях.

Целью данной работы является получение асимптотических разложений для НОРМД и ОМП, справедливых при большом объеме выборки n в случае однопараметрического экспоненциального семейства решетчатых распределений. Случай однопараметрического экспоненциального семейства абсолютно непрерывных распределений изучен в [1]. Задача получения асимптотических разложений среднеквадратических ошибок НОРМД и ОМП при большом объеме выборки n в случае однопараметрического экспоненциального семейства распределений была рассмотрена в работе [2]. Более общие результаты, чем в [2], для НОРМД представлены в [3].

2 Описание модели наблюдений, основные предположения и обозначения

Имеется $\mathbf{X} = (X_1, X_2, \dots, X_n)$ — повторная выборка, элементы которой являются независимыми случайными величинами, имеющими то же самое

* Работа выполнена при финансовой поддержке Министерства образования и науки РФ (проект 2096).

¹ Пермский государственный национальный исследовательский университет, chichagov@psu.ru

распределение, что и наблюдаемая случайная величина ξ .

Распределение случайной величины ξ принадлежит экспоненциальному семейству решетчатых распределений, определяемых набором вероятностей

$$\mathbf{P}(\xi = x) = \exp \{ \Phi_1[a] T[x] - \kappa[\Phi_1[a]] + d[x] \}, \quad x \in \mathbb{X}_\xi \subset \mathbf{Z}, \quad (2)$$

с параметром $a = \mathbf{E}T[\xi] \in \mathbb{A}$. Здесь $d[x]$, $T[x]$ и $\Phi_1[a]$ — заданные известные функции; $\kappa[\theta]$ — кумулянтное преобразование распределения.

Будем предполагать, что распределение (2) удовлетворяет следующим условиям регулярности.

(A₁). Носитель распределения \mathbb{X}_ξ содержится в \mathbf{Z} , но ни в какой подрешетке \mathbf{Z} и не зависит от параметра a .

(A₂). Пусть $\tilde{\Theta}$ — множество значений θ , удовлетворяющих соотношению

$$\sum_{x \in \mathbb{X}_G} \exp \{ \theta T[x] + d[x] \} < \infty.$$

Параметрическое множество $\mathbb{A} = \text{Int} \{ a = \kappa'[\theta] : \theta \in \tilde{\Theta} \}$ не пусто.

(A₃). $\Phi_1[a]$ — бесконечное число раз дифференцируемая функция, причем $\Phi_1'[a] > 0$ при $a \in \mathbb{A}$.

Далее в работе будем придерживаться следующих обозначений:

$G[a]$ — заданная параметрическая функция, $a \in \mathbb{A}$;

$S_n = \sum_{i=1}^n T(X_i)$ — минимальная достаточная статистика распределения (2);

$\widehat{G}[a|S_n]$ — НОРМД функции $G[a]$;

$\widetilde{G}[a|S_n]$ — ОМП функции $G[a]$;

$\varphi[x] = (1/\sqrt{2\pi}) \exp[-x^2/2]$ — плотность стандартного нормального распределения;

$$\Phi[x] = \int_{-\infty}^x \varphi[t] dt;$$

$U^{(k)}[a] = d^k U[a]/da^k$ — производная k -го порядка от функции $U[a]$;

$U[a]^j = (U[a])^j$ — j -я степень функции $U[a]$;

$[x]$ — дробная часть x ;

$\ell_0[x] = 1$, $\ell_1[x] = [x] - 1/2$, $\ell_2[x] = (1/2)([x]^2 - [x] + 1/6)$, $0 \leq x < 1$, — функции, используемые в формуле суммирования Эйлера–Маклорена;

$I(D)$ — индикатор события D .

3 Основной результат и его доказательство

Теорема. Пусть выполнены условия (A₁)–(A₃), а последовательности среднеквадратических ошибок $\{ \mathbf{V}\widehat{G}[a|S_n] \}$ и $\{ \mathbf{E}(\widetilde{G}[a|S_n] - G[a])^2 \}$ равномерно ограничены начиная с некоторого $n \geq n_0$. Тогда при $n \rightarrow \infty$ и $G'[a] \neq 0$ справедливы разложения:

$$\begin{aligned} \mathbf{E} \left| \widehat{G}[a|S_n] - G[a] \right| &= \\ &= \frac{2\varphi[0] |G'[a]|}{\sqrt{n\Phi_1'[a]}} \left\{ 1 - \frac{\Phi_1'[a]}{n} \ell_2[na + \epsilon] + \right. \\ &+ \frac{1}{n\Phi_1'[a]} \left(\frac{G''[a]^2}{8G'[a]^2} + \frac{G''[a]\Phi_1''[a]}{4G'[a]\Phi_1'[a]} - \frac{\Phi_1''[a]^2}{12\Phi_1'[a]^2} - \right. \\ &\left. \left. - \frac{G^{(3)}[a]}{6G'[a]} + \frac{\Phi_1^{(3)}[a]}{24\Phi_1'[a]} \right) \right\} + o\left(\frac{1}{n^{3/2}}\right), \quad (3) \end{aligned}$$

где

$$\epsilon = \frac{G''[a]}{2G'[a]\Phi_1'[a]};$$

$$\begin{aligned} \mathbf{E} \left| \widetilde{G}[a|S_n] - G[a] \right| &= \\ &= \frac{2\varphi[0] |G'[a]|}{\sqrt{n\Phi_1'[a]}} \left\{ 1 - \frac{\Phi_1'[a]}{n} \ell_2[na] + \right. \\ &+ \frac{1}{n\Phi_1'[a]} \left(- \frac{G''[a]\Phi_1''[a]}{6G'[a]\Phi_1'[a]} - \frac{\Phi_1''[a]^2}{12\Phi_1'[a]^2} + \frac{G^{(3)}[a]}{3G'[a]} + \right. \\ &\left. \left. + \frac{\Phi_1^{(3)}[a]}{24\Phi_1'[a]} \right) \right\} + o\left(\frac{1}{n^{3/2}}\right). \quad (4) \end{aligned}$$

Если $G'[a] = 0$, $G''[a] \neq 0$, то при $n \rightarrow \infty$

$$\mathbf{E} \left| \widehat{G}[a|S_n] - G[a] \right| = \sqrt{\frac{2}{\pi e}} \frac{|G''[a]|}{n\Phi_1'[a]} + o\left(\frac{1}{n}\right); \quad (5)$$

$$\mathbf{E} \left| \widetilde{G}[a|S_n] - G[a] \right| = \frac{|G''[a]|}{2n\Phi_1'[a]} + o\left(\frac{1}{n}\right). \quad (6)$$

Примечание 1. Если из разложений (3) и (4) исключить слагаемые, содержащие в качестве одного из множителей функцию $\ell_2[\cdot]$, то в результате получим асимптотические разложения средних абсолютных ошибок НОРМД и ОМП функции $G[a]$, соответствующие случаю, когда распределение случайной величины ξ абсолютно непрерывно и принадлежит экспоненциальному семейству. Этот результат впервые получен в [1].

Примечание 2. Поскольку $\sqrt{2/(\pi e)} < 1/2$, то при достаточно большом объеме выборки n и $G'[a] = 0$,

но $G'''[a] \neq 0$ средняя абсолютная ошибка НОРМД меньше средней абсолютной ошибки ОМП функции $G[a]$ при любом $a \in \mathbb{A}$.

Примечание 3. Предположение о равномерной ограниченности последовательности дисперсий $\{VG[a|S_n]\}$ выполнено, если найдется некоторое $L \in \mathbb{N}$, для которого $VG[a|S_L] < \infty$ [3, лемма 5.11]. Доказательство теоремы. Приведем сначала основные обозначения и соотношения, которые будут использоваться для доказательства утверждения теоремы.

Положим

$$Z_n = \frac{S_n - na}{b\sqrt{n}}$$

при $b = \sqrt{VT[\xi]}$.

При $|z| \leq \sqrt{2\alpha \ln n}$ для любых $\alpha > 0$, $0 < \delta_1 < 0,5$, $0 < \delta_2 < 0,5$ верны следующие асимптотические разложения функций $\hat{G}[a; z]$ и $\tilde{G}[a; z]$, определяющих НОРМД $\hat{G}[a; Z_n] = \hat{G}[a|S_n]$ и ОМП $\tilde{G}[a; Z_n] = \tilde{G}[a|S_n]$ параметрической функции $G[a]$:

$$\hat{G}[a; z] - G[a] = \frac{G'[a]z}{\sqrt{n\Phi_1'[a]}} + \frac{G''[a]H_2[z]}{2n\Phi_1'[a]} + \frac{G'''[a]H_3[z]}{6n^{3/2}\Phi_1'[a]^{3/2}} + \frac{G''[a]\Phi_1''[a]z}{2n^{3/2}\Phi_1'[a]^{5/2}} + O(n^{-2+\delta_1}); \quad (7)$$

$$\tilde{G}[a; z] - G[a] = \frac{G'[a]z}{\sqrt{n\Phi_1'[a]}} + \frac{G''[a]z^2}{2n\Phi_1'[a]} + \frac{G'''[a]z^3}{6n^{3/2}\Phi_1'[a]^{3/2}} + O(n^{-2+\delta_2}). \quad (8)$$

Справедливость разложения (7) установлена в следствии 6.2 из [3]. В условиях теоремы ОМП функции $G[a]$ равна $\hat{G}[a; Z_n] = G[S_n/n]$, а несмещенно оцениваемая функция $G[a]$ имеет производные всех порядков (см., например, [3, следствие 5.3]). Поэтому разложение (8) нетрудно получить с помощью формулы Тейлора, примененной к функции $G[s/n]$ в точке a при $z = (s - na)/(b\sqrt{n})$:

$$\begin{aligned} \tilde{G}_n[a; z] &= G\left[\frac{s}{n}\right] = \\ &= \sum_{j=0}^3 \frac{G^{(j)}[a]}{j!} \left(\frac{s}{n} - a\right)^j + \frac{G^{(4)}[\eta]}{4!} \left(\frac{s}{n} - a\right)^4 = \\ &= \sum_{j=0}^3 \frac{G^{(j)}[a]}{j!} \left(\frac{bz}{\sqrt{n}}\right)^j + \frac{G^{(4)}[\eta]}{4!} \left(\frac{bz}{\sqrt{n}}\right)^4 = \\ &= \sum_{j=0}^3 \frac{G^{(j)}[a]}{j!} \left(\frac{z}{\sqrt{n\Phi_1'[a]}}\right)^j + O\left(\frac{\ln^2 n}{n^2}\right), \end{aligned}$$

где η принадлежит интервалу с концами s/n и a . При этом использован тот факт, что для распределения (2)

$$b = \frac{1}{\sqrt{\Phi_1'[a]}}.$$

Кроме того, потребуются приближения Эджворта 2-го и 3-го порядка

$$p_2[z, n] = \varphi[z] \left(1 + \frac{\rho_3 H_3[z]}{6\sqrt{n}} + \frac{\rho_4 H_4[z]}{24n} + \frac{\rho_3^2 H_6[z]}{72n} \right); \quad (9)$$

$$p_3[z, n] = \varphi[z] \left(1 + \frac{\rho_3 H_3[z]}{6\sqrt{n}} + \frac{\rho_4 H_4[z]}{24n} + \frac{\rho_3^2 H_6[z]}{72n} + \frac{\rho_5 H_5[z]}{5!n^{3/2}} + \frac{\rho_3 \rho_4 H_7[z]}{3!4!n^{3/2}} + \frac{\rho_3^3 H_9[z]}{(3!)^4 n^{3/2}} \right), \quad (10)$$

где $H_j[x]$ — полином Чебышёва–Эрмита порядка j , а ρ_3 и ρ_4 — коэффициенты асимметрии и эксцесса случайной величины $T[\xi]$, которые, как отмечается в [2], определяются выражениями:

$$\rho_3 = -\frac{\Phi_1''[a]}{\Phi_1'[a]^{3/2}}; \quad \rho_4 = \frac{3\Phi_1''[a]^2}{\Phi_1'[a]^3} - \frac{\Phi_1'''[a]}{\Phi_1'[a]^2}. \quad (11)$$

Формула, определяющая явный вид нормированного кумулянта ρ_5 , не приводится, поскольку в утверждении теоремы его нет.

Рассмотрим сначала случай $G'[a] \neq 0$. В этом случае соотношения (7) и (8) удобнее представить иначе:

$$\begin{aligned} \hat{G}[a; z] - G[a] &= \\ &= \frac{G'[a]}{\sqrt{n\Phi_1'[a]}} \left\{ h_1[z, n] + O(n^{-3/2+\delta_1}) \right\}; \quad (12) \\ \tilde{G}[a; z] - G[a] &= \\ &= \frac{G'[a]}{\sqrt{n\Phi_1'[a]}} \left\{ h_2[z, n] + O(n^{-3/2+\delta_2}) \right\}, \end{aligned}$$

где

$$\begin{aligned} h_1[z, n] &= z + \frac{AH_2[z]}{\sqrt{n}} + \frac{B_1 H_3[z] + B_2 z}{n}; \\ h_2[z, n] &= z + \frac{Az^2}{\sqrt{n}} + \frac{Bz^3}{n}; \\ A &= \frac{G''[a]}{2G'[a]\sqrt{\Phi_1'[a]}}; \quad B = \frac{G'''[a]}{6G'[a]\Phi_1'[a]}; \\ B_1 &= \frac{G'''[a]}{6G'[a]\Phi_1'[a]}; \quad B_2 = \frac{G''[a]\Phi_1''[a]}{2G'[a]\Phi_1'[a]^2}. \end{aligned}$$

Поскольку $\mathbf{E}|Z_n|^m < \infty$ при любом $m \in \mathbf{N}$, то по теореме Аносовой [4, с. 309] для любого $\alpha > 0$

$$\mathbf{P}\left(|Z_n| \geq \sqrt{2\alpha \ln n}\right) = \frac{1}{n^\alpha \sqrt{\pi \alpha \ln n}} \left(1 + \mathbf{O}\left(\frac{1}{\ln n}\right)\right).$$

Поэтому в силу неравенства Коши–Буняковского при $j = 1, 2$ и равномерной ограниченности последовательности $\{\mathbf{E}|Z_n^m|\}$ для любого $m \in \mathbf{N}$, а также $\{\mathbf{V}\widehat{G}[a|S_n]\}$ и $\{\mathbf{E}(\widehat{G}[a|S_n] - G[a])^2\}$ начиная с некоторого $n \geq n_0$ верны оценки:

$$\begin{aligned} \mathbf{E}\left|Z_n^m I(|Z_n| > \sqrt{2\alpha \ln n})\right| &\leq \\ &\leq \sqrt{\mathbf{E}|Z_n|^{2m} \mathbf{P}\left(|Z_n| \geq \sqrt{2\alpha \ln n}\right)} = \\ &= \mathbf{o}\left(n^{-\alpha/2}\right); \end{aligned} \quad (13)$$

$$\begin{aligned} \mathbf{E}\left\{\left|\widehat{G}[a|S_n] - G[a]\right| I(|Z_n| > \sqrt{2\alpha \ln n})\right\} &\leq \\ &\leq \sqrt{\mathbf{V}\widehat{G}[a|S_n] \mathbf{P}\left(|Z_n| \geq \sqrt{2\alpha \ln n}\right)} = \\ &= \mathbf{o}\left(n^{-\alpha/2}\right); \end{aligned} \quad (14)$$

$$\begin{aligned} \mathbf{E}\left\{\left|\widetilde{G}[a|S_n] - G[a]\right| I(|Z_n| > \sqrt{2\alpha \ln n})\right\} &\leq \\ &\leq \sqrt{\mathbf{E}\left(\widetilde{G}[a|S_n] - G[a]\right)^2 \mathbf{P}\left(|Z_n| \geq \sqrt{2\alpha \ln n}\right)} = \\ &= \mathbf{o}\left(n^{-\alpha/2}\right). \end{aligned} \quad (15)$$

Чтобы подчеркнуть согласованность изменений переменных s и z , будем в дальнейшем использовать обозначение:

$$z_{s,n} = \frac{s - na}{b\sqrt{n}}.$$

При $n \rightarrow \infty$ в условиях (\mathbf{A}_1) – (\mathbf{A}_3) по теореме А.4.3 из [5] верны соотношения:

$$\sum_{s \in \mathbf{Z}} |z_{s,n}|^j \frac{\varphi[z_{s,n}]}{b\sqrt{n}} = \int_{-\infty}^{\infty} |z|^j \varphi[z] dz + \mathbf{O}\left(n^{-1/2}\right),$$

$$j \in \{0, 1, \dots\}, \quad (16)$$

а по теореме 22.3 из [5]:

$$\sum_{s \in \mathbf{Z}} \left| \mathbf{P}(S_n = s) - \frac{p_3[z_{s,n}, n]}{b\sqrt{n}} \right| = \mathbf{o}\left(n^{-3/2}\right). \quad (17)$$

Из (9), (10) и (16) следует, что

$$\begin{aligned} \sum_{s \in \mathbf{Z}} \left| \frac{p_3[z_{s,n}, n]}{b\sqrt{n}} - \frac{p_2[z_{s,n}, n]}{b\sqrt{n}} \right| &= \\ &= \frac{1}{n^{3/2}} \sum_{s \in \mathbf{Z}} \frac{\varphi[z_{s,n}]}{b\sqrt{n}} \left| \frac{\rho_5 H_5[z_{s,n}]}{5!} + \frac{\rho_3 \rho_4 H_7[z_{s,n}]}{3!4!} + \right. \\ &\quad \left. + \frac{\rho_3^3 H_9[z_{s,n}]}{(3!)^4} \right| = \mathbf{O}\left(n^{-3/2}\right), \end{aligned}$$

а потому с учетом (17) имеем:

$$\sum_{s \in \mathbf{Z}} \left| \mathbf{P}(S_n = s) - \frac{p_{n,2}[z_{s,n}]}{b\sqrt{n}} \right| = \mathbf{O}\left(n^{-3/2}\right). \quad (18)$$

Отсюда при $j = 1, 2$ получим

$$\begin{aligned} \sum_{s \in \mathbf{Z}} I(|z_{s,n}| \leq \sqrt{2\alpha \ln n}) \left| h_j[z_{s,n}, n] (\mathbf{P}(S_n = s) - \right. \\ \left. - \frac{p_{n,2}[z_{s,n}]}{b\sqrt{n}}) \right| &= \mathbf{O}\left(\sqrt{\ln n}\right) \times \\ &\times \sum_{s \in \mathbf{Z}} I(|z_{s,n}| \leq \sqrt{2\alpha \ln n}) \left| \mathbf{P}(S_n = s) - \right. \\ &\quad \left. - \frac{p_{n,2}[z_{s,n}]}{b\sqrt{n}} \right| = \mathbf{O}\left(\frac{\sqrt{\ln n}}{n^{3/2}}\right). \end{aligned} \quad (19)$$

В [1] показано, что на промежутке $[-\sqrt{2\alpha \ln n}; \sqrt{2\alpha \ln n}]$ функции $h_1[z, n]$ и $h_2[z, n]$ при $n \rightarrow \infty, \alpha > 0$ имеют единственные корни, соответственно $z_1 = A/\sqrt{n} + \mathbf{O}(n^{-3/2})$ и $z_2 = 0$. При этом $h_j[z, n] < 0$, если $-\sqrt{2\alpha \ln n} < z < z_j$, и $h_j[z, n] > 0$, если $z_j < z < \sqrt{2\alpha \ln n}, j = 1, 2$.

Используя упомянутые выше свойства функции $h_1[z, n]$, а также (12), (13) и (14), при $\alpha \geq 4$ получим:

$$\begin{aligned} \mathbf{E}\left|\widehat{G}[a; Z_n] - G[a]\right| &= \\ &= \mathbf{E}\left\{\left|\widehat{G}[a; Z_n] - G[a]\right| I(|Z_n| \leq \sqrt{2\alpha \ln n})\right\} + \\ &+ \mathbf{E}\left\{\left|\widehat{G}[a; Z_n] - G[a]\right| I(|Z_n| > \sqrt{2\alpha \ln n})\right\} = \\ &= \frac{|G'[a]|}{\sqrt{n\Phi_1'[a]}} \mathbf{E}\left\{\left|h_1[Z_n, n]\right| + \right. \\ &\quad \left. + \mathbf{O}\left(n^{-3/2+\delta_1}\right)\right\} I(|Z_n| \leq \sqrt{2\alpha \ln n}) + \\ &\quad + \mathbf{o}\left(n^{-\alpha/2}\right) = \\ &= \frac{|G'[a]|}{\sqrt{n\Phi_1'[a]}} \mathbf{E}\left\{\left|h_1[Z_n, n]\right| I(|Z_n| \leq \sqrt{2\alpha \ln n})\right\} + \\ &\quad + \mathbf{O}\left(n^{-2+\delta_1}\right), \end{aligned} \quad (20)$$

где $0 < \delta_1 < 0,5$.

Аналогично можно показать, что при $\alpha \geq 4$, $0 < \delta_2 < 0,5$

$$\begin{aligned} & \mathbf{E} \left| \tilde{G}[a; Z_n] - G[a] \right| = \\ & = \frac{|G'[a]|}{\sqrt{n\Phi_1'[a]}} \mathbf{E} \left\{ |h_2[Z_n, n]| I \left(|Z_n| \leq \sqrt{2\alpha \ln n} \right) \right\} + \\ & \quad + \mathbf{O} \left(n^{-2+\delta_2} \right). \quad (21) \end{aligned}$$

Далее всюду будем предполагать, что $\alpha \geq 4$.

Найдем асимптотические разложения главных членов (20) и (21), воспользовавшись теоремой А.4.3 из [5]. В соответствии с этой теоремой для любой функции Шварца f , определенной на \mathbf{R} , и каждого борелевского множества \mathbf{A} верна оценка:

$$\left| \sum_{s \in \mathbf{S}_A} \frac{f[z_{s,n}]}{b\sqrt{n}} - \int_{\mathbf{A}} d\Lambda[z] \right| = \mathbf{O} \left(n^{-3/2} \right), \quad (22)$$

если положить

$$\begin{aligned} \mathbf{S}_A &= \{s \in \mathbf{Z} : z_{s,n} \in \mathbf{A}\}; \\ F[z] &= \int_{-\infty}^z f[x] dx; \\ \Lambda[z] &= \sum_{i=0}^2 \left(-\frac{1}{b\sqrt{n}} \right)^i \ell_i[bz\sqrt{n} + an] F^{(i)}[z]. \end{aligned}$$

Поскольку функции $f_j[x] = h_j[x, n]p_2[x, n]$, $j = 1, 2$, являются функциями Шварца, то, полагая

$$\begin{aligned} A_{j,1} &= [z_j; \sqrt{2\alpha \ln n}]; \quad A_{j,2} = [-\sqrt{2\alpha \ln n}; z_j]; \\ \mathbf{S}_{j,1} &= \{s \in \mathbf{Z} : z_{s,n} \in A_{j,1}\}; \\ \mathbf{S}_{j,2} &= \{s \in \mathbf{Z} : z_{s,n} \in A_{j,2}\}; \\ F_j[z] &= \int_{-\infty}^z f_j[x] dx = \int_{-\infty}^z h_j[x, n]p_2[x, n] dx; \quad (23) \\ \Lambda_j[z] &= \sum_{i=0}^2 \left(-\frac{1}{b\sqrt{n}} \right)^i \ell_i[bz\sqrt{n} + an] F_j^{(i)}[z] \quad (24) \end{aligned}$$

и применяя (19) и (22), получим при $j = 1, 2$

$$\begin{aligned} & \mathbf{E} \left| h_j[Z_n, n] I \left(|Z_n| \leq \sqrt{2\alpha \ln n} \right) \right| = \\ & = \sum_{s \in \mathbf{Z}} |h_j[z_{s,n}, n]| I \left(|z_{s,n}| \leq \sqrt{2\alpha \ln n} \right) \times \\ & \quad \times \left(\mathbf{P} (S_n = s) - \frac{p_2[z_{s,n}, n]}{b\sqrt{n}} \right) + \\ & + \sum_{s \in \mathbf{Z}} |h_j[z_{s,n}, n]| I \left(|z_{s,n}| \leq \sqrt{2\alpha \ln n} \right) \frac{p_2[z_{s,n}, n]}{b\sqrt{n}} = \end{aligned}$$

$$\begin{aligned} & = \mathbf{O} \left(\frac{\sqrt{\ln n}}{n^{3/2}} \right) + \\ & + \sum_{s \in \mathbf{Z}} |h_j[z_{s,n}, n]| I \left(|z_{s,n}| \leq \sqrt{2\alpha \ln n} \right) \frac{p_2[z_{s,n}, n]}{b\sqrt{n}} = \\ & = \sum_{s \in \mathbf{S}_{j,1}} h_j[z_{s,n}, n] \frac{p_2[z_{s,n}, n]}{b\sqrt{n}} - \\ & - \sum_{s \in \mathbf{S}_{j,2}} h_j[z_{s,n}, n] \frac{p_2[z_{s,n}, n]}{b\sqrt{n}} + \mathbf{O} \left(\frac{\sqrt{\ln n}}{n^{3/2}} \right) = \\ & = \int_{A_{j,1}} d\Lambda_j[z] - \int_{A_{j,2}} d\Lambda_j[z] + \mathbf{O} \left(\frac{\sqrt{\ln n}}{n^{3/2}} \right). \quad (25) \end{aligned}$$

Чтобы оценить значения интегралов $\int_{A_{j,1}} d\Lambda_j[z]$ и $\int_{A_{j,2}} d\Lambda_j[z]$, представим подынтегральные функции в (23) следующим образом:

$$\begin{aligned} h_1[x, n]p_2[x, n] &= \varphi[x] \left\{ x + \frac{\rho_3 x H_3[x]}{6\sqrt{n}} + \right. \\ & \quad + \frac{\rho_4 x H_4[x]}{24n} + \frac{\rho_3^2 x H_6[x]}{72n} + \frac{A H_2[x]}{\sqrt{n}} + \\ & \quad \left. + \frac{A \rho_3 H_2[x] H_3[x]}{6n} + \frac{B_1 H_3[x] + B_2 x}{n} + P_1[x, n] \right\} = \\ & = \varphi[x] \left\{ x + \frac{\rho_3 (H_4[x] + 3H_2[x])/6 + A H_2[x]}{\sqrt{n}} + \right. \\ & \quad + \frac{\rho_4 (H_5[x] + 4H_3[x])}{24n} + \frac{\rho_3^2 (H_7[x] + 6H_5[x])}{72n} + \\ & \quad + \frac{A \rho_3 (H_5[x] + 6H_3[x] + 6x)}{6n} + \frac{B_1 H_3[x] + B_2 x}{n} + \\ & \quad \left. + P_1[x, n] \right\}; \\ h_2[x, n]p_2[x, n] &= \varphi[x] \left\{ x + \frac{\rho_3 x H_3[x]}{6\sqrt{n}} + \right. \\ & \quad + \frac{\rho_4 x H_4[x]}{24n} + \frac{\rho_3^2 x H_6[x]}{72n} + \frac{A x^2}{\sqrt{n}} + \frac{A \rho_3 x^2 H_3[x]}{6n} + \\ & \quad + \frac{B x^3}{n} + P_2[x, n] \left. \right\} = G'[a] \varphi[x] \left\{ x + \right. \\ & \quad + \frac{\rho_3 (H_4[x] + 3H_2[x])/6 + A (H_2[x] + 1)}{\sqrt{n}} + \\ & \quad + \frac{\rho_4 (H_5[x] + 4H_3[x])}{24n} + \frac{\rho_3^2 (H_7[x] + 6H_5[x])}{72n} + \\ & \quad + \frac{A \rho_3 (H_5[x] + 7H_3[x] + 6x)}{6n} + \frac{B (H_3[x] + 3x)}{n} + \\ & \quad \left. + P_2[x, n] \right\}, \end{aligned}$$

где

$$P_1[x, n] = \frac{AH_2[x]}{n^{3/2}} \left(\frac{\rho_4 H_4[x]}{24} + \frac{\rho_3^2 H_6[x]}{72} \right) + \frac{B_1 H_3[x] + B_2 x}{n^{3/2}} \left(\frac{\rho_3 H_3[x]}{6} + \frac{\rho_4 H_4[x]}{24} + \frac{\rho_3^2 H_6[x]}{72} \right);$$

$$P_2[x, n] = \frac{Ax^2}{n^{3/2}} \left(\frac{\rho_4 H_4[x]}{24} + \frac{\rho_3^2 H_6[x]}{72} \right) + \frac{Bx^3}{n^{3/2}} \left(\frac{\rho_3 H_3[x]}{6} + \frac{\rho_4 H_4[x]}{24} + \frac{\rho_3^2 H_6[x]}{72} \right).$$

Учитывая, что

$$\varphi[\sqrt{2\alpha \ln n}] = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{2\alpha \ln n}{2} \right\} = \mathbf{O}(n^{-\alpha});$$

$$\varphi^{(r)}[x] = (-1)^r H_r[x] \varphi[x] \text{ при } r = 0, 1, 2, \dots,$$

получим сначала асимптотические разложения функций $F_j[z]$, $j = 1, 2$, и их производных, справедливые при значениях аргумента $|z| \leq \sqrt{2\alpha \ln n}$:

$$F_1[z] = -\varphi[z] \left\{ 1 + \frac{\rho_3(H_3[z] + 3z)/6 + Az}{\sqrt{n}} + \frac{\rho_4(H_4[z] + 4H_2[z])}{24n} + \frac{\rho_3^2(H_6[z] + 6H_4[z])}{72n} + \frac{A\rho_3(H_4[z] + 6H_2[z] + 6)}{6n} + \frac{B_1 H_2[z] + B_2}{n} \right\} + \mathbf{O}(n^{-3/2}); \quad (26)$$

$$F_1'[z] = \varphi[z] \left\{ z + \frac{\rho_3(H_4[z] + 3H_2[z])/6 + AH_2[z]}{\sqrt{n}} \right\} + \mathbf{O}(n^{-1}); \quad (27)$$

$$F_1''[z] = -\varphi[z]H_2[z] + \mathbf{O}(n^{-1/2}); \quad (28)$$

$$F_1^{(3)}[z] = \varphi[z]H_3[z] + \mathbf{O}(n^{-1/2}); \quad (29)$$

$$F_2[z] = \frac{A\Phi[z]}{\sqrt{n}} - \varphi[z] \left\{ 1 + \frac{\rho_3(H_3[z] + 3z)/6 + Az}{\sqrt{n}} + \frac{\rho_4(H_4[z] + 4H_2[z])}{24n} + \frac{\rho_3^2(H_6[z] + 6H_4[z])}{72n} + \frac{A\rho_3(H_4[z] + 7H_2[z] + 6)}{6n} + \frac{B(H_2[z] + 3)}{n} \right\} + \mathbf{O}(n^{-3/2}),$$

$$F_2'[z] = \varphi[z] \left\{ z + \frac{\rho_3(H_4[z] + 3H_2[z])/6 + A(H_2[z] + 1)}{\sqrt{n}} \right\} + \mathbf{O}(n^{-1});$$

$$F_2''[z] = -\varphi[z]H_2[z] + \mathbf{O}(n^{-1/2});$$

$$F_2^{(3)}[z] = \varphi[z]H_3[z] + \mathbf{O}(n^{-1/2}),$$

а затем асимптотические разложения этих же функций в точках $z = \pm\sqrt{2\alpha \ln n}$ и $z = 0$:

$$F_1[\pm\sqrt{2\alpha \ln n}] = \mathbf{O}(n^{-\alpha}) + \mathbf{O}(n^{-3/2}) = \mathbf{O}(n^{-3/2}); \quad (30)$$

$$F_2[\sqrt{2\alpha \ln n}] = \frac{A\Phi[\sqrt{2\alpha \ln n}]}{\sqrt{n}} + \mathbf{O}(n^{-\alpha}) + \mathbf{O}(n^{-3/2}) = \frac{A}{\sqrt{n}} - \frac{A\varphi[\sqrt{2\alpha \ln n}]}{\sqrt{2n\alpha \ln n}} + \mathbf{O}(n^{-3/2}) = \frac{A}{\sqrt{n}} + \mathbf{O}(n^{-3/2}); \quad (31)$$

$$F_2[-\sqrt{2\alpha \ln n}] = \frac{A\Phi[-\sqrt{2\alpha \ln n}]}{\sqrt{n}} + \mathbf{O}(n^{-\alpha}) + \mathbf{O}(n^{-3/2}) = \frac{A(1 - \Phi[\sqrt{2\alpha \ln n}])}{\sqrt{n}} + \mathbf{O}(n^{-\alpha}) + \mathbf{O}(n^{-3/2}) = \mathbf{O}(n^{-3/2}); \quad (32)$$

$$F_j'[\pm\sqrt{2\alpha \ln n}] = \mathbf{O}(n^{-\alpha}) + \mathbf{O}(n^{-1}) = \mathbf{O}(n^{-1}), \quad j = 1, 2; \quad (33)$$

$$F_j''[\pm\sqrt{2\alpha \ln n}] = \mathbf{O}(n^{-\alpha}) + \mathbf{O}(n^{-1/2}) = \mathbf{O}(n^{-1/2}), \quad j = 1, 2. \quad (34)$$

Используя равенства (26)–(29), оценим значения функции $F_1[z]$ и ее производных в точке $z = 0$ и точке $z = \zeta$, находящейся в интервале с концами 0 и z_1 :

$$F_1[0] = -\varphi[0] \left\{ 1 + \frac{B_2 - B_1 + A\rho_3/2}{n} + \frac{\rho_3^2 - \rho_4}{24n} \right\} + \mathbf{O}(n^{-3/2}); \quad (35)$$

$$F_1'[0] = -\varphi[0] \frac{A}{\sqrt{n}} + \mathbf{O}(n^{-1}); \quad (36)$$

$$F_1^{(2)}[0] = \varphi[0] + \mathbf{O}\left(\frac{1}{\sqrt{n}}\right); \quad (37)$$

$$F_1^{(3)}[\zeta] = \mathbf{O}\left(n^{-1/2}\right). \quad (38)$$

Используя равенства (35)–(38), найдем разложения функции $F_1[z]$ и ее производных в точке $z = z_1$:

$$F_1^{(2)}[z_1] = F_1^{(2)}[0] + F_1^{(3)}[\zeta_1]z_1 = \varphi[0] + \mathbf{O}\left(\frac{1}{\sqrt{n}}\right); \quad (39)$$

$$\begin{aligned} F_1'[z_1] &= F_1'[0] + F_1^{(2)}[0]z_1 + F_1^{(3)}[\zeta_2]\frac{z_1^2}{2} = \\ &= -\varphi[0]\frac{A}{\sqrt{n}} + \mathbf{O}(n^{-1}) + \left(\varphi[0] + \mathbf{O}(n^{-1/2})\right) \times \\ &\quad \times \left(\frac{A}{n} + \mathbf{O}(n^{-3/2})\right) = \mathbf{O}(n^{-1}); \quad (40) \end{aligned}$$

$$\begin{aligned} F_1[z_1] &= F_1[0] + F_1'[0]z_1 + F_1^{(2)}[0]\frac{z_1^2}{2} + F_1^{(3)}[\zeta_3]\frac{z_1^3}{6} = \\ &= -\varphi[0]\left\{1 + \frac{B_2 - B_1}{n} + \frac{A\rho_3}{2n} + \frac{\rho_3^2 - \rho_4}{24n}\right\} + \\ &\quad + \mathbf{O}(n^{-3/2}) + \\ &+ \left(-\varphi[0]\frac{A}{\sqrt{n}} + \mathbf{O}(n^{-1})\right)\left(\frac{A}{\sqrt{n}} + \mathbf{O}(n^{-3/2})\right) + \\ &+ \left(\varphi[0] + \mathbf{O}(n^{-1/2})\right)\frac{1}{2}\left(\frac{A}{\sqrt{n}} + \mathbf{O}(n^{-3/2})\right)^2 + \\ &\quad + \mathbf{O}(n^{-2}) = \\ &= -\varphi[0]\left\{1 + \frac{B_2 - B_1}{n} + \frac{A\rho_3 + A^2}{2n} + \frac{\rho_3^2 - \rho_4}{24n}\right\} + \\ &\quad + \mathbf{O}(n^{-3/2}), \quad (41) \end{aligned}$$

где ζ_1, ζ_2 и ζ_3 — значения из интервала с концами 0 и z_1 .

Теперь оценим значения функции $F_2[z]$ и ее производных в точке $z = 0$:

$$\begin{aligned} F_2[0] &= \frac{A}{2\sqrt{n}} - \\ &- \varphi[0]\left\{1 + \frac{1}{n}\left(2B + \frac{A\rho_3}{3} + \frac{\rho_3^2}{24} - \frac{\rho_4}{24}\right)\right\} + \\ &\quad + \mathbf{O}(n^{-3/2}); \quad (42) \end{aligned}$$

$$F_2'[0] = \mathbf{O}(n^{-1}); \quad (43)$$

$$F_2''[0] = \varphi[0] + \mathbf{O}\left(\frac{1}{\sqrt{n}}\right). \quad (44)$$

Используя оценки (30), (33), (34), (39)–(41), а также (20), (23)–(25), найдем асимптотическое разложение средней абсолютной ошибки НОРМД:

$$\begin{aligned} \mathbf{E}\left|\widehat{G}[a; Z_n] - G[a]\right| &= \frac{|G'[a]|}{\sqrt{n\Phi_1'[a]}} \times \\ &\times \sum_{i=0}^2 \left(-\frac{1}{b\sqrt{n}}\right)^i \ell_i[na + b\sqrt{2n\alpha \ln n}] F_1^{(i)}[\sqrt{2\alpha \ln n}] - \\ &- \frac{2|G'[a]|}{\sqrt{n\Phi_1'[a]}} \sum_{i=0}^2 \left(-\frac{1}{b\sqrt{n}}\right)^i \ell_i[na + z_1 b\sqrt{n}] F_1^{(i)}[z_1] + \\ &+ \frac{|G'[a]|}{\sqrt{n\Phi_1'[a]}} \sum_{i=0}^2 \left(-\frac{1}{b\sqrt{n}}\right)^i \ell_i[na - \\ &- b\sqrt{2n\alpha \ln n}] F_1^{(i)}[-\sqrt{2\alpha \ln n}] + \mathbf{O}(n^{-2+\delta_1}) = \\ &= \frac{2\varphi[0]|G'[a]|}{\sqrt{n\Phi_1'[a]}} \left\{1 + \frac{A^2 + A\rho_3}{2n} + \frac{B_2 - B_1}{n} + \right. \\ &\quad \left. + \frac{\rho_3^2 - \rho_4}{24n} - \frac{\ell_2[Ab + na]}{b^2 n}\right\} + \mathbf{O}(n^{-2+\delta_1}). \quad (45) \end{aligned}$$

Проделав аналогичные выкладки, но с использованием формул (31)–(34), (42)–(44), а также (21), (23)–(25), найдем асимптотическое разложение средней абсолютной ошибки ОМП:

$$\begin{aligned} \mathbf{E}\left|\widetilde{G}[a; Z_n] - G[a]\right| &= \\ &= \frac{|G'[a]|}{\sqrt{n\Phi_1'[a]}} \sum_{i=0}^2 \left(-\frac{1}{b\sqrt{n}}\right)^i \ell_i[na + \\ &+ b\sqrt{2n\alpha \ln n}] F_2^{(i)}[\sqrt{2\alpha \ln n}] - \\ &- \frac{2|G'[a]|}{\sqrt{n\Phi_1'[a]}} \sum_{i=0}^2 \left(-\frac{1}{b\sqrt{n}}\right)^i \ell_i[na] F_2^{(i)}[0] + \\ &+ \frac{|G'[a]|}{\sqrt{n\Phi_1'[a]}} \sum_{i=0}^2 \left(-\frac{1}{b\sqrt{n}}\right)^i \ell_i[na - \\ &- b\sqrt{2n\alpha \ln n}] F_2^{(i)}[-\sqrt{2\alpha \ln n}] + \mathbf{O}(n^{-2+\delta_2}) = \\ &= \frac{2\varphi[0]|G'[a]|}{\sqrt{n\Phi_1'[a]}} \left\{1 + \frac{2B}{n} + \frac{A\rho_3}{3n} + \frac{\rho_3^2 - \rho_4}{24n} - \right. \\ &\quad \left. - \frac{\ell_2[na]}{b^2 n}\right\} + \mathbf{O}(n^{-2+\delta_2}). \quad (46) \end{aligned}$$

Подставляя в (45) и (46) выражения коэффициентов асимметрии и эксцесса (11), завершим вывод асимптотических разложений (3) и (4).

Рассмотрим теперь случай $G'[a] = 0$, $G''[a] \neq 0$. В этом случае можно ограничиться более компактным вариантом разложений (7) и (8):

$$\widehat{G}[a; z] - G[a] = \frac{G''[a] H_2[z]}{2n\Phi_1'[a]} + \mathbf{O}\left(n^{-3/2+\delta_3}\right); \quad (47)$$

$$\widetilde{G}[a; z] - G[a] = \frac{G''[a] z^2}{2n\Phi_1'[a]} + \mathbf{O}\left(n^{-3/2+\delta_4}\right), \quad (48)$$

где $0 < \delta_3 < 0,5$, $0 < \delta_4 < 0,5$, $|z| \leq \sqrt{2\alpha \ln n}$.

Из (47), используя (14), (18) и теорему А.4.3 из [5] при $n \rightarrow \infty$, последовательно получим:

$$\begin{aligned} \mathbf{E} \left| \widehat{G}[a; Z_n] - G[a] \right| &= \\ &= \frac{|G''[a]|}{2n\Phi_1'[a]} \mathbf{E} \left| H_2[Z_n] I\left(|Z_n| \leq \sqrt{2\alpha \ln n}\right) \right| + \\ &\quad + \mathbf{O}\left(n^{-3/2+\delta_3}\right) = \frac{|G''[a]|}{2n\Phi_1'[a]} \times \\ &\quad \times \sum_{s \in \mathbf{Z}} |H_2[z_{s,n}]| I\left(|z_{s,n}| \leq \sqrt{2\alpha \ln n}\right) \frac{\varphi[z_{s,n}]}{b\sqrt{n}} + \\ &\quad + \mathbf{O}\left(n^{-3/2+\delta_3}\right) = \\ &= \frac{|G''[a]|}{2n\Phi_1'[a]} \int_{-\sqrt{2\alpha \ln n}}^{\sqrt{2\alpha \ln n}} |H_2[z]| \varphi[z] dz + \mathbf{O}\left(n^{-3/2+\delta_3}\right) = \\ &= \frac{|G''[a]|}{n\Phi_1'[a]} \sqrt{\frac{2}{\pi e}} + \mathbf{O}\left(n^{-3/2+\delta_3}\right), \end{aligned}$$

что завершает доказательство (5).

Справедливость (6) установим, используя (48), (15), (18) и теорему А.4.3 из [5] при $n \rightarrow \infty$:

$$\begin{aligned} \mathbf{E} \left| \widetilde{G}[a; Z_n] - G[a] \right| &= \\ &= \frac{|G''[a]|}{2n\Phi_1'[a]} \mathbf{E} \left\{ Z_n^2 I\left(|Z_n| \leq \sqrt{2\alpha \ln n}\right) \right\} + \\ &\quad + \mathbf{O}\left(n^{-3/2+\delta_4}\right) = \frac{|G''[a]|}{2n\Phi_1'[a]} \times \\ &\quad \times \sum_{s \in \mathbf{Z}} z_{s,n}^2 I\left(|z_{s,n}| \leq \sqrt{2\alpha \ln n}\right) \frac{\varphi[z_{s,n}]}{b\sqrt{n}} + \\ &\quad + \mathbf{O}\left(n^{-3/2+\delta_4}\right) = \\ &= \frac{|G''[a]|}{2n\Phi_1'[a]} \int_{-\sqrt{2\alpha \ln n}}^{\sqrt{2\alpha \ln n}} z^2 \varphi[z] dz + \mathbf{O}\left(n^{-3/2+\delta_4}\right) = \\ &= \frac{|G''[a]|}{2n\Phi_1'[a]} + \mathbf{O}\left(n^{-3/2+\delta_4}\right). \end{aligned}$$

Теорема доказана.

4 Экспериментальная оценка точности асимптотических разложений средней абсолютной ошибки

С целью выяснения возможности выбора лучшей из оценок с помощью асимптотических разложений (3) и (4) была проведена серия вычислений в предположении, что наблюдаемая случайная величина ξ имеет пуассоновское распределение, определяемое выражением:

$$\mathbf{P}(\xi = x) = \frac{a^x}{x!} e^{-a}, \quad x = 0, 1, 2, \dots,$$

с минимальной достаточной статистикой $S_n = \sum_{i=1}^n X_i$, также имеющей пуассоновское распределение:

$$\mathbf{P}(S_n = s) = \frac{(na)^s}{s!} e^{-na}, \quad s = 0, 1, 2, \dots$$

Для двух параметрических функций $G_1[a] = a^4$ и $G_2[a] = e^{-a}$ приближенное значение выражений

$$\begin{aligned} \Delta_j[a] &= \mathbf{E} \left| \widehat{G}_j[a|S_n] - G_j[a] \right| - \\ &\quad - \mathbf{E} \left| \widetilde{G}_j[a|S_n] - G_j[a] \right|, \quad j = 1, 2, \end{aligned}$$

вычислялось с помощью разложений (3) и (4) по формулам:

$$\Delta_1^*[a] = \frac{4a^{5/2}}{n^{3/2}\sqrt{2\pi}} \left(-\frac{25}{4} + \left[na + \frac{3}{2} \right] - \left[na + \frac{3}{2} \right]^2 - [na] + [na]^2 \right);$$

$$\Delta_2^*[a] = \frac{e^{-a}}{n^{3/2}\sqrt{2\pi a}} \left(\frac{10a - 9a^2}{12} + \left[na - \frac{a}{2} \right] - \left[na - \frac{a}{2} \right]^2 - [na] + [na]^2 \right),$$

а точное значение — по формулам:

$$\begin{aligned} \Delta_1[a] &= \\ &= \sum_{s=0}^{\infty} \left| \frac{s(s-1)(s-2)(s-3)}{n^4} - a^4 \right| \mathbf{P}(S_n = s) - \\ &\quad - \sum_{s=0}^{\infty} \left| \left(\frac{s}{n} \right)^4 - a^4 \right| \mathbf{P}(S_n = s) = \\ &= 2a^4 \sum_{s=\lceil s_1 \rceil - 3}^{\lceil s_1 \rceil} \mathbf{P}(S_n = s) - 2a^4 \sum_{s=\lceil na \rceil - 3}^{\lceil na \rceil} \mathbf{P}(S_n = s) + \end{aligned}$$

Таблица 1 Значения (в %) относительной погрешности вычисления (49) в зависимости от значений параметров n, a

a	n					
	10	20	50	100	200	500
0,12	30,4	9,50	4,87	2,44	1,221	0,488
0,24	9,50	5,08	2,44	1,22	0,610	0,244
0,49	5,47	2,42	0,47	0,60	0,299	0,119
0,79	3,37	1,48	0,28	0,37	0,185	0,074
1,00	2,93	1,46	0,59	0,29	0,146	0,059
2,29	1,15	0,50	0,09	0,13	0,064	0,026
4,93	0,63	0,15	0,04	0,06	0,030	0,012
10,0	0,29	0,15	0,06	0,03	0,015	0,006
20,0	0,15	0,07	0,03	0,01	0,007	0,003

Таблица 2 Значения (в %) относительной погрешности вычисления (50) в зависимости от значений параметров n, a

a	n					
	10	20	50	100	200	500
0,12	3,87	2,73	0,089	0,04	0,02	0,009
0,24	4,72	1,97	0,51	0,25	0,13	0,050
0,49	5,21	2,75	1,31	0,49	0,24	0,098
0,79	8,01	4,60	8,07	0,77	0,38	0,154
1,00	10,8	5,62	2,30	1,16	0,58	0,234
2,29	2,46	1,06	0,29	0,23	0,12	0,046
4,93	5,19	2,86	1,16	0,57	0,29	0,115
10,0	14,1	7,29	2,97	1,49	0,75	0,300
20,0	29,6	15,6	6,43	3,24	1,63	0,653

$$\begin{aligned}
 & + \frac{6a^3}{n} \left(2 \sum_{s=0}^{\lceil na \rceil - 3} \mathbf{P}(S_n = s) - 1 \right) + \\
 & + \frac{7a^2}{n^2} \left(2 \sum_{s=0}^{\lceil na \rceil - 2} \mathbf{P}(S_n = s) - 1 \right) + \\
 & + \frac{a}{n^3} \left(2 \sum_{s=0}^{\lceil na \rceil - 1} \mathbf{P}(S_n = s) - 1 \right), \quad (49)
 \end{aligned}$$

$$\begin{aligned}
 & = 2 \sum_{s=0}^{\lceil s_2 \rceil} \left\{ \left(\frac{n-1}{n} \right)^s - e^{-a} \right\} \mathbf{P}(S_n = s) - \\
 & - 2 \sum_{s=0}^{\lceil na \rceil} \left\{ e^{-s/n} - e^{-a} \right\} \mathbf{P}(S_n = s) + \\
 & + \exp \left\{ -na \left(1 - e^{-1/n} \right) \right\} - e^{-a}, \quad (50)
 \end{aligned}$$

где $s_2 = -a/(\ln[1 - 1/n])$.

где $\lceil x \rceil$ — целая часть числа x , а s_1 — вещественное решение уравнения

$$s(s-1)(s-2)(s-3) = (na)^4$$

при $s > 3$ и

$$\begin{aligned}
 \Delta_2[a] & = \sum_{s=0}^{\infty} \left| \left(\frac{n-1}{n} \right)^s - e^{-a} \right| \mathbf{P}(S_n = s) - \\
 & - \sum_{s=0}^{\infty} \left| e^{-s/n} - e^{-a} \right| \mathbf{P}(S_n = s) =
 \end{aligned}$$

Вычисления проводились при следующих значениях параметров: $n = 10, 20, 50, 100, 200$ и 500 ; $a = 0,12, 0,24, 0,49, 0,79, 1, 2,29, 4,93, 10$ и 20 . Результаты вычислений приведены в табл. 1 и 2.

Из табл. 1 и 2 видно, что относительная погрешность вычисления выражений $\Delta_j[a]$, $j = 1, 2$, с помощью асимптотических разложений (3) и (4) при всех значениях входных параметров не превосходит 31%, причем с ростом объема выборки убывает со скоростью $1/n$.

При увеличении значения оцениваемой параметрической функции также наблюдается тенденция к уменьшению относительной погрешности.

Таким образом, применение асимптотических разложений (3) и (4) позволяет получить ответ на вопрос о том, какая из двух оценок имеет меньшее значение средней абсолютной ошибки.

Литература

1. Федосеева Н. П., Чичагов В. В. Сравнение UMVUE и MLE с помощью абсолютной функции потерь в случае однопараметрического экспоненциального семейства непрерывных распределений // Статистические методы оценивания и проверки гипотез. — Пермь: ПГНИУ, 2012. С. 96–109.
2. Hwang T.-Y., Hu C.-Y. More comparisons of MLE with UMVUE for exponential families // Ann. Inst. Statist. Math., 1990. Vol. 42. P. 65–75.
3. Чичагов В. В. Асимптотические разложения высокого порядка для несмещенных оценок и их дисперсий в модели однопараметрического экспоненциального семейства // Информатика и её применения, 2015. Т. 9. Вып. 3. С. 75–87.
4. Петров В. В. Суммы независимых случайных величин. — М.: Наука, 1972. 416 с.
5. Бхаттачария Р. Н., Рао Р. Р. Аппроксимация нормальным распределением и асимптотические разложения / Пер. с англ. — М.: Наука, 1982. 286 с. (Bhattacharya R. N., Rao R. R. Normal approximation and asymptotic expansions. New York, NY, USA: Wiley, 1976. 316 p.)

Поступила в редакцию 22.06.16

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

V. V. Chichagov

Perm State University, 15 Bukireva Str., Perm 614990, Russian Federation

Abstract: The paper considers a model of duplicate sampling with the fixed size n from a lattice distribution belonging to the natural one-parameter exponential family. Asymptotic expansions of the mean absolute errors of the uniformly minimum variance unbiased estimator (UMVUE) and the maximum likelihood estimator (MLE) of the given parametric function are obtained in the case of infinite size of the sample. The case when $G'[a] = 0$ and $G''[a] \neq 0$ was studied separately. The relative error in calculating the difference in the mean absolute error UMVUE and MLE was evaluated in the case of the Poisson distribution for the two parametric functions. This error was received via the asymptotic expansions. It was found that the asymptotic results with a sufficiently large sample size allows one to compare UMVUE and MLE using such indicator of quality assessment as the mean absolute error.

Keywords: exponential family; lattice distribution; unbiased estimator; maximum likelihood estimator; asymptotic expansion

DOI: 10.14357/19922264160309

Acknowledgments

The research was financially supported by the Russian Ministry of Education and Science (project No. 2096).

References

1. Fedoseeva, N. P., and V. V. Chichagov. 2012. Sravnenie UMVUE i MLE s pomoshch'yu absolyutnoy funktsii poter' v sluchae odnoparametricheskogo eksponentsial'nogo semeystva nepreryvnykh raspredeleniy [Comparison of UMVUE and MLE by the absolute loss function for the one-parameter exponential family of continuous distributions]. *Statisticheskie metody otsenivaniya i proverki gipotez* [Statistical methods of estimation and testing hypotheses]. Perm': Perm State University Publishing House. 96–109.

2. Hwang, T.-Y., and C.-Y. Hu. 1990. More comparisons of MLE with UMVUE for exponential families. *Ann. Inst. Statist. Math.* 42:65–75.
3. Chichagov, V. V. 2015. Asimptoticheskie razlozheniya vysokogo poryadka dlya nesmeshchennykh otsenok i ikh dispersiy v modeli odnparametricheskogo eksponentsial'nogo semeystva [Higher-order asymptotic expansions of unbiased estimators and their variances on the one-parameter exponential family model]. *Informatika i ee Primeneniya — Inform. Appl.* 9(3):75–87.
4. Petrov, V. V., and A. A. Brown. 1975. *Sums of independent random variables*. Berlin: Springer-Verlag. 182 p.
5. Bhattacharya, R. N., and R. R. Rao. 1976. *Normal approximation and asymptotic expansions*. New York, NY: Wiley. 316 p.

Received June 22, 2016

Contributor

Chichagov Vladimir V. (b. 1955) — Candidate of Science (PhD) in physics and mathematics, associate professor, Perm State University, 15 Bukireva Str., Perm 614990, Russian Federation; chichagov@psu.ru

ЗАВИСИМЫЕ ОТ КОЭФФИЦИЕНТА БАЛАНСА ХАРАКТЕРИСТИКИ В БАЙЕСОВСКИХ МОДЕЛЯХ С КОМПАКТНЫМ НОСИТЕЛЕМ АПРИОРНЫХ РАСПРЕДЕЛЕНИЙ*

А. А. Кудрявцев¹

Аннотация: Приведены вероятностные распределения некоторых характеристик, зависящих от коэффициента баланса, т. е. от отношения параметра, «препятствующего» функционированию системы, к параметру, «способствующему» функционированию. В теории массового обслуживания для модели $M|M|1$ такие характеристики интерпретируются как среднее число заявок в системе, коэффициент готовности, вероятность «непотери» вызова, а для дискретных экспоненциальных моделей роста надежности — как предельная надежность системы. В рамках байесовского подхода предполагается, что исходные параметры случайны и имеют априорные распределения с компактным носителем.

Ключевые слова: байесовский подход; системы массового обслуживания; надежность; смешанные распределения; распределения с компактным носителем

DOI: 10.14357/19922264160310

1 Введение

В работе [1] был предложен подход, позволяющий частично унифицировать терминологию, используемую при постановке задач, относящихся к байесовским моделям теории массового обслуживания и надежности. Суть этого подхода заключается в определении параметров, «препятствующих» и «способствующих» функционированию некоторой сложной системы, и рассмотрению их частного (называемого коэффициентом баланса), от величины которого зависит эффективность работы системы. В классических задачах обслуживания и надежности коэффициент баланса входит в большинство основных формул, характеризующих функционирование системы.

Обозначим один из параметров, «препятствующих»/«способствующих» функционированию, через λ , а второй — через μ . Через $\rho = \lambda/\mu$ обозначим коэффициент баланса. В рамках байесовского подхода предполагается, что λ и μ — случайные величины с некоторыми известными априорными распределениями. Ниже будут рассмотрены вероятностные характеристики случайных величин $\pi = 1/(1 + \rho)$ и $N = \rho/(1 - \rho)$.

В задачах массового обслуживания для моделей $M|M|1$ величину ρ , равную отношению параметра входящего потока к параметру обслуживания, принято называть коэффициентом загрузки, вели-

чину π — коэффициентом готовности и вероятностью «непотери» вызова, величину N — средним числом заявок в системе. В дискретных экспоненциальных моделях роста надежности ρ имеет смысл отношения параметра «эффективности» к параметру «дефективности», а под величиной π понимают предельную надежность системы.

Далее приводятся результаты для распределения величин π и N в случае, когда носителями распределений λ и μ являются отрезки на положительной полупрямой. При применении изложенных ниже результатов к надежностным постановкам необходимо ограничивать правые концы носителей распределений единицей [2].

2 Основные результаты

Пусть λ и μ — независимые абсолютно непрерывные случайные величины, причем $P(\lambda \in [a_\lambda, b_\lambda]) = 1$, $0 < a_\lambda < b_\lambda$, и не существует множества $S \subset [a_\lambda, b_\lambda]$ положительной меры Лебега тако- го, что $P(\lambda \in S) = 0$, а для случайной величины μ выполнены аналогичные требования с параметрами a_μ и b_μ . Во всех последующих выкладках будем предполагать, что $x > 0$.

В статье [1] было сформулировано следующее утверждение.

* Исследование выполнено при поддержке Российского научного фонда (проект 14-11-00397).

¹ Московский государственный университет им. М. В. Ломоносова, факультет вычислительной математики и кибернетики; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, nubigena@mail.ru

Теорема 1. Пусть независимые абсолютно непрерывные случайные величины λ и μ имеют соответственно носители распределений $[a_\lambda, b_\lambda]$ и $[a_\mu, b_\mu]$, $0 < a_\lambda < b_\lambda$, $0 < a_\mu < b_\mu$, и плотности $f_\lambda(x)$ и $f_\mu(x)$. Тогда случайная величина $\rho = \lambda/\mu$ имеет функцию распределения

$$\begin{aligned}
 F_\rho(x) = & \mathbf{1} \left(\frac{a_\lambda}{b_\mu} < x \leq \min \left\{ \frac{a_\lambda}{a_\mu}, \frac{b_\lambda}{b_\mu} \right\} \right) \times \\
 & \times \int_{a_\lambda/x}^{b_\mu} \int_{a_\lambda}^{xy} f_\lambda(u) f_\mu(y) \, dudy + \\
 & + \mathbf{1} \left(\frac{a_\lambda}{a_\mu} < x \leq \frac{b_\lambda}{b_\mu} \right) \int_{a_\mu}^{b_\mu} \int_{a_\lambda}^{xy} f_\lambda(u) f_\mu(y) \, dudy + \\
 & + \mathbf{1} \left(\frac{b_\lambda}{b_\mu} < x \leq \frac{a_\lambda}{a_\mu} \right) \times \\
 & \times \left[\int_{a_\lambda/x}^{b_\lambda/x} \int_{a_\lambda}^{xy} f_\lambda(u) f_\mu(y) \, dudy + \int_{b_\lambda/x}^{b_\mu} f_\mu(y) \, dy \right] + \\
 & + \mathbf{1} \left(\max \left\{ \frac{a_\lambda}{a_\mu}, \frac{b_\lambda}{b_\mu} \right\} < x \leq \frac{b_\lambda}{a_\mu} \right) \times \\
 & \times \left[\int_{a_\mu}^{b_\lambda/x} \int_{a_\lambda}^{xy} f_\lambda(u) f_\mu(y) \, dudy + \int_{b_\lambda/x}^{b_\mu} f_\mu(y) \, dy \right] + \\
 & + \mathbf{1} \left(x > \frac{b_\lambda}{a_\mu} \right). \quad (1)
 \end{aligned}$$

Основываясь на теореме 1 и соотношении

$$F_\pi(x) = 1 - F_\rho \left(\frac{1-x}{x} \right),$$

несложно убедиться в справедливости следующего утверждения.

Следствие 1. Пусть коэффициент баланса $\rho = \lambda/\mu$ имеет функцию распределения (1). Тогда случайная величина $\pi = 1/(1 + \rho)$ имеет функцию распределения

$$\begin{aligned}
 F_\pi(x) = & 1 - \mathbf{1} \left(x < \frac{a_\mu}{b_\lambda + a_\mu} \right) - \\
 & - \mathbf{1} \left(\frac{a_\mu}{b_\lambda + a_\mu} \leq x < \min \left\{ \frac{a_\mu}{a_\lambda + a_\mu}, \frac{b_\mu}{b_\lambda + b_\mu} \right\} \right) \times \\
 & \times \left[\int_{a_\mu}^{b_\lambda x/(1-x)} \int_{a_\lambda}^{(1-x)y/x} f_\lambda(u) f_\mu(y) \, dudy + \int_{b_\lambda x/(1-x)}^{b_\mu} f_\mu(y) \, dy \right] - \\
 & - \mathbf{1} \left(\frac{a_\mu}{a_\lambda + a_\mu} \leq x < \frac{b_\mu}{b_\lambda + b_\mu} \right) \times
 \end{aligned}$$

$$\begin{aligned}
 & \times \left[\int_{a_\lambda x/(1-x)}^{b_\lambda x/(1-x)} \int_{a_\lambda}^{(1-x)y/x} f_\lambda(u) f_\mu(y) \, dudy + \int_{b_\lambda x/(1-x)}^{b_\mu} f_\mu(y) \, dy \right] - \\
 & - \mathbf{1} \left(\frac{b_\mu}{b_\lambda + b_\mu} \leq x < \frac{a_\mu}{a_\lambda + a_\mu} \right) \times \\
 & \times \int_{a_\mu}^{b_\mu} \int_{a_\lambda}^{(1-x)y/x} f_\lambda(u) f_\mu(y) \, dudy - \\
 & - \mathbf{1} \left(\max \left\{ \frac{a_\mu}{a_\lambda + a_\mu}, \frac{b_\mu}{b_\lambda + b_\mu} \right\} \leq x < \frac{b_\mu}{a_\lambda + b_\mu} \right) \times \\
 & \times \int_{a_\lambda x/(1-x)}^{b_\mu} \int_{a_\lambda}^{(1-x)y/x} f_\lambda(u) f_\mu(y) \, dudy.
 \end{aligned}$$

Замечание 1. Для вычисления моментов случайной величины π удобно использовать формулу для плотности $f_\pi(x)$, которая получается аналогично следствию 1 при помощи соотношения

$$f_\pi(x) = \frac{1}{x^2} f_\rho \left(\frac{1-x}{x} \right)$$

из соответствующего утверждения статьи [1]. Для сокращения изложения опустим это следствие.

Замечание 2. Особый интерес представляет случай, в котором плотности априорных распределений с ограниченным носителем могут быть представлены в виде полинома

$$f(x) = \sum_{i=0}^n c_i x^i \cdot \mathbf{1}(x \in [a, b]).$$

В работах [3, 4] можно соответственно найти примеры вычисления вероятностных и моментных характеристик π в случаях, когда рассматриваются смеси двух равномерных ($n = 0$) и параболических распределений ($n = 2$).

Для вычисления функции распределения случайной величины $N = \rho/(1 - \rho)$ достаточно воспользоваться теоремой 1 и формулой

$$F_N(x) = F_\rho \left(\frac{x}{1+x} \right).$$

При этом, в отличие от следствия 1, в общем случае нельзя сформулировать соответствующее утверждение в терминах индикаторов неравенств относительно аргумента функции распределения x . Это связано с тем, что, как показано в [5], функция распределения $F_N(x)$ может быть несобственной, т. е. $\lim_{x \rightarrow +\infty} F_N(x) < 1$. При рассмотрении характеристик распределения случайной величины N , таким

образом, существенную роль играет величина «дефекта»

$$\delta = 1 - F_N(+\infty) \equiv P(\rho \geq 1) \equiv 1 - F_\rho(1),$$

для которого справедливо следующее утверждение.

Теорема 2. Пусть независимые абсолютно непрерывные случайные величины λ и μ имеют соответственно носители распределений $[a_\lambda, b_\lambda]$ и $[a_\mu, b_\mu]$, $0 < a_\lambda < b_\lambda$, $0 < a_\mu < b_\mu$. Тогда для «дефекта» δ функции распределения $F_N(x)$ справедливо

$$\delta = \begin{cases} 1, & \text{если } b_\mu < a_\lambda; \\ 0, & \text{если } b_\lambda < a_\mu; \\ 1 - EF_\lambda(\mu) & \text{в противном случае.} \end{cases}$$

Доказательство. Для нахождения величины «дефекта» δ нужно для всевозможных комбинаций взаимного расположения чисел a_λ , b_λ , a_μ , b_μ вычислить $1 - F_\rho(1)$, воспользовавшись теоремой 1. Случаи, когда отрезки $[a_\lambda, b_\lambda]$ и $[a_\mu, b_\mu]$ не пересекаются, — тривиальные.

Пусть $a_\lambda < a_\mu < b_\lambda < b_\mu$. Тогда

$$\begin{aligned} 1 - \delta = F_\rho(1) &= \mathbf{1} \left(\max \left\{ \frac{a_\lambda}{a_\mu}, \frac{b_\lambda}{b_\mu} \right\} < 1 \leq \frac{b_\lambda}{a_\mu} \right) \times \\ &\times \left[\int_{a_\mu}^{b_\lambda} \int_{a_\lambda}^y f_\lambda(u) f_\mu(y) du dy + \int_{b_\lambda}^{b_\mu} f_\mu(y) dy \right] = \\ &= \int_{a_\mu}^{b_\lambda} F_\lambda(y) f_\mu(y) dy + \int_{b_\lambda}^{b_\mu} f_\mu(y) dy = \\ &= EF_\lambda(\mu) - \int_{b_\lambda}^{b_\mu} F_\lambda(y) f_\mu(y) dy + \int_{b_\lambda}^{b_\mu} f_\mu(y) dy = EF_\lambda(\mu). \end{aligned}$$

Остальные случаи взаимного расположения концов отрезков рассматриваются аналогично и приводят к тому же результату.

Замечание 3. При помощи классических методов теории вероятностей имеет смысл рассматривать только случай $\delta = 0$. В частности, при помощи соотношения

$$f_N(x) = \frac{1}{(1+x)^2} f_\rho \left(\frac{x}{1+x} \right)$$

можно найти плотность и моменты случайной величины N . В остальных случаях изучать свойства распределения N можно при помощи квантилей порядка $(0, 1 - \delta)$.

Литература

1. Кудрявцев А. А. Байесовские модели массового обслуживания и надежности: априорные распределения с компактным носителем // Информатика и её применения, 2016. Т. 10. Вып. 1. С. 67–71.
2. Кудрявцев А. А., Шоргин С. Я. Байесовские модели в теории массового обслуживания и надежности. — М.: ФИЦ ИУ РАН, 2015. 76 с.
3. Кудрявцев А. А., Соколов И. А., Шоргин С. Я. Байесовская рекуррентная модель роста надежности: равномерное распределение параметров // Информатика и её применения, 2013. Т. 7. Вып. 2. С. 55–59.
4. Кудрявцев А. А., Палионная С. И. Байесовская рекуррентная модель роста надежности: параболическое распределение параметров // Информатика и её применения, 2016. Т. 10. Вып. 2. С. 80–83.
5. Кудрявцев А. А., Шоргин С. Я. Байесовские модели массового обслуживания и надежности: характеристики среднего числа заявок в системе $M|M|1|_\infty$ // Информатика и её применения, 2010. Т. 4. Вып. 3. С. 16–21.

Поступила в редакцию 29.06.16

CHARACTERISTICS DEPENDENT ON THE BALANCE COEFFICIENT IN BAYESIAN MODELS WITH COMPACT SUPPORT OF *A PRIORI* DISTRIBUTIONS

A. A. Kudryavtsev^{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: Distributions of some characteristics dependent on the balance coefficient, which is defined as a ratio of two parameters that are interpreted as the parameter “obstructing” the functioning of the system and the

parameter “conducting” the functioning of the system, are presented. In the queuing theory for $M|M|1$ models, such characteristics are interpreted as an average amount of claims in the system, the readiness coefficient, the probability that the claim will not be lost and as the marginal system’s reliability for the discrete exponential reliability model. In the framework of the Bayesian approach, it is supposed that initial parameters are random and have *a priori* distributions with compact support.

Keywords: Bayesian approach; mass service theory; reliability theory; mixed distributions; distributions with compact support

DOI: 10.14357/19922264160310

Acknowledgments

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

References

1. Kudryavtsev, A. A. 2016. Bayesovskie modeli massovogo obsluzhivaniya i nadezhnosti: Apriornye raspredeleniya s kompaktnym nositelem [Bayesian queueing and reliability models: *A priori* distributions with compact support]. *Informatika i ee Primeneniya — Inform. Appl.* 10(1):67–71.
2. 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: FIC IU RAN. 76 p.
3. Kudryavtsev, A. A., I. A. Sokolov, and S. Ya. Shorgin. 2013. Bayesovskaya rekurrentnaya model rosta nadezhnosti: Ravnomernoe raspredelenie parametrov [Bayesian recurrent model of reliability growth: Homogeneous distribution of parameters]. *Informatika i ee Primeneniya — Inform. Appl.* 7(2):55–59.
4. Kudryavtsev, A. A., and S. I. Palionnaia. 2016. Bayesovskaya rekurrentnaya model rosta nadezhnosti: Parabolicheskoe raspredelenie parametrov [Bayesian recurrent model of reliability growth: Parabolic distribution of parameters]. *Informatika i ee Primeneniya — Inform. Appl.* 10(2):80–83.
5. Kudryavtsev, A. A., and S. Ya. Shorgin. 2010. Bayesovskie modeli massovogo obsluzhivaniya i nadezhnosti: Kharakteristiki srednego chisla zayavok v sisteme $M|M|1|\infty$ [Bayesian queueing and reliability models: Average number of claims characteristics in $M|M|1|\infty$ system]. *Informatika i ee Primeneniya — Inform. Appl.* 4(3):16–21.

Received June 29, 2016

Contributor

Kudryavtsev Alexey A. (b. 1978) — Candidate of Sciences (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; 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; nubigena@mail.ru

«ВИРТУАЛЬНЫЙ КОНСИЛИУМ» — ИНСТРУМЕНТАЛЬНАЯ СРЕДА ПОДДЕРЖКИ ПРИНЯТИЯ СЛОЖНЫХ ДИАГНОСТИЧЕСКИХ РЕШЕНИЙ*

И. А. Кириков¹, А. В. Колесников², С. В. Листопад³, С. Б. Румовская⁴

Аннотация: Рассматривается проблема принятия индивидуального решения при диагностике пациентов в амбулаторно-поликлинических учреждениях на примере диагностики артериальной гипертензии (АГ). Предлагается повысить качество принятия индивидуального решения за счет консультаций с системой поддержки принятия решения — «Виртуальным консилиумом», моделирующим коллективный интеллект врачей стационара многопрофильного больничного учреждения. Приведены результаты проектирования и экспериментального исследования лабораторного прототипа «Виртуального консилиума».

Ключевые слова: система поддержки принятия решения; виртуальный консилиум; функциональная гибридная интеллектуальная система

DOI: 10.14357/19922264160311

1 Введение

Степень исследования, понимания и качества диагностики проблемных сред и их окружения отражена в научной картине мира, онтологизирующей его представления и делающей рассуждения и целенаправленную деятельность «зависимыми» от них. В искусственном интеллекте понятие «картина мира» соответствует понятие «модель внешнего мира» М. Г. Гаазе-Раппопорта и Д. А. Поспелова [1].

Новая картина мира складывается из многочисленных теорий и взглядов: «ноосфера», «разумный мир» (В. И. Вернадский, Н. Н. Моисеев, А. В. Поздняков); «мир диалектики» — мир диалога разных логик (Е. Л. Доценко); социальная парадигма искусственного интеллекта («The society of mind») М. Минского; системно-организационный подход в искусственном интеллекте В. Б. Тарасова; теория иерархических многоуровневых систем М. Месаровича, Д. Мако и И. Такахару и др. — и укладывается в семь постулатов [2]: (1) признание гетерогенности мира и любого объекта, разнообразия жизни; (2) неопределенность границ объектов и связь «всего со всем»; (3) относительность любой иерархии и горизонтальные связи; (4) дополнитель-

ность и сотрудничество; (5) полицентризм; (6) отнесенность знания; (7) соответствие управления сложности объекта.

Сложная задача диагностики АГ (СЗДАГ) — задача-система, включающая диагностические и технологические подзадачи, повышающие эффективность обработки симптоматической информации о пациенте. Разнообразие подзадач СЗДАГ с различными характеристическими свойствами требует разнообразия соответствующих методов принятия решений, системного анализа, искусственного интеллекта и инженерии знаний.

Анализ результатов влияния новой картины мира на ментальную составляющую врачебной практики и медицинской информатики [3] показал, что, несмотря на стремление биомедицины к гетерогенности восприятия организма человека и процесса его диагностики в рамках семипостулатной картины мира, человек по-прежнему остается «расчлененным» объектом познания, что сформировало «узких» специалистов, поглощенных решением частных задач. Новый тип ученого «прагматика-фактолога» утратил системное мышление, перестал задумываться над тем, что делается «вокруг» и какое значение могут иметь добытые им факты для понимания работы организма в целом. В этой связи

* Работа выполнена при частичной поддержке РФФИ (проект 16-07-00272 А).

¹ Калининградский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, baltbipiran@mail.ru

² Балтийский Федеральный университет имени И. Канта, Калининградский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, avkolesnikov@yandex.ru

³ Калининградский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, ser-list-post@yandex.ru

⁴ Калининградский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, sophiyabr@gmail.com

очевидна необходимость перехода от методов «конкурентной» диагностики к системному мышлению и методам гетерогенной диагностики.

В [3–5] представлены результаты системного анализа СЗДАГ, следуя проблемно-структурной (ПС) методологии, этапы 1–5 [6]: идентификация, редукция сложной задачи, спецификация диагностических подзадач, выбор методов их решения, а также проверка неоднородности сложной задачи диагностики. Работы [3–5] подтвердили релевантность применения междисциплинарных инструментов для решения СЗДАГ, моделирующих разнообразие информации, сотрудничество, дополнительность и относительность знаний, сочетающих методы и методики системного анализа диагностической проблемы с динамическим синтезом метода ее решения и имитацией работы искусственного гетерогенного коллектива — «виртуального консилиума».

Разнообразие — признак, проявление гетерогенности. Следствие закона необходимого разнообразия У. Р. Эшби констатирует, что управление обеспечивается, если разнообразие средств управляющего не меньше разнообразия управляемой им ситуации. Для отображения в информатике ситуативного разнообразия в естественных гетерогенных системах в [6] введены модели «гетерогенная, неоднородная задача» и «гомогенная, однородная задача», а сам закон трактуется так: только разнообразная, скоординированная клиническая деятельность, элементы которой в комбинации решают одну задачу, сделает результат диагностики качественно лучше в обществе с новой научной картиной мира. Специфике такой работы соответствует коллективный труд экспертов в малых группах за круглым столом — консилиумы, совещания, естественные гетерогенные системы для решения сложных задач [3], где на первый план выходят знания и опыт лица, принимающего решения (ЛПР), и экспертов.

Настоящая работа — продолжение работ [3–5, 7] и имеет целью представить: (1) результаты исследования процесса диагностики АГ в лечебно-профилактических больничных учреждениях (ЛПУ) широкого профиля — предлагается повысить эффективность и качество индивидуальных диагностических решений в ЛПУ широкого профиля амбулаторно-поликлинического характера (рис. 1, а) за счет внедрения информационной технологии «Виртуальный консилиум», моделирующей коллективное обсуждение; (2) архитектуру «Виртуального консилиума» и результаты лабораторных экспериментов с его интегрированными моделями (первые результаты лабораторных экспериментов приведены в [7]).

2 Диагностика артериальной гипертензии в многопрофильном стационарном больничном учреждении и в амбулаторно-поликлиническом учреждении

В [8, 9] представлены результаты исследования процесса диагностики АГ в Калининградской клинической областной больнице (КОКБ) (см. рис. 1, б) и ее Диагностическом центре (см. рис. 1, а).

Для формирования полного дифференциального диагноза АГ коллективом врачей во главе с лечащим врачом, ЛПР-кардиологом, в стационаре привлекаются до тринадцати врачей-экспертов — носителей знаний из различных разделов медицины: невролог, нефролог, сосудистый хирург, уролог, психолог, педиатр, акушер-гинеколог, онколог, окулист, врачи функциональной диагностики, эндокринолог, терапевт, кардиолог.

Для исследований выбраны шесть специалистов (см. рис. 1, б), решающих двенадцать функциональных подзадач (рис. 2), возникающих в 90% случаев диагностики АГ, каждый из которых формирует промежуточные заключения о состоянии объекта диагностики в своей области медицинских знаний.

Полученные исходные данные об объекте диагностики разнородны (содержатся в медицинской карте): количественные, визуально-графические параметры (детерминированные переменные), лингвистические четкие и нечеткие переменные. Лицо, принимающее решение, изучает в медицинской карте симптомы и частные диагностические мнения врачей-экспертов, множество которых подбирает сам, и ставит заключительный диагноз. Врачам-экспертам доступны симптомы и мнения других врачей-экспертов из медицинской карты. Лицо, принимающее решение, и врачи-эксперты обследуют пациента и формируют диагностические заключения согласно нормативным документам, например [10]. В ЛПУ широкого профиля (см. рис. 1, а) ЛПР — это врач общей практики или терапевт (иногда кардиолог, но зачастую без опыта работы, к которому направляет терапевт сразу же при выявлении повышенного артериального давления), это врач «прагматик-фактолог» [9], объединяющий в себе роли врача-ЛПР и врачей-экспертов узкой специализации.

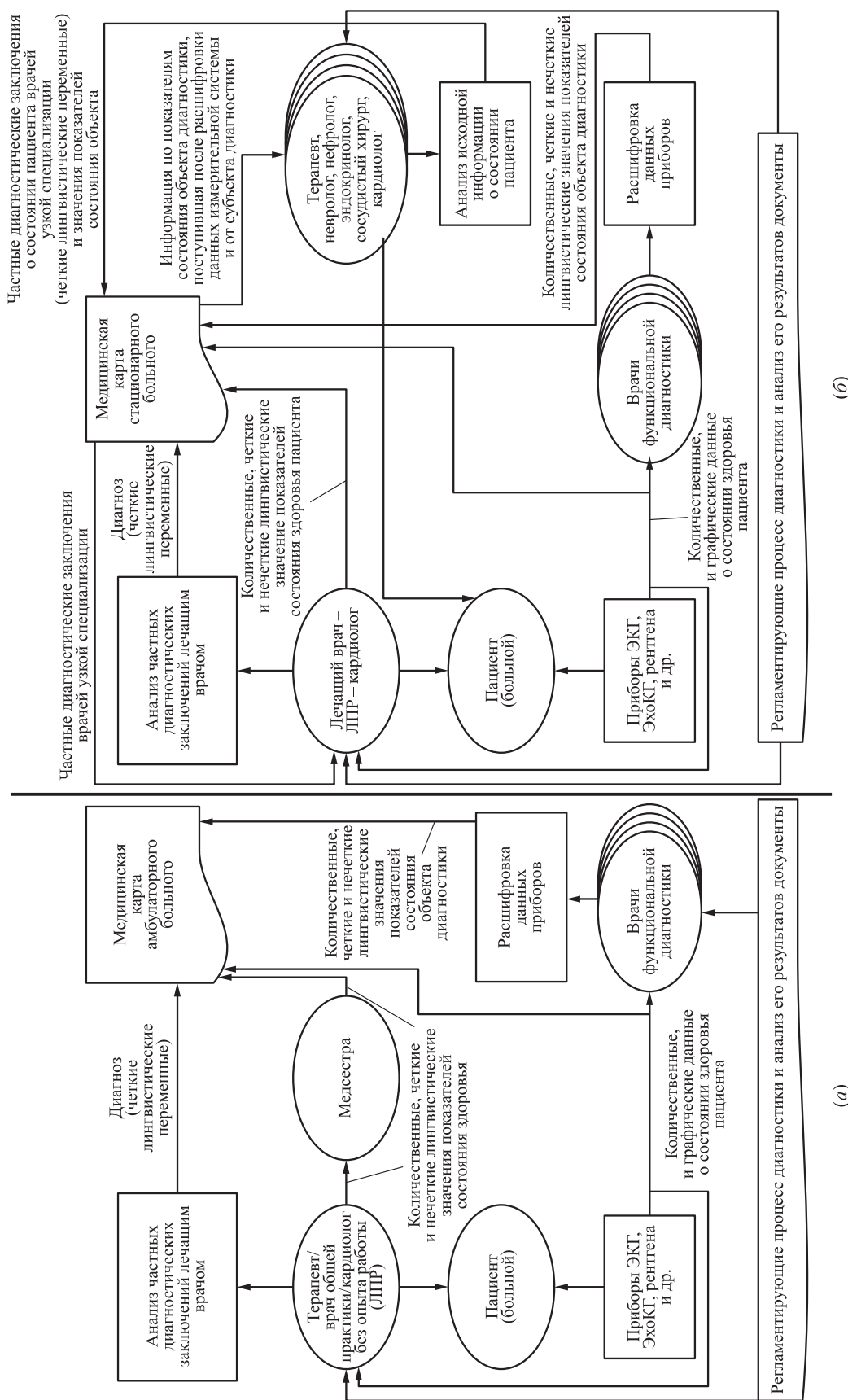


Рис. 1 Концептуальная модель процесса диагностики АГ: (а) в амбулаторно-поликлиническом учреждении; (б) в многопрофильном стационарном больничном учреждении

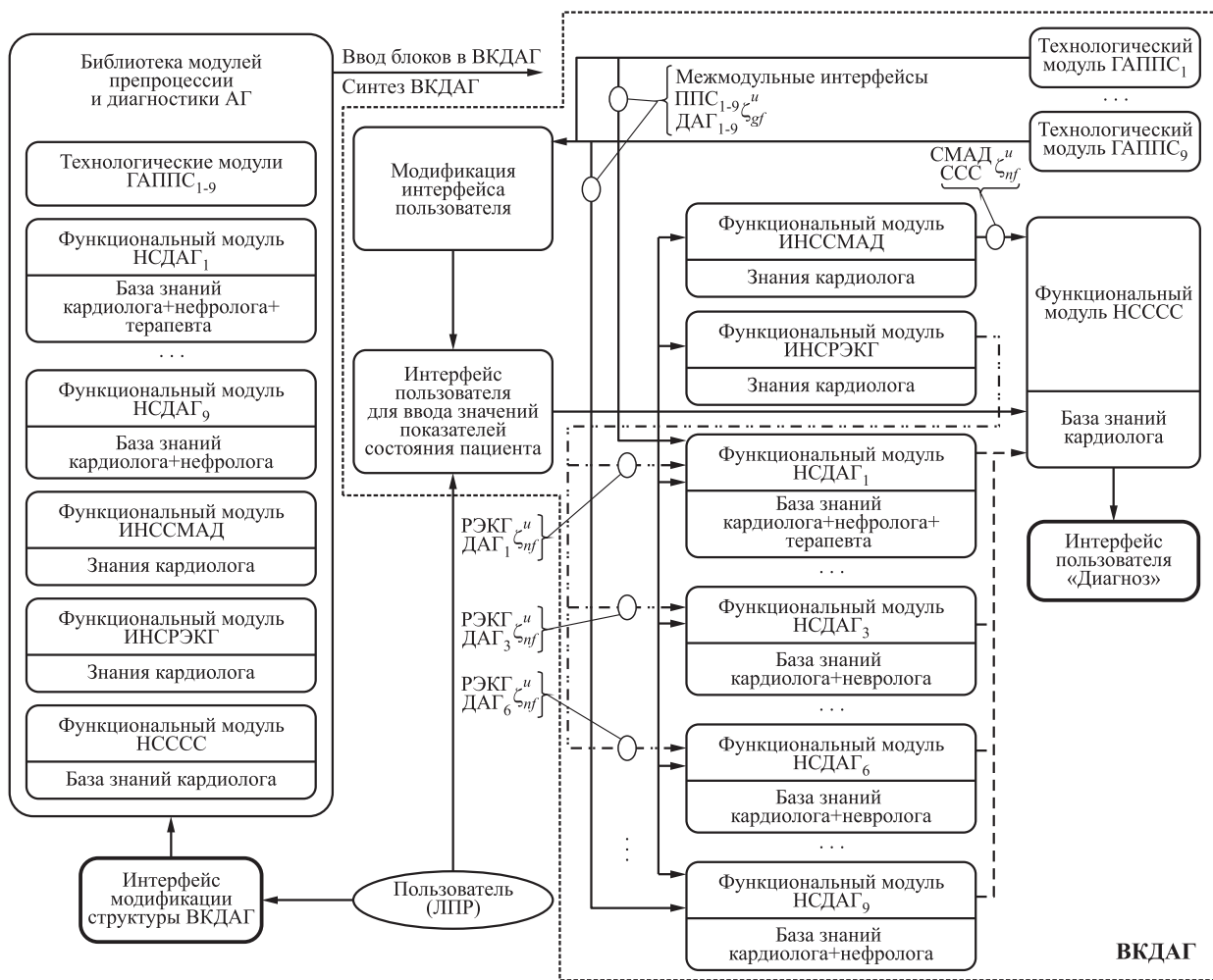


Рис. 2 Архитектура ВКДАГ

Исследования диагностического процесса на материалах Диагностического центра КОКБ по модели на рис. 1, а показали, что 70% пациентов с АГ амбулаторно-поликлинического учреждения не знают о своем заболевании, в то время как в стационарных медицинских учреждениях (см. рис. 1, б) практически в 100% случаев имеет место как адекватное проведение, так и отображение в медицинских картах симптоматических данных обследования с подтверждением диагноза лабораторно-инструментальными методами исследования.

В этой связи предлагается повысить эффективность и качество индивидуальных диагностических решений в ЛПУ широкого профиля амбулаторно-поликлинического характера (см. рис. 1, а) за счет внедрения информационной технологии «Виртуальный консилиум» (см. рис. 2), моделирующей коллективное обсуждение, обладающего синергией, опытом и знаниями в решении подзадач диагностики АГ в стационаре (см. рис. 1, б).

3 Инструментальная среда «Виртуальный консилиум для диагностики артериальной гипертензии»

Инструментальная среда «Виртуальный консилиум», архитектура которой представлена на рис. 2, а структура в [7], ограничена пациентами старше 18 лет, без особых состояний, нет распознавания снимков, не предусматривается назначение лечения и не диагностируется ряд симптоматических артериальных гипертензий.

Архитектура «Виртуального консилиума для диагностики артериальной гипертензии» (ВКДАГ) включает межмодульные интерфейсы ζ^u для модулей, реализованных посредством различных методологий гибридных интеллектуальных систем (ГиИС) (генетические алгоритмы (g), нечеткие сис-

Таблица 1 Описание блоков архитектуры ВКДАГ

Наименование блока	Функции	Вход	Выход
Технологический модуль i -й	Организация эффективной обработки данных и знаний, выбирается для включения в функциональную ГиИС — построение информативного набора признаков для диагностики	Популяция индивидуумов, накладывающихся как маска на i -й функциональный модуль	Наилучшая особь с оптимальным набором признаков — накладывается как маска на i -й функциональный модуль
Функциональный модуль i -й	Классификация состояния здоровья пациента в рамках i -й диагностической подзадачи, выбирается для включения в функциональную ГиИС	Подмножество i -е симптомов с интерфейса пользователя	Частное i -е заключение о состоянии здоровья пациента
Функциональный модуль НСССС, моделирующий ЛПР	Формирование заключительного диагноза АГ (всегда в составе «Виртуального консилиума»)	Подмножество симптомов с интерфейса пользователя, множество выходов функциональных модулей	Заключительный диагноз АГ
Функциональный модуль ИНСРЭКГ	Классификация патологического состояния пациента по его электрокардиограмме	Рассмотрены подробно в [4]	
Функциональный модуль ИНССМАД	Прогноз нормальных значений суточного мониторинга артериального давления и вычисление отклонения		
Интерфейс модификации структуры ВКДАГ	Исключение из диагностики модулей, решающих не интересующие пользователя подзадачи	Выбранные пользователем подзадачи диагностики	Функциональная ГиИС, синтезированная посредством алгоритма из [4]
Интерфейс пользователя «Диагноз»	Визуализация результатов диагностики и корректировка их пользователем	Заключительный диагноз от функционального модуля НСССС	Отчет, содержащий множество симптомов и диагноз
Интерфейс пользователя	Ввод информации о состоянии здоровья пациента	Множество значений показателей состояния здоровья пациента	Показатели состояния здоровья пациента, распределенные по функциональным модулям
Модификация интерфейса пользователя	Деактивация элементов на интерфейсе пользователя для ввода значений показателей состояния здоровья	Множество выходов технологических модулей	Частично деактивированный интерфейс пользователя

темы (f), искусственные нейронные сети (n)). В библиотеке модулей диагностики и препроцессии хранятся заранее инициализированные в программной среде функциональные и технологические модели. По умолчанию все модули включены в структуру «Виртуального консилиума», их описание представлено в табл. 1. «Виртуальный консилиум» (см. рис. 2) запускает интерфейс пользователя, ЛПР-врача — «Интерфейс модификации структуры ВКДАГ», посредством которого включаются функциональные

и технологические модули в работу системы: модуль «Анализ СМАД», модуль «Распознавание ЭКГ», модули технологических подзадач из группы «Построение информативного набора признаков (симптомов) при диагностике заболеваний» и модули подзадач из группы «Диагностика критериев оценки сердечно-сосудистого риска и вторичной АГ у пациента» (ДАГ₁, . . . , ДАГ₉): диагностики поражений органов-мишеней, факторов риска, цереброваскулярных болезней, метаболического синдрома и сахарного диабета, заболеваний пери-

ферических артерий, ишемической болезни сердца, эндокринной АГ, паренхиматозной нефропатии и реноваскулярной АГ соответственно. Все выбранные i -е технологические модули запускаются, решают соответствующую подзадачу и передают информацию на блок «Модификация интерфейса пользователя». Он деактивирует показатели состояния здоровья на «Интерфейсе пользователя для ввода значений показателей состояния здоровья пациента» и корректирует работу i -го функционального модуля подзадач ДАГ₁, ДАГ₉. Далее активируется откорректированный интерфейс, вводятся симптомы, которые передаются функциональным нечетким модулям, решающим подзадачи ДАГ₁, ДАГ₉ (моделируют принятие решения экспертами, врачами смежных специальностей — кардиологом как экспертом, неврологом, нефрологом, терапевтом, эндокринологом, урологом). Последние в свою очередь передают информацию о патологиях, выявленных ими у пациента, функциональному модулю НСССС (моделирует принятие решения ЛПР — врачом-кардиологом), решающему подзадачу «Оценка степени и стадии артериальной гипертензии, степени риска сердечно-сосудистых заболеваний».

В библиотеке ВКДАГ есть еще два функциональных модуля (см. табл. 1), включающихся в работу консилиума посредством «Интерфейса модификации структуры ВКДАГ»:

- (1) ИНСРЭКГ, передающий информацию на модули диагностики поражений органов-мишеней (на рис. 2 — это НСДАГ₁), цереброваскулярных болезней (НСДАГ₃) и ишемической болезни сердца (НСДАГ₆);
- (2) ИНССМАД, формирующий информацию о нормальных значениях суточного артериального давления на функциональный модуль НСССС.

4 Экспериментальное лабораторное исследование программной реализации прототипа инструментальной среды «Виртуальный консилиум»

Экспериментальное лабораторное исследование программной реализации исследовательского прототипа функциональной гибридной интеллектуальной системы ВКДАГ для поддержки приня-

тия сложных диагностических решений необходимо для подтверждения его релевантности [3–5, 7] реальной ситуации диагностики АГ. В [4] представлена информация по особенностям функциональных и технологических моделей гетерогенного модельного поля ВКДАГ, а в [7] — информация по их инициализации в среде MATLAB-Simulink, результаты исследований качества работы каждой модели гетерогенного модельного поля «Виртуального консилиума» автономно, а также подтверждена их релевантность работе экспертов — врачей узкой специализации, что предотвращает распространение ошибок работы автономных моделей на работу интегрированной модели.

В настоящей работе приведены результаты исследования качества интегрированных моделей, синтезированных «Виртуальным консилиумом» и моделирующих дополнительную и сотрудничество, которые имитируют коллективные рассуждения специалистов при постановке диагноза.

В табл. 2 представлены критерии и результаты тестирования интегрированных моделей «Виртуального консилиума» с различными комбинациями знаний врачей, классифицирующих патологическое состояние пациента. Порядок работы моделей гетерогенного модельного поля ВКДАГ: запускаются модели первой очереди — модели технологических элементов ГАППС_{1–9}, корректирующие множества входных переменных моделей НСДАГ_{1–9} и НСССС; обработка информации передается функциональным элементам: модели второй очереди «отправляют» информацию на модели третьей, пятой, шестой и седьмой очередей — ИНСРЭКГ (модель, решающая задачу распознавания электрокардиограммы (ЭКГ)), ИНССМАД (формирует оптимальные множества показателей суточного давления), НСДАГ₉, НСДАГ₂ и НСДАГ₆; третья очередь содержит модели НСДАГ₄ и НСДАГ₅, передающие выходную информацию на вход моделей четвертой и седьмой очередей; четвертая очередь содержит модель НСДАГ₈, передающую информацию модели пятой очереди НСДАГ₁, которая в свою очередь передает информацию НСДАГ₃ (шестая очередь); от НСДАГ₃ передается информация НСДАГ₇ (седьмая очередь); последней запускается модель НСССС, формирующая заключительный диагноз, на вход которой передается выходная информация функциональных моделей второй–седьмой очередей.

Таким образом: (1) без знаний кардиолога, или нефролога, или эндокринолога среднеквадратическая ошибка наибольшая — 0,697; 0,448 и 0,211 соответственно, и объясняется это тем, что кардиолог играет ключевую роль в обработке информации, поступающей от других врачей

Таблица 2 Параметры и результаты тестирования интегрированных моделей

Наименование параметров и результатов тестирования	Значения параметров и результатов тестирования
Объем тестовой выборки ВКДАГ, интегрирующего знания всех шести врачей	800 наблюдений — 500 с диагнозами эссенциальной АГ и 300 с диагнозами вторичной АГ
Объем тестовой выборки ВКДАГ, интегрирующего знания менее шести врачей	400 наблюдений — 200 с диагнозами эссенциальной АГ и 200 с диагнозами вторичной АГ
Источник формирования тестовой выборки	Архив медицинских карт пациентов 1-го кардиологического отделения КОКБ
Элемент тестирующей последовательности	Содержит множество нечетких лингвистических переменных и вектор образа электрокардиограммы (может отсутствовать)
Эталонный диагноз	Результаты деятельности лечащего врача-кардиолога, подводящего общий итог — дифференциальный диагноз АГ
Критерии тестирования	Среднеквадратическая ошибка f классификации состояния здоровья пациента [7]
f (шесть врачей)	0,0837
f (без кардиолога)	0,697
f (без нефролога)	0,448 (в остальных 55,2% случаях диагноз не вызовет доверия)
f (без терапевта)	0,151
f (без невролога)	0,149
f (без эндокринолога)	0,211 (в остальных 78,9% случаях диагноз не вызовет доверия)
f (без сосудистого хирурга)	0,0798
f (без знаний терапевта, невролога, нефролога, эндокринолога, сосудистого хирурга)	0,711
f (без знаний терапевта, невролога, эндокринолога, сосудистого хирурга)	0,485
f (без знаний невролога, эндокринолога, сосудистого хирурга)	0,334
f (без знаний невролога, сосудистого хирурга)	0,167

и от лабораторных исследований, и в постановке заключительного диагноза, а нефролог и эндокринолог — в исключении вторичной АГ; (2) знания врача — сосудистого хирурга не влияют на результаты работы «Виртуального консилиума», и объясняется это тем, что знания сосудистого хирурга, касающиеся диагностики АГ, составляют только 20% базы знаний нечеткой системы, распознающей заболевания периферических артерий (ассоциативные клинические состояния), встречающихся не более чем у 10% населения [11], и в тестовую выборку не попала ни одна карта с данными заболеваниями; (3) чем больше численный состав «Виртуального консилиума», тем с меньшей среднеквадратической ошибкой он классифицирует состояние здоровья пациента; (4) «Виртуальный консилиум» в составе шести врачей диагностирует АГ со среднеквадратической ошибкой постановки диагноза $f = 0,0837$, т. е. дает диагноз, верный в 84% случаях.

Поскольку «Виртуальный консилиум» разработан на основе всероссийских [9] и международных рекомендаций по диагностике АГ и сопутствующих заболеваний, которых должен

придерживаться каждый врач в своей практике, при переносе ВКДАГ в другое больничное учреждение необходимо предоставить врачам данного учреждения протоколы подтверждения диагностических правил всех баз знаний экспериментальными данными из архива КОКБ для ознакомления и внесения при необходимости коррективов в связи с возможными особенностями их контингента пациентов, а также возможных требований по устранению ограничений системы со стороны персонала нового больничного учреждения. Значительной корректировки баз знаний не потребуется.

Таким образом, лабораторные эксперименты с прототипом «Виртуального консилиума» дали обнадеживающие результаты.

Верное решение получено в 84% случаев. В амбулаторно-клинических учреждениях диагноз не выявляется у 70% пациентов в основном по причине инертности врачей, недостатка опыта врачей узкой специализации и нехватки кадров в ЛПУ широкого профиля, что по результатам экспериментов может быть устранено с помощью применения ВКДАГ во время приема пациентов с подозрением на АГ.

5 Заключение

Лабораторно подтверждена эффективность предлагаемого подхода для проектирования диагностических систем как гетерогенных искусственных диагностических систем со свойствами дополненности, сотрудничества и относительности знаний, синтезирующих интегрированные методы и модели, разнообразие которых устраняет разнообразие диагностической информации об организме человека — «Виртуальных консилиумов», моделирующих работу коллектива врачей в многопрофильном стационарном больничном учреждении (на примере КОКБ) и внедрение которых повышает эффективность и качество индивидуальных диагностических решений в амбулаторно-поликлиническом учреждении широкого профиля (на примере Диагностического центра КОКБ), где заключение о состоянии больного из-за проблемы с кадрами узкой специализации принимает чаще всего один специалист — терапевт или врач общей практики, иногда кардиолог, но без опыта работы.

Литература

1. Гаазе-Ранпорт М. Г., Поспелов Д. А. От амебы до работа: модели поведения. — М.: Наука, 1987. 288 с.
2. Колесников А. В., Кириков И. А., Листопад С. В. и др. Решение сложных задач коммивояжера методами функциональных гибридных интеллектуальных систем. — М.: ИПИ РАН, 2011. 295 с.
3. Кириков И. А., Колесников А. В., Румовская С. Б. Исследование сложной задачи диагностики артериальной гипертензии в методологии искусственных гетерогенных систем // Системы и средства информатики, 2013. Т. 23. № 2. С. 81–99. doi: 10.14357/08696527130208.
4. Кириков И. А., Колесников А. В., Румовская С. Б. Функциональная гибридная интеллектуальная система для поддержки принятия решений при диагностике артериальной гипертензии // Системы и средства информатики, 2014. Т. 24. № 1. С. 153–179. doi: 10.14357/08696527140110.
5. Колесников А. В., Румовская С. Б., Листопад С. В., Кириков И. А. Системный анализ в решении сложных диагностических задач // Системный анализ и информационные технологии (САИТ-2015): Тр. VI Междунар. конф. — М.: ИСА РАН, 2015. Т. 1. С. 157–167.
6. Колесников А. В., Кириков И. А. Методология и технология решения сложных задач методами функциональных гибридных интеллектуальных систем. — М.: ИПИ РАН, 2007. 387 с.
7. Кириков И. А., Колесников А. В., Румовская С. Б. Исследование лабораторного прототипа искусственной гетерогенной системы для диагностики артериальной гипертензии // Системы и средства информатики, 2014. Т. 24. № 3. С. 131–143. doi: 10.14357/08696527140309.
8. Румовская С. Б. Методы и средства информатики для диагностики артериальной гипертензии в лечебно-профилактических учреждениях широкого профиля // Задачи современной информатики (ЗСИ-2015): Тр. 2-й молодежной научной конф. — М.: ФИЦ ИУ РАН, 2015. С. 168–174.
9. Кириков И. А., Румовская С. Б. Гетерогенная диагностика артериальной гипертензии // Информатика, управление и системный анализ (ИУСА-2016): Тр. 4-й Всеросс. научной конф. молодых ученых с международным участием. — Тверь: ТвГТУ, 2016. Т. 1. С. 180–188.
10. Комитет экспертов ВНОК. Диагностика и лечение артериальной гипертензии. Российские рекомендации // Системные гипертензии, 2010. Вып. 3. С. 5–26.
11. Галимзянов Ф. В. Заболевания периферических артерий (клиника, диагностика, лечение) // Международный журнал экспериментального образования, 2014. Вып. 8. С. 113–114.

Поступила в редакцию 18.06.16

“VIRTUAL COUNCIL” — SOURCE ENVIRONMENT SUPPORTING COMPLEX DIAGNOSTIC DECISION MAKING

I. A. Kirikov¹, A. V. Kolesnikov^{1,2}, S. V. Listopad¹, and S. B. Rumovskaya¹

¹Kaliningrad Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 5 Gostinaya Str., Kaliningrad 236000, Russian Federation

²Immanuel Kant Baltic Federal University, 14 Nevskogo Str., Kaliningrad 236041, Russian Federation

Abstract: The paper considers the problem of individual decision making during diagnostics of patients in outpatient clinics by the example of arterial hypertension diagnostics. It is proposed to raise the quality of individual decision

making by means of consultations with the “Virtual council” decision support system, which models the work of physician councils in inpatient multifield clinics. The results of development and experimental research of the laboratory prototype of “Virtual council” are presented.

Keywords: decision support system; virtual council; functional hybrid intellectual system

DOI: 10.14357/19922264160311

Acknowledgments

The work was performed with partial support of the Russian Foundation for Basic Research (grant No. 16-07-00272 A).

References

1. Gaaze-Rappoport, M. G., and D. A. Pospelov. 1987. *Ot ameby do robota: Modeli povedeniya* [From ameba to robotic machine: Behavior model] Moscow: Nauka. 288 p.
2. Kolesnikov, A. V., I. A. Kirikov, S. V. Listopad, et al.. 2011. *Reshenie slozhnykh zadach kommivoyazhera metodami funktsional'nykh gibridnykh intellektual'nykh sistem* [Solving of the complex traveling salesman problem by means of functional hybrid intellectual systems]. Moscow: IPI RAN. 295 p.
3. Kirikov, I. A., A. V. Kolesnikov, and S. B. Rumovskaya. 2013. Issledovanie slozhnoy zadachi diagnostiki arterial'noy gipertenzii v metodologii iskusstvennykh geterogennykh sistem [Research of the complex problem at diagnosing of the arterial hypertension within the methodology of artificial heterogeneous systems]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 23(2):81–99. doi: 10.14357/08696527130208.
4. Kirikov, I. A., A. V. Kolesnikov, and S. B. Rumovskaya. 2014. Funktsional'naya gibridnaya intellektual'naya sistema dlya podderzhki prinyatiya resheniya pri diagnostike arterial'noy gipertenzii [Functional hybrid intelligent decision support system for diagnosing of the arterial hypertension]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 24(1):153–179. doi: 10.14357/08696527140110.
5. Kolesnikov, A. V., I. A. Kirikov, S. V. Listopad, and S. B. Rumovskaya. 2015. Sistemnyy analiz v reshenii slozhnykh diagnosticheskikh zadach [Systems analysis for solving complex diagnostic tasks]. *Tr. 6-y Mezhdunar. konf. “Sistemnyy analiz i informatsionnye tekhnologii”* [6th Conference (International) “Systems Analysis and Information Technology” Proceedings]. Moscow. 1:157–167.
6. Kolesnikov, A. V., and I. A. Kirikov. 2007. *Metodologiya i tekhnologiya resheniya slozhnykh zadach metodami funktsional'nykh gibridnykh intellektual'nykh sistem* [Methodology and technology for solving of complex problems using the methodology of functional hybrid artificial systems]. Moscow: IPI RAN. 387 p.
7. Kirikov, I. A., A. V. Kolesnikov, and S. B. Rumovskaya. 2014. Issledovanie laboratornogo prototipa iskusstvennoy geterogennoy sistemy dlya diagnostiki arterial'noy gipertenzii [Research of the laboratory prototype of the artificial heterogeneous system for diagnosing of the arterial hypertension]. *Sistemy i Sredstva informatiki — Systems and Means of Informatics* 24(3):131–143. doi: 10.14357/08696527140309.
8. Rumovskaya, S. B. 2015. Metody i sredstva informatiki dlya diagnostiki arterial'noy gipertenzii v lechenno-profilakticheskikh uchrezhdeniyakh shirokogo profilya [Methods and tools of informatics for diagnostics of arterial hypertension in multiskilled medical preventive institution]. *Tr. 2-y molodezhnoy nauchnoy konf. “Zadachi sovremennoy informatiki”* [2nd Youth Conference “Tasks of Modern Informatics” Proceedings]. Moscow: FRC “Computer Science and Control” RAS. 168–174.
9. Kirikov, I. A., and S. B. Rumovskaya. 2016. Geterogen'naya diagnostika arterial'noy gipertenzii [Heterogeneous diagnostics of arterial hypertension]. *Tr. 4-y Vseross. nauchnoy konf. molodykh uchenykh s mezhdunarodnym uchastiem “Informatika, upravlenie i sistemnyy analiz”* [4th Youth Conference (International) “Informatics, Control and Systems Analysis” Proceedings]. Tver: Tver State Technical University. 1:180–188.
10. Komitet ekspertov VNOK [Committee of experts of All-Russia Scientific Society of Cardiologists]. 2010. Diagnostika i lechenie arterial'noy gipertenzii. Rossiyskie rekomendatsii [Diagnosing and treatment of arterial hypertension. Russian recommendation]. *Sistemnye gipertenzii* [Systemic Hypertension] 3:5–26.
11. Galimzyanov, F. V. 2014. Zabolevaniya perifericheskikh arteriy (Klinika, diagnostika, lechenie) [Peripheral vascular disease (Clinic, diagnostics, treatment)]. *Mezhdunarodnyy zhurnal eksperimental'nogo obrazovaniya* [Int. J. Research Education] 8:113–114.

Received June 18, 2016

Contributors

Kirikov Igor A. (b. 1955) — Candidate of Sciences (PhD) in technology; director, Kaliningrad Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 5 Gostinaya Str., Kaliningrad 236000, Russian Federation; baltbipiran@mail.ru

Kolesnikov Alexander V. (b. 1948) — Doctor of Sciences in technology; professor, Department of Telecommunications, Immanuel Kant Baltic Federal University, 14 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 Sciences (PhD) in technology; scientist, Kaliningrad Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 5 Gostinaya Str., Kaliningrad 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

ВАРИАНТ СОЗДАНИЯ ЛОКАЛЬНОЙ СИСТЕМЫ КООРДИНАТ ДЛЯ СИНХРОНИЗАЦИИ ИЗОБРАЖЕНИЙ ВЫБРАННЫХ СНИМКОВ

О. П. Архипов¹, П. О. Архипов², И. И. Сидоркин³

Аннотация: Рассмотрены проблемы сравнения пар изображений, имеющих искажения поворота и сдвига сцен друг относительно друга. Разработан алгоритм создания локальной системы координат (ЛСК) для пар сравниваемых изображений.

Ключевые слова: алгоритм; методика; локальная система координат; цветное изображение; синхронизация; пиксель; цветное пятно; фильтрация

DOI: 10.14357/19922264160312

1 Введение

Важным этапом обработки кадров видеопотока является построение ЛСК для синхронизации обрабатываемых изображений. Под синхронизацией понимается процедура совмещения пары обрабатываемых кадров путем смещения одного изображения относительно другого для достижения совпадения одинаковых устойчивых робастных структур. В качестве общих робастных структур могут выступать границы объектов, имеющих на полутонных изображениях, и центры одинаковых по площади цветных пятен соответствующих цветных изображений. В случае необходимости сравнения пары кадров, полученных с различных точек съемки либо с отличным углом съемки, в результате чего изображения оказались смещены относительно друг друга, синхронизация может стать единственно возможным решением для осуществления возможности машинного сравнения изображений.

В данной статье описывается процесс создания ЛСК для синхронизации пар обрабатываемых изображений. Актуальность работы обусловлена необходимостью сравнения пар изображений, которые были получены с разных точек съемки, что привело к искажениям поворота и смещения.

Целью данной работы является разработка варианта создания ЛСК для синхронизации пар изображений выбранных снимков. Основная идея работы состоит в том, что для синхронизации двух изображений необходимо отыскать на этих изображениях

робастные структуры, которые повторялись бы на каждом из этих изображений, а затем выполнить создание ЛСК с сохранением лишь общей части обрабатываемой пары изображений. Предполагается, что даже будучи смещенными друг относительно друга и/или повернутыми на произвольный угол, данные изображения, имеющие общую совпадающую часть, могут быть синхронизированы путем создания ЛСК и преобразованием одного из изображений. Независимо от угла поворота и смещения изображений, имеющих общую часть, робастные структуры данных изображений будут совпадать.

2 Обзор аналогов

Одними из наиболее распространенных методов определения геометрического рассогласования изображений являются корреляционные методы [1, 2]. Данные методы позволяют рассчитать коэффициент корреляции для всех возможных вариантов смещения изображений друг относительно друга и выбрать одно пиковое значение, которое будет соответствовать наибольшему совпадению двух сравниваемых изображений. Еще одним примером определения взаимного сдвига изображений являются статические методы, в основе которых лежит процесс вычисления евклидовой меры взаимного рассогласования изображений [3]. Однако данные методы являются весьма чувствительными к шумам на изображениях, которые являются их неотъем-

¹Орловский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, arkhypov12@yandex.ru

²Орловский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, agraul@mail.ru

³Орловский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, voronecburgsiti@mail.ru

лемой частью, и, что более существенно, они не позволяют выполнить согласование изображений, имеющих искажение поворота.

Так как при съемке изображений нестационарной камерой получаемые изображения имеют именно искажения сдвига и поворота, то перечисленные выше методы не могут быть использованы для синхронизации таких изображений. В данной статье предлагается метод, основанный на выявлении робастных характеристик, имеющих сходство на обоих обрабатываемых изображениях, который позволит выполнять синхронизацию изображений, подвергнутых искажениям сдвига и поворота.

3 Создание локальной системы координат для синхронизации изображений выбранных снимков

Цветное изображение представляется в виде двумерной последовательности пикселей вида

$$\text{Image}_i = \left\{ \begin{array}{l} p_{i,1,1}(\mathbf{R}, \mathbf{G}, \mathbf{B}), p_{i,1,2}(\mathbf{R}, \mathbf{G}, \mathbf{B}), \dots \\ \dots \dots \dots \\ p_{i,h,1}(\mathbf{R}, \mathbf{G}, \mathbf{B}), p_{i,h,2}(\mathbf{R}, \mathbf{G}, \mathbf{B}), \dots \\ \dots, p_{i,1,w}(\mathbf{R}, \mathbf{G}, \mathbf{B}) \\ \dots \dots \\ \dots, p_{i,h,w}(\mathbf{R}, \mathbf{G}, \mathbf{B}) \end{array} \right\},$$

$$i \in [1, 2], w \in [1, W_i], h \in [1, H_i],$$

где Image_i — изображение снимка i ; p — пиксели с цветовыми координатами $(\mathbf{R}, \mathbf{G}, \mathbf{B})$; W_i и H_i — ширина и высота изображения снимка i в пикселях.

Для сравнения цветных изображений предлагается использовать цветные пятна изображений и робастные характеристики этих изображений. Для этого необходимо выполнить процедуру получения полутоновых изображений для получения наборов робастных характеристик каждого из обрабатываемых изображений вида

$$\text{Im}_i = Q(\varphi_{a,1}(\text{Image}_i), \varphi_{a,2}(\text{Image}_i), \varphi_{a,3}(\text{Image}_i)),$$

$$i \in [1, 2],$$

где Im_i — полутоновое изображение снимка i ; Q — функция объединения полутоновых преобразований; φ — функция выполнения полутоновых преобразований [4].

Перед выполнением сегментации цветных изображений необходимо выполнить огрубление цветовых составляющих изображений до 256 цветов, что позволит получить более удобные для сегментации изображения с четким контрастированием цвето-

вых пятен [5]. Процедура аппроксимации изображений выполняется в два этапа:

- (1) аппроксимация изображений до 4096 цветов вида

$$\text{Img}_i = \{ \Psi_{\text{app}_{4096}}(\text{Image}_i, \text{Pal}_{4096}) \}, i \in [1, 2],$$

где Img_i — аппроксимированное до 4096 цветов изображение снимка i ; $\Psi_{\text{app}_{4096}}$ — функция получения множества $(\mathbf{R}, \mathbf{G}, \mathbf{B})$ -пикселей в результате аппроксимации к 4096 цветам; Pal_{4096} — палитра 4096 цветов;

- (2) аппроксимация изображений до 256 цветов вида

$$\text{Img}_i = \{ \Psi_{\text{app}_{256}}(\text{Image}_i, \text{Pal}_{256}) \}, i \in [1, 2],$$

где $\Psi_{\text{app}_{256}}$ — функция получения множества $(\mathbf{R}, \mathbf{G}, \mathbf{B})$ -пикселей в результате аппроксимации к 256 цветам; Pal_{256} — палитра 256 цветов.

Полученные в результате выполнения двух этапов аппроксимации изображения должны быть сегментированы с целью формирования последовательности цветных пятен каждого изображения. Последовательность цветных пятен изображений можно представить в виде:

$$\Psi_{\text{segm}_i} = \{ \Psi_{i,j} \} = \left\{ \begin{array}{l} \psi_{i,1}(p_{i,1,1}), \dots, \psi_{i,1}(p_{i,1,t}), \dots \\ \dots \dots \dots \\ \psi_{i,n}(p_{i,h,1}), \dots, \psi_{i,n}(p_{i,h,t}), \dots \\ \dots, \psi_{i,u}(p_{i,1,w-g}), \dots, \psi_{i,k}(p_{i,1,w}) \\ \dots \dots \dots \\ \dots, \psi_{i,j}(p_{i,h,w-d}), \dots, \psi_{i,j}(p_{i,h,w}) \end{array} \right\}$$

$$i \in [1, 2], j \in [0, J_j], t \in [1, T_i],$$

$$d \in [1, D_i], g \in [1, G_i], u \in [1, U_i],$$

$$w \in [1, W_i], h \in [1, H_i], n \in [1, N_i],$$

$$N_i \leq J_i, U_i \leq J_i, T_i \leq J_i, D_i < W_i,$$

где Ψ_{segm_i} — множество сегментов изображения снимка i ; $\psi_{i,j}$ — сегмент с номером j цветного изображения снимка i ; J_i — максимальное количество цветных сегментов изображения снимка i .

Полученное множество цветных пятен представляет собой последовательность из всех цветоразличимых сегментов обрабатываемых изображений [6–8]. Для осуществления синхронизации изображений путем определения и сопоставления робастных структур пары обрабатываемых изображений необходимо выполнить фильтрацию полученного множества цветных пятен. Сравнительно маленькие и большие по площади пятна на изображениях не позволяют выполнить синхронизацию изображений из-за того, что маленькие пятна могут с большой вероятностью повторяться на паре обрабатываемых изображений либо вовсе пропа-

дать, а большие пятна могут оказаться на границах изображений, что приведет к невозможности точного определения их центров. Следовательно, необходимо выбрать для дальнейшего рассмотрения только цветные пятна средних размеров, имеющих граничные точки, полученные в результате полутоновых преобразований.

Средняя величина пятен определяется как суммарная величина значений площадей всех цветных пятен изображения, деленная на количество цветных пятен данного изображения.

Ввиду того что точное совпадение размеров пятен маловероятно, выбирать следует при фильтрации пятна, площадь которых будет принадлежать промежутку от $AveSize/2$ до $3AveSize$, где $AveSize$ — средний размер цветного пятна изображения. Выбор данного интервала увеличивает количество цветных пятен, подлежащих рассмотрению, и повышает вероятность успешной синхронизации изображений.

Отфильтрованная последовательность цветных пятен обрабатываемых изображений может быть представлена в виде:

$$\theta_{\text{segm}_i} = \{\theta_{i,k}\} \subset \Psi_{\text{segm}_i} : i \in [1, 2], \\ k \in [0, K], j \in [0, J_i], K_i \leq J_i,$$

где θ_{segm_i} — множество цветных сегментов, оставшихся после фильтрации цветных сегментов изображения снимка i .

После того как получена последовательность цветных пятен, удовлетворяющая всем условиям фильтрации: площадь и наличие граничных точек, необходимо выполнить процедуру сравнения множеств θ_{segm_i} для первого и второго изображений. Сравнение производится путем определения совпадения площадей цветных пятен и взаимного удаления данных пятен от других на каждом изображении. При этом, найдя цветное пятно $\theta_{1,b} : b \leq K_1$, принадлежащее первому изображению, и цветное пятно $\theta_{2,v} : v \leq K_2$, площадь которого соответствует $\theta_{1,b}$, необходимо выполнить проверку удаленности от всех цветных пятен на каждом изображении относительно данных пятен. Если $Y_{\text{segm}} = \emptyset$, то данная пара заносится в множество как совпадающие друг с другом сегменты. Если расстояния до большей части пятен, которые были занесены в Y_{segm} , совпадают, то необходимо добавить в множество Y_{segm} данную пару цветных пятен, иначе они будут признаны как ошибочно выбранные совпадающими друг с другом. Получаемую последовательность можно представить в виде:

$$Y_{\text{segm}} = \{\tau_m\} \subset \theta_{\text{segm}_i} : i \in [1, 2], \\ m \in [0, M_i], K_1 \geq M_i \leq K_2,$$

где Y_{segm} — множество цветных сегментов, выбранных в качестве совпадающих на паре изображений выбранных снимков.

Поскольку первый элемент множества Y_{segm} не может быть проверен на удаленность от других цветных пятен, а второй и последующий сравниваются только с предыдущими элементами, то необходимо выполнить дополнительную проверку удаленности каждого элемента множества Y_{segm} для каждого изображения, тем самым исключив возможные случайные ошибки. Результирующее множество сегментов, имеющих соответствие на паре обрабатываемых изображений, может быть представлено в виде:

$$S_{\text{segm}} = \{s_c\} \subset Y_{\text{segm}} : c \in [0, C_i], K_1 \geq C_i \leq K_2,$$

где S_{segm} — множество цветных сегментов, оставшихся после фильтрации цветных сегментов, выбранных в качестве совпадающих на паре изображений выбранных снимков.

Успешная синхронизация изображений возможна, если имеется три и более сегментов, имеющих соответствие на паре обрабатываемых изображений. При этом чем дальше данные цветные пятна будут располагаться друг от друга, тем выше точность синхронизации и создания ЛСК. Если число соответствующих друг другу сегментов три и более, то выполняется процедура создания ЛСК для обрабатываемой пары изображений путем определения угла поворота одного изображения относительно другого и вычисления расстояний до краев общей области с переносом пикселей каждого из изображений. Таким образом, при выполнении условия наличия минимум трех цветных пятен, которые совпадают на паре обрабатываемых изображений, за счет выполнения процедуры

$$\text{Im}_-s_i = \text{Qs}_i(\text{Image}, S_{\text{segm}}), i \in [1, 2],$$

где Im_-s_i — изображения, преобразованные к ЛСК; Qs_i — процедура построения ЛСК для изображения снимка i ; получаем два новых изображения, которые будут иметь новые ЛСК, совпадающие на обоих изображениях.

На рис. 1 приведена схема алгоритма создания ЛСК для синхронизации изображений выбранных кадров.

Процессы на рис. 1: 1 — получение полутоновых представлений изображений Image_1 и Image_2 ; 2 — аппроксимация изображений Image_1 и Image_2 к палитре 4096 цветов; 3 — аппроксимация изображений Img_1 и Img_2 к палитре 256 цветов; 4 — сегментация изображений Img_1 и Img_2 ; 5 — фильтрация цветных сегментов Ψ_{segm_1} и Ψ_{segm_2} ;

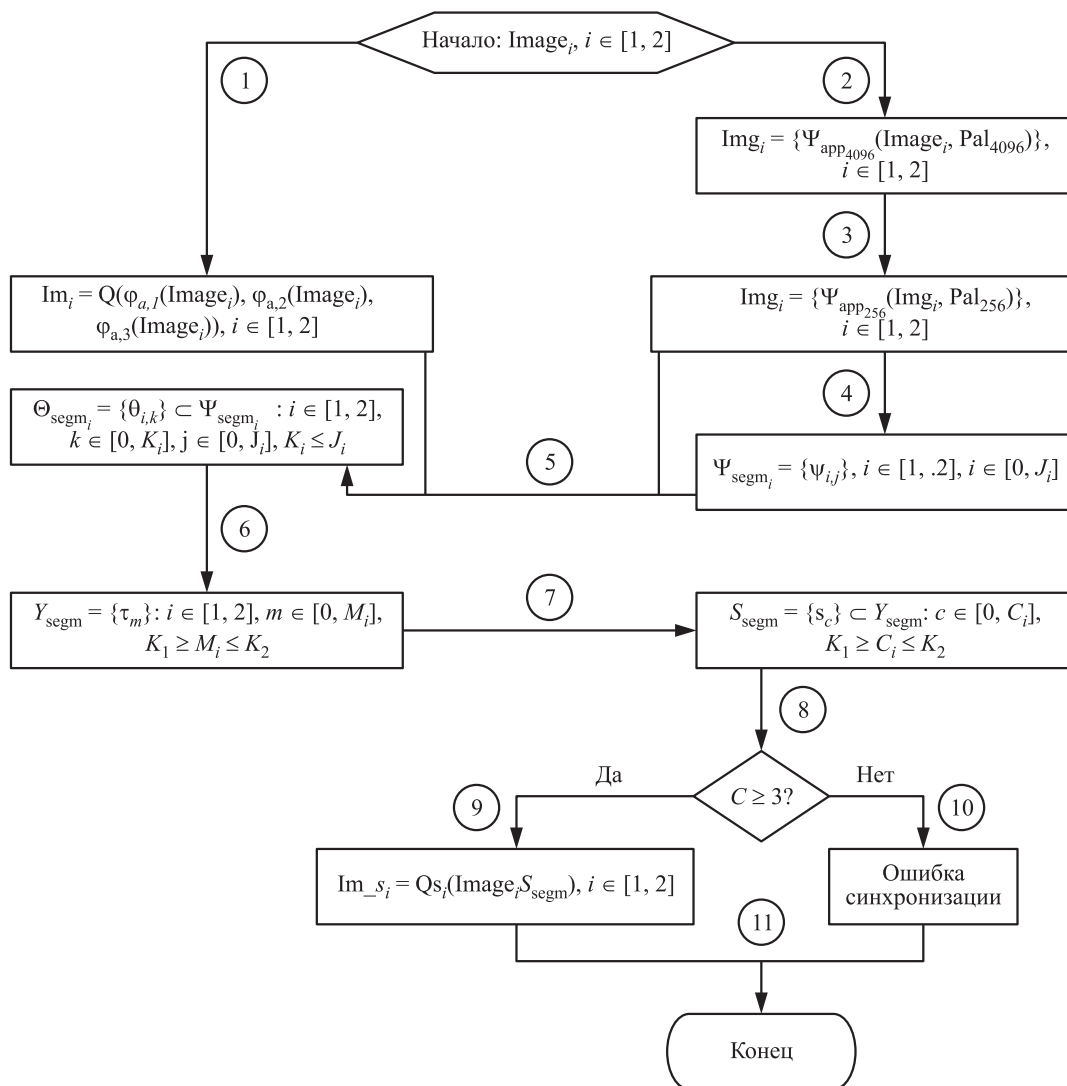


Рис. 1 Схема алгоритма создания ЛСК для синхронизации изображений выбранных кадров

6 — определение соответствия цветных сегментов θ_{segm_1} и θ_{segm_2} пары изображений Img_1 и Img_2 вне зависимости от угла поворота и смещения; 7 — фильтрация обнаруженных сегментов множества Y_{segm} ; 8 — определение числа совпадающих пятен на паре сравниваемых изображений и возможности построения ЛСК; 9 — построение ЛСК; 10 — выдача ошибки построения ЛСК; 11 — завершение работы.

4 Результаты вычислительных экспериментов

Для тестирования предлагаемого варианта создания ЛСК для синхронизации изображений выбранных снимков была выбрана пара кадров, представленных на рис. 2.

После выполнения сегментации и фильтрации сегментов оставшиеся сегменты на изображениях первого и второго снимков выделены зеленым цветом. Для наглядности вручную были отмечены красными линиями и стрелками части изображений, по которым наиболее отчетливо видно относительное смещение изображений (рис. 3).

После выполнения вышеописанных процедур было обнаружено 6 совпадающих цветных сегментов, за счет которых была построена ЛСК и преобразованы оба изображения к виду, пригодному для сравнения. Полученные изображения представлены на рис. 4, а и 4, б.

5 Заключение

В данной статье был предложен вариант создания ЛСК для синхронизации изображений выбран-



(a)



(б)

Рис. 2 Изображения первого (a) и второго (a) снимков



(a)



(б)

Рис. 3 Сегменты для синхронизации, визуализация относительного смещения изображений



(a)



(б)

Рис. 4 Преобразованные изображения первого (a) и второго (б) снимков

ных снимков. В результате его тестирования были получены положительные результаты для изображений, имеющих искажение сдвига и поворота.

Литература

1. *Прэйтт У.* Цифровая обработка изображений / Пер. с англ. — М.: Мир, 1982. (Pratt W. K. Digital image processing. —Wiley-Interscience Publication, 1978. 750 p.)
2. *Форсайт Д., Понс Ж.* Компьютерное зрение. Современный подход. — М.: Вильямс, 2004. 928 с.
3. *Гонзалес Р., Вудс Р.* Цифровая обработка изображений / Пер. с англ. — М.: Техносфера, 2005. 1070 с. (Gonzalez R. C., Woods R. E. Digital image processing. — Wiley-Interscience Publication, 2002. 793 p.)
4. *Архипов О. П., Зыкова З. П.* Применение полутоновых представлений при анализе изменений цветных изображений // Информатика и её применения, 2014. Т. 8. Вып. 3. С. 90–99.
5. *Архипов О. П., Зыкова З. П.* Интеграция гетерогенной информации о пикселях и их цветовосприятии // Информатика и её применения, 2010. Т. 4. Вып. 4. С. 14–25.
6. *Архипов О. П., Зыкова З. П.* Функциональное описание индивидуального цветовосприятия // Известия ОрелГТУ. Сер. Информационные системы и технологии, 2010. № 5. С. 5–12.
7. *Архипов О. П., Зыкова З. П.* RGB-характеризация пространства цветовосприятия // Системы и средства информатики, 2010. Вып. 20. № 1. С. 72–89.
8. *Архипов О. П., Зыкова З. П.* Равноконтрастные градиционные преобразования ступенчатых тоновых шкал // Информационные системы и технологии, 2011. № 4. С. 39–46.

Поступила в редакцию 12.07.16

THE OPTION TO CREATE A LOCAL COORDINATE SYSTEM FOR SYNCHRONIZATION OF SELECTED IMAGES

O. P. Arkhipov, P. O. Arkhipov, and I. I. Sidorkin

Orel Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 137 Moskovskoe Sh., Orel 302025, Russian Federation

Abstract: While comparing pairs of images, in most cases, the problem of misalignment of images arises in which one image is distortion of translation and rotation relative to another image. Such an image is quite difficult to compare in automatic mode. Existing methods of image pairs synchronize a large number of constraints due to which most of them are rarely used. The proposed option to create a local coordinate system for synchronizing images is based on the analysis of color spots presented on the images. It is assumed that successful synchronization of two images on their total amount of colored spots is to be found that match on the data images. For comparison, it is suggested to use colored spots and distances between spots. In order to successfully synchronize, one needs at least three colored spots, which would coincide in all modes of filtration. The experiments show acceptable results of synchronization.

Keywords: algorithm; local coordinate system; color image; synchronization; pixel; colored spot; filtration

DOI: 10.14357/19922264160312

References

1. Pratt, W. K. 1978. *Digital image processing*. Wiley-Interscience Publication. 750 p.
2. Forsyth, D. A., and J. Ponce. 2002. *Computer vision: A modern approach*. Prentice Hall Professional Technical Reference. 720 p.
3. Gonzalez, R. C., and R. E. Woods. 2002. *Digital image processing*. Wiley-Interscience Publication. 793 p.
4. Arkhipov, O. P., and Z. P. Zykova. 2014. Primenenie polutonykh predstavleniy pri analize izmeneniy tsvetnykh izobrazheniy [The use of half-tone representations in the analysis of changes in color images]. *Informatika i ee Primeneniya — Inform. Appl.* 8(3):90–99.
5. Arkhipov, O. P., and Z. P. Zykova. 2010. Integratsiya geterogennoy informatsii o pikselyakh i ikh tsvetovospriyatii [Integration of heterogeneous information about pixels and their color perception]. *Informatika i ee Primeneniya — Inform. Appl.* 4(4):14–25.
6. Arkhipov, O. P., and Z. P. Zykova. 2010. Funktsional'noe opisaniye individual'nogo tsvetovospriyatiya [Characteristics of color perceptual space]. *Izvestiya OrTGU. Ser. Informatsionnye sistemy i tekhnologii* [Herald of Oryol Technical State University. Ser. information systems and technologies] 5:5–12.

7. Arhipov, O. P., and Z. P. Zykova. 2010. RGB-kharakterizatsiya prostranstva tsvetospriyatiya [RGB-characterization of color space]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 1(20): 72–89.
8. Arhipov, O. P., and Z. P. Zykova. 2011. Ravnokontrastnye gradatsionnye preobrazovaniya stupenchatykh tonovykh shkal [Equal contrast graded transformation of step tinted scales]. *Informatsionnye Sistemy i Tekhnologii* [Information Systems and Technologies] 4:39–46.

Received July 12, 2016

Contributors

Arhipov Oleg P. (b. 1948) — Candidate of Science (PhD) in technology, Director, Oryol Branch of Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 137 Moskovskoe Sh., Oryol 302025, Russian Federation; arhipov12@yandex.ru

Arhipov Pavel O. (b. 1979) — Candidate of Science (PhD) in technology, senior scientist, Oryol Branch of Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 137 Moskovskoe Sh., Oryol 302025, Russian Federation; arpaul@mail.ru

Sidorkin Ivan I. (b. 1990) — engineer-researcher, Orel Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 137 Moskovskoe Sh., Orel 302025, Russian Federation; voronecburgsiti@mail.ru

УСКОРЕННЫЙ АЛГОРИТМ СТЕРЕОСОПОСТАВЛЕНИЯ НА ОСНОВЕ ГЕОДЕЗИЧЕСКИХ ВСПОМОГАТЕЛЬНЫХ КОЭФФИЦИЕНТОВ

О. А. Яковлев¹, А. В. Гасилов²

Аннотация: Среди локальных алгоритмов стереосопоставления качественные результаты дают алгоритмы, использующие концепцию вспомогательных коэффициентов. В данной работе предложена модификация алгоритма стереосопоставления на основе геодезических расстояний. Предлагаемая модификация существенно снижает вычислительные затраты, незначительно уступая в качестве сопоставления оригинальному подходу, что подтверждается приведенными результатами экспериментов. Рассматриваемый алгоритм опирается на сегментацию одного из изображений стереопары, используя геодезические вспомогательные коэффициенты для вычисления цвета каждого сегмента. Такое преобразование изображения делает возможным применение принципа частичных сумм при вычислении стоимости сопоставления, что и является основным источником прироста производительности.

Ключевые слова: стереосопоставление; сегментация; геодезические вспомогательные коэффициенты

DOI: 10.14357/19922264160313

1 Введение

Задача стереосопоставления двух изображений состоит в установлении для каждой точки одного изображения соответствующей ей точки на втором изображении. В данной работе рассматривается самый распространенный вариант исходных данных — пара ректифицированных изображений. В паре таких изображений естественным образом можно выделить левое (I_L) и правое (I_R). На ректифицированных изображениях эпиполярные прямые параллельны оси Ox , следовательно, соответствующие точки имеют равные координаты по оси Oy . Пусть $p \in I_L$ и $p' \in I_R$ — соответствующие точки, тогда диспарантностью пикселя p будем называть величину $\text{disp}(p) = p_x - p'_x$. Значения $\text{disp}(p)$ для каждого $p \in I_L$ образуют матрицу, называемую картой диспарантности. Карта диспарантности выступает результатом стереосопоставления изображений и может быть использована, например, для трехмерной реконструкции сцены.

Среди алгоритмов стереосопоставления выделяют глобальные и локальные алгоритмы. Глобальные алгоритмы решают задачу оптимизации, вычисляя оптимальную по некоторому критерию карту диспарантности, что обеспечивает высокое качество результата. Однако вычислительная сложность глобальных алгоритмов не позволяет применять их в системах реального времени.

Локальные алгоритмы являются представителями класса жадных алгоритмов, имеют высокую производительность, но качество вычисленных карт диспарантности у них значительно ниже, чем у глобальных алгоритмов. В связи с постоянным совершенствованием аппаратных средств большинство исследований направлено на улучшение качества сопоставления изображений. Как следствие, многие локальные алгоритмы приобрели черты глобальных, и их применение в условиях ограниченных вычислительных ресурсов становится невозможным. Алгоритм сопоставления изображений с помощью геодезических вспомогательных коэффициентов [1] является ярким примером локального алгоритма, обладающего как высоким качеством, так и высокой вычислительной сложностью. В настоящей работе рассматривается модификация этого алгоритма, позволяющая добиться роста производительности с незначительной потерей качества.

2 Стереосопоставление на основе геодезических вспомогательных коэффициентов

Рассмотрим неориентированный граф, вершинами которого являются пиксели изображения и каждая вершина соединена ребром с 8 соседними.

¹Орловский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, maucra@gmail.com

²Орловский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, gasilov.av@yandex.ru

Стоимостью ребра будем считать евклидово расстояние между двумя пикселями как RGB-векторами. Стоимость кратчайшего пути между двумя вершинами такого графа будем называть геодезическим расстоянием между соответствующими пикселями изображения.

В работе [1] предлагается использовать в качестве вспомогательного коэффициента величину, обратную геодезическому расстоянию:

$$w(p, q) = e^{-g(p, q)/\gamma}, \quad (1)$$

где $g(p, q)$ — геодезическое расстояние между p и q ; γ — положительный коэффициент.

Коэффициент γ позволяет регулировать скорость убывания $w(p, q)$ при возрастании $g(p, q)$. Рассмотрим функцию $f(x) = e^{-x/\gamma}$. Ее производная $f'(x) = -(1/\gamma)e^{-x/\gamma}$. Таким образом, чем меньше γ , тем быстрее убывает $f(x)$.

Для вычисления $\text{disp}(p)$ с помощью вспомогательных коэффициентов используется сканирующее окно, размер которого подбирается экспериментально. Пусть Win_p — сканирующее окно с центром в точке p , тогда стоимость диспарантности d для пикселя p можно вычислить как

$$c(p, d) = \sum_{q \in \text{Win}_p} w(p, q) \text{diff}(q, (q_x - d, q_y)),$$

где $\text{diff}(p, q)$ — величина цветоразличия пикселя p левого изображения и пикселя q правого.

Для вычисления диспарантности каждого пикселя будем использовать локальную оптимизацию: в качестве $\text{disp}(p)$ возьмем такое значение d , при котором достигается минимум стоимостной функции $c(p, d)$, т. е.

$$\text{disp}(p) = \underset{d \leq D}{\text{argmin}} c(p, d),$$

где D — максимально допустимая диспарантность.

Пусть W и H — ширина и высота изображения соответственно; s — размер сканирующего окна. Тогда вычислительная сложность построения карты диспарантности описанным способом составит $O(WHs^2D)$.

3 Ускоренный алгоритм стереосопоставления

Нетрудно понять, что вспомогательные коэффициенты позволяют оценить вероятность принадлежности двух пикселей одной поверхности сцены. Более грубую оценку принадлежности двух пикселей одной поверхности дает сегментация изображения [2]. Сегментация и вспомогательные коэффициенты могут быть скомбинированы, что позволит

снизить вычислительную сложность сопоставления изображений.

Пусть $S_1 \cup S_2 \cup S_3 \cup \dots \cup S_m$ — сегментация левого изображения. Опорной точкой pivot_i сегмента S_i будем называть его медиану как упорядоченного по значениям координат списка точек. Цветом сегмента i будем называть средневзвешенный вспомогательными коэффициентами цвет составляющих его пикселей:

$$C_i = \frac{\sum_{p \in S_i} w(p, \text{pivot}_i) I_L(p)}{\sum_{p \in S_i} w(p, \text{pivot}_i)},$$

где $I_L(p)$ — цвет пикселя p на изображении.

На данном этапе предполагается, что такой вариант использования коэффициентов принадлежности не приведет к значительному снижению качества по сравнению с оригинальным алгоритмом. Параметр γ в выражении (1) должен быть согласован с уровнем сегментации, иначе теряется смысл применения вспомогательных коэффициентов — значение C_i будет близко к среднему цвету сегмента. Высокому уровню сегментации соответствует малое значение γ и наоборот. Рисунок 1 иллюстрирует влияние параметра γ на значения вспомогательных коэффициентов.

Пусть $s(p)$ — номер сегмента, которому принадлежит пиксель p , тогда в качестве стоимостной функции будем рассматривать величину

$$c(p, d) = \sum_{q \in \text{Win}_p} \|C_{s(q)} - I_R(q_x - d, q_y)\|_2, \quad (2)$$

где $I_R(x, y)$ — цвет пикселя (x, y) на правом изображении.

Используя стоимостную функцию (2), можно значительно сократить вычислительные затраты на построение карты диспарантности — стоимостная функция для каждой позиции сканирующего окна может быть вычислена за $O(1)$ с помощью частичных сумм.

Рассмотрим метод частичных сумм. Имеется матрица $A = (a_{i,j})$, состоящая из $n \times m$ элементов, и необходимо многократно вычислять сумму подматриц этой матрицы. Пусть

$$\text{ps}_{r,c} = \sum_{i=1}^r \sum_{j=1}^c a_{i,j}.$$

Используя принцип динамического программирования, величину ps можно рассчитать за время порядка $O(nm)$ согласно соотношению:

$$\text{ps}_{r,c} = \text{ps}_{r-1,c} + \text{ps}_{r,c-1} - \text{ps}_{r-1,c-1} + a_{r,c}. \quad (3)$$

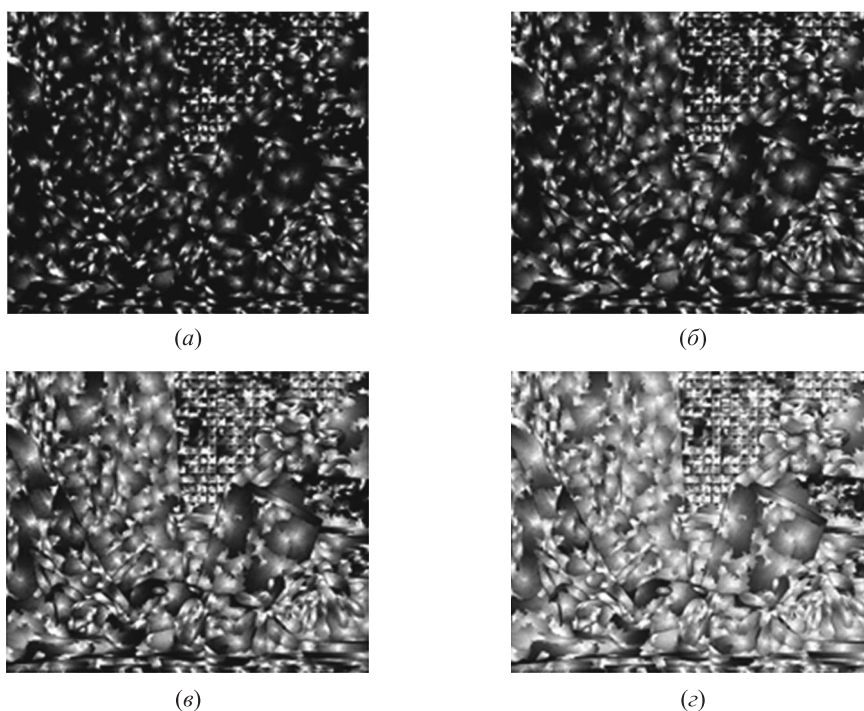


Рис. 1 Вспомогательные коэффициенты при различных значениях γ : (а) $\gamma = 10$; (б) 15; (в) 30; (д) $\gamma = 50$

Сумму на произвольной подматрице $(r_1, c_1) - (r_2, c_2)$ можно вычислить по формуле включения-исключения:

$$\begin{aligned} \text{sum}(r_1, c_1, r_2, c_2) = \\ = \text{ps}_{r_2, c_2} - \text{ps}_{r_2, c_1-1} - \text{ps}_{r_1-1, c_2} + \text{ps}_{r_1-1, c_1-1}. \end{aligned} \quad (4)$$

Воспользуемся методом частичных сумм для вычисления стоимостной функции $c(p, d)$. Пусть $\text{sum}_d(x_1, y_1, x_2, y_2)$ — суммарная стоимость диспарантности d для всех пикселей прямоугольника $(x_1, y_1) - (x_2, y_2)$, т. е.

$$\begin{aligned} \text{sum}_d(x_1, y_1, x_2, y_2) = \\ = \sum_{i=x_1}^{x_2} \sum_{j=y_1}^{y_2} \|C_{s(i,j)} - I_R(x-d, y)\|_2, \end{aligned}$$

где $s(i, j)$ — номер сегмента, которому принадлежит пиксель (i, j) .

Обозначив

$$\text{ps}_d(x, y) = \sum_{i \leq x} \sum_{j \leq y} \|C_{s(i,j)} - I_R(x-d, y)\|_2,$$

согласно выражению (4) имеем:

$$\begin{aligned} \text{sum}_d(x_1, y_1, x_2, y_2) = \\ = \text{ps}_d(x_2, y_2) - \text{ps}_d(x_2 - 1, y_1 - 1) - \\ - \text{ps}_d(x_1 - 1, y_2 - 1) + \text{ps}_d(x_1 - 1, y_1 - 1). \end{aligned}$$

С помощью выражения (3) величину ps_d можно вычислить за время порядка $O(WH)$. Значение стоимостной функции для сканирующего окна размером $s = 2h + 1$ выражается через sum_d :

$$c(p, d) = \text{sum}_d(p_x - h, p_y - h, p_x + h, p_y + h).$$

Таким образом, стоимостная функция для всевозможных позиций сканирующего окна при фиксированной диспарантности рассчитывается за время $O(WH)$; следовательно, вычислительная сложность построения карты диспарантности есть $O(WHD)$.

Для определения ошибочных значений диспарантности воспользуемся стандартным критерием согласованности. Пусть $\text{Disp}_{L,R}$ — карта диспарантности для левого изображения относительно правого, а $\text{Disp}_{R,L}$ — карта диспарантности для правого изображения относительно левого. В соответствии с приведенным ранее определением, значения диспарантности для правого изображения должны быть неположительными.

Пусть $\text{Disp}_{L,R}(x, y)$ и $\text{Disp}_{R,L}(x', y)$, где $x' = x - \text{Disp}_{L,R}$, являются корректными значениями диспарантности. Тогда выполняется равенство:

$$\text{Disp}_{L,R}(x, y) = -\text{Disp}_{R,L}(x', y).$$

На основании этого равенства будем производить фильтрацию карты диспарантности: если равенство для некоторого пикселя не выполняется, будем считать, что его диспарантность не установлена.

Таблица 1 Результаты расчета геодезических расстояний алгоритмом Vorgefors

Число итераций	Средняя ошибка	Максимальная ошибка	Время расчета, с	Время расчета алгоритмом Дейкстры, с
2	0,0022	0,35	0,22	0,15
3	$4,5 \cdot 10^{-4}$	0,15	0,32	
5	$4 \cdot 10^{-5}$	0,057	0,53	
10	10^{-8}	10^{-4}	1,05	

Ввиду того что при вычислении диспарантности задействуется некоторая окрестность пикселя, вполне вероятным является случай незначительного рассогласования значений $\text{Disp}_{L,R}$ и $\text{Disp}_{R,L}$. Для корректной обработки таких случаев следует ослабить критерий согласованности до

$$|\text{Disp}_{L,R}(x, y) + \text{Disp}_{R,L}(x', y)| \leq 1.$$

4 Вычисление геодезических расстояний

В работе [1] для вычисления геодезических расстояний используется итерационный алгоритм Vorgefors [3], позволяющий найти приближение геодезического расстояния с заданной точностью. Вычислительная сложность алгоритма Vorgefors линейна относительно размера сегмента.

Для точного вычисления геодезических расстояний можно воспользоваться алгоритмом Дейкстры, который позволяет вычислить длины кратчайших путей от одной вершины графа до остальных. В качестве вершин графа будем использовать пиксели одного сегмента, а ребро между двумя вершинами будет существовать, если соответствующие им пиксели являются смежными в восьми направлениях. Стоимость ребра — евклидово расстояние между цветами пикселей в системе RGB. Очевидно, что длина кратчайшего пути в таком графе будет геодезическим расстоянием по определению.

Описанный граф обладает важным свойством: число ребер линейно зависит от числа вершин. Такое свойство дает возможность реализовать алгоритм Дейкстры с использованием двоичной кучи [4], снизив вычислительную сложность до $O(|V| \log |V|)$, где $|V|$ — число вершин графа, т. е. размер сегмента.

В ходе экспериментов было выявлено, что асимптотически лучшее быстродействие алгоритма Vorgefors проявляется лишь на областях достаточно большого размера, наличие которых в сегментации реальных снимков маловероятно. В табл. 1 представлены результаты расчета геодезических расстояний алгоритмом Vorgefors с разным числом итераций для изображения размером 450×375 .

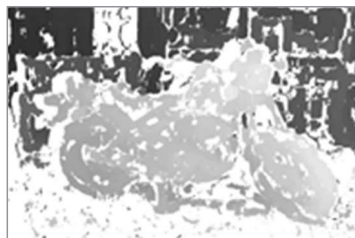
Таким образом, применение алгоритма Vorgefors не оправдано при расчете геодезических расстояний внутри сегмента.

5 Результаты

Алгоритм был реализован на языке C++ в соответствии с приведенным описанием. Реализация алгоритма доступна в открытом репозитории (<https://github.com/helgui/FastGSW>). Оценка результата стереосопоставления проводилась по методике, описанной в работе [5], с помощью предоставленного авторами работы программного обеспечения Middlebury Stereo Evaluation SDK. Набор входных данных [6] состоит из 14 стереопар различной конфигурации. В ходе экспериментов проводилось сравнение ускоренного алгоритма



(a)



(б)



(в)

Рис. 2 Результаты стереосопоставления: (a) левое изображение стереопары; (б) результат работы ускоренного алгоритма; (в) результат работы оригинального алгоритма

Таблица 2 Параметры запуска алгоритмов

Параметр	Значение для оригинального алгоритма	Значение для ускоренного алгоритма
Размер сканирующего окна	31	15
Коэффициент γ	10	5
Алгоритм сегментации	Не используется	Сегментация на основе графа [7]

Таблица 3 Результаты экспериментов

Название теста	Используемый алгоритм	Значения диспаратности с ошибкой более 2 единиц, %	Неустановленные значения диспаратности, %	Время выполнения, с
Adirondack	Оригинальный	9,98	66,72	562,19
	Ускоренный	10,26	60,44	6,39
Jadeplant	Оригинальный	4,05	40,03	1063,23
	Ускоренный	11,79	60,23	6,31
Motorcycle	Оригинальный	5,72	31,98	557,48
	Ускоренный	6,92	39,03	6,58
Motorcycle E	Оригинальный	14,46	82,6	556,19
	Ускоренный	8,31	84,86	6,77
Piano	Оригинальный	13,71	26,21	478,00
	Ускоренный	8,55	47,49	6,17
Piano L	Оригинальный	10,77	81,84	479,30
	Ускоренный	7,49	84,33	6,17
Pipes	Оригинальный	5,24	36,68	578,16
	Ускоренный	10,7	38,23	6,45
Playroom	Оригинальный	10,19	61,9	597,20
	Ускоренный	11,08	47,81	6,02
Playtable	Оригинальный	10,86	42,08	497,01
	Ускоренный	10,04	54,52	5,67
Playtable P	Оригинальный	7,15	36,33	498,76
	Ускоренный	9,98	54,4	5,83
Recycle	Оригинальный	8,78	31,9	491,86
	Ускоренный	8,4	53,29	6,35
Shelves	Оригинальный	16,64	42,83	475,90
	Ускоренный	10,19	58,21	6,52
Teddy	Оригинальный	3,11	30,12	233,55
	Ускоренный	7,1	34,48	3,36
Vintage	Оригинальный	13,06	32,35	1283,93
	Ускоренный	7,9	52,3	6,38

ма с оригинальным (рис. 2). Параметры запуска алгоритмов приведены в табл. 2. Следует заметить, что параметры для оригинального алгоритма заданы в соответствии с рекомендациями авторов. Результаты экспериментов сведены в табл. 3.

Главный критерий оценки качества — доля пикселей от общего числа, вычисленное значение диспаратности для которых отличается более чем на 2 единицы от истинного значения. Разница между средними значениями этого параметра для оригинального и ускоренного алгоритмов составила 0,33%.

При оценке также использовался дополнительный критерий — доля пикселей, отброшенных при фильтрации. Разница между средними значениями составила 8,4% в пользу оригинального алгоритма.

Из табл. 3 видно, что ускоренный алгоритм имеет многократное превосходство в производительности. Исходя из приведенных ранее асимптотических оценок сложности, нетрудно понять, что разница во времени расчета при увеличении размера сканирующего окна будет расти пропорционально квадрату этого размера.

6 Заключение

В данной работе была описана модификация алгоритма [1] и представлена его реализация. Результаты экспериментов позволили убедиться в том, что путем незначительной потери качества был получен существенный прирост производительности. Также очевидным преимуществом рассмотренного метода является эффективность интеграции с системами компьютерного зрения, которые уже используют сегментацию при анализе изображений.

Единственным недостатком предложенного подхода является меньшая плотность вычисленных карт диспаратности, что может свидетельствовать о недостаточной совместимости алгоритма с фильтрацией по критерию согласованности.

Литература

1. Hosni A., Bleyer M., Gelautz M., Rhemann C. Local stereo matching using geodesic support weights // 16th IEEE Conference (International) on Image Processing. — IEEE Press, 2009. P. 2093–2096.
2. Shapiro L. G., Stockman G. C. Computer vision. — Upper Saddle River, NJ, USA: Prentice Hall, 2001. 580 p.
3. Borgefors G. Distance transformations in digital images // Comput. Vision Graph. Image Proc., 1986. Vol. 34. No. 3. P. 344–371.
4. Кормен Т. Х., Лейзерсон Ч. И., Ривест Р. Л., Штайн К. Алгоритмы: построение и анализ / Пер. с англ. — 3-е изд. — М.: Вильямс, 2013. 1328 с. (Cormen T. H., Leiserson C. E., Rivest R. L., Stein C. Introduction to algorithms. — 3rd ed. — MIT Press, 2009. 1312 p.).
5. Scharstein D., Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms // IJCV, 2002. Vol. 47. No. 1/2/3. P. 7–42.
6. Scharstein D., Hirschmüller H., Kitajima Y., Krathwohl G., Nesci N., Wang X., Westling P. High-resolution stereo datasets with subpixel-accurate ground truth // German Conference on Pattern Recognition. — Springer, 2014. P. 31–42.
7. Felzenszwalb P. F., Huttenlocher D. P. Efficient graph-based image segmentation // IJCV, 2004. Vol. 59. No. 2. P. 167–181.

Поступила в редакцию 14.07.16

SPEEDED-UP STEREO MATCHING USING GEODESIC SUPPORT WEIGHTS

O. A. Yakovlev and A. V. Gasilov

Orel Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 137 Moskovskoe Sh., Orel 302025, Russian Federation

Abstract: In local stereo matching, the algorithms based on the adaptive support weights have good-quality results. This paper presents a modified version of the local matching with geodesic support. The proposed algorithm considerably reduces computation time at the cost of insignificant loss of quality that is experimentally confirmed with Middlebury Stereo Evaluation SDK. The key idea is to combine geodesic support weights and image segmentation for recoloring the reference image. This transformation makes it possible to use partial sums for matching cost computation.

Keywords: stereo matching; segmentation; geodesic support weights

DOI: 10.14357/19922264160313

References

1. Hosni, A., M. Bleyer, M. Gelautz, and C. Rhemann. 2009. Local stereo matching using geodesic support weights. *16th IEEE Conference (International) on Image Processing*. IEEE Press. 2093–2096.
2. Shapiro, L. G., and G. C. Stockman. 2001. *Computer vision*. Upper Saddle River, NJ: Prentice Hall. 580 p.
3. Borgefors, G. 1986. Distance transformations in digital images. *Comput. Vision Graph. Image Proc.* 34:344–371.
4. Cormen, T. H., Ch. E. Leiserson, R. L. Rivest, and C. Stein. 2009. *Introduction to algorithms*. 3rd ed. MIT Press. 1312 p.

5. Scharstein, D., and R. Szeliski. 2002. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *IJCV* 47(1-3):7–42.
6. Scharstein, D., H. Hirschmüller, Y. Kitajima, G. Krathwohl, N. Nesić, X. Wang, and P. Westling. 2014. High-resolution stereo datasets with subpixel-accurate ground truth. *German Conference on Pattern Recognition*. Springer. 31–42.
7. Felzenszwalb, P. F., and D. P. Huttenlocher. 2004. Efficient graph-based image segmentation. *IJCV* 59:167–181.

Received July 14, 2016

Contributors

Yakovlev Oleg A. (b. 1992) — research engineer, Orel Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 137 Moskovskoe Sh., Orel 302025, Russian Federation; maucra@gmail.com

Gasilov Artur V. (b. 1992) — research engineer, Orel Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 137 Moskovskoe Sh., Orel 302025, Russian Federation; gasilov.av@ya.ru

К ВОПРОСУ ОБ УМЕНЬШЕНИИ ОБЪЕМА ПОРЦИЙ УЧЕБНОГО МАТЕРИАЛА ПРИ ЭЛЕКТРОННОМ ОБУЧЕНИИ*

А. А. Федосеев¹

Аннотация: Предпринята попытка анализа электронного предъявления учебного материала как автоматизированного процесса. Проанализированы причины сокращения продолжительности видеолекций для массовых открытых онлайн-курсов, а также аналогичного сокращения необходимого времени работы с мультимедиа электронными образовательными ресурсами и параграфами электронных учебников. Показано, что причиной для таких сокращений является не столько сама продолжительность, сколько объем предъявляемого учебного материала, который может быть усвоен за один сеанс. Для определения пределов этого объема количество предъявляемой информации, измеренное в новых понятиях и связанных с ними уже усвоенных понятий-линков, сравнивается с предельным количеством элементов, обрабатываемых в оперативной памяти человека одновременно. В результате делается вывод о том, что за требованием сокращения продолжительности лекций стоит необходимое ограничение объема предъявляемой учебной информации. Учет этого обстоятельства позволил сформулировать понятие комплекта заданий и сделать предложение относительно процедуры автоматизированного обучения. Статья публикуется в порядке обсуждения.

Ключевые слова: электронные средства обучения; микрообучение; понятие; линк; «кошелек Миллера»; комплект заданий; автоматизированное обучение

DOI: 10.14357/19922264160314

1 Введение

Нормальная продолжительность лекции в российской высшей школе — два академических часа, что составляет 90 мин. Бывают сдвоенные лекции продолжительностью 180 мин. В других странах примерно то же самое, но предполагается некоторое пространство для маневра продолжительностью лекции.

Если вузовскую лекцию записать на видео, то, казалось бы, должна получиться основа учебного материала для одного из новых Массовых открытых онлайн-курсов (МООК), которые приобрели известность и популярность начиная с 2012 г. [1]. Чтобы лекционная основа превратилась в заготовку такого курса, следует добавить к ней механизм обратной связи для фиксации факта освоения материала лекции слушателями. Для этого в зависимости от целей и материала курса применяются различные методы.

Однако опыт применения МООК и электронных ресурсов (в том числе электронных учебников) показал, что продолжительные видеолекции или параграфы электронных учебников, требу-

ющие значительного по времени внимания, для дистанционного обучения не годятся. Эмпирически оказалось, что учебные материалы должны быть сформированы таким образом, чтобы работа с ними не превышала 15 мин, а лучше — еще меньше. Эта ситуация породила термин *microlearning* (микрообучение) [2] и метафоры «информация на один укус» и «внимание на чайную ложку» (в прямом переводе). При этом материал для микрообучения — это не порубленные на части длинные лекции, а самостоятельные, логически полные и связанные короткие видеоролики. Таким образом, считается, что для заочного обучения, основанного на информационно-коммуникационных технологиях (ИКТ), новые знания должны подаваться мелкими порциями (*microlearning*), работа с каждой из которых занимает не более нескольких минут.

Попытки объяснить необходимость предъявления новых знаний мелкими порциями сводятся к соображению, что при более длинных лекциях внимание слушателя рассеивается и материал перестает усваиваться. Так, исследование лекций почти пятисот МООК показало [3], что интерес к лекции резко уменьшается уже на шестой минуте просмотра.

*Работа выполнена в рамках Программы фундаментальных научных исследований в Российской Федерации на долгосрочный период (2013–2020 годы). Тема № 34.2. Когнитивные мультимедиа и интерактивность в образовании в условиях мобильного Интернета.

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

ра. Не опровергая это исследование, попробуем не согласиться с тем, что именно продолжительность лекции является критической в дистанционном обучении. Внимание студентов точно так же рассеивается и во время очной лекции в аудитории, однако девяностоминутный стандарт существует столетиями, и никто не собирается изменять его в угоду непоседливости студентов. И это при том, что отвлечение внимания во время очной лекции чревато полной утерей понимания изложения и невозможностью его восстановления, поскольку лекция существует только в момент ее прочтения. Что касается заочных учебных материалов, доставляемых средствами ИКТ, потеря внимания не наносит существенного вреда слушателю, поскольку существует возможность получить повторно ту часть материала, которая не была воспринята с первой попытки. По-видимому, дело не только в продолжительности лекции, а в чем-то еще.

2 Отличие автоматизированного процесса от «ручного»

Предъявление слушателю, студенту или учащемуся учебного материала с использованием средств ИКТ является автоматизированным процессом в отличие от изложения материала лектором или учителем. Учитель или лектор, излагая материал, самостоятельно определяет необходимую продолжительность цикла обучения, так же как и условия перехода к изложению следующего материала. Теоретически это означает, что учитель должен убедиться в готовности учащихся к дальнейшему учению. На практике это не всегда возможно. Учитель не в состоянии опросить всех учащихся, не все из них присутствовали на прошлом уроке, кое-кто, возможно, не выполнил домашнее задание. Тем не менее учитель хорошо знаком со своими подопечными и в состоянии принять решение о моменте, когда можно предъявлять следующий учебный материал в рамках существующего учебного плана. Ситуация в высшей школе несколько более свободна, поскольку предполагается (не всегда обоснованно) ответственное отношение студентов к учению.

Если какая-то операция осуществляется автоматизированно, в данном случае это операция предъявления средствами ИКТ учебного материала слушателям (учащимся, студентам) для усвоения, то естественно потребовать наличия автоматического сигнала о ее нормальном завершении. Этот принцип удачно реализован в Академии Салмана Хана (<https://www.khanacademy.org>). После корот-

кой видеолекции учащемуся предлагается выполнить ряд заданий. При возникновении сложностей предусмотрена возможность воспользоваться подсказками. Только после того, как все задания выполнены — что является автоматическим сигналом о завершении процесса восприятия и закрепления учебного материала, учащемуся становится доступной следующая порция видеоматериалов.

По договору с Академией Хана студенты Масачусетского технологического института должны были готовить учебные видеоматериалы для их использования Академией. В прочитанной для студентов в 2012 г. лекции (<https://www.youtube.com/watch?v=VA273i3z7Mk&nohtml5=False>) Салман Хан, в частности, настаивал, что ролики должны быть по возможности короткими. Некоторые лекции ему самому удалось сделать трехминутными, но, к сожалению, не все. То же самое относится и к заданиям, выполнение которых должно засвидетельствовать освоение материала микролекции. Задания также должны быть простыми и быстро-выполнимыми. Как видим, подход Академии Хана вполне согласуется с концепцией микрообучения.

У организаторов дистанционного курса нет иной возможности понять, воспринята ли видеолекция, как получить сигнал о том, что все задания, относящиеся к этой лекции, выполнены. Поскольку МООК относятся к высшему или дополнительному образованию, заранее предполагается большая мотивированность и ответственность слушателей. Что касается электронных ресурсов и учебников для общеобразовательной школы, то там ситуация другая. Полагаться на сознательность и ответственность учащихся не приходится. Поэтому и методы контроля усвоения знаний должны быть более строгими. Очевидно, что учитель может существовать (и на самом деле существует) в условиях, когда часть его учеников знает весь заданный материал, кто-то знает его частично, а некоторые вообще ничего не знают. Он примерно представляет уровень знаний каждого ученика и, поскольку их не так уж много, может своим индивидуальным вниманием (применяя различные педагогические приемы, в том числе формирующее обучение) до некоторой степени компенсировать разноразность знаний учащихся.

Если предъявление нового учебного материала передается технологиям, то ситуация изменяется и учащихся приходится подгонять (автоматизированно) под единый уровень. В противном случае от автоматизации не будет толка. Что с того, что ученикам предъявили некий электронный образовательный ресурс — кто-то посмотрел и изучил, а кто-то и компьютер не включал, — подчищать все придется учителю вручную.

Таким образом, автоматизация предъявления учебного материала неизбежно тянет за собой и автоматизированный процесс контроля усвоения этого материала. Иначе автоматизированный процесс оказывается незавершенным и восприятие учебного материала приходится проверять традиционными способами (контрольный опрос учащихся), т. е. учебный процесс возвращается к своей традиционной форме и смысл автоматизации пропадает.

3 Как человек воспринимает новую информацию

У каждого человека со временем складывается собственная система знаний. Как показано в [4], сначала закладываются первичные элементарные понятия, затем по мере поступления новой информации присутствующие в ней понятия (факты, образы, связи, категории — все что угодно) выражаются через уже имеющиеся, усвоенные элементы, после чего новая информация встраивается в систему знаний. Для объяснения этого процесса все уже устоявшиеся элементы, нужные для объяснения нового понятия, в [4] предлагается называть линками (в переводе — связями). Таким образом, каждое новое понятие оказывается связанным с некоторым количеством линков и в таком виде остается в долговременной памяти человека. Когда изучается новое понятие, например закон Ома, линками являются понятия напряжения, тока и сопротивления. Однако, когда закон Ома встроится в систему знаний, в памяти он будет представлен одним укрупненным понятием, определяющим весь закон. Если обладателю знания закона Ома понадобится использовать его для формирования какого-либо нового понятия, то этот закон будет использован как линк. Более крупным понятием-линком может быть вся электротехника и даже вся физика. Здесь важно, что при восприятии новой информации каждое понятие обрастает соответствующими линками и таким образом укрупняется. Кстати, при изучении закона Ома учащимся потребуются старый комплексный линк под названием «алгебраические преобразования», иначе никак не справиться с вычислением тока или, наоборот, напряжения. Естественно, что этот линк уже находится в памяти как единое целое, объединяющее все изученные алгебраические преобразования.

В западной литературе, например в [5], для объяснения аналогичных процессов используется понятие чанка (chunk — кусок, ломоть).

Особенности восприятия информации человеком в соответствии с открытием, сделанным

Джорджем Миллером [6] (остроумно названным «кошелек Миллера»), не позволяют обрабатывать в оперативной памяти одновременно более семи плюс-минус двух элементов. Таким образом, количество элементов предназначенного для усвоения учебного материала должно быть в пределах семи (плюс-минус два). В [7] на большом практическом материале показано, что количество элементов, предъявляемых в новом учебном материале, должно быть не менее трех и не более пяти. Если количество таких элементов три и менее, то интерес слушателей не пробуждается, поскольку они воспринимают материал слишком простым и потому не заслуживающим внимания. Если число элементов учебного материала более пяти, то интерес учащихся пропадает уже из-за того, что они теряют смысл предъявляемого материала и не понимают его. Очевидно, что количество одновременно обрабатываемых элементов в диапазоне от трех до пяти вполне согласуется с законом Миллера.

Отождествляя элементы обрабатываемой оперативной памятью человека информации с понятиями и линками, можно заключить, что эффективное восприятие может произойти, если их количество в новом учебном материале не менее трех и не более пяти. Проблема в том, что в понятиях и линках никто учебный материал не измеряет. Интуитивно отмечается, что более короткий материал усваивается успешнее, причем проверка усвоения осуществляется сравнительно просто. Сколько понятий и связанных с ними линков можно ввести на лекции продолжительностью 3 мин и сколько — за 90 мин? И далее: как убедиться, что весь 90-минутный материал усвоен? Каков объем выполненных заданий должен убедить в том, что материал на самом деле усвоен?

Таким образом, теперь можно быть уверенным, что дело не только в регулировании продолжительности лекции, но и в ограничении объема предъявляемой информации.

4 Понятие комплекта заданий

Следующим шагом на пути автоматизации учебного процесса становится предъявление учащемуся таких заданий, выполнение которых гарантированно свидетельствовало бы о полном усвоении этого материала. Как упомянуто ранее, рекомендации на этот счет касаются только количества и сложности заданий: заданий должно быть по возможности много, но они должны быть простыми. Попробуем связать это с теми понятиями и линками [4], количество которых определяет предельный объем учебного материала. В примере про закон Ома есть

новое понятие — закон Ома, есть три линка: напряжение, ток и сопротивление — и есть комплексный линк «алгебраические преобразования». Всего — пять. Никакие параллельные и последовательные соединения, электродвижущая сила и полное сопротивление цепи, а также удельное сопротивление и поперечное сечение проводника не должны входить в материал о законе Ома, поскольку это дополнительные понятия и им место в дальнейших порциях учебного материала, которые будут предьявляться позже.

Какими должны быть задания, чтобы полностью удостовериться в усвоении материала «закон Ома»? Во-первых, это вопросы на понимание каждого нового понятия. Новое понятие одно: закон Ома. Во-вторых, вопросы на понимание действия каждого уже известного понятия (линка) в законе Ома. Таких понятий три: напряжение, ток и сопротивление. У алгебраических преобразований нет специфических взаимодействий с законом Ома. Этот линк проявит себя, когда понадобится составлять формулы для вычислений. В-третьих, вопросы на понимание взаимодействия каждой пары величин в рамках изучаемого закона. Таких пар три: ток и напряжение, ток и сопротивление, напряжение и сопротивление. В-четвертых, задания на вычисление каждой величины при известных двух других. Их тоже три. В-пятых, определение зависимостей каждой величины от изменений двух других. Их может быть шесть, если одна из независимых величин является аргументом, а вторая — параметром. Итого в данном примере насчитывается минимум 16 заданий пяти различных типов. Почему минимум? Потому что меньше нельзя: не все аспекты порции знаний будут проверены. Абольше 16 заданий вполне может быть. Например, чтобы проверить усвоение новых понятий, может понадобиться более одного вопроса на понимание для каждого понятия. Задачи на вычисления могут быть сформулированы, что называется, «в лоб», а могут иметь завуалированную структуру, с тем чтобы учащийся догадался, как решить задачу.

Таким образом, опираясь на использованные в порции учебного материала понятия и линки, можно сформировать некоторый набор заданий, которыми можно проверить усвоение всех аспектов заключенного в этой порции знания. При этом рекомендации оказываются соблюденными: количество заданий существенно превышает обычную норму. Так, количество вопросов для самопроверки, размещаемых после параграфа с учебным материалом, как правило, не превышает трех—четырёх. Число задач, задаваемых на дом, не превышает пяти—шести. И это при том, что количество информации в параграфах учебников не нормировано

понятиями и связанными с ними линками и существенно превышает рекомендуемый уровень.

Будем этот необходимый объем заданий, определенный по понятиям и линкам нового учебного материала, называть комплектом. Каждой порции учебного материала соответствует свой комплект заданий.

Количество необходимых заданий для выявления полного усвоения материала — еще один довод в пользу мелких порций электронных лекций, электронных образовательных мультимедиа-ресурсов и параграфов электронных учебников. При увеличении числа понятий и линков учебного материала стремительно возрастает размер необходимого комплекта заданий, поскольку необходимо проверять понимание взаимодействия каждого с каждым.

5 Автоматизированная диагностика

Поскольку комплект заданий охватывает проверку усвоения всех аспектов соответствующего учебного материала, то возникает возможность автоматизированной диагностики возникших пробелов в знании. Таким образом, практическое применение понятия комплекта заданий позволяет осуществлять автоматизированную диагностику пробелов в знаниях учащихся. Поскольку заданиями охвачены все аспекты учебного материала по возможности с исчерпывающей полнотой, невыполнение отдельных заданий с очевидностью указывает на те разделы материала, к которым они относятся.

6 Автоматизированное обучение

Согласно [8], обучение — это процесс, при котором исправляются ошибки восприятия учащихся, не позволившие им выполнить задания с первой попытки. В этом смысле обучение является индивидуальным и итеративным процессом, в котором на каждой итерации в соответствии с допущенными учащимися ошибками каждый из них должен получить новый учебный материал, специально разработанный для устранения его пробелов в восприятии. После проработки этого материала вновь совершается попытка выполнить задания. Если задания оказываются снова не выполненными в какой-то части, то должна осуществиться следующая итерация создания нового учебного материала и предьявления его учащемуся для проработки. Этот процесс не применяется в полной

мере на практике в учебных заведениях, поскольку в них не предусмотрено систематических индивидуальных занятий. В какой-то степени эти положения теории отрабатываются репетиторами или родственниками учащихся.

Разработчик образовательного ресурса или электронного учебника, доставляемого средствами ИКТ, имеет возможность заранее предусмотреть и заготовить специальные дополнительные учебные материалы, направленные на преодоление непонимания в выполнении заданий. Рассмотрим для примера, какие должны быть заготовлены корректирующие материалы по закону Ома. Во-первых, это материалы, направленные на формирование правильного понимания каждого из участвующих в законе Ома линков. Их три. Это напряжение, ток и сопротивление. Во-вторых, должны также присутствовать материалы, корректирующие неправильное понимание пар линков. Их тоже три. Неправильное понимание самого закона Ома также должно исправляться корректирующим материалом. И наконец, в-третьих, неправильное использование алгебраических преобразований должно привести к отсылке ученика к соответствующему разделу математики. Итого получилось семь корректирующих материалов. Надо сказать, что в той или иной степени эти разделы излагались в основном учебном материале. Речь идет только о более подробном их изложении для полной ликвидации неправильного понимания. Эти материалы должны быть предъявлены каждому учащемуся соответственно допущенным им ошибкам. После проработки специального материала учащийся вновь получает соответствующий комплект заданий. При выполнении всех заданий делается заключение об усвоении материала, а при невыполнении — повторяется процедура предъявления специального учебного материала, соответствующего новым допущенным ошибкам. По-видимому, для предотвращения возможности бесконечного цикла следует предусмотреть общение с учителем после нескольких циклов с одними и теми же типами ошибок.

По существу, описанная процедура является не чем иным, как автоматизированным обучением. А это означает, что электронная доставка учебных материалов мелкими порциями с выделением элементов информации, включая комплекты заданий, дает возможность не только убедиться в усвоении или неусвоении предъявленного материала, но и обеспечить приемлемый уровень его усвоения всеми учащимися.

7 Заключение

Анализ причин сокращения продолжительности учебных материалов, предъявляемых средствами ИКТ, показал, что дело скорее в объеме предъявляемой информации, которая может быть воспринята за один сеанс, чем в продолжительности как таковой. Организаторы MOOC, создатели электронных образовательных ресурсов и электронных учебников интуитивно уменьшают порции предъявляемой учебной информации, поскольку автоматизированный процесс не позволяет откладывать усвоение материала «на потом». Введение в практику электронного обучения регулирования количества понятий-линков и формирования комплектов заданий может позволить довести процедуры электронного обучения до полного усвоения материала всеми слушателями или учащимися в темпе учебного процесса и тем самым повысить эффективность и результативность методов дистанционного обучения.

Литература

1. Богданова Д. А. Большой прорыв: от открытых образовательных ресурсов — к Массовым Открытым Онлайн-курсам // Дистанционное и виртуальное обучение, 2013. № 4. С. 35–47.
2. Fernandez J. The microlearning trend: Accommodating cultural and cognitive shifts // Learning Solutions Magazine, 2014. December 1. <http://www.learningsolutionsmag.com/articles/1578/the-microlearning-trend-accommodating-cultural-and-cognitive-shifts>.
3. Guo P. J., Kim J., Rubin R. How video production affects student engagement: An empirical study of MOOC videos. — MIT Computer Science and Artificial Intelligence Laboratory, 2014. 10 p. <https://groups.csail.mit.edu/uid/other-pubs/las2014-pguo-engagement.pdf>.
4. Карпенко М. П. Телеобучение. — М.: СГА, 2008. 800 с.
5. Chase W. G., Simon H. A. Perception in chess // Cognitive Psychol., 1973. No. 4. P. 55–61.
6. Miller G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information // Psychol. Rev., 1956. Vol. 63. P. 81–97.
7. Петрова В. Метод 3-4-5, чтобы все запоминать! Освойте новую технологию запоминания. — Montreal: Accent Graphics Communications, 2014. 169 с.
8. Писарев В. Е., Писарева Т. Е. Теория педагогики. — Воронеж: Кварта, 2009. 611 с.

Поступила в редакцию 27.04.16

WHAT IS BEHIND THE CONCEPT OF “KNOWLEDGE IN SMALL PACKAGES”

A. A. Fedoseev

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

Abstract: An attempt has been made to analyze the electronic presentation of educational material as automated process. The reasons for the reduction of the length of video lectures for massive open online courses and other educational electronic resources as well as the requirement of reducing the time required for multimedia electronic educational resources and paragraphs of electronic tutor books are analyzed. It is shown that the reason for these requirements is not the resource duration, but rather the volume of the educational material that can be learned in one session. To determine the limits of this volume the amount of presented information, measured in terms of new concepts and related already learned concepts-links, was compared with the limited number of items processed in the human memory simultaneously. As a result, it is concluded that the requirement of reducing the length of lectures is necessary to limit the scope of the educational information presented. This circumstance has made it possible to formulate the concept of a task set and to make a proposal for automated training procedures. The paper is published in order to discuss this problem.

Keywords: e-learning tools; microlearning; concept; link; “Miller purse;” task set; automated training

DOI: 10.14357/19922264160314

Acknowledgments

The work was done under the Program of Fundamental Scientific Research in the Russian Federation for the long term (2013–2020). Subject No. 34.2. Cognitive multimedia and interactivity in education in the conditions of mobile Internet.

References

1. Bogdanova, D. A. 2013. Bol'shoy proryv: Ot otkrytykh obrazovatel'nykh resursov — k Massovym Otkrytym Onlaynovym Kursam [The big breakthrough: From open educational resources — to massive open online courses]. *Distantionnoe i virtual'noe obuchenie* [Distance and Virtual Learning] 4:35–47.
2. Fernandez, J. December 1, 2014. The microlearning trend: Accommodating cultural and cognitive shifts. *Learning Solutions Magazine*. Available at: <http://www.learningsolutionsmag.com/articles/1578/the-microlearning-trend-accommodating-cultural-and-cognitive-shifts> (accessed March 31, 2016).
3. Guo, P. J., J. Kim, and R. Rubin. 2014. How video production affects student engagement: An empirical study of MOOC videos. MIT Computer Science and Artificial Intelligence Laboratory. 10 p. Available at: <https://groups.csail.mit.edu/uid/other-pubs/las2014-pguo-engagement.pdf> (accessed March 31, 2016).
4. Karpenko, M. P. 2008. *Teleobuchenie* [Teleeducation]. Moscow: SGA. 800 p.
5. Chase, W. G., and H. A. Simon. 1973. Perception in chess. *Cognitive Psychol.* (4):55–61.
6. Miller, G. A. 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev.* 63:81–97.
7. Petrova, V. 2014. *Metod 3-4-5, chtoby vse zapominat'! Osvoyte novuyu tekhnologiyu zapominaniya* [Method 3-4-5 to remember everything! Master the new memory technology]. Montreal: Accent Graphics Communications. 169 p.
8. Pisarev, V. E., and T. E. Pisareva. 2009. *Teoriya pedagogiki* [Pedagogy theory]. Voronezh: KVARTA. 611 p.

Received April 27, 2016

Contributor

Fedoseev Andrei A. (b. 1946) — Candidate of Science (PhD) in technology, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilova Str., Moscow 119333, Russian Federation; a.fedoseev@ipiran.ru

ГУМАНИТАРНЫЕ АСПЕКТЫ ПРОБЛЕМЫ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

К. К. Колин¹

Аннотация: Анализируются гуманитарные аспекты проблемы информационной безопасности (ИБ), которая рассматривается как важнейший компонент национальной и глобальной безопасности. Показано, что в современных условиях становления глобального информационного общества и усиления геополитического противоборства в информационном пространстве ИБ государства, человека и общества становится глобальной проблемой дальнейшего развития цивилизации, при этом гуманитарные компоненты этой проблемы выдвигаются на первый план. Рассмотрена структура гуманитарных проблем ИБ и первоочередные меры по их решению в России.

Ключевые слова: глобальная безопасность; гуманитарные проблемы; информационная безопасность; информационная культура; информационная этика; национальная безопасность

DOI: 10.14357/19922264160315

1 Информационная безопасность как гуманитарная проблема

Исследования показывают, что обеспечение ИБ государства, человека и общества сегодня становится одной из глобальных и стратегически важных проблем дальнейшего развития цивилизации в XXI в., при этом на первый план выдвигаются гуманитарные аспекты этой проблемы, которые необходимо обязательно учитывать при ее анализе и решении [1–3]. Возрастание роли гуманитарных аспектов данной проблемы обусловлено следующими тенденциями развития современного общества:

1. Процесс информатизации общества принял глобальный характер и сегодня охватывает практически все страны и регионы мира, при этом новые средства и технологии для работы с информацией получают массовое распространение и становятся атрибутами профессиональной и бытовой культуры для все большей части населения. Их использование повышает качество жизни, дает существенную экономию социального времени, создает новые стереотипы поведения и общения миллионов людей, изменяет их традиционные представления о личном и общественном богатстве [4] и даже о пространстве и времени.
2. По оценкам ряда специалистов, человечество вступило в информационную эпоху своего развития [5], оно активно формирует новую среду обитания и в результате этого само изменяется вместе с этой средой, оказывающей на человека

существенно большее влияние, чем это ожидалось ранее [6].

3. Глобальная информатизация создает для государства, человека и общества не только новые возможности, но также и новые проблемы, одной из которых является проблема ИБ. Исследования показывают, что эта проблема является многоаспектной и комплексной, а ее гуманитарные аспекты недостаточно исследованы, не учитываются в принятой ООН стратегии устойчивого развития [7], а в системе образования на необходимом уровне не изучаются [8].

Гуманитарный аспект проблемы ИБ состоит в том, что именно человек является творцом всех информационных ресурсов, систем и технологий информационного общества. Поэтому их качество и безопасность использования во многом определяются качествами самого человека. При этом речь идет не только о надежности и эффективности работы этих средств и систем, но и об их воздействии на человека, общество и окружающую природу.

2 Онтологическая двойственность гуманитарных проблем информационной безопасности и их антропологические аспекты

Исследования показали, что отличительной особенностью гуманитарных проблем ИБ является

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

ся их *онтологическая двойственность*. Она состоит в том, что человек в этой проблеме выступает не только как *объект защиты* от внешних информационных угроз, но также и как основной *источник этих угроз* для своего внешнего окружения. Кроме того, в современных информационных системах различного назначения, сложность которых неуклонно возрастает, именно человек становится основным фактором риска для их безопасного функционирования. Эта особенность данной проблемы также является принципиально важной для ее понимания и исследования.

Необходимо отметить, что в последние годы гуманитарные аспекты проблемы ИБ стали все более заметно проявлять себя не только на социальном и психологическом, но также и на биологическом уровне природы человека. Так, например, исследования американских, немецких и российских ученых показали, что воздействие на человека интенсивных потоков информации, которые являются характерными для информационного общества, приводят к изменениям нейронной структуры головного мозга человека, которые существенным образом изменяют его интеллектуальные и психические способности, социальное поведение, коммуникабельность и самооценку своих поступков [6]. Это означает, что проблемы ИБ сегодня необходимо изучать комплексно, с учетом также и антропологических аспектов этих проблем.

Автору представляется, что этому должно содействовать формирование *информационной антропологии* — новой научной дисциплины, которая

начала изучаться в России с 2011 г. Структура предметной области этой дисциплины рассмотрена в работах [9, 10].

3 Структура гуманитарных проблем информационной безопасности

Структура основных гуманитарных проблем ИБ в сжатом виде представлена в табл. 1. В ней отражены четыре группы этих проблем, каждая из которых связана с определенным видом деятельности современного общества.

4 Социальные проблемы информационной безопасности

Информационная преступность. В числе социальных проблем ИБ проблема информационной преступности стала изучаться одной из первых. При этом она связывалась, главным образом, с проблемой несанкционированного доступа к информации, хранящейся и циркулирующей в компьютерных информационных системах. Эта проблема стала проявлять себя уже в начале 1990-х гг. в связи с развитием процесса информатизации общества и его распространением на финансово-экономическую сферу.

Для противодействия этой угрозе достаточно быстро стали создаваться различные системы ин-

Таблица 1 Структура гуманитарных проблем ИБ

Группа проблем	Краткое содержание проблемы
1. Геополитические проблемы	Технологии «мягкой силы» в геополитике [11, 12] Электронная слежка за политическими лидерами «Глобальное наблюдение» за населением Информационные и «гибридные» войны [13]
2. Социальные проблемы	Информационная преступность Информационное неравенство Манипуляции общественным сознанием [14] Виртуализация общества
3. Культурологические проблемы	Глобализация и культура Новая информационная культура общества Электронная культура Многоязычие в киберпространстве
4. Антропологические проблемы	Энергоинформационная безопасность Интеллектуальная безопасность Информационные факторы деструктивного поведения Информационные болезни Информационная видеэкология

формационной защиты компьютерных систем и сетей, которые широко используются и в настоящее время. Тем не менее информационная преступность остается актуальной проблемой и сейчас, причем наибольшую опасность представляют уже не столько атаки хакеров с целью хищения финансовых средств из банков, сколько несанкционированный доступ к конфиденциальной информации и персональным данным отдельных категорий граждан в компьютерных системах, их копирование и последующее распространение.

Проблема информационного неравенства. Анализ основных тенденций развития глобального процесса информатизации общества показал, что этот процесс создает для развития цивилизации не только новые возможности, но также и новые проблемы. Одной из них является **проблема информационного неравенства** [15].

Суть этой проблемы заключается в том, что в процессе становления информационного общества электронные информационные ресурсы, а также новые средства, сети и информационные технологии оказываются в различной степени доступными для отдельных людей, организаций, стран и регионов мирового сообщества. При этом те люди, организации, страны и регионы, которые оказываются способными эффективно использовать возможности новой информационной среды общества для своего интеллектуального развития и решения других проблем, получают существенные преимущества перед другими субъектами мирового сообщества, которые при этом вытесняются на обочину процесса развития цивилизации.

Так, например, объем продаж товаров и услуг через сеть Интернет еще в 2000 г. превысил сумму в 1 трлн долл. США. Однако основную долю прибыли от этих продаж получили лишь те страны, в которых эта сеть была в достаточной степени развитой и доступной для населения.

Что же касается современных средств информатики и новых информационных технологий, то их массовое использование создает беспрецедентные возможности не только для научно-технического, но и для социально-экономического развития общества. При этом формируется совершенно новый, информационный уклад жизни и производственной деятельности многих миллионов людей.

Системные исследования проблемы информационного неравенства проводятся в Институте проблем информатики РАН уже более 25 лет. Их результаты опубликованы в ряде статей и монографий [15–17] и неоднократно докладывались на международных конференциях. На основе этих результатов в 1997 г. в ЮНЕСКО была направлена аналитическая записка, в которой была представле-

на российская концепция трактовки содержания проблемы информационного неравенства как новой комплексной проблемы глобального масштаба. В ней было показано, что принятый в тот период времени инструментально-технологический подход к этой проблеме является недостаточным, так как он не учитывает целого ряда важных факторов гуманитарного характера. В их числе такие факторы, как уровень информационной, в том числе лингвистической, культуры человека и общества, а также уровень их общей образованности, который в значительной мере определяет мотивацию активной деятельности людей в новом информационном пространстве.

Дальнейшее развитие процесса информатизации общества показало, что эта концепция является более адекватной реальности, и поэтому она сегодня находит все больше сторонников как в России, так и в других странах. Так, например, если в 1997 г., когда в ежегодном докладе Программы развития ООН было введено понятие «информационной бедности», она определялась исходя из возможностей доступа людей к современным информационно-телекоммуникационным технологиям, то в 2005 г., на втором этапе Международной встречи по проблемам глобального информационного общества в Тунисе, эта проблема трактовалась уже не как проблема «цифрового разрыва» (*digital divide*), а именно как *глобальная проблема информационного неравенства*, с учетом указанных выше ее гуманитарных аспектов.

В работе [16] показано, что в структуре этой проблемы целесообразно различать следующие три основных аспекта.

1. *Личностно-социальный аспект*, который связан с проблемой социальной адаптации человека в новой, быстро изменяющейся информационной среде. Именно здесь возникает новая форма социального неравенства людей — *информационное неравенство*. Снизить остроту этой проблемы призвана перспективная система образования, которая должна предоставить возможность всем членам общества получать необходимые знания и умения, для того чтобы правильно ориентироваться в новом информационном пространстве и эффективно использовать его возможности.
2. *Социально-экономический аспект*, который связан с национальной политикой той или иной страны в области развития информационной среды отдельных регионов и страны в целом, их информационной инфраструктуры, средств и методов доступа к информационным ресурсам и информационным коммуникациям,

а также в области развития и практического использования информационных технологий и информационного законодательства. Решение этих проблем должно являться одним из важнейших направлений государственной политики в научно-технической, экономической и социальной сферах современного общества.

3. *Геополитический аспект*, который связан с неравномерностью развития процесса информатизации в различных странах и регионах мира, что объясняется не только различиями в научно-техническом и экономическом потенциалах этих стран, но также и уровнем развития образования в этих странах, а также степенью понимания их политическими лидерами основных тенденций и закономерностей современного этапа развития цивилизации.

При изучении проблемы информационного неравенства в контексте задач обеспечения ИБ необходимо учитывать, что процесс информатизации общества оказывает на него как позитивное, так и негативное воздействие. С одной стороны, он повышает эффективность общественного производства и содействует созданию новых рабочих мест, в том числе для людей с ограниченными возможностями, повышает качество жизни населения. Но, с другой стороны, появление все более сложной информационной техники и технологий, электронных офисов и роботизированных производств требует от людей более высокого уровня квалификации и интеллекта. А поскольку система образования во многих странах не обеспечивает этих требований, происходит дальнейшее социальное расслоение общества, которое усиливает в нем социальную напряженность. Именно поэтому проблема информационного неравенства и должна сегодня квалифицироваться как одна из актуальных глобальных проблем, тесно связанных с обеспечением национальной и глобальной безопасности.

Для решения этой проблемы необходимо проведение адекватной государственной и международной политики в области развития информационной инфраструктуры общества, в правовой сфере, а также в сфере образования и культуры. Во многих странах это сегодня осуществляется на уровне целевых национальных программ развития информационного общества.

Проблема виртуализации общества. В последние годы в обществе стала все более заметной принципиально новая тенденция социальных изменений, которая получила название *виртуализации общества*. Суть ее заключается в том, что во многих жизненно важных сферах общества — в экономике,

политике, культуре, науке и образовании — происходит замещение реальных вещей и действий их симулякрами — искусственными образами, которые являются лишь символами этих вещей и действий [18–20].

Другими словами, современное человечество активно формирует вокруг себя новый, иллюзорный мир символов, который существует параллельно с реальным физическим миром и становится такой же неотъемлемой частью нашего бытия, как и физическая реальность.

Казалось бы, ну и что здесь плохого? Ведь на то и дано природой человеку сознание и развитое воображение, чтобы он мог при помощи этих двух своих особенных качеств моделировать процессы реального мира и таким образом лучше познавать этот мир, прогнозировать возможное развитие в нем различных процессов.

Оказывается, все гораздо сложнее. Погружаясь все глубже в мир виртуальности, человек не только подменяет реальные вещи и действия их образами и символами, но также и *формирует новые ценности*, которые затем оказывают влияние на него самого. А это уже принципиально новый социально-психологический феномен, и, как показывают исследования, его прогнозируемые последствия далеко неоднозначны.

Понятие виртуальности. Термин «виртуальный» происходит от латинского слова *virtualis* — возможный, вероятный, т. е. такой, который может проявиться при определенных условиях, но реально не существует [20].

В современном русском языке понятие «виртуальный» имеет несколько смысловых значений. Сначала это понятие использовали физики для обозначения элементарных частиц, имеющих очень малое время существования. Затем этот термин стал проникать на страницы научной и популярной литературы для обозначения искусственной реальности, создаваемой в сознании человека при помощи новейших средств компьютерной техники и кибернетических систем. Эта искусственная реальность и получила название *виртуальной реальности*.

Однако в данной работе речь идет о совсем другом феномене, который напрямую не связан с компьютерной техникой и кибернетическими устройствами. Имеется в виду то новое явление общественной жизни, которое проявляется в устойчивой тенденции отхода все большего числа людей от традиционных условий своего существования, основанных на личном общении с другими людьми. Оно подменяется принципиально новыми процессами информационных коммуникаций, где присутствуют лишь символы и образы реального мира,

которые постепенно заменяют человеку этот мир и все больше изолируют его от этого мира.

Виртуализация общества как глобальный процесс.

Феномен виртуализации общества стал объектом внимания ученых совсем недавно, не более 15 лет тому назад. Попытки его анализа практически одновременно предприняли А. Бюль и М. Поэту в Германии, а также канадские ученые М. Вейнстен и А. Крокер. В России одним из первых эту проблему стал изучать социолог из Санкт-Петербургского государственного университета Д. В. Иванов [18]. Результаты исследований показали, что здесь мы имеем дело с принципиально новым процессом глобального масштаба, который отражает новые трансформации в современном обществе. Эти трансформации еще мало изучены, но уже сегодня понятно, что они имеют достаточно серьезные последствия.

Каковы же причины возникновения процесса виртуализации общества? На этот счет сегодня существуют различные точки зрения. Западные ученые эти причины связывают в основном с развитием процессов информатизации общества и все более широким распространением новых информационных и телекоммуникационных технологий. Так, например, согласно мнению Бюля, виртуализация общества представляет собой технический процесс создания своеобразного *виртуального общества*, которое существует как бы «параллельно» с реальным обществом, не оказывая при этом на него существенного влияния.

Принципиально иной позиции придерживается Иванов, который считает, что причины виртуализации общества находятся в нем самом и заключаются в *изменении социальной природы самого общества*. Что же касается информатизации, компьютеризации и виртуализации общества, то эти процессы являются следствиями, а не причинами вышеуказанных изменений. Именно поэтому виртуализация общества и должна рассматриваться как некая глубинная социальная тенденция трансформации самого общества, связанная с общими закономерностями его развития, а вовсе не как результат развития научно-технического прогресса.

В соответствии с этой точкой зрения, которую разделяет и автор настоящей работы, изучение процессов виртуализации общества и их возможных последствий является сегодня весьма актуальной проблемой. Ее решение позволило бы не только лучше понять существо и закономерности тех глобальных процессов, которые происходят сегодня в мировом сообществе, но также и выработать рациональную стратегию адаптации человека и общества к новым условиям их существования в XXI в., которые становятся все более динамичными.

5 Культурологические проблемы информационной безопасности

Глобализация и культура. Исследования показали, что процессы глобализации общества оказывают существенное влияние на его культуру [21]. Развитию процессов глобализации общества содействует его все более масштабная информатизация, которая несет за собой не только новые средства и технологии стран Запада, но также и их языки, манеру одеваться, стереотипы поведения и общения.

В работе [22] показано, что с информационной точки зрения процессы глобализации общества оказывают на него двоякое воздействие. С одной стороны, развитие информационных коммуникаций существенно повышает *информационную связанность* мирового сообщества, содействует распространению новых знаний и технологий, способов организации производства и борьбы с болезнями. И этот результат является позитивным с точки зрения перспектив дальнейшего безопасного развития цивилизации.

Но, с другой стороны, деградация национальных культур снижает уровень *культурного разнообразия* общества, делает его более однородным и, следовательно, менее приспособленным к противодействию глобальным вызовам и угрозам XXI в.

Кроме того, разрушаются духовные ценности национальных культур, а вместо них насаждаются новые ценности потребительского общества. Этот процесс является одной из глобальных угроз для безопасного развития цивилизации, что уже признается не только российскими, но и западными учеными [22, 23].

Новая информационная культура общества. Проблема формирования новой информационной культуры общества, которая должна быть адекватной условиям жизни и деятельности людей в новой информационной среде их обитания, была поставлена в России академиком А. П. Ершовым еще в 1988 г. [23]. Он показал, что эта проблема будет глобальной, стратегически важной и социально значимой для развития цивилизации в XXI в. Однако системные исследования этой проблемы начались лишь в 2011 г., когда в Германии на русском языке была издана первая монография, специально посвященная этой проблеме [24]. В ней было показано, что для комплексного изучения проблем становления и развития новой информационной культуры общества должна быть сформирована специальная научная дисциплина — *информационная культурология*.

Таблица 2 Культурологические проблемы ИБ

Группа проблем	Краткое содержание проблемы
1. Глобализация и культура	Деградация национальных культур Этнос и нация в культурологической перспективе Национальное единство в условиях глобализации Развитие человеческих ресурсов в информационном обществе
2. Человек в информационном обществе	Новая структура занятости населения Усиление технократии Новые формы информационного неравенства Урбанизация в информационном обществе
3. Языки в новом информационном пространстве	Информационная бедность и лингвистическая культура Сокращение мирового русскоязычного пространства Многоязычие в киберпространстве Технологии автоматизированного перевода текстов и речи
4. Электронная культура	Безопасность электронной информационной техники Массовое обучение пользователей Формирование культуры ИБ

В данной монографии была предложена структура предметной области этой дисциплины, рассмотрены ее основные задачи и перспективы развития, показана их связь с проблемами обеспечения ИБ.

В 2015 г. эта монография в существенно переработанном виде была издана и в России [25]. При этом культурологическим аспектам проблемы ИБ в ней посвящен отдельный раздел, включающий пять глав. Состав этих проблем представлен в табл. 2.

Таким образом, можно утверждать, что Россия сегодня является лидером в области изучения проблем информационной культуры в комплексной постановке с учетом взаимосвязи с проблемами ИБ.

6 Антропологические проблемы информационной безопасности

Энергоинформационная безопасность в информационном обществе. Современная промышленная и технологическая революция существенным образом изменили энергоинформационное поле нашей планеты. Мощные электростанции, крупные промышленные производства, высоковольтные линии электропередачи, городские здания и сооружения — все эти объекты создают вокруг себя достаточно интенсивные электромагнитные поля, которые постоянно окружают современного человека и воздействуют на его организм.

Развитие информационного общества усиливает это воздействие и делает его глобальным. Ведь средства телевидения и мобильной связи сегодня

имеются практически в каждой семье и регулярно используются как взрослыми, так и детьми, причем их количество и интенсивность использования продолжают возрастать.

Какое воздействие оказывает электромагнитное излучение этих средств на организм человека? Каков допустимый уровень этого воздействия на детей, взрослых, а также на зародышей, еще находящихся в утробе матери? Какими могут быть последствия этого воздействия? На все эти вопросы пока нет удовлетворительных ответов, так как данная проблема системно не изучается. А ведь она является глобальной и представляет серьезную угрозу не только для человека, но и для всей биосферы нашей планеты.

Так, например, одним из тревожных признаков является сокращение численности пчел, которое в последние годы наблюдается во многих странах, но причина его пока не выявлена. Возможно, это связано с развитием средств мобильной связи.

Впервые проблема энергоинформационной безопасности была поставлена автором настоящей статьи в работе [16]. Эта публикация не привлекла к себе внимания специалистов в области глобальных экологических проблем, однако надеяться, что она сама собой решится, также нет оснований. Ведь подавляющая часть объектов энергетики и связи находится сегодня в собственности частных компаний, заинтересованных главным образом в получении прибыли, а не в решении проблем энергоинформационной безопасности человека и общества.

Поколение Next и новая проблема интеллектуальной безопасности. Исследования последних лет показывают, что информатизация общества оказывает сильное воздействие не только на социальные

аспекты повседневной жизни и профессиональной деятельности людей, но также на их психику, образ мышления и даже на развитие головного мозга. Так, например, американские психологи Г. Смолл и Г. Ворган в своей монографии [6] утверждают, что новое поколение людей информационной эпохи, которое уже получило название «поколения Next», будет обладать совсем другой психикой и образом мышления по сравнению с людьми старшего поколения. При этом весьма вероятно, что нейронная структура головного мозга у этих людей будет отличаться от той, которая существует в настоящее время.

Свою гипотезу авторы указанной монографии аргументируют следующим образом. Согласно теории эволюции Чарльза Дарвина, развитие головного мозга человека происходит в результате его приспособления к изменениям окружающей среды. Эта общая закономерность действует и сегодня, в условиях стремительного развития процесса информатизации и формирования глобального информационного общества. А поскольку наиболее радикальные и быстрые перемены происходят именно в информационной сфере общества, мозг человека начинает приспосабливаться к этим изменениям путем адекватных изменений в организации своей структуры. И этот феномен является вполне закономерным. Вероятнее всего, в ближайшие годы он будет только нарастать.

Проблема здесь заключается в том, что указанные изменения в природе человека происходят слишком быстро, на протяжении жизни одного поколения людей. Для психологов это оказалось полной неожиданностью. Ведь таких радикальных изменений природа человека не испытывала никогда, а по своей значимости они сопоставимы, пожалуй, лишь с феноменом появления членораздельной речи.

Но ведь и масштабы современной информационной революции также являются беспрецедентными в истории человечества. Их значимость и возможные последствия еще в необходимой мере не исследованы. Это нам еще предстоит сделать в будущем.

Отличительные черты людей эпохи Интернета. Всех нас удивляет, как быстро и легко дети осваивают современную достаточно сложную информационную технику и новые информационные технологии. Специалисты по возрастной психологии знают, что мозг ребенка является очень пластичным, поэтому дети легко осваивают и новые языки, и новую технику, и новые стереотипы поведения людей в информационном обществе. При этом у них вырабатываются совсем другие, отличные от традиционных, формы мыслительной деятель-

ности, обусловленные повседневным использованием новых информационных технологий. Их мозг становится способным к обработке больших объемов информации, а также к быстрой реакции на зрительные образы.

Развитию этих способностей в значительной мере содействует активное использование компьютерных поисковых систем, компьютерные игры, а также общение по электронной почте. Социологические исследования развития интеллектуального уровня людей показывают, что IQ среднего человека в XXI в. стремительно растет.

Вполне возможно, что новые информационные технологии развивают интеллект точно так же, как это делают головоломки, игра в шахматы и изучение новых языков. Наблюдения показывают, что люди, часто использующие Интернет, как правило, быстрее находят выход из сложных положений и в повседневной жизни. Ведь каждый день, отыскивая для себя в сети нужную информацию, они тренируют те мозговые центры, которые связаны с оперативным решением практических задач.

Однако с развитием интеллекта и логического мышления у нового поколения людей эпохи Интернет не так все однозначно. Здесь есть и достаточно серьезные негативные факторы.

Угроза психологического расслоения человечества в информационном обществе. Исследования показывают, что постоянное использование компьютерных информационных технологий влечет за собой не только положительные, но и отрицательные последствия для психики человека. Одно из них — это так называемое «клиповое мышление» [6].

Суть этого феномена состоит в том, что частое использование сети Интернет уменьшает способность человека к концентрации мысли, созерцанию и абстрактному мышлению. Его мозг начинает постепенно привыкать к получению информации в готовом виде, которую уже не нужно анализировать, поэтому и процесс мышления у таких людей становится фрагментарным, «клиповым».

Таким образом, вместо мыслителя человек превращается в своего рода сортировщика готовой информации, при этом те зоны мозга, которые отвечают за абстрактное мышление, постепенно деградируют, и в будущем, вполне возможно, они могут совсем атрофироваться. Как же он сможет решать те новые глобальные проблемы XXI в. [26], которые, как снежная лавина, нарастают уже сегодня? В этом, по мнению автора, и состоит суть новой глобальной проблемы *интеллектуальной безопасности* [15].

Тревогу вызывает тот факт, что указанные изменения психики чаще всего наблюдаются у молодого

го поколения людей, вырастающих в современную информационную эпоху. Так, например, в Японии, одной из наиболее информационно развитых стран мира, многие школьники младших классов сегодня не умеют считать в уме, так как вместо этого используют калькуляторы в своих смартфонах или компьютерах. На это уже обратили свое внимание японские преподаватели, которые специально заставляют таких школьников считать в уме и даже сдавать соответствующие экзамены.

Таким образом, на наших глазах вырастает новое поколение людей, которые будут обладать совсем другой психикой и другим типом мышления. Их отличительной чертой будет рассеянное внимание, когда человек следит за всем сразу, ни на чем не сосредотачиваясь. Они будут хуже нас общаться между собой в обычной, не компьютерной реальности, так как их мозг будет все больше утрачивать те базовые механизмы, которые управляют контактами с другими людьми.

Их память будет все меньше использоваться для запоминания фактографической и другой информации, так как ее «кибернетическими протезами» станут персональные компьютеры, смартфоны и электронные базы данных сети Интернет. Эти люди, вероятнее всего, будут запоминать не саму информацию, а метainформацию, т. е. информацию о том, в какой папке компьютерной памяти она хранится или же в какой электронной библиотеке ее можно найти.

Следует ожидать, что представители «поколения Next» будут еще меньше, чем наши современники, читать художественную литературу, в особенности классическую, историческую и научную. Зачем им это делать, если есть Интернет и Википедия? Театры, консерватории и музеи, скорее всего, им также будут неинтересны. Ведь с их содержанием можно будет познакомиться в электронной сети, не выходя из дома.

В информационном обществе важную роль играет *сетевое общение*, которое существенным образом расширяет возможность контактов с другими людьми, повышает их оперативность и экономит массу социального времени. Но ведь при этом возникает риск забыть о том, что на самом деле представляет собой дружба между людьми в реальном мире, для которой необходимо реальное общение.

Вполне возможно, что в информационном обществе появится также и *новый вид одиночества*. Это ситуация, когда телевизор и другие средства информационной техники выключены и человек остается один в уже мало привычном для него реальном мире. Ведь уже давно известно, что нигде люди не чувствуют себя так одинокими, как в большом го-

роде, когда они часто не знакомы даже с теми, кто живет в соседней квартире.

Информационные факторы деструктивного поведения человека. В числе новых направлений исследования антропологических аспектов проблем ИБ следует отметить работы российского композитора В. С. Дашкевича, где рассматривается влияние на деятельность головного мозга человека той акустической и, в частности, музыкальной среды, в которой он обитает. В них показано, что одной из причин повышения уровня деструктивности поведения людей в современном обществе, является его *музыкальная культура* [27].

Таким образом, музыкальная культура общества также должна стать одним из объектов тех перспективных исследований, которые должны проводиться в интересах изучения гуманитарных проблем ИБ.

7 Заключение

Доктрина и Стратегия информационной безопасности Российской Федерации. В настоящее время концептуальные основы ИБ России определяет Доктрина информационной безопасности РФ, принятая еще в 2000 г. По своему содержанию она представляет собой развернутый и достаточно хорошо продуманный документ, многие положения которого являются актуальными и в настоящее время. Тем не менее он требует корректировки, так как за последние годы ситуация в данной области изменилась в худшую для России сторону и, кроме того, появились новые информационные вызовы и угрозы как национального, так и глобального масштаба. Поэтому еще в декабре 2015 г. Советом Безопасности РФ был подготовлен проект новой Доктрины информационной безопасности, утверждение которой ожидается в 2016 г. Однако до сих пор текст этого документа не опубликован и на общественном уровне еще не обсуждался.

Представляется, что для эффективного противодействия комплексу современных угроз для ИБ России необходимо также разработать и принять *Стратегию информационной безопасности РФ на период до 2030 г.* В ней должны быть определены конкретные задачи в этой области, сроки их решения и количественные показатели, необходимое правовое, организационное и кадровое обеспечение.

Отметим, что вопрос о необходимости разработки и принятия Стратегии информационной безопасности России на среднесрочный период неоднократно ставился Институтом проблем информатики РАН как в научных публикациях, так и в Аналитических материалах, которые Российская академия наук представляла для включения

в ежегодный Доклад Президенту России «О состоянии национальной безопасности РФ и мерах по ее укреплению». Сегодня пришло время, когда Стратегия информационной безопасности России является крайне необходимой.

Необходимо также пересмотреть содержание научных дисциплин ВАК РФ по тематике ИБ, предусмотрев в них наиболее актуальные гуманитарные аспекты, некоторые из которых были рассмотрены в данной работе.

Литература

1. *Коллин К. К.* Информационная безопасность как гуманитарная проблема // Открытое образование, 2006. № 1(54). С. 86–93.
2. *Соколов И. А., Коллин К. К.* Развитие информационного общества в России и актуальные проблемы информационной безопасности // Информационное общество, 2009. № 4-5. С. 98–106.
3. *Коллин К. К.* Гуманитарные проблемы информационной безопасности // Информационные технологии, 2012. № 12. Приложение. С. 1–32.
4. *Тоффлер Э.* Революционное богатство. — М.: АСТ Москва, 2008. 569 с.
5. *Кастельс М.* Информационная эпоха: экономика, общество и культура. — М.: ГУ ВШЭ, 2000. 458 с.
6. *Смолл Г., Врган Г.* Мозг онлайн. Человек в эпоху Интернета. — М.: Колибри, 2011. 352 с.
7. *Коллин К. К.* Половинчатая стратегия: критический анализ новой Стратегии ООН в области устойчивого развития // Партнерство цивилизаций, 2016. № 1-2. С. 33–41.
8. *Соколов И. А., Коллин К. К.* Новый этап информатизации общества и актуальные проблемы образования // Информатика и её применения, 2008. Т. 2. Вып. 1. С. 67–76.
9. *Коллин К. К.* Информационная антропология: поколение Next и угроза психологического расслоения человечества в информационном обществе // Вестник Челябинской государственной академии культуры и искусств, 2011. № 4. С. 32–36.
10. *Коллин К. К.* Информационная антропология: предмет и задачи нового направления в науке и образовании // Вестник Кемеровского государственного университета культуры и искусств, 2011. № 17. С. 17–32.
11. *Смирнов А. И., Кохтюлина И. Н.* Глобальная безопасность и «мягкая сила 2.0»: вызовы и возможности для России. — М.: ВНИИГеосистем, 2012. 276 с.
12. *Шабалов М. П.* «Мягкая сила» в современной геополитике // Стратегические приоритеты, 2014. № 4. С. 27–43.
13. *Кошкин Р. П.* Россия и мир: новые приоритеты в геополитике. — М.: Стратегические приоритеты, 2015. 236 с.
14. *Кара-Мурза С. Г.* Манипуляции сознанием. — М.: Алгоритм, 2000. 688 с.
15. *Коллин К. К.* Информационное неравенство — новая проблема XXI века // Социология, социальность, современность. — М.: Союз, 1998. Вып. 5. С. 99–101.
16. *Коллин К. К.* Информационные проблемы социально-экономического развития общества. — М.: Союз, 1995. 72 с.
17. *Коллин К. К.* Глобальные проблемы информатизации: информационное неравенство // Alma mater (Вестник высшей школы), 2000. № 6. С. 27–32.
18. *Иванов Д. В.* Виртуализация общества. — СПб.: Петербургское востоковедение, 2000. 96 с.
19. *Коллин К. К.* Виртуализация общества — новая угроза для его стабильности // Синергетическая парадигма. Человек и общество в условиях нестабильности. — М.: РАГС, 2003. С. 449–462.
20. *Коллин К. К.* Виртуализация общества // Большая Российская энциклопедия, 2006. Т. 5. С. 370.
21. *Коллин К. К.* Глобализация и культура // Вестник Библиотечной Ассамблеи Евразии, 2004. № 1. С. 12–15.
22. *Коллин К. К.* Информатизация общества и глобализация. — Красноярск, СФУ, 2011. 52 с.
23. *Ершов А. П.* Информатизация: от компьютерной грамотности школьников к информационной культуре общества // Коммунист, 1988. № 2. С. 82–92.
24. *Коллин К. К., Урсул А. Д.* Информационная культурология: предмет и задачи нового научного направления. — Saarbruchen, Germany: LAP Lambert Academic Publishing, 2011. 249 с.
25. *Коллин К. К., Урсул А. Д.* Информация и культура. Введение в информационную культурологию. — М.: Стратегические приоритеты, 2015. 300 с.
26. *Коллин К. К.* Глобальные угрозы развитию цивилизации в XXI веке // Стратегические приоритеты, 2014. № 1. С. 6–30.
27. *Дашкевич В. С.* Великое культурное одичание. Арт-анализ. — М.: Russian Chess House, 2013. 717 с.

Поступила в редакцию 18.07.16

HUMANITARIAN ASPECTS OF INFORMATION SECURITY

K. K. Kolin

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

Abstract: The paper analyzes humanitarian aspects of information security, which is regarded as the most important component of national and global security. It is shown that in modern conditions, the formation of the global information society and strengthening of geopolitical confrontation in the information space of information security of state, individual, and society is becoming a global problem for the further development of civilization. The humanitarian component of this problem comes to the forefront. The structure of humanitarian problems of information security is described. Priority measures for their solution in Russia are proposed.

Keywords: global security; humanitarian issues; information security; information culture; information ethics; national security

DOI: 10.14357/19922264160315

References

1. Kolin, K. K. 2006. Informatsionnaya bezopasnost' kak gumanitarnaya problema [Information security as a humanitarian problem]. *Otkrytoe Obrazovanie* [Open Education] 1:86–93.
2. Sokolov, I. A., and K. K. Kolin. 2009. Razvitiye informatsionnogo obshchestva v Rossii i aktual'nye problemy informatsionnoy bezopasnosti [Information society development in Russia and problems of information security]. *Informatsionnoe Obshchestvo* [Information Society] 4-5:98–106.
3. Kolin, K. K. 2012. Gumanitarnye problemy informatsionnoy bezopasnosti [Since the humanitarian problems of information security]. *Informatsionnye Tekhnologii* [Information Technology] 12(App.):1–32.
4. Toffler, E. 2008. *Revolutsionnoe bogatstvo* [Revolutionary wealth]. Moscow: ACT Moskva. 569 p.
5. Kastel's, M. 2000. *Informatsionnaya epokha: Ekonomika, obshchestvo i kul'tura* [The information age: Economy, society, and culture]. Moscow: GU VShE. 458 p.
6. Smoll, G., and G. Vrgan. 2011. *Mozg onlayn. Chelovek v epokhu Interneta* [Brain online. People in the Internet age]. Moscow: Kolibri. 352 p.
7. Kolin, K. K. 2016. Polovinchataya strategiya: Kriticheskiy analiz novoy Strategii OON v oblasti ustoychivogo razvitiya [Since a half-hearted strategy: A critical analysis of the new UN Strategy on sustainable development]. *Partnerstvo Tsivilizatsiy* [Partnership of Civilizations] 1-2:33–41.
8. Sokolov, I. A., and K. K. Kolin. 2008. Novyy etap informatizatsii obshchestva i aktual'nye problemy obrazovaniya [The new stage of the society informatization and actual problems of education]. *Informatika i ee Primeneniya — Inform. Appl.* 2(1):67–76.
9. Kolin, K. K. 2011. Informatsionnaya antropologiya: Pokolenie Next i ugroza psikhologicheskogo rassloeniya chelovechestva v informatsionnom obshchestve [Information anthropology: The Next generation and the threat of psychological stratification of humanity in the information society]. *Vestnik Chelyabinskoy gosudarstvennoy akademii kul'tury i iskusstv* [Bulletin of the Chelyabinsk State Academy of Culture and Arts] 4:32–36.
10. Kolin, K. K. 2011. Informatsionnaya antropologiya: Predmet i zadachi novogo napravleniya v nauke i obrazovanii [Information anthropology: The subject and objectives of the new direction in science and education]. *Vestnik Kemerovskogo gosudarstvennogo un-ta kul'tury i iskusstv* [Bulletin of Kemerovo State University of Culture and Arts] 17:17–32.
11. Smirnov, A. I., and I. N. Kokhtyulina. 2012. Global'naya bezopasnost' i “myagkaya sila 2.0.” Vyzovy i vozmozhnosti dlya Rossii [Global security and “soft power 2.0.” Challenges and opportunities for Russia]. Moscow: VNIIGeosistem. 276 p.
12. *Shabalov, M. P.* 2014. “Myagkaya sila” v sovremennoy geopolitike [“Soft power” in contemporary geopolitics]. *Strategicheskie Prioritety* [Strategic Priorities] 4:27–43.
13. *Koshkin, R. P.* 2015. Rossiya i mir: Noveye prioritety v geopolitike [Russia and the world: New priorities in geopolitics]. Moscow: Strategicheskie prioritety [Strategic Priorities]. 236 p.
14. Kara-Murza, S. G. 2000. *Manipulyatsii soznaniem* [Manipulation of consciousness]. Moscow: Algoritm. 688 p.
15. Kolin, K. K. 1998. Informatsionnoe neravenstvo — novaya problema XXI veka [Information inequality is a new problem of the XXI century]. *Sotsiologiya, sotsial'nost', sovremennost'* [Sociology, sociality, modernity]. Moscow: Soyuz. 5:99–101.
16. Kolin, K. K. 1995. *Informatsionnye problemy sotsial'no-ekonomicheskogo razvitiya obshchestva* [Information problems of socio-economic development of society]. Moscow: Soyuz. 72 p.
17. Kolin, K. K. 2000. Global'nye problemy informatizatsii: Informatsionnoe neravenstvo [The global problem of informatization: Information inequality]. *Alma mater (Vestnik vysshneyshkoly)* [Alma Mater (Bulletin of High School)] 6:27–32.
18. Ivanov, D. V. 2000. Virtualizatsiya obshchestva [Virtualization of society]. Saint Petersburg: Peterburgskoe vos-tokovedenie. 96 p.

19. Kolin, K. K. 2003. Virtualizatsiya obshchestva – novaya ugroza dlya ego stabil'nosti [Since virtualization companies — a new threat to its stability]. *Sinergeticheskaya paradigma. Chelovek i obshchestvo v usloviyakh nestabil'nosti* [Synergetic paradigm. Man and society in conditions of instability]. Moscow: RAGS. 449–462.
20. Kolin, K. K. 2006. Virtualizatsiya obshchestva [Virtualization of society]. *Bol'shaya Rossiyskaya entsiklopediya* [Great Russian Encyclopedia]. 5:370.
21. Kolin, K. K. 2004. Globalizatsiya i kul'tura [Globalization and culture]. *Vestnik Bibliotechnoy Assamblei Evrazii* [Bulletin of the Library Assembly of Eurasia] 1:12–15.
22. Kolin, K. K. 2011. *Informatizatsiya obshchestva i globalizatsiya* [Information society and globalization]. Krasnoyarsk: SFU. 52 p.
23. Ershov, A. P. 1998. Informatizatsiya: Ot komp'yuternoy gramotnosti shkol'nikov k informatsionnoy kul'ture obshchestva [Informatization: From computer literacy to information culture society]. *Kommunist* [Communist] 2:82–92.
24. Kolin, K. K., and A. D. Ursul. 2011. *Informatsionnaya kul'turologiya: Predmet i zadachi novogo nauchnogo napravleniya* [Information studies: Subject and problems of a new scientific direction]. Saarbruchen, Germany: LAP Lambert Academic Publishing. 249 p.
25. Kolin, K. K., and A. D. Ursul. 2015. *Informatsiya i kul'tura. Vvedenie v informatsionnuyu kul'turologiyu* [Information and culture. Introduction to information studies]. Moscow: Strategicheskie priority. 300 p.
26. Kolin, K. K. 2014. Global'nye ugrozy razvitiyu tsivilizatsii v XXI veke [The global threat to the development of civilization in the XXI century]. *Strategicheskie Priority* [Strategic Priorities] 1:6–30.
27. Dashkevich, V. S. 2013. *Velikoe kul'turnoe odichanie. Art-analiz* [The Great cultural savagery. Art-analysis]. Moscow: Russian Chess House. 717 p.

Received July 18, 2016

Contributor

Kolin Konstantin K. (b. 1935) — Doctor of Science in technology, professor, Honored scientist of RF, principal scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; kolinkk@mail.ru

Андреев Сергей Дмитриевич (р. 1984) — кандидат технических наук, доцент Российского университета дружбы народов

Архипов Олег Петрович (р. 1948) — кандидат технических наук, директор Орловского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

Архипов Павел Олегович (р. 1979) — кандидат технических наук, старший научный сотрудник Орловского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

Гасилов Артур Владимирович (р. 1992) — инженер-исследователь Орловского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

Грушо Александр Александрович (р. 1946) — доктор физико-математических наук, профессор, заведующий лабораторией Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Грушо Николай Александрович (р. 1982) — кандидат физико-математических наук, старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Забейало Михаил Иванович (р. 1956) — кандидат физико-математических наук, доцент, заведующий лабораторией Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Захарова Татьяна Валерьевна (р. 1962) — кандидат физико-математических наук, старший преподаватель 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

Кириков Игорь Александрович (р. 1955) — кандидат технических наук, директор Калининградского филиала Федерального исследовательского центра

«Информатика и управление» Российской академии наук

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

Коллин Константин Константинович (р. 1935) — доктор технических наук, профессор, заслуженный деятель науки РФ, главный научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Кривенко Михаил Петрович (р. 1946) — доктор технических наук, профессор, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

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; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Кучерявый Евгений Андреевич (р. 1974) — кандидат технических наук, профессор Национального исследовательского университета Высшая школа экономики

Леонтьев Николай Дмитриевич (р. 1988) — аспирант кафедры математической статистики факультета вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова

Листопад Сергей Викторович (р. 1984) — кандидат технических наук, научный сотрудник Калининградского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

Наумов Валерий Арсентьевич (р. 1950) — кандидат физико-математических наук, научный руководитель Исследовательского института инноваций, г. Хельсинки, Финляндия

Омётов Александр Ярославич (р. 1991) — аспирант Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича

Румовская София Борисовна (р. 1985) — программист 1-й 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 народов; ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Сидоркин Иван Игоревич (р. 1990) — инженер-исследователь Орловского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

Синицын Игорь Николаевич (р. 1940) — доктор технических наук, профессор, заслуженный деятель науки РФ, заведующий отделом Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Тимонина Елена Евгеньевна (р. 1952) — доктор технических наук, профессор, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Тюрликов Андрей Михайлович (р. 1957) — доктор технических наук, профессор, заведующий кафедрой Санкт-Петербургского государственного университета аэрокосмического приборостроения

Ушаков Владимир Георгиевич (р. 1952) — доктор физико-математических наук, профессор кафедры математической статистики факультета вычислительной математики и кибернетики Московского государственного университета им. М.В. Ломоносова; старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Федосеев Андрей Алексеевич (р. 1946) — кандидат технических наук, доцент, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Чичагов Владимир Витальевич (р. 1955) — кандидат физико-математических, доцент Пермского государственного национального исследовательского университета

Шестаков Олег Владимирович (р. 1976) — доктор физико-математических, доцент кафедры математической статистики факультета вычислительной математики и кибернетики Московского государственного университета им. М.В. Ломоносова; старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Яковлев Олег Альбертович (р. 1992) — инженер-исследователь Орловского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

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 slovary 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 slovarya informatsionnogo monitoringa [Technology of development of semantic dictionary of information monitoring system]. PhD Thesis. Moscow: IPI RAN. 23 p.
- State standards and patents:**
GOST 8.586.5-2005. 2007. Metodika vypolneniya izmereniy. Izmerenie raskhoda i kolichestva zhidkostey i gazov s pomoshch’yu standartnykh suzhayushchikh ustroystv [Method of measurement. Measurement of flow rate and volume of liquids and gases by means of orifice devices]. M.: Standardinform Publs. 10 p.
Bolshakov, M. V., A. V. Kulakov, A. N. Lavrenov, and M. V. Palkin. 2006. Sposob orientirovaniya po krenu letatel’nogo apparata s opticheskoy 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>