

Информатика и её применения

Том 12 Выпуск 2 Год 2018

СОДЕРЖАНИЕ

Об устойчивости конфигурационных графов в случайной среде М. М. Лери, Ю. Л. Павлов	2
Несмещенная оценка риска стабилизированной жесткой пороговой обработки в модели с долгосрочной зависимостью О. В. Шестаков	11
Минимизация ошибок вычисления вейвлет-коэффициентов при решении обратных задач А. А. Кудрявцев, О. В. Шестаков	17
Достаточные условия эргодичности приоритетных систем массового обслуживания А. В. Мистрюков, В. Г. Ушаков	24
Вероятностная модель влияния книги заказов на процесс цены Е. В. Быковец, В. В. Лаврентьев, Л. В. Назаров	29
Максимальные ветвящиеся процессы в случайной среде А. В. Лебедев	35
Иерархический метод порождения метаданных для управления сетевыми соединениями А. А. Грушо, Е. Е. Тимонина, С. Я. Шоргин	44
Модели управления риском в гауссовских стохастических системах А. Н. Тырсин, А. А. Сурина	50
Алгоритм визуализации плоского ядра вероятностной меры С. Н. Васильева, Ю. С. Кан	60
Математическая модель оптимальной триангуляции А. А. Батенков, Ю. А. Маньяков, А. В. Гасилов, О. А. Яковлев	69
Автоматическое извлечение метаданных из научных PDF-документов А. В. Огальцов, О. Ю. Бахтеев	75
Решение проблемы оптимального управления запасом непрерывного продукта при постоянно происходящем потреблении в стохастической полумарковской модели П. В. Шнурков, А. Ю. Егоров	83
Влияние плотности связей на кластеризацию и порог перколяции при распространении информации в социальных сетях Д. О. Жуков, Т. Ю. Хватова, С. А. Лесько, А. Д. Зальцман	90
Дискретный анализ в синтаксическом анализе Я. М. Мирзабеков, Ш. Б. Шихиев	98
Machine translation of Russian connectives into French: Errors and quality failures V. Nuriev, N. Buntman, and O. Inkova	105
Об авторах	114
Правила подготовки рукописей	116
Requirements for manuscripts	119

ОБ УСТОЙЧИВОСТИ КОНФИГУРАЦИОННЫХ ГРАФОВ В СЛУЧАЙНОЙ СРЕДЕ*

М. М. Лери¹, Ю. Л. Павлов²

Аннотация: Рассматриваются случайные конфигурационные графы, степени вершин которых независимы и имеют дискретное степенное распределение со случайным параметром. Законом распределения параметра является усеченное гамма-распределение. Изучается устойчивость таких графов к двум типам разрушающего воздействия: случайному и целенаправленному. Функционирование графов происходит в случайной среде, когда значения параметра распределения степеней вершин выбираются для каждой вершины отдельно. Проведен сравнительный анализ последствий разрушающих воздействий в таких моделях и в графах с общим для всех вершин распределением степеней, индуцированным усреднением по распределению параметра. Обсуждаются условия, при которых исследование поведения графов в случайной среде может быть сведено к изучению эволюции графов с усредненным распределением степеней. Проведен сравнительный анализ последствий этих двух типов разрушающих воздействий.

Ключевые слова: конфигурационные графы; степенное распределение; гамма-распределение; устойчивость; модель лесного пожара; имитационное моделирование

DOI: 10.14357/19922264180201

1 Введение

Широкое использование случайных графов для моделирования топологии и динамики развития таких сложных сетей коммуникаций, как Интернет, социальные и телекоммуникационные сети (см., например, [1, 2]), привело к появлению новых важных направлений исследования подобных структур. Одним из направлений является исследование устойчивости случайных графов к различным разрушающим воздействиям [1, 3, 4], имеющим как случайный, так и целенаправленный характер.

Наблюдения за реальными сетями показали (см., например, [1, 5, 6]), что степени вершин соответствующих графов можно считать независимыми одинаково распределенными случайными величинами, имеющими дискретное степенное распределение. Одними из наиболее подходящих для моделирования сетей видов случайных графов являются так называемые конфигурационные графы [7] со случайными независимыми степенями.

Будем рассматривать конфигурационные графы, состоящие из N вершин, степени которых являются независимыми одинаково распределенными случайными величинами, имеющими следующее дискретное степенное распределение:

$$P\{\xi = k\} = k^{-\tau} - (k+1)^{-\tau}, \\ k = 1, 2, \dots; \tau > 0, \quad (1)$$

где ξ — случайная величина, равная степени произвольной вершины. Случайная величина ξ принимает целые положительные значения, равные числу полуребер вершины, т. е. ребер, для которых смежные вершины еще не определены. Все полуребра занумерованы в произвольном порядке. Поскольку сумма степеней вершин конфигурационного графа должна быть четной, в случае нечетной суммы к равновероятно выбранной вершине добавляется еще одно ребро. Построение графа завершается путем попарного равновероятного соединения полуребер друг с другом для образования ребер. Ясно, что такое строение графа допускает наличие петель, кратных ребер и циклов.

Во многих работах отмечается, что для реальных сетей коммуникаций обычно можно считать, что τ фиксировано и принадлежит интервалу $(1, 2)$ (см., например, [1, 2, 5, 6]). Но в некоторых задачах, как показали последние исследования [8, 9], представляют интерес и модели, в которых $\tau > 2$. Известны также модели, в которых распределения степеней вершин могут изменяться при возрастании числа вершин и даже быть случайными [10].

В статье [11] при $N \rightarrow \infty$ рассматривалась динамика степенной структуры конфигурационного

*Финансовое обеспечение исследований осуществлялось из средств федерального бюджета на выполнение государственного задания КарНЦ РАН (Институт прикладных математических исследований КарНЦ РАН) и при частичной финансовой поддержке РФФИ (проект № 16-01-0005а). Исследования выполнены на научном оборудовании Центра коллективного пользования ФИЦ «Карельский научный центр РАН».

¹Институт прикладных математических исследований КарНЦ РАН, ФИЦ «Карельский научный центр РАН», leri@krc.karelia.ru

²Институт прикладных математических исследований КарНЦ РАН, ФИЦ «Карельский научный центр РАН», pavlov@krc.karelia.ru

графа, степени вершин которого имеют распределение (1), а параметр τ является случайной величиной, равномерно распределенной на отрезке $[a, b]$, $0 < a < b < \infty$. Однако более естественным представляется предположение, что распределение τ , заданное на конечном интервале, является унимодальным.

В настоящей работе предполагается, что τ имеет усеченное гамма-распределение на отрезке $[a, b]$ с параметрами $(2, \lambda)$. Плотность этого распределения имеет следующий вид:

$$f_\lambda(x) = \frac{\lambda^2 x e^{-\lambda x}}{F_{2,\lambda}(b) - F_{2,\lambda}(a)}, \quad x \in [a, b], \quad (2)$$

где $F_{2,\lambda}(a)$ и $F_{2,\lambda}(b)$ — значения функции гамма-распределения с параметрами 2 и λ в точках a и b соответственно.

Главной целью данной работы было изучение устойчивости таких конфигурационных графов к разрушающим воздействиям при условии, что эволюция графов происходит в случайной среде. Это означает, что степени вершин имеют распределение (1), в котором значение параметра τ является случайным и выбирается отдельно для каждой вершины из распределения с плотностью (2). Заметим, что, используя (1) и (2), можно усреднить распределение случайной величины ξ и получить, что

$$P\{\xi = k\} = \frac{\lambda^2}{F_{2,\lambda}(b) - F_{2,\lambda}(a)} \times \int_a^b x e^{-\lambda x} (e^{-x \ln k} - e^{-x \ln(k+1)}) dx, \quad (3)$$

где $k = 1, 2, \dots$. Понятно, что проще изучать поведение таких графов, в которых степени всех вершин имеют общее распределение (3). Поэтому была поставлена задача сравнить устойчивость графов в случайной среде и в случае распределения (3), когда случайной среды не возникает. Это позволит выявить условия, при которых изучение случайных графов в случайной среде можно заменить на исследование графов с общим распределением степеней вершин.

Заметим, что одним из интенсивно развивающихся направлений современной теории вероятностей является изучение случайных процессов в случайной среде, что отражено в многочисленных публикациях как в нашей стране, так и за рубежом (см., например, [12]). Однако авторам неизвестны статьи других исследователей, посвященные случайным графам, формирующимся в случайной среде. В данной работе (см. также материалы конференций [11, 13]) такие задачи, по-видимому, рассматриваются впервые.

2 Разрушение графов

Под разрушением графов далее понимается последовательное удаление вершин вместе с инцидентными им ребрами. Известно (см., например, [6]), что если $\tau \in (1, 2)$, то конфигурационный граф имеет единственную гигантскую компоненту связности, объем которой при $N \rightarrow \infty$ пропорционален числу вершин. Объем любой другой компоненты является бесконечно малой величиной по сравнению с объемом гигантской компоненты. Если $\tau > 2$, то в графе отсутствует гигантская компонента, но, как показали исследования [8, 9], в практически важных случаях и при таких значениях параметра τ наибольшая по объему компонента связности содержит значительно больше вершин, чем любая другая. Ниже наибольшую компоненту будем называть максимальной независимо от того, является ли она гигантской или нет. Обозначим η долю вершин графа, входящих в максимальную компоненту; η_2 — долю вершин, входящих во вторую по объему компоненту. В [8] был предложен критерий разрушения графа, согласно которому граф считается разрушенным, если произошло событие $A : \{\eta \leq 2\eta_2\}$. Рассматривались два вида разрушения: случайное и целенаправленное. При случайном разрушении вершины, подлежащие последовательному удалению из графа, выбираются равновероятно, тогда как в случае целенаправленной атаки удаляются вершины с максимальными степенями.

Обозначим $p = P\{A\}$ вероятность события A , а r — доля удаленных из графа вершин. Для указанных двух видов разрушений были найдены зависимости p и η от r и λ . В этих регрессионных моделях выражение $\xi \sim F[a, b]$ означает, что рассматривались графы, степени вершин которых имеют распределение (3), а $\tau \sim G[a, b]$ означает, что процесс разрушения происходит в условиях случайной среды. Имитационные эксперименты по разрушению графов проводились для моделей объемом от 1000 до 10 000 вершин с шагом 1000, с параметром $0,3 \leq \lambda \leq 2,5$ (с шагом 0,2 и 0,3) для трех интервалов $[a, b]$: (1, 2), (1, 3) и [2, 3]. Для каждой тройки $([a, b], N, \lambda)$ генерировалось по 100 моделей, что в целом составило 30 000 экспериментов. Конфигурационные графы с параметром τ из интервала (1, 2), согласно [5, 6], хорошо подходят для описания сложных телекоммуникационных сетей. Графы с параметром $\tau \in [2, 3]$ показали [8, 9, 14] свою устойчивость к «лесному пожару» (см. разд. 3), а интервал (1, 3] был рассмотрен в качестве обобщения.

Таким образом, в случае целенаправленного разрушения были получены следующие регрессионные модели:

$$\begin{aligned} \tau \sim G(1, 2) : \eta &= 49,3 + 2,91\lambda - 7,76r - 7,21 \ln r, \\ p &= -0,09 - 0,084\lambda + 0,082r^{1,64}; \\ \xi \sim F(1, 2) : \eta &= 48,94 + 2,74\lambda - 7,55r - 7,44 \ln r, \\ p &= -0,09 - 0,09\lambda + 0,086r^{1,6}; \\ \tau \sim G[2, 3] : \eta &= -3,94 + 1,53\lambda + 3,3r - 5,64 \ln r, \\ p &= 1,25 - 0,12\lambda + 0,37 \ln r; \\ \xi \sim F[2, 3] : \eta &= -3,77 + 1,52\lambda + 3,13r - 5,58 \ln r, \\ p &= 1,25 - 0,13\lambda + 0,36 \ln r; \\ \tau \sim G(1, 3] : \eta &= 26,94 + 10,34\lambda - 7,1r - 8,78 \ln r, \\ p &= 0,012 - 0,29\lambda + 0,34r; \\ \xi \sim F(1, 3] : \eta &= 26,7 + 10,48\lambda - 7,05r - 8,79 \ln r, \\ p &= 0,01 - 0,3\lambda + 0,35r. \end{aligned}$$

Коэффициенты детерминации всех приведенных выше моделей имеют значения не ниже 0,94. Значения r в моделях зависимости вероятности разрушения графа p от λ и r имеют следующие ограничения:

$$\begin{aligned} 1,13 + 0,44\lambda \leq r \leq 4,85 + 0,22\lambda &\text{ при } \tau \sim G(1, 2); \\ 1,11 + 0,47\lambda \leq r \leq 4,9 + 0,24\lambda &\text{ при } \xi \sim F(1, 2); \\ 0,03 + 0,02\lambda \leq r \leq 0,54 + 0,16\lambda^{1,4} &\text{ при } \tau \sim G[2, 3]; \\ 0,026 + 0,019\lambda \leq r \leq 0,53 + 0,18\lambda^{1,5} &\text{ при } \xi \sim F[2, 3]; \\ -0,035 + 0,85\lambda \leq r \leq 2,91 + 0,85\lambda &\text{ при } \tau \sim G(1, 3]; \\ -0,03 + 0,86\lambda \leq r \leq 2,83 + 0,86\lambda &\text{ при } \xi \sim F(1, 3], \end{aligned}$$

где $R^2 = 0,99$ для всех моделей. Ясно, что $p = 0$ для значений r , меньших нижней границы, и $p = 1$ для r , больших верхней границы. То же самое верно и для соответствующих моделей, приведенных ниже.

Для процесса случайных разрушений были получены следующие модели:

$$\begin{aligned} \tau \sim G(1, 2) : \eta &= 83,02 + 2,19\lambda - 1,45r, \\ p &= -0,75 - 0,035\lambda + 0,00037r^2; \\ \xi \sim F(1, 2) : \eta &= 82,98 + 2,21\lambda - 1,44r, \\ p &= -0,72 - 0,026\lambda + 0,00036r^2; \\ \tau \sim G[2, 3] : \eta &= 18,97 + 1,82\lambda - 0,28r - 2,1 \ln r, \\ p &= 0,13 - 0,077\lambda + 0,00035r^2; \\ \xi \sim F[2, 3] : \eta &= 18,95 + 1,82\lambda - 0,28r - 2,13 \ln r, \\ p &= 0,13 - 0,075\lambda + 0,00035r^2; \\ \tau \sim G(1, 3] : \eta &= 68,63 + 6,24\lambda - 1,33r, \\ p &= -0,54 - 0,074\lambda + 0,00034r^2; \\ \xi \sim F(1, 3] : \eta &= 68,65 + 6,28\lambda - 1,33r, \\ p &= -0,55 - 0,086\lambda + 0,00036r^2 \end{aligned}$$

с коэффициентами детерминации не ниже 0,92 и следующими ограничениями на значения r в моделях для вероятности разрушения p :

$$\begin{aligned} 45,04 + 1,02\lambda \leq r \leq 68,8 + 0,68\lambda &\text{ при } \tau \sim G(1, 2); \\ 44,73 + 0,79\lambda \leq r \leq 69,12 + 0,52\lambda &\text{ при } \xi \sim F(1, 2); \\ -0,65 + 0,6\lambda^{3,6} \leq r \leq 49,9 + 2,08\lambda &\text{ при } \tau \sim G[2, 3]; \\ -0,57 + 0,43\lambda^{3,9} \leq r \leq 50 + 2,03\lambda &\text{ при } \xi \sim F[2, 3]; \\ 40 + 2,5\lambda \leq r \leq 67,3 + 1,57\lambda &\text{ при } \tau \sim G(1, 3]; \\ 39,23 + 2,77\lambda \leq r \leq 65,65 + 1,75\lambda &\text{ при } \xi \sim F(1, 3]. \end{aligned}$$

Регрессионные модели обычно нивелируют отклонения, поэтому было проведено сравнение исходного материала, использованного для построения этих моделей. Для каждой тройки параметров $([a, b], \lambda, r)$ вычислялась разница между такими основными характеристиками, как объем максимальной компоненты η и вероятность разрушения графа p , полученными для графов в случайной среде и для графов с распределением (3) степеней вершин. На рис. 1 и 2 приведены примеры результатов вычисления $\Delta\eta = \eta_G - \eta_F$ (см. рис. 1) и $\Delta p = p_G - p_F$ (см. рис. 2), где η_G — объем максимальной компоненты для графов в случайной среде; η_F — объем максимальной компоненты для графов с распределением (3) степеней вершин; p_G и p_F — вероятности разрушения графа для графов в случайной среде и с распределением (3) соответственно.

Эти и аналогичные результаты, полученные для других интервалов $[a, b]$, показывают, что при обоих видах процесса разрушения графа (целенаправленном и случайном) разница в динамике изменения объема гигантской компоненты не превышает 3%. Что касается вероятности разрушения графа, здесь уже в некоторых случаях (см., например, рис. 2, б) разница может превышать 5%.

Полученные результаты показывают, что регрессионные модели закономерностей, описывающих разрушения максимальной компоненты в случайной среде и с распределением (3) степеней вершин, практически совпадают. Это означает, что исследование таких процессов разрушения в случайной среде может быть заменено исследованием более простого случая с усредненным распределением степеней вершин при изучении процесса целенаправленного разрушения графов. Однако при исследовании вероятностей случайных разрушений в случае такой замены следует учесть возможность заметных отклонений при достаточно больших значениях параметра.

3 «Лесные пожары»

Второй вид процесса разрушения известен как «модель лесного пожара» (см., например, [15, 16]), которая также может быть использована при изучении банковских кризисов [17]. В этом случае вер-

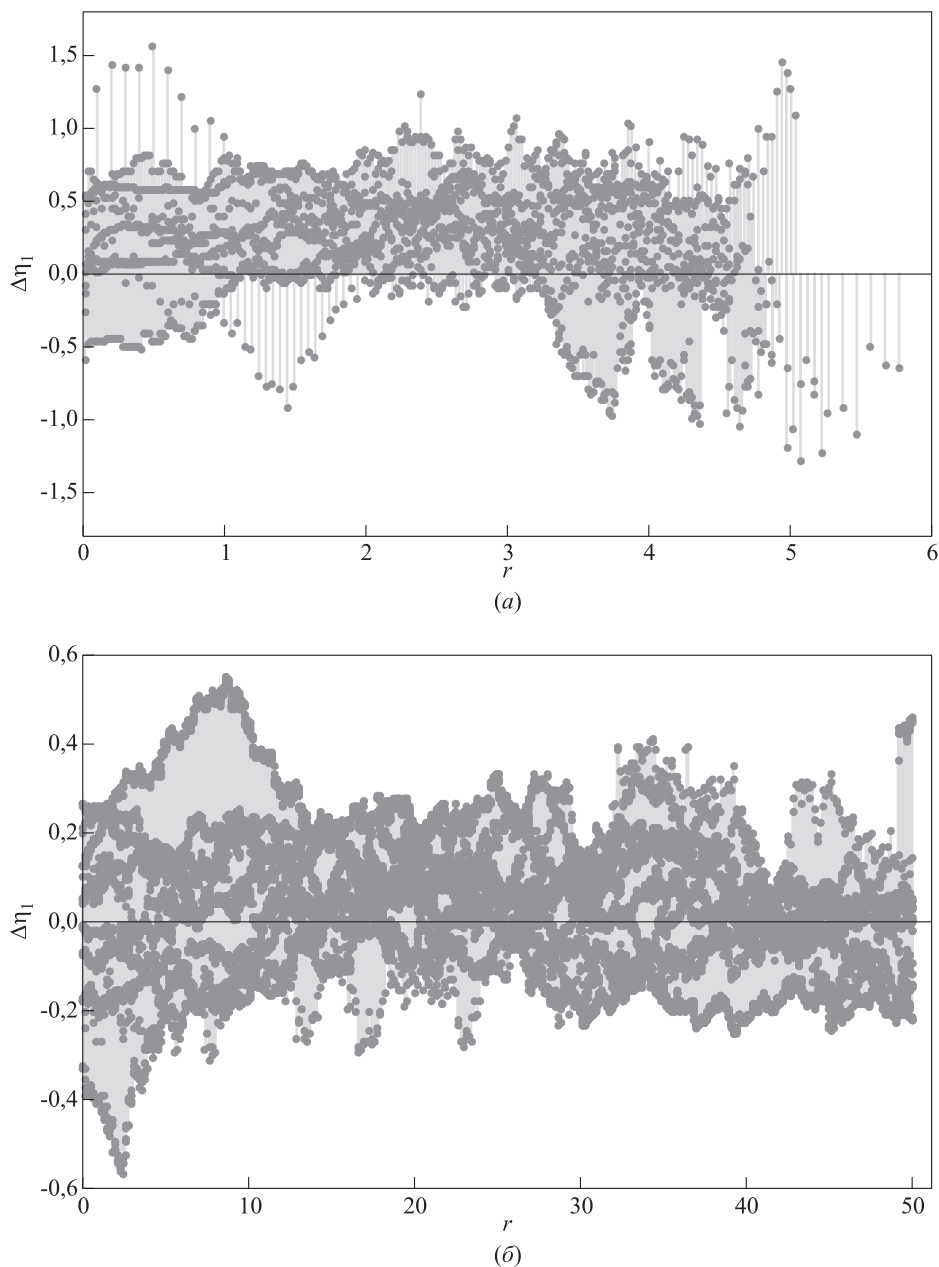


Рис. 1 Сравнение динамики объемов максимальных компонент: (а) при целенаправленном разрушении для интервала (1, 2); (б) при случайном разрушении для интервала [2, 3]

шины конфигурационного графа интерпретируются как деревья, растущие на некоторой ограниченной территории, а ребра графа — это возможные пути перехода «огня» от одной вершины к другой.

Рассматривалась задача сравнения устойчивости к «лесному пожару» двух типов моделей конфигурационных графов:

(1) графов, степени вершин которых имеют распределение (1) с параметром τ , имеющим распределение с плотностью (2) (графы в случайной среде);

(2) графов, распределение степеней вершин которых имеет вид (3).

Для согласования топологии конфигурационного графа с моделью «лесного пожара» в предыдущих работах (см., например, [8, 9, 14]) было предложено использовать вспомогательные графы, вершины которых располагаются в узлах квадратной целочисленной решетки. Ребра вспомогательного графа соединяют вершины по принципу «ближайшего соседа». Таким образом, степень любой вершины, находящейся внутри решетки, не пре-

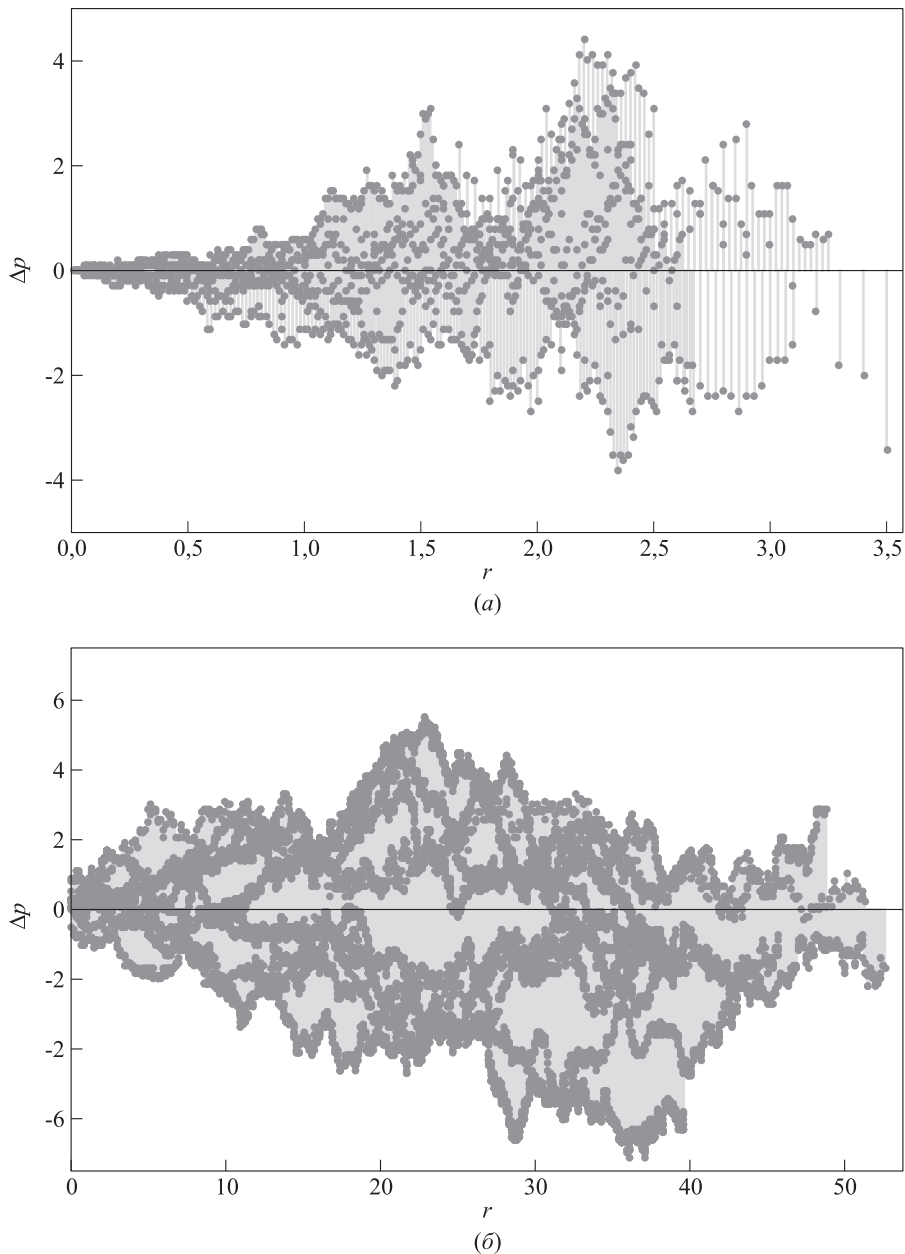


Рис. 2 Сравнение вероятностей разрушения графа: (а) при целенаправленном разрушении для интервала (1, 3]; (б) при случайном разрушении для интервала [2, 3]

вышает 8. Вспомогательные графы были использованы для установления зависимости между параметром λ распределения степеней вершин и числом N исходных вершин конфигурационного графа.

Для нахождения такой зависимости были вычислены средние значения $m = m(N)$ степеней вершин (см. также [9]) при различных значениях $N \leq 10\,000$ и различных вариантах размещения N вершин на решетке размера 100×100 (табл. 1), при этом, учитывая явный вид матема-

тического ожидания ξ , использовалось соотношение:

$$m \approx \int_a^b x \frac{\lambda^2 x e^{-\lambda x} dx}{F_{2,\lambda}(b) - F_{2,\lambda}(a)} = \frac{1}{\lambda} \frac{\Gamma(3)}{F_{2,\lambda}(b) - F_{2,\lambda}(a)} (F_{3,\lambda}(b) - F_{3,\lambda}(a)).$$

На основании приведенных в табл. 1 данных были найдены регрессионные зависимости λ от N (рис. 3) для трех интервалов $[a, b]$:

Таблица 1 Данные для нахождения зависимости между N и λ

m	N	λ		
		(1,2)	[2,3]	(1,3]
1,01	3350	0,824	0,638	0,722
1,21	3600	0,904	0,699	0,795
1,33	3750	0,949	0,733	0,836
1,42	3900	0,981	0,758	0,866
1,5	4000	1,009	0,779	0,892
1,6	4200	1,043	0,805	0,924
2	5000	1,17	0,902	1,043
2,66	4780	1,357	1,043	1,221
3	4489	1,444	1,108	1,306
4	5578	1,679	1,284	1,536
5	6700	1,888	1,439	1,745
5,33	7500	1,952	1,487	1,811
6	8350	2,078	1,58	1,939
6,93	8680	2,244	1,701	2,109
7	8911	2,256	1,71	2,122
8	10000	2,422	1,831	2,294

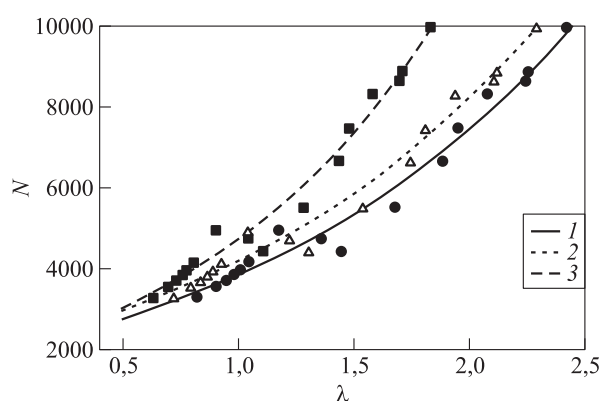


Рис. 3 Зависимости N от λ на интервалах (1, 2) (1), [2, 3] (2) и (1, 3] (3)

$$N(\lambda) = \begin{cases} 1986,36e^{0,66\lambda} & \text{для (1, 2);} \\ 1934,98e^{0,89\lambda} & \text{для [2, 3];} \\ 2143,92e^{0,67\lambda} & \text{для (1, 3]} \end{cases} \quad (4)$$

с коэффициентами детерминации $R^2 = 0,99$ для всех трех моделей.

Процесс «лесного пожара» состоит в следующем. Возгорание начинается с некоторой вершины графа, которая может быть либо выбрана равновероятно (случайное возгорание), либо имеет максимальную степень, т. е. происходит целенаправленный поджог. После того как первая вершина загорелась, огонь переходит по ребрам на инцидентные вершины с вероятностью $0 < p \leq 1$. Вероятность p либо фиксирована для всех ребер графа, либо равномерно распределена на интервале $(0, 1]$ и генерируется отдельно для каждого ребра.

Имитационное моделирование процесса «лесного пожара» проводилось для моделей с параметром $0,5 \leq \lambda \leq 2,0$, изменяющимся с шагом 0,1, для трех интервалов $[a, b]$: (1, 2), (1, 3] и [2, 3]. Объемы графов вычислялись из уравнений (4).

Рассмотрим сначала результаты экспериментов, полученные в случае фиксированных значений вероятности перехода огня по ребру графа $0 < p \leq 1$, задаваемых с шагом 0,01. Для каждой тройки параметров $([a, b], \lambda, p)$ было построено по 100 моделей графов. Найдены зависимости числа n оставшихся после пожара вершин графа от параметров λ и p для трех интервалов $[a, b]$. В случае направленного поджога были получены следующие регрессионные модели:

$$\begin{aligned} \tau \sim G(1, 2) : \quad n &= -1477,5 + 2320,14\lambda - \\ &\quad - 1016,68 \ln \lambda - 1199,93 \ln p; \\ \xi \sim F(1, 2) : \quad n &= -1479,46 + 2322,17\lambda - \\ &\quad - 1016,65 \ln \lambda - 1199,88 \ln p; \\ \tau \sim G[2, 3] : \quad n &= -4608,64 + 9630,65\lambda - \\ &\quad - 5139,23 \ln \lambda - 1301,38p; \\ \xi \sim F[2, 3] : \quad n &= -4657,43 + 9671,31\lambda - \\ &\quad - 5184,17 \ln \lambda - 1276,98p; \\ \tau \sim G(1, 3] : \quad n &= -1125,4 + 2473,31\lambda - \\ &\quad - 1087,46 \ln \lambda - 1209,57 \ln p; \\ \xi \sim F(1, 3] : \quad n &= -1180 + 2521,48\lambda - \\ &\quad - 1146,44 \ln \lambda - 1209,79 \ln p, \end{aligned}$$

коэффициенты детерминации которых не ниже 0,95. Аналогично при рассмотрении случайного возгорания модели были следующими:

$$\begin{aligned} \tau \sim G(1, 2) : \quad n &= 1395,8 + 3255\lambda - \\ &\quad - 1444 \ln \lambda - 3848,76p; \\ \xi \sim F(1, 2) : \quad n &= 1434,29 + 3218,54\lambda - \\ &\quad - 1403,52 \ln \lambda - 3837,88p; \\ \tau \sim G[2, 3] : \quad n &= -5657,21 + 10335,37\lambda - \\ &\quad - 5563,53 \ln \lambda - 236,79p; \\ \xi \sim F[2, 3] : \quad n &= -5674,23 + 10349,4\lambda - \\ &\quad - 5578,1 \ln \lambda - 230,24p; \\ \tau \sim G(1, 3] : \quad n &= 1612,9 + 3465,12\lambda - \\ &\quad - 1469,7 \ln \lambda - 3539,2p; \\ \xi \sim F(1, 3] : \quad n &= 1443,68 + 3615,87\lambda - \\ &\quad - 1632,72 \ln \lambda - 3541,65p \end{aligned}$$

со значениями коэффициентов детерминации не ниже 0,89.

Для каждой тройки значений $([a, b], \lambda, p)$ было проведено попарное сравнение выборок, состоящих каждая из 100 значений n — числа оставшихся

Таблица 2 Сравнение числа оставшихся после пожара вершин ($0 < p \leq 1$ фиксировано)

Вид возгорания	Интервал $[a, b]$		
	(1,2)	[2,3]	(1,3)
Целенаправленный поджог	0,25%	6,19%	4,87%
Случайное возгорание	4,06%	4,94%	5,87%

после пожара вершин. Таким образом, всего было сформировано 1600 выборок для каждого интервала. Для всех пар выборок (в случайной среде и с распределением (3) степеней вершин) проверялась гипотеза о равенстве средних с уровнем значимости 0,05. Вследствие отсутствия нормальности был использован критерий сравнения Манна–Уитни. В табл. 2 приведены доли случаев, в которых гипотеза отвергалась для целенаправленного поджога и случайного возгорания.

Далее рассматривался процесс случайного распространения огня на сравниваемых графах, т.е. случай, когда вероятность p перехода огня по ребру графа не фиксирована, а является случайной величиной, равномерно распределенной на интервале $(0, 1]$. Аналогично были построены регрессионные модели зависимостей числа оставшихся в графе вершин n от параметра λ . В случае целенаправленного поджога были получены следующие зависимости:

$$\begin{aligned} \tau \sim G(1, 2) : n &= -228,81 + 2161,35\lambda - 937,28 \ln \lambda; \\ \xi \sim F(1, 2) : n &= -235,57 + 2166,73\lambda - 942,78 \ln \lambda; \\ \tau \sim G[2, 3] : n &= -5680,61 + 10162,21\lambda - \\ &\quad - 5458,21 \ln \lambda; \\ \xi \sim F[2, 3] : n &= -5715,7 + 10202,6\lambda - \\ &\quad - 5507,79 \ln \lambda; \\ \tau \sim G(1, 3) : n &= 202,28 + 2293,47\lambda - 999,13 \ln \lambda; \\ \xi \sim F(1, 3) : n &= 151,96 + 2336,82\lambda - 1054,33 \ln \lambda, \end{aligned}$$

все коэффициенты детерминации которых примерно равны 0,99. Аналогично при рассмотрении случайного возгорания модели были следующими:

$$\begin{aligned} \tau \sim G(1, 2) : n &= -1082,11 + 3896,2\lambda - \\ &\quad - 1956,43 \ln \lambda; \\ \xi \sim F(1, 2) : n &= -660,53 + 3545,96\lambda - \\ &\quad - 1531,34 \ln \lambda; \\ \tau \sim G[2, 3] : n &= -5680,61 + 10162,21\lambda - \\ &\quad - 5458,21 \ln \lambda; \\ \xi \sim F[2, 3] : n &= -5881,93 + 10467,04\lambda - \\ &\quad - 5649,63 \ln \lambda; \end{aligned}$$

Таблица 3 Сравнение числа оставшихся после пожара вершин

Вид возгорания	Интервал $[a, b]$		
	(1,2)	[2,3]	(1,3)
Целенаправленный поджог	0	1	2
Случайное возгорание	1	1	1

$$\begin{aligned} \tau \sim G(1, 3) : n &= -1118,8 + 4536,45\lambda - \\ &\quad - 2417,32 \ln \lambda; \\ \xi \sim F(1, 3) : n &= -351,88 + 3824,35\lambda - \\ &\quad - 1716,98 \ln \lambda \end{aligned}$$

со значениями коэффициентов детерминации не ниже 0,89.

По аналогии со случаем фиксированного значения p для каждой пары значений $([a, b], \lambda)$ было проведено попарное сравнение числа оставшихся после пожара вершин на двух типах графов. Вследствие того, что для каждого интервала $[a, b]$ здесь имеется только 16 пар $([a, b], \lambda)$, в табл. 3 приведена не доля, а число случаев принятия альтернативной гипотезы.

Таким образом, проведенное исследование процесса разрушения типа «лесного пожара» показало, что с достаточно высокой достоверной вероятностью исследование такого процесса на графах в случайной среде может быть заменено изучением более простого случая с усредненным распределением степеней вершин. Понятно, что выводы этой статьи о возможности упрощения методов изучения процессов разрушения графов в случайной среде путем усреднения распределений степеней вершин относятся только к рассмотренным распределениям степеней и видам разрушения. В других случаях, особенно в более общих и ответственных, следует, видимо, проводить дополнительные исследования.

Литература

1. *Durrett R.* Random graph dynamics. — Cambridge: Cambridge University Press, 2007. 221 p.
2. *Hofstad R.* Random graphs and complex networks. — Cambridge: Cambridge University Press, 2017. Vol. 1. 337 p.
3. *Cohen R., Erez K., Ben-Avraham D., Havlin S.* Resilience of the Internet to random breakdowns // Phys. Rev. Lett., 2000. Vol. 85. P. 4626–4628.
4. *Bollobas B., Riordan O.* Robustness and vulnerability of scale-free random graphs // Internet Math., 2004. Vol. 1. No. 1. P. 1–35.

5. Faloutsos C., Faloutsos P., Faloutsos M. On power-law relationships of the internet topology // *Comp. Comm. R.*, 1999. Vol. 29. P. 251–262.
6. Reittu H., Norros I. On the power-law random graph model of massive data networks // *Perform. Evaluation*, 2004. Vol. 55. P. 3–23.
7. Bollobas B. A. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs // *Eur. J. Combust.*, 1980. Vol. 1. P. 311–316.
8. Leri M., Pavlov Yu. Power-law random graphs' robustness: Link saving and forest fire model // *Austrian J. Stat.*, 2014. Vol. 43. No. 4. P. 229–236.
9. Leri M., Pavlov Y. Forest fire models on configuration random graphs // *Fundam. Inform.*, 2016. Vol. 145. Iss. 3. P. 313–322.
10. Bianconi G., Barabasi A.-L. Bose–Einstein condensation in complex networks // *Phys. Rev. Lett.*, 2001. Vol. 86. Iss. 24. P. 5632–5635.
11. Pavlov Y. On random graphs in random environment // *Computer data analysis and modeling: Theoretical and applied stochastics: 11th Conference (International) Proceedings*. — Minsk, 2016. P. 177–180.
12. Afanasyev V. I., Boinghoff C., Kersting G., Vatutin V. A. Limit theorems for weekly branching processes in random environment // *J. Theor. Probab.*, 2012. Vol. 25. No. 3. P. 703–732.
13. Leri M., Pavlov Y. On robustness of configuration graphs in random environment // *4th Russian Finnish Symposium on Discrete Mathematics Proceedings / Eds. J. Karhumäki, A. Saarela*. — TUCS lecture notes ser. — Turku, Finland: Turku Centre for Computer Science, 2017. Vol. 26. P. 112–115.
14. Леру М. М. Пожар на конфигурационном графе со случайными переходами огня по ребрам // *Информатика и её применения*, 2015. Т. 9. Вып. 3. С. 67–73.
15. Drossel B., Schwabl F. Self-organized critical forest-fire model // *Phys. Rev. Lett.*, 1992. Vol. 69. P. 1629–1632.
16. Bertoin J. Fires on trees // *Ann. I. H. Poincaré Pr.*, 2012. Vol. 48. No. 4. P. 909–921.
17. Arinaminparty N., Kapadia S., May R. Size and complexity in model financial systems // *P. Natl. Acad. Sci. USA*, 2012. Vol. 109. No. 45. P. 18338–18343.

Поступила в редакцию 20.10.17

ON THE ROBUSTNESS OF CONFIGURATION GRAPHS IN A RANDOM ENVIRONMENT

M. M. Leri and Yu. L. Pavlov

Institute of Applied Mathematical Research of the Karelian Research Centre of the Russian Academy of Sciences,
11 Pushkinskaya Str., Petrozavodsk 185910, Russian Federation

Abstract: The paper considers configuration graphs with vertex degrees being independent identically distributed random variables following the power-law distribution with a random parameter. The parameter of the vertex degree distribution follows the truncated gamma distribution. The authors study the robustness of such graphs to the two types of destruction processes: random and targeted. The graphs function in a random environment where the values of the vertex degree distribution parameter are chosen separately for each vertex. A comparative analysis of destruction effects on these models and on graphs with the degree distribution common for all vertices and induced by averaging over the distribution parameter has been performed. The conditions under which the study of the graphs' behavior in a random environment can be reduced to the study of the evolution of graphs with an averaged vertex degree distribution are discussed. A comparative analysis of destruction effects of the two types of destruction processes has been performed.

Keywords: configuration graphs; power-law distribution; gamma distribution; robustness; forest fire model; simulation

DOI: 10.14357/19922264180201

Acknowledgments

The study was carried out under state order to the Karelian Research Centre of the Russian Academy of Sciences (Institute of Applied Mathematical Research KarRC RAS) and partly supported by the Russian Foundation for Basic Research, grant No. 16-01-0005a. The research was carried out using the equipment of the Core Facility of the Karelian Research Centre of the Russian Academy of Sciences.

References

1. Durrett, R. 2007. *Random graph dynamics*. Cambridge: Cambridge University Press. 221 p.
2. Hofstad, R. 2011. *Random graphs and complex networks*. Cambridge: Cambridge University Press. Vol. 1. 363 p.
3. Cohen, R., K. Erez, D. Ben-Avraham, and S. Havlin. 2000. Resilience of the Internet to random breakdowns. *Phys. Rev. Lett.* 85:4626–4628.
4. Bollobas, B., and O. Riordan. 2004. Robustness and vulnerability of scale-free random graphs. *Internet Math.* 1(1):1–35.
5. Faloutsos, C., P. Faloutsos, and M. Faloutsos. 1999. On power-law relationships of the internet topology. *Comp. Comm. R.* 29:251–262.
6. Reittu, H., and I. Norros. 2004. On the power-law random graph model of massive data networks. *Perform. Evaluation* 55:3–23.
7. Bollobas, B. A. 1980. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *Eur. J. Combust.* 1:311–316.
8. Leri, M., and Yu. Pavlov. 2014. Power-law random graphs' robustness: Link saving and forest fire model. *Austrian J. Stat.* 43(4):229–236.
9. Leri, M., and Yu. Pavlov. 2016. Forest fire models on configuration random graphs. *Fundam. Inform.* 145(3):313–322.
10. Bianconi, G., and A.-L. Barabasi. 2001. Bose–Einstein condensation in complex networks. *Phys. Rev. Lett.* 86(24):5632–5635.
11. Pavlov, Y. 2016. On random graphs in random environment. *11th Conference (International) "Computer Data Analysis and Modeling: Theoretical and Applied Stochastics" Proceedings*. Minsk. 177–180.
12. Afanasyev, V. I., C. Boinghoff, G. Kersting, and V. A. Vatutin. 2012. Limit theorems for weekly branching processes in random environment. *J. Theor. Probab.* 25(3):703–732.
13. Leri, M., and Y. Pavlov. 2017. On robustness of configuration graphs in random environment. *4th Russian Finnish Symposium on Discrete Mathematics Proceedings*. Eds. J. Karhumäki and A. Saarela. TUCS lecture notes ser. Turku, Finland: Turku Centre for Computer Science. 26:112–115.
14. Leri, M. M. 2015. Pozhar na konfiguratsionnom grafe so sluchaynymi perekhodami ognya po rebram [Forest fire on configuration graph with random fire propagation]. *Informatika i ee Primeneniya — Inform. Appl.* 9(3):67–73.
15. Drossel, B., and F. Schwabl. 1992. Self-organized critical forest-fire model. *Phys. Rev. Lett.* 69:1629–1632.
16. Bertoin, J. 2012. Fires on trees. *Ann. I. H. Poincaré Pr.* 48(4):909–921.
17. Arinaminparty, N., S. Kapadia, and R. May. 2012. Size and complexity in model financial systems. *P. Natl. Acad. Sci. USA* 109:18338–18343.

Received October 20, 2017

Contributors

Leri Marina M. (b. 1969) — Candidate of Science (PhD) in technology, scientist, Institute of Applied Mathematical Research of the Karelian Research Centre of the Russian Academy of Sciences, 11 Pushkinskaya Str., Petrozavodsk 185910, Russian Federation; leri@krc.karelia.ru

Pavlov Yuri L. (b. 1949) — Doctor of Science in physics and mathematics, principal scientist, Institute of Applied Mathematical Research of the Karelian Research Centre of the Russian Academy of Sciences, 11 Pushkinskaya Str., Petrozavodsk 185910, Russian Federation; pavlov@krc.karelia.ru

НЕСМЕЩЕННАЯ ОЦЕНКА РИСКА СТАБИЛИЗИРОВАННОЙ ЖЕСТКОЙ ПОРОГОВОЙ ОБРАБОТКИ В МОДЕЛИ С ДОЛГОСРОЧНОЙ ЗАВИСИМОСТЬЮ*

О. В. Шестаков¹

Аннотация: Методы подавления шума в сигналах и изображениях, основанные на процедуре пороговой обработки коэффициентов вейвлет-разложения, стали популярными благодаря своей простоте, скорости и возможности адаптации к функциям сигналов, имеющим на разных участках различную степень регулярности. Анализ погрешностей этих методов является важной практической задачей, поскольку дает возможность оценивать качество как самих методов, так и используемого для обработки оборудования. Рассматривается предложенный недавно стабилизированный метод жесткой пороговой обработки, в котором устранены основные недостатки мягкой и жесткой пороговой обработки, и исследуются статистические свойства этого метода. В модели данных с аддитивным гауссовским шумом проводится анализ несмещенной оценки среднеквадратичного риска. В предположении о том, что шумовые коэффициенты обладают долгосрочной зависимостью, приводятся условия, при которых имеет место сильная состоятельность и асимптотическая нормальность несмещенной оценки риска. Полученные результаты дают возможность строить асимптотические доверительные интервалы для погрешностей пороговой обработки, используя только наблюдаемые данные.

Ключевые слова: вейвлеты; пороговая обработка; несмещенная оценка риска; коррелированный шум; асимптотическая нормальность

DOI: 10.14357/19922264180202

1 Введение

В задачах статистической обработки данных часто предполагается, что наблюдения независимы. Однако существует множество физических процессов, демонстрирующих долгосрочную зависимость, при которой корреляции между наблюдениями убывают столь медленно, что ряд из них не сходится. Такая долгосрочная зависимость часто наблюдается, например, при исследовании геофизических процессов, в которых она принимает форму длительных периодов больших или маленьких значений наблюдений. Схожие явления демонстрируют помехи в коммуникационных каналах.

При анализе и обработке сигналов, регистрируемых при изучении таких процессов, широко применяются методы вейвлет-анализа. К данным применяется вейвлет-преобразование и осуществляется пороговая обработка получившихся вейвлет-коэффициентов [1]. Наличие шумовой составляющей в сигнале неизбежно приводит к погрешностям в оценке сигнала. При использовании метода мягкой пороговой обработки можно

построить статистическую оценку среднеквадратичной погрешности (риска) [2]. Свойства этой оценки в моделях с независимым и коррелированным шумом исследовались в работах [3–7]. Показано, что при определенных условиях оценка риска является сильно состоятельной и асимптотически нормальной.

Однако при мягкой пороговой обработке в оценке функции сигнала появляется дополнительное смещение.

При жесткой пороговой обработке используется разрывная пороговая функция, что приводит к появлению дополнительных артефактов, отсутствию устойчивости при выборе порога и невозможности построения несмещенной оценки среднеквадратичного риска. В работе [8] предложен стабилизированный вариант жесткой пороговой обработки, позволяющий обойти указанные недостатки.

В данной работе исследуются статистические свойства оценки среднеквадратичного риска стабилизированной жесткой пороговой обработки в модели с коррелированным шумом. Предполагается, что при обработке используется «универсальный»

*Работа выполнена при частичной финансовой поддержке РФФИ (проект 16-07-00736).

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.ru

порог. При определенных условиях на гладкость функции сигнала показано, что оценка риска, как и в случае мягкой пороговой обработки, является асимптотически нормальной и сильно состоятельной. Данные свойства служат обоснованием использования этой оценки при построении доверительных интервалов для теоретического средне-квадратичного риска.

2 Модель данных с долгосрочной зависимостью

В данной работе предполагается, что функция сигнала f задана на отрезке $[0, 1]$ и равномерно регулярна по Липшицу с некоторым показателем $\gamma > 0$. Вейвлет-разложение функции f представляет собой ряд

$$f = \sum_{j,k \in \mathbb{Z}} \langle f, \psi_{jk} \rangle \psi_{jk}, \quad (1)$$

где $\psi_{jk}(t) = 2^{j/2} \psi(2^j t - k)$, а $\psi(t)$ — некоторая материнская вейвлет-функция. Индекс j в (1) называется масштабом, а индекс k — сдвигом. Функция ψ должна удовлетворять определенным требованиям, однако ее можно выбрать таким образом, чтобы она обладала некоторыми полезными свойствами, например была M раз дифференцируемой, имела заданное число M нулевых моментов и достаточно быстро убывала на бесконечности. Известно [1], что если $M \geq \gamma$, то найдется такая константа $C_f > 0$, что

$$\langle f, \psi_{jk} \rangle \leq \frac{C_f}{2^{j(\gamma+1/2)}}. \quad (2)$$

Всюду далее предполагается, что используются вейвлеты Мейера [1], обладающие нужным числом нулевых моментов и непрерывных производных.

На практике регистрируются дискретные отсчеты функции сигнала $f_i = f(i/2^J)$, $i = 1, \dots, 2^J$ (считается, что число этих отсчетов равно 2^J для некоторого $J > 0$). Дискретное вейвлет-преобразование представляет собой умножение вектора из значений f_i на ортогональную матрицу, определяемую вейвлет-функцией ψ [1]. При этом дискретные вейвлет-коэффициенты связаны с непрерывными коэффициентами разложения в (1) следующим образом: $\mu_{jk} \approx 2^{J/2} \langle f, \psi_{jk} \rangle$ [1]. Это приближение тем точнее, чем больше J .

В реальных наблюдениях всегда присутствует шум. В данной работе рассматривается модель коррелированного шума. Пусть $\{e_i, i \in \mathbb{Z}\}$ — стационарный гауссовский процесс с ковариационной последовательностью $r_k = \text{cov}(e_i, e_{i+k})$. Будем полагать, что e_i имеют нулевое среднее и единичную

дисперсию. Также предположим, что автоковариационная функция шума убывает со скоростью $r_k \sim Ak^{-\alpha}$, где $0 < \alpha < 1$, что соответствует долгосрочной зависимости между наблюдениями [9].

Рассмотрим следующую модель данных:

$$Y_i = f_i + e_i, \quad i = 1, \dots, 2^J.$$

Для $t \in [0, 1]$ определим наблюдаемый процесс

$$Y_J(t) = \frac{1}{2^J} \sum_{j=1}^{[2^J t]} Y_i = F_J(t) + \frac{1}{2^J} \sum_{i=1}^{[2^J t]} e_i,$$

где $F_J(t) = 1/2^J \sum_{i=1}^{[2^J t]} f(i/2^J)$ — «суммарный сигнал». Положим $\tau^2 = 2A/((1-\alpha)(2-\alpha))$ (без ограничения общности далее предполагается, что $\tau = 1$) и $H = 1 - \alpha/2 \in (1/2, 1)$.

Определим дробное броуновское движение $\mathbf{B}_H(t)$ как гауссовский процесс на \mathbb{R} с нулевым средним и ковариационной функцией

$$r(s, t) = \frac{V_H}{2} (|s|^{2H} + |t|^{2H} - |t - s|^{2H}), \quad s, t \in \mathbb{R},$$

где

$$V_H = D(\mathbf{B}_H(1)) = \frac{-\Gamma(2-2H) \cos(\pi H)}{\pi H(2H-1)}.$$

Лемма 5.1 из [10] показывает, что

$$2^{\alpha J/2} (Y_J(t) - F_J(t)) \Rightarrow \mathbf{B}_H(t), \quad t \in [0, 1].$$

Таким образом, полагая $\epsilon = 2^{-J/2}$, можно аппроксимировать наблюдаемый процесс $Y_J(t)$ с помощью $Y(t)$ для $t \in [0, 1]$:

$$Y(t) = F(t) + \epsilon^\alpha \mathbf{B}_H(t). \quad (3)$$

Применяя к (3) вейвлет-разложение и аппроксимируя его дискретным вейвлет-преобразованием, приходим к следующей модели дискретных вейвлет-коэффициентов [9, 11]:

$$X_{jk} = \mu_{jk} + 2^{(1-\alpha)(J-j)/2} z_{jk}, \quad (4)$$

где

$$z_{jk} = 2^{j(1-\alpha)/2} \int \psi_{jk} d\mathbf{B}_H.$$

Шумовые коэффициенты z_{jk} имеют стандартное нормальное распределение, но не являются независимыми. Дисперсия коэффициентов модели (4) имеет вид:

$$\sigma_j^2 = 2^{(1-\alpha)(J-j)}.$$

3 Стабилизированная жесткая обработка

Для подавления шума и построения оценки функции сигнала к коэффициентам X_{jk} обычно применяется функция жесткой пороговой обработки $\rho_H(x, T_j) = y\mathbf{1}(|x| > T_j)$ или функция мягкой пороговой обработки $\rho_S(x, T_j) = \text{sgn}(x)(|x| - T_j)_+$ с порогом T_j , который может зависеть от масштаба j , но не зависит от сдвига k . Смысл пороговой обработки заключается в удалении достаточно маленьких коэффициентов, которые считаются шумом.

Как уже отмечалось, функция ρ_H разрывна, что приводит к отсутствию устойчивости, а функция ρ_S приводит к появлению дополнительного смещения в оценке функции сигнала. В работе [8] предложен альтернативный метод пороговой обработки, являющийся сглаженным аналогом жесткой пороговой обработки. В этом методе оценки μ_{jk} вычисляются по формулам:

$$\hat{\mu}_{jk} = \mathbb{E}[\rho_H(X_{jk} + \lambda\xi_{jk}, T_j)|X_{jk}],$$

где случайные величины ξ_{jk} имеют стандартное нормальное распределение и не зависят от X_{jk} , а $\lambda > 0$ — параметр стабилизации, отвечающий за степень сглаживания. Вычисляя математическое ожидание, получаем:

$$\hat{\mu}_{jk} = X_{jk} \left[\Phi\left(-\frac{T_j + X_{jk}}{\lambda}\right) + 1 - \Phi\left(\frac{T_j - X_{jk}}{\lambda}\right) \right] + \lambda \left[\phi\left(\frac{T_j - X_{jk}}{\lambda}\right) - \phi\left(\frac{T_j + X_{jk}}{\lambda}\right) \right].$$

Достоинством такого метода является бесконечная дифференцируемость $\hat{\mu}_{jk}$ по X_{jk} , что приводит к более устойчивым оценкам [8]. Заметим также, что при $\lambda \rightarrow 0$ получается обычный метод жесткой пороговой обработки.

Среднеквадратичная погрешность (риск) описанного метода определяется по формуле:

$$R_J(f) = \sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} \mathbb{E}(\hat{\mu}_{jk} - \mu_{jk})^2.$$

В [8] показано, что

$$\begin{aligned} \mathbb{E}(\hat{\mu}_{jk} - \mu_{jk})^2 &= \\ &= \mathbb{E} \left[(X_{jk} - \hat{\mu}_{jk})^2 + 2\sigma_j^2 \frac{\partial}{\partial X_{jk}} \hat{\mu}_{jk} \right] - \sigma_j^2, \end{aligned}$$

где

$$\begin{aligned} \frac{\partial}{\partial X_{jk}} \hat{\mu}_{jk} &= \\ &= \left[\Phi\left(-\frac{T_j + X_{jk}}{\lambda}\right) + 1 - \Phi\left(\frac{T_j - X_{jk}}{\lambda}\right) \right] + \\ &+ \frac{T_j}{\lambda} \left[\phi\left(\frac{T_j - X_{jk}}{\lambda}\right) + \phi\left(\frac{T_j + X_{jk}}{\lambda}\right) \right]. \end{aligned}$$

Таким образом, величина

$$\begin{aligned} \hat{R}_J(f) &= \\ &+ \sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} \left[(X_{jk} - \hat{\mu}_{jk})^2 + 2\sigma_j^2 \frac{\partial}{\partial X_{jk}} \hat{\mu}_{jk} - \sigma_j^2 \right] \quad (5) \end{aligned}$$

является несмещенной оценкой среднеквадратичного риска $R_J(f)$, не зависящей от ненаблюдаемых «чистых» значений μ_{jk} .

В данной работе параметр λ предполагается фиксированным, а в качестве T_j для каждого масштаба j выбирается «универсальный» порог $T_j = \sigma_j \sqrt{2 \ln 2^J}$, который позволяет достичь хороших результатов при подавлении шума и обеспечивает близость среднеквадратичного риска к минимальному как при жесткой, так и при мягкой пороговой обработке [12]. В следующем разделе устанавливаются свойства асимптотической нормальности и сильной состоятельности оценки (5). Эти свойства служат обоснованием использования $\hat{R}_J(f)$ при построении доверительных интервалов для $R_J(f)$.

4 Статистические свойства оценки среднеквадратичного риска

Покажем, что оценка (5) является асимптотически нормальной.

Теорема 1. Пусть $\alpha > 1/2$, а функция f задана на отрезке $[0, 1]$ и равномерно регулярна по Липшицу с показателем $\gamma > (4\alpha - 2)^{-1}$. Тогда при стабилизированной жесткой пороговой обработке с «универсальными» порогами T_j имеет место сходимость по распределению

$$P \left(\frac{\hat{R}_J(f) - R_J(f)}{D_J} < x \right) \rightarrow \Phi(x) \text{ при } J \rightarrow \infty. \quad (6)$$

Здесь $\Phi(x)$ — функция распределения стандартного нормального закона; $D_J^2 = C_\alpha 2^J$, где C_α — константа, зависящая только от α и выбранного вейвлет-базиса.

Замечание. На практике, например при построении асимптотических доверительных интервалов

для среднеквадратичного риска, необходимо знать константу C_α . В отличие от случая независимых наблюдений эта константа зависит от выбранного вейвлет-базиса (напомним, что по предположению базис строится на основе вейвлетов Мейера). Константу C_α можно вычислить достаточно точно, пользуясь методикой работы [13].

Доказательство. Поскольку в силу условий теоремы $(2\gamma + 1)^{-1} < 1 - (2\alpha)^{-1}$, можно выбрать такое p , что $(2\gamma + 1)^{-1} < p < 1 - (2\alpha)^{-1}$. Обозначим слагаемые в (5) через F_{jk} и запишем дробь под вероятностью в (6) в виде:

$$\frac{\widehat{R}_J(f) - R_J(f)}{D_J} = \frac{1}{D_J} \sum_{j=0}^{[pJ]} \sum_{k=0}^{2^j-1} [F_{jk} - \mathbb{E}F_{jk}] + \frac{1}{D_J} \sum_{j=[pJ]+1}^{J-1} \sum_{k=0}^{2^j-1} [F_{jk} - \mathbb{E}F_{jk}]. \quad (7)$$

Учитывая вид T_j , можно убедиться, что существует такая константа $C_F > 0$, что

$$|F_{jk} - \mathbb{E}F_{jk}| \leq C_F J 2^{(J-j)(1-\alpha)} \text{ п. в.} \quad (8)$$

Таким образом,

$$\left| \sum_{i=0}^{[pJ]} \sum_{l=0}^{2^i-1} [F_{il} - \mathbb{E}F_{il}] \right| \leq C_F \sum_{i=0}^{[pJ]} \sum_{l=0}^{2^i-1} J 2^{(J-i)(1-\alpha)} \leq C'_F J 2^{J(1-\alpha+\alpha p)} \text{ п. в.}$$

с некоторой константой $C'_F > 0$. Следовательно, так как $1 - \alpha + \alpha p < 1/2$, первая сумма в (7) стремится к нулю п. в. при $J \rightarrow \infty$.

Повторяя рассуждения работ [9, 11], можно убедиться, что последовательность $\{F_{jk}\}$ обладает свойством ρ -перемешивания и, следовательно, обладает свойством α -перемешивания [14].

Далее, рассуждая как в работе [6] и используя (2), можно показать, что при выполнении условий теоремы существует такая константа $C_\alpha > 0$, зависящая только от α и выбранного вейвлет-базиса, что

$$\lim_{J \rightarrow \infty} \frac{1}{D_J^2} D \left[\sum_{j=[pJ]+1}^{J-1} \sum_{k=0}^{2^j-1} [F_{jk} - \mathbb{E}F_{jk}] \right] = 1.$$

Кроме того, легко видеть, что

$$\sup_{J>0} \frac{1}{D_J^2} \sum_{j=[pJ]+1}^{J-1} \sum_{k=0}^{2^j-1} D F_{jk} < \infty.$$

Наконец, выполнено условие Линдеберга: для любого $\varepsilon > 0$

$$\frac{1}{D_J^2} \sum_{j=[pJ]+1}^{J-1} \sum_{k=0}^{2^j-1} \mathbb{E} (F_{jk} - \mathbb{E}F_{jk})^2 \mathbf{1} (|F_{jk} - \mathbb{E}F_{jk}| > \varepsilon D_J) \rightarrow 0 \text{ при } J \rightarrow \infty. \quad (9)$$

Действительно, так как

$$|F_{jk} - \mathbb{E}F_{jk}| \leq C_F J 2^{(J-j)(1-\alpha)} \text{ п. в. ,}$$

а $D_J^2 = C_\alpha 2^J$, то при $\alpha > 1/2$, начиная с некоторого J , все индикаторы в (9) обращаются в ноль.

Таким образом, выполнены все условия теоремы 2.1 из работы [15] и справедливо (6). Теорема доказана.

Оценка риска $\widehat{R}_J(f)$ также является сильно состоятельной, причем при более слабых ограничениях на α и регулярность f .

Теорема 2. Пусть функция $f \in L^2([0, 1])$. Тогда

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

при любом $\lambda > 1/2$ в случае $1/2 \leq \alpha < 1$ и любом $\lambda > 1 - \alpha$ в случае $0 < \alpha < 1/2$.

Принимая во внимание оценку (8), доказательство этого утверждения практически полностью повторяет доказательство соответствующего свойства оценки риска в работе [7].

Литература

1. Mallat S. A wavelet tour of signal processing. — New York, NY, USA: Academic Press, 1999. 857 p.
2. Donoho D., Johnstone I. M. Adapting to unknown smoothness via wavelet shrinkage // J. Am. Stat. Assoc., 1995. Vol. 90. P. 1200–1224.
3. Маркин А. В. Предельное распределение оценки риска при пороговой обработке вейвлет-коэффициентов // Информатика и её применения, 2009. Т. 3. Вып. 4. С. 57–63.
4. 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. № 1. С. 26–34.
5. Шестаков О. В. Асимптотическая нормальность оценки риска пороговой обработки вейвлет-коэффициентов при выборе адаптивного порога // Докл. РАН, 2012. Т. 445. № 5. С. 513–515.
6. Ерошенко А. А., Шестаков О. В. Асимптотические свойства оценки риска при пороговой обработке вейвлет-коэффициентов в модели с коррелированным шумом // Информатика и её применения, 2014. Т. 8. Вып. 1. С. 36–44.
7. Шестаков О. В. Сходимость почти всюду оценки риска пороговой обработки вейвлет-коэффициентов в модели с коррелированным шумом // Вестн. Моск.

- ун-та. Сер. 15: Вычисл. матем. и киберн., 2016. № 3. С. 19–22.
8. Huang H.-C., Lee T.C.M. Stabilized thresholding with generalized sure for image denoising // IEEE 17th Conference (International) on Image Processing Proceedings. — IEEE, 2010. P. 1881–1884.
 9. Johnstone I.M., Silverman B.W. Wavelet threshold estimates for data with correlated noise // J. Roy. Stat. Soc. B, 1997. Vol. 59. P. 319–351.
 10. Taqqu M.S. Weak convergence to fractional Brownian motion and to the Rosenblatt process // Z. Wahrscheinlichkeit., 1975. Vol. 31. P. 287–302.
 11. Johnstone I.M. Wavelet shrinkage for correlated data and inverse problems adaptivity results // Stat. Sinica, 1999. Vol. 9. P. 51–83.
 12. Donoho D., Johnstone I.M. Ideal spatial adaptation via wavelet shrinkage // Biometrika, 1994. Vol. 81. No. 3. P. 425–455.
 13. Ерошенко А. А. Статистические свойства оценок сигналов и изображений при пороговой обработке коэффициентов в вейвлет-разложениях. Дис. . . . канд. физ.-мат. наук. — М.: МГУ, 2015. 82 с.
 14. Bradley R. C. Basic properties of strong mixing conditions. A survey and some open questions // Probab. Surveys, 2005. Vol. 2. P. 107–144.
 15. Peligrad M. On the asymptotic normality of sequences of weak dependent random variables // J. Theor. Probab., 1996. Vol. 9. No. 3. P. 703–715.

Поступила в редакцию 09.10.17

UNBIASED RISK ESTIMATE OF STABILIZED HARD THRESHOLDING IN THE MODEL WITH A LONG-RANGE DEPENDENCE

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: De-noising methods for processing signals and images, based on the thresholding of wavelet decomposition coefficients, have become popular due to their simplicity, speed, and the ability to adapt to signal functions that have a different degree of regularity at different locations. An analysis of inaccuracies of these methods is an important practical task, since it makes it possible to evaluate the quality of both the methods themselves and the equipment used for processing. The present author considers the recently proposed stabilized hard thresholding method which avoids the main disadvantages of the popular soft and hard thresholding techniques. The statistical properties of this method are studied. In the model with an additive Gaussian noise, the author analyzes the unbiased risk estimate. Assuming that the noise coefficients have a long-range dependence, the author formulates the conditions under which strong consistency and asymptotic normality of the unbiased risk estimate take place. The results obtained make it possible to construct asymptotic confidence intervals for the threshold processing errors using only observable data.

Keywords: wavelets; thresholding; unbiased risk estimate; correlated noise; asymptotic normality

DOI: 10.14357/19922264180202

Acknowledgments

The work was partly supported by the Russian Foundation for Basic Research (project 16-07-00736).

References

1. Mallat, S. 1999. *A wavelet tour of signal processing*. New York, NY: Academic Press. 857 p.
2. Donoho, D., and I.M. Johnstone. 1995. Adapting to unknown smoothness via wavelet shrinkage. *J. Am. Stat. Assoc.* 90:1200–1224.
3. 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.
4. Markin, A. V., and O. V. Shestakov. 2010. Consistency of risk estimation with thresholding of wavelet coefficients. *Mosc. Univ. Comput. Math. Cybern.* 34(1):22–30.

5. Shestakov, O. V. 2012. Asymptotic normality of adaptive wavelet thresholding risk estimation. *Dokl. Math.* 86(1):556–558.
6. Eroshenko, A. A., and O. V. Shestakov. 2014. Asimptoticheskie svoystva otsenki riska pri porogovoy obrabotke veyvlet-koeffitsientov v modeli s korrelirovannym shumom [Asymptotic properties of wavelet thresholding risk estimate in the model of data with correlated noise]. *Informatika i ee Primeneniya — Inform. Appl.* 8(1):36–44.
7. Shestakov, O. V. 2016. Almost everywhere convergence of a wavelet thresholding risk estimate in a model with correlated noise. *Mosc. Univ. Comput. Math. Cybern.* 40(3):114–117.
8. Huang, H.-C. and T. C. M. Lee. 2010. Stabilized thresholding with generalized sure for image denoising. *IEEE 17th Conference (International) on Image Processing Proceedings*. IEEE. 1881–1884.
9. Johnstone, I. M., and B. W. Silverman. 1997. Wavelet threshold estimates for data with correlated noise. *J. Roy. Stat. Soc. B* 59:319–351.
10. Taqqu, M. S. 1975. Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Z. Wahrscheinlichkeit.* 31:287–302.
11. Johnstone, I. M. 1999. Wavelet shrinkage for correlated data and inverse problems adaptivity results. *Stat. Sinica* 9:51–83.
12. Donoho, D., and I. M. Johnstone. 1994. Ideal spatial adaptation via wavelet shrinkage. *Biometrika* 81(3):425–455.
13. Eroshenko, A. A. 2015. Statisticheskie svoystva otsenok signalov i izobrazheniy pri porogovoy obrabotke koeffitsientov v veyvlet-razlozheniyakh [Statistical properties of signal and image estimates under thresholding of coefficients in wavelet decompositions]. Moscow: MSU. PhD Diss. 82 p.
14. Bradley, R. C. 2005. Basic properties of strong mixing conditions. A survey and some open questions. *Probab. Surveys* 2:107–144.
15. Peligrad, M. 1996. On the asymptotic normality of sequences of weak dependent random variables. *J. Theor. Probab.* 9(3):703–715.

Received October 9, 2017

Contributor

Shestakov Oleg V. (b. 1976) — Doctor of Sciences 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

МИНИМИЗАЦИЯ ОШИБОК ВЫЧИСЛЕНИЯ ВЕЙВЛЕТ-КОЭФФИЦИЕНТОВ ПРИ РЕШЕНИИ ОБРАТНЫХ ЗАДАЧ*

А. А. Кудрявцев¹, О. В. Шестаков²

Аннотация: Статистические обратные задачи возникают во многих прикладных областях, включая медицину, астрономию, биологию, физику плазмы, химию и т.п. При этом в наблюдаемых данных всегда присутствуют погрешности, связанные с несовершенством оборудования, фоновыми шумами, дискретизацией данных и др. Для уменьшения этих погрешностей необходимо применять специальные методы регуляризации, позволяющие строить приближенные устойчивые решения обратных задач. Классические методы регуляризации базируются на использовании оконного сингулярного разложения. Однако при таком подходе учитывается лишь вид оператора, участвующего в формировании наблюдаемых данных, и никак не учитываются свойства самого объекта наблюдения. Для линейных однородных операторов эта проблема решается с помощью специальных методов вейвлет-анализа, позволяющих адаптироваться одновременно к виду оператора и локальным особенностям функции, описывающей объект. В данной работе рассматривается задача обращения линейного однородного оператора при наличии шума в наблюдаемых данных с помощью пороговой обработки коэффициентов вейвлет-разложения наблюдаемой функции. Вычисляются асимптотически оптимальные пороги и порядки функции потерь при минимизации усредненной вероятности ошибки вычисления вейвлет-коэффициентов.

Ключевые слова: вейвлеты; пороговая обработка; линейный однородный оператор; функция потерь

DOI: 10.14357/19922264180203

1 Введение

Математические модели, лежащие в основе многих прикладных задач анализа и обработки сигналов, предполагают решение задачи обращения некоторого линейного оператора. Например, в медицинских исследованиях бывает необходимо с помощью неинвазивных методов получить изображение структуры какого-либо внутреннего органа. В этом случае возникает задача реконструкции изображения путем обращения проекционного оператора. Как правило, эта задача является некорректно поставленной, и для ее решения требуется использовать методы регуляризации, поскольку в реальных наблюдениях всегда присутствует шум. Классические методы, основанные на сингулярном разложении в сочетании с линейной регуляризацией, адаптируются к виду оператора, но не принимают в расчет свойства функции сигнала, которую необходимо оценить. При рассмотрении моделей с линейными однородными операторами с этими

недостатками позволяют справиться методы вейвлет-разложения, предложенные в работах [1, 2].

В данной работе рассматривается метод нелинейной регуляризации при обращении линейных однородных операторов, основанный на пороговой обработке, и вычисляются параметры этого метода, асимптотически оптимальные в смысле функции потерь, основанной на вероятностях ошибок вычисления коэффициентов вейвлет-разложения.

2 Метод обращения линейных однородных операторов

Линейный оператор K называется однородным с показателем $\alpha > 0$, если

$$K [f(a(x - x_0))] = a^{-\alpha} (Kf) [a(x - x_0)]$$

для любого $x_0 \in \mathbb{R}$ и любого $a > 0$.

*Работа выполнена при частичной финансовой поддержке РФФИ (проект 18-07-00252).

¹Московский государственный университет им. М. В. Ломоносова, факультет вычислительной математики и кибернетики, nubigena@mail.ru

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

Для построения оценки функции сигнала f в данной работе рассматривается метод вейвлет-вейвлет-разложения (vaguelette-wavelet decomposition). Функция Kf представляется в виде ряда

$$Kf = \sum_{j,k \in \mathbb{Z}} \langle Kf, \psi_{j,k} \rangle \psi_{j,k}$$

по некоторому вейвлет-базису $\{\psi_{j,k}\}$. Здесь $\psi_{j,k}(x) = 2^{j/2} \psi(2^j x - k)$; $\psi(x)$ — заданная материнская вейвлет-функция. Обозначим $\beta_{j,k} = \|K^{-1} \psi_{j,k}\|$. Тогда в силу однородности оператора $\beta_{j,k} = \beta_{0,0} 2^{\alpha j}$ [2]. Функция f представляется в виде ряда:

$$f = \sum_{j,k \in \mathbb{Z}} \beta_{j,k} \langle Kf, \psi_{j,k} \rangle u_{j,k}, \quad (1)$$

где $u_{j,k} = K^{-1} \psi_{j,k} / \beta_{j,k}$. Функции $u_{j,k}$ называются «вейвлетами». Последовательность $\{u_{j,k}\}$ не является ортонормированной, однако если выполнены некоторые условия гладкости на $K^* \psi$ и $K^{-1} \psi$, то последовательность $\{u_{j,k}\}$ образует устойчивый базис в $L^2(\mathbb{R})$ [3]. Формула (1) и представляет собой метод обращения, называемый вейвлет-вейвлет-разложением.

3 Модель данных

Рассмотрим следующую модель данных:

$$X_i = (Kf)_i + z_i, \quad i = 1, \dots, 2^J,$$

где X_i — наблюдаемые данные; K — некоторый линейный однородный оператор; f — истинная функция сигнала; z_i — случайные погрешности измерения, которые предполагаются независимыми и имеющими одинаковое гауссово распределение с нулевым средним и дисперсией σ^2 .

Дискретное вейвлет-преобразование представляет собой умножение вектора значений функции Kf на ортогональную матрицу W , определяемую вейвлет-функцией ψ [4]. При этом дискретные вейвлет-коэффициенты приближенно равны $2^{J/2} \langle Kf, \psi_{jk} \rangle$. Это приближение тем точнее, чем больше J . Таким образом, для дискретных вейвлет-коэффициентов разложения Kf принимается следующая модель:

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

где $\mu_{j,k} = 2^{J/2} \langle Kf, \psi_{jk} \rangle$, а случайные величины $\zeta_{j,k}$ в силу ортогональности W также независимы и имеют нормальное распределение с нулевым средним и дисперсией σ^2 .

4 Функция потерь для метода пороговой обработки

Одним из самых популярных методов подавления шума является пороговая обработка, смысл которой заключается в обнулении коэффициентов, чьи абсолютные значения не превышают заданного порога.

Через $\hat{Y}_{j,k}$ обозначим оценку вейвлет-коэффициента, которая получается с помощью пороговой функции $\rho_T(x)$ с порогом T : $\hat{Y}_{j,k} = \rho_T(Y_{j,k})$. В данной работе рассматриваются функции жесткой пороговой обработки $\rho_T^{(h)}(x) = x \mathbf{1}(|x| > T)$ и мягкой пороговой обработки $\rho_T^{(s)}(x) = \text{sign}(x)(|x| - T)_+$.

Рассмотрим функцию потерь, основанную на вероятностях ошибок вычисления вейвлет-коэффициентов. Пусть (ξ, η) — двумерная случайная величина, не зависящая от всех $\zeta_{j,k}$ и имеющая дискретное равномерное распределение на множестве индексов $j = 0, \dots, J-1, k = 0, \dots, 2^j - 1$. Для заданного критического уровня $\varepsilon > 0$ определим функцию потерь:

$$r_J(f) = \text{EP} \left(\left| \beta_{\xi,\eta} \hat{Y}_{\xi,\eta} - \beta_{\xi,\eta} \mu_{\xi,\eta} \right| > \varepsilon \mid \xi, \eta \right) = \frac{\sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} \text{P} \left(\left| \beta_{j,k} \hat{Y}_{j,k} - \beta_{j,k} \mu_{j,k} \right| > \varepsilon \right)}{2^J},$$

т. е. $r_J(f)$ представляет собой усредненную вероятность того, что ошибка вычисления вейвлет-коэффициента превысит критический уровень ε . Такое определение функции потерь является обобщением определения, предложенного в работе [5]. В той же работе показано, что оценки, целью которых является минимизация функции потерь, основанной на вероятностях ошибок вычисления вейвлет-коэффициентов, дают сравнимые, а иногда и лучшие результаты, чем оценки, минимизирующие среднеквадратичный риск.

Целью данной работы является поиск асимптотически оптимального порога и оценивание максимального порядка функции потерь r_J пороговой обработки наблюдаемого сигнала Kf в классе регулярных по Липшицу функций $\text{Lip}(\gamma, L)$, где $\gamma > 0$ — показатель; $L > 0$ — константа Липшица [4]:

$$R_J = \sup_{Kf \in \text{Lip}(\gamma, L)} r_J(f), \quad (2)$$

т. е. поиск порога, асимптотически оптимального в минимаксном смысле. Подробное исследование поведения асимптотически оптимального порога для среднеквадратичного риска можно найти в работах [6, 7]. Также в [8] был предложен метод поиска адаптивного оптимального порога, с помощью ко-

того можно оценить среднеквадратичный риск пороговой обработки конкретной функции. Этот метод основан на построении несмещенной оценки риска, статистические свойства которой подробно исследованы в работах [9, 10]. Для задачи оценивания функции сигнала по прямым зашумленным наблюдениям асимптотически оптимальные значения порогов, минимизирующие функцию потерь, основанную на вероятностях ошибок, в том числе для моделей с негауссовым распределением шума, найдены в работах [11, 12].

Заметим, что «разумный» порог должен возрастать по J со скоростью, не выше чем $C_T \sqrt{\ln 2^J}$, где C_T — положительная константа, зависящая от оператора K (подробнее см. [2, 7]).

Далее символом \asymp обозначается порядок рассматриваемой величины по J , т.е. $a_J \asymp b_J$, если начиная с некоторого J выполнено $C_1 b_J \leq a_J \leq C_2 b_J$ для некоторых положительных констант C_1 и C_2 .

Везде далее дополнительно будем предполагать, что вейвлет-функция имеет M нулевых моментов ($M \geq \gamma$), M раз непрерывно дифференцируема и достаточно быстро убывает на бесконечности вместе со своими производными [4].

5 Жесткая пороговая обработка

Пусть $\hat{Y}_{j,k} = \rho_T^{(h)}(Y_{j,k})$. Пусть функция $g(J) > 0$ сколь угодно медленно убывает по J к нулю.

Теорема 1. Пусть $\gamma > \alpha$. При выборе асимптотически оптимального порога для жесткой пороговой обработки функция потерь (2) начиная с некоторого J удовлетворяет неравенствам:

$$C_1^{(h)} \cdot 2^{-(2\gamma-2\alpha)J/(2\gamma-2\alpha+1)} g(J) \leq R_J \leq C_2^{(h)} \cdot 2^{-((2\gamma-2\alpha)/(2\gamma-2\alpha+1))J} J^{1/(2\gamma-2\alpha+1)},$$

где $C_1^{(h)}$ и $C_2^{(h)}$ — некоторые положительные константы. Для асимптотически оптимального значения порога, минимизирующего порядок функции потерь (2) при жесткой пороговой обработке, справедливо неравенство:

$$T_*^{(h)} - T_2^{(h)} \leq T \leq T_*^{(h)} - T_1^{(h)},$$

где

$$T_*^{(h)} = \sigma \sqrt{\frac{4\gamma - 4\alpha}{2\gamma - 2\alpha + 1} \ln 2^J};$$

$$T_1^{(h)} = \sigma \sqrt{\frac{2\gamma - 2\alpha + 1}{4\gamma - 4\alpha} \frac{\ln \left((\ln 2^J)^{1/2} g(J) \right)}{\sqrt{\ln 2^J}}};$$

$$T_2^{(h)} = \sigma \sqrt{\frac{2\gamma - 2\alpha + 1}{4\gamma - 4\alpha} \frac{\ln \left((\ln 2^J)^{1/2} J^{1/(2\gamma-2\alpha+1)} \right)}{\sqrt{\ln 2^J}}}$$

Доказательство. При выполнении приведенных выше ограничений на вейвлет-функцию справедливо неравенство [4]:

$$|\mu_{j,k}| \leq \frac{A 2^{J/2}}{2^{j(\gamma+1/2)}},$$

$$j = 0, \dots, J-1, \quad k = 0, \dots, 2^j - 1, \quad (3)$$

где A — некоторая константа, зависящая от γ и L и не зависящая от J .

Поскольку $\gamma > \alpha$, неравенство (3) позволяет разбить все множество индексов $\{0, \dots, J-1\}$ на три класса в зависимости от величины $|\mu_{j,k}|$. Пусть индексы j_1 и j_2 ($j_1 < j_2$) таковы, что

$$|\mu_{j,k}| \leq \begin{cases} 2^{-\alpha j} (g(J))^{-(\gamma-\alpha+1/2)}, & j_1 \leq j \leq j_2 - 1; \\ 2^{-\alpha j} J^{-1/2}, & j_2 \leq j \leq J-1. \end{cases}$$

При этом в силу (3)

$$\left. \begin{aligned} j_1 &= \frac{J}{2\gamma - 2\alpha + 1} + \log_2 g(J) + \frac{\log_2 A}{\gamma - \alpha + 1/2}; \\ j_2 &= \frac{J}{2\gamma - 2\alpha + 1} + \frac{\log_2 J}{2\gamma - 2\alpha + 1} + \\ &\quad + \frac{\log_2 A}{\gamma - \alpha + 1/2}. \end{aligned} \right\} (4)$$

Разобьем сумму в числителе (2) на три составляющие:

$$\begin{aligned} &\sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} \mathbb{P} \left(\left| \beta_{j,k} \hat{Y}_{j,k} - \beta_{j,k} \mu_{j,k} \right| > \varepsilon \right) = \\ &= \sum_{j=0}^{j_1-1} \sum_{k=0}^{2^j-1} \mathbb{P} \left(\left| \beta_{j,k} \hat{Y}_{j,k} - \beta_{j,k} \mu_{j,k} \right| > \varepsilon \right) + \\ &+ \sum_{j=j_1}^{j_2-1} \sum_{k=0}^{2^j-1} \mathbb{P} \left(\left| \beta_{j,k} \hat{Y}_{j,k} - \beta_{j,k} \mu_{j,k} \right| > \varepsilon \right) + \\ &\sum_{j=j_2}^{J-1} \sum_{k=0}^{2^j-1} \mathbb{P} \left(\left| \beta_{j,k} \hat{Y}_{j,k} - \beta_{j,k} \mu_{j,k} \right| > \varepsilon \right) \equiv \\ &\equiv S_1 + S_2 + S_3. \end{aligned} \quad (5)$$

Рассмотрим S_3 . Заметим, что для любого $\varepsilon > 0$ найдется такое $J_0 = J_0(\varepsilon)$, что $2^{-\alpha j} J^{-1/2} \leq \varepsilon / \beta_{j,k}$ и $\varepsilon / \beta_{j,k} \leq cT$, $0 < c < 1$, для всех $J > J_0$, и найдется такое $J_1 = J_1(\varepsilon, c) \geq J_0$, что для всех $J > J_1$ имеют место соотношения:

$$\begin{aligned} \mu_{j,k} + \frac{\varepsilon}{\beta_{j,k}} &\geq 0; & \mu_{j,k} - \frac{\varepsilon}{\beta_{j,k}} &\leq 0; \\ \mu_{j,k} + \frac{\varepsilon}{\beta_{j,k}} &\leq T; & \mu_{j,k} - \frac{\varepsilon}{\beta_{j,k}} &\geq -T \end{aligned}$$

для $j_2 \leq j \leq J-1$. Следовательно, для одного слагаемого из S_3 имеем при $J > J_1$:

$$\begin{aligned}
 & \mathbb{P} \left(\left| \beta_{j,k} \hat{Y}_{j,k} - \beta_{j,k} \mu_{j,k} \right| > \varepsilon \right) = \\
 & = \mathbb{P} \left(\left| \rho_T^{(h)}(Y_{j,k}) - \mu_{j,k} \right| > \frac{\varepsilon}{\beta_{j,k}} \right) = \\
 & = \mathbb{P} \left(\mathbf{1}(|Y_{j,k}| > T) Y_{j,k} > \mu_{j,k} + \frac{\varepsilon}{\beta_{j,k}} \right) + \\
 & + \mathbb{P} \left(\mathbf{1}(|Y_{j,k}| > T) Y_{j,k} < \mu_{j,k} - \frac{\varepsilon}{\beta_{j,k}} \right) = \\
 & = \mathbb{P} \left(Y_{j,k} > T, Y_{j,k} > \mu_{j,k} + \frac{\varepsilon}{\beta_{j,k}} \right) + \\
 & + \mathbb{P} \left(Y_{j,k} < -T, Y_{j,k} < \mu_{j,k} - \frac{\varepsilon}{\beta_{j,k}} \right) = \\
 & = \mathbb{P}(|Y_{j,k}| > T) \asymp \\
 & \asymp \frac{\exp\{-(T + \mu_{j,k})^2 / (2\sigma^2)\}}{T} + \\
 & + \frac{\exp\{-(T - \mu_{j,k})^2 / (2\sigma^2)\}}{T}, \quad (6)
 \end{aligned}$$

поскольку $Y_{j,k}$ имеют нормальное распределение со средним $\mu_{j,k}$ и дисперсией σ^2 , а для стандартной нормальной функции распределения при достаточном больших x справедливо соотношение:

$$1 - \Phi(x) \asymp \frac{\Phi'(x)}{x}.$$

Заметим, что при $j_2 \leq j \leq J - 1$ величины $|\mu_{j,k}|$ не влияют на порядок правой части (6). Учитывая, что число слагаемых в S_3 имеет порядок 2^J , получаем:

$$\begin{aligned}
 S_3 & = \sum_{j=j_2}^{J-1} \sum_{k=0}^{2^j-1} \mathbb{P} \left(\left| \beta_{j,k} \hat{Y}_{j,k} - \beta_{j,k} \mu_{j,k} \right| > \varepsilon \right) \asymp \\
 & \asymp \frac{2^J \exp\{-T^2 / (2\sigma^2)\}}{T}. \quad (7)
 \end{aligned}$$

Найдем верхнюю оценку для функции потерь (2) при жесткой пороговой обработке. Для этого предположим, что все слагаемые в суммах S_1 и S_2 из (5) отделены от нуля некоторой константой. Поэтому из (4) получаем, что

$$\begin{aligned}
 S_1 + S_2 & = \sum_{j=0}^{j_2-1} \sum_{k=0}^{2^j-1} \mathbb{P} \left(\left| \beta_{j,k} \hat{Y}_{j,k} - \beta_{j,k} \mu_{j,k} \right| > \varepsilon \right) \asymp \\
 & \asymp 2^{j_2} \asymp 2^{J/(2\gamma-2\alpha+1)} J^{1/(2\gamma-2\alpha+1)}. \quad (8)
 \end{aligned}$$

Порог $T_*^{(h)} - T_2^{(h)}$ обеспечивает равенство порядков правых частей (7) и (8) и, таким образом, является нижней границей (с точностью до величины порядка $O(1/\sqrt{\ln 2^J})$) для асимптотически оптимального в смысле функции потерь R_J порога.

Теперь найдем нижнюю границу для функции потерь (2). Заметим, что найдется такая функция $Kf \in \text{Lip}(\gamma, L)$, что в неравенстве (3) будет достигаться равенство для $0 \leq j \leq j_1 - 1$ [4]. Следовательно, существует такое $J_2 > 0$, что для всех $\varepsilon > 0$ и $J > J_2$ выполняется $|\mu_{j,k}| > \varepsilon/\beta_{j,k}$ при $0 \leq j \leq j_1 - 1$. Тогда

$$\begin{aligned}
 & \mathbb{P} \left(\left| \beta_{j,k} \hat{Y}_{j,k} - \beta_{j,k} \mu_{j,k} \right| > \varepsilon \right) \geq \\
 & + \geq \mathbb{P} \left(|Y_{j,k} - \mu_{j,k}| > \frac{\varepsilon}{\beta_{j,k}} \right) \geq 2 - 2\Phi \left(\frac{\varepsilon}{\sigma\beta_{j,k}} \right).
 \end{aligned}$$

В этом случае порядок суммы S_1 в (5) равен числу слагаемых, т. е.

$$\begin{aligned}
 S_1 & = \sum_{j=0}^{j_1-1} \sum_{k=0}^{2^j-1} \mathbb{P} \left(\left| \beta_{j,k} \hat{Y}_{j,k} - \beta_{j,k} \mu_{j,k} \right| > \varepsilon \right) \asymp 2^{j_1} \asymp \\
 & \asymp 2^{J/(2\gamma-2\alpha+1)} g(J). \quad (9)
 \end{aligned}$$

Приравняем порядки S_1 и S_3 из (9) и (7). В этом случае порог равен $T_*^{(h)} - T_1^{(h)}$.

Заметим, что сумма S_2 в данных рассуждениях не присутствует. Это означает, что истинное значение R_J имеет порядок не ниже данного, т. е. рассматриваемый порядок является нижней оценкой для истинного порядка функции потерь, а $T_*^{(h)} - T_1^{(h)}$ — верхней границей для асимптотически оптимального порога T , поскольку, чтобы не удалить важные компоненты функции сигнала, следует выбирать наименьший порог, не ухудшающий порядок функции потерь.

Теорема доказана.

Замечание 1. Пороги $T_*^{(h)} - T_2^{(h)}$ и $T_*^{(h)} - T_1^{(h)}$ имеют одинаковую возрастающую по J компоненту $T_*^{(h)}$, причем $|T_1^{(h)} - T_2^{(h)}|$ стремится к нулю. Это означает, что истинное значение порога, минимизирующего порядок функции потерь при жесткой пороговой обработке, также будет иметь главную часть $T_*^{(h)}$.

6 Мягкая пороговая обработка

Пусть $\hat{Y}_{j,k} = \rho_T^{(s)}(Y_{j,k})$. Пусть функция $g_1(J) > 0$ сколь угодно медленно убывает по J к нулю, а $g_2(J) > 0$ неограниченно возрастает по J , причем

$$\ln g_2(J) = o\left(\sqrt{\ln 2^J}\right), \quad J \rightarrow \infty. \quad (10)$$

Теорема 2. Пусть $\gamma > \alpha - 1/2$. При выборе асимптотически оптимального порога для мягкой пороговой

обработки функция потерь (2), начиная с некоторого J , удовлетворяет неравенствам:

$$C_1^{(s)} \cdot 2^{-(2\gamma-2\alpha)J/(2\gamma-2\alpha+1)} g_1(J) \leq R_J \leq C_2^{(s)} \cdot 2^{-(2\gamma-2\alpha)J/(2\gamma-2\alpha+1)} g_2(J),$$

где $C_1^{(s)}$ и $C_2^{(s)}$ — некоторые положительные константы. Для асимптотически оптимального значения порога, минимизирующего порядок функции потерь (2) при мягкой пороговой обработке, справедливо неравенство:

$$T_*^{(s)} - T_2^{(s)} \leq T \leq T_*^{(s)} - T_1^{(s)},$$

где

$$T_*^{(s)} = \sigma \sqrt{\frac{4\gamma - 4\alpha}{2\gamma - 2\alpha + 1} \ln 2^J};$$

$$T_i^{(s)} = \sigma \sqrt{\frac{2\gamma - 2\alpha + 1}{4\gamma - 4\alpha} \frac{\ln \left((\ln 2^J)^{1/2} g_i(J) \right)}{\sqrt{\ln 2^J}}}, \quad i = 1, 2.$$

Доказательство. Разобьем, как в разд. 5, все множество индексов $\{0, \dots, J-1\}$ на три класса в зависимости от величины $|\mu_{j,k}|$. Пусть индексы j_1 и j_2 ($j_1 < j_2$) таковы, что

$$|\mu_{j,k}| \leq \begin{cases} 2^{-\alpha j} (g_1(J))^{-(\gamma-\alpha+1/2)}, & j_1 \leq j \leq j_2 - 1; \\ 2^{-\alpha j} (g_2(J))^{-(\gamma-\alpha+1/2)}, & j_2 \leq j \leq J-1. \end{cases}$$

При этом в силу (3) для $i = 1, 2$

$$j_i = \frac{J}{2\gamma - 2\alpha + 1} + \log_2 g_i(J) + \frac{\log_2 A}{\gamma - \alpha + 1/2}.$$

Разобьем сумму в числителе (2) на три суммы (5).

Рассмотрим S_3 . Зафиксируем некоторое положительное число ε . Заметим, что для любого $\varepsilon > 0$ найдется такое $J_0 = J_0(\varepsilon)$, что $2^{-\alpha j} (g_2(J))^{-(\gamma-\alpha+1/2)} \leq \varepsilon/\beta_{j,k}$ для всех $J > J_0$. Следовательно, имеет место соотношение $|\mu_{j,k}| \leq \varepsilon/\beta_{j,k}$ для $j_2 \leq j \leq J-1$. Для одного слагаемого из S_3 имеем при $J > J_0$:

$$\begin{aligned} & \mathbb{P} \left(\left| \beta_{j,k} \hat{Y}_{j,k} - \beta_{j,k} \mu_{j,k} \right| > \varepsilon \right) = \\ & = \mathbb{P} \left(\left| \rho_T^{(s)}(Y_{j,k}) - \mu_{j,k} \right| > \frac{\varepsilon}{\beta_{j,k}} \right) = \\ & = \mathbb{P} \left(\mathbf{1}(Y_{j,k} > T) (Y_{j,k} - T) + \right. \\ & \left. + \mathbf{1}(Y_{j,k} < -T) (Y_{j,k} + T) - \mu_{j,k} \right| > \frac{\varepsilon}{\beta_{j,k}} \Big) = \\ & = \mathbb{P} \left(Y_{j,k} > T, Y_{j,k} > T + \mu_{j,k} + \frac{\varepsilon}{\beta_{j,k}} \right) + \end{aligned}$$

$$\begin{aligned} & + \mathbb{P} \left(Y_{j,k} < -T, Y_{j,k} < -T + \mu_{j,k} - \frac{\varepsilon}{\beta_{j,k}} \right) = \\ & = \mathbb{P} \left(|Y_{j,k} - \mu_{j,k}| > T + \frac{\varepsilon}{\beta_{j,k}} \right). \quad (11) \end{aligned}$$

Поскольку $\beta_{j,k}$ возрастает как 2^{cJ} и $Y_{j,k}$ имеют нормальное распределение со средним $\mu_{j,k}$ и дисперсией σ^2 , из (11) имеем:

$$S_3 \asymp \frac{2^J \exp\{-T^2/(2\sigma^2)\}}{T}. \quad (12)$$

Найдем верхнюю оценку для функции потерь (2) при мягкой пороговой обработке. Для этого предположим, что все слагаемые в суммах S_1 и S_2 из (5) отделены от нуля некоторой константой. В этом случае получаем:

$$S_1 + S_2 \asymp 2^{J/(2\gamma-2\alpha+1)} g_2(J). \quad (13)$$

Порог $T_*^{(s)} - T_2^{(s)}$ обеспечивает равенство порядков правых частей (12) и (13) и, таким образом, является нижней границей (с точностью до величины порядка $O(1/\sqrt{\ln 2^J})$) для асимптотически оптимального в смысле функции потерь R_J порога.

Теперь найдем нижнюю границу для функции потерь (2). Пользуясь рассуждениями, приведенными в разд. 5, ввиду убывания к нулю $g_1(J)$, для любого $\varepsilon > 0$ и достаточно больших J найдется такая функция $Kf \in \text{Lip}(\gamma, L)$, что в неравенстве (3) будет достигаться равенство и выполняться $|\mu_{j,k}| > \varepsilon/\beta_{j,k}$ при $0 \leq j \leq j_1 - 1$. Тогда поскольку

$$\begin{aligned} & \mathbb{P} \left(\left| \beta_{j,k} \hat{Y}_{j,k} - \beta_{j,k} \mu_{j,k} \right| > \varepsilon \right) = \\ & = \mathbb{P} \left(|\mu_{j,k}| > \frac{\varepsilon}{\beta_{j,k}}, |Y_{j,k}| \leq T \right) + \\ & + \mathbb{P} \left(|Y_{j,k} - T - \mu_{j,k}| > \frac{\varepsilon}{\beta_{j,k}}, Y_{j,k} > T \right) + \\ & + \mathbb{P} \left(|Y_{j,k} + T - \mu_{j,k}| > \frac{\varepsilon}{\beta_{j,k}}, Y_{j,k} < -T \right), \end{aligned}$$

учитывая, что $Y_{j,k}$ имеют нормальное распределение с максимумом плотности, достигаемым в точке $\mu_{j,k}$, получаем:

$$\mathbb{P} \left(\left| \beta_{j,k} \hat{Y}_{j,k} - \beta_{j,k} \mu_{j,k} \right| > \varepsilon \right) \geq 2 - 2\Phi \left(\frac{\varepsilon}{\sigma\beta_{0,0}} \right).$$

Отсюда следует оценка, аналогичная (9):

$$S_1 \asymp 2^{J/(2\gamma-2\alpha+1)} g_1(J). \quad (14)$$

Приравняем порядки (12) и (14). В этом случае порог равен $T_*^{(s)} - T_1^{(s)}$. Воспользовавшись рассуждениями, приведенными в разд. 5, получаем, что

порог $T_*^{(s)} - T_1^{(s)}$ является верхней оценкой для истинного значения асимптотически оптимального порога T .

Теорема доказана.

Замечание 2. Пороги $T_*^{(s)} - T_1^{(s)}$ и $T_*^{(s)} - T_2^{(s)}$ имеют одинаковую возрастающую по J компоненту $T_*^{(s)}$, причем, поскольку выполнено (10), $|T_1^{(s)} - T_2^{(s)}|$ стремится к нулю. Это означает, что истинное значение порога, минимизирующего порядок функции потерь при мягкой пороговой обработке, также будет иметь главную часть $T_*^{(s)}$.

Замечание 3. В случае мягкой пороговой обработки функции $g_i(J)$, $i = 1, 2$, фигурирующие в верхней и нижней оценках, имеют сколь угодно малую скорость сходимости по J (в отличие от логарифмической функции в одной из оценок при жесткой пороговой обработке). Это позволяет сделать вывод о том, что верхняя оценка функции потерь и нижняя оценка асимптотически оптимального порога при мягкой пороговой обработке точнее, чем при жесткой.

Замечание 4. Если дополнительно предположить, что вейвлет-функция ψ имеет компактный носитель, то требование равномерной регулярности по Липшицу можно заменить на требование кусочной регулярности (подробнее см. [7]).

Литература

1. Donoho D. Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition // Appl. Comput. Harmon. A., 1995. Vol. 2. P. 101–126.
2. Abramovich F., Silverman B. W. Wavelet decomposition approaches to statistical inverse problems // Biometrika, 1998. Vol. 85. No. 1. P. 115–129.
3. Lee N. Wavelet-vaguelette decompositions and homogeneous equations. — West Lafayette, IN, USA: Purdue University, 1997. PhD Thesis. 103 p.
4. Mallat S. A wavelet tour of signal processing. — New York, NY, USA: Academic Press, 1999. 857 p.
5. Sadasivan J., Mukherjee S., Seelamantula C. S. An optimum shrinkage estimator based on minimum-probability-of-error criterion and application to signal denoising // 39th IEEE Conference (International) on Acoustics, Speech and Signal Processing Proceedings. — Piscataway, NJ, USA: IEEE, 2014. P. 4249–4253.
6. Donoho D., Johnstone I. M. Minimax estimation via wavelet shrinkage // Ann. Stat., 1998. Vol. 26. No. 3. P. 879–921.
7. Jansen M. Noise reduction by wavelet thresholding. — Lecture notes in statistics ser. — New York, NY, USA: Springer Verlag, 2001. Vol. 161. 217 p.
8. Donoho D., Johnstone I. Adapting to unknown smoothness via wavelet shrinkage // J. Am. Stat. Assoc., 1995. Vol. 90. P. 1200–1224.
9. 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. № 1. С. 26–34.
10. Шестаков О. В. Асимптотическая нормальность оценки риска пороговой обработки вейвлет-коэффициентов при выборе адаптивного порога // Докл. РАН, 2012. Т. 445. № 5. С. 513–515.
11. Кудрявцев А. А., Шестаков О. В. Асимптотическое поведение порога, минимизирующего усредненную вероятность ошибки вычисления вейвлет-коэффициентов // Докл. РАН, 2016. Т. 468. № 5. С. 487–491.
12. Кудрявцев А. А., Шестаков О. В. Асимптотически оптимальная пороговая обработка вейвлет-коэффициентов в моделях с негауссовым распределением шума // Докл. РАН, 2016. Т. 471. № 1. С. 11–15.

Поступила в редакцию 25.02.18

MINIMIZATION OF ERRORS OF CALCULATING WAVELET COEFFICIENTS WHILE SOLVING INVERSE PROBLEMS

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

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

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

Abstract: Statistical inverse problems arise in many applied fields, including medicine, astronomy, biology, plasma physics, chemistry, etc. At the same time, there are always errors in the observed data due to imperfect equipment, background noise, data discretization, and other reasons. To reduce these errors, it is necessary to apply special regularization methods that allow constructing approximate stable solutions of inverse problems. The classical

regularization methods are based on the use of windowed singular value decomposition. However, this approach takes into account only the type of operator involved in the formation of observable data and does not take into account the properties of the object of observation. For linear homogeneous operators, this problem is solved with the help of special methods of wavelet analysis, which allow adapting simultaneously to the form of the operator and local features of the function describing the object. In this paper, the authors consider the problem of inverting a linear homogeneous operator in the presence of noise in the observational data by thresholding the wavelet expansion coefficients of the observed function. The asymptotically optimal thresholds and orders of the loss function are calculated when minimizing the averaged probability of error of wavelet coefficient calculation.

Keywords: wavelets; thresholding; linear homogeneous operator; loss function

DOI: 10.14357/19922264180203

Acknowledgments

The work was partly supported by the Russian Foundation for Basic Research (project 18-07-00252).

References

1. Donoho, D. 1995. Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition. *Appl. Comput. Harmon. A.* 2:101–126.
2. Abramovich, F., and B. W. Silverman. 1998. Wavelet decomposition approaches to statistical inverse problems. *Biometrika* 85(1):115–129.
3. Lee, N. 1997. Wavelet-vaguelette decompositions and homogenous equations. West Lafayette, IN: Purdue University. PhD Thesis. 103 p.
4. Mallat, S. 1999. *A wavelet tour of signal processing*. New York, NY: Academic Press. 857 p.
5. Sadasivan, J., S. Mukherjee, and C. S. Seelamantula. 2014. An optimum shrinkage estimator based on minimum-probability-of-error criterion and application to signal denoising. *39th IEEE Conference (International) on Acoustics, Speech and Signal Processing Proceedings*. Piscataway, NJ: IEEE. 4249–4253.
6. Donoho, D., and I. M. Johnstone. 1998. Minimax estimation via wavelet shrinkage. *Ann. Stat.* 26(3):879–921.
7. Jansen, M. 2001. *Noise reduction by wavelet thresholding*. Lecture notes in statistics ser. New York, NY: Springer Verlag. Vol. 161. 217 p.
8. Donoho, D., and I. Johnstone. 1995. Adapting to unknown smoothness via wavelet shrinkage. *J. Am. Stat. Assoc.* 90:1200–1224.
9. Markin, A. V., and O. V. Shestakov. 2010. Consistency of risk estimation with thresholding of wavelet coefficients. *Mosc. Univ. Comput. Math. Cybern.* 34(1):22–30.
10. Shestakov, O. V. 2012. Asymptotic normality of adaptive wavelet thresholding risk estimation. *Dokl. Math.* 86(1):556–558.
11. Kudryavtsev, A. A., and O. V. Shestakov. 2016. Asymptotic behavior of the threshold minimizing the average probability of error in calculation of wavelet coefficients. *Dokl. Math.* 93(3):295–299.
12. Kudryavtsev, A. A., and O. V. Shestakov. 2016. Asymptotically optimal wavelet thresholding in the models with non-Gaussian noise distributions. *Dokl. Math.* 94(3):615–619.

Received February 25, 2018

Contributors

Kudryavtsev Alexey A. (b. 1978) — Candidate of Science (PhD) in physics and mathematics, associate professor, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; nubigena@mail.ru

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

ДОСТАТОЧНЫЕ УСЛОВИЯ ЭРГОДИЧНОСТИ ПРИОРИТЕТНЫХ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ*

А. В. Мистрюков¹, В. Г. Ушаков²

Аннотация: Известные результаты по эргодичности приоритетных систем массового обслуживания получены в предположении, что входящие потоки требований всех приоритетов являются пуассоновскими. В данной работе это требование ослаблено, а именно: найдены достаточные условия эргодичности систем массового обслуживания с двумя классами приоритетов, в которых только поток требований высшего приоритета является пуассоновским. Исследованы системы с относительным приоритетом и тремя разновидностями абсолютного приоритета: с дообслуживанием, обслуживанием заново и потерей прерванного требования низшего приоритета. Для получения искомым условий для последовательных времен ожидания в очереди требований каждого приоритета получены рекуррентные соотношения, известные как рекурсия Линдли. Полученная цепь Маркова исследуется методом пробных функций. Найдены достаточные условия, при которых исследуемая цепь является харрисовой и, следовательно, имеет стационарное распределение.

Ключевые слова: относительный приоритет; абсолютный приоритет; эргодичность; метод пробных функций; время ожидания; рекурсия Линдли

DOI: 10.14357/19922264180204

1 Введение

Проблема нахождения условий эргодичности традиционна для теории массового обслуживания. Эти условия важны для приложений, поскольку они определяют соотношения между параметрами модели, при выполнении которых не возникает бесконечно больших очередей. Существует обширная литература по эргодической теории случайных процессов (см., например, [1, 2]). Известно большое число достаточных условий эргодичности различных классов случайных процессов. Среди них особое место занимают марковские процессы. Большая часть условий эргодичности марковских процессов формулируется в терминах свойств переходной функции. Для теории массового обслуживания, однако, нужны условия, выраженные через параметры исследуемой системы (входящие потоки, длительности обслуживания и т. п.), получение которых в качестве следствия из общих результатов (особенно для сложных, в частности приоритетных, систем) является нетривиальной задачей.

При изучении эргодичности приоритетных систем обычно накладывались ограничения либо на время между поступлениями требований в систему, либо на время их обслуживания, а именно: предпо-

лагалось, что эти времена имеют экспоненциальное распределение (см., например, [3]). Работ, в которых исследуется эргодичность приоритетных систем, не относящихся к классам $M/G/1$ и $G/M/1$, практически нет.

При исследовании различных систем массового обслуживания часто удается найти случайный процесс w_n с дискретным временем, характеризующий работу системы (длина очереди, время ожидания, число обслуженных требований и т. п.), который удовлетворяет рекуррентным соотношениям вида:

$$w_{n+1} = \max(0, w_n + s_n - t_n).$$

В теории массового обслуживания такие соотношения принято называть рекурсией Линдли. Различные асимптотические свойства этого соотношения, в том числе условия эргодичности, достаточно хорошо изучены (см., например, [4, 5]); приложения к теории массового обслуживания можно найти в монографиях [6, 7].

В данной статье получено рекуррентное соотношение Линдли для времен ожидания неприоритетных требований и найдены достаточные условия эргодичности систем обслуживания с относительным и тремя разновидностями абсолютного приорите-

*Работа выполнена при частичной финансовой поддержке РФФИ (проект 18-07-00678).

¹Факультет вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова, unff08@rambler.ru

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

та при условии, что лишь приоритетный входящий поток является пуассоновским.

2 Определения и вспомогательные утверждения

Пусть на измеримом пространстве $(X, \sigma(X))$ задана марковская цепь $x_n, n \in N_0 = \{0, 1, 2, \dots\}$, с переходными вероятностями

$$P(x, A) = \mathbf{P}(x_{n+1} \in A | x_n = x), \quad x \in X, \quad A \in \sigma(X).$$

Определение 1. Инвариантной мерой марковской цепи с переходными вероятностями $P(x, A)$ называется вероятностная мера π на $\sigma(X)$, для которой справедливо равенство:

$$\pi(A) = \int_X \pi(dx)P(x, A) \text{ для любого } A \in \sigma(X).$$

Определение 2. Марковская цепь называется эргодической, если она имеет единственную инвариантную меру.

Для того чтобы сформулировать достаточные условия эргодичности марковских цепей, введем следующие обозначения:

$$Pf(x) = \mathbf{E}(f(x_{n+1}) | x_n = x) = \int_X f(y)P(x, dy),$$

где f — функция, определенная на X ;

$$\tau_A(x) = \inf(n : x_n \in A | x_0 = x), \quad A \in \sigma(X),$$

есть время первого попадания в множество A из точки x , а

$$\mathbf{E}_x(\tau_A) = \mathbf{E}(\inf(n : x_n \in A | x_0 = x))$$

есть среднее время первого достижения множества A из точки x .

Теорема 1. Пусть марковская цепь удовлетворяет следующим условиям: существуют множество $A \subset X, p > 0$ и вероятностная мера ν на $\sigma(X)$ такие, что

- (а) $\mathbf{P}(\tau_A(x) < \infty) = 1$ для всех $x \in A^c$;
- (б) $\sup_{x \in A} \mathbf{E}_x(\tau_A) < \infty$;
- (в) $P_m(x, B) \geq p\nu(B)$ для некоторого $m \in N$ и всех $B \in \sigma(X), x \in A$.

Тогда она имеет единственную инвариантную меру π . Если, кроме того, цепь неперiodична, то переходные вероятности сходятся к инвариантной мере по вариации.

Марковскую цепь, удовлетворяющую условиям теоремы 1, называют харрисовой.

Теорема 2. Пусть существуют неотрицательная функция $w(x), A \subset X$ и $\varepsilon > 0$ такие, что $Pw(x) \leq w(x) - \varepsilon$ для всех $x \in A^c$ и $\sup_{x \in A} Pw(x) < \infty$. Тогда $\sup_{x \in A} \mathbf{E}_x(\tau_A) < \infty$.

Доказательства теорем 1 и 2 можно найти, например, в [1, 2].

Лемма 1. Пусть выполнены условия теоремы 2. Тогда

$$\mathbf{E}_x(\tau_A) \leq \frac{w(x)}{\varepsilon} \text{ для всех } x \in A^c.$$

3 Условия эргодичности систем с относительным и абсолютным приоритетами

Рассматривается одноканальная система массового обслуживания с неограниченным числом мест для ожидания и двумя потоками требований. Первый поток — пуассоновский, второй — рекуррентный с абсолютно непрерывной функцией распределения интервалов между поступлениями требований. Требования первого потока имеют один из четырех видов приоритета перед требованиями второго потока:

- (1) относительный приоритет;
- (2) абсолютный приоритет с дообслуживанием прерванного требования;
- (3) абсолютный приоритет с потерей прерванного требования;
- (4) абсолютный приоритет с обслуживанием заново прерванного требования.

Требования одного приоритета обслуживаются в порядке их поступления в систему.

Обозначим через $s_1^{(i)}, s_2^{(i)}, \dots$ и $t_1^{(i)}, t_2^{(i)}, \dots$ последовательные времена обслуживания и интервалы между поступлениями требований i -го потока, $s^{(i)} = \mathbf{E}s_1^{(i)}, t^{(i)} = \mathbf{E}t_1^{(i)}, i = 1, 2, p(s) = \mathbf{E}e^{-st_1^{(2)}}, \alpha = (t^{(1)})^{-1}$.

Пусть $w_n^{(2)}$ — время ожидания до начала обслуживания n -м требованием второго потока (нумерация производится в порядке поступления в систему).

Теорема 3. Для эргодичности последовательности $w_n^{(2)}$ достаточно существования второго момента времени обслуживания приоритетных требований и выполнения условий:

- (1) для систем обслуживания с относительным и абсолютным с дообслуживанием прерванного требования приоритетами:

$$\frac{s^{(1)}}{t^{(1)}} + \frac{s^{(2)}}{t^{(2)}} < 1;$$

- (2) для систем обслуживания с абсолютным приоритетом и обслуживанием заново прерванного требования:

$$\frac{(1-p(\alpha))}{p(\alpha)} \frac{t^{(1)}}{t^{(2)}} + \frac{s^{(1)}}{t^{(1)}} < 1;$$

- (3) для систем обслуживания с абсолютным приоритетом и потерей прерванного требования:

$$(1-p(\alpha)) \frac{t^{(1)}}{t^{(2)}} + \frac{s^{(1)}}{t^{(1)}} < 1.$$

Доказательство. Для всех рассматриваемых приоритетных дисциплин справедливы следующие рекуррентные соотношения:

$$\begin{aligned} w_{n+1}^{(2)} &= I(t_{n+1}^{(2)} < w_n^{(2)} + T(s_n^{(2)})) \times \\ &\quad \times (w_n^{(2)} + T(s_n^{(2)}) - t_{n+1}^{(2)}) + \\ &\quad + I(t_{n+1}^{(2)} > w_n^{(2)} + T(s_n^{(2)})) \times \\ &\quad \times T^*(t_{n+1}^{(2)}, w_n^{(2)}, T(s_n^{(2)})). \end{aligned}$$

Здесь $T(s_n^{(2)})$ — интервал времени с момента поступления на обслуживание n -го требования второго приоритета до первого после этого момента освобождения системы от этого требования и требований более высокого приоритета; $I(A)$ — индикатор события A ; случайная величина $T^*(t_{n+1}^{(2)}, w_n^{(2)}, T(s_n^{(2)}))$ имеет такое же распределение, как время до первого освобождения системы после момента времени $t_{n+1}^{(2)} - w_n^{(2)} + T(s_n^{(2)})$, если в систему поступает только поток приоритетных требований и в начальный момент система свободна от них. В литературе, посвященной анализу приоритетных систем массового обслуживания, случайное время $T(s_n^{(2)})$ называют циклом обслуживания [3, 8]. Если поток приоритетных требований является пуассоновским, то распределение и моменты цикла

обслуживания известны. В частности, математические ожидания равны: $s^{(2)}t^{(1)}/(t^{(1)} - s^{(1)})$, $((1-p(\alpha))/p(\alpha))(t^{(1)})^2/(t^{(1)} - s^{(1)})$ и $(1 - p(\alpha))(t^{(1)})^2/(t^{(1)} - s^{(1)})$ для систем с относительным и абсолютным с дообслуживанием, абсолютным с обслуживанием заново и абсолютным с потерей приоритетами соответственно. Нетрудно найти и распределение случайной величины $T^*(t_{n+1}^{(2)}, w_n^{(2)}, T(s_n^{(2)}))$ при любых фиксированных $t_{n+1}^{(2)}$, $T(s_n^{(2)})$ и $w_n^{(2)}$, но оно в данной ситуации не понадобится. Заметим только, что при любых $u \geq 0$, $v \geq 0$, $w \geq 0$ и $u - w + v \geq 0$ существует постоянная c такая, что $\mathbf{E}T^*(u, w, v) < c$.

При сделанных предположениях последовательность $w_n^{(2)}$ является неразложимой непериодической однородной цепью Маркова. Покажем, что она эргодична, если $\mathbf{E}T(s_n^{(2)}) - \mathbf{E}t_{n+1}^{(2)} < 0$. Для этого воспользуемся результатами теорем 1 и 2.

Пусть $w_0 > 0$ таково, что

$$\begin{aligned} \varepsilon &= -E(T(s_n^{(2)}) - t_{n+1}^{(2)}) \times \\ &\quad \times I(T(s_n^{(2)}) - t_{n+1}^{(2)} > -w_0) > 0. \end{aligned}$$

Положим $A = [0, \max(w_0, c)]$. Тогда для любого $w \in A^c$ имеем:

$$\begin{aligned} \mathbf{E}(w_{n+1}^{(2)} | w_n^{(2)} = w) &= \\ &= \mathbf{E}I(t_{n+1}^{(2)} < w + T(s_n^{(2)}))(w + T(s_n^{(2)}) - t_{n+1}^{(2)}) + \\ &+ \mathbf{E}I(t_{n+1}^{(2)} > w + T(s_n^{(2)}))T^*(t_{n+1}^{(2)}, w, T(s_n^{(2)})) < \\ &< -\varepsilon + w \mathbf{E}I(t_{n+1}^{(2)} < w + T(s_n^{(2)})) + \\ &+ c \mathbf{E}I(t_{n+1}^{(2)} > w + T(s_n^{(2)})) < w - \varepsilon. \end{aligned}$$

Очевидно, что неравенство

$$\sup_{w \in A} \mathbf{E}(w_{n+1}^{(2)} | w_n^{(2)} = w) < \infty$$

также выполняется. Таким образом, условия теоремы 2, а значит, и условия (а) и (б) теоремы 1 выполнены. Покажем, что и условие (в) теоремы 1 выполнено. Обозначим

$$\tau_0(x) = \inf_{n \geq 1} (n w_n^{(2)} = 0, w_0^{(2)} = x).$$

Так как для любого x

$$\mathbf{E}(\tau_0(x)) < \infty,$$

то существует n такое, что вероятность p перехода за n шагов из состояния x в состояние 0 больше 0. В качестве меры ν из формулировки теоремы 1 можно взять дельта-меру, сосредоточенную в нуле. Таким образом, все условия теоремы 1 выполнены.

Литература

1. Walters P. An introduction to ergodic theory. — New York, NY, USA: Springer-Verlag, 1982. 259 p.
2. Боровков А. А. Эргодичность и устойчивость случайных процессов. — М.: Эдиториал УРСС, 1999. 450 с.
3. Гнеденко Б. В., Даниелян Э. А., Димитров Б. Н., Климов Г. П., Матвеев В. Ф. Приоритетные системы обслуживания. — М.: Изд-во Московского ун-та, 1973. 448 с.
4. Meyn S., Tweedie R. Markov chains and stochastic stability. — New York, NY, USA: Springer Verlag, 1993. 562 p.
5. Asmussen S. Subexponential asymptotics for stochastic processes: Extremal behavior, stationary distributions and first passage probabilities // Ann. Appl. Probab., 1998. Vol. 8. P. 354–374.
6. Cohen J. W. The single server queue. — Amsterdam: North-Holland Publishing Co., 1982. 694 p.
7. Asmussen S. Applied probability and queues. — New York, NY, USA: Springer-Verlag, 2003. 439 p.
8. Матвеев В. Ф., Ушаков В. Г. Системы массового обслуживания. — М.: Изд-во Московского ун-та, 1984. 240 с.

Поступила в редакцию 31.01.18

SUFFICIENT ERGODICITY CONDITIONS FOR PRIORITY QUEUES

A. V. Mistryukov¹ 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: Known results in ergodicity of priority queues are based on the assumption that interarrival times in each queue have exponential distribution. The aim of this paper is to relax this assumption, namely, to establish sufficient conditions of ergodicity for queues with two priority classes $GI|GI|1$ under assumption that interarrival times only in high priority class queue have exponential distribution. Queues with nonpreemptive and different kinds of preemptive priority are considered. To formulate desired conditions, the authors use Lindley’s recursion for waiting times of each priority class queue. Using Lyapunov–Foster criteria, the authors obtain sufficient conditions for a given recursion to be a Harris-ergodic Markov chain, meaning to have a unique invariant measure, to which its transition probabilities converge in total variation.

Keywords: head of the line priority; preemptive priority; ergodicity; the method of test functions; waiting time; Lindley recursion

DOI: 10.14357/19922264180204

Acknowledgments

This work was partly supported by the Russian Foundation for Basic Research (project No. 18-07-00678).

References

1. Walters, P. 1982. *An introduction to ergodic theory*. New York, NY: Springer-Verlag. 259 p.
2. Borovkov, A. A. 1999. *Ergodichnost’ i ustoychivost’ sluchaynykh protsessov* [Ergodicity and stability of stochastic processes]. Moscow: Editorial URSS. 450 p.
3. Gnedenko, B. V., E. A. Danielyan, B. N. Dimitrov, G. P. Klimov, and V. F. Matveev. 1973. *Prioritetnye sistemy obsluzhivaniya* [Priority queueing systems]. Moscow: Izd-vo Moskovskogo un-ta. 448 p.
4. Meyn, S., and R. Tweedie. 1993. *Markov chains and stochastic stability*. New York, NY: Springer Verlag. 562 p.
5. Asmussen, S. 1998. Subexponential asymptotics for stochastic processes: Extremal behavior, stationary distributions and first passage probabilities. *Ann. Appl. Probab.* 8:354–374.
6. Cohen, J. W. 1982. *The single server queue*. Amsterdam: North-Holland Publishing Co. 694 p.
7. Asmussen, S. 2003. *Applied probability and queues*. New York, NY: Springer-Verlag. 439 p.
8. Matveev, V. F., and V. G. Ushakov. 1984. *Sistemy massovogo obsluzhivaniya* [Queueing systems]. Moscow: Izd-vo Moskovskogo un-ta. 240 p.

Received January 31, 2018

Contributors

Mistryukov Andrey V. (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; unf08@rambler.ru

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

ВЕРОЯТНОСТНАЯ МОДЕЛЬ ВЛИЯНИЯ КНИГИ ЗАКАЗОВ НА ПРОЦЕСС ЦЕНЫ

Е. В. Быковец¹, В. В. Лаврентьев², Л. В. Назаров³

Аннотация: Рассматривается модель книги заказов, в которой заказы на покупку и продажу образуют два независимых процесса Кокса. Предложен механизм влияния поступающих заказов на цену актива на основе физической модели абсолютно упругого соударения. В этой модели цена представляет собой материальную точку с некоторой массой, движущуюся по прямой без трения. Приходящие заказы на покупку и уходящие заказы на продажу упруго сталкиваются с ней и придают дополнительный импульс в одном направлении, а приходящие заказы на продажу и уходящие заказы на покупку — в противоположном. Получена функциональная предельная теорема для процесса цены при высокой интенсивности входящего потока заказов, позволяющая аппроксимировать его некоторым процессом Леви.

Ключевые слова: лимитные заявки; абсолютно упругий удар; модель книги заказов; процесс цены; процесс Кокса; функциональная предельная теорема

DOI: 10.14357/19922264180205

1 Введение

Рассмотрим некоторый торгуемый на бирже актив, в отношении которого могут приходиться два вида запросов: на покупку и на продажу. Список таких запросов формирует книгу заказов для данного актива. Информация, содержащаяся в книге, позволяет делать прогнозы относительно возможного движения цены рассматриваемого актива. Особый интерес эта информация начала представлять с развитием высокочастотной торговли.

В работе рассматривается модель, которая описывает влияние книги заказов на цену актива. Базовой моделью для исследования была выбрана модель книги заказов, близкая к описанной в [1]. Основное отличие состоит в следующем: на бирже торгуемый актив имеет цену, которая размещается в узлах сетки nh , где n — некоторое целое число; h — тик, т.е. минимальное изменение цены. Однако высокая частота узлов сетки позволяет считать, что рассматриваемый актив может иметь произвольную цену, равно как и заявки на покупку и продажу торгуемого актива. С учетом этого рассматриваем следующую модель влияния заявок на цену актива, используя физическую аналогию. Рассмотрим материальную точку массой M , которая может двигаться по прямой (числовой оси) в любом направлении без трения. При этом, связывая мо-

дель физическую и математическую, будем считать, что текущее положение на оси — это текущая цена $X(t)$. Будем далее для краткости называть ценой и саму указанную материальную точку массой M , т.е. будем говорить о скорости цены, импульсе цены и т.п. Каждый заказ на продажу (поступающий по цене A_i не ниже, чем $X(t)$) придает цене дополнительный импульс в направлении от A_i к $X(t)$. Заказы живут экспоненциальное время, после чего уходят из книги за счет исполнения или отмены. Уход заказа из книги придает цене дополнительный импульс той же абсолютной величины, что и при ее поступлении, но противоположного направления. С заказами на продажу все аналогично, только приходят они с ценой, не превосходящей $X(t)$.

Цель данной работы состоит в том, чтобы выяснить, какой процесс движения цены порождает такая система при интенсивном потоке приходящих заказов. Близкая задача решалась авторами в работе [2], но там рассматривался другой механизм влияния поступающих заказов на цену. Здесь следует упомянуть и работу [3], в которой также изучается связь механизма функционирования книги заказов на микроуровне с процессом цены.

2 Описание модели

Будем рассматривать работу книги заказов на временном интервале $t \in [0, T]$. В начальный мо-

¹Московский государственный университет им. М.В. Ломоносова, факультет вычислительной математики и кибернетики, eugene.bykovets@stud.cs.msu.ru

²Московский государственный университет им. М.В. Ломоносова, факультет вычислительной математики и кибернетики, lavrent@cs.msu.ru

³Московский государственный университет им. М.В. Ломоносова, факультет вычислительной математики и кибернетики, nazarov@cs.msu.ru

мент времени $t = 0$ заказов в книге нет. Считаем, что поток приходящих заказов является процессом Кокса следующего вида:

$$\{N(\Lambda(t)) = N_1(\Lambda(t)), t \geq 0\},$$

где N_1 — пуассоновский процесс, интенсивность которого равна 1; $\Lambda(t)$ — стартующий из нуля случайный процесс, у которого траектории являются неубывающими и непрерывными справа функция-ми, а также справедливо $\mathbb{P}(\Lambda(t) < \infty)$.

Каждый заказ, приходящий в книгу, находит-ся в ней некоторое случайное время. Более точ-но, время нахождения конкретного заказа в книге является случайной величиной, распределенной по экспоненциальному закону.

Для каждого приходящего заказа определен на-бор параметров (h_i, γ_i, η_i) , где h_i — абсолютное значение разности между ценой заказа и текущей ценой; γ_i — разность между скоростью приходяще-го заказа и текущей скоростью цены; η_i — время пребывания заказа в книге.

Случайные величины h_i, γ_i и $\eta_i, i = 1, 2, \dots$, независимы в совокупности и не зависят от потока заявок, а η_i распределены экспоненциально с пара-метром μ . Параметр γ_i фактически определяет тип заказа. Для заказов на покупку γ_i положительны, для заказов на продажу — отрицательны.

В качестве физической модели влияния заказа на цену возьмем модель абсолютно упругого удара. Считаем, что i -й заказ, поступающий в момент t , — это материальная точка массой $m_0(h_i) > 0$, которая движется по той же прямой, что и цена, и имеет в момент времени t скорость, равную $u_i = v_{i-1} + \gamma_i$, где v_{i-1} — скорость цены до столкновения с i -м заказом. В момент t происходит их упругое соударение. На распределения h_i и γ_i наложим следующие ограничения:

$$\mathbb{E}\gamma_i = 0; \mathbb{E}\gamma_i^2 = \bar{\gamma} < \infty; \mathbb{E}m_0(h_i)^2 = \bar{m} < \infty. \quad (1)$$

Смысл первого условия заключается в том, что раз-ности между скоростями приходящих заказов и те-кущей скоростью актива для заказов на покупку и продажу в среднем равны. Остальные ограниче-ния имеют технический характер и лишь постули-руют конечность соответствующих моментов.

Функция m_0 является убывающей на интерва-ле $(0, \infty)$, поскольку воздействие заказа на цену тем больше, чем ближе его цена к текущей цене акти-ва. Это соответствует реальному положению дел на рынке, где заказы на уровнях, близких к текущей цене, выставляются более ответственно, так как мо-гут быть тут же удовлетворены. В то же время заказы на более удаленных уровнях чаще ставятся для дез-ориентации других участников рынка и снимаются

при приближении к ним цены. Иными словами, они не отражают реальный спрос.

3 Процесс цены

Рассмотрим точку на числовой прямой, которая представляет собой текущую цену актива. Положим M — масса точки; v_i — текущая скорость цены, по-лученная после удара i -й частицы, полагаем $v_0 = 0$; u_i — скорость i -го заказа до соударения; $m_0(h_i)$ — масса i -го заказа. Как было сказано в предыдущем разделе, данная точка (исследуемая цена) в опре-деленные моменты времени абсолютно упруго со-ударяется с другими частицами (заказами). В этом случае есть возможность выразить скорость точки после удара i -й частицы через массу точки и массу частицы, а также их скорости до столкновения (это следует из закона сохранения импульса и закона сохранения энергии, см. [4, гл. 4, § 28]):

$$v_i = -v_{i-1} + 2 \frac{Mv_{i-1} + m_0(h_i)u_i}{M + m_0(h_i)}.$$

Обозначим $\Delta v_i = v_i - v_{i-1}$, тогда

$$\begin{aligned} \Delta v_i &= -2v_{i-1} + 2 \frac{Mv_{i-1} + m_0(h_i)u_i}{M + m_0(h_i)} = \\ &= \frac{2m_0(h_i)}{M + m_0(h_i)} (u_i - v_{i-1}). \end{aligned}$$

Фактически Δv_i показывает изменение скорости цены после соударения с i -м заказом.

Пусть в начальный момент книга заказов пус-та, а начальная скорость $v_0 = 0$. Далее полагаем, что заказ с номером i приходит в момент време-ни τ_{i0} и уходит в момент времени τ_{i1} . В итоге получаем, что скорость цены является случайным процессом $\{V(t), t \geq 0\}$ с кусочно-постоянными траекториями:

$$V(t) = \sum_{i=1}^{N_1(\Lambda(t))} \Delta v_i \mathbb{I}_{\{\tau_{i0} \leq t \leq \tau_{i1}\}}(t).$$

Тогда изменение цены за время T будет иметь вид:

$$X(T) = \int_0^T V(t) dt = \sum_{i=1}^{N_1(\Lambda(T))} X_i(T), \quad (2)$$

где $X_i(T)$ — изменение цены на интервале $[0, T]$ за счет удара i -го заказа:

$$\begin{aligned} X_i(T) &= \int_0^T \Delta v_i \mathbb{I}_{\{\tau_{i0} \leq t \leq \tau_{i1}\}}(t) dt = \\ &= \frac{2m_0(h_i)}{M + m_0(h_i)} \int_0^T (u_i - v_{i-1}) \mathbb{I}_{\{\tau_{i0} \leq t \leq \tau_{i1}\}}(t) dt. \end{aligned}$$

Поскольку по определению $u_i - v_{i-1} = \gamma_i$, то последнее выражение можем переписать в виде $(a \wedge b = \min(a, b))$:

$$\begin{aligned} X_i(T) &= \frac{2m_0(h_i)\gamma_i}{M + m_0(h_i)} \int_0^T \mathbb{I}_{\{\tau_{i0} \leq t \leq \tau_{i1}\}}(t) dt = \\ &= \frac{2m_0(h_i)\gamma_i}{M + m_0(h_i)} (T \wedge \tau_{i1} - T \wedge \tau_{i0}) = \\ &= \frac{2m_0(h_i)\gamma_i}{M + m_0(h_i)} (T \wedge (\tau_{i0} + \eta_i) - T \wedge \tau_{i0}). \end{aligned}$$

Строго говоря, случайные величины $\{X_i(T), i = 1, 2, \dots\}$ не являются независимыми, но суммы в (2) можно представить в виде сумм независимых случайных величин.

Рассмотрим распределение вектора моментов прихода заявок $\tau_0 = \{\tau_{10}, \dots, \tau_{n0}\}$. По свойству пуассоновского потока при $N_1(\Lambda(T)) = n$ распределение τ_0 есть распределение вариационного ряда выборки из n независимых случайных величин, равномерно распределенных на $[0, \Lambda(T)]$. Поскольку значение конечной суммы при перестановке слагаемых не меняется, далее будем считать, что в каждой из сумм (2) τ_{i0} независимы и равномерно распределены на $[0, \Lambda(T)]$, а следовательно, случайные величины $\{X_i(T), i = 1, 2, \dots\}$ также независимы.

Изучим асимптотические свойства моментов $X_i(T)$.

Лемма 1. Пусть случайная величина ξ равномерно распределена на $[0, T]$, η_0 не зависит от ξ и имеет экспоненциальное распределение с параметром μ и

$$s = T \wedge (\xi + \eta_0) - \xi.$$

Тогда

(1) $s \stackrel{d}{=} \xi \wedge \eta_0$;

(2) моменты случайной величины s обладают следующими асимптотическими свойствами:

$$\lim_{\mu \rightarrow \infty} \mu \mathbb{E}s = 1; \quad \lim_{\mu \rightarrow \infty} \mu^2 \mathbb{E}s^2 = 2; \quad \lim_{\mu \rightarrow \infty} \mu^2 \mathbb{D}s = 1.$$

Доказательство. Вычислим математическое ожидание случайной величины s с учетом независимости ξ и η_0 . По определению

$$s = T \wedge (\xi + \eta_0) - \xi = (T - \xi) \wedge \eta_0.$$

Справедливость первого утверждения леммы следует из независимости ξ и η_0 и одинаковой распределенности ξ и $T - \xi$. Таким образом, математическое ожидание s есть

$$\mathbb{E}s = \mathbb{E}(\xi \wedge \eta_0).$$

При вычислении моментов неоднократно будет требоваться значение интеграла

$$\int_0^T y^n e^{-\mu y} dy = \frac{n!}{\mu^{n+1}} F_{n+1}(T),$$

где F_{n+1} — функция распределения Эрланга $(n+1)$ -го порядка:

$$F_{n+1}(x) = 1 - e^{-\mu x} \sum_{i=1}^n \frac{\mu^i T^i}{i!}.$$

Вычислим $\mathbb{E}(\xi \wedge \eta_0)$:

$$\begin{aligned} \mathbb{E}(\xi \wedge \eta_0) &= \frac{\mu}{T} \int_0^\infty e^{-\mu y} dy \int_0^T (x \wedge y) dx = \\ &= \frac{\mu}{T} \int_0^\infty e^{-\mu y} dy \left\{ \mathbb{I}_{\{y < T\}} \left[\int_0^y x dx + \int_y^T y dx \right] + \right. \\ &\quad \left. + \mathbb{I}_{\{y \geq T\}} \int_0^T x dx \right\} = \frac{\mu}{T} \left[T \int_0^T y e^{-\mu y} dy - \right. \\ &\quad \left. - \frac{1}{2} \int_0^T y^2 e^{-\mu y} dy + \frac{T^2}{2} \int_T^\infty e^{-\mu y} dy \right] = \\ &= \left[\frac{1}{\mu} - \frac{1}{\mu^2 T} \right] + \left[\frac{T}{2} + \frac{1}{\mu^2 T} \right] e^{-\mu T}. \end{aligned}$$

И, соответственно,

$$\begin{aligned} \lim_{\mu \rightarrow \infty} \mu \mathbb{E}s &= \\ &= \lim_{\mu \rightarrow \infty} \left\{ \left[1 - \frac{1}{\mu T} \right] + \left[\frac{T}{2} + \frac{1}{\mu^2 T} \right] \mu e^{-\mu T} \right\} = 1. \end{aligned}$$

Вычислим $\mathbb{E}s^2 = \mathbb{E}(\xi \wedge \eta_0)^2$:

$$\begin{aligned} \mathbb{E}(\xi \wedge \eta_0)^2 &= \frac{\mu}{T} \int_0^\infty e^{-\mu y} dy \int_0^T (x \wedge y)^2 dx = \\ &= \frac{\mu}{T} \int_0^\infty e^{-\mu y} dy \left\{ \mathbb{I}_{\{y < T\}} \left[\int_0^y x^2 dx + \int_y^T y^2 dx \right] + \right. \\ &\quad \left. + \mathbb{I}_{\{y \geq T\}} \int_0^T x^2 dx \right\} = \frac{\mu}{T} \left[T \int_0^T y^2 e^{-\mu y} dy - \right. \\ &\quad \left. - \frac{2}{3} \int_0^T y^3 e^{-\mu y} dy + \frac{T^3}{3} \int_T^\infty e^{-\mu y} dy \right] = \\ &= \left[\frac{2}{\mu^2} - \frac{4}{\mu^3 T} \right] + \left[\frac{2}{\mu^2} + \frac{4}{\mu^3 T} \right] e^{-\mu T}. \end{aligned}$$

Отсюда получаем асимптотику второго момента

$$\begin{aligned} \lim_{\mu \rightarrow \infty} \mu^2 \mathbb{E}s^2 &= \\ &= \lim_{\mu \rightarrow \infty} \left\{ \left[2 - \frac{4}{T\mu} \right] + \left[\frac{2}{\mu^2} + \frac{4}{\mu^3 T} \right] \mu^2 e^{-\mu T} \right\} = 2 \end{aligned}$$

и дисперсии

$$\begin{aligned} \lim_{\mu \rightarrow \infty} \mu^2 \mathbb{D}s &= \lim_{\mu \rightarrow \infty} \mu^2 [\mathbb{E}s^2 - (\mathbb{E}s)^2] = \\ &= \lim_{\mu \rightarrow \infty} \mu^2 \mathbb{E}s^2 - \lim_{\mu \rightarrow \infty} (\mu \mathbb{E}s)^2 = 1. \end{aligned}$$

Утверждение леммы доказано.

Рассмотрим следующую последовательность:

$$\left\{ X_n(t) = \sum_{i=1}^{N_1(\Lambda_n(t))} X_{ni}(t), t \in [0, T] \right\}. \quad (3)$$

При этом каждому члену $\{X_n\}$ соответствует процесс $\{\Lambda_n(t), t \in [0, T]\}$, параметр μ_n и функция массы ударяющей частицы (приходящего заказа) m_{n0} . При увеличении n будем увеличивать интенсивность входящего потока заявок $\Lambda_n(t)$ и уменьшать время пребывания заказов в книге посредством увеличения μ_n ($\Lambda_n(t) \Rightarrow \infty, \mu_n \rightarrow \infty$ при $n \rightarrow \infty$). Будем также уменьшать влияние отдельного заказа на цену:

$$m_{n0} = \alpha_n m_0, \quad \alpha_n > 0, \quad \alpha_n \rightarrow 0, \quad n \rightarrow \infty.$$

Получим асимптотические свойства моментов случайных величин $X_{n1}(T)$ при установленных параметрических зависимостях. Аргумент T у них одинаков и для краткости будем его опускать.

Лемма 2. Пусть $\mu_n \rightarrow \infty, \alpha_n \rightarrow 0, k_n = \mu_n^2/\alpha_n^2$. Тогда

- (1) $k_n \mathbb{E}X_{n1} \rightarrow 0, k_n \mathbb{D}X_{n1} \rightarrow 8\overline{m}\gamma/M^2, n \rightarrow \infty;$
- (2) *Выполняется условие Линдеберга, т.е. для любого $\varepsilon > 0$*

$$\lim_{n \rightarrow \infty} k_n \mathbb{E} [X_{n1}^2 \mathbb{I}(|X_{n1}| > \varepsilon)] = 0.$$

Доказательство. Как было показано выше,

$$X_{n1} = \frac{2m_{n0}(h_1)\gamma_1}{M + m_{n0}(h_1)} s_n,$$

где $s_n \stackrel{d}{=} \xi \wedge \eta_{n0}$ и η_{n0} распределена экспоненциально с параметром μ_n , а величины h_1, ξ, η_{n0} и γ_1 независимы. Так как в соответствии с (1) $\mathbb{E}\gamma_1 = 0$, то $k_n \mathbb{E}X_{n1} = 0$ для любого n . Проверим соотношение для дисперсии:

$$\begin{aligned} k_n \mathbb{D}X_{n1} &= \frac{\mu_n^2}{\alpha_n^2} \mathbb{E} \left[\frac{2m_{n0}(h_1)\gamma_1}{M + m_{n0}(h_1)} s_n \right]^2 = \\ &= 4\mathbb{E}\gamma_1^2 \mathbb{E} \left[\frac{m_{n0}(h_1)}{\alpha_n} \frac{1}{M + m_{n0}(h_1)} \right]^2 \mu_n^2 \mathbb{E}s_n^2 = \\ &= 4\overline{\gamma} \mathbb{E} \left[\frac{m_0(h_1)}{M + m_{n0}(h_1)} \right]^2 \mu_n^2 \mathbb{E}s_n^2. \end{aligned}$$

Последовательность случайных величин $\{[m_0(h_1)/(M + m_{n0}(h_1))]^2\}$ мажорируется интегрируемой случайной величиной $[m_0(h_1)/M]^2$ и поточечно сходится к ней, так как $\alpha_n \rightarrow 0, n \rightarrow \infty$. Поэтому

$$\begin{aligned} \lim_{n \rightarrow \infty} k_n \mathbb{D}X_{n1} &= \lim_{n \rightarrow \infty} 4\overline{\gamma} \mathbb{E} \left[\frac{m_0(h_1)}{M} \right]^2 \mu_n^2 \mathbb{E}s_n^2 = \\ &= \frac{4\overline{m}\gamma}{M^2} \lim_{n \rightarrow \infty} \mu_n^2 \mathbb{E}s_n^2 = \frac{8\overline{m}\gamma}{M^2}. \end{aligned}$$

Докажем справедливость условия Линдеберга. Рассмотрим

$$\begin{aligned} \frac{\mu_n}{\alpha_n} |X_{n1}| &= \frac{\mu_n}{\alpha_n} \left| \frac{2m_{n0}(h_1)\gamma_1}{M + m_{n0}(h_1)} s_n \right| = \\ &= 2 \frac{m_{n0}(h_1)}{\alpha_n} \frac{1}{M + m_{n0}(h_1)} |\gamma_1| \mu_n s_n \leq \\ &\leq 2 \frac{m_0(h_1)}{M} |\gamma_1| \mu_n s_n \leq \\ &\leq \frac{2m_0(h_1)|\gamma_1|\mu_n \eta_{n0}}{M} \stackrel{d}{=} \frac{2m_0(h_1)|\gamma_1|\hat{\eta}}{M}, \end{aligned}$$

где $\hat{\eta}$ распределена экспоненциально с параметром 1. Распределение случайной величины

$$Y_n = \frac{2m_0(h_1)|\gamma_1|\mu_n \eta_{n0}}{M}$$

не зависит от n и согласно (1) имеет конечный второй момент, поэтому

$$\begin{aligned} k_n \mathbb{E} [X_{n1}^2 \mathbb{I}(|X_{n1}| > \varepsilon)] &= \\ &= k_n \mathbb{E} \left[X_{n1}^2 \mathbb{I}(\sqrt{k_n}|X_{n1}| > \sqrt{k_n}\varepsilon) \right] \leq \\ &\leq \mathbb{E} \left[Y_n^2 \mathbb{I}(|Y_n| > \sqrt{k_n}\varepsilon) \right] = \mathbb{E} \left[Y_1^2 \mathbb{I}(|Y_1| > \sqrt{k_n}\varepsilon) \right]. \end{aligned}$$

Последнее математическое ожидание стремится к нулю при $n \rightarrow \infty$ по теореме Лебега о мажорируемой сходимости. Утверждение леммы доказано.

Сформулируем доказанную в работе [5] функциональную центральную предельную теорему, устанавливающую условия, при которых процессы вида (3) сходятся к некоторому предельному процессу X в пространстве Скорохода $\mathcal{D} = (D[0, 1], d_0)$ (см. [6, гл. 3]). Позднее были получены более сильные результаты, касающиеся сходимости обобщенных процессов Кокса (3) (см., например, [7]), но достаточно будет приводимого ниже утверждения.

Теорема [5]. Пусть для некоторой неограниченно возрастающей последовательности чисел $\{k_n\}_{n \geq 1}$ выполнены условия:

(1) существуют числа $a \in \mathbb{R}$ и $\sigma > 0$ такие, что

$$k_n \mathbb{E} X_{n1} \rightarrow a; \quad k_n \mathbb{D} X_{n1} \rightarrow \sigma^2 (n \rightarrow \infty);$$

(2) условие Линдберга, т. е. для любого $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} k_n \mathbb{E} [(X_{n1} - a_n)^2 \mathbb{I}(|X_{n1} - a_n| > \varepsilon)] = 0,$$

где $\mathbb{I}(A)$ — индикатор события A ; $a_n = \mathbb{E} X_{n1}$;

(3) существует безгранично делимая случайная величина U такая, что $\mathbb{P}(U = 0) < 1$, $\mathbb{P}(U \geq 0) = 1$, $\mathbb{E} U^2 < \infty$ и

$$k_n^{-1} \Lambda_n(1) \Rightarrow U, \quad n \rightarrow \infty;$$

(4) $\sup_n k_n^{-2} \mathbb{E} \Lambda_n(1)^2 < \infty$.

Тогда обобщенные процессы Кокса $\{X_n\}$ слабо сходятся в пространстве Скорохода \mathcal{D} к процессу Леви X такому, что

$$X(1) \stackrel{d}{=} \sigma \sqrt{U} N(0, 1) + aU,$$

где $N(0, 1)$ — случайная величина, имеющая стандартное нормальное распределение, при этом не зависящая от U .

Для последовательности $\{k_n = \mu_n^2 / \alpha_n^2\}$ при $a = 0$ первые два условия теоремы выполняются по лемме 2. Таким образом, достаточно наложить определенные условия на интенсивность входящего потока заявок, чтобы была справедлива следующая теорема.

Теорема 1. Пусть $\mu_n \rightarrow \infty$, $\alpha_n \rightarrow 0$, $k_n = \mu_n^2 / \alpha_n^2$, $\sup_n k_n^{-2} \mathbb{E} \Lambda_n(1)^2 < \infty$ и существует безгранично делимая случайная величина U такая, что

$$\mathbb{P}(U = 0) < 1; \quad \mathbb{P}(U \geq 0) = 1; \quad \mathbb{E} U^2 < \infty$$

и

$$k_n^{-1} \Lambda_n(1) \Rightarrow U, \quad n \rightarrow \infty.$$

Тогда обобщенные процессы Кокса $\{X_n\}$ слабо сходятся в пространстве Скорохода \mathcal{D} к процессу Леви X такому, что

$$X(1) \stackrel{d}{=} \sigma \sqrt{U} N(0, 1),$$

где $\sigma = 8\overline{m\gamma} / M^2$, а $N(0, 1)$ — случайная величина со стандартным нормальным распределением, независимая от U .

4 Заключение

В настоящей работе была предложена модель механизма влияния поступающих заказов на цену актива на основе физической модели абсолютно упругого соударения частиц.

Была установлена справедливость функциональной предельной теоремы, на основании результатов которой можно аппроксимировать процесс цены при интенсивном потоке входящих заявок процессом Леви, приращения которого являются смесью нормальных законов и поддаются более точному анализу. Такая аппроксимация дает также возможность оценки риска динамических стратегий [8].

Литература

1. Kukanov A. Stochastic models of limit order markets. — Columbia University, 2013. Ph.D. Thesis. 131 p.
2. Лаврентьев В. В., Назаров Л. В. Процесс движения цены, порожденный непрерывной моделью книги заказов // Вестн. Тверского государственного ун-та. Сер. Прикладная математика, 2015. № 4. С. 55–63.
3. Korolev V. Yu., Chertok A. V., Korchagin A. Yu., Zeifman A. I. Modeling high-frequency order flow imbalance by functional limit theorems for two-sided risk processes // Appl. Math. Comput., 2015. Vol. 253. P. 224–241.
4. Сивухин Д. В. Общий курс физики. — В 5 т. — Т. 1. Механика. — 4-е изд. — М.: МФТИ, 2005. 560 с.
5. Кащеев Д. Е. Моделирование динамики финансовых временных рядов и оценивание производных ценных бумаг: Дис. . . . канд. физ.-мат. наук. — Тверь: ТвГУ, 2001. 191 с.
6. Биллингсли П. Сходимость вероятностных мер / Пер. с англ. — М.: Наука, 1977. 353 с. (Billingsley P. Convergence of probability measures. — New York, NY, USA: John Wiley & Sons, Inc., 1977. 277 p.)
7. Korolev V. Yu., Chertok A. V., Korchagin A. Yu., Kossova E. V., Zeifman A. I. A note on functional limit theorems for compound Cox processes // J. Math. Sci., 2016. Vol. 218. No. 2. P. 182–194.
8. Balasnov Y., Doynikov A., Lavrent'ev V., Nazarov L. Estimating risk of dynamic trading strategies from high frequency data flow // Advances in data mining: Applications and theoretical aspects / Ed. P. Pernert. — Lecture notes in computer science ser. — Springer, 2015. Vol. 9165. P. 153–165.

Поступила в редакцию 07.12.17

A PROBABILITY MODEL OF THE INFLUENCE OF THE ORDER BOOK ON THE PRICE PROCESS

L. V. Nazarov, V. V. Lavrentyev, and E. V. Bykovets

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

Abstract: The Limit Order Book model is considered, with buy and sell orders arriving as two independent Cox processes. It includes the price impact model built on the basis of a physical model of perfectly elastic collision. Price is treated as a particle of some mass, moving along a straight line without friction. The incoming buy orders and outgoing sell orders hit the price giving it additional momentum in one direction, while incoming sell orders and outgoing buy orders do the same in the opposite direction. A functional limit theorem for the price process is obtained at a high intensity of incoming order flow, which allows approximation by some Lévy process

Keywords: limit orders; perfectly elastic collision; limit order book model; price process; Cox process; functional limit theorem

DOI: 10.14357/19922264180205

References

1. Kukanov, A. 2013. Stochastic models of limit order markets. Columbia University. Ph.D. Thesis. 131 p.
2. Lavrent'ev, V.V., and L.V. Nazarov. 2015. Protsess dvizheniya tseny, porozhdenny nepreryvnoy model'yu knigi zakazov [Price process, generated by the continuous model of the order book]. *Vestnik Tverskogo gosudarstvennogo un-ta. Ser. Prikladnaya matematika* [Bull. of the Tverskoy State University. Ser. Appl. Math.] 4:55–63.
3. Korolev, V.Yu., A.V. Chertok, A.Yu. Korchagin, and A. I. Zeifman. 2015. Modeling high-frequency order flow imbalance by functional limit theorems for two-sided risk processes. *Appl. Math. Comput.* 253:224–241.
4. Sivukhin, D. V. 2005. *Obshchiy kurs fiziki. Mekhanika* [General course of physics. Mechanics]. 4 ed. Moscow: MIPT Publ. Vol. 1. 560 p.
5. Kashcheev, D.E. 2001. Modelirovanie dinamiki finansovykh vremennykh ryadov i otsenivanie proizvodnykh tsennykh bumag [Modeling of dynamics of financial time series and estimation of derivative securities]. Tver'. PhD Thesis. 191 p.
6. Billingsley, P. 1977. *Convergence of probability measures*. New York, NY: John Wiley & Sons, Inc. 277 p.
7. Korolev, V.Yu., A.V. Chertok, A.Yu. Korchagin, E. V. Kossova, and A. I. Zeifman. 2016. A note on functional limit theorems for compound Cox processes. *J. Math. Sci.* 218(2):182–194.
8. Balasanov, Y., A. Doynikov, V. Lavrent'ev, and L. Nazarov. 2015. Estimating risk of dynamic trading strategies from high frequency data flow. *Advances in data mining: Applications and theoretical aspects*. Ed. P. Perner. Lecture notes in computer science ser. Springer. 9165:153–165.

Received December 7, 2017

Contributors

Bykovets Eugene V. (b. 1994) — MSc student, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; eugene.bykovets@stud.cs.msu.ru

Lavrentyev Victor V. (b. 1955) — Candidate of Science (PhD) in physics and mathematics, scientist, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; lavrent@cs.msu.ru

Nazarov Leonid V. (b. 1957) — Candidate of Science (PhD) in physics and mathematics, senior scientist, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; nazarov@cs.msu.ru

МАКСИМАЛЬНЫЕ ВЕТВЯЩИЕСЯ ПРОЦЕССЫ В СЛУЧАЙНОЙ СРЕДЕ

А. В. Лебедев¹

Аннотация: Работа продолжает многолетние исследования автора по теории максимальных ветвящихся процессов (МВП), которые получаются из классических ветвящихся процессов путем замены операции суммирования чисел потомков на максимум. Можно сказать, что в каждом поколении выживают потомки только одной частицы, у которой их больше всего. Ранее автором было проведено обобщение процессов с целочисленными значениями до процессов с произвольными неотрицательными значениями, исследованы их свойства и доказаны предельные теоремы. Далее были введены и изучались процессы с несколькими типами частиц. В настоящей работе вводится понятие МВП в случайной среде (МВПСС) (с одним типом частиц) и важного случая «степенной» случайной среды (МВПССС). В последнем случае изучены свойства МВП и доказана эргодическая теорема. В качестве приложений рассмотрены вентильные бесконечнолинейные системы массового обслуживания.

Ключевые слова: максимальные ветвящиеся процессы; случайная среда; эргодическая теорема; устойчивые распределения; теория экстремумов

DOI: 10.14357/19922264180206

1 Введение

Классическими объектами исследования в теории случайных процессов являются ветвящиеся процессы Гальтона–Ватсона (с одним типом частиц и дискретным временем) [1]. Их экстремальные аналоги называются максимальными ветвящимися процессами, а именно: суммирование чисел потомков частиц (при определении численности очередного поколения) заменяется на максимум.

Напомним историю вопроса. Максимальные ветвящиеся процессы были введены и изучались Дж. Ламперти [2, 3] в 1970–1972 гг., однако в дальнейшем были совершенно всеми заброшены (хотя и упомянуты в обзоре [4]). Новый этап исследований МВП был начат А. В. Лебедевым с 2001 г. Было проведено обобщение процессов с целочисленных значений на произвольные неотрицательные [5]. Изучались МВП сначала с одним типом частиц (см. обзор [6]), а затем с несколькими типами (многотипные) [7]. Наиболее полно на данный момент результаты и библиография представлены в диссертации [8, гл. 4 и 5]. До недавнего времени исследования автора по этой тематике оставались одиночными. Лишь в 2012 г. они были неожиданно подхвачены в работе зарубежных ученых О. Айдогмуса, А. П. Гхоша, С. Гхоша и А. Ройтерштейна [9], где введены раскрашенные МВП. Их отличие от многотипных МВП заключается в том, что типы (цвета) частиц определяются уже после формиро-

вания поколения, случайным образом, причем тип влияет на дальнейшую плодовитость. Другим отличием от подхода автора стало рассмотрение только процессов, уходящих в бесконечность.

Напомним основные понятия и факты об МВП с одним типом частиц.

Рассмотрим случайные процессы со значениями в \mathbb{Z}_+ , заданные стохастически рекуррентными формулами вида:

$$Z_{n+1} = \bigvee_{m=1}^{Z_n} \xi_{m,n}, \quad (1)$$

где через \bigvee обозначена операция взятия максимума; $\xi_{m,n}$, $m \geq 1$, $n \geq 0$, — независимые случайные величины с общим распределением F на \mathbb{Z}_+ . Полагаем (как и в случае суммирования), что результат взятия максимума «ноль раз» (при $Z_n = 0$) равен нулю.

Можно сказать (по аналогии с процессами Гальтона–Ватсона), что в МВП в каждом поколении выживают потомки только одной частицы, у которой их больше всего. Понятно также, что множество возможных значений МВП (при $n \geq 1$) совпадает с множеством возможных значений числа потомков. Из (1) следует, что процесс является однородной цепью Маркова на этом множестве.

Другую интерпретацию МВП можно предложить в теории массового обслуживания, рассмотрев вентильные бесконечнолинейные системы. Так на-

¹Московский государственный университет им. М. В. Ломоносова, кафедра теории вероятностей механико-математического факультета, avlebed@yandex.ru

зывают системы с бесконечным числом приборов, в которых доступ заявок к обслуживанию регулируется вентилем. Предполагается, что вентиль открыт только в том случае, когда все приборы свободны. Заявки поступают в очередь с бесконечным числом мест ожидания, а обслуживание происходит по стадиям. В начале стадии, когда вентиль открывается, все заявки из очереди мгновенно получают доступ к приборам и далее обслуживаются параллельно и независимо до полного освобождения всех приборов. В момент освобождения всех приборов вентиль вновь открывается для новой партии заявок (пришедших за это время) и следующей стадии.

Заметим, что система очень проста в управлении: нет необходимости постоянного учета входящих и уходящих заявок, свободных и занятых приборов и т. п. Распределение заявок по приборам (которые на тот момент все свободны) производится однократно и одновременно в начале каждой стадии. Другое преимущество вентильной системы может проявиться в ситуации, когда заявки в очереди и обслуживающие приборы каким-то образом разделены между собой, а установление связи сопряжено с затратами. Например, может оказаться невыгодно (или невозможно) держать постоянно включенным канал передачи данных, а предпочтительней недолгие подключения время от времени.

Разумеется, любая бесконечнолинейная система является лишь приближением для случая, когда реальное число приборов велико. Вместе с тем, имеет смысл изучение таких систем для оценки различных характеристик качества обслуживания (которые для любой системы с конечным числом приборов могут быть только хуже).

Рассмотрим подобную систему с дискретным временем, и пусть в каждый момент времени поступает ровно по одной заявке. Тогда в силу параллельности работы приборов время обслуживания очередной партии заявок (а значит, и число заявок в следующей) равно максимуму из времен их обслуживания. Таким образом, обозначая через Z_n длительность n -й стадии, а через $\xi_{m,n}$ — времена обслуживания заявок на ней, получаем в точности (1).

Вентильные бесконечнолинейные системы с непрерывным временем и пуассоновским входным потоком изучались в [10–13] (другими методами) и автором в [14, 15]. В этом случае необходимо оговорить, что происходит, если вентиль открывается при пустой очереди. Наиболее естественно предположить, что система ждет поступления новой заявки, с которой и начинается следующая стадия.

Отметим, что модели с параллельной обработкой данных в последние годы стали весьма популярны в связи с развитием технологии облачных

вычислений. При этом возникает необходимость изучать максимумы случайных величин. В качестве недавней работы на эту тему отметим [16].

Максимальные ветвящиеся процессы были введены в [2] (в связи с моделями дальнедействующей перколяции), и там же были получены критерии их возвратности, а именно (в предположении $F(0) = 0$): при выполнении условия

$$\limsup_{x \rightarrow +\infty} x(1 - F(x)) < e^{-\gamma}, \quad (2)$$

где $\gamma = 0,577\dots$ — константа Эйлера, цепь $\{Z_n\}$ положительно возвратна и, напротив, при

$$\liminf_{x \rightarrow +\infty} x(1 - F(x)) > e^{-\gamma}$$

имеет место $Z_n \rightarrow +\infty$, $n \rightarrow \infty$, почти наверное (п. н.).

Далее в [3] рассмотрен критический случай

$$x(1 - F(x)) \rightarrow e^{-\gamma}, \quad x \rightarrow +\infty,$$

с учетом дальнейших членов разложения хвоста на бесконечности. Показано, что если

$$(e^\gamma x(1 - F(x)) - 1) \ln x \rightarrow d, \quad x \rightarrow +\infty,$$

то процесс возвратен при $d < \pi^2/12$ и уходит в бесконечность п. н. при $d > \pi^2/12$.

Согласно (1) и предположению о случае $Z_n = 0$ процесс имеет переходные вероятности

$$\mathbf{P}(Z_{n+1} \leq j | Z_n = i) = F^i(j), \quad i, j \in \mathbb{Z}_+$$

(где полагаем $0^0 = 1$), что подсказало рассмотреть в [5] цепи Маркова на произвольном измеримом множестве $T \subset \mathbb{R}_+$ с переходными вероятностями

$$\mathbf{P}(Z_{n+1} \leq y | Z_n = x) = F^x(y), \quad x, y \in T, \quad (3)$$

где F также сосредоточено на T .

Подобные процессы могут рассматриваться как самостоятельно, так и в качестве предельных (в каком-либо смысле) для МВП на \mathbf{Z}_+ (нормированных определенным образом). Например, они могут пригодиться для описания поведения вентильных бесконечнолинейных систем, в том числе предельного при большой загрузке и т. п.

В частности, для вентильной бесконечнолинейной системы с непрерывным временем и пуассоновским входным потоком последовательность длительностей стадий на протяжении периода занятости удовлетворяет (3) с $F(x) = \exp\{-\lambda \bar{B}(x)\}$, $x \geq 0$, где λ — интенсивность входного потока; $B(x)$ — функция распределения времени обслуживания одной заявки, $\bar{B}(x) = 1 - B(x)$. Действительно, обозначим длительность n -й стадии через Z_n .

При условии $Z_n = x$ число заявок, поступивших на данной стадии, пуассоновское с параметром λx , а Z_{n+1} есть максимум этого (случайного) числа независимых случайных величин с распределением B . Таким образом,

$$\begin{aligned} \mathbf{P}(Z_{n+1} \leq y | Z_n = x) &= \sum_{k=0}^{\infty} \frac{(\lambda x)^k}{k!} e^{-\lambda x} B(y)^k = \\ &= \exp \{-\lambda x \bar{B}(y)\} = F(y)^x. \end{aligned}$$

Для описания системы на всем протяжении времени требуются уже МВП с иммиграцией в момент обнуления и т. п. [8, § 4.3]

Далее будем говорить об МВП на T и обозначать их $\text{МВП}(T)$. Заметим, что аналогичным обобщением для процессов Гальтона–Ватсона являются процессы Иржины [17] (с непрерывным множеством состояний и дискретным временем), однако в данном случае T может быть любым измеримым подмножеством \mathbb{R}_+ .

Формула (1) для таких МВП в общем случае уже не имеет места, но в соответствии с (3) они допускают эквивалентное представление стохастической рекуррентной последовательностью вида

$$Z_{n+1} = \begin{cases} F^{-1}(U_{n+1}^{1/Z_n}), & Z_n > 0; \\ 0, & Z_n = 0, \end{cases} \quad n \geq 0,$$

где $F^{-1}(y) = \inf\{x : F(x) \geq y\}$ и $U_n, n \geq 1$, — независимые равномерно распределенные на $(0, 1)$ случайные величины, а $Z_0 \geq 0$ и не зависит от них. При этом распределение F по-прежнему называется распределением числа (непосредственных) потомков.

В [5] был получен ряд свойств МВП (свойство преобразования подобия, ассоциированность, монотонность по параметрам, условие вырождения) и доказана теорема эргодичности.

Заметим, что для МВП нуль всегда является поглощающим состоянием. Так, для $\text{МВП}(\mathbb{Z}_+)$ при условии (2) и $F(0) > 0$ это приводит к вырождению п. н. [2]. Для $\text{МВП}(T)$, если $F(0) = 0$, нуль можно просто исключить из множества состояний, рассматривая процесс с ненулевым начальным условием. Однако, если нуль является предельной точкой T , остается возможность асимптотической сходимости к нему при $n \rightarrow \infty$. Следующая теорема дает достаточные условия для того, чтобы исключить уход процесса как в нуль, так и в бесконечность, и сделать его эргодическим. Здесь и далее имеется в виду эргодичность по Харрису [18, гл. 1].

Теорема А [5, теорема 1]. *Если для $\text{МВП}(T)$, $T \subset (0, +\infty)$, выполнено (2) и*

$$\liminf_{x \rightarrow 0} x(-\ln F(x)) > e^{-\gamma},$$

то процесс эргодический.

Определим теперь МВПСС.

В приложениях «случайная среда» может описывать различные факторы природного, технического или общественного характера, меняющиеся со временем случайным образом и оказывающие влияние на систему, например ускоряющие или замедляющие работу системы массового обслуживания.

Пусть заданы последовательность $F_l, l \geq 1$, распределений на \mathbb{Z}_+ , набор независимых случайных величин $\xi_{m,n,l}, m \geq 1, n \geq 0, l \geq 1$, с распределениями $F_l, l \geq 1$, и последовательность независимых случайных величин $\nu_n, n \geq 0$, с общим распределением G на \mathbb{N} , не зависящих от $\xi_{m,n,l}, m \geq 1, n \geq 0, l \geq 1$. Тогда МВПСС(\mathbb{Z}_+) можно определить с помощью стохастически рекуррентной формулы вида

$$Z_{n+1} = \bigvee_{m=1}^{Z_n} \xi_{m,n,\nu_n}. \quad (4)$$

Случайную среду здесь отражают случайные величины $\nu_n, n \geq 1$, от которых зависит выбор распределения $F_l, l \geq 1$, числа потомков частиц на каждом шаге.

Согласно (4) и предположению о случае $Z_n = 0$ процесс имеет переходные вероятности

$$\mathbf{P}(Z_{n+1} \leq j | Z_n = i) = \mathbf{E}(F_\nu^i(j)), \quad i, j \in \mathbb{Z}_+,$$

где ν имеет распределение G , что подсказывает рассмотреть процесс на произвольном измеримом множестве $T \subset \mathbb{R}_+$ с переходными вероятностями

$$\mathbf{P}(Z_{n+1} \leq y | Z_n = x) = \mathbf{E}(F_\nu^x(y)), \quad x, y \in T, \quad (5)$$

где имеется в виду семейство распределений $F_s, s > 0$, на T и случайная величина ν с распределением G на $(0, +\infty)$. Таким образом, определяем МВПСС(T).

Подобные процессы могут рассматриваться как самостоятельно, так и в качестве предельных (в каком-либо смысле) для МВПСС(\mathbb{Z}_+), нормированных определенным образом.

Далее будем изучать класс процессов со степенным семейством $F_s(x) = F^s(x), s > 0$, где F — распределение на T . В таком случае будем говорить о «степенной» случайной среде, а соответствующие процессы обозначим МВПССС(T).

Введем преобразование Лапласа–Стилтьеса $\varphi(u) = \mathbf{E}e^{-uw}$. В этом случае формула (5) преобразуется к виду:

$$\mathbf{P}(Z_{n+1} \leq y | Z_n = x) = \varphi(-x \ln F(y)), \quad x, y \in T. \quad (6)$$

Поставим задачу изучения свойств таких процессов.

2 Основные свойства

Прежде всего отметим, что для МВПССС также имеет место конструктивное представление, следующее из (6), а именно:

$$Z_{n+1} = \begin{cases} F^{-1} \left(\exp \left\{ -\frac{\varphi^{-1}(U_{n+1})}{Z_n} \right\} \right), & Z_n > 0; \\ 0, & Z_n = 0, \end{cases} \quad n \geq 0, \quad (7)$$

где $U_n, n \geq 1$, — независимые равномерно распределенные на $(0, 1)$ случайные величины, $Z_0 \geq 0$ и не зависит от них.

Докажем ряд свойств МВПССС по аналогии с [5] в форме утверждений.

Утверждение 1. Если $\{Z_n\}$ является МВПССС(T) с $F(x)$, то $\{\lambda Z_n\}$ при любом $\lambda > 0$ является МВПССС(λT) с $F^{1/\lambda}(x/\lambda)$.

Доказательство. Используем подстановку в (6). Действительно,

$$\begin{aligned} \mathbf{P}(\lambda Z_{n+1} \leq y | \lambda Z_n = x) &= \\ &= \mathbf{P}\left(Z_{n+1} \leq \frac{y}{\lambda} | Z_n = \frac{x}{\lambda}\right) = \\ &= \varphi\left(-\left(\frac{x}{\lambda}\right) \ln F\left(\frac{y}{\lambda}\right)\right) = \varphi\left(-x \ln F^{1/\lambda}\left(\frac{y}{\lambda}\right)\right). \quad \square \end{aligned}$$

Из данного свойства следует замкнутость класса МВПССС(\mathbb{R}_+) относительно умножения на $\lambda > 0$ и замкнутость МВПССС(Z_+) при $\lambda \in \mathbb{N}$.

Лемма 1. Для любых чисел $Z'_0 \leq Z''_0$ и $U'_n \leq U''_n, n \geq 1$, где $Z'_0, Z''_0 \geq 0$ и $U'_n, U''_n \in (0, 1)$, числовые последовательности $\{Z'_n\}$ и $\{Z''_n\}$, построенные по формуле (7), удовлетворяют условию $Z'_n \leq Z''_n$ для всех $n \geq 0$.

Доказательство. По условию $Z'_0 \leq Z''_0$. Пусть верно $Z'_n \leq Z''_n$ для некоторого $n \geq 0$. Тогда, поскольку F^{-1} — неубывающая функция, из $U'_{n+1} \leq U''_{n+1}$ и формулы (7) получаем $Z'_{n+1} \leq Z''_{n+1}$. Утверждение леммы верно по принципу математической индукции. \square

Напомним понятие ассоциированности случайных величин [19, 20].

Функцию многих переменных $f(x)$, где $x = (x_1, \dots, x_n)$, назовем монотонно неубывающей, если из $x'_i \leq x''_i, 1 \leq i \leq n$, следует $f(x') \leq f(x'')$. Случайные величины набора $\zeta = (\zeta_1, \dots, \zeta_n)$ называются ассоциированными, если $\text{cov}(f(\zeta), g(\zeta)) \geq 0$ для всех тех монотонно неубывающих f и g , для которых эта ковариация существует. Говорят, что случайный процесс или поле $\{\zeta(t) : t \in T\}$ ассоциированы, если ассоциированы их значения $\zeta(t_1), \dots, \zeta(t_n)$ для любого конечного множества $\{t_1, \dots, t_n\} \subset T$.

Согласно [19] и [20, теорема 1.8], независимые случайные величины ассоциированы; монотонно неубывающие функции от ассоциированных случайных величин также обладают этим свойством.

Утверждение 2. Любой МВПССС ассоциирован.

Доказательство. Достаточно заметить, что по лемме 1 любые Z_{i_1}, \dots, Z_{i_m} представляют собой неубывающие функции от независимых случайных величин Z_0 и $U_n, 1 \leq n \leq \max\{i_1, \dots, i_m\}$. \square

Для установления монотонности по параметрам введем отношение (частичного) порядка между распределениями: $F_1 \prec F_2$, если $F_1(x) \geq F_2(x), \forall x$. Заметим, что из $F_1 \prec F_2$ следует $F_1^{-1}(y) \leq F_2^{-1}(y), \forall y \in (0, 1)$.

Обозначим через $Z = Z(F, G, H)$ МВПССС, где Z_0 имеет распределение H .

Утверждение 3. Если $F' \prec F'', G' \prec G''$ и $H' \prec H''$, то можно построить процессы $Z' = Z(F', G', H')$ и $Z'' = Z(F'', G'', H'')$ на одном вероятностном пространстве так, что $Z'_n \leq Z''_n$ для всех $n \geq 0$ п. н.

Доказательство. Пусть U_0 — равномерно распределенная на $(0, 1)$ случайная величина, не зависящая от $U_n, n \geq 1$. Полагая $Z'_0 = (H')^{-1}(U_0)$ и $Z''_0 = (H'')^{-1}(U_0)$, получаем $Z'_0 \leq Z''_0$. Пусть верно $Z'_n \leq Z''_n$ для некоторого $n \geq 0$. Из $G' \prec G''$ следует $\varphi'(u) \geq \varphi''(u), \forall u > 0$, и $(\varphi'(v))^{-1} \geq (\varphi''(v))^{-1}, \forall v \in (0, 1)$. Из $F' \prec F''$ следует $(F')^{-1}(y) \leq (F'')^{-1}(y), \forall y \in (0, 1)$. По формуле (7) получаем $Z'_{n+1} \leq Z''_{n+1}$. Утверждение 3 доказано по принципу математической индукции. \square

Обозначим предельное распределение Z_n при $n \rightarrow \infty$ через Ψ (если оно существует).

Утверждение 4. Если для двух МВПССС(T) верно $F' \prec F'', G' \prec G''$, то $\Psi' \prec \Psi''$.

Доказательство. Возьмем произвольные $H' = H''$ на T и построим процессы на одном вероятностном пространстве согласно утверждению 3. Тогда из $Z'_n \leq Z''_n$ п. н. следует $\mathbf{P}(Z'_n \leq x) \geq \mathbf{P}(Z''_n \leq x), n \geq 1$, откуда при $n \rightarrow \infty$ получаем $\Psi'(x) \geq \Psi''(x), x > 0$. \square

3 Эргодическая теорема

Напомним функцию распределения Гумбеля $\Lambda(x) = \exp\{-e^{-x}\}$, играющую важную роль в стохастической теории экстремумов.

Если функция распределения $F(x)$ непрерывна и строго возрастает, а $F(0) = 0$, то (7) преобразованием $\zeta_n = \Lambda^{-1}(F(Z_n))$ приводится к форме общей (нелинейной) авторегрессии первого порядка

$$\zeta_{n+1} = f(\zeta_n) + \eta_{n+1}, \quad n \geq 0, \quad (8)$$

где $f(u) = \ln F^{-1}(\Lambda(u))$ и независимые случайные величины $\eta_n = -\ln \varphi^{-1}(U_n)$, $n \geq 1$, имеют функцию распределения $\varphi(e^{-x})$.

Эргодичность подобных моделей изучалась, например, в [18, § 8.4] и [21].

Обозначим $\delta = \mathbf{E}\eta_1$ и пусть это среднее существует. Имеем

$$\varphi(e^{-x}) = \mathbf{E} \exp \{-\nu e^{-x}\} = \mathbf{E}\Lambda(x - \ln \nu),$$

откуда следует

$$\delta = \gamma + \mathbf{E} \ln \nu, \tag{9}$$

поскольку распределение Гумбеля Λ имеет математическое ожидание γ .

Заметим, что δ как математическое ожидание при функции распределения $\varphi(e^{-x})$ может быть также представлено формулой

$$\begin{aligned} \delta &= \int_0^{+\infty} (1 - \varphi(e^{-x})) dx - \int_{-\infty}^0 \varphi(e^{-x}) dx = \\ &= \int_0^{\infty} (1 - \varphi(e^{-x})) dx - \int_0^{\infty} \varphi(e^x) dx, \end{aligned} \tag{10}$$

когда все интегралы сходятся.

Пример 1. Пусть $F(x) = \exp\{-(x/c)^{-\beta}\}$, $x, c, \beta > 0$ (распределение Фреше), тогда МВПССС допускает конструктивное представление:

$$Z_{n+1} = W_{n+1} Z_n^{1/\beta}, \tag{11}$$

где W_n , $n \geq 1$, независимы и имеют функцию распределения $\varphi((x/c)^{-\beta})$, а (8) переписывается в форме линейной авторегрессии:

$$\zeta_{n+1} = \frac{\zeta_n}{\beta} + \ln c + \eta_{n+1}, \quad n \geq 0. \tag{12}$$

При $\beta < 1$ процесс $\{\zeta_n\}$ уходит в $\pm\infty$ в зависимости от знака начального условия ζ_0 . При $\beta > 1$ процесс $\{\zeta_n\}$ эргодический. При $\beta = 1$ имеем простое случайное блуждание, уходящее в $+\infty$ при $c > e^{-\delta}$, в $-\infty$ при $c < e^{-\delta}$ и осциллирующее между $\pm\infty$ при $c = e^{-\delta}$.

Отсюда легко получить результаты для $\{Z_n\}$, если учесть, что $Z_n = F^{-1}(\Lambda(\zeta_n))$, откуда $Z_n \rightarrow 0$ при $\zeta_n \rightarrow -\infty$ и $Z_n \rightarrow +\infty$ при $\zeta_n \rightarrow +\infty$, а эргодичность сохраняется.

Данный пример наводит на мысль, что верна следующая теорема.

Теорема 1. Если для МВПССС(T), $T \subset (0, +\infty)$, выполнены условия

$$\liminf_{x \rightarrow +\infty} x(-\ln F(x)) < e^{-\delta}; \tag{13}$$

$$\liminf_{x \rightarrow 0} x(-\ln F(x)) > e^{-\delta}, \tag{14}$$

где δ из (9) существует, то процесс эргодический.

Доказательство. Без ограничения общности можно считать $T = (0, +\infty)$. Прежде всего заметим, что условия (13) и (14) эквивалентны утверждению о существовании таких $0 < x_1 < x_2$ и $0 < c_2 < e^{-\delta} < c_1$, что $F(x) \leq e^{-c_1/x}$ при $x \leq x_1$ и $F(x) \geq e^{-c_2/x}$ при $x \geq x_2$.

В качестве функции Ляпунова рассмотрим

$$g(x) = \left(\ln \left(\frac{x}{x_2} \right) \right)_+ + \left(\ln \left(\frac{x_1}{x} \right) \right)_+,$$

где $y_+ = \max\{0, y\}$. Обозначим

$$\mu(x) = \mathbf{E}(g(Z_{n+1}) | Z_n = x) - g(x),$$

тогда

$$\begin{aligned} \mu(x) &= \mathbf{E} \left\{ \ln \left(\frac{Z_{n+1}}{x_2} \right)_+ | Z_n = x \right\} + \\ &+ \mathbf{E} \left\{ \ln \left(\frac{x_1}{Z_{n+1}} \right)_+ | Z_n = x \right\} - g(x) = \\ &= \int_0^{\infty} \mathbf{P} \left(\ln \left(\frac{Z_{n+1}}{x_2} \right) > y | Z_n = x \right) dy + \\ &+ \int_0^{\infty} \mathbf{P} \left(\ln \left(\frac{x_1}{Z_{n+1}} \right) > y | Z_n = x \right) dy - g(x) = \\ &= \int_0^{\infty} (1 - \varphi(-x \ln F(x_2 e^y))) dy + \\ &+ \int_0^{\infty} \varphi(-x \ln F(x_1 e^{-y})) dy - g(x) \leq \\ &\leq \int_0^{\infty} \left(1 - \varphi \left(\left(\frac{c_2 x}{x_2} \right) e^{-y} \right) \right) dy + \\ &+ \int_0^{\infty} \varphi \left(\left(\frac{c_1 x}{x_1} \right) e^y \right) dy - g(x) = \\ &= \int_{-\ln(c_2 x/x_2)}^{\infty} (1 - \varphi(e^{-z})) dz + \int_{\ln(c_1 x/x_1)}^{\infty} \varphi(e^z) dz - g(x) = \\ &= \int_0^{\infty} (1 - \varphi(e^{-z})) dz + \int_0^{\infty} \varphi(e^z) dz + \\ &+ \int_{-\ln(c_2 x/x_2)}^0 (1 - \varphi(e^{-z})) dz - \int_0^{\ln(c_1 x/x_1)} \varphi(e^z) dz - \\ &- \left(\left(\ln \left(\frac{x}{x_2} \right) \right)_+ + \left(\ln \left(\frac{x_1}{x} \right) \right)_+ \right). \end{aligned}$$

Обозначим полученную оценку сверху для $\mu(x)$ через $\mu^*(x)$ и пусть

$$\tilde{\mu}(x) = \int_{-\ln(c_2x/x_2)}^0 (1 - \varphi(e^{-z})) dz - \int_0^{\ln(c_1x/x_1)} \varphi(e^z) dz - \left(\left(\ln\left(\frac{x}{x_2}\right) \right)_+ + \left(\ln\left(\frac{x_1}{x}\right) \right)_+ \right),$$

тогда

$$\mu(x) \leq \mu^*(x) = \int_0^\infty (1 - \varphi(e^{-z})) dz + \int_0^\infty \varphi(e^{-z}) dz + \tilde{\mu}(x). \quad (15)$$

При $x \geq x_2$ имеем:

$$\begin{aligned} \tilde{\mu}(x) &= \int_{-\ln(c_2x/x_2)}^0 (1 - \varphi(e^{-z})) dz - \int_0^{\ln(c_1x/x_1)} \varphi(e^z) dz - \\ &\quad - \ln\left(\frac{c_2x}{x_2}\right) + \ln c_2 = \\ &= - \int_{-\ln(c_2x/x_2)}^0 \varphi(e^{-z}) dz - \int_0^{\ln(c_1x/x_1)} \varphi(e^z) dz + \ln c_2 \rightarrow \\ &\quad \rightarrow -2 \int_0^\infty \varphi(e^z) dz + \ln c_2, \quad x \rightarrow +\infty, \end{aligned}$$

откуда в силу (10) и (15) верно

$$\mu(x) \leq \mu^*(x) \rightarrow \delta + \ln c_2 = \ln\left(\frac{c_2}{e^{-\delta}}\right) < 0, \quad x \rightarrow +\infty.$$

При $x \leq x_1$ имеем:

$$\begin{aligned} \tilde{\mu}(x) &= \int_{-\ln(c_2x/x_2)}^0 (1 - \varphi(e^{-z})) dz - \int_0^{\ln(c_1x/x_1)} \varphi(e^z) dz + \\ &\quad + \ln\left(\frac{c_1x}{x_1}\right) - \ln c_1 = - \int_{-\ln(c_2x/x_2)}^0 \varphi(e^{-z}) dz - \\ &\quad - \int_0^{\ln(c_1x/x_1)} \varphi(e^{-z}) dz - \ln c_1 \rightarrow \\ &\quad \rightarrow -2 \int_0^\infty (1 - \varphi(e^{-z})) dz - \ln c_2, \quad x \rightarrow 0, \end{aligned}$$

откуда в силу (10) и (15) верно

$$\mu(x) \leq \mu^*(x) \rightarrow -\delta - \ln c_1 = -\ln\left(\frac{c_1}{e^{-\delta}}\right) < 0, \quad x \rightarrow 0.$$

Следовательно, существуют такие $\varepsilon > 0$ и $0 < v_1 \leq v_2$, что $\mu(x) \leq -\varepsilon$ при $x \notin V = [v_1, v_2]$. Кроме того,

$$\sup_{x \in V} \mathbf{E}(g(Z_{n+1})|Z_n = x) < \infty.$$

Таким образом, условия Ляпунова [18, § 4.2] выполнены.

Проверим теперь условие перемешивания.

Пусть $0 < u_1 \leq u \leq u_2$, $0 < a < b$. В силу выпуклости и убывания $\varphi(s)$ верно

$$\begin{aligned} \varphi(-u \ln F(b)) - \varphi(u \ln F(a)) &\geq \\ &\geq \frac{u_1}{u_2} (\varphi(-u_2 \ln F(b)) - \varphi(-u_2 \ln F(a))), \end{aligned}$$

откуда следует, что для любого измеримого B выполняется

$$\mathbf{P}(Z_{n+1} \in B|Z_n = x) \geq \frac{v_1}{v_2} \mathbf{P}(Z_{n+1} \in B|Z_n = v_2), \quad \forall x \in V. \quad (16)$$

Кроме того, любой МВПССС неприводим и аperiodичен (иначе говоря, из любого состояния $x \in T$ можно попасть в любое множество $B \subset T$, причем за один шаг). Из условий Ляпунова и (16) по [18, § 2, теорема 2] следует эргодичность МВПССС. \square

Отметим, что при $\nu = 1$ п. н. теорема 1 сводится к теореме А, поскольку тогда $\delta = \gamma$.

В некоторых случаях δ удобнее вычислять по определению, чем по формуле (9).

Пример 2. Пусть ν имеет показательное распределение со средним θ , тогда $\varphi(u) = (1 + \theta u)^{-1}$ и

$$\varphi(e^{-x}) = \frac{1}{1 + \theta e^{-x}} = \frac{1}{1 + e^{-(x - \ln \theta)}},$$

т. е. получаем логистическое распределение с параметром сдвига $\ln \theta$. Следовательно, $\delta = \ln \theta$.

Пример 3. Пусть ν имеет строго устойчивое распределение с $\varphi(u) = e^{-cu^\alpha}$, $c > 0$, $0 < \alpha < 1$, тогда

$$\varphi(e^{-x}) = \exp\{-ce^{-\alpha x}\} = \Lambda(\alpha x - \ln c),$$

откуда

$$\delta = \frac{\gamma + \ln c}{\alpha}.$$

Это, кстати, дает удобный способ вычисления среднего логарифма от устойчивой случайной величины. С помощью (9) получаем:

$$\mathbf{E} \ln \nu = \frac{\gamma(1 - \alpha) + \ln c}{\alpha}.$$

Для эргодического МВПССС(T) обозначим случайную величину с предельным распределением через \tilde{Z} . В некоторых случаях удается найти числовые характеристики этого распределения.

Пример 4. Пусть ν имеет строго устойчивое распределение с $\varphi(u) = e^{-u^\alpha}$, $0 < \alpha < 1$, и $F(x) = \exp\{-x^{-\beta}\}$, $x, \beta > 0$, тогда получаем представление (11), где $W_n, n \geq 1$, имеют функцию распределения Фреше $\exp\{-x^{-\alpha\beta}\}$, $x > 0$, и

$$\mathbf{E}W_1^s = \Gamma\left(1 - \frac{s}{\alpha\beta}\right), \quad 0 < s < \alpha\beta.$$

Из представления (11) и эргодичности процесса следует, что предельным является распределение следующего бесконечного произведения, которое сходится п. н.:

$$\tilde{Z} \stackrel{d}{=} \prod_{n=0}^{\infty} W_n^{1/\beta^n},$$

откуда

$$\mathbf{E}\tilde{Z}^s = \prod_{n=1}^{\infty} \Gamma\left(1 - \frac{s}{\alpha\beta^n}\right), \quad 0 < s < \alpha\beta.$$

Очевидно, утверждение 4 верно и в тех случаях, когда одно или оба предельных распределения сосредоточены в нуле. Отсюда, в частности, можно получить условие вырождения для процессов с $F(0) > 0$. Здесь под вырождением понимается обращение процесса в нуль начиная с некоторого (случайного) момента.

Следствие 1. Если для МВПССС(\mathbb{R}_+) выполнено (13) и $F(0) > 0$, то процесс вырождается п. н. **Доказательство.** Заметим, что для любых $C > 0$ и $n \geq 0$ по формуле (6) верно

$$\begin{aligned} \mathbf{P}(Z_{n+1} = 0) &= \mathbf{E}\varphi(-Z_n \ln F(0)) \geq \\ &\geq \mathbf{P}(Z_n = 0) + \varphi(-C \ln F(0))\mathbf{P}(0 < Z_n \leq C). \end{aligned}$$

Последовательность $\mathbf{P}(Z_n = 0)$ монотонно неубывает и ограничена, а значит, стремится к некоторому пределу $p_0 \in (0, 1]$. Получаем:

$$\sum_{n=0}^{\infty} \mathbf{P}(0 < Z_n \leq C) \leq \frac{p_0}{\varphi(-C \ln F(0))} < \infty,$$

так что по лемме Бореля–Кантелли Z_n попадает в $(0, C]$ конечное число раз п. н. при любом $C > 0$,

что может означать либо вырождение, либо уход в бесконечность при $n \rightarrow \infty$.

Пусть $F^*(x) = F(x) \exp\{-x^{-2}\}\mathbf{I}(x > 0)$, тогда $F \prec F^*$. Согласно утверждению 3 можно построить МВПССС(\mathbb{R}_+) с F^* такой, что $Z_n \leq Z_n^*$ п. н. Заметим, что F^* удовлетворяет условиям теоремы 1, так что $\mathbf{P}(Z_n^* \rightarrow +\infty) = 0$, а следовательно, и $\mathbf{P}(Z_n \rightarrow +\infty) = 0$. Таким образом, происходит вырождение п. н. \square

До сих пор предполагалось, что δ конечно. Однако предельное распределение может существовать в некоторых случаях и при $\delta = +\infty$, что соответствует сверхтяжелым хвостам G .

Пример 5. Пусть $F(x) = \exp\{-x^{-\beta}\}$, $x > 0$, $\beta > 1$, и $G(x) = 1 - 1/\ln x$, $x \geq e$, тогда по тауберовой теореме $1 - \varphi(u) \sim -1/\ln u$, $u \rightarrow 0$, откуда

$$\mathbf{P}(\eta_1 > x) = 1 - \varphi(e^{-x}) \sim \frac{1}{x}, \quad x \rightarrow +\infty. \quad (17)$$

Вместе с тем, поскольку $\nu \geq e$ п. н., то $\varphi(u) \leq e^{-eu}$, $u \geq 0$, откуда

$$\mathbf{P}(\eta_1 < -x) = \varphi(e^x) \leq \exp\{-e^{x+1}\}, \quad x > 0. \quad (18)$$

Из (12) следует, что существование предельного распределения эквивалентно сходимости следующего случайного ряда п. н.:

$$\tilde{\zeta} \stackrel{d}{=} \sum_{n=0}^{\infty} \beta^{-n} \eta_n, \quad (19)$$

где $\tilde{\zeta} = \Lambda^{-1}(F(\tilde{Z}))$. Для любого $1/\beta < \varepsilon < 1$ в силу (17) и (18) получаем:

$$\sum_{n=0}^{\infty} \mathbf{P}(|\eta_n| > (\varepsilon\beta)^n) < \infty,$$

откуда по лемме Бореля–Кантелли следует, что события $A_n = \{\beta^{-n}|\eta_n| > \varepsilon^n\}$ происходят не более чем конечное число раз; следовательно, ряд (19) сходится п. н.

4 Заключение

В работе введены МВПСС (с одним типом частиц). Разобран случай «степенной» случайной среды, для него изучен ряд свойств, доказана эргодическая теорема, приведены примеры. Отмечена связь МВП с бесконечнолинейными системами массового обслуживания.

Дальнейшие исследования могут быть связаны с изучением более широкого класса случайных сред и процессов с иммиграцией.

Литература

1. *Harris T.* Теория ветвящихся случайных процессов // Пер. с англ. — М.: Мир, 1966. 356 с. (*Harris T.* The theory of branching processes. — Berlin: Springer-Verlag, 1963. 230 p.)
2. *Lamperti J.* Maximal branching processes and long-range percolation // *J. Appl. Probab.*, 1970. Vol. 7. No. 1. P. 89–96.
3. *Lamperti J.* Remarks on maximal branching processes // Теория вероятностей и ее применения, 1972. Т. 17. № 1. С. 46–54.
4. *Vatutin V. A., Zubkov A. M.* Branching processes. II // *J. Sov. Math.*, 1993. Vol. 67. No. 6. P. 3407–3485.
5. *Лебедев А. В.* Максимальные ветвящиеся процессы с неотрицательными значениями // Теория вероятностей и ее применения, 2005. Т. 50. № 3. С. 564–570.
6. *Лебедев А. В.* Максимальные ветвящиеся процессы // Современные проблемы математики и механики, 2009. Т. 4. № 1. С. 93–106.
7. *Лебедев А. В.* Максимальные ветвящиеся процессы с несколькими типами частиц // Вестн. Моск. ун-та. Сер. 1: Математика. Механика, 2012. № 3. С. 8–13.
8. *Лебедев А. В.* Неклассические задачи стохастической теории экстремумов: Дис. . . . докт. физ.-мат. наук. — М.: МГУ, 2016.
9. *Aydogmus O., Ghosh A. P., Ghosh S., Roitershtein A.* Coloured maximal branching process // Теория вероятностей и ее применения, 2014. Т. 59. № 4. С. 790–800.
10. *Browne S., Coffman E. G., Gilbert E. N., Wright P. E.* Gated, exhaustive, parallel service // *Probab. Eng. Inform. Sc.*, 1992. Vol. 2. No. 2. P. 217–239.
11. *Browne S., Coffman E. G., Gilbert E. N., Wright P. E.* The gated infinite-server queue: Uniform service times // *SIAM J. Appl. Math.*, 1992. Vol. 52. No. 6. P. 1751–1762.
12. *Tan X., Knessl Ch.* Heavy traffic asymptotics for a gated, infinite-server queue with uniform service times // *SIAM J. Appl. Math.*, 1994. Vol. 54. No. 6. P. 1768–1779.
13. *Pinotsi D., Zazanis M. A.* Stability conditions for gated $M|G|\infty$ queues // *Probab. Eng. Inform. Sc.*, 2004. Vol. 18. No. 1. P. 103–110.
14. *Лебедев А. В.* Вентильная бесконечнолинейная система с неограниченными временами обслуживания и большой загрузкой // Проблемы передачи информации, 2003. Т. 39. № 3. С. 87–94.
15. *Лебедев А. В.* Вентильная бесконечнолинейная система с большой загрузкой и степенным хвостом // Проблемы передачи информации, 2004. Т. 40. № 3. С. 62–68.
16. *Горбунова А. В., Зарядов И. С., Матюшенков С. И., Самуйлов К. Е., Шоргин С. Я.* Аппроксимация времени отклика системы облачных вычислений // Информатика и её применения, 2015. Т. 9. Вып. 3. С. 32–38.
17. *Jirina M.* Stochastic branching processes with continuous state space // *Czech. Math. J.*, 1958. Vol. 8. No. 2. P. 292–313.
18. *Боровков А. А.* Эргодичность и устойчивость случайных процессов. — М.: УРСС, 1999. 440 с.
19. *Esary J., Prochan F., Walkup D.* Association of random variables with applications // *Ann. Math. Stat.*, 1967. Vol. 38. No. 5. P. 1466–1474.
20. *Булинский А. В., Шашкин А. П.* Предельные теоремы для ассоциированных случайных полей и родственных систем. — М.: Физматлит, 2008. 480 с.
21. *Bhattacharya R. N., Lee C.* Ergodicity of nonlinear first order autoregressive models // *J. Theor. Probab.*, 1995. Vol. 8. No. 1. P. 207–219.

Поступила в редакцию 07.08.17

MAXIMAL BRANCHING PROCESSES IN RANDOM ENVIRONMENT

A. V. Lebedev

Department of Probability Theory, Faculty of Mechanics and Mathematics, M. V. Lomonosov Moscow State University, Main Building, 1 Leninskiye Gory, Moscow 119991, Russian Federation

Abstract: The work continues the author’s long research in the theory of maximal branching processes that are obtained from classical branching processes by replacing the sum of offspring numbers by the maximum. One can say that the next generation is formed by the offspring of the most productive particle. Earlier, the author generalized processes with integer values up to processes with arbitrary nonnegative values, investigated their properties, and proved the limit theorems. Further, maximal branching processes with several types of particles were introduced and studied. In this paper, the author introduces the concept of maximal branching processes in random environment (with one type of particles) and an important case of the “power” random environment. In the latter case, the basic properties of maximal branching processes are studied and the ergodic theorem is proved. As an application, the author considers gated infinite-server queues.

Keywords: maximal branching processes; random environment; ergodic theorem; stable distributions; extreme value theory

DOI: 10.14357/19922264180206

References

1. Harris, T. 1963. *The theory of branching processes*. Berlin: Springer-Verlag. 230 p.
2. Lamperti, J. 1970. Maximal branching processes and long-range percolation. *J. Appl. Probab.* 7(1): 89–96.
3. Lamperti, J. 1972. Remarks on maximal branching processes. *Theor. Probab. Appl.* 17(1): 44–53.
4. Vatutin, V. A., and A. M. Zubkov. 1993. Branching processes. II. *J. Sov. Math.* 67(6):3407–3485.
5. Lebedev, A. V. 2006. Maximal branching processes with nonnegative values. *Theor. Probab. Appl.* 50(3):482–488.
6. Lebedev, A. V. 2009. Maksimal'nye vetyashchiesya protsessy [Maximal branching processes]. *Sovremennye problemy matematiki i mekhaniki* [Modern Problems of Mathematics and Mechanics] 4(1):93–106.
7. Lebedev, A. V. 2012. Maximal branching processes with several types of particles. *Mosc. Univ. Math. Bull.* 67(3):97–101.
8. Lebedev, A. V. 2016. Neklassicheskie zadachi stokhasticheskoy teorii ekstremumov [Nonclassical problems of extreme value theory]. Moscow. DSc Diss. 257 p.
9. Aydogmus, O., A. P. Ghosh, S. Ghosh, and A. Roitershtein. 2015. Colored maximal branching process. *Theor. Probab. Appl.* 59(4): 663–672.
10. Browne, S., E. G. Coffman, E. N. Gilbert, and P. E. Wright. 1992. Gated, exhaustive, parallel service. *Probab. Eng. Inform. Sc.* 2(2): 217–239.
11. Browne, S., E. G. Coffman, E. N. Gilbert, and P. E. Wright. 1992. The gated infinite-server queue: Uniform service times. *SIAM J. Appl. Math.* 52(6):1751–1762.
12. Tan, X., and Ch. Knessl. 1994. Heavy traffic asymptotics for a gated, infinite-server queue with uniform service times. *SIAM J. Appl. Math.* 54(6):1768–1779.
13. Pinotsi, D., and M. A. Zazanis. 2004. Stability conditions for gated $M|G|\infty$ queues. *Probab. Eng. Inform. Sc.* 18(1):103–110.
14. Lebedev, A. V. 2003. The gated infinite-server queue with unbounded service times and heavy traffic. *Probl. Inf. Transm.* 39(3):309–316.
15. Lebedev, A. V. 2004. Gated infinite-server queue with heavy traffic and power tail. *Probl. Inf. Transm.* 40(3):237–242.
16. Gorbunova, A. V., I. S. Zaryadov, S. I. Matyushenko, K. E. Samouylov, and S. Ya. Shorgin. 2015. Approksimatsiya vremeni otklika sistemy oblachnykh vychisleniy [The approximation of response time of a cloud computing system]. *Informatika i ee Primeneniya — Inform. Appl.* 9(4):32–38.
17. Jiřina, M. 1958. Stochastic branching processes with continuous state space. *Czech. Math. J.* 8(2):292–313.
18. Borovkov, A. A. 1998. *Ergodicity and stability of stochastic processes*. Wiley. 618 p.
19. Esary, J., F. Prochan, and D. Walkup. 1967. Association of random variables with applications. *Ann. Math. Stat.* 38(5):1466–1474.
20. Bulinski, A. V., and A. P. Shashkin. 2007. *Limit theorems for associated random fields and related systems*. World Scientific Publishing. 448 p.
21. Bhattacharya, R. N., and C. Lee. 1995. Ergodicity of nonlinear first order autoregressive models. *J. Theor. Probab.* 8(1):207–219.

Received August 7, 2017

Contributor

Lebedev Alexey V. (b. 1971) — Doctor of Science in physics and mathematics, associate professor, Department of Probability Theory, Faculty of Mechanics and Mathematics, M. V. Lomonosov Moscow State University, Main Building, 1 Leninskiye Gory, Moscow 119991, Russian Federation; avlebed@yandex.ru

ИЕРАРХИЧЕСКИЙ МЕТОД ПОРОЖДЕНИЯ МЕТАДАННЫХ ДЛЯ УПРАВЛЕНИЯ СЕТЕВЫМИ СОЕДИНЕНИЯМИ*

А. А. Грушо¹, Е. Е. Тимонина², С. Я. Шоргин³

Аннотация: Важным классом угроз для распределенных информационно-вычислительных систем (РИВС) является возможность организации нелегальных информационных взаимодействий и, наоборот, запрет на разрешенные информационные взаимодействия. Для предотвращения этого класса угроз предложена организация управления соединениями с помощью метаданных. Метаданные формируются на основе математических моделей бизнес-процессов, которые в РИВС представлены совокупностью информационных технологий (ИТ). Разобраны два метода построения метаданных для ИТ. Предложенный подход основан на иерархической декомпозиции ИТ и задач.

Ключевые слова: информационная безопасность; распределенные информационно-вычислительные системы; метаданные; иерархическая декомпозиция; составные задачи

DOI: 10.14357/19922264180207

1 Введение

Важным классом угроз для РИВС является возможность организации нелегальных информационных взаимодействий и, наоборот, запрет на разрешенные информационные взаимодействия [1–4]. Для предотвращения этого класса угроз в [2–4] предложена организация управления соединениями с помощью метаданных, а именно: зашифрованное взаимодействие двух хостов разрешено тогда, когда данное соединение служит для реализации производственной ИТ.

Вместе с изоляцией задач и системой идентификаторов решаемых задач данный подход обеспечивает достаточно высокую защищенность РИВС [2, 3].

Метаданные формируются на основе математических моделей бизнес-процессов, которые в РИВС представлены совокупностью ИТ. Часть информации, связанная с функционированием ИТ, используется для организации разрешительной системы соединения хостов в сети.

Информационная безопасность в таком подходе основана на следующих принципах:

- (1) политики безопасности (ПБ) формулируются на уровне моделей бизнес-процессов и моделей ИТ [5] и выражаются в терминах языков, используемых на этих уровнях;

- (2) основная идея управления соединениями в сети с помощью метаданных состоит в разделении плоскости логики и функционала решения задач от плоскости передачи данных через сеть [6]. При таком разделении безопасность соединений определяется проекцией ПБ верхней плоскости на нижнюю.

Соединения хостов разрешаются, когда на них находятся задачи, взаимодействующие по логике ИТ. Тем самым решается проблема возможности реализации разрешенных информационных ВЗАИМОДЕЙСТВИЙ и существенного ограничения возможности существования нелегальных информационных потоков.

2 Моделирование задач

Для формирования метаданных необходимо построить последовательность переходов от моделей бизнес-процессов к моделям ИТ, из которых выделяются метаданные.

Основой этих моделей является понятие задачи. Простейшее определение задачи — это преобразование входных данных в выходные. Преобразование состоит в выполнении некоторых функций. Значит, задача должна реализовать некоторый функционал. Для задачи A обозначим функционал

* Работа поддержана РНФ (проект 16-11-10227).

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

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

³Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, sshorgin@ipiran.ru

преобразования $\pi(A)$. Для выполнения преобразования задачи A необходимо обеспечить $\pi(A)$ исходными данными и значениями конфигурационных параметров функционала. Эти данные надо получить извне. Кроме того, необходимо распределить результаты преобразования входной информации или сохранить эти данные в памяти. Объединение этих функций организации решения задачи A будем обозначать через $\varphi(A)$. Таким образом, задача A определяется как пара $(\pi(A), \varphi(A))$.

Далее обобщим это определение на случай, когда задача A состоит из нескольких задач $\{B_1, B_2, \dots, B_s\}$.

Согласно введенному выше определению $A = (\pi(A), \varphi(A))$, а $B_i = (\pi(B_i), \varphi(B_i))$, $i = 1, \dots, s$.

Определение 1. Можно отождествить составную задачу A с множеством задач $\{B_1, B_2, \dots, B_s\}$ тогда и только тогда, когда выполняются следующие условия:

- (1) функционал множества $\{\pi(B_i), i = 1, \dots, s\}$ покрывает функционал $\pi(A)$;
- (2) объект $\varphi(A)$ имеет возможность взаимодействия с каждым из объектов $\{\varphi(B_i), i = 1, \dots, s\}$:
 - (а) для получения и передачи всех входных данных для каждой задачи B_i , $i = 1, \dots, s$;
 - (б) распределения всех выходных данных задач B_i , $i = 1, \dots, s$, которые должны быть определены в задаче A ;
 - (в) определения порядка выполнения задач B_i , $i = 1, \dots, s$, для решения задачи A ;
 - (г) связи задач B_i , $i = 1, \dots, s$, между собой.

Это определение сходно с понятием клиент-серверной архитектуры, в которой клиенты B_i организуют свою работу под управлением сервера $\varphi(A)$.

Введенное определение связано с иерархической декомпозицией задач как множеством подзадач. Развитием этого взгляда на задачу является понимание того, что сами подзадачи исходной задачи могут представлять собой множество задач. Например, имеется составная задача $A = \{B_1, B_2, \dots, B_s\}$. В то же время каждая $B_i = \{C_{i,j}\}$. Тогда возникает иерархическая декомпозиция задачи A на подзадачи. Такая иерархическая декомпозиция в теории искусственного интеллекта называется редукцией [7]. В [7] приведены алгоритмы построения редукции. При этом необходимо, чтобы все задачи были однозначно идентифицированы.

Отметим, что $\pi(C_{i,j})$ однозначно определяет названием задачи $C_{i,j}$, а объект $\varphi(C_{i,j})$ связан с объектом $\varphi(B_i)$. В связи с этим будем использовать следующее обозначение: $\varphi_X(C)$ — это объект

организации решения задачи C как подзадачи задачи X .

Задачи в иерархической декомпозиции могут повторяться. При этом если значения объектов $\pi(\cdot)$ для всех повторений одной задачи одинаковы, то значения объектов $\varphi(\cdot)$ для всех повторений различны, поэтому повторяющиеся задачи однозначно идентифицируются.

Повторное использование задач можно реализовать тремя способами.

Предположим, что задача является совместно используемым объектом. Тогда возможно параллельное обращение к этой задаче от различных ее повторений. Например, система управления базами данных может обрабатывать параллельно несколько запросов от разных источников. Тогда возможно, что такая задача A характеризуется обращением к ней с помощью разных объектов $\varphi_X(A)$.

Вместо параллельного обращения возможно создание очереди использования задачи A . Эта очередь будет состоять из цепочки объектов $\varphi_X(A)$ для разных X .

Третий способ связан с созданием копий для каждого экземпляра задачи A .

Таким образом, каждая составная задача A , представленная в виде множества подзадач, образует дерево подзадач $L(A)$ с корнем в A . Это дерево взаимно однозначно определяет дерево $L(\varphi(A))$. Вершинами дерева $L(\varphi(A))$ являются объекты $\varphi_X(C)$, где X — вершина дерева, порождающая подзадачу C .

В этом дереве находится вся информация о взаимодействиях задач. В принципе, каждое такое дерево можно рассматривать как максимальные метаданные, совокупность которых определяет порядок решения задачи A , поэтому это дерево может использоваться как разрешительная информация для соединений в сети. Однако эти данные сложны для использования в реальном времени.

Модель РИВС можно представить как составную задачу, включающую все задачи, которые можно решать с помощью этой системы. Можно считать, что объект $\varphi(\text{РИВС})$ способен выделять и конфигурировать различные наборы задач. Тогда модель ИТ — это подмножество задач модели РИВС, объединенных моделью бизнес-процесса и конкретными конфигурациями этих задач. Объект $\varphi(\text{ИТ})$ определяет также порядок преобразования исходных данных для реализации этой ИТ.

Определим задачу \mathfrak{M} для ИТ следующим образом. Эта задача связывает физические и системные ресурсы РИВС с множеством задач, которые определены деревом задач, построенным для данной ИТ, а именно: \mathfrak{M} отображает множество задач

дерева задач ИТ во множество компьютеров и сетевого оборудования. Конфигурирование компьютеров и сетевого оборудования должно обеспечивать связность решения задач данной ИТ. Поскольку дерево задач ИТ L изоморфно дереву $L(\varphi)$ этой технологии, то для функционирования этой ИТ необходимо вместе с объектами π разместить и согласовать взаимодействия объектов из дерева $L(\varphi)$.

Часть этих взаимодействий должна быть реализована через сеть. Традиционно расположение одной задачи на другом хосте задается в объекте φ . Это дает возможность хосту инициировать соединение с другим хостом, где находится необходимая задача.

Однако реализация соединения от хоста может порождать нелегальные информационные взаимодействия [1–4]. Также такой подход переносит бизнес-логику на уровень низшей плоскости. Поэтому подход, связанный с метаданными, предлагает не передавать информацию о размещении задач на уровень хостов (хотя в памяти могут храниться следы прошедших реализаций ИТ).

Для разделения плоскостей необходимо ввести для ИТ задачу \mathfrak{M} , к которой задача A на хосте $H(A)$ всегда может обратиться через защищенный канал с просьбой о разрешении взаимодействия с задачей B [2]. Задача \mathfrak{M} на основании метаданных определяет возможность разрешения взаимодействия с задачей B , исходя из текущей стадии реализации ИТ.

При разрешении взаимодействия задача \mathfrak{M} по защищенному каналу передает на хост $H(B)$ информацию о необходимости взаимодействия с задачей A на хосте $H(A)$. Также на хост $H(B)$ передается ключ для связи с хостом $H(A)$ и другие параметры организации соединения. Хосту $H(A)$ также передается ключ для связи с хостом $H(B)$ и другие параметры организации соединения.

3 Информационная безопасность в распределенных информационно-вычислительных системах

Рассмотрим случай, когда порядок взаимодействий в ИТ жестко определен. При этом информация $L(\varphi)$ для ИТ не позволяет привлекать для вычислительного процесса экземпляры задач, не связанных с данным фрагментом дерева $L(\varphi)$.

Теорема 1. *В ИТ, представленной деревом $L(\varphi)$, для того чтобы все разрешенные жестко определенные взаимодействия задач были разрешены, а остальные были запрещены, необходимо и достаточно, чтобы*

для каждой составной задачи A в объекте $\varphi(A)$ однозначно (без избыточности) был определен порядок всех взаимодействий непосредственных подзадач задачи A и исключены взаимодействия, не соответствующие этому порядку.

Доказательство. Докажем достаточность. Каждая составная задача A в ИТ определяет поддерево в $L(\varphi)$ с корнем в A . Если строить дерево вниз, то число слоев будем называть глубиной (это соответствует высоте при построении дерева вверх). Доказательство достаточности будем вести по индукции по значениям глубины. Пусть составное дерево задачи A имеет глубину $h = 1$. Множество взаимодействий всех подзадач решает задачу A . По условию теоремы для такой задачи, представленной деревом $\varphi(A)$, все разрешенные взаимодействия задач разрешены, а остальные — запрещены, т. е. для ИТ, определяемой этой задачей A , достаточность доказана.

Предположим, что достаточность доказана для любых $\varphi(A)$ глубиной меньше h . Пусть глубина $\varphi(A)$ равна h . Предположим, что задачи A_1, A_2, \dots, A_t являются составными и расположены на глубине $h - 1$. Тогда порождаемые ими деревья имеют глубину 1 и для каждого из них доказана достаточность. Для дерева глубиной $h - 1$ достаточность выполняется по предположению индукции. Тогда в исходном дереве порядок всех взаимодействий однозначно определен и разрешен. Остальные взаимодействия отсутствуют.

Докажем необходимость. Каждая составная задача A может рассматриваться как ИТ. Тогда по условию для дерева $\varphi(A)$ однозначно определен порядок взаимодействий всех подзадач. Этот порядок порождает единственный порядок взаимодействий подзадач A . Это следует из того, что любое поддерево, замыкающееся на подзадачу A , можно считать выполнением этой подзадачи и это выполнение не влияет на порядок взаимодействий непосредственных подзадач A . При этом возможно неоднократное возвращение к любой из подзадач A в этом однозначном порядке. Теорема доказана.

Следствие 1. *Дерево $L(\varphi)$ можно рассматривать как метаданные для ИТ в том смысле, что соединения в сети могут однозначно определяться в условиях теоремы 1.*

4 Модель допустимых взаимодействий

Пусть метаданные для каждой составной задачи A — это множества идентификаторов непосредственных подзадач A , помеченных также иденти-

фикатором ИТ. Эти множества легко вычисляются по дереву $L(\varphi)$ для всех задач и подзадач ИТ и определяют метаданные ИТ. Обозначим для любой составной задачи A множество идентификаторов ее непосредственных подзадач из ИТ через $c(A)$, включая идентификатор самой A .

Определение 2. Назовем простыми допустимыми взаимодействиями задачи A из ИТ любые взаимодействия задач с идентификаторами из $c(A)$.

Тогда дерево $L(\varphi)$ порождает семейство множеств простых допустимых взаимодействий. Это семейство определяет метаданные.

Таким образом, при обращении к составной задаче A согласно метаданным рассматриваемой ИТ разрешаются любые простые допустимые взаимодействия из множества $c(A)$. При этом сама задача A входит также во множество подзадач какой-то задачи C верхнего уровня, что позволяет ей взаимодействовать с задачами, определяемыми $c(C)$, но другие задачи из $c(A)$ это делать не могут. Движение по связанным множествам $c(\cdot)$ потенциально позволяет задаче A снабжать и получать любые данные от любых других задач ИТ. Построенные метаданные значительно проще в реализации и позволяют объектам $\varphi(A)$ изменять порядок взаимодействий внутри множества, определяемого $c(A)$.

Определение 3. Любую связную последовательность множеств $c(A)$ для A из ИТ назовем допустимыми взаимодействиями.

Лемма 1. Все жестко определенные реализации ИТ входят в множество допустимых взаимодействий.

Доказательство. Все жестко определенные реализации ИТ для каждой составной задачи образуют множество простых допустимых взаимодействий. Тогда по теореме 1 все жестко определенные реализации ИТ входят в множество допустимых взаимодействий.

На множествах $c(A)$ как на вершинах построим граф метаданных G . Ребро соединяет $c(A)$ с $c(B)$ тогда и только тогда, когда B является подзадачей A .

Лемма 2. Граф G является деревом, изоморфным дереву $L(\varphi)$.

Доказательство. Множества вершин графа G и графа $L(\varphi)$ взаимно однозначно соответствуют друг другу. Вершина $c(A)$ соответствует $\varphi(A)$.

Ребро $(c(A), c(B))$ взаимно однозначно соответствует связи объектов $\varphi(A)$ и $\varphi(B)$. Лемма доказана.

Пусть ПБ в каждой вершине $\varphi(A)$ дерева $L(\varphi)$ определяет возможность сквозной передачи информации через $\varphi(A)$. Тогда согласно лемме 2 метаданные G могут не заметить нарушение ПБ

только для взаимодействий внутри множеств $c(A)$ для задач рассматриваемой ИТ.

Отметим, что в множествах $c(A)$ всегда присутствует идентификатор ИТ. Это позволяет не допускать нелегальных взаимодействий задач разных ИТ.

Из лемм 1, 2 и последних замечаний следует теорема 2.

Теорема 2. Метаданные G не разрешают недопустимых взаимодействий.

Следствие 2. Если соединения в сети определяются метаданными G , то недопустимых взаимодействий хостов нет.

Таким образом, метаданные G разрешают изменение порядка взаимодействий в сети только в рамках изменения порядка допустимых взаимодействий задач ИТ.

5 Заключение

В работе разобраны два метода построения метаданных для ИТ.

Первый метод связан с жесткой последовательностью взаимодействий задач в ИТ. Второй метод неоднозначно определяет порядок взаимодействий задач, но проще для вычислений [4] и может быть реализован в реальном времени в разрешительной системе установления соединений в сети [2]. Оба метода могут реализовать требования ПБ по безопасности взаимодействий.

Предложенный подход основан на иерархической декомпозиции ИТ и задач, где каждая задача A представляется парой $(\pi(A), \varphi(A))$. Такое представление позволяет абстрагироваться от существа алгоритмов и сложности их взаимодействий. Однако остается открытой проблема практического построения объектов $\varphi(A)$. Если эта проблема для ИТ может быть решена несложно, то будет получен эффективный метод построения метаданных для управления сетями.

Литература

1. 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.
2. Grusho A. A., Timonina E. E., Shorgin S. Ya. Modelling for ensuring information security of the distributed information systems // 31th European Conference on Modelling and Simulation Proceedings. — Dudweiler, Germany: Digitaldruck Pirrot GmbH, 2017. P. 656–660. http://www.scs-europe.net/dlib/2017/ecms2017acceptedpapers/0656-probstat_ECMS2017_0026.pdf.

3. Grusho A., Grusho N., Zabezhailo M., Zatsarinny A., Timonina E. Information security of SDN on the basis of meta data // *Computer network security* / Eds. J. Rak, J. Bay, I. V. Kotenko, *et al.* — Lecture notes in computer science ser. — Springer, 2017. Vol. 10446. P. 339–347. https://link.springer.com/chapter/10.1007/978-3-319-65127-9_27.
4. Grusho A., Timonina E., Shorgin S. Security models based on stochastic meta data // *Analytical and computational methods in theory probability* / Eds. V. Rykov, N. Singpurwalla, A. Zubkov. — Lecture notes in computer science ser. — Springer, 2017. Vol. 10684. P. 388–400. https://link.springer.com/chapter/10.1007/978-3-319-71504-9_32.
5. Грушо А.А., Применко Э.А., Тимонина Е.Е. Теоретические основы компьютерной безопасности. — М.: Академия, 2009. 272 с.
6. Grusho A.A., Abaev P.O., Shorgin S.Ya., Timonina E.E. Graphs for information security control in software defined networks // *AIP Conf. Proc.*, 2017. Vol. 1863. P. 090002-1–090002-4. <http://aip.scitation.org/doi/pdf/10.1063/1.4992267>.
7. Нильсон Н. Искусственный интеллект. Методы поиска решений / Пер. с англ. — М.: Мир, 1973. 272 с. (*Nilsson N.J. Problem-solving methods in artificial intelligence.* — New York, NY, USA: McGraw-Hill Publ. Co., 1971. 255 p.)

Поступила в редакцию 21.02.18

HIERARCHICAL METHOD OF META DATA GENERATION FOR CONTROL OF NETWORK CONNECTIONS

A. A. Grusho, E. E. Timonina, and S. Ya. Shorgin

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

Abstract: An important class of threats for distributed information systems is the possibility of organization of illegal information interactions and, vice versa, prohibition of the allowed information interactions. For preventing these threats, creation of connections control with the help of meta data is proposed. Meta data are created on the basis of mathematical models of business processes which are provided in distributed information systems by a set of information technologies. In the paper, two methods of meta data creation for information technologies are considered. The suggested approach is based on hierarchical decomposition of information technologies and tasks.

Keywords: information security; distributed information systems; meta data; hierarchical decomposition; composite tasks

DOI: 10.14357/19922264180207

Acknowledgments

The paper was supported by the Russian Science Foundation (project 16-11-10227).

References

1. Grusho, A., M. Zabezhailo, and A. Zatsarinny. 2015. Kontrol' i upravlenie informatsionnymi potokami v oblachnoy srede [Information flow monitoring and control in cloud computing environment]. *Informatika i ee Primeneniya — Inform. Appl.* 9(4):95–101.
2. Grusho, A.A., E. E. Timonina, and S.Ya. Shorgin. 2017. Modelling for ensuring information security of the distributed information systems. *31th European Conference on Modelling and Simulation Proceedings*. Digitaldruck Pirrot GmbH Dudweiler, Germany. 656–660. Available at: http://www.scs-europe.net/dlib/2017/ecms2017acceptedpapers/0656-probstat_ECMS2017_0026.pdf (accessed February 20, 2018).
3. Grusho, A., N. Grusho, M. Zabezhailo, A. Zatsarinny, and E. Timonina. 2017. Information security of SDN on the basis of meta data. *Computer network security*. Eds. J. Rak, J. Bay, I. V. Kotenko, *et al.* Lecture notes in computer science ser. Springer. 10446:339–347. Available at: https://link.springer.com/chapter/10.1007/978-3-319-65127-9_27 (accessed February 20, 2018).
4. Grusho, A., E. Timonina, and S. Shorgin. 2017. Security models based on stochastic meta data. *Analytical and computational methods in theory probability*. Eds. V. Rykov, N. Singpurwalla, and A. Zubkov. Lecture notes in computer science ser. Springer. 10684:388–400. Available at: https://link.springer.com/chapter/10.1007/978-3-319-71504-9_32 (accessed February 20, 2018).
5. Grusho, A., Ed. Применко, and E. Timonina. 2009. *Teoreticheskie osnovy komp'yuternoy bezopasnosti* [Theoretical bases of computer security]. Moscow: Academy. 272 p.
6. Grusho, A.A., P.O. Abaev, S.Ya. Shorgin, and E.E. Timonina. 2017. Graphs for information security con-

- trol in software defined networks. *AIP Conf. Proc.* 1863:090002. 4 p. Available at: <http://aip.scitation.org/doi/pdf/10.1063/1.4992267> (accessed February 20, 2018).
7. Nilsson, N.J. 1971. *Problem-solving methods in artificial intelligence*. New York, NY: McGraw-Hill Publ. Co. 255 p.

Received February 21, 2018

Contributors

Grusho Alexander A. (b. 1946) — Doctor of Science in physics and mathematics, professor, principal scientist, Institute of Informatics Problems, Federal Research Center “Computer Sciences and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; grusho@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 119133, Russian Federation; eltimon@yandex.ru

Shorgin Sergey Ya. (b. 1952) — Doctor of Science in physics and mathematics, professor, principal scientist, Institute of Informatics Problems, Federal Research Center “Computer Sciences and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; sshorgin@ipiran.ru

МОДЕЛИ УПРАВЛЕНИЯ РИСКОМ В ГАУССОВСКИХ СТОХАСТИЧЕСКИХ СИСТЕМАХ*

А. Н. Тырсин¹, А. А. Сурина²

Аннотация: Описан новый подход к исследованию риска многомерных стохастических систем. Он основан на гипотезе о том, что риском можно управлять за счет изменения вероятностных свойств компонент многомерной стохастической системы, в качестве которых используют факторы риска. Исследован случай гауссовских стохастических систем, описываемых случайными векторами, имеющими многомерное нормальное распределение. Как показало моделирование, не учтенные в явном виде многомерность системы и взаимная коррелированность ее компонент могут привести к существенному занижению фактического риска. Приведены результаты расчета вероятности опасного исхода в зависимости от числовых характеристик многомерной гауссовской случайной величины — ковариационной матрицы и вектора математических ожиданий. Выполнена апробация предложенной модели на примере анализа популяционного риска сердечно-сосудистых заболеваний. Описаны модели управления риском в виде задач его минимизации или достижения заданного уровня. Управляющими переменными являются числовые характеристики случайного вектора — ковариационная матрица и вектор математических ожиданий. Проведена апробация метода управления риском с помощью статистического моделирования методом Монте Карло.

Ключевые слова: риск; модель; стохастическая система; случайный вектор; управление; нормальное распределение

DOI: 10.14357/19922264180208

1 Введение

Уже не вызывает сомнений наличие общемировой тенденции быстрого роста ущерба от природных катаклизмов, техногенных катастроф, террористических актов и экономических потрясений. Многие авторы отмечают, что темпы роста ущерба значительно превосходят темпы роста экономики [1–3]. Это можно объяснить постоянным возрастанием риска в условиях научно-технической революции и форсированного развития техносферы [4]. Очевидно, что для снижения ущерба от природных катаклизмов, техногенных катастроф, террористических актов и экономических потрясений необходимо повысить безопасность функционирования соответствующих систем, а значит, снизить риск. Для этого необходимы адекватные модели и эффективные методы управления риском систем.

Реальные системы, как правило, являются многомерными, их функционирование во многом носит стохастический характер, у них часто можно выделить десятки различных факторов риска [2]. При решении задачи управления риском необходимо опираться на модель риска.

Обычно моделирование риска сводится к выделению опасных исходов, количественному заданию последствий от их наступления и оцениванию вероятностей этих исходов [5, с. 37–43]. При этом вклад компонент многомерной системы объединяют и рассматривают уже одномерную систему как случайную величину [5, с. 148–156; 6, с. 82–87].

Но вопрос взаимного влияния опасных ситуаций, вызванных разными элементами многомерной системы, мало исследован, чаще всего им пренебрегают или существенно упрощают, считая разные опасные исходы взаимно независимыми, и пренебрегают вероятностью их одновременного наступления.

Для относительно простых объектов, когда можно априори указать все опасные исходы, при наличии статистической информации или экспертных оценок о шансах их появления в целом данный подход дает приемлемые на практике результаты. Обычно здесь удается накопить достаточную статистику для оценивания вероятностей наступления опасных исходов, а форма взаимосвязи между элементами системы является достаточно простой и может быть описана, например, с помощью ло-

*Работа выполнена при финансовой поддержке РФФИ (проект 17-01-00315а).

¹Уральский федеральный университет имени первого Президента России Б. Н. Ельцина; Институт экономики Уральского отделения Российской академии наук, at2001@yandex.ru

²Южно-Уральский государственный университет (национальный исследовательский университет), dallila87@mail.ru

гико-вероятностных моделей риска [7] в рамках теории структурно-сложных систем [8].

Однако у сложных систем структуру взаимодействия между элементами обычно не удается описать с помощью логико-вероятностных моделей — стохастические связи между элементами не позволяют их адекватно моделировать с помощью алгебры логики (AND, OR, NOT), а изменения состояния элементов и самой системы носят непрерывный характер. Понятия опасных исходов также могут размываться, делая невозможным их конкретное выделение. К таким системам можно отнести, например, социально-экономические системы, включая территориальные и региональные системы, живые системы, например человека с точки зрения состояния здоровья.

Таким образом, несмотря на большое число исследований, взаимному влиянию элементов и различных факторов риска на безопасность сложных многомерных систем уделяется недостаточно внимания. Во многих случаях, когда нет возможности явно связать разные факторы риска в виде логико-вероятностной модели, их корреляция при расчете риска не учитывается, поэтому проблематика исследований в области анализа риска, особенно в части создания эффективных моделей описания и управления риском сложных многомерных систем, в настоящее время становится одной из актуальных.

В [9, 10] предложен подход к моделированию риска, согласно которому стохастическую систему представляют в виде случайного вектора со взаимно коррелированными компонентами, а в качестве управляющих переменных используют его числовые характеристики. Целью статьи является описание моделей управления риском на основе данного подхода.

2 Модель риска в гауссовских стохастических системах

Пусть S — некоторая многомерная стохастическая система. Выделим в этой системе факторы риска X_1, X_2, \dots, X_m . В результате получим представление системы в виде случайного вектора $\mathbf{X} = (X_1, X_2, \dots, X_m)$ с некоторой плотностью вероятности $p_{\mathbf{X}}(\mathbf{x})$.

Вместо общепринятого выделения конкретных опасных ситуаций будем задавать геометрические области неблагоприятных исходов. Они могут выглядеть произвольным образом в зависимости от конкретной задачи и определяются на основе имеющейся априорной информации. Для определен-

ности опишем предлагаемый подход на примере распространенной концепции нежелательных событий как больших и маловероятных отклонений случайной величины относительно ее математического ожидания. Тогда опасными ситуациями будем считать большие и маловероятные отклонения выборочных значений x_{ij} любой из компонент X_j относительно математических ожиданий $\mu_j = M[X_j]$, $j = 1, 2, \dots, m$. Вероятность неблагоприятного исхода для каждой из компонент X_j зададим как

$$P(D_j) = P(X_j \in D_j) = P(X_j \notin \bar{D}_j),$$

$$\bar{D}_j = \{x : \mu_j - A_{1j}\sigma_j < x < \mu_j + A_{2j}\sigma_j\},$$

где σ_j — среднее квадратическое отклонение случайной величины X_j ; A_{1j} и A_{2j} — заданные нижний и верхний пороговые уровни (в единицах σ_j), т.е. область благоприятных исходов ограничена диапазоном $(\mu_j - A_{1j}\sigma_j; \mu_j + A_{2j}\sigma_j)$.

Теперь необходимо задать многомерную область опасных ситуаций D , учтя взаимное влияние компонент на появление неблагоприятных исходов. Она равна $D = \mathbf{R}^m \setminus \bar{D}$, где \bar{D} — область допустимых значений факторов риска. Опишем область \bar{D} . Это можно сделать различными способами. Наиболее оправданным с геометрической точки зрения представляется задать ее в виде внутренней области m -осного эллипсоида

$$\bar{D} = \left\{ \mathbf{x} = (x_1, x_2, \dots, x_m) : \sum_{j=1}^m \frac{(x_j - \mu'_j)^2}{A_j^2 \sigma_j^2} < 1 \right\}$$

с центром в точке $\boldsymbol{\mu}' = (\mu'_1, \mu'_2, \dots, \mu'_m)$, $\mu'_j = \mu_j + A_j \sigma_j$, $A_j = (A_{1j} + A_{2j})/2$, $j = 1, 2, \dots, m$. Тогда для случайного вектора \mathbf{X} вероятность неблагоприятного исхода будет равна

$$P(D) = P(\mathbf{X} \in D), \quad D = \left\{ \mathbf{x} = (x_1, x_2, \dots, x_m) : \sum_{j=1}^m \frac{(x_j - \mu_j)^2}{A_j^2 \sigma_j^2} \geq 1 \right\}. \quad (1)$$

Заметим, что в (1) область D неблагоприятных исходов представляет собой внешнюю область m -осного эллипсоида, у которого полуоси по каждой из координат равны $A_j \sigma_j$ соответственно, т.е. по каждой j -й оси эта область соответствует одномерному случаю D_j . Очевидно, когда исход не лежит на одной из осей, событие D может реализоваться и при отсутствии рискованных отклонений по всем компонентам (возможны ситуации $\mathbf{X} \in D$ и $\forall j X_j \notin D_j$).

Задав функцию последствий от опасных ситуаций в виде $g(\mathbf{x})$, получим модель для количественной оценки риска:

$$r(\mathbf{X}) = \int \dots \int_{\mathbf{R}^m} g(\mathbf{x}) p_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}.$$

Если принять

$$g(\mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in D; \\ 0, & \mathbf{x} \notin D, \end{cases} \quad (2)$$

то $r(\mathbf{X}) = P(\mathbf{X} \in D)$, т. е. риск оцениваем как вероятность неблагоприятного исхода. Если на ранней стадии исследования системы сложно достаточно точно описать функцию $g(\mathbf{x})$, то формула (2) становится оценкой $P(D)$ и является удобным начальным приближением модели риска.

Рассмотрим далее наиболее распространенный частный случай, когда \mathbf{X} имеет совместное нормальное распределение с плотностью вероятности

$$p_{\mathbf{X}}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^m |\Sigma|}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \mathbf{a})^T \Sigma^{-1} (\mathbf{x} - \mathbf{a}) \right\},$$

где $\mathbf{a} = (a_1, a_2, \dots, a_m)^T$ — вектор математических ожиданий; $\Sigma = \{\sigma_{ij}\}_{m \times m}$ — ковариационная матрица.

Использование гауссовского случайного вектора опирается на центральную предельную теорему [11]. Как показала апробация на ряде примеров,

такая идеализация не столь критична, и если есть какие-либо основания считать, что плотности вероятностей компонент вектора \mathbf{X} имеют более вытянутые хвосты, то это можно скорректировать за счет соответствующего задания функции $g(\mathbf{x})$.

Исследуем влияние многомерности и коррелированности факторов риска на вероятность появления неблагоприятных исходов.

Пример 1. Для наглядности рассмотрим двумерный гауссовский случайный вектор (X_1, X_2) с плотностью вероятности

$$p_{X_1, X_2}(x_1, x_2) = \frac{e^{-Q(x_1 - a_1, x_2 - a_2)/2}}{2\pi\sigma_1\sigma_2\sqrt{1 - \rho^2}}. \quad (3)$$

Здесь

$$Q(y_1, y_2) = \frac{1}{1 - \rho^2} \left(\frac{y_1^2}{\sigma_1^2} - \frac{2\rho y_1 y_2}{\sigma_1 \sigma_2} + \frac{y_2^2}{\sigma_2^2} \right),$$

где $y_i = x_i - a_i, i = 1, 2; \rho = \sigma_{12}/(\sigma_1\sigma_2)$ — коэффициент корреляции между X_1 и X_2 .

На рис. 1 показаны примеры реализаций стандартного нормального случайного вектора (X_1, X_2) для некоррелированных ($\rho = 0$) и коррелированных ($\rho = 0,9$) компонент. Видим, что увеличение тесноты корреляционной связи между компонентами приводит к вытягиванию диаграммы рассеяния и увеличению вероятности появления больших отклонений случайного вектора.

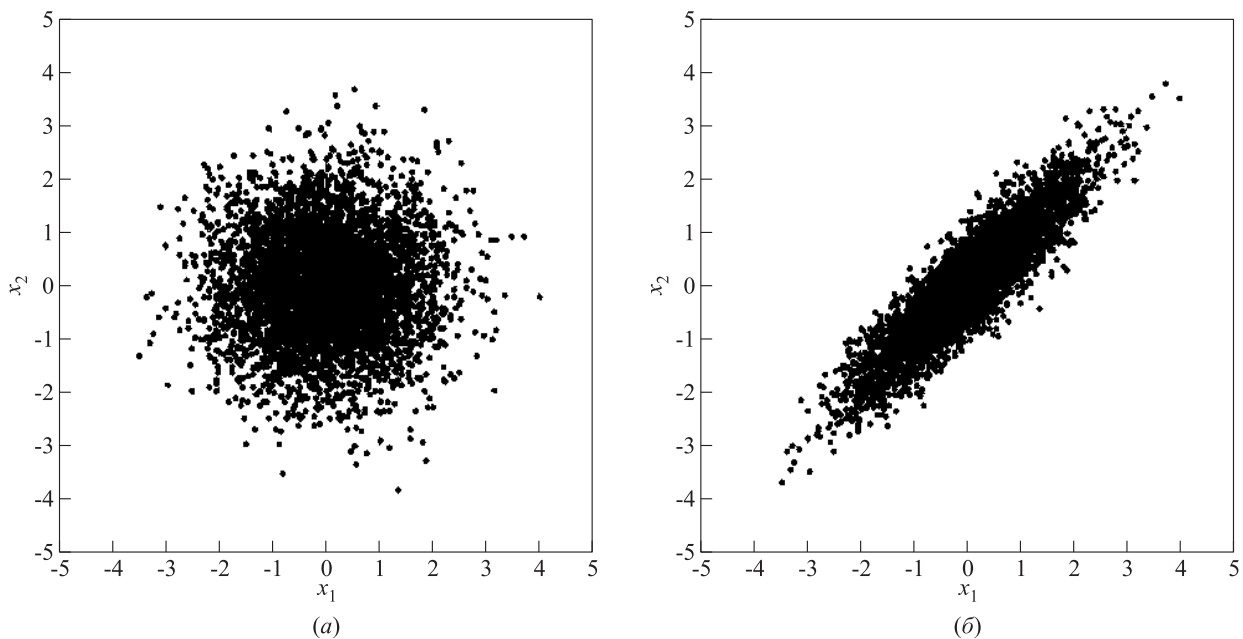


Рис. 1 Реализации стандартного нормального случайного вектора: (а) $\rho = 0$; (б) $\rho = 0,9$

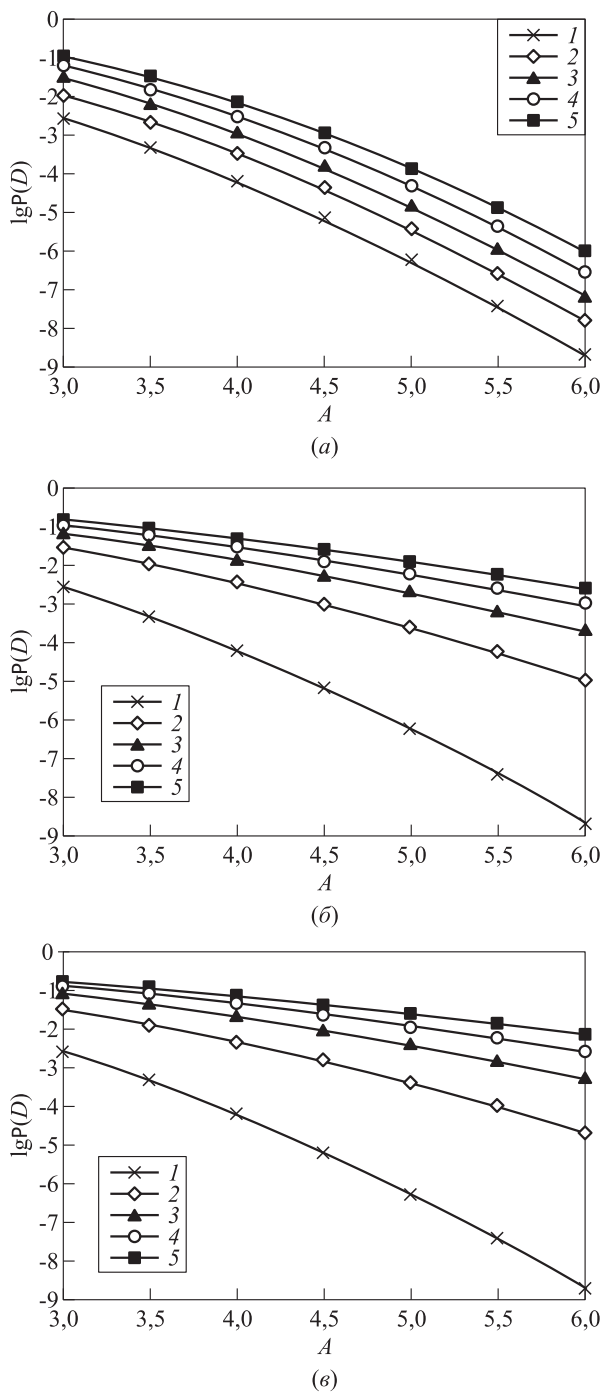


Рис. 2 Зависимости $\lg P(D)$ от порогового уровня A : (а) $D_e(\mathbf{X}) = 0$; (б) $D_e(\mathbf{X}) = 0,5$; (в) $D_e(\mathbf{X}) = 1$; 1 – $m = 1$; 2 – 2; 3 – 3; 4 – 4; 5 – $m = 5$

Пример 2. Зададим для определенности размерность вектора \mathbf{X} от 1 до 5. В [12] введен коэффициент тесноты совместной линейной корреляционной связи компонент случайного вектора \mathbf{X} , равный $D_e(\mathbf{X}) = 1 - |\mathbf{R}_X|^{1/m}$, где \mathbf{R}_X — корреляционная матрица

случайного вектора \mathbf{X} . Очевидно, что $0 \leq D_e(\mathbf{X}) \leq 1$. Случай $D_e(\mathbf{X}) = 0$ соответствует независимости компонент X_1, X_2, \dots, X_m , а при $D_e(\mathbf{X}) = 1$ имеем строгую линейную зависимость компонент.

Рассмотрим три случая: $D_e(\mathbf{X}) = 0$, $D_e(\mathbf{X}) = 0,5$ и $D_e(\mathbf{X}) = 1$. Результаты расчета вероятности неблагоприятного исхода (1) приведены на рис. 2. Для большей наглядности примем $A_1 = A_2 = \dots = A_m = A$.

Анализ графиков на рис. 2 говорит о следующем. Увеличение размерности m и тесноты корреляционной связи между компонентами случайного вектора \mathbf{X} приводит к резкому росту вероятности неблагоприятного исхода.

Особенно важным оказалось то, что даже относительно малая теснота корреляционной связи ($D_e(\mathbf{X}) = 0,5$), которая почти всегда наблюдается на практике, уже приводит к значительному росту вероятности $P(D)$. Эффект усиливается с увеличением значений A_j , что соответствует менее вероятным, но более опасным неблагоприятным исходам. Например, при $A = 6$ вероятность неблагоприятного исхода более чем в 7000 раз выше у коррелированной системы ($D_e(\mathbf{X}) = 1$) по сравнению с некоррелированной ($D_e(\mathbf{X}) = 0$). Поэтому при моделировании риска необходимо учитывать как фактор многомерности, так и тесноту корреляционных связей.

3 Апробация модели риска на примере анализа популяционного риска сердечно-сосудистых заболеваний

Одной из малоизученных проблем в медицине является комплексная оценка популяционного здоровья одновременно по нескольким факторам риска в их взаимосвязи. Это объясняется тем, что неясно, как учитывать вклад каждого фактора риска в общую оценку состояния здоровья. Обычно в таких случаях используются экспертные оценки, которые нельзя считать в полной мере объективными [13, 14].

Исследуем динамику изменения с возрастом популяционного риска сердечно-сосудистых заболеваний по основным биологическим факторам риска, к которым относят артериальную гипертензию, дислипидемию, повышенный уровень глюкозы в крови и избыточную массу тела [15].

В качестве биологических параметров, характеризующих эти факторы риска, используют уровень

общего холестерина (ОХС), систолическое артериальное давление (САД), индекс массы тела (ИМТ), уровень глюкозы (УГ).

Статистический материал получен в результате комплексного сплошного углубленного клинико-эпидемиологического обследования мужской сельской популяции с гнездовой выборкой. Всего было обследовано 1402 мужчины одного из сел Челябинской области, что составило 93% от списочного состава села. Для всех пациентов был проведен необходимый комплекс клинических, лабораторных и инструментальных методов обследования для квалифицированного заключения о состоянии здоровья. Работу проводила бригада специалистов, состоящая из сотрудников кафедры госпитальной терапии и семейной медицины Челябинской государственной медицинской академии и врачей Челябинской областной клинической больницы № 1 [16].

Пороговые значения биологических параметров, характеризующих основные биологические факторы риска, при превышении которых риск сердечно-сосудистых осложнений резко возрастает (ниже этого значения — норма), в соответствии с [15] равны: САД — 140 мм рт. ст.; ИМТ — 25 кг/м²; ОХС — 5 ммоль/л; УГ — 5,5 ммоль/л. Исследование проводилось следующим образом.

Было сформировано четыре группы по возрастам: 18–24 года, 25–34 года, 35–44 года и 45–54 года. Проверка по критерию согласия χ^2 -Пирсона статистической гипотезы о соответствии каждой группы наблюдений для всех факторов риска нормальному распределению на уровне значимости 0,05 не была отклонена. Поэтому считаем, что имеем гауссовскую стохастическую систему размерности $m = 4$.

Затем для каждой группы были определены средние значения и ковариационные матрицы. Вычисление вероятности $P(D)$ можно выполнять двумя способами — с помощью численного интегрирования для малых размерностей ($m \leq 4$) или методом статистических испытаний Монте Карло [17] при размерности $m > 4$. Результаты расчета приведены в таблице.

Видим, что наблюдается тенденция роста риска возникновения сердечно-сосудистых осложнений.

Значения вероятностей риска возникновения сердечно-сосудистых осложнений

Возраст, лет	Вероятность
18–24	0,70
25–34	0,78
35–44	0,95
45–54	0,98

Полученные в целом высокие значения вероятностей $P(D)$ соответствуют фактическому состоянию здоровья.

Как показали результаты комплексного сплошного углубленного клинико-эпидемиологического обследования, в обследованной популяции здоровых лиц в возрасте старше 34 лет практически не оказалось.

4 Модели управления риском

Введенная модель риска позволяет на практике осуществлять управление стохастической системой с целью его снижения.

Пример 3. Проиллюстрируем данный подход на простейшем примере гауссовской стохастической системы с размерностью $m = 2$. На рис. 3 показана возможность уменьшения вероятности неблагоприятного исхода $P(D)$, а значит, и риска за счет варьирования параметров плотности $p_{\mathbf{x}}(\mathbf{x})$. Область неблагоприятного исхода D расположена выше линии в правом верхнем углу.

Видим, что возможные варианты изменения параметров случайного вектора (X_1, X_2) : уменьшение ковариации (или коэффициента корреляции), изменение математических ожиданий случайных величин, уменьшение дисперсий σ_1^2 или σ_2^2 — могут привести к снижению вероятности $P(D)$.

Суть управления риском гауссовской стохастической системы состоит в следующем. Задав функцию последствий от опасных ситуаций $g(\mathbf{x})$ и введя ограничения на допустимые значения элементов ковариационной матрицы $G(\Sigma)$ и средних значений компонент системы $H(\mathbf{a})$, сформулируем задачу минимизации риска с переменными Σ и \mathbf{a} :

$$r(\Sigma, \mathbf{a}) = \int \cdots \int_{\mathbf{R}^m} g(\mathbf{x}) p_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} \rightarrow \min_{\Sigma, \mathbf{a}},$$

$$\Sigma \in G(\Sigma), \mathbf{a} \in H(\mathbf{a}). \quad (4)$$

Задача (4) является задачей нелинейного программирования. Ее можно решить разными методами. Одним из них является метод барьерных функций (внутренних штрафных функций) [18]. Его основная идея состоит в приведении задачи поиска условного экстремума к последовательности задач нахождения безусловного экстремума вспомогательной функции:

$$F(\mathbf{X}, b_k) = r(\Sigma, \mathbf{a}) + P(\Sigma, \mathbf{a}, b_k),$$

где $P(\Sigma, \mathbf{a}, b_k)$ — штрафная функция; b_k — параметр штрафа.

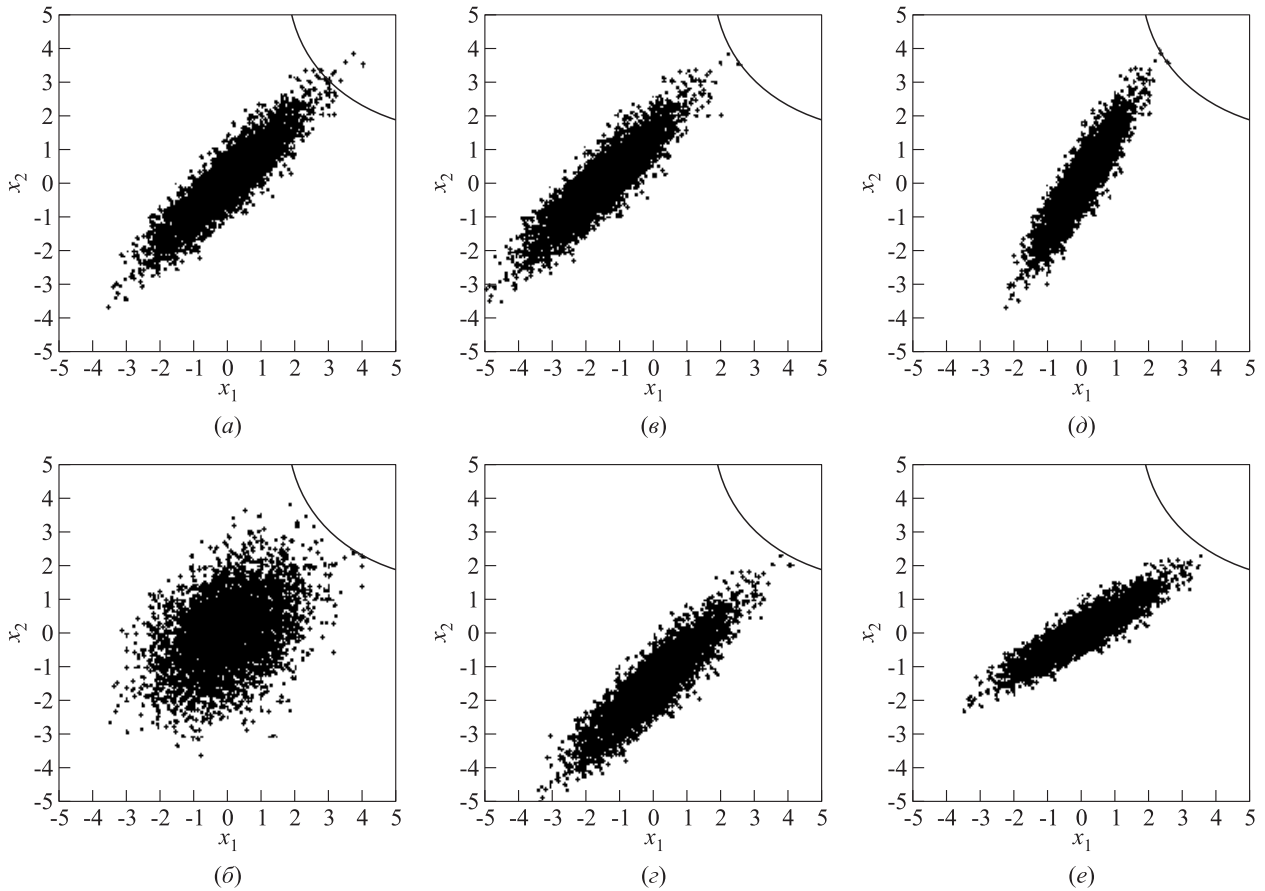


Рис. 3 Снижение риска: (а) исходное состояние; (б) за счет уменьшения корреляции; (в) и (г) за счет изменения математических ожиданий случайных величин X_1 или X_2 соответственно; (д) и (е) за счет уменьшения дисперсии σ_1^2 или σ_2^2 соответственно

Пример 4. Рассмотрим двумерный гауссовский случайный вектор с плотностью вероятности (3). Задача минимизации будет выглядеть как

$$r(\Sigma, \mathbf{a}) = \iint_{\mathbf{R}^2} \frac{g(x_1, x_2)}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-Q(x_1-a_1, x_2-a_2)/2} dx_1 dx_2 \rightarrow \min_{\Sigma, \mathbf{a}}$$

с ограничениями

$$\begin{aligned} \sigma_1^2\sigma_2^2 &> \sigma_{12}^2; \\ a_i^- &< a_i < a_i^+, \quad i = 1, 2; \\ \sigma_{ij}^- &< \sigma_{ij} < \sigma_{ij}^+, \quad j = 1, 2. \end{aligned}$$

Зададим конкретные значения: $a_1^- = a_2^- = -3$; $a_1^+ = a_2^+ = 3$; $\sigma_{12}^- = 0,1$; $\sigma_{12}^+ = 3$. Для определенности считаем, что $\forall \mathbf{x} \notin D \quad g(\mathbf{x}) = 0$. Получаем задачу

$$r(\Sigma, \mathbf{a}) = \iint_{\mathbf{R}^2} \frac{g(x_1, x_2)}{1,2\pi} \times e^{-((x_1-2)^2 - 1,6x_1x_2 + (x_2-2)^2)/(2 \cdot 0,6^2)} dx_1 dx_2 \rightarrow \min_{\Sigma, \mathbf{a}}$$

с ограничениями

$$\begin{aligned} \sigma_1^2\sigma_2^2 &> \sigma_{12}^2; \\ -3 &< a_1, a_2 < 3; \\ 0,1 &< \sigma_1, \sigma_2 < 3; \\ 0,1 &< \sigma_2 < 3. \end{aligned}$$

Выберем в качестве штрафной функции обратную:

$$P(\Sigma, \mathbf{a}, b^k) = -b^k \sum_{j=1}^m \frac{1}{t_j(\Sigma, \mathbf{a})}.$$

Тогда с учетом всех ограничений вспомогательная функция примет вид:

$$F(\mathbf{X}, b_k) = \int\int_{\mathbf{R}^2} \frac{1}{1,2\pi} e^{-((x_1-2)^2-1,6x_2x_2+(x_2-2)^2)/(2\cdot 0,6^2)} dx - b^k \left(\frac{1}{\sigma_1} + \frac{1}{\sigma_2} + \frac{1}{\sigma_1^2\sigma_2^2 - \sigma_{12}^2} + \frac{1}{a_1 + 3} + \frac{1}{3 - a_1} + \frac{1}{a_2 + 3} + \frac{1}{3 - a_1} + \frac{1}{\sigma_1 - 0,1} + \frac{1}{3 - \sigma_1} + \frac{1}{\sigma_2 - 0,1} + \frac{1}{3 - \sigma_2} \right) \rightarrow \min_{\Sigma, \mathbf{a}}.$$

Поиск минимума вспомогательной функции находим с помощью покоординатного спуска. Начальные значения параметров:

$$\Sigma^0 = \begin{pmatrix} 1 & 0,8 \\ 0,8 & 1 \end{pmatrix}; \mathbf{a}^0 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}; b^0 = 10.$$

Задача минимизации была решена при значениях параметров: $a_1 = 2,6; a_2 = 0,6; \sigma_1 = 2,9; \sigma_2 = 2,9; \sigma_{12} = 1,39 \cdot 10^{-16}$. При этом минимум целевой функции с точностью до 0,001: $r(\Sigma^*, \mathbf{a}^*) = 0,041$.

Задача (4) полезна лишь в качестве первого приближения модели управления риском, так как не учитывает ограничений, связанных с затратами на изменения варьируемых параметров относительно своих начальных значений.

Если ввести ограничения на затраты, связанные с изменением переменных Σ и \mathbf{a} , то получим задачу:

$$\left. \begin{aligned} r(\Sigma, \mathbf{a}) &= \int \dots \int_{\mathbf{R}^m} g(\mathbf{x}) p_{\mathbf{X}}(\mathbf{x}) dx \rightarrow \min_{\Sigma, \mathbf{a}}, \\ \Sigma &\in G(\Sigma), \mathbf{a} \in H(\mathbf{a}), \\ a_i &= a_i^0 + \delta_i, v_i(\delta_i) \leq V_i, i = 1, \dots, m, \\ \sigma_{ij} &= \sigma_{ij}^0 + \Delta_{ij}, w_{ij}(\Delta_{ij}) \leq W_{ij}, \\ & i, j = 1, \dots, m, \end{aligned} \right\} (5)$$

где $v_i(\delta_i)$ — функция затрат на изменение среднего значения i -й компоненты, имеющей начальное значение a_i^0 ; V_i — предельная величина затрат на это изменение; $w_{ij}(\Delta_{ij})$ — функция затрат на изменение ковариации между i -й и j -й компонентами; σ_{ij}^0 — начальное значение ковариации; W_{ij} — предельная величина затрат на это изменение.

Минимизация риска не всегда может быть приемлемым управлением. Альтернативой является достижение приемлемого риска r^* при минимальных изменениях числовых характеристик гауссовской системы \mathbf{X} . Здесь возможны два варианта постановки задачи.

Во-первых, на основе (5) получаем альтернативный вариант:

$$\sum_{i=1}^m v_i(\delta_i) + \sum_{i=1}^m \sum_{j=i}^m w_{ij}(\Delta_{ij}) \rightarrow \min_{\Sigma, \mathbf{a}},$$

$$\Sigma \in G(\Sigma), \mathbf{a} \in H(\mathbf{a}),$$

$$\delta_i = a_i - a_i^0, \Delta_{ij} = \sigma_{ij} - \sigma_{ij}^0, i, j = 1, \dots, m,$$

$$r(\Sigma, \mathbf{a}) = r^*.$$

Во-вторых, если сложно задать функции затрат $v_i(\cdot)$ и $w_{ij}(\cdot)$, то можно минимизировать суммарное изменение Σ и \mathbf{a} , перейдя к задаче

$$\sum_{i=1}^m \alpha_i \delta_i^2 + \sum_{i=1}^m \sum_{j=i}^m \beta_{ij} \Delta_{ij}^2 \rightarrow \min_{\Sigma, \mathbf{a}},$$

$$\Sigma \in G(\Sigma), \mathbf{a} \in H(\mathbf{a}),$$

$$\delta_i = a_i - a_i^0, \Delta_{ij} = \sigma_{ij} - \sigma_{ij}^0, i, j = 1, \dots, m,$$

$$r(\Sigma, \mathbf{a}) = r^*,$$

где α_i и β_{ij} — весовые коэффициенты.

5 Заключение

1. Предложен новый подход к исследованию риска сложных систем. В его основе лежит моделирование системы в виде многомерной случайной величины, компоненты которой являются факторами риска.
2. Для гауссовских стохастических систем предложены модели управления риском на основе его минимизации или достижения заданного уровня, используя в качестве управляющих переменных числовые характеристики случайного вектора — вектор математических ожиданий и ковариационную матрицу.
3. В настоящее время обычно при исследовании риска сложных многомерных систем не выделяют в явном виде их компоненты и их коррелированность. Как показало моделирование, неучет в явном виде многомерности системы и взаимной коррелированности ее компонент может привести к существенному занижению фактического риска. Усиление тесноты корреляционной связи между факторами риска приводит к значительному росту вероятности одновременного принятия ими опасных значений.
4. Предложенная гипотеза об управлении риском сложной системы на основе изменения числовых характеристик ее математической модели в форме случайного вектора носит предварительный характер. Необходимо выполнить апробацию данного подхода на ряде задач.

Литература

1. Гор А. Земля на чаше весов. В поисках новой общей цели // Новая постиндустриальная волна на Западе: Антология / Пер. с англ. — М.: Academia, 1999. С. 557–571. (*Gore A. Earth in the balance. Forging a new common purpose.* — London: Earthscan Publications Ltd., 1992.)
2. Воробьев Ю. Л., Малинецкий Г. Г., Махутов Н. А. Управление риском и устойчивое развитие: Человеческое измерение // Известия вузов. Прикладная нелинейная динамика, 2000. Т. 8. № 6. С. 12–26.
3. Порфирьев Б. Н. Снижение природных рисков экономического развития России: роль государства // Актуальные проблемы гражданской защиты: Мат-лы 11-й Междунар. науч.-практич. конф. по проблемам защиты населения и территорий от чрезвычайных ситуаций. — Н. Новгород: Вектор-ТиС, 2006. С. 44–50. gov.mari.ru/debzn/omgo/46.djvu.
4. Порфирьев Б. Н. Управление в чрезвычайных ситуациях. Итоги науки и техники. Проблемы безопасности: чрезвычайные ситуации. Т. 1. — М.: ВИНТИ, 1991. 204 с.
5. Вишняков Я. Д., Радаев Н. Н. Общая теория рисков. — 2-е изд., испр. — М.: Академия, 2008. 368 с.
6. Акимов В. А., Лесных В. В., Радаев Н. Н. Риски в природе, техносфере, обществе и экономике. — М.: Деловой экспресс, 2004. 352 с.
7. Соложенцев Е. Д. Сценарное логико-вероятностное управление риском в бизнесе и технике. — 2-е изд. — СПб.: Бизнес-пресса, 2006. 560 с.
8. Рябинин И. А. Надежность и безопасность структурно-сложных систем. — СПб.: Политехника, 2000. 248 с.
9. Тырсин А. Н. О моделировании риска в системах критических инфраструктур // Экономические и технические аспекты безопасности строительных критических инфраструктур: Тезисы Междунар. конф. — Екатеринбург: УрФУ, 2015. С. 205–208. http://elar.urfu.ru/bitstream/10995/33468/1/safety_2015.pdf.
10. Тырсин А. Н., Сурина А. А. Моделирование риска в многомерных стохастических системах // Вестн. Томского государственного университета. Управление, вычислительная техника и информатика, 2017. № 2(39). С. 65–72.
11. Гнеденко Б. В. Курс теории вероятностей. — 8-е изд., испр. и доп. — М.: Едиториал УРСС, 2005. 448 с.
12. Pena D., Rodriguez J. Descriptive measures of multivariate scatter and linear dependence // J. Multivariate Anal., 2003. Vol. 85. P. 361–374.
13. Кирьянов Б. Ф., Токмачёв М. С. Математические модели в здравоохранении. — Великий Новгород: НовГУ им. Ярослава Мудрого, 2009. 279 с.
14. Цинкер М. Ю., Кирьяков Д. А., Камалтдинов М. Р. Применение комплексного индекса нарушения здоровья населения для оценки популяционного здоровья в Пермском крае // Известия Самарского научного центра РАН, 2013. Т. 15. № 3(6). С. 1988–1992.
15. Профилактика хронических неинфекционных заболеваний. Рекомендации. — М., 2013. 128 с. <http://www.webmed.irkutsk.ru/doc/pdf/prevent.pdf>.
16. Тырсин А. Н., Калев О. Ф., Яшин Д. А., Лебедева О. В. Оценка состояния здоровья популяции на основе энтропийного моделирования // Математическая биология и биоинформатика, 2015. Т. 10. Вып. 1. С. 206–219. doi: 10.17537/2015.10.206.
17. Михайлов Г. А., Войтишек А. В. Численное статистическое моделирование. Методы Монте-Карло. — М.: Академия, 2006. 368 с.
18. Пантелеев А. В., Летова Т. А. Методы оптимизации в примерах и задачах. — 3-е изд., стер. — М.: Высшая школа, 2008. 544 с.

Поступила в редакцию 21.08.17

A MODEL OF RISK MANAGEMENT IN GAUSSIAN STOCHASTIC SYSTEMS

A. N. Tyrsin^{1,2} and A. A. Surina³

¹Ural Federal University named after first President of Russia B. N. Yeltsin, 19 Mira Str., Ekaterinburg 620002, Russian Federation

²Institute of Economics, Ural Branch of the Russian Academy of Sciences, 29 Moskovskaya Str., Yekaterinburg 620014, Russian Federation

³Institute of Natural Sciences, South Ural State University, 87 Lenin Ave., Chelyabinsk 454080, Russian Federation

Abstract: A new approach to research of risk of multidimensional stochastic systems is described. It is based on a hypothesis that the risk can be managed by changing probabilistic properties of a component of a multidimensional stochastic system. The case of Gaussian stochastic systems described by random vectors having the multidimensional normal distribution is investigated. Modeling has shown that multidimensionality of a system and relative

correlation of components unaccounted in an explicit form, can lead to essential understating of risk factors. Results of calculation of the probability of a dangerous outcome depending on numerical characteristics of a multidimensional Gaussian random variable (a covariance matrix and a vector of mathematical expectations) are given. Approbation of the suggested model is executed by the example of the analysis of the risk of cardiovascular diseases in population. Models of risk management in the form of a minimization problem or achievement of the given level are described. Control variables are the numerical characteristics of a random vector covariance matrix and a vector of mathematical expectations. Approbation of the method of risk management was carried out by means of statistical model operation by the Monte-Carlo method.

Keywords: risk; model; stochastic system; random vector; control; normal distribution

DOI: 10.14357/19922264180208

Acknowledgments

The work was supported by the Russian Foundation for Basic Research (project 17-01-00315a).

References

- Gore, A. 1992. *Earth in the balance. Forging a new common purpose*. London: Earthscan Publications Ltd. 440 p.
- Vorob'ev, Yu. L., G. G. Malinetskiy, and N. A. Makhutov. 2000. Upravlenie riskom i ustoychivoe razvitiye: chelovecheskoe izmerenie [Management of risk and sustainable development: Human measurement]. *Izvestiya vuzov. Prikladnaya nelineynaya dinamika* [Proceedings of the Universities. Applied Nonlinear Dynamics] 8(6):12–26.
- Porfir'ev, B. N. 2006. Snizhenie prirodnykh riskov ekonomicheskogo razvitiya Rossii: rol' gosudarstva [The reduction of natural risks of economic development of Russia: The role of the state]. *Aktual'nye problemy grazhdanskoj zashchity: Mat-ly 11-y Mezhdunar. nauch.-praktich. konf. po problemam zashchity naseleniya i territoriy ot chrezvychaynykh situatsiy* [Actual Problems of Civil Protection: 11th Scientific and Practical Conference (International) on Problems of Protection of the Population and Territories from Emergency Situations Proceedings]. N. Novgorod: Vector-TiS. 44–50. Available at: <http://gov.mari.ru/debzn/omgo/46.djvu> (accessed August 7, 2017).
- Porfir'ev, B. N. 1991. *Upravlenie v chrezvychaynykh situatsiyakh. T. 1. Itogi nauki i tekhniki. Problemy bezopasnosti: chrezvychaynye situatsii* [Management in emergency situations. Vol. 1. The results of science and technology. Security concerns: Emergency situations]. Moscow: VINITI. 204 p.
- Vishnyakov, Ya. D., and N. N. Radaev. 2008. *Obshhaya teoriya riskov* [Common theory of risks]. 2nd ed. Moscow: Academy. 368 p.
- Akimov, V. A., V. V. Lesnykh, and N. N. Radaev. 2004. *Riski v prirode, tekhnosfere, obshchestve i ekonomike* [Risks in the nature, technosphere, society, and the economy]. Moscow: Business Express. 352 p.
- Solozhentsev, E. D. 2006. *Stsenarnoe logiko-veroyatnostnoe upravlenie riskom v biznese i tekhnike* [Scenario logic and probabilistic management of risk in business and engineering]. 2nd ed. St. Petersburg: Biznes pressa. 560 p.
- Ryabinin, I. A. 2000. *Nadezhnost' i bezopasnost' strukturalno-slozhnykh sistem* [Reliability and safety of the structural and composite systems]. St. Petersburg: Polytechnique. 248 p.
- Tyrsin, A. N. 2015. O modelirovanii riska v sistemakh kritichnykh infrastruktur [About model operation of risk in the systems of critical infrastructures]. *Economic and Technical Aspects of Safety of Civil Engineering Critical Infrastructures Conference (International) Abstracts*. Ekaterinburg: Ural Federal University. 205–208. Available at: http://elar.ufu.ru/bitstream/10995/33468/1/safety_2015.pdf (accessed August 7, 2017).
- Tyrsin, A. N., and A. A. Surina. 2017. Modelirovanie riska v mnogomernykh stokhasticheskikh sistemakh [Modeling of risk in multidimensional stochastic systems]. *Vestn. Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika* [Bull. Tomsk State University. Management, Computer Facilities, and Informatics] 2(39):65–72.
- Gnedenko, B. V. 2005. *Kurs teorii veroyatnostey* [Course of probability theory]. 8th ed. Moscow: Editorial URSS. 448 p.
- Pena, D., and J. Rodriguez. 2003. Descriptive measures of multivariate scatter and linear dependence. *J. Multivariate Anal.* 85:361–374.
- Kir'yanov, B. F., and M. S. Tokmachev. 2009. *Matematicheskie modeli v zdravookhranении* [Mathematical models in health care]. Veliky Novgorod: NovSU. 279 p.
- Tsinker, M. Yu., D. A. Kir'yakov, and M. R. Kamaltdinov. 2013. Primenenie kompleksnogo indeksa narusheniya zdorov'ya naseleniya dlya otsenki populyatsionnogo zdorov'ya v Permskom krae [The integrated index of health situation of the population to assess population health in the Perm region]. *Izvestiya Samarskogo nauchnogo tsentra RAN* [Proceedings of the Samara Scientific Center of RAS] 15(3(6)):1988–1992.
- Profilaktika khronicheskikh neinfektsionnykh zabolevaniy. Rekomendatsii [Prevention of pre-existing conditions. Recommendations]. Moscow. 128 p. Available at: <http://www.webmed.irkutsk.ru/doc/pdf/prevent.pdf> (accessed August 7, 2017).
- Tyrsin, A. N., O. F. Kalev, D. A. Yashin, and O. V. Lebedeva. 2015. Otsenka sostoyaniya zdorov'ya populyatsii

- na osnove entropiynogo modelirovaniya [Assessment of health status of a population on the basis of entropy modeling]. *Math. Biol. Bioinf.* 10(1):206–219. doi: 10.17537/2015.10.206.
17. Mihaylov, G. A., and A. V. Voytishchek. 2006. *Chislennoe statisticheskoe modelirovanie. Metody Monte-Karlo* [Numerical statistical model operation. Monte-Carlo methods]. Moscow: Akademy. 368 p.
 18. Panteleev, A. V., and T. A. Letova. 2008. *Metody optimizatsii v primerakh i zadachakh* [Optimization methods in examples and tasks]. 3rd ed. Moscow: Higher School. 544 p.

Received August 21, 2017

Contributors

Tyrsin Alexander N. (b. 1961) – Doctor of Science in technology, Head of Department of Applied Mathematics, Ural Federal University named after first President of Russia B. N. Yeltsin, 19 Mira Str., Ekaterinburg 620002, Russian Federation; senior scientist, Institute of Economics, Ural Branch of the Russian Academy of Sciences, 29 Moskovskaya Str., Yekaterinburg 620014, Russian Federation; at2001@yandex.ru

Surina Alfiya A. (b. 1987) – PhD student, Institute of Natural Sciences, South Ural State University, 87 Lenin Ave., Chelyabinsk 454080, Russian Federation; dallila87@mail.ru

АЛГОРИТМ ВИЗУАЛИЗАЦИИ ПЛОСКОГО ЯДРА ВЕРОЯТНОСТНОЙ МЕРЫ*

С. Н. Васильева¹, Ю. С. Кан²

Аннотация: Предлагается алгоритм построения многогранной аппроксимации ядра вероятностной меры для двумерного случайного вектора с независимыми компонентами. Ядро является одним из важных понятий, используемых в алгоритмах решения задач стохастического программирования с вероятностными критериями. Наиболее эффективно ядро применяется в случаях, когда постановки указанных задач имеют свойство линейности по отношению к случайным параметрам. В силу линейности максимум по случайным параметрам определяется путем перебора всех вершин аппроксимирующего многогранника. Предложенный в статье алгоритм основан на построении пересечения конечного числа доверительных полупространств, параметры которых оцениваются методом Монте Карло. Результатом работы предложенного алгоритма является определение множества вершин аппроксимирующего многогранника. Аппроксимация ядра является их выпуклой оболочкой. Приводятся результаты расчетов для ряда типовых непрерывных законов распределения.

Ключевые слова: задача квантильной оптимизации; метод линеаризации; ядро вероятностной меры

DOI: 10.14357/19922264180209

1 Введение

Ядро вероятностной меры заданного уровня α для случайного вектора определяется как пересечение всех выпуклых, замкнутых, α -доверительных множеств в пространстве реализаций этого случайного вектора. Это понятие играет ключевую роль в теоретических аспектах задач стохастического программирования с квантильным критерием качества [1], называемых ниже задачами квантильной оптимизации. Последние с прикладной точки зрения моделируют принятие решений в условиях неопределенности с учетом риска или требований надежности.

Впервые α -ядро было введено в рассмотрение в [2] для гауссовского случая с целью доказательства асимптотической точности доверительного метода решения задач минимизации функции квантили при использовании доверительного множества в виде эллипсоида, являющегося множеством уровня плотности вероятности. Исследованию свойств ядра посвящены работы [3, 4]. Следует также отметить, что α -ядро может использоваться в задачах определения стартовой точки для численных методов оптимизации функции квантили [5, 6].

Квантильный критерий представляет собой α -квантиль распределения некоторой функции потерь, зависящей от вектора оптимизируемой стра-

тегии и случайных параметров. Среди задач квантильной оптимизации можно выделить важный класс задач, в которых функция потерь линейна по случайным параметрам. В этот класс включаются задачи оптимизации портфеля ценных бумаг с учетом риска по квантильному критерию, впервые рассмотренные в [7]. В этом случае задача минимизации квантильного критерия при выполнении некоторых условий регулярности ядра эквивалентна минимаксной задаче для функции потерь [1], где внутренний максимум берется по реализациям случайных параметров, а внешний минимум — по оптимизируемой стратегии.

Таким образом, стохастическая задача квантильной оптимизации равносильна минимаксной со специально задаваемым множеством неопределенности, по которому производится максимизация, и это множество совпадает с α -ядром.

Эта минимаксная задача исследовалась в [8] для случая, когда функция потерь является линейной и по стратегии. В этой же статье предложены общие алгоритмические схемы аппроксимации α -ядра выпуклыми многогранниками, с использованием которых указанная минимаксная задача сводится к задаче линейного программирования с большим числом ограничений. На основе одной из этих схем и разработан предлагаемый ниже алгоритм, реализованный в программном пакете MATLAB.

*Результаты работы получены в рамках выполнения государственного задания Минобрнауки № 2.2461.2017/ПЧ, а также при финансовой поддержке РФФИ (проект 15-08-02833а).

¹Московский авиационный институт (национальный исследовательский университет), sofia_mai@mail.ru

²Московский авиационный институт (национальный исследовательский университет), yu_kan@mail.ru

Во многих математических моделях прикладных задач встречается ситуация, когда случайные параметры являются в некотором смысле малыми. В этих случаях возможно использование метода линеаризации, точное описание которого изложено в [9]. В соответствии с этим методом исходную нелинейную функцию потерь можно линеаризовать по случайным параметрам и получить линейную по случайным параметрам модель. Тем самым задача квантильной оптимизации с нелинейной функцией потерь может быть приближенно аппроксимирована вышеупомянутой минимаксной задачей, в которой роль множества неопределенности играет α -ядро.

К настоящему времени масштабные исследования геометрии α -ядра для типовых законов распределения не проводились. Предлагаемый ниже программно реализованный алгоритм предназначен для того, чтобы частично ликвидировать этот пробел.

В разд. 2 вводятся основные определения и утверждения, описывается рассматриваемая задача квантильной оптимизации, приведены некоторые свойства ядра, продемонстрирована роль ядра вероятностной меры в задачах квантильной оптимизации. В разд. 3 представлено описание алгоритма построения аппроксимации ядра. В разд. 4 представлены результаты расчетов α -ядер для двумерных логнормального и экспоненциального распределений, построенные с использованием пакета MATLAB.

2 Постановка задачи и предварительные результаты

Определение 1. Множество $S_\alpha \subset \mathbb{R}^t$ называется α -доверительным множеством для случайного вектора \mathbf{X} с реализациями $\mathbf{x} \in \mathbb{R}^t$, если $P(\mathbf{X} \in S_\alpha) \geq \alpha$.

Определение 2. α -ядро может быть определено как

$$K_\alpha = \bigcap_{S \in E_\alpha} S,$$

где E_α — семейство всех выпуклых замкнутых α -доверительных множеств.

Рассматривается задача о построении границы α -ядра в случае, когда вектор X имеет размерность $m = 2$ и его компоненты независимы с абсолютно непрерывным заданным законом распределения.

Поскольку для большинства распределений не удастся аналитически построить границу ядра, то вместо нахождения точной границы будем искать ее аппроксимацию.

Функцию $\Phi(\mathbf{X}, \mathbf{u})$, зависящую от вектора случайных параметров \mathbf{X} и вектора стратегии $\mathbf{u} \in U \subset \mathbb{R}^n$, будем называть функцией потерь.

Введем необходимые определения и утверждения из [1].

Определение 3. Функция вероятности для функции потерь $\Phi(\mathbf{X}, \mathbf{u})$ имеет вид:

$$P_\varphi(\mathbf{u}) \triangleq P\{\Phi(\mathbf{X}, \mathbf{u}) \leq \varphi\},$$

где P — вероятность; φ — допустимый уровень потерь.

Определение 4. Функция квантили для функции потерь $\Phi(\mathbf{X}, \mathbf{u})$ определяется следующим образом:

$$\varphi_\alpha(\mathbf{u}) \triangleq \min\{\varphi : P_\varphi(\mathbf{u}) \geq \alpha\} \triangleq [\Phi(\mathbf{X}, \mathbf{u})]_\alpha.$$

Задача квантильной оптимизации имеет вид:

$$[\Phi(\mathbf{X}, \mathbf{u})]_\alpha \rightarrow \min_{\mathbf{u} \in U}. \quad (1)$$

α -ядро допускает следующее представление:

$$K_\alpha = \bigcap_{\|c\|=1} \{z \in \mathbb{R}^t : c^T z \leq [c^T \mathbf{X}]_\alpha\},$$

где c — вектор внешней нормали к границе α -доверительного полупространства; $\|\cdot\|$ — евклидова норма. Таким образом, α -ядро является пересечением всех замкнутых α -доверительных полупространств.

Определим аппроксимацию α -ядра как пересечение конечного числа замкнутых α -доверительных полуплоскостей:

$$V_{\alpha N} = \bigcap_{j=1}^N \{\mathbf{x} : c_j^T \mathbf{x} \leq [c_j^T \mathbf{X}]_\alpha\}. \quad (2)$$

В случае использования точного значения квантили $[c_j^T \mathbf{X}]_\alpha$ полученный многогранник является внешней аппроксимацией α -ядра. В общем случае явно определить эту величину не удастся, поэтому для ее вычисления ниже, в разд. 5, будет использована выборочная оценка квантили.

Рассмотрим функцию потерь вида:

$$\Phi(\mathbf{X}, \mathbf{u}) = a(\mathbf{u}) + b^T(\mathbf{u})\mathbf{X}. \quad (3)$$

Введем следующее

Определение 5 [1]. α -ядро K_α называется регулярным, если всякое замкнутое полупространство, содержащее это ядро, автоматически является α -доверительным.

Теорема 1 [1]. Если случайный вектор \mathbf{X} имеет регулярное α -ядро K_α , то для любого детерминированного вектора $\mathbf{u} \in R^n$ справедливо

$$[\mathbf{u}^T \mathbf{X}]_\alpha = \max_{\mathbf{x} \in K_\alpha} \mathbf{u}^T \mathbf{x}.$$

В соответствии с теоремой 1 задача квантильной оптимизации (1) для функции (3) может быть сведена к задаче:

$$a(\mathbf{u}) + \max_{\mathbf{x} \in K_\alpha} (b^T(\mathbf{u})\mathbf{x}) \rightarrow \min_{\mathbf{u} \in U}. \quad (4)$$

После замены многогранника на его аппроксимацию получаем:

$$\varphi_\alpha^N(u) = a(\mathbf{u}) + \max_{j=\overline{1, N}} (b^T(\mathbf{u})\mathbf{v}_j) \rightarrow \min_{\mathbf{u} \in U},$$

где v_j — вершины $V_{\alpha N}$, $j = \overline{1, N}$. Функция $\varphi_\alpha^N(u)$ представляет собой оценку исходного квантильного критерия, при нахождении которой исходный максимум по случайному вектору из множества K_α заменен на максимум по его аппроксимации $V_{\alpha N}$.

Максимум по вершинам выпуклого многогранного множества может быть найден как максимум из значений функции в вершинах этого множества $\max_{j=\overline{1, N}}$.

Сходимость

$$\min_{u \in U} \varphi_\alpha^N(u) - \min_{u \in U} \varphi_\alpha(u) \xrightarrow{N \rightarrow \infty} 0$$

доказана в [8] для случая, когда $a(\mathbf{u})$ и $b(\mathbf{u})$ линейны по \mathbf{u} и множество U компактно.

Задача (1) для функции потерь, нелинейно зависящей от вектора случайных параметров, является сложно разрешимой. В [9] доказана справедливость использования линейаризованной по случайным параметрам функции потерь в случае малых случайных параметров. В качестве вектора малых случайных параметров используется вектор \mathbf{X}^μ , составленный из покоординатных произведений вектора малых параметров μ и случайного вектора \mathbf{X} , т.е. $\mathbf{X}^\mu = M\mathbf{X}$, где $M = \text{diag}(\mu_1, \mu_2, \dots, \mu_t)$. Функция потерь имеет вид $\Phi(\mathbf{X}^\mu, \mathbf{u})$. Разложение этой функции потерь в ряд Тейлора по вектору \mathbf{X}^μ в окрестности нуля справедливо, если функция является дважды непрерывно дифференцируемой в окрестности нуля, а также непрерывна по $\mathbf{u} \in U$ вместе со своими частными производными по \mathbf{X}^μ до второго порядка включительно. Разложение может быть представлено следующим выражением:

$$\Phi(\mathbf{X}^\mu, \mathbf{u}) = a(\mathbf{u}) + b^T(\mathbf{u})\mathbf{X}^\mu + r_1(\mathbf{X}^\mu, \mathbf{u}),$$

где $a(\mathbf{u}) = \Phi(\mathbf{0}, \mathbf{u})$ — значение функции потерь в нуле; $b(\mathbf{u}) = (\partial\Phi(\mathbf{x}^\mu, \mathbf{u})/\partial\mathbf{x}^\mu)|_{\mathbf{x}^\mu=0}$ — вектор значений покоординатных производных в нуле; $r_1(\mathbf{X}^\mu, \mathbf{u}) = (X_1^\mu \partial_1 \Phi(\theta\mathbf{X}^\mu, \mathbf{u}) + \dots + X_t^\mu \partial_t \Phi(\theta\mathbf{X}^\mu, \mathbf{u}))/2$ — остаточный член ряда Тейлора в форме Лагранжа, $\theta \in (0, 1)$; $\partial_i \Phi(\mathbf{X}^\mu, \mathbf{u}) = \partial\Phi(\mathbf{X}^\mu, \mathbf{u})/\partial X_i^\mu$ — частная производная функции $\Phi(\mathbf{X}^\mu, \mathbf{u})$ по i -й координате случайного вектора \mathbf{X}^μ .

Под линейаризованной моделью подразумевается линейная часть разложения исходной функции в ряд Тейлора по случайным параметрам в окрестности нуля. Линейаризованную модель можно представить в виде (3):

$$\Phi_l(\mathbf{X}^\mu, \mathbf{u}) = a(\mathbf{u}) + b^T(\mathbf{u})\mathbf{X}^\mu.$$

Такая замена приводит к следующему результату:

$$\min_{u \in U} [\Phi(\mathbf{X}^\mu, \mathbf{u})]_\alpha = \min_{u \in U} [\Phi_l(\mathbf{X}^\mu, \mathbf{u})]_\alpha + O(\|\mu\|^2). \quad (5)$$

Результат (5) обоснован в работе [9] для случая, когда множество U является компактом, а носитель случайного вектора X ограничен, и для случая, когда множество U конечно, а носитель вектора случайных параметров X неограничен.

Учитывая (4), заключаем, что с точностью до $O(\|\mu\|^2)$ задача

$$[\Phi(\mathbf{X}, \mathbf{u})]_\alpha \rightarrow \min_{u \in U}$$

может быть аппроксимирована минимаксной задачей

$$a(\mathbf{u}) + \max_{\mathbf{x} \in K_\alpha} b^T(\mathbf{u})M\mathbf{x} \rightarrow \min_{u \in U},$$

если, конечно, ядро K_α регулярно и справедливо соотношение (5).

В данной статье рассматривается задача построения сколь угодно точной аппроксимации ядра многогранником.

3 Алгоритм построения аппроксимации α -ядра

Для построения аппроксимации ядра сгенерируем выборку X_1, X_2, \dots, X_k случайного вектора \mathbf{X} . Рассмотрим вспомогательную линейную функцию потерь вида

$$Z(c, \mathbf{X}) = c^T \mathbf{X},$$

где c — единичный вектор нормали, т.е. $\|c\| = 1$. С учетом этого обозначения аппроксимация ядра (2) может быть записана в виде:

$$V_{\alpha N} = \bigcap_{j=1}^N \{ \mathbf{x} : c_j^T \mathbf{x} \leq [Z(c_j, \mathbf{X})]_\alpha \}.$$

Функцию квантили $[Z(c, \mathbf{X})]_\alpha$ будем оценивать методом Монте Карло.

Используя сгенерированную ранее выборку случайного вектора \mathbf{X} , построим выборку значений функции потерь $Z(c_j, \mathbf{X})$ по формуле:

$$Z_i^j = c_j^T \mathbf{X}_i, \quad i = \overline{1, k}.$$

Для каждого номера j построим вариационный ряд этой выборки:

$$Z_{(1)}^j \leq Z_{(2)}^j \leq \dots \leq Z_{(k)}^j.$$

Выборочная оценка \widehat{Z}_α^j функции квантили $[Z(c_j, X)]_\alpha$ вычисляется по формуле из [1]:

$$\widehat{Z}_\alpha^j = \begin{cases} Z_{(\alpha k)}^j, & \alpha k \in \mathbb{N}; \\ Z_{([\alpha k]+1)}^j, & \alpha k \notin \mathbb{N}, \end{cases}$$

где $[\cdot]$ — целая часть от числа.

Точность выборочной оценки обоснована следующей теоремой [10].

Теорема 2. Если $\alpha \in (0, 1)$ и случайная величина Φ имеет плотность вероятности $p(\varphi)$, непрерывную в некоторой окрестности точки φ_α , причем $p(\varphi_\alpha) > 0$, то

$$\widehat{\Phi}_\alpha - \varphi_\alpha = \frac{\widehat{P}(\varphi_\alpha) - \alpha}{p(\varphi_\alpha)} + o_p(k^{-1/2}),$$

где $\widehat{P}(\varphi_\alpha)$ — значение выборочной оценки вероятности в точке φ_α ; φ_α — точное значение квантили; $k^{1/2} o_p(k^{-1/2}) \rightarrow 0$ по вероятности при $k \rightarrow \infty$.

Как уже отмечалось выше, аппроксимация α -ядра строится путем пересечения конечного числа полуплоскостей, определяемых своими векторами нормалей. В работе [8] предложен алгоритм, который позволяет построить густую сеть векторов c_j , $j = \overline{1, N}$. При этом векторы из этого множества в двумерном случае образуют неравномерную сетку на единичной окружности с центром в начале координат.

Этот недостаток алгоритма из [8] можно устранить в случае двумерного распределения случайного вектора. Новый алгоритм включает в себя следующую последовательность шагов.

1. Нанесем равномерную сетку из заданного числа N точек c_j , $j = \overline{1, N}$, на поверхность единичной окружности.
2. Сгенерируем выборку заданного объема n для рассматриваемого двумерного распределения случайного вектора \mathbf{X} .
3. Для каждого вектора нормали c_j , $j = \overline{1, N}$, сформируем вариационный ряд $Z_{(1)}^j, Z_{(2)}^j, \dots, Z_{(k)}^j$ и построим выборочную оценку квантили $[Z(c_j, \mathbf{X})]_\alpha$ в соответствии с процедурой, описанной в начале данного раздела.
4. Найдем точки пересечения границ всех пар рассматриваемых доверительных полуплоскостей $\{\mathbf{x} : c_j^T \mathbf{x} \leq \widehat{Z}_\alpha^j\}$.

После этого производится проверка их принадлежности остальным полуплоскостям. Точки, удовлетворяющие всем ограничениям, являются вершинами аппроксимирующего многогранника. За счет использования выборочной оценки квантили, не совпадающей в общем случае с точным значением квантили, число вершин многогранника может оказаться меньше N . Вершины аппроксимирующего многогранника отображаются на итоговом рисунке, соседние вершины соединяются прямойлинейным отрезком.

4 Программная реализация алгоритма визуализации α -ядра

Для реализации алгоритма была использована программная среда MATLAB. Для удобства использования программы с помощью встроенного инструмента GUIDE был создан пользовательский интерфейс программы, вид которого представлен на рис. 1.

В окне программы пользователь имеет возможность выбрать или ввести следующие параметры: вид распределения (среди предложенных); параметры (параметр) распределения; объем выборки; число доверительных полуплоскостей; число аппроксимаций; параметр α для каждой из аппроксимаций. В меню выбора доступны для выбора следующие распределения: нормальное; равномерное; логнормальное; экспоненциальное; распределение Коши. На одном рисунке может быть построено не более трех аппроксимаций. Виды линий границы указаны в легенде. Для удобства все необходимые данные изначально заполнены. Для начала работы алгоритма программы необходимо нажать кнопку «Build approximation», после чего справа от меню появится рисунок с выбранным пользователем числом аппроксимаций. После этого параметры в области меню могут быть изменены. После нажатия кнопки «Build approximation» будет показан рисунок, построенный для новых параметров.

На рис. 1, б показано окно программы. На рисунке построены три аппроксимации ядер для различных заданных α . В качестве распределения выбрано экспоненциальное распределение с параметром $\lambda = 1$.

Для генерации выборки распределения используются стандартные функции из пакета MATLAB Statistics Toolbox.

Для удобства была создана функция Kernel, которая по заданным параметрам находит аппроксимацию и возвращает массив значений. На вход необходимо подать следующие данные: массив ре-

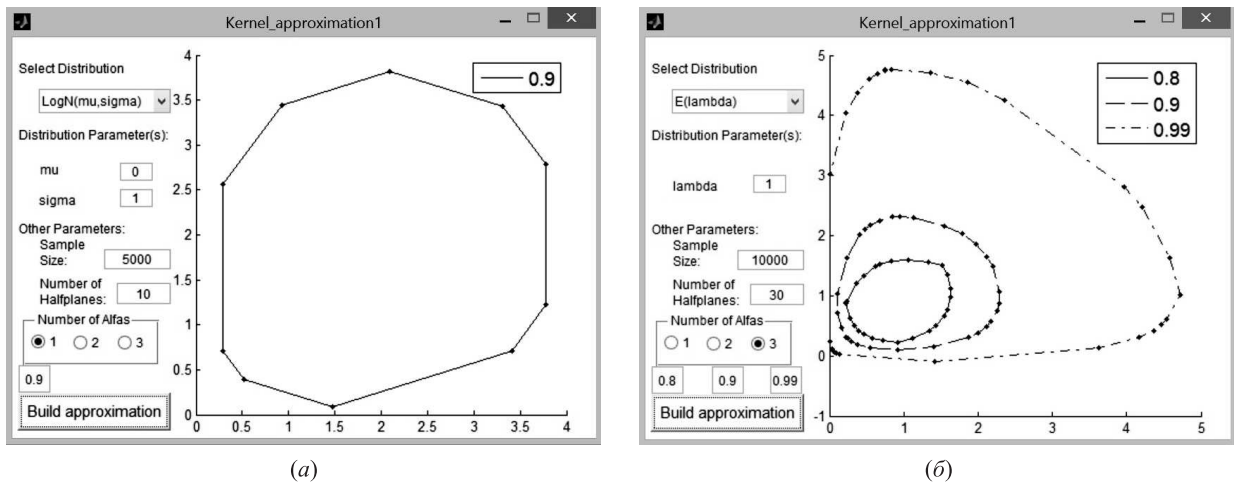


Рис. 1 Пример работы программы: (а) логнормальное распределение, $\alpha = 0,9$; (б) экспоненциальное распределение при $\alpha = 0,8, 0,9$ и $0,99$

лизаций случайной величины, уровень α , объем выборки, число полуплоскостей.

После генерации точек всех ядер, используется стандартная функция plot, которая по точкам строит границу ядра.

5 Результаты расчетов

В данном разделе представлены результаты для экспоненциального и логнормального распределений компонент случайного вектора X . На рис. 2–5 точками обозначены вершины многогранников, аппроксимирующих α -ядро. Как было отмечено выше, число вершин не превосходит N .

С помощью описанной выше программы построим α -ядра для объема выборки 10^6 для $\alpha = 0,9$. Результаты работы алгоритма построения аппроксимации α -ядра при $N = 8, 16$ и 70 для случая, когда компоненты случайного вектора X независимы и $X_i \sim \log N(0, 1), i = 1, 2$, показаны на рис. 2.

Эмпирически установлено, что для выборок порядка $k = 10^6$ и выше выборочная оценка достаточно точна, поэтому уменьшения числа вершин аппроксимирующего многогранника не наблюдается.

Аналогичные результаты построения аппроксимации α -ядра при $N = 8, 16$ и 70 для случая, когда компоненты случайного вектора X независимы и $X_i \sim E(1), i = 1, 2$, показаны на рис. 3.

Для экспоненциального распределения при объеме выборки порядка $k = 10^6$ и выше выборочная оценка квантили также достаточно точна.

Разработанный программный модуль позволяет строить на одном рисунке одновременно две или

три аппроксимации для различных значений α . На рис. 4 продемонстрированы границы ядер для $\alpha \in \{0,7; 0,9; 0,99\}$.

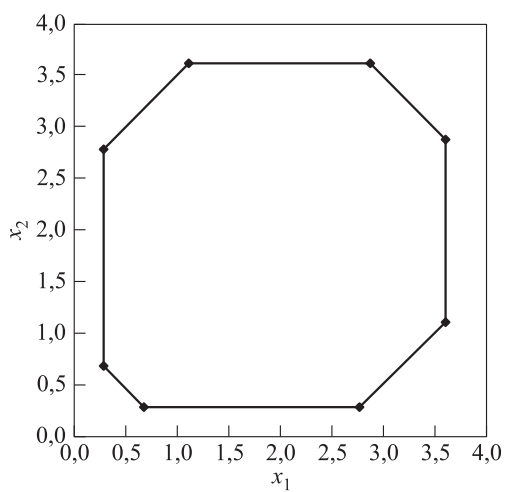
На рис. 5 показано, что при уменьшении объема выборки в случае логнормального распределения при $\alpha = 0,9$ число вершин аппроксимирующего многогранника уменьшается.

В частности, при $k = 10^4$ число вершин аппроксимирующего многогранника уменьшилось на 20.

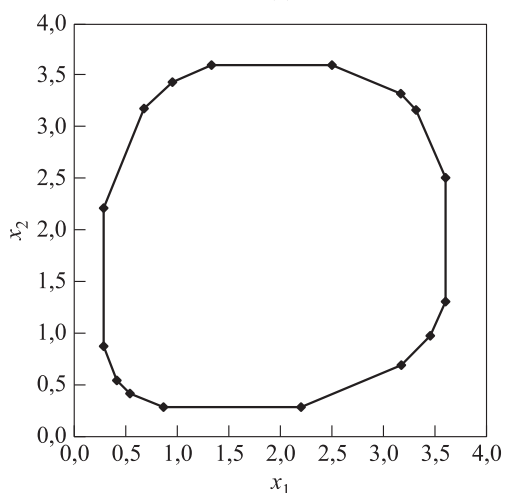
При использовании выборки объема $k = 10^3$ снижение точности выборочной оценки квантили приводит к тому, что остается всего лишь 38 вершин многогранника из 70 возможных.

За счет того, что процедура построения выборочной оценки квантили производится для всех единичных векторов нормали и пересечение ограничений, задающих доверительные полуплоскости, производится попарно, приходится производить значительную часть расчетов и операций, не влияющих на конечный результат. В случае, когда объем выборки увеличить невозможно или пользователю необходимо увеличить скорость работы алгоритма, рекомендуется уменьшить число пересекаемых полупространств. Число вершин многогранника при этом уменьшится, но незначительно. Если пользователю необходимо увеличить точность аппроксимации, то необходимо увеличить параметр n для повышения точности выборочной оценки квантили.

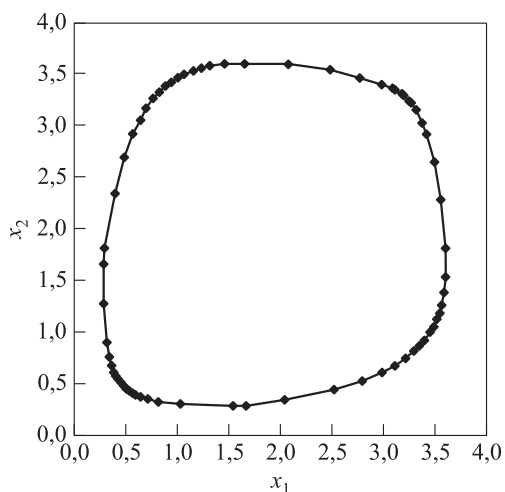
Соотношения параметров α, N и k не могут быть заданы априори, поскольку они зависят как от вида распределения, так и от параметров распределения. Число вершин многогранника является случайной величиной. Для рассматриваемого распределения при фиксированном α могут быть по-



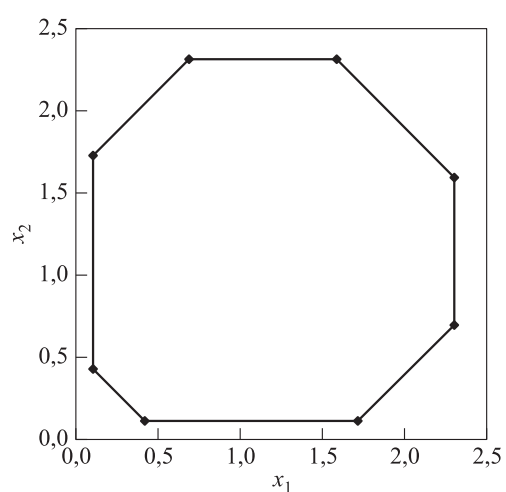
(a)



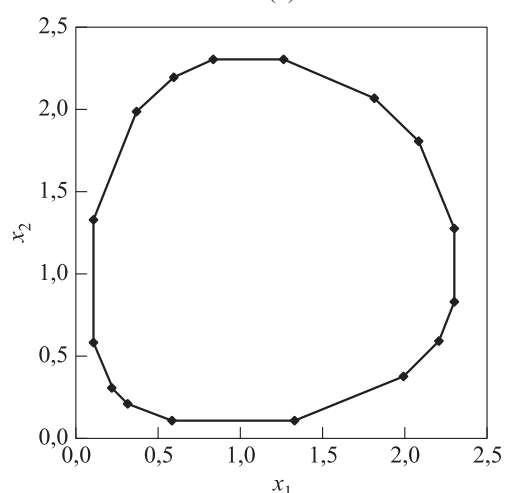
(б)



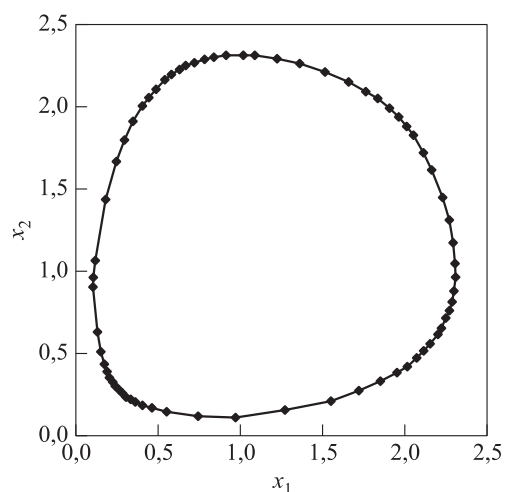
(в)



(a)



(б)



(в)

Рис. 2 Логнормальное распределение, $k = 10^6$, $\alpha = 0,9$:
(a) $N = 8$; (б) 16; (в) $N = 70$

Рис. 3 Экспоненциальное распределение, $k = 10^6$,
 $\alpha = 0,9$: (a) $N = 8$; (б) 16; (в) $N = 70$

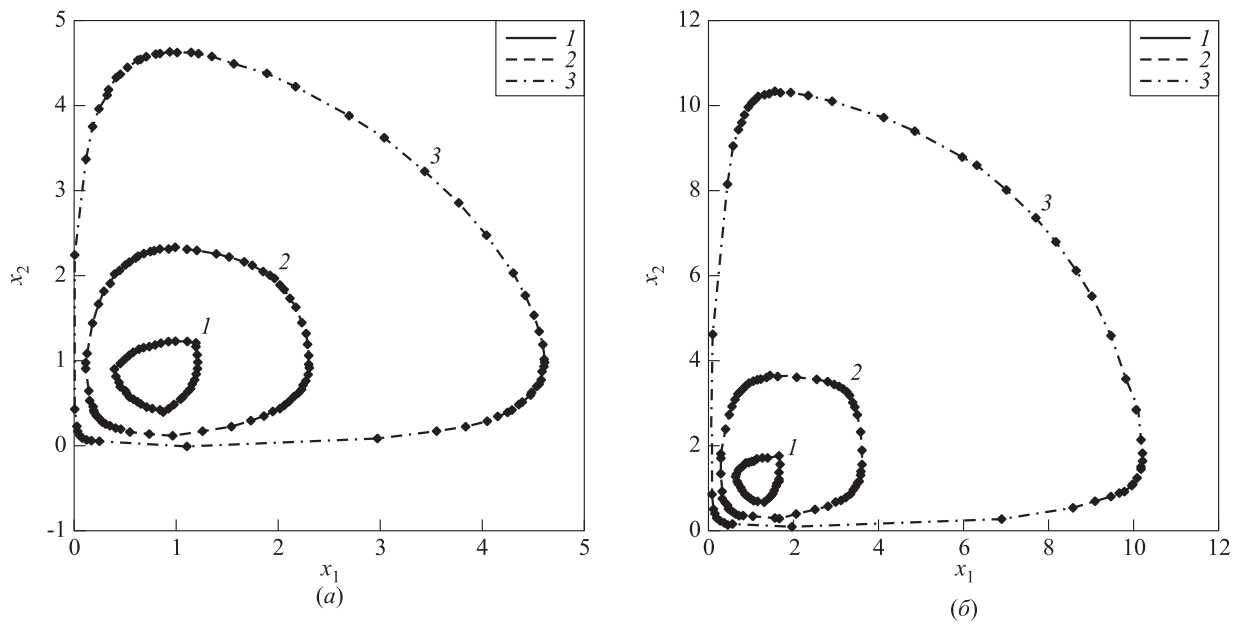


Рис. 4 Экспоненциальное (а) и логнормальное (б) распределения, 70 точек: 1 — $\alpha = 0,7$; 2 — $0,9$; 3 — $\alpha = 0,99$

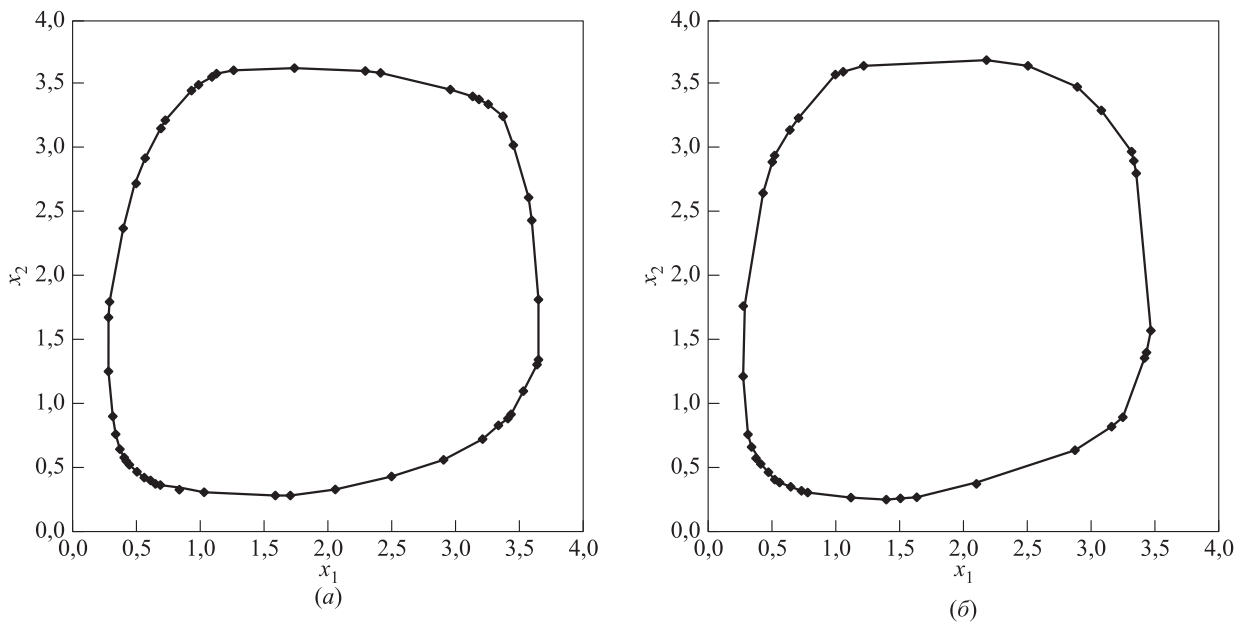


Рис. 5 Логнормальное распределение, 70 точек, $\alpha = 0,9$: (а) выборка 10^4 ; (б) выборка 10^3

строены выборочные оценки при фиксированных параметрах N и k для числа вершин (например, выборочного среднего, выборочной дисперсии и т. д.) в зависимости от параметров N и k .

6 Заключение

Предложенный алгоритм позволяет более точно (по сравнению с методом, предложенным в [8]) аппроксимировать α -ядро. Этот эффект достигается

за счет того, что используется равномерная сеть точек на окружности, задающих векторы нормалей.

На рисунках показано, как уменьшается число точек многогранника при уменьшении объема выборки распределения случайного вектора.

При решении практических задач при заданном α параметры N и k необходимо выбирать так, чтобы число вершин многогранника было как можно ближе к N .

Помимо графических иллюстраций разработанный программный модуль позволяет построить систему линейных ограничений, задающих аппроксимацию α -ядра, которую можно использовать в алгоритмах решения задач с квантильным критерием качества. Кроме того, определяются координаты вершин многогранника, удовлетворяющие данной системе ограничений.

Литература

1. Кибзун А. И., Кан Ю. С. Задачи стохастического программирования с вероятностными критериями. — М.: Физматлит, 2009. 372 с.
2. Малышев В. В., Кибзун А. И. Анализ и синтез высокоточного управления летательными аппаратами. — М.: Машиностроение, 1987. 303 с.
3. Кан Ю. С., Русяев А. В. Задача квантильной минимизации с билинейной функцией потерь // Автоматика и телемеханика, 1998. № 7. С. 67–75.
4. Кан Ю. С., Суринов Р. Т. О неравенстве треугольника для критерия VaR // Моделирование и анализ безопасности и риска в сложных системах: Тр. Междунар. научной школы МАБР-2004. — СПб.: СПбГУАП, 2004. С. 243–248.
5. Kibzun A. I., Kurbakovskiy V. Yu. Guaranteeing approach to solving quantile optimization problem // Ann. Oper. Res., 1991. Vol. 30. P. 81–93.
6. Кибзун А. И., Курбаковский В. Ю. Численные алгоритмы квантильной оптимизации и их применение к решению задач с вероятностными ограничениями // Изв. РАН, Техническая кибернетика, 1992. № 1. С. 75–81.
7. Van Moesèke P. Stochastic linear programming // Yale Econ. Essays, 1965. Vol. 5. P. 197–253.
8. Васильева С. Н., Кан Ю. С. Метод решения задачи квантильной оптимизации с билинейной функцией потерь // Автоматика и телемеханика, 2015. № 9. С. 83–101.
9. Васильева С. Н., Кан Ю. С. Метод линеаризации для решения задачи квантильной оптимизации с функцией потерь, зависящей от вектора малых случайных параметров // Автоматика и телемеханика, 2017. № 7. С. 95–109.
10. Bahadur R. R. A note on quantiles in large samples // Ann. Math. Stat., 1996. Vol. 37. P. 577–580.

Поступила в редакцию 26.04.17

A VISUALIZATION ALGORITHM FOR THE PLANE PROBABILITY MEASURE KERNEL

S. N. Vasil'eva and Yu. S. Kan

Moscow Aviation Institute (National Research University), 4 Volokolamskoe Shosse, Moscow 125993, Russian Federation

Abstract: The authors propose an algorithm for constructing a probability measure kernel polyhedral approximation for a two-dimensional random vector with independent components. The kernel is one of the important concepts used in algorithms for solving stochastic programming problems with probabilistic criteria. The kernel is most effectively used in cases when the statements of the indicated problems have the property of linearity with respect to random parameters. Because of linearity, the maximum in random parameters is determined by searching all vertices of the approximating polyhedron. The authors propose an algorithm for constructing a polyhedral approximation of the kernel of a probability measure for a two-dimensional random vector with independent components. The algorithm is based on construction of the intersection of a finite number of confidence half-spaces, the parameters of which are estimated by the Monte-Carlo method. The result of the proposed algorithm is the definition of the set of vertices of the approximating polyhedron. Approximation of the nucleus is their convex hull. The results of calculations for a number of typical continuous distribution laws are presented.

Keywords: quantile optimization problem; linearization method; probability measure kernel

DOI: 10.14357/19922264180209

Acknowledgments

The work was supported by the Russian Ministry of Education and Science (Project No. 2.2461.2017/PCh) and by the Russian Foundation for Basic Research (grant No. 15-08-02833a).

References

1. Kibzun, A. I., and Yu. S. Kan. 2009. *Zadachi stokhasticheskogo programirovaniya s veroyatnostnymi kriteriyami* [Stochastic programming problems with probabilistic criteria]. Moscow: Fizmatlit. 372 p.
2. Malyshev, V. V., and A. I. Kibzun. 1987. *Analiz i sintez vysokotochnogo upravleniya letatel'nymi apparatami* [Analysis and synthesis of high-precision aircraft control]. Moscow: Mashinostroenie. 303 p.
3. Kan, Yu. S., and A. V. Rusyaev. 1998. Quantile minimization with bilinear loss function. *Automat. Rem. Contr.* 59(7(1)):960–966.
4. Kan, Yu. S., and R. T. Surinov. 2004. O neravenstve treugol'nika dlya kriteriya VaR [On the triangle inequality for the VAR criterion]. *Modeling and Analysis of Safety and Risk in Complex Systems: Scientific School (International) MASR-2004 Proceedings*. St. Petersburg: SPbGUAP. 243–248.
5. Kibzun, A. I., and V. Yu. Kurbakovskiy. 1991. Guaranteeing approach to solving quantile optimization problem. *Ann. Oper. Res.* 30:81–93.
6. Kibzun, A. I., and V. Yu. Kurbakovskij. 1992. Chislennye algoritmy kvantil'noy optimizatsii i ikh primeneniye k resheniyu zadach s veroyatnostnymi ogranicheniyami [Numerical algorithms for quantitative optimization and their application to solving problems with probability constraints]. *J. Comput. Sys. Sc. Int.* 1:75–81.
7. Van Moesèke, P. 1965. Stochastic linear programming. *Yale Econ. Essays* 5:197–253.
8. Vasil'eva, S. N., and Yu. S. Kan. 2015. A method for solving quantile optimization problems with a bilinear loss function. *Automat. Rem. Contr.* 76(9):1582–1597.
9. Vasil'eva, S. N., and Yu. S. Kan. 2017. Linearization method for solving quantile optimization problems with loss function depending on a vector of small random parameters. *Automat. Rem. Contr.* 78(7):1248–1260.
10. Bahadur, R. R. 1996. A note on quantiles in large samples. *Ann. Math. Stat.* 37:577–580.

Received April 26, 2017

Contributors

Vasil'eva Sofia N. (b. 1993) — PhD student, Moscow Aviation Institute (National Research University), 4 Volokolamskoe Shosse, Moscow 125993, Russian Federation; sofia_mai@mail.ru

Kan Yuri S. (b. 1960) — Doctor of Science in physics and mathematics, Professor, Moscow Aviation Institute (National Research University), 4 Volokolamskoe Shosse, Moscow 125993, Russian Federation; yu_kan@mail.ru

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ОПТИМАЛЬНОЙ ТРИАНГУЛЯЦИИ

А. А. Батенков¹, Ю. А. Маньяков², А. В. Гасилов³, О. А. Яковлев⁴

Аннотация: Формализована постановка задачи синтеза оптимальной плоской выпуклой триангуляции. Данная задача относится к вопросам приложений информатики и весьма актуальна в таких областях, как компьютерная графика и геоинформационные системы. Математическая модель представлена в трех вариантах: экстремальная задача с бесконечным числом ограничений, минимаксная задача со связанными переменными и экстремальная задача с дополнительным ограничением на пересечения отрезков триангуляции, но конечным числом всех ограничений. Путем введения идемпотентных ограничений на булевы переменные исходная целочисленная задача погружена в общую задачу математического программирования на континуальном множестве решений. Произведен сравнительный анализ решений, получаемых жадным алгоритмом на основе представленной математической модели и алгоритмом триангуляции Делоне.

Ключевые слова: математическая модель; триангуляция; триангуляция Делоне

DOI: 10.14357/19922264180210

1 Введение

Задача построения триангуляции весьма востребована в таких разделах информатики, как компьютерная графика, распознавание образов и геоинформационные системы, для моделирования поверхностей и решения пространственных задач, задач навигации автономных роботизированных систем. Кроме того, триангуляция может быть использована в алгоритмах трехмерной реконструкции [1].

Отсутствие формализованной постановки задачи синтеза триангуляции в виде математической модели привело к появлению множества эвристических приближенных алгоритмов [2, 3] по различным критериям. Этому также способствовало подозрение на NP-полноту задачи поиска оптимальной триангуляции.

Наиболее применима в настоящее время триангуляция, предложенная советским математиком Б. Н. Делоне, которая имеет вычислительную сложность порядка $O(N \log N)$ [2, 4] (N — число точек триангуляции), но слабый критерий оптимизации [5]. Прежде всего это обусловлено ориентированностью триангуляции Делоне на алгоритмический аспект решаемой задачи, а никак не на поиск оптимальной триангуляции, под которой понимается триангуляция с минимальной суммой

длин ребер среди всех возможных триангуляций, построенных на тех же исходных точках [6]. Также необходимо отметить, что алгоритмы триангуляции Делоне лишь улучшают некоторую заданную первичную триангуляцию, поиск которой требует дополнительных вычислительных затрат.

2 Идентификация пересечения отрезков на множестве прямых

Так как под триангуляцией понимается планарный граф, все внутренние области которого являются треугольниками, то центральное место в математической модели должна занимать идентификация пересечений отрезков прямых, входящих в триангуляцию. Известные алгоритмы обнаружения пересечения отрезков [7] используют их параметрическое задание, что затрудняет использование этих алгоритмов при формулировке экстремальных задач, поэтому необходим функционал, желательнее с булевым множеством области значений [8].

С этой целью используем оконную функцию в виде:

$$\text{Ok}(z, \text{sn}, \text{sv}) = (\Phi(z - \text{sn} - \text{sign}(\text{sv} - \text{sn})\text{zap}) - \Phi(z - \text{sv} + \text{sign}(\text{sv} - \text{sn})\text{zap})) \text{sign}(\text{sv} - \text{sn}),$$

¹ Орловский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, batenkov1957@mail.ru

² Орловский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, maniakov_yuri@mail.ru

³ Орловский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, gasilov.av@yandex.ru

⁴ Орловский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, maucra@gmail.com

где sn и sv — верхняя и нижняя границы окна; $\Phi(x)$ — функция Хевисайда [9]; $sign(x)$ — функция знака [10]; zar — вспомогательная константа.

Сформируем переопределенную систему линейных уравнений с двумя неизвестными:

$$a_k x + b_k y = c_k, \quad k = \overline{1, D}, \quad (1)$$

где

$$a_k = \frac{1}{x_j - x_i}; \quad b_k = -\frac{1}{y_j - y_i};$$

$$c_k = \frac{x_i}{x_j - x_i} - \frac{y_i}{y_j - y_i}; \quad i, j = \overline{1, N};$$

$$AZ = B;$$

D — число прямых, соединяющих каждую пару точек триангуляции: $D = N(N - 1)/2$; A — матрица коэффициентов размерности $D \times 2$; $Z^T = (x, y)$ — вектор неизвестных размерности 2×1 ; B — вектор свободных членов размерности $D \times 1$.

Введем вектор неизвестных Y размерности $D \times 1$, определяющий активные уравнения системы (1), если $Y \neq 0$, и свяжем их с (1):

$$\text{diag}(Y)AZ = \text{diag}(Y)B, \quad (2)$$

где $\text{diag}(Y)$ — диагональная матрица с вектором \vec{Y} на главной диагонали.

Решим систему (2) методом наименьших квадратов:

$$Z = (M^T M)^{-1} M^T C,$$

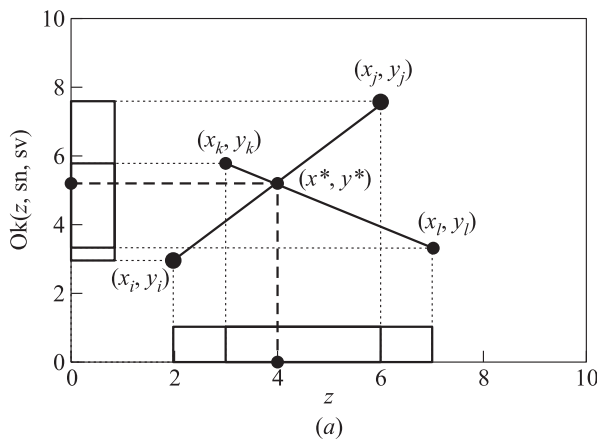
где $M = \text{diag}(Y)A$; $C = \text{diag}(Y)B$.

Анализ рис. 1 приводит к выводу, что точка пересечения отрезков существует, когда равны единице оконные функции:

$$(x_i, y_i), (x_j, y_j); (x_k, y_k), (x_l, y_l); \quad i, j, k, l \in \overline{1, N};$$

$$\text{OK1}(x^*, y^*)_R = \text{Ok}(x^*, x_i, x_j) \cdot \text{Ok}(y^*, y_i, y_j);$$

$$R \in \overline{1, D};$$



$$\text{OK2}(x^*, y^*)_S = \text{Ok}(x^*, x_k, x_l) \cdot \text{Ok}(y^*, y_k, y_l);$$

$$S \in \overline{1, D}.$$

В результате получаем матрицу идентификации пересечения пар отрезков в зависимости от вектора \vec{Y} при известных координатах точек триангуляции на плоскости:

$$\text{Cp}(Y) =$$

$$= \left\{ \left(\begin{array}{l} (\text{OK1}((M^T M)^{-1} M^T C) \bullet Y) \\ (\text{OK2}((M^T M)^{-1} M^T C) \bullet Y)^T \end{array} \right) \bullet \text{EDW} \right\}, \quad (3)$$

где EDW — верхняя треугольная матрица размерности $D \times D$:

$$\text{EDW} = \begin{bmatrix} 0 & 1 & 1 & \dots & 1 \\ 0 & 0 & 1 & \dots & 1 \\ 0 & 0 & 0 & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix};$$

(\bullet) — произведение Адамара [11].

3 Постановка задачи оптимальной триангуляции

Использование функционала пересечения отрезков на основе (3) порождает множество постановок задач оптимальной триангуляции.

Здесь приведем перспективную математическую модель оптимальной триангуляции с дополнительным ограничением на пересечения отрезков триангуляции:

$$W^T X \Rightarrow \min_{X, Y} \quad (4)$$

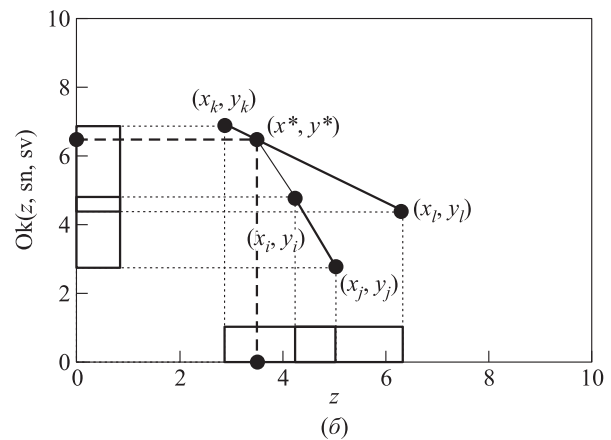


Рис. 1 Пересекающиеся (а) и непересекающиеся (б) отрезки

при ограничениях

$$\text{Ed}^T X = 3(N - 1) - N_c; \quad (5)$$

$$X^T P X = 0; \quad (6)$$

$$X_i = 0 \text{ или } 1, \quad (7)$$

где $W^T = (w_1, w_2, \dots, w_D)$ — вектор длин отрезков между парами исходных точек триангуляции; $\text{Ed}_R = 1, R = \overline{1, D}$ — единичный вектор; N — число точек триангуляции; N_c — число точек выпуклой оболочки.

Целевая функция (4) обеспечивает минимальную суммарную длину ребер триангуляции. Ограничение (5) сформировано на основе леммы о числе ребер триангуляции [6]. Условие (7) определяет булевый характер искомым переменных X . В такой постановке введено дополнительное ограничение (6) на пересечения ребер триангуляции с помощью предварительно вычисленной матрицы $[P]$, формирующей множество пар пересекающихся отрезков для исходного множества точек триангуляции. Матрица $[P]$ является верхней треугольной булевой матрицей размерности $D \times D$ вида:

$$P = \begin{matrix} & 1 \dots l \dots m \dots n \dots D \\ 1 & \left[\begin{array}{cccccc} 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 & / 1 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & \\ i & 0 & \dots & 1 & \dots & 0 & \dots & 0 & \dots & 0 & / 1 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & \\ j & 0 & \dots & 0 & \dots & 1 & \dots & 0 & \dots & 0 & / 1 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & \\ k & 0 & \dots & 0 & \dots & 0 & \dots & 1 & \dots & 0 & / 1 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & \\ D & 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 1 & / 0 \end{array} \right]. \end{matrix}$$

Такая матрица идентифицирует пересечение (j -го, m -го) и (k -го, n -го) отрезков.

Для формирования матрицы $[P]$ применим (3). Тогда RS -й элемент определится как

$$P_{RS} = \text{Cp}(Y^{RS}),$$

где Y^{RS} — вектор с компонентами $Y_R^{RS} = 1, Y_S^{RS} = 1, Y_i^{RS} = 0, i = \overline{1, D}, i \neq R, S$.

Решение задачи (4)–(7) было найдено жадным алгоритмом.

Условия булевости (7) приводят к NP-полноте решаемой задачи. Однако его можно преобразовать к континуальному множеству решений задачи (4)–(7) с помощью ограничений на идемпотентность переменных [12]:

$$X_i^2 - X_i = 0, \quad i = \overline{1, D}. \quad (8)$$

Так, в последующих работах планируется погрузить дискретную математическую модель (4)–(7) в непрерывную путем замены ограничений (7) на (8). Это позволит применить градиентные методы поиска [13] оптимальной триангуляции, что должно существенно снизить вычислительные затраты.

4 Апробирование математических моделей

Произведем сравнительный анализ решений, получаемых при использовании разработанной математической модели и алгоритма триангуляции Делоне (рис. 2).

Апробирование математических моделей и их сравнение проводилось при следующих исходных данных. Известны координаты 50 точек двумерных проекций на плоскости трехмерного изображения поверхности в пространстве.

Используя представленные координаты, были найдены одиннадцать точек выпуклой оболочки.

Для двух триангуляций были найдены по 136 ребер для каждой из них. Общая длина для жадной триангуляции на основе разработанной модели составила 9 667,487, а для триангуляции Делоне — 10 096,338. При этом графы триангуляций различаются 13 уникальными ребрами.

Для сравнения пар получаемых решений применялась нормированная величина:

$$L(x, y) = \left| \frac{x - y}{\max(x, y)} \right| \cdot 100\%, \quad x > 0, y > 0, \quad (9)$$

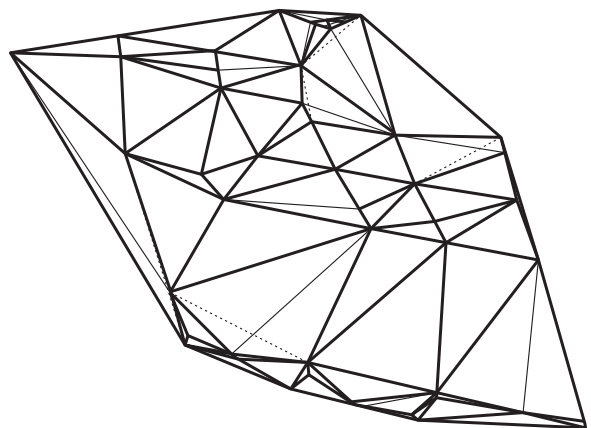


Рис. 2 Сравнение триангуляций как множеств ребер: толстые линии — общие ребра; тонкая сплошная линия — ребра, принадлежащие только триангуляции Делоне; тонкая пунктирная линия — ребра, присутствующие только в жадной триангуляции

где x, y — суммарные длины ребер для сравниваемых триангуляций.

Выигрыш по суммарной длине ребер составил 4,436%.

Далее отметим использование найденных триангуляций для решения задач навигации и задач трехмерной реконструкции. Здесь может быть использована задача коммивояжера с целью формирования оптимального маршрута обхода заданных опорных точек триангуляции. Это позволит существенно снизить размерность решаемой NP-полной задачи коммивояжера. Так, для 50 точек размерность без триангуляции составит 1225 ребер, а с триангуляцией — 136 ребер.

При исследовании математических моделей были найдены маршруты мажоритарными алгоритмами на основе сравниваемых триангуляций. Здесь выигрыш разработанной модели при использовании параметра (9) составил 14,79%.

5 Выводы

1. Центральное место в работе занимает поиск функционала пересечения отрезков с булевым множеством значений.
2. Математическая модель оптимальной триангуляции представлена в виде постановки экстремальной задачи с учетом ограничений на пересечение отрезков.
3. Математическая модель может позволить преодолеть предполагаемую NP-полноту задачи оптимальной триангуляции за счет применения градиентных методов непрерывной оптимизации.
4. Разработанные варианты математической модели оптимальной триангуляции позволяют обосновать применение эвристических приближенных алгоритмов триангуляции, а также использование большего числа расширенных классов алгоритмов оптимизации.
5. Даже в случае жадного алгоритма разработанная математическая модель приводит к меньшему значению целевого функционала на 4,436% по сравнению с алгоритмом триангуляции Делоне.
6. Использование решения задачи триангуляции в качестве исходных данных в задаче коммивояжера позволяет снизить ее размерность с $N(N-1)/2$ до $3(N-1) - N_c$.
7. Применение в задаче навигации жадного алгоритма коммивояжера к разработанной модели триангуляции позволило получить выигрыш

в 14,79% по сравнению с применением его к триангуляции Делоне.

Следующие два вывода предопределяют направления дальнейших прикладных исследований предлагаемой математической модели.

8. Так как ребра двумерной триангуляции представляют собой отрезки прямых, то и при восстановлении их в трехмерном пространстве они также являются отрезками прямых. Это свойство может быть использовано для трехмерной реконструкции концов выделенных отрезков триангуляции.
9. При трехмерной реконструкции для выделения отрезков использование решения задачи коммивояжера в условиях разработанной модели триангуляции может обеспечить меньшие искажения по сравнению с условиями триангуляции Делоне.

Литература

1. Гасилов А. В., Яковлев О. А. Создание реалистичных наборов данных для алгоритмов трехмерной реконструкции с помощью виртуальной съемки компьютерной модели // Системы и средства информатики, 2016. Т. 26. № 2. С. 98–107.
2. Костюк Ю. Л., Фукс А. Л. Приближенное вычисление оптимальной триангуляции // Геоинформатика: Теория и практика, 1998. Вып. 1. С. 61–66.
3. Лисин А. В., Файзуллин Р. Т. Эвристический алгоритм поиска приближенного решения задачи Штейнера, основанный на физических аналогиях // Компьютерная оптика, 2013. Т. 37. № 4. С. 503–510.
4. *Preparata F., Shamos M.* Вычислительная геометрия: Введение / Пер. с англ. — М.: Мир, 1989. 478 с. (*Preparata F., Shamos M.* Computational geometry. An introduction. — New York — Berlin — Heidelberg — Tokyo: Springer-Verlag, 1985. 398 p.)
5. Скворцов А. В., Мирза Н. С. Алгоритмы построения и анализа триангуляции. — Томск: ТГУ, 2006. 168 с.
6. Скворцов А. В. Триангуляция Делоне и ее применение. — Томск: ТГУ, 2002. 128 с.
7. Ивановский С. А., Симончик С. К. Алгоритмы вычислительной геометрии. Пересечение отрезков: метод заметания плоскости // Компьютерные инструменты в образовании, 2007. № 4. С. 18–33.
8. Игошин В. И. Математическая логика и теория алгоритмов. — М.: Академия, 2008. 448 с.
9. Волков И. К., Канатников А. Н. Интегральные преобразования и операционное исчисление. — 2-е изд. — М.: МГТУ им. Н. Э. Баумана, 2002. 228 с.
10. Бронштейн И. Н., Семендяев К. А. Справочник по математике. — М.: Наука, 1964. 608 с.

11. Потехина Е. А. Применение произведения Адамара в задаче перечисления упорядоченных разбиений // Научные технологии, 2012. Т. 13. № 8. С. 58–65.
12. Гантмахер Ф. Р. Теория матриц. — 2-е изд., доп. — М.: Наука, 1966. 576 с.
13. Химмельблау Д. Прикладное нелинейное программирование / Пер. с англ. — М.: Мир, 1975. 526 с. (Himmelblau D. M. Applied nonlinear programming. — New York, NY, USA: McGraw-Hill Book Co., 1972. 498 p.)

Поступила в редакцию 24.08.17

MATHEMATICAL MODEL OF OPTIMAL TRIANGULATION

A. Batenkov, Yu. Maniakov, A. Gasilov, and O. Yakovlev

Orel Branch of the Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 137 Moskovskoe Shosse, Orel 302025, Russian Federation

Abstract: The problem of synthesis of optimal planar convex triangulation is formalized. This problem arises in different applications of informatics problems and is very actual for its sections such as computer graphics and geographical information systems. The mathematical model is represented as an extremum problem with infinite number of constraints, as a minimax problem with bound variables, and as an extremum problem with additional constraints on line segments intersections of triangulation with limited number of constraints. By putting idempotent limitations on Boolean variables, the initial integer-valued problem could be solved as a general mathematical programming problem on a continuum set of answers. In addition, the comparison of results obtained by the greedy algorithm based on the represented model and Delaunay triangulation is provided.

Keywords: mathematical model; triangulation; Delaunay triangulation

DOI: 10.14357/19922264180210

References

1. Gasilov, A. V., and O. A. Yakovlev. 2016. Sozдание реалистичных наборов данных для алгоритмов трехмерной реконструкции с помощью виртуальной стрельбы [Generating realistic structure-from-motion datasets with a virtual shooting]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 26(2):98–107.
2. Kostyuk, Yu. L., and A. L. Fuks. 1998. Priblizhennoe vychislenie optimal'noy triangulyatsii [Approximate calculation of optimal triangulation]. *Geoinformatika: teoriya i praktika* [Geoinformatics: Theory and practice] 1:61–66.
3. Lisin, A. V., and R. T. Faizullin. 2013. Evristicheskiy algoritm poiska priblizhennogo resheniya zadachi Shteynera, osnovanny na fizicheskikh analogiyakh [Heuristic algorithm for finding approximate solution of Steiner problem based on physical analogies]. *Computer Optics* 37(4):503–510.
4. Preparata, F., and M. Shamos. 1985. *Computational geometry: An introduction*. — New York — Berlin — Heidelberg — Tokyo: Springer-Verlag. 398 p.
5. Skvortsov, A. V., and N. S. Mirza. 2006. *Algoritmy postroeniya i analiza triangulyatsii* [Generation and analysis algorithms of triangulation]. Tomsk: Tomsk State University. 168 p.
6. Skvortsov, A. V. 2002. *Triangulyatsiya Delone i ee primeneniye* [Delaunay triangulation and its application]. Tomsk: Tomsk State University. 128 p.
7. Ivanovskiy, S. A., and S. K. Simonchik. 2007. Algoritmy vychislitel'noy geometrii. Peresechenie otrezkov: metod zametaniya ploskosti [Computational geometry algorithms: Plane covering method]. *Komp'yuternye instrumenty v obrazovanii* [Computer instruments in education] 4:18–33.
8. Igoshin, V. I. 2008. *Matematicheskaya logika i teoriya algoritmov* [Mathematical logic and algorithms' theory]. Moscow: Akademiya. 448 p.
9. Volkov, I. K., and A. N. Kanatnikov. 2002. *Integral'nye preobrazovaniya i operatsionnoe ischislenie* [Integral transformation and operational calculus]. 2nd ed. Moscow: MSTU. 228 p.
10. Bronshtein, I. N., and K. A. Semendyaev. 1964. *Spravochnik po matematike* [Mathematics handbook]. Moscow: Nauka. 608 p.
11. Potekhina, E. A. 2012. Primeneniye proizvedeniya Adamara v zadache perechisleniya uporyadochennykh razbieniye [Application of Adamar's multiplication for enumeration of ordered decompositions problem]. *Naukoemkie tekhnologii* [Science Technologies] 13(8):58–65.
12. Gantmakher, F. R. 1966. *Teoriya matrits* [The matrix theory]. 2nd ed. Moscow: Nauka. 576 p.
13. Himmelblau, D. M. 1972. *Applied nonlinear programming*. New York, NY: McGraw-Hill Book Co. 498 p.

Received August 24, 2017

Contributors

Batenkov Alexander A. (b. 1957) — research engineer, Orel Branch of the Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 137 Moskovskoe Shosse, Orel 302025, Russian Federation; batenkov1957@mail.ru

Maniakov Yuri A. (b. 1984) — senior scientist, Orel Branch of the Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 137 Moskovskoe Shosse, Orel 302025, Russian Federation; maniakov_yuri@mail.ru

Yakovlev Oleg A. (b. 1992) — junior scientist, Orel Branch of the Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 137 Moskovskoe Shosse, Orel 302025, Russian Federation; maucra@gmail.com

Gasilov Artur A. (b. 1992) — junior scientist, Orel Branch of the Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 137 Moskovskoe Shosse, Orel 302025, Russian Federation; gasilov.av@ya.ru

АВТОМАТИЧЕСКОЕ ИЗВЛЕЧЕНИЕ МЕТАДААННЫХ ИЗ НАУЧНЫХ PDF-ДОКУМЕНТОВ*

А. В. Огальцов¹, О. Ю. Бахтеев²

Аннотация: Исследуется извлечение метаданных документа. Рассматриваются научные PDF-документы на русском языке. Особенностью формата PDF является разнообразие расположения текста на страницах документа. Это создает трудности для автоматического извлечения метаданных. Предложенный метод извлечения метаданных основан на рассмотрении текстовых блоков, полученных при помощи PDF-парсера, как объектов в задаче машинного обучения. Признаковое пространство содержит не только текстовые признаки, но и признаки, связанные с форматированием и расположением блока, которые получены из PDF-парсера. В работе измерено качество классификации предложенного алгоритма и проведено сравнение с базовым алгоритмом.

Ключевые слова: извлечение метаданных; обработка естественного языка; признаки форматирования; извлечение информации; метаописания

DOI: 10.14357/19922264180211

1 Введение

Метаданные — это информация об объекте, расположенном в каком-либо репозитории. Репозитории, такие как электронные библиотеки, Semantic Web или Open Archives, могут содержать миллионы документов. Авторы достаточно редко представляют метаданные, а электронные документы, собранные в сети Интернет автоматически, имеют метаданные еще реже; извлечение метаданных из документов вручную представляет собой трудную задачу. По этой причине возникает необходимость в системах автоматического извлечения метаданных из самих документов. В данной статье рассматривается извлечение *Заголовка, Автора, Аннотации, Содержания, Библиографии*.

Задача извлечения метаданных может быть решена с помощью правил. Преимущество этого подхода в том, что его достаточно легко реализовать без размеченного корпуса. Однако ни правила, ни регулярные выражения не могут обеспечить достаточного покрытия всевозможных стилей форматирования, так как сам процесс создания правил содержит в себе предположение о схожести форматирования исследуемых документов [1–3]. Недостатки правил заключаются также в необходимости экспертных знаний и достаточно большого времени, необходимого для их создания.

Помимо правил существуют еще несколько подходов к решению задачи извлечения метаданных. Основной из них — рассматривать текстовые блоки,

полученные из pdf-парсера, как объекты. Ставится задача классификации рассматриваемого блока с признаковым пространством, содержащим информацию о форматировании и расположении блока, а также текстовую и контекстную информацию.

Наиболее известные алгоритмы обучения с учителем, применяемые в данной задаче: Support Vector Machine (SVM) [4, 5], Hidden Markov Models (HMMs) [6] и Conditional Random Fields (CRFs) [7].

Существует семейство подходов, основанных на нейронных сетях [8, 9]. Здесь объектом является изображение PDF-документа. Для обработки изображения применяются OCR (Optical Character Recognition) системы, но они показывают низкую скорость и точность работы, поэтому в данной статье этот подход не используется.

Качество работы алгоритмов извлечения метаданных часто измеряется на небольшой коллекции достаточно однотипных документов, например из одной области знания или одного журнала [5, 10].

Научная новизна данной работы заключается в том, что авторы не ограничиваются типом *Статья*, а работают с документами различных типов (*Статья, Автореферат, Диссертация, Монография, Учебное пособие*) на русском языке. Оценка качества алгоритмов производится при помощи таких метрик, как точность, полнота и F-мера, на коллекции PDF-документов различных жанров, полученной авторами из открытых источников.

* Работа выполнена при поддержке РФФИ (проект 18-07-01441).

¹ Высшая школа экономики; ЗАО «Антиплагиат», ogaltsov@ap-team.ru

² Московский физико-технический институт; ЗАО «Антиплагиат», bahteev@ap-team.ru

Проведено сравнение результатов с базовым SVM-алгоритмом [5]. По сравнению с базовым алгоритмом использовалось больше небинарных признаков и расширенная информация о расположении блока.

2 Формальная постановка задачи

Задана выборка

$$\mathcal{D} = \{(\mathbf{x}_i, y_i)\}, \quad i = 1, \dots, m,$$

состоящая из множества пар «объект–класс», $\mathbf{x}_i \in \mathbb{R}^n$. Каждый объект \mathbf{x}_i принадлежит одному из Z классов с меткой $y_i \in \mathbb{Y} = \{1, \dots, Z\}$.

Пусть также задано множество моделей \mathfrak{F} , среди которых производится поиск подходящей модели классификации объектов выборки \mathcal{D} . Требуется найти отображение $\hat{f} \in \mathfrak{F} : \mathbb{R}^d \rightarrow \mathbb{Y}$, которое

бы минимизировало эмпирический риск на выборке \mathcal{D} :

$$\hat{f} = \arg \min_{f \in \mathfrak{F}} \sum_{\mathbf{x}_i, y_i \in \mathcal{D}} [f(\mathbf{x}_i) \neq y_i].$$

3 Описание метода и признаков

Для получения блоков документ был обработан парсером. Был выбран парсер Apache PDFBox [11], так как он показал наилучшую скорость работы и возможность настройки. Парсер возвращает текстовый блок, который является прямоугольником с текстом внутри. Все признаки блока можно разделить на четыре группы (описание всех признаков содержится в табл. 1):

- (1) **текстовые признаки.** Данная группа извлекалась из текста напрямую. Признаки этой группы проверяют наличие в строке ключевых слов и собирают различные текстовые статистики;

Таблица 1 Описание признаков

	Признаки	Описание
Текстовые признаки	isUpper	Написан ли весь текст в верхнем регистре
	blockLen	Число токенов в блоке
	hasDigit	Есть ли в блоке цифра
	NumOfPoints	Число точек
	NumOfCommas	Число запятых
	NumOfUpCase	Число символов в верхнем регистре
	UppcaseFraction	Доля символов в верхнем регистре
	NumOfDigits	Число цифр
	NumOfBibSymbols	Число символов из словаря частых слов библиографии
	hasYear	Есть ли год
	hasAnno	Есть ли слово аннотация
	isEmpty	Является ли пустым (без текста)
	hasKeyW	Есть ли слова из словаря ключевых слов
Признаки форматирования	Max/Mean/MedianAdv	Максимальное, среднее и медианное значение толщины символов блока
	Max/Mean/MedianCharH	Максимальное, среднее и медианное значение высоты символов блока
	fontIndex	Индекс семейства шрифтов
Признаки расположения	normBlockH/W/A	Нормированные высота, ширина и площадь блока
	verticalPos	Позиция сверху страницы
	leftIndent	Левый отступ
	rightIndent	Правый отступ
	relativeBlockCenter	Позиция центра блока
Контекстные признаки	sameFontAbove/Below	Блок сверху/снизу имеет такой же индекс шрифта
	sameCenterAbove/Below	Блок сверху/снизу имеет такой же центр
	sameLeftIndentAbove/Below	Блок сверху/снизу имеет такой же левый отступ
	hasYearInNeigh	В фиксированном окне вокруг блока есть год
	isEmptyAbove/Below	Блок сверху/снизу пуст
	hasAnnoInNeigh	В фиксированном окне вокруг блока есть слова из словаря аннотации
	isCharHSmallerThanInNeigh	Средний размер шрифта блока меньше, чем средний размер шрифта в фиксированном окне

- (2) **признаки форматирования.** В эту группу входят такие признаки, как размер шрифта, семейство шрифтов и др.;
- (3) **признаки расположения.** В этой группе содержится информация о положении блока на странице, такая как правый и левый отступ, координата центра блока и др.;
- (4) **контекстные признаки.** Эти признаки описывают контекст рассматриваемого блока, как предложено в [10].

4 Эксперимент

4.1 Описание коллекции

Коллекция содержит 550 размеченных PDF-документов из различных областей знаний и разных типов (табл. 2).

Одна из основных характеристик задачи — сильная несбалансированность классов (табл. 3). Видно, что общее число блоков в табл. 3 варьируется для категорий метаданных. Это связано с эвристиками, примененными на этапе обработки документа. Например, для заголовка, автора и аннотации обрабатывалась только первая и вторая страницы документа, так как обработка занимает достаточно много времени и нет необходимости обрабатывать всю диссертацию, чтобы извлечь данные категории. Для библиографии извлекались последние $k = \max(0, 1p, 1)$ страниц документа, для содержания — первые $n = \max(0, 1p, 4)$ страниц документа, где p — число страниц в документе. Количество блоков с автором меньше, чем число работ. Это вызвано ошибками работы парсера, когда блок с именем автора не извлекался.

Таблица 2 Распределение жанров

Жанр	Число документов
Статья	148
Автореферат	161
Диссертация	196
Монография, пособие	45

Таблица 3 Распределение блоков по категориям

Категория	Число блоков данной категории	Полное число извлеченных блоков
Заголовок	957	76 986
Автор	432	76 986
Аннотация	1386	76 986
Библиография	45 323	316 015
Содержание	5866	309 681

4.2 Описание эксперимента

Для обоснованного выбора алгоритма была снижена размерность данных с помощью t-SNE [12]. Визуализация представлена на рис. 1. Видно, что в пространстве сниженной размерности данные не являются линейно разделимыми, но есть однородные кластеры. Этот факт послужил одним из обоснований применения ансамбля решающих деревьев. Также этот алгоритм показал наилучшее качество во время экспериментов.

Для проведения экспериментов выборка была разделена на обучающую и тестовую в пропорции 80/20 стратифицированно в соответствии с табл. 3. Качество измерялось стандартными метриками бинарной классификации: точность, полнота и F-мера. Результаты приведены в табл. 4. Качество работы алгоритма в данной работе сравнивалось с базовым алгоритмом, взятым из статьи [5]. Отличия признакового пространства базового алгоритма от предложенного:

- в базовом алгоритме есть признаки, собирающие различные html-теги, так как использовался другой pdf-парсер. В проведенном эксперименте такие признаки не собирались;
- в отличие от базового, в предложенном алгоритме использовались небинарные текстовые статистики, больше признаков расположения и контекстные признаки.

В базовом алгоритме использовался SVM и извлекалась только первая страница статей узкой направленности. Базовый алгоритм показал низ-

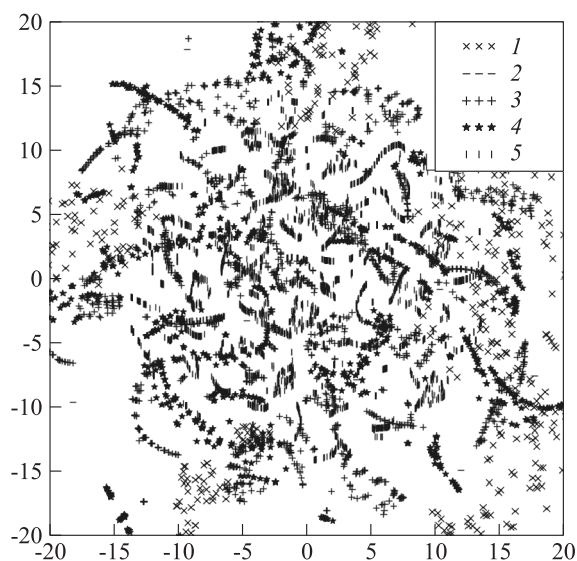


Рис. 1 Визуализация данных при помощи t-SNE: 1 — заголовки; 2 — автор; 3 — библиография; 4 — содержание; 5 — другое

Таблица 4 Результаты эксперимента

Категория	Метод	P	R	F
Заголовок	Предложенный	0,74	0,79	0,76
	Предложенный только на текстовых признаках	0,67	0,66	0,66
	Базовый	0,20	0,77	0,32
Автор	Предложенный	0,78	0,71	0,74
	Предложенный только на текстовых признаках	0,45	0,74	0,56
	Базовый	0,33	0,75	0,46
Библиография	Предложенный	0,76	0,85	0,80
	Предложенный только на текстовых признаках	0,59	0,85	0,69
	Базовый	—	—	—
Содержание	Предложенный	0,72	0,74	0,73
	Предложенный только на текстовых признаках	0,71	0,69	0,70
	Базовый	—	—	—
Аннотация	Предложенный	0,73	0,71	0,72
	Предложенный только на текстовых признаках	0,22	0,54	0,31
	Базовый	0,09	0,63	0,16

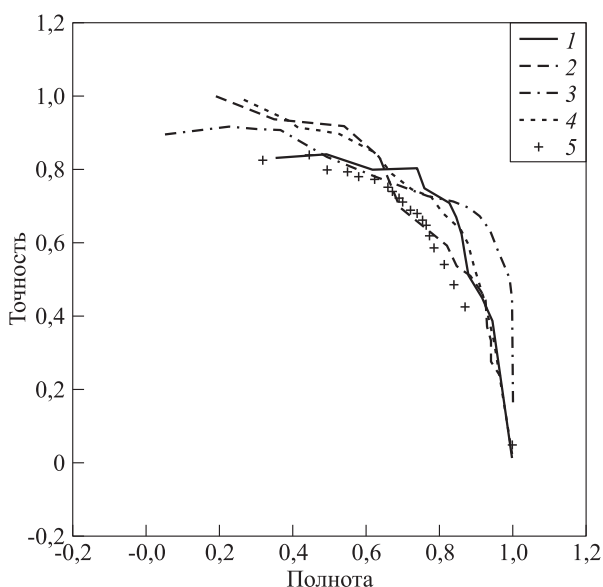


Рис. 2 Кривые полнота–точность: 1 — заголовок; 2 — автор; 3 — библиография; 4 — содержание; 5 — аннотация

кое качество, так как в текущей постановке эксперимента стиль форматирования документов очень сильно варьируются.

В табл. 4 приведены результаты эксперимента, в котором признаковое пространство состояло только из текстовых статистик. Видно, что расширение признакового пространства информацией о расположении и форматировании блока заметно повышает качество классификации.

Из-за несбалансированности классов была проведена редукция блоков, относящихся к классу *Прочее*. Если 150 элементов до и после блока *Прочее* не содержали других классов, то такой блок удалялся.

Проводились эксперименты с порогом вероятности классификации и весами классов. Эксперименты показали, что порог классификации 0,5 и вес редкого класса, равный двум, демонстрирует наилучший результат для *Содержания*, *Библиографии*, *Аннотации*. Для *Заголовка* и *Автора* перенос порога классификации в 0,6 и 0,65 соответственно в комбинации с весом редкого класса, равным 20, дали повышение точности (+0,05 и +0,18) с падением полноты (−0,02 и −0,16), но с небольшим повышением F-меры. Кривая полнота–точность изображена на рис. 2.

Был проведен анализ значимости признаков, которая была получена стандартным для алгоритма Random Forest способом [13]. Три первых по значимости признака для каждого класса приведены в табл. 5. Видно, что почти в каждом классе

Таблица 5 Результаты анализа значимости признаков

Категория	Признак	Значимость
Заголовок	NumOfUpCase	0,65
	verticalPos	0,06
	MeanAdv	0,06
Автор	UppcaseFraction	0,34
	NumOfUpCase	0,23
	verticalPos	0,11
Содержание	NumOfPoints	0,41
	blockLen	0,18
	NumOfDigits	0,08
Библиография	blockLen	0,14
	relativeBlockCenter	0,12
	normBlockW	0,10
Аннотация	verticalPos	0,10
	normBlockA	0,10
	MeanAdv	0,09

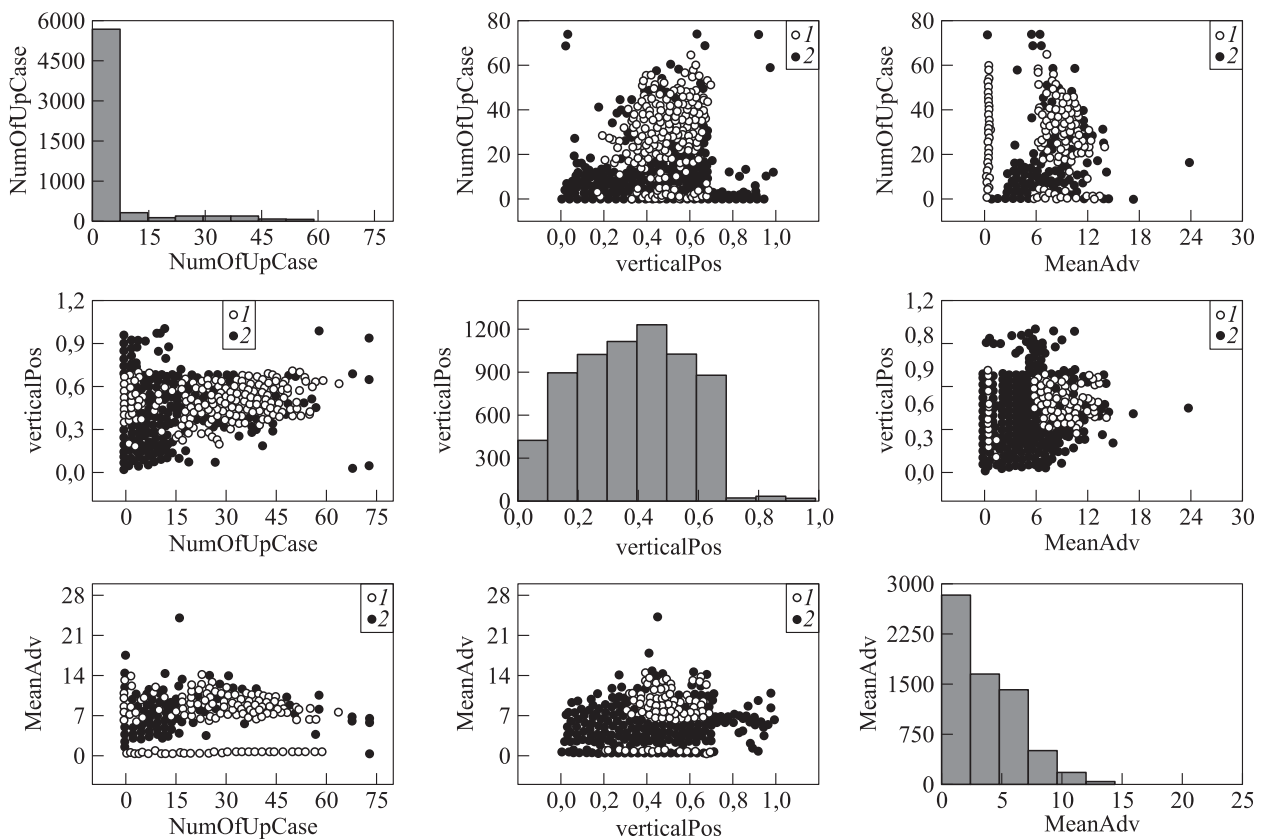


Рис. 3 Визуализация данных для заголовка (1 — заголовок; 2 — прочее) в проекции на пространство значимых признаков

признаки расположения и форматирования вошли в число значимых. Данные были визуализированы в проекции на пространство значимых признаков. На рис. 3 и 4 показаны примеры таких визуализаций для категорий *Заголовки* и *Авторы*. На рисунках видны сегменты, где преимущественно расположен целевой класс.

4.3 Анализ ошибок

С целью выявления основных ошибок предложенного алгоритма был произведен анализ результатов.

Заголовок. В табл. 5 видно, что для категории *Заголовки* самым значимым признаком является число символов в верхнем регистре. Поэтому ложно-положительные примеры наблюдаются вблизи заголовка по расположению и с большим количеством символов в верхнем регистре.

Примеры:

- ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ УПРАВЛЕНИЯ
- 05.13.11 МАТЕМАТИЧЕСКОЕ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ

Основной причиной ложно-отрицательных примеров оказались длинные заголовки с форматированием, близким к обыкновенному тексту. Примеры:

- *Французское органное искусство Барокко: музыка, органостроение, исполнительство*
- *Нейтральный внешнеполитический курс Нидерландов: от Мюнстерского мира 1648 г. до конца Первой мировой войны*
- *Музыкальная жизнь Москвы XIX столетия и ее отражение в фортепианной практике*

Автор. Если извлеченный блок располагался вверху страницы и содержал имя и фамилию, то наблюдался ложно-положительный пример. Примеры такого рода наблюдались, когда блок содержал комбинацию точек и символов в верхнем регистре, схожую с именем и фамилией.

Примеры:

- *им. М. В. Келдыша РАН*
- *(ПЖ). Одним*

Ложно-отрицательные примеры возникали, когда блок с автором находился в очень нетипичном месте.

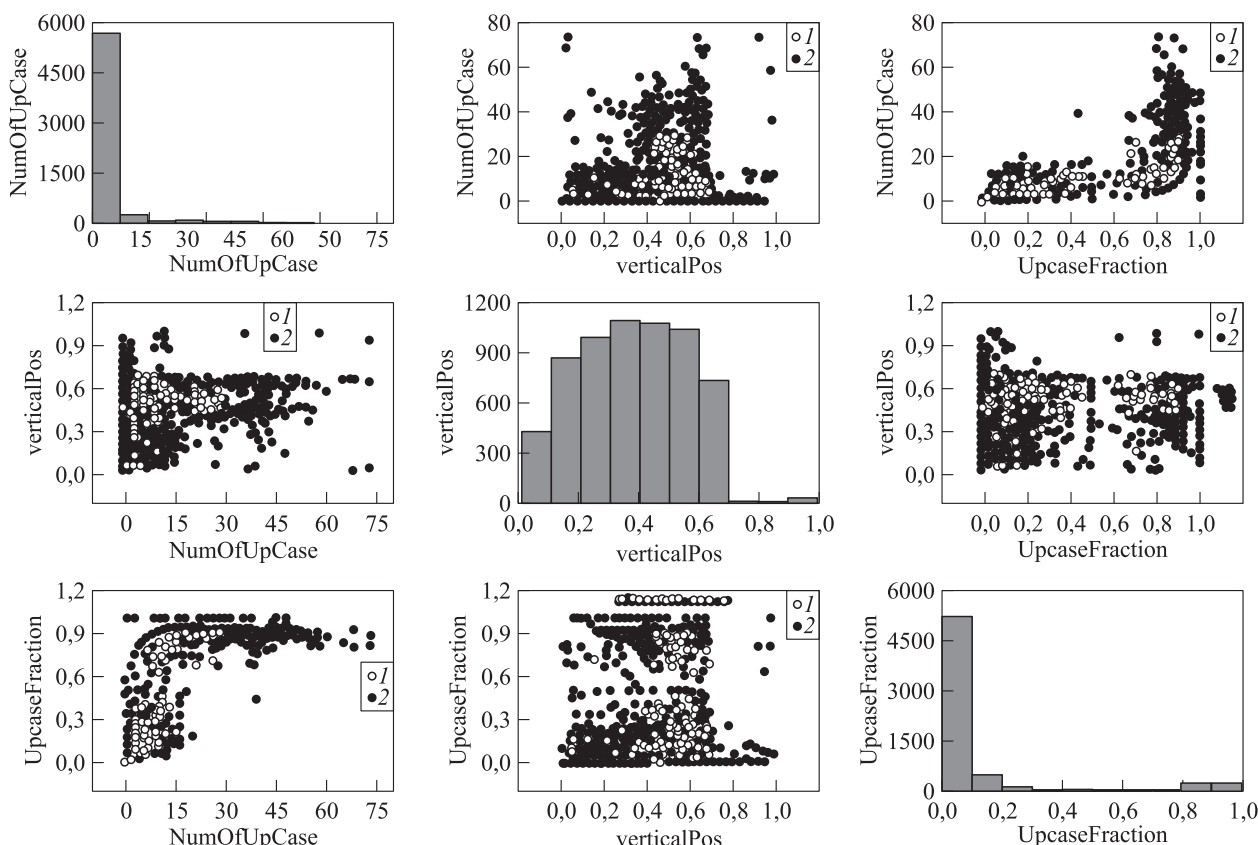


Рис. 4 Визуализация данных для автора (1 — автор; 2 — прочее) в проекции на пространство значимых признаков

- Огальцов А. В. (если находится где-то в очень нетипичном месте страницы)

Аннотация. Самые частые ошибки обоих типов в этой категории возникали в связи с тем, что блоки с аннотацией легко спутать с обычным текстом.

Содержание. Ложно-положительные примеры — это в основном блоки с большим числом точек, скорее всего вызванных ошибками обработки.

Примеры:

- В ТОМСКОЙ ОБЛАСТИ
- Случай $d(k)=0$

Ложно-отрицательные примеры вызваны форматированием, близким к обычному тексту.

- Функции ввода, вывода и работы с символами
- Циклы, блоки и присваивания

Библиография. Ложно-положительные примеры этой категории похожи на библиографию расположением и форматированием.

Примеры:

- Тел.: (3472) 1234567, 1234567

- Печать офсетная. Гарнитура NewtonС. Бумага офсетная

- как отмечают А. Я. Гаев и др. 2005, нереально.

Ложно-отрицательные примеры были вызваны сильной схожестью с обычным текстом.

Большое число блоков с обычным текстом делает возможным ситуацию сходства этих блоков с целевыми по ряду признаков, поэтому одним из направлений улучшения является снижение несбалансированности выборки различными методами.

5 Заключение

В работе была исследована возможность применения методов машинного обучения к задаче извлечения метаданных. Использовано расширенное признаковое пространство, включающее не только текстовые признаки, но также признаки форматирования и расположения, которые оказались важными в решении проблемы большой вариативности стилей форматирования документа. Измерено качество предлагаемого алгоритма на коллекции научных PDF-документов на русском языке, взятых из открытых источников. Проведен анализ ошибок.

Показано, что предложенный алгоритм значительно превосходит базовый на документах различных типов.

Будущие исследования будут направлены на расширение признакового пространства и на обзор возможностей новых PDF-парсеров. Также будет исследована возможность применения графических моделей, таких как условные случайные поля.

Авторы выражают свою благодарность к.ф.-м.н. Чеховичу Ю. В. за ценные советы при планировании исследования и рекомендации по оформлению статьи.

Литература

1. Bergmark D. Automatic extraction of reference linking information from online documents. — Ithaca, NY, USA: Cornell University, 2000. 20 p.
2. Klink S., Dengel A., Kieninger T. Document structure analysis based on layout and textual features // Workshop (International) on Document Analysis Systems Proceedings. — Boston, MA, USA, 2000. P. 99–111.
3. Mao S., Kim J. W., Thoma G. R. A dynamic feature generation system for automated metadata extraction in preservation of digital materials // 1st Workshop (International) on Document Image Analysis for Libraries Proceedings. — Palo Alto, CA, USA: IEEE Computer Society, 2004. P. 225–232.
4. Han H., Giles C. L., Manavoglu E., Zha H., Zhang Z., Fox E. A. Automatic document metadata extraction using support vector machines // 3rd ACM/IEEE-CS Joint Conference on Digital Libraries Proceedings, 2003. P. 37–48.
5. Kovacević A., Ivanović D., Milosavljević B., Konjović Z., Surla D. Automatic extraction of metadata from scientific publications for CRIS systems // Program, 2011. Vol. 45. Iss. 4. P. 376–396.
6. Seymore K., McCallum A., Rosenfeld R. Learning hidden Markov model structure for information extraction // AAI 99 Workshop on Machine Learning for Information Extraction Proceedings, 1999. P. 37–42.
7. Councill I., Giles C. L., Kan M.-Y. ParsCit: An open-source CRF reference string parsing package // 6th Conference (International) on Language Resources and Evaluation Proceedings. — Marrakech, Morocco, 2008. P. 661–667.
8. Rangoni Y., Belaïd A. Document logical structure analysis based on perceptive cycles // IAPR Workshop on Document Analysis Systems, 2006. P. 117–128.
9. Rangoni Y., Belaïd A., Vajda S. Labelling logical structures of document images using a dynamic perceptive neural network // Int. J. Doc. Anal. Recog., 2012. Vol. 15. Iss. 1. P. 45–55.
10. Tao X., Tang Z., Xu C. Document page structure learning for fixed-layout e-books using conditional random fields // Document Recognition and Retrieval XXI: Proc. SPIE, 2014. Vol. 9021. P. 1–9.
11. Apache PDFBox. <http://pdfbox.apache.org>.
12. Van der Maaten L. J. P., Hinton G. E. Visualizing high-dimensional data using t-SNE // J. Mach. Learn. Res., 2008. Vol. 9. P. 2579–2605.
13. Breiman L. Random forests // Mach. Learn., 2001. Vol. 45. P. 5–32.

Поступила в редакцию 20.12.17

AUTOMATIC METADATA EXTRACTION FROM SCIENTIFIC PDF DOCUMENTS

A. V. Ogaltsov^{1,2} and O. Y. Bakhteev^{2,3}

¹National Research University Higher School of Economics, 20 Myasnitskaya Str., Moscow 101000, Russian Federation

²Antiplagiat JSC, 33 Varshavskoe Shosse, Moscow 117105, Russian Federation

³Moscow Institute of Physics and Technology, 9 Institutskiy Per., Dolgoprudny, Moscow Region 141700, Russian Federation

Abstract: The authors investigate the task of metadata extraction. The authors consider scientific PDF documents in Russian. One of the features of PDF is a rich layout. It is difficult to extract metadata due to this fact. The authors propose a method based on considering blocks from pdf-parser as objects in a machine learning task. The feature space is constructed not only of text statistics but also of formatting and positioning features of the block. The authors performed computational experiments and compared their approach with the baseline.

Keywords: metadata extraction; natural language processing; layout features; information retrieval; metadescrptions

DOI: 10.14357/19922264180211

Acknowledgments

The paper was supported by the Russian Foundation for Basic Research (project 18-07-01441).

References

1. Bergmark, D. 2000. Automatic extraction of reference linking information from online documents. Ithaca, NY: Cornell University. 20 p.
2. Klink, S., A. Dengel, and T. Kieninger. 2010. Document structure analysis based on layout and textual features. *Workshop (International) on Document Analysis Systems Proceedings*. Boston, MA. 99–111.
3. Mao, S., J. W. Kim, and G. R. Thoma. 2004. A dynamic feature generation system for automated metadata extraction in preservation of digital materials. *1st Workshop (International) on Document Image Analysis for Libraries Proceedings*. Palo Alto, CA: IEEE Computer Society. 225–232.
4. Han, H., C. L. Giles, E. Manavoglu, H. Zha, Z. Zhang, and E. A. Fox. 2003. Automatic document metadata extraction using support vector machines. *3rd ACM/IEEE-CS Joint Conference on Digital Libraries Proceedings*. 37–48.
5. Kovacević, A., D. Ivanović, B. Milosavljević, Z. Konjović, and D. Surla. 2011. Automatic extraction of metadata from scientific publications for CRIS systems. *Program* 45(4):376–396.
6. Seymore, K., A. McCallum, and R. Rosenfeld. 1999. Learning hidden Markov model structure for information extraction. *AAAI 99 Workshop on Machine Learning for Information Extraction Proceedings*. 37–42.
7. Councill, I., C. L. Giles, and M.-Y. Kan. 2008. ParsCit: An open-source CRF reference string parsing package. *6th Conference (International) on Language Resources and Evaluation Proceedings*. Marrakech, Morocco. 661–667.
8. Rangoni, Y., and A. Belaïd. 2006. Document logical structure analysis based on perceptive cycles. *IAPR Workshop on Document Analysis Systems*. 117–128.
9. Rangoni, Y., A. Belaïd, and S. Vajda. 2012. Labelling logical structures of document images using a dynamic perceptive neural network *Int. J. Doc. Anal. Recog.* 15(1):45–55.
10. Tao, X., Z. Tang, and C. Xu. 2013. Document page structure learning for fixed-layout e-books using conditional random fields. *Document Recognition and Retrieval XXI: Proc. SPIE* 9021:1–9.
11. Apache PDFBox. Available at: <http://pdfbox.apache.org/> (accessed June 19, 2018).
12. Van der Maaten, L. J. P., and G. E. Hinton. 2008. Visualizing high-dimensional data using t-SNE. *J. Mach. Learn. Res.* 9:2579–2605.
13. Breiman, L. 2001. Random forests. *Mach. Learn.* 45:5–32.

Received December 20, 2017

Contributors

Ogaltsov Alexander V. (b. 1993) — student, National Research University Higher School of Economics, 20 Myasnikskaya Str., Moscow 101000, Russian Federation; researcher, Antiplagiat JSC, 33 Varshavskoe Shosse, Moscow 117105, Russian Federation; avogaltsov@edu.hse.ru

Bakhteev Oleg Y. (b. 1993) — PhD student, Moscow Institute of Physics and Technology, 9 Institutskiy Per., Dolgoprudny, Moscow Region 141700, Russian Federation; researcher, Antiplagiat JSC, 33 Varshavskoe Shosse, Moscow 117105, Russian Federation; bakhteev@phystech.edu

РЕШЕНИЕ ПРОБЛЕМЫ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ ЗАПАСОМ НЕПРЕРЫВНОГО ПРОДУКТА ПРИ ПОСТОЯННО ПРОИСХОДЯЩЕМ ПОТРЕБЛЕНИИ В СТОХАСТИЧЕСКОЙ ПОЛУМАРКОВСКОЙ МОДЕЛИ

П. В. Шнурков¹, А. Ю. Егоров²

Аннотация: Теоретически обосновано решение задачи оптимального управления запасом в рассматриваемой полумарковской модели. Для достижения этой цели произведены формальные аналитические преобразования полученных авторами ранее интегральных представлений для основных вероятностных характеристик модели. Данные преобразования сделали возможным применение теоремы об аналитическом представлении стационарного стоимостного показателя эффективности управления полумарковским процессом в форме дробно-линейного интегрального функционала. Используется общая теорема об экстремуме дробно-линейного интегрального функционала, доказанная П. В. Шнурковым. Эта теорема позволяет свести поставленную задачу оптимального управления запасом к задаче исследования на глобальный экстремум некоторой заданной функции от конечного числа действительных неотрицательных переменных, которая может быть эффективно решена на практике с использованием известных численных методов.

Ключевые слова: управление запасом; полумарковский случайный процесс; стационарный стоимостный показатель; оптимальное управление стохастическими системами; дробно-линейный интегральный функционал

DOI: 10.14357/19922264180212

1 Введение

Настоящая работа является второй и завершающей частью исследования стохастической полумарковской модели управления запасом непрерывного продукта при постоянно происходящем потреблении, первая часть которого была изложена в статье [1].

Характеризуя данное исследование в целом, можно отметить следующее.

В первой части была построена и проанализирована стохастическая полумарковская модель управления запасом непрерывного продукта. Основным результатом первой части является получение явных аналитических представлений для вероятностных характеристик построенной полумарковской модели, необходимых для решения задачи оптимального управления.

Вторая часть посвящена непосредственному решению задачи оптимального управления, которая формально представляет собой задачу нахождения глобального экстремума стационарного стоимостного показателя эффективности управления, заданного на множестве наборов вероятностных рас-

пределений, которые определяют стратегию управления.

Напомним в краткой форме читателю описание системы управления запасом и соответствующей стохастической модели, предложенной в работе [1].

Исследуемая система представляет собой хранилище максимальной вместимости τ , из которого производится непрерывное потребление продукта. Рассматриваемая стохастическая модель представляет собой пару случайных процессов $(x(t), \zeta(t))$ в которой основной процесс $x(t)$ описывает текущий объем запаса в системе в момент времени t . Помимо основного случайного процесса $x(t)$ вводится вспомогательный полумарковский случайный процесс $\zeta(t)$ с конечным множеством состояний $\{0, 1, \dots, N_0\}$, называемый в данном исследовании сопровождающим. Такой полумарковский процесс определяется состоянием основного процесса в моменты пополнения запаса t_n , $n = 0, 1, 2, \dots$. В эти моменты происходят изменения состояния сопровождающего процесса и случайный выбор решения (управления), которое определяется как реализация случайной величины u_n с распределением $G_i(u)$ при условии $\zeta(t_n) = \zeta_n = i$. В рассмат-

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

²Национальный исследовательский университет «Высшая школа экономики», ayuegorov@hse.ru

риваемой модели случайная величина u_n представляет собой время от момента очередного пополнения до момента заказа на следующее пополнение запаса. Период времени между последовательными моментами пополнения запаса $[t_n, t_{n+1})$ разделяется на два этапа. На первом этапе, называемом периодом чистого потребления, скорость потребления фиксирована и равна величине α_c , а длительность такого периода представляет собой случайную величину u_n , имеющую распределение $G_i(u)$, если $\zeta(t_n) = \zeta_n = i$. Второй этап по содержанию представляет собой время выполнения заказа или период задержки поставки. Длительность этого этапа является известной детерминированной величиной h . Потребление продукта на втором этапе продолжается с заданной скоростью α_w , причем $\alpha_w \leq \alpha_c$.

Пополнение запаса происходит мгновенно в моменты $t_n, n = 0, 1, 2, \dots$. Процедура пополнения является стохастической и имеет двухэтапный характер. Полное формализованное описание процедуры пополнения и всей полумарковской модели в целом приведено в работе [1] и приложении к ней.

2 Показатель эффективности управления и общая постановка оптимизационной задачи

Определим показатель эффективности управления. Обозначим через $V(t), t \geq 0$, некоторый аддитивный функционал, имеющий смысл прибыли, полученной в результате функционирования системы за период времени $[0, t]$. Данный функционал связан с исходным случайным процессом $x(t)$ и сопровождающим полумарковским процессом $\zeta(t)$. Формальное задание такого функционала производится по схеме, описанной в классических работах [2, 3]. Известно, что при достаточно общих условиях, сформулированных в указанных работах (см. также современные исследования для общих управляемых полумарковских моделей [4]), имеет место следующее утверждение, называемое эргодической теоремой для аддитивного функционала:

$$I = \lim_{t \rightarrow \infty} \frac{EV(t)}{t} = \frac{\sum_{i=0}^{N_0} d_i \pi_i}{\sum_{i=0}^{N_0} T_i \pi_i}. \quad (1)$$

Дадим определения величин, входящих в правую часть соотношения (1):

$\Delta V_n = V(t_{n+1}) - V(t_n)$ — приращение функционала прибыли за период времени между последовательными моментами изменения

состояния сопровождающего полумарковского процесса $\zeta(t)$;

$d_i = E[\Delta V_n | \zeta_n = i]$ — условное математическое ожидание прибыли, образующейся за период между последовательными моментами изменения состояния сопровождающего процесса при условии, что в этот период указанный процесс находился в состоянии i ;

$T_i = E[t_{n+1} - t_n | \zeta_n = i]$ — условное математическое ожидание длительности периода между последовательными моментами изменения состояния сопровождающего процесса при условии, что в этот период указанный процесс находился в состоянии i ;

$\Pi = \{\pi_0, \dots, \pi_{N_0}\}$ — стационарное распределение цепи Маркова $\{\zeta_n\}$, вложенной в данный полумарковский процесс $\zeta(t)$.

Следуя уже упомянутому классическому исследованию [2, 3], будем считать величину I показателем эффективности управления в рассматриваемой модели. Данный показатель имеет смысл средней удельной прибыли в установившемся (стационарном) режиме функционирования системы.

Задача оптимального управления в рассматриваемой модели формализуется в виде задачи нахождения безусловного глобального экстремума функционала $I = I(G_0, G_1, \dots, G_{N_0})$ на множестве наборов вероятностных распределений $\{G_0(\cdot), G_1(\cdot), \dots, G_{N_0}(\cdot)\}$.

3 Интегральные преобразования вероятностных характеристик модели

В работе [1] были доказаны три основных утверждения, сформулированных в виде теорем, в которых были представлены явные аналитические выражения для величин d_i и $T_i, i = 0, 1, \dots, N_0$, а также для вероятностей перехода цепи Маркова, вложенной в сопровождающий полумарковский процесс $\zeta(t)$. Полученные аналитические выражения имеют форму двойных интегралов от вероятностных распределений $G_i(\cdot)$ и $B_i(\cdot), i = 0, 1, \dots, N_0$.

Вероятностные распределения $B_i(\cdot), i = 0, 1, \dots, N_0$, связаны с процедурой пополнения запаса и предполагаются заданными. Распределения $G_0(\cdot), G_1(\cdot), \dots, G_{N_0}(\cdot)$ описывают управления и должны быть определены в результате решения задачи оптимизации. При этом указанные двойные интегралы устроены таким образом, что внутренний из них представляет собой интеграл по

мере, задаваемой распределением $G_i(\cdot)$, а внешний — по мере, задаваемой распределением $B_i(\cdot)$, $i = 0, 1, \dots, N_0$.

Для решения поставленной задачи оптимального управления необходимо провести аналитические преобразования упомянутых вероятностных характеристик, которые в целом заключаются в изменении порядка интегрирования в соответствующих двойных интегралах. Проводимые преобразования основаны на теореме Фубини [5] и связаны с конкретным видом областей интегрирования.

Приведем полученные результаты в форме утверждений об аналитических представлениях характеристик d_i и T_i , $i = 0, 1, \dots, N_0$, и вероятностей перехода вложенной цепи Маркова $\{\zeta_n\}$. Полные аналитические обоснования этих утверждений приведены в приложении к настоящей работе [6].

Утверждение 1. Вспомогательные вероятностные характеристики, связанные с вероятностями перехода вложенной цепи Маркова $\{\zeta_n\}$, введенные в работе [1], представимы в следующем виде:

$$p_{ik}^{(+)} = \int_{(i-k)L/\alpha_c - \alpha_w h/\alpha_c}^{(i-k)L/\alpha_c - \alpha_w h/\alpha_c} [B_i(\alpha_c u + (k+1)L + \alpha_w h) - B_i(iL)] dG_i + \int_{(i+1-k)L/\alpha_c - \alpha_w h/\alpha_c}^{(i-k)L/\alpha_c - \alpha_w h/\alpha_c} [B_i((i+1)L) - B_i(\alpha_c u + kL + \alpha_w h)] dG_i(u) \text{ при } k = 0, \dots, k_i^{(+)} - 1;$$

$$p_{ik_i^{(+)}}^{(+)} = \int_0^{(i+1-k_i^{(+)})L/\alpha_c - \alpha_w h/\alpha_c} [B_i((i+1)L) - B_i(\alpha_c u + k_i^{(+)}L + \alpha_w h)] dG_i(u) \text{ при } k = k_i^{(+)};$$

$$p_{ik}^{(+)} = 0 \text{ при } k > k_i^{(+)};$$

$$p_{ik}^{(-)} = \int_{(i+k)L/\alpha_c - \alpha_w h/\alpha_c}^{(i+k+1)L/\alpha_c - \alpha_w h/\alpha_c} [B_i(\alpha_c u_0 - kL + \alpha_w h) - B_i(iL)] dG_i(u) + \int_{(i+k+1)L/\alpha_c - \alpha_w h/\alpha_c}^{(i+k+2)L/\alpha_c - \alpha_w h/\alpha_c} [B_i((i+1)L) - B_i(\alpha_c u - (k+1)L + \alpha_w h)] dG_i(u) \text{ при } k = k_i^{(-)} + 1, \dots, N_1 - 1;$$

$$p_{ik_i^{(-)}}^{(-)} = \int_0^{(i+k_i^{(-)}+2)L/\alpha_c - \alpha_w h/\alpha_c} [B_i((i+1)L) - B_i(\alpha_c u - Lk_i^{(-)} + \alpha_w h)] dG_i(u) \text{ при } k = k_i^{(-)};$$

$$p_{iN_1}^{(-)} = \int_{(i+N_1)L/\alpha_c - \alpha_w h/\alpha_c}^{(i+N_1+1)L/\alpha_c - \alpha_w h/\alpha_c} [B_i(\alpha_c u - LN_1 + \alpha_w h) - B_i(iL)] dG_i(u) + [B_i((i+1)L) - B_i(iL)] \times \left[1 - G_i\left(\frac{(i+N_1+1)L}{\alpha_c} - \frac{\alpha_w h}{\alpha_c}\right) \right] \text{ при } k = N_1, (i+N_1)L > \alpha_w h;$$

$$p_{iN_1}^{(-)} = \int_{iL}^{(i+1)L} \left[\int_0^\infty dG_i(u) \right] dB_i(x) = [B_i((i+1)L) - B_i(iL)] [1 - g_{0i}] \text{ при } k = N_1, (i+N_1)L \leq \alpha_w h;$$

$$p_{ik}^{(-)} = 0 \text{ при } k = 0, 1, \dots, k_i^{(-)} - 1,$$

где $k_i^{(\pm)}$ — граничные значения дискретной величины k , которые зависят от известных исходных параметров системы α_w, h и L :

$$k = k_i^{(+)}(\alpha_w, h, L) = \max\left(0, i - \frac{\alpha_w h}{L}\right);$$

$$k = k_i^{(-)}(\alpha_w, h, L) = \left\lfloor \min\left(0, i - \frac{\alpha_w h}{L}\right) \right\rfloor.$$

Утверждение 2. Выражения для математических ожиданий длительностей пребывания сопровождающего процесса $\zeta(t)$ в различных состояниях имеют вид:

$$T_i = \int_0^\infty (u [B_i((i+1)L) - B_i(iL)]) dG_i(u) + h = \int_0^\infty u dG_i(u) + h, \quad i = 0, 1, \dots, N_0.$$

Утверждение 3. Вспомогательные характеристики, связанные с математическими ожиданиями приращений функционала прибыли $d_{ik}^{(+)}$ и $d_{ik}^{(-)}$, введенными в работе [1], представимы в виде линейных комбинаций двойных интегралов по распределениям $G_i(\cdot)$ и $B_i(\cdot)$, $i = 0, 1, \dots, N_0$, таких, что внутренний интеграл представляет собой интеграл по мере, задаваемой распределением $B_i(\cdot)$, а внешний — по мере, задаваемой распределением $G_i(\cdot)$, $i = 0, 1, \dots, N_0$.

Явные представления для характеристик $d_{ik}^{(+)}$ имеют следующий вид:

$$d_{ik}^{(+)} = \int_{\frac{(i-k)L}{\alpha_c} - \frac{\alpha_w h}{\alpha_c}}^{\frac{(i-k)L}{\alpha_c} - \frac{\alpha_w h}{\alpha_c}} \left[\int_{iL}^{\alpha_c u + (k+1)L + \alpha_w h} D_{ik}^{(+)}(x, u) dB_i(x) \right] dG_i(u) + \int_{\frac{(i-k)L}{\alpha_c} - \frac{\alpha_w h}{\alpha_c}}^{\frac{(i+1-k)L}{\alpha_c} - \frac{\alpha_w h}{\alpha_c}} \left[\int_{\alpha_c u + kL + \alpha_w h}^{(i+1)L} D_{ik}^{(+)}(x, u) dB_i(x) \right] dG_i(u)$$

при $k = 0, \dots, k_i^{(+)} - 1$;

$$d_{ik_i^{(+)}}^{(+)} = \int_0^{\frac{(i+1-k_i^{(+)})L}{\alpha_c} - \frac{\alpha_w h}{\alpha_c}} \left[\int_{\alpha_c u + k_i^{(+)}L + \alpha_w h}^{(i+1)L} D_{ik_i^{(+)}}^{(+)}(x, u) dB_i(x) \right] dG_i(u)$$

при $k = k_i^{(+)}$;

$$d_{ik}^{(+)} = 0 \text{ при } k > k_i^{(+)};$$

$$D_{ik}^{(+)}(x, u) = \int_0^u [\alpha_c g_1(x - \alpha_c t) - c_1(x - \alpha_c t)] dt + \int_0^h [\alpha_w g_1(x - \alpha_c u - \alpha_w t) - c_1(x - \alpha_c u - \alpha_w t)] dt.$$

Отметим, что работа по аналитическому преобразованию всех вспомогательных вероятностных характеристик, полученных в теоремах 1–3 работы [1], потребовала детального исследования всех возможных вариантов соотношений исходных параметров модели. При этом выражения для характеристик $d_{ik}^{(-)}$ имеют структуру, аналогичную структуре характеристик $d_{ik}^{(+)}$, представленных в утверждении 3, однако более громоздки по форме и связаны с анализом большего числа вариантов в силу особенностей областей интегрирования. В связи с этим явные представления для указанных характеристик не могут быть приведены в основном тексте настоящей работы.

Конкретные формулы для характеристик $d_{ik}^{(-)}$ и необходимые пояснения к проведенным аналитическим преобразованиям можно найти в приложении [6].

4 Аналитическое представление показателя эффективности управления в форме дробно-линейного интегрального функционала

В результате проведения интегральных преобразований, описанных в предыдущем разделе, основные вероятностные характеристики рассматриваемой полумарковской модели представлены в следующей форме:

$$p_{ij} = \int_0^\infty p_{ij}(u_i) dG_i(u_i), \quad i, j \in \{0, 1, \dots, N_0\}; \quad (2)$$

$$T_i = \int_0^\infty T_i(u_i) dG_i(u_i), \quad i \in \{0, 1, \dots, N_0\}; \quad (3)$$

$$d_i = \int_0^\infty d_i(u_i) dG_i(u_i), \quad i \in \{0, 1, \dots, N_0\}. \quad (4)$$

При этом подынтегральные функции $p_{ij}(u_i)$, $d_i(u_i)$ и $T_i(u_i)$ заданы явно и не зависят от вероятностных распределений $G_i(u_i)$, $i = 0, 1, \dots, N_0$.

Замечание 1. Условные математические ожидания d_i непосредственно определяются через вспомогательные вероятностные характеристики $d_{ik}^{(+)}$ и $d_{ik}^{(-)}$, полученные в разд. 3 (см. Утверждение 3), на основе формулы (7) работы [1]. Вероятности перехода p_{ij} , $i, j \in \{0, 1, \dots, N_0\}$, определяются аналогичным образом при помощи интегральных представлений для вспомогательных вероятностных характеристик $p_{ik}^{(+)}$ и $p_{ik}^{(-)}$, $i, k \in \{0, 1, \dots, N_0\}$, полученных в разд. 3 (см. Утверждение 1), на основе формулы (4) работы [1].

Теперь возникает возможность использовать теорему об аналитическом представлении стационарного показателя эффективности управления полумарковским процессом в форме дробно-линейного интегрального функционала, доказанную в работе [7].

Приведем формулировку указанной теоремы. Заметим предварительно, что вероятностная модель сопровождающего управляемого полумарковского процесса $\zeta(t)$, рассматриваемого в настоящем исследовании, полностью соответствует общей модели управления полумарковским процессом, рассматриваемой в работе [7].

Теорема 1. Стационарный стоимостный показатель средней удельной прибыли, который изначально задается равенством (1), может быть представлен в виде

дробно-линейного интегрального функционала от вероятностных распределений $G_i(u_i)$, $i \in \{0, \dots, N_0\}$. Данный функционал аналитически определяется следующими формулами:

$$I = I(G_0(\cdot), \dots, G_{N_0}(\cdot)) = \int_0^\infty \dots \int_0^\infty A_d(u_0, \dots, u_{N_0}) dG_0(u_0) \dots \dots dG_{N_0}(u_{N_0}) \Big/ \int_0^\infty \dots \int_0^\infty B_d(u_0, \dots, u_{N_0}) dG_0(u_0) \dots \dots dG_{N_0}(u_{N_0}), \quad (5)$$

где

$$\left. \begin{aligned} A_d(u_0, \dots, u_{N_0}) &= \sum_{i=0}^{N_0} d_i(u_i) \hat{D}^i(u_0, \dots, u_{i-1}, u_{i+1}, \dots, u_{N_0}); \\ B_d(u_0, \dots, u_{N_0}) &= \sum_{i=0}^{N_0} T_i(u_i) \hat{D}^i(u_0, \dots, u_{i-1}, u_{i+1}, \dots, u_{N_0}). \end{aligned} \right\} \quad (6)$$

При этом функции $\hat{D}^i(u_0, \dots, u_{i-1}, u_{i+1}, \dots, u_{N_0})$, $i = 0, 1, \dots, N_0$, входящие в правые части соотношений (6), представляются в виде:

$$\hat{D}^i(u_0, \dots, u_{i-1}, u_{i+1}, \dots, u_{N_0}) = (-1)^{N_0+i+2} \sum_{\alpha^{N_0,i}} (-1)^{\delta(\alpha^{N_0,i})} \times \times \hat{D}^i(\alpha^{N_0,i}, u_0, \dots, u_{i-1}, u_{i+1}, \dots, u_{N_0}), \quad (7)$$

где $\alpha^{N_0,i} = (\alpha_0, \dots, \alpha_{i-1}, \alpha_{i+1}, \dots, \alpha_{N_0})$ — произвольная перестановка чисел $(0, \dots, i-1, i+1, \dots, N_0)$, $\delta(\alpha^{N_0,i})$ — число инверсий в перестановке $\alpha^{N_0,i}$,

$$\hat{D}^i(\alpha^{N_0,i}, u_0, \dots, u_{i-1}, u_{i+1}, \dots, u_{N_0}) = \tilde{p}_{0,\alpha_0}(u_0) \dots \tilde{p}_{i-1,\alpha_{i-1}}(u_{i-1}) \tilde{p}_{i+1,\alpha_{i+1}}(u_{i+1}) \dots \dots \tilde{p}_{N_0,\alpha_{N_0}}(u_{N_0}), \quad (8)$$

где

$$\tilde{p}_{k,\alpha_k}(u_k) = \begin{cases} p_{kk}(u_k) - 1, & \alpha_k = k; \\ p_{k\alpha_k}(u_k), & \alpha_k \neq k, \end{cases} \quad k = 0, \dots, i-1, i+1, \dots, N_0.$$

Замечание 2. Функции $d_i(u_i)$ и $T_i(u_i)$, $i = 0, 1, \dots, N_0$, входящие в правые части равенств (6), определяются явными представлениями для характеристик d_i и T_i , $i = 0, 1, \dots, N_0$, полученными в разд. 3. Указанные функции формально совпадают с подынтегральными выражениями

в соответствующих интегральных представлениях для d_i и T_i (см. формулы (3) и (4)) и получают-ся при подстановке в эти представления вырожден-ного вероятностного распределения $G_i(u)$, со-средоточенного в фиксированной точке $u = u_i$, $i = 0, 1, \dots, N_0$. Функции $p_{ij}(u_i)$, входящие в пра-вые части равенств (7) и (8), определяются явными представлениями для вероятностей перехода p_{ij} , полученными в разд. 3. Указанные функции форм-ально совпадают с подынтегральными вырази-ниями в интегральных представлениях для вероят-ностей p_{ij} (см. формулу (2)).

Замечание 3. Функции $d_i(u_i)$ могут быть также определены непосредственно через вспомога-тельные вероятностные характеристики $d_{ik}^{(+)}$ и $d_{ik}^{(-)}$, по-лученные в разд. 3 (см. Утверждение 3). Для этого необходимо использовать соотношение (7) рабо-ты [1] и подставить в соответствующие выраже-ния для вспомогательных характеристик вырож-денные вероятностные распределения $G_i(u)$, как было указано в Замечании 2. Функции $p_{ij}(u_i)$, $i, j \in \{0, 1, \dots, N_0\}$, определяются непосредственно аналогичным образом при помощи интегральных представлений для вспомогательных вероятност-ных характеристик $p_{ik}^{(+)}$ и $p_{ik}^{(-)}$, $i, k \in \{0, 1, \dots, N_0\}$, полученных в разд. 3 (см. Утверждение 1). При этом используется соотношение (4) работы [1] и приме-няется формальный прием, описанный в замеча-нии 2.

5 Теоретическое решение задачи оптимального управления

Согласно теореме 1, стационарный стоимост-ный показатель (1) можно представить в дробно-линейной интегральной форме (5), которая явно выражает его зависимость от вероятностных рас-пределений $G_0(t), \dots, G_{N_0}(t)$, задающих стратегию управления. Таким образом, задача поиска опти-мального управления сводится к экстремальной за-даче для дробно-линейного интегрального функ-ционала.

Проблема нахождения безусловного экстремума для дробно-линейного интегрального функционала подробно изучена в работах П. В. Шнуркова [8, 9]. В частности, доказано, что если основная функ-ция дробно-линейного интегрального функциона-ла, определяемая формулой

$$C_d(u_0, \dots, u_{N_0}) = \frac{A_d(u_0, \dots, u_{N_0})}{B_d(u_0, \dots, u_{N_0})}, \quad (9)$$

достигает глобального экстремума на множестве

$$U^{N_0} = \{(u_0, \dots, u_{N_0}), u_i \in [0; \infty), i \in \{0, \dots, N_0\}\}$$

в некоторой фиксированной точке $u_* = (u_{0*}, \dots, u_{N_0*})$, то решение экстремальной задачи для функционала $I(G_0(\cdot), \dots, G_{N_0}(\cdot))$ на множестве наборов вероятностных распределений $\{G_0(\cdot), \dots, G_{N_0}(\cdot)\}$ существует и достигается на наборе вырожденных распределений $G_0^*(\cdot), G_1^*(\cdot), \dots, G_{N_0}^*(\cdot)$, сосредоточенных в точках $(u_{0*}, \dots, u_{N_0*})$ соответственно. В поставленной задаче основная функция дробно-линейного интегрального функционала задается аналитически формулами (6)–(9). Таким образом, задача оптимального управления запасом в рассматриваемой полумарковской модели получила полное теоретическое решение.

6 Итоги теоретического исследования проблемы оптимального управления

Существенная особенность полученных результатов заключается в том, что для заданных исходных параметров модели может быть найдено конкретное решение задачи оптимального управления численным методом. Такая возможность была впервые реализована в диссертационной работе А. В. Иванова [10], в которой была разработана программа численного нахождения оптимальной стратегии управления запасом для полумарковской модели с периодическим прекращением потребления, теоретическое исследование которой было произведено в работе [7]. Методика этого численного решения была в краткой форме описана в работе [11]. Отметим также, что в заключительной части работы [9] была указана принципиальная возможность создания универсальных программных продуктов, предназначенных для численного решения задач оптимального управления полумарковскими процессами с конечным множеством состояний.

Литература

1. Шнурков П. В., Егоров А. Ю. Разработка и предварительное исследование стохастической полумарковской модели управления запасом непрерывного продукта при постоянно происходящем потреблении // Информатика и её применения, 2018. Т. 12. Вып. 1. С. 109–117.
2. Джемелл В. Управляемые полумарковские процессы // Кибернетический сборник. Новая серия. — М.: Мир, 1967. Вып. 4. С. 97–134.
3. Майн Х., Осаки С. Марковские процессы принятия решений / Пер. с англ. — М.: Наука, 1977. 176 с. (Mine H., Osaki S. Markovian decision processes. — New York, NY, USA: Elsevier, 1970. 142 p.).
4. Luque-Vasquez F., Herndndez-Lerma O. Semi-Markov control models with average costs // Appl. Math., 1999. Vol. 26. No. 3. P. 315–331.
5. Халмош П. Теория меры / Пер. с англ. — М.: Изд-во ИЛ, 1953. 282 с. (Halmos P. R. Measure theory. — Princeton, NJ, USA: Van Nostrand, 1950. 304 p.).
6. Шнурков П. В., Егоров А. Ю. Приложение к статье «Решение проблемы оптимального управления запасом непрерывного продукта при постоянно происходящем потреблении в стохастической полумарковской модели», 2018. 21 с. http://www.ipiran.ru/publications/app_var6.docx.
7. Шнурков П. В., Иванов А. В. Анализ дискретной полумарковской модели управления запасом непрерывного продукта при периодическом прекращении потребления // Дискретная математика, 2014. Т. 26. № 1. С. 143–154.
8. Шнурков П. В. О решении проблемы безусловного экстремума для дробно-линейного интегрального функционала на множестве вероятностных мер // Докл. Академии наук. Сер. Математика, 2016. Т. 470. № 4. С. 387–392.
9. Шнурков П. В., Горшенин А. К., Белоусов В. В. Аналитическое решение задачи оптимального управления полумарковским процессом с конечным множеством состояний // Информатика и её применения, 2016. Т. 10. Вып. 4. С. 72–88.
10. Иванов А. В. Анализ дискретной полумарковской модели управления запасом непрерывного продукта при периодическом прекращении потребления: Дис. . . . канд. физ.-мат. наук. — М.: НИУ ВШЭ, 2014. 120 с.
11. Gorshenin A. K., Belousov V. V., Shnourkoff P. V., Ivanov A. V. Numerical research of the optimal control problem in the semi-Markov inventory model // AIP Conf. Proc., 2015. Vol. 1648. P. 250007-1–250007-4.

Поступила в редакцию 19.02.18

SOLUTION TO THE PROBLEM OF OPTIMAL CONTROL OF A STOCHASTIC SEMI-MARKOV MODEL OF CONTINUOUS SUPPLY OF PRODUCT MANAGEMENT UNDER THE CONDITION OF CONSTANTLY HAPPENING CONSUMPTION

P. V. Shnurkov and A. Y. Egorov

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

Abstract: The solution of the optimal control problem in the semi-Markov model in question is theoretically justified. To achieve this goal, formal analytic transformations of the integral representations obtained by the authors earlier for the basic probabilistic characteristics of the model were carried out. These transformations made it possible to use the theorem on the analytical representation of the stationary value of the management effectiveness of a semi-Markov process in the form of a fractional-linear integral functional. In the sequel, the authors use the general theorem on the extremum of a fractional-linear integral functional, proved by P. V. Shnurkov. This theorem makes it possible to reduce the problem of optimal reserve management to the problem of investigating the global extremum of a given function from a finite number of real nonnegative variables that can be effectively solved in practice using the known numerical methods.

Keywords: inventory management; semi-Markov stochastic process; stationary value index; optimal control of stochastic systems; fractional-linear integral functional

DOI: 10.14357/19922264180212

References

1. Shnurkov, P. V., and A. Y. Egorov. 2018. Razrabotka i predvaritel'noe issledovanie stokhasticheskoy polumarkovskoy modeli upravleniya zapasom nepreryvnogo produkta pri postoyanno proiskhodyashchem potreblenii [Development and preliminary study stochastic semi-Markov model of continuous supply of product management at constantly happening consumption]. *Informatika i ee Primeneniya — Inform. Appl.* 12(1):109–119.
2. Jewell, W. S. 1963. Markov-renewal programming. *Oper. Res.* 11:938–971.
3. Mine, H., and S. Osaki. 1970. *Markovian decision processes*. New York, NY: Elsevier. 142 p.
4. Luque-Vásquez, F., and O. Herndandez-Lerma. 1999. Semi-Markov control models with average costs. *Appl. Math.* 26(3):315–331.
5. Halmos, P. R. 1950. *Measure theory*. Princeton, NJ: Van Nostrand. 304 p.
6. Shnurkov, P. V., and A. Y. Egorov. 2018. Prilozhenie k stat'e "Reshenie problemy optimal'nogo upravleniya zapasom nepreryvnogo produkta pri postoyanno proiskhodyashchem potreblenii v stokhasticheskoy polumarkovskoy modeli" [Appendix to article "Solution to the problem of optimal control stochastic semi-Markov model of continuous supply of product management at constantly happening consumption"]. 21 p. Available at: http://www.ipiran.ru/publications/app_var6.docx (accessed April 25, 2018).
7. Shnurkov, P. V., and A. V. Ivanov. 2015. Analysis of a discrete semi-Markov model of continuous inventory control with periodic interruptions of consumption. *Discrete Math. Appl.* 25(1):59–67.
8. Shnurkov, P. V. 2016. Solution of the unconditional extremum problem for a linear-fractional integral functional on a set of probability measures. *Dokl. Math.* 94(2):550–554.
9. Shnurkov, P. V., A. K. Gorshenin, and V. V. Belousov. 2016. Analiticheskoe reshenie zadachi optimal'nogo upravleniya polumarkovskim protsessom s konechnym mnozhestvom sostoyaniy [Analytical solution of the optimal control task of a semi-Markov process with finite set of states]. *Informatika i ee Primeneniya — Inform. Appl.* 10(4):72–88.
10. Ivanov, A. V. 2014. Analiz diskretnoy polumarkovskoy modeli upravleniya zapasom nepreryvnogo produkta pri periodicheskom prekrashchenii potrebleniya [Analysis of a discrete semi-Markov control model of continuous product inventory in a periodic cessation of consumption]. Moscow: HSE Publishing House. PhD Thesis. 120 p.
11. Gorshenin, A. K., V. V. Belousov, P. V. Shnurkov, and A. V. Ivanov. 2015. Numerical research of the optimal control problem in the semi-Markov inventory model. *AIP Conf. Proc.* 1648:250007-1–250007-4.

Received February 19, 2018

Contributors

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

Egorov Artem Y. (b. 1992) — Master student, Moscow Institute of Electronics and Mathematics, National Research University Higher School of Economics, 34 Tallinskaya Str., Moscow 123458, Russian Federation; ayuegorov@edu.hse.ru

ВЛИЯНИЕ ПЛОТНОСТИ СВЯЗЕЙ НА КЛАСТЕРИЗАЦИЮ И ПОРОГ ПЕРКОЛЯЦИИ ПРИ РАСПРОСТРАНЕНИИ ИНФОРМАЦИИ В СОЦИАЛЬНЫХ СЕТЯХ*

Д. О. Жуков¹, Т. Ю. Хватова², С. А. Лесько³, А. Д. Зальцман⁴

Аннотация: Рассматриваются вопросы применения новых теоретических подходов к описанию процессов передачи и обработки информации в социотехнических системах и сетях социальных связей на основе теории перколяции. Величина порога перколяции случайной сети зависит от ее плотности. В сетях, имеющих случайную структуру, пороги перколяции как в задаче узлов, так и в задаче связей при большой плотности сети достигают величины насыщения, причем величина насыщения порога перколяции в задаче связей больше, чем в задаче узлов. С точки зрения информационного влияния сети, имеющей случайную структуру, увеличение плотности связей оказывается более предпочтительным, чем наличие небольшого числа отдельных «центральных» узлов, имеющих множество связей. В практическом плане полученные результаты могут быть применены в междисциплинарных исследованиях, включая информатику, математическое моделирование и экономику, с привлечением социологических данных для прогнозирования поведения и управления группами людей в сетевых сообществах. Полученные результаты дополняют и расширяют применение методов и подходов, принятых в классической информатике, на описание социальных и социотехнических систем, что может быть полезно для широкого круга исследователей, занимающихся изучением социальных сетевых структур.

Ключевые слова: теория перколяции; структура социальной сети; плотность связей; кластеризация сети; порог перколяции

DOI: 10.14357/19922264180213

1 Введение

Изначально, в классическом представлении, информатика (*фр.* informatique, *англ.* computer science) — это наука о методах и процессах сбора, хранения, обработки, передачи, анализа и оценки информации с применением компьютерных технологий, обеспечивающих возможность ее использования для принятия решений. Однако стремительное развитие сетевых технологий, и в первую очередь Интернета, привело к появлению обширного класса социотехнических систем, ярким представителем которых выступают социальные сети. Наличие человеческого фактора приводит к тому, что для описания протекающих в них процессов уже недостаточно классических методов и моделей, принятых в информатике, для принятия решений требуется использование междисциплинарных подходов и расширение устоявшихся классических представлений, особенно при исследовании процессов распространения и обработки информации

в социальных сетях. В этом плане необходимо говорить о том, что применение методов и моделей информатики должно быть расширено на социотехнические и социальные системы точно так же, как в свое время Норберт Винер определил кибернетику как науку об управлении в живом и неживом [1].

Исследование процессов распространения информации и кластеризации узлов (кластер — группа связанных между собой узлов, выделенных по определенным свойствам или типам) в сетях социальных связей, имеющих случайную топологию, является очень важной и актуальной задачей для экономики, рекламы, маркетинга, социологии, политологии и т. д., что подтверждается достаточно большим числом работ по данной тематике [2–4].

Сеть социальных связей можно определить как совокупность информационных каналов каждого человека, связывающих его с другими членами сообщества, а также средств массовой информации (т. е. не только взаимодействие между членами определенной группы в социальной сети). Иными

* Работа выполнена при финансовой поддержке РФФИ (проект 16-29-09458 офи.м).

¹Московский технологический университет (МИРЭА), zhukov.do@mirea.ru

²Санкт-Петербургский политехнический университет Петра Великого, khvatova.ty@spbstu.ru

³Московский технологический университет (МИРЭА), sergey@testor.ru

⁴Московский технологический университет (МИРЭА), ad.zaltcman@gmail.com

словами, сеть социальных связей является социальной сетью, помещенной в информационную среду (средства массовой информации, книги, газеты, журналы и т. д.).

Средства массовой информации (радио, телевидение, интернет-ресурсы, социальные сети и др.), а также книги, газеты, журналы оказывают существенное влияние на состояние отдельных узлов социальной сети (выбор предпочтений и поведенческие реакции людей), которое может меняться с течением времени за счет как информационной среды, так и взаимодействия пользователей между собой.

В связи с этим возникает ряд вопросов. Во-первых, как может происходить изменение в сети доли узлов, находящихся в том или ином состоянии? Во-вторых, как эти узлы связываются между собой в подгруппы (кластеризация сети)? В-третьих, как информационные процессы зависят от всей сети в целом для различных состояний узлов?

Узлами социальной сети являются отдельные люди, а ребрами — коммуникативные связи, число которых может иметь произвольное значение.

В такой сети распространение информации может одновременно происходить множеством путей через разные узлы сети. Отдельный узел сети может получать от другого узла некоторую информацию (рекламные предложения, идеи, политические взгляды, профессиональные сведения, мнения и др.) и передавать ее другим узлам, если имеет согласную с поступившей информацией позицию (является активным проводником), или блокирует ее в противном случае.

Для моделирования и анализа информационных процессов, протекающих в социальных сетях со случайной структурой, возможно применение методов теории перколяции [5–7], которая может позволить, например, ответить на следующие вопросы:

1. Как происходит кластеризация сети на группы связанных между собой определенными взглядами людей в зависимости, например, от среднего числа связей на узел?
2. При какой доле людей (узлов сети) с определенными взглядами может создаваться условие для беспрепятственного распространения этих взглядов между двумя любыми произвольно выбранными узлами (протекание или перколяция)?

Наиболее распространенными задачами теории перколяции являются *решеточные задачи*: задача связей и задача узлов. В задаче связей ищут ответ на вопрос: какую долю связей нужно удалить (перерезать), чтобы сетка распалась на две части?

В задаче узлов блокируют узлы (удаляют узел, перерезая все входящие в узел связи) и ищут, при какой доле заблокированных узлов сетка распадется.

Кластеризация социальной сети. Если переход любого узла социальной сети (индивидуума) из одного состояния в другое рассматривать как случайный процесс (с некоторой вероятностью перехода, определяемой множеством случайных факторов, в том числе зависящих от влияния средств массовой информации), то вероятность перехода будет влиять на средний размер кластера пользователей сети (группы напрямую связанных между собой узлов).

Перколяция в социальной сети. В теории перколяции доля проводящих (неблокированных) узлов, при которой возникает проводимость между двумя различными произвольно выбранными узлами сети, называется порогом перколяции (протекания).

Доля узлов социальной сети, находящихся в том или ином состоянии, может быть выявлена путем социологических опросов, что предоставляет возможность оценить, насколько социальная сеть близка к порогу перколяции, а также управлять ее состояниями.

2 Обзор существующих моделей описания характеристик и анализа структуры социальных сетей

Наиболее часто для проведения исследований операций и процессов в сетевых структурах сегодня используются готовые средства анализа, например инструменты анализа социальных сетей (SNA — social network analysis), позволяющие получить количественные характеристики параметров графа сети, таких как «центральность», «промежуточность», «плотность» (среднее число связей, приходящихся на один узел сети) [8]. Центральность характеризует степень влияния данного узла на всю сеть. Промежуточность характеризует степень включенности объекта в маршруты связей между другими участниками сети. Промежуточность показывает, насколько часто данный узел встречается на кратчайших путях между другими узлами.

Использование готовых инструментов имеет как свои преимущества, так и ряд недостатков. К преимуществам можно отнести то, что готовые инструменты социального сетевого анализа позволяют сравнивать между собой однотипные сети по количественным характеристикам. Основные недостатки заключаются в том, что они не позволяют создавать новые, более информативные модели.

В работе [9] были изучены статистические свойства реальных социальных сетей работников домохозяйств и их мера фрагментации после удаления некоторых долей узлов или ссылок из сети.

В работе [10] рассматривается взаимосвязь перколяционного перехода и выживаемости узлов в сложной сети с 3 млн связей, построенная вокруг примерно 300 тыс. фирм (узлы). Характерная особенность этих реальных сетей заключается в том, что они являются масштабируемыми и степень их масштабируемости асимптотически следует степенному закону. Эта функция подразумевает, что каждая из таких сетей состоит из нескольких крупных центральных узлов с тысячами связей, множеством промежуточных узлов и еще большим числом очень мелких узлов с несколькими связями [11]. Масштабируемые сети кластеризуются на отдельные блоки при достаточно высокой плотности связей, если узлы удаляются в порядке убывания степени центральности [12]. При случайном удалении узлов и связей масштабируемые сети плохо кластеризуются, даже при очень низкой плотности узлов, и перколяционные переходы не отмечаются. В данном исследовании с помощью точного численного расчета [10] доказано существование перколяционного перехода в сложных сетях при случайном удалении узлов, когда плотность сети очень низкая, но не нулевая.

3 Постановка задачи и методика проведения исследований

Несмотря на существенный прогресс в исследовании информационных процессов, протекающих в социальных сетях, и использование для этого теории перколяции, еще очень многие задачи остаются нерешенными. Большинство исследователей уделяют большое внимание изучению степени промежуточности, плотности сети, средней длине пути (близости), центральности и т. д. Вместе с тем никто из исследователей не обращал еще целенаправленно внимание на изучение таких вопросов, как влияние плотности (среднего числа связей в расчете на один узел) сети на ее кластеризацию и величину порога перколяции (как в задаче узлов, так и в задаче связей) как в масштабируемых, так и в случайных сетях. На взгляд авторов, необходимо учитывать в целом всю совокупность свойств сети, которые определяют порог перколяции и кластеризацию; значение имеют как узлы, у которых много связей и они являются значимыми или центральными, так и с малым числом связей.

Следует отметить, что существующие программные инструменты информационного анализа со-

циальных сетей (SNA) в данном случае не могут быть использованы, поскольку они не позволяют создавать случайные сети с произвольной плотностью связей и изучать их кластеризацию и перколяцию при блокировании узлов.

Для изучения случайных сетей с множеством связей аналитических моделей описания перколяционных процессов не существует, и их исследование возможно только методами численного моделирования [13, 14] с использованием специально разработанного программного обеспечения. Для этого необходимо сначала построить структурную модель социальной сети, состоящую из большого числа (например, в данном исследовании — 1 млн) узлов. Затем выбрать пару произвольных узлов и с помощью методов численного моделирования определить, при какой доле неблокированных узлов (для задачи узлов) в рассматриваемой сети появляется свободный путь между рассматриваемыми узлами (или, наоборот, исчезает при блокировании). Затем аналогичным образом эта процедура проводится для других произвольных пар узлов. После этого со статистическим усреднением результатов по отдельным экспериментам необходимо провести определение среднего значения порога перколяции по всем рассматриваемым парам узлов [14, 15]. В задаче связей используется практически такой же алгоритм исследования, однако блокируются не узлы, а связи.

4 Исследование кластеризации узлов социальных сетей и достижения порога перколяции

4.1 Перколяция (протекание) информации в сетях со случайной структурой

Задача связей (разрываются связи между узлами) при определении порогов перколяции в сети, имеющей случайную структуру, была решена авторами ранее в работе [15]. Результаты проведенного численного моделирования [14, 15] зависимости средней вероятности невозможности передачи информации между двумя произвольно выбранными узлами случайной сети с множеством путей между узлами и различным средним числом связей на один узел от вероятности блокирования (разрыва) связи представлены на рис. 1.

Кривая 1 построена для сети, у которой среднее число связей в расчете на один узел составляет 3,99, кривая 2 — 5,99, кривая 3 — 7,97, кривая 4 —

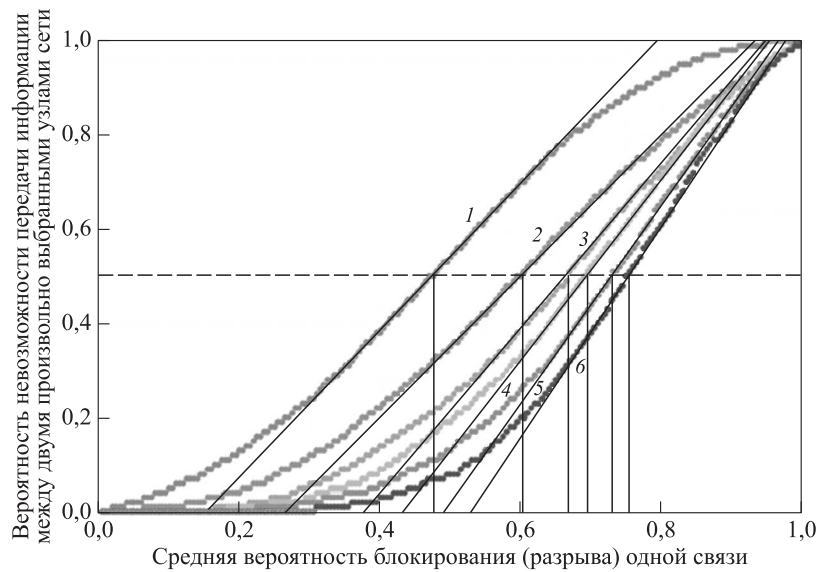


Рис. 1 Зависимость вероятности невозможности передачи информации между двумя произвольно выбранными узлами случайной сети от вероятности блокирования (разрыва) одной связи

9,93, кривая 5 — 13,86 и кривая 6 — 15,79 связей. Основной задачей исследования является определение порогов перколяции для сетей с различным средним числом связей в расчете на один узел, но в задаче узлов при этом удаляются узлы, а в задаче связей — связи. Можно определить линейные участки в центре кривых 1–6 и экстраполировать их до пересечения с осью абсцисс (см. рис. 1), значения величин которых можно условно принять за величину нижний границы порога перколяции данной сети.

На рис. 2 представлена зависимость величины нижней границы порога перколяции от среднего числа связей на один узел данной сети, найденная по описанной выше методике.

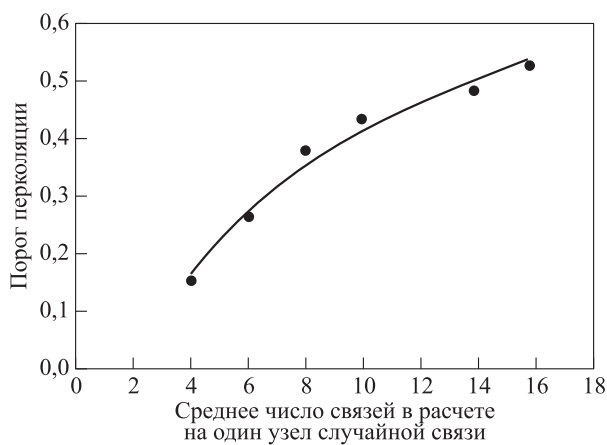


Рис. 2 Зависимость порога перколяции в случайной сети от среднего числа связей (плотности сети) на один ее узел в задаче связей

Представленная на рис. 2 зависимость хорошо линеаризуется в координатах $\ln P(x)$ от $z = 1/x$ (натуральный логарифм порога перколяции — величина, обратная среднему числу связей x , приходящихся на один узел) и позволяет получить линейную зависимость $y = -6,581z - 0,203$ со значением коэффициента корреляции, равным 0,992.

Здесь говорится о пороге перколяции как невозможности передачи информации при некоторой вероятности разрыва связи. Например, если в данном случае для сети с плотностью связей 3,99 порог перколяции равен 0,16, то это значит, что если разорвать меньше 16% связей, то сеть информацию передает, а если 16% или больше, то передача информации в сети в целом прекращается. Но это будет нижняя оценка. Заметим, что другие методы будут давать большее значение величины порога (верхнюю оценку). Например, если за основу взять вероятность невозможности передачи информации, равную 0,5, которая показана на рис. 1 горизонтальной штриховой линией, то получим следующие значения порогов перколяции: для среднего числа связей 3,99 — 0,46; для 5,99 — 0,60; для 7,97 — 0,66; для 9,93 — 0,69; для 13,86 — 0,73; для 15,79 — 0,75.

Задача нахождения порога перколяции не для разорванных связей, а для заблокированных узлов в случайной сети была решена в работе [13]. Результаты численного моделирования нахождения порога перколяции для случайных сетей с множеством путей между узлами и различным средним числом связей на один узел также хорошо линеаризуются в координатах $\ln P(x)$ от $z = 1/x$ (нату-

ральный логарифм порога перколяции — величина, обратная среднему числу связей x (плотность сети), приходящихся на один узел) и позволяют получить линейную зависимость $y = 4,39z - 2,41$ со значением коэффициента корреляции, равным 0,95.

Здесь речь идет о пороге перколяции уже как о возможности передачи информации при некоторой вероятности активации узлов (узел становится проводящим). Например, если в данном случае для сети с плотностью связей 4,70 порог перколяции равен 0,27, то это значит, что если будет активировано меньше 27% узлов, то сеть информацию не передаст, а если 27% или больше, то передача информации в сети в целом возникает.

Проанализируем полученные данные. Выберем для задачи узлов и задачи связей в качестве примера четыре значения плотности сети (среднего числа связей, приходящихся на один узел): 5, 10, 50 и 100. Для задачи узлов используем уравнение $y = 4,39z - 2,41$ и при $z = 1/5 = 0,2$ получим $y = -1,532$, и величина порога перколяции будет равна 0,22 (в данном случае доля проводящих узлов, при которой появляется проводимость). Расчет по уравнению $y = -6,581z - 0,203$ для задачи связей при $z = 0,2$ дает величину доли разорванных связей, при которой исчезает проводимость всей сети в целом, равную 0,22 (т.е., если разорвем 22% связей и более, проводимость исчезнет).

Полученные результаты позволяют сделать ряд важных выводов для перколяционных процессов в случайных сетях. Они приводятся в разд. 5.

4.2 Кластеризация социальной сети

Поскольку проводимость узлов в большей степени определяет решение информационных задач всей сетью в целом, то будем рассматривать задачу узлов. Переход любого узла социальной сети из одного состояния в другое можно рассматривать как случайный процесс (с некоторой вероятностью перехода, определяемой множеством случайных факторов), и эта вероятность должна влиять на средний размер кластера (группа напрямую связанных между собой узлов). В проведенных авторами численных экспериментах [13, 15] было исследовано влияние вероятности перехода узла из одного состояния в другое (например, желания голосовать «за» или «против») на средний размер кластера пользователей, находящихся в данном состоянии (в долях от общего числа всех индивидуумов) в случайных сетях с различным средним числом связей на один узел и общим числом узлов, равном 1 млн. Можно рассмотреть два предельных случая: первый — небольшое среднее число связей на один узел социальной сети; второй — большое среднее

число связей. Исследования показали, что с ростом среднего числа связей при фиксированной вероятности воздействия размер кластера увеличивается. Аналогичная ситуация наблюдается и при большом среднем числе связей на один узел случайной сети.

С ростом среднего числа связей при фиксированной вероятности воздействия размер кластера увеличивается, а скорость роста кластеризации узлов, находящихся в данном состоянии, наиболее сильно увеличивается в области значений вероятности перехода единичных узлов от 0,4 до 0,6, а при малых и высоких значениях возрастает не так сильно.

5 Выводы

1. Как в задаче связей, так и в задаче узлов величина порога перколяции случайной сети зависит от ее плотности (среднего числа узлов, приходящихся на один узел).
2. В сетях, имеющих случайную структуру, пороги перколяции как в задаче узлов, так и в задаче связей при большой плотности сети (среднее число связей на один узел) практически достигают величины насыщения (0,24 для задачи связей и 0,10 для задачи узлов) и затем слабо зависят от нее. Величина насыщения порога перколяции в задаче связей почти в 2,5 раза больше, чем в задаче узлов.
3. С точки зрения создания проводимости случайной сети в целом образование проводящих связей в задаче связей менее эффективно, чем образование проводящих узлов в задаче узлов (например, при плотности сети, равной 5, для возникновения проводимости необходимо иметь долю проводящих узлов, равную 0,22, в то время как доля проводящих связей должна быть равной 0,78).
4. С ростом среднего числа связей (плотности сети) при фиксированной вероятности воздействия размер кластера увеличивается, а скорость роста кластеризации узлов, находящихся в данном состоянии, наиболее сильно увеличивается в области значений вероятности перехода единичных узлов от 0,4 до 0,6, а при малых и высоких значениях возрастает не так сильно.
5. С точки зрения информационного влияния сети, имеющей случайную структуру, увеличение плотности связей оказывает большее влияние, чем наличие отдельных «центральных» узлов, имеющих множество связей.
6. В практическом плане полученные результаты могут быть применены в междисциплинар-

ных исследованиях, включая информатику, математическое моделирование и экономику, с привлечением социологических данных для прогнозирования поведения и управления группами людей в сетевых сообществах. Кроме того, полученные результаты дополняют и расширяют применение методов и подходов, принятых в классической информатике, на описание социальных и социотехнических систем, что может быть полезно для широкого круга исследователей, занимающихся изучением социальных сетевых структур.

Литература

1. Винер Н. Кибернетика, или Управление и связь в животном и машине // Пер. с англ. — 2-е изд. — М.: Наука, 1983. 344 с. (*Wiener N.* Cybernetics: Or control and communication in the animal and the machine. — 2nd ed. — MIT Press, 1961. 212 p.)
2. Баканова С. А., Силкина Г. Ю. Процессы распространения знаний в параметризованной сети информационных обменов // Научно-технические ведомости Санкт-Петербургского государственного политехнического университета. Экономические науки, 2015. № 2(216). С. 133–146.
3. Сулимов П. А. Методы машинного обучения для предсказания распространения инфекции в сети // Вестн. НГУЭУ, 2016. № 1. С. 285–306.
4. Торопов Б. А. Модель независимых каскадов распространения репоста в онлайн-социальной сети // Кибернетика и программирование, 2016. № 5. С. 61–67.
5. Тарасевич Ю. Ю. Перколяция: теория, приложения, алгоритмы. — М.: Эдиториал УРСС, 2002. 112 с.
6. Лесько С. А., Жуков Д. О., Самойло И. В. Математическое моделирование перколяционных процессов передачи данных и потери работоспособности в информационно-вычислительных сетях с 2D и 3D регулярной и случайной структурой // Качество. Инновации. Образование, 2013. № 6(97). С. 42–50.
7. Лесько С. А., Жуков Д. О., Самойло И. В., Брукс Д. У. Алгоритмы построения сетей и моделирования потери их работоспособности в результате кластеризации заблокированных узлов // Качество. Инновации. Образование, 2013. № 12(103). С. 82–87.
8. Павлековская И. В. Применение метода анализа социальных сетей в моделировании процессов распространения информации и знаний в организации // Научно-техническая информация. Сер. 2: Информационные процессы и системы, 2007. № 3. С. 30–36.
9. Chen Y., Paul G., Cohen R., Yavlin S., Borgatti S. P., Liljeros F., Stanley H. E. Percolation theory and fragmentation measures in social networks // Physica A, 2006. Vol. 378. No. 1. P. 11–19.
10. Kawamoto H., Takayasu H., Jensen H. J., Takayasu M. Precise calculation of a bond percolation transition and survival rates of nodes in a complex network // PLoS One, 2015. <https://doi.org/10.1371/journal.pone.0119979>.
11. Barabási A. L., Albert R. Emergence of scaling in random networks // Science, 1999. Vol. 286. P. 509–512.
12. Albert R., Jeong H., Barabási A. L. Error and attack tolerance of complex networks // Nature, 2000. Vol. 406. P. 378–382.
13. Zhukov D., Lesko S. Percolation models of information dissemination in social networks // IEEE Conference (International) on Smart City/SocialCom/SustainCom together with DataCom Proceedings. — IEEE, 2015. P. 213–216.
14. Block M., Khvatova T., Zhukov D., Lesko S. Studying the structural topology of the knowledge sharing network // 11th European Conference on Management, Leadership and Governance Proceedings. — Lisbon, Portugal: Academic Conferences and Publishing International Ltd., 2015. P. 20–27.
15. Khvatova T., Block M., Zhukov D., Lesko S. How to measure trust: The percolation model applied to intra-organisational knowledge sharing networks // J. Knowl. Manag., 2016. Vol. 20. No. 5. P. 918–935.

Поступила в редакцию 04.07.17

THE INFLUENCE OF THE CONNECTIONS' DENSITY ON CLUSTERIZATION AND PERCOLATION THRESHOLD DURING INFORMATION DISTRIBUTION IN SOCIAL NETWORKS

D. O. Zhukov¹, T. Yu. Khvatova², S. A. Lesko¹, and A. D. Zaltsman¹

¹Moscow Technological University (MIREA), 78 Vernadskogo Ave., Moscow 119454, Russian Federation

²Peter the Great St. Petersburg Polytechnic University, 29 Polytechnicheskaya Str., St. Petersburg 195251, Russian Federation

Abstract: The paper is focused on applying new theoretical approaches to describing the processes of information transmission and transformation in sociotechnical systems and in social networks based on the percolation theory. Percolation threshold of a random network depends on its density. In networks with random structure, in both the

task of bonds and the task of nodes, percolation thresholds reach saturation when the network's density is high. The saturation value of a percolation threshold is higher in the task of bonds. From the point of information influence of a random network, increasing the average connection's density within the network turns out to be more preferable than fostering a small number of separate 'central nodes' with numerous connections. The results obtained in this study can be applied in interdisciplinary research in such areas as informatics, mathematic modeling, and economics involving certain sociological survey data for forecasting behavior and managing groups of individuals in network communities. This research enhances and enlarges the scope of methods and approaches applied in classic informatics for describing social and sociotechnical systems, which can be useful for a wide range of researchers engaged into studying social network structures.

Keywords: percolation theory; social network structure; connections' density; network clusterisation; percolation threshold

DOI: 10.14357/19922264180213

Acknowledgments

This research was performed with financial support of the Russian Foundation for Basic Research (project No. 16-29-09458 ofi_m) "Developing percolation topological models for describing virtual social systems, their participants' clusterization into groups according to their views, stochastic dynamics of influence distribution, and for managing transitions."

References

1. Wiener, N. 1961. *Cybernetics: Or control and communication in the animal and the machine*. 2nd ed. MIT Press. 212 p.
2. Bakanova, S. A., and G. Yu. Silkina. 2015. Knowledge dissemination process in parametrized networks of enterprises. *St. Petersburg State Polytechnical University J. Economics* 2(216):133–146.
3. Sulimov, P. A. 2016. Metody mashinnogo obucheniya dlya predskazaniya rasprostraneniya infektsii v seti [Methods of machine learning for forecasting an infection distribution within a network]. *Vestnik NGUEU* [NGUEU J.] 1:285–306.
4. Toropov, B. A. 2016. Model' nezavisimyykh kaskadov rasprostraneniya reposta v onlaynovoy sotsial'noy seti [The model of independent cascades of a repost distribution in an online social network]. *Kibernetika i programirovanie* [Cybernetics and Programming] 5:61–67.
5. Tarasevich, Yu. 2002. *Perkolyatsii: teoriya, prilozheniya, algoritmy* [Percolations: Theory, applications, algorithms]. Moscow: Editorial URSS. 112 p.
6. Les'ko, S., D. Zhukov, and I. Samoylo. 2013. Matematicheskoe modelirovanie perkolyatsionnykh protsessov peredachi dannykh i poteri rabotosposobnosti v informatsionno-vychislitel'nykh setyakh s 2D i 3D regul'arnoy i sluchaynoy strukturoy [Mathematic modeling of percolation processes of data transmission and operability loss in informational-computational networks with 2D and 3D regular and random structure]. *Kachestvo. Innovatsii. Obrazovanie* [Quality. Innovations. Education] 11(97):42–50.
7. Les'ko, S., D. Zhukov, I. Samoylo, and D. Bruks. 2013b. Algoritmy postroeniya setey i modelirovaniya poteri ikh rabotosposobnosti v rezul'tate klasterizatsii blokirovannykh uzlov [Algorithms of network construction and modeling of their operability loss as a result of blocked nodes clusterization]. *Kachestvo. Innovatsii. Obrazovanie* [Quality. Innovations. Education] 12(103):25–33.
8. Pavlekovskaya, I. V. 2007. The use of social network analysis in modeling the organizational processes of information and knowledge circulation. *Automatic Documentation Math. Linguistics* 41(2):65–72.
9. Chen, Y., G. Paul, R. Cohen, S. Yavlin, S. P. Borgatti, F. Liljeros, and H. E. Stanley. 2006. Percolation theory and fragmentation measures in social networks. *Physica A* 378(1):11–19.
10. Kawamoto, H., H. Takayasu, H. J. Jensen, and M. Takayasu. 2015. Precise calculation of a bond percolation transition and survival rates of nodes in a complex network. *PLoS One*. Available at: <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0119979> (accessed December 16, 2017).
11. Barabási, A. L., and R. Albert. 1999. Emergence of scaling in random networks. *Science* 286:509–512.
12. Albert, R., H. Jeong, and A. L. Barabási. 2000. Error and attack tolerance of complex networks. *Nature* 406:378–382.
13. Zhukov, D., and S. Lesko. 2015. Percolation models of information dissemination in social networks. *IEEE Conference (International) on Smart City/SocialCom/SustainCom together with DataCom Proceedings*. IEEE. 213–216.
14. Block, M., T. Khvatova, D. Zhukov, and S. Lesko. 2015. Studying the structural topology of the knowledge sharing network. *11th European Conference on Management, Leadership and Governance Proceedings*. Lisbon, Portugal: Academic Conferences and Publishing International Ltd. 20–27.
15. Khvatova, T., M. Block, D. Zhukov, and S. Lesko. 2016. How to measure trust: The percolation model applied to intra-organisational knowledge sharing networks. *J. Knowl. Manag.* 20(5):918–935.

Received July 4, 2017

Contributors

Zhukov Dmitry O. (b. 1965) — Doctor of Science in technology, professor, Head of Department, Moscow Technological University (MIREA), 78 Vernadskogo Ave., Moscow 119454, Russian Federation; zhukov_do@mirea.ru

Khvatova Tatiana Yu. (b. 1971) — Doctor of Science in economics, professor, Peter the Great St. Petersburg Polytechnic University, 29 Polytechnicheskaya Str., St. Petersburg 195251, Russian Federation; khvatova.ty@spbstu.ru

Lesko Sergey A. (b. 1983) — Candidate of Science (PhD) in technology, associate professor, Moscow Technological University (MIREA), 78 Vernadskogo Ave., Moscow 119454, Russian Federation; sergey@testor.ru

Zaltsman Anastasia D. (b. 1989) — lecturer, Moscow Technological University (MIREA), 78 Vernadskogo Ave., Moscow 119454, Russian Federation; ad.zaltcman@gmail.com

ДИСКРЕТНЫЙ АНАЛИЗ В СИНТАКСИЧЕСКОМ АНАЛИЗЕ

Я. М. Мирзабеков¹, Ш. Б. Шихиев²

Аннотация: Дано неформальное определение синтаксиса в терминах дискретной математики и теории графов. Приведен алгоритм выделения семантически замкнутых (самодостаточных по значению) фрагментов предложения. Отсутствие формального определения понятия семантики слова и сочетания словоформ затрудняет компьютерную обработку текста, в частности процесс его сегментации. Показано, каким образом можно классифицировать выражения по семантическому признаку, используя их синтаксические особенности. Классификация выражений по тому, каким вопросам они отвечают, является самым доступным и упрощенным способом группировки выражений по семантическому признаку. Описан алгоритм, распознающий указатели места, т. е. выражения, отвечающие на вопрос «где?». На конкретных примерах рассмотрены задача анализа выражений и обратная ей задача синтеза — прикладная задача распознавания выражений.

Ключевые слова: естественный язык; дискретная математика; теория графов; синтаксис; словоформа; морфологический параметр; согласованные определения; несогласованные определения; лексика; семантика

DOI: 10.14357/19922264180214

1 Введение

Рассматриваемая задача имеет целью предложить новые подходы к решению задачи *синтаксического анализа*, и, в частности, предлагается алгоритм решения известных в литературе задач Text Mining по *извлечению понятий из текста* [1]. Предлагаемый в статье алгоритм отличается от известных авторам работ тем, что он опирается на *конструктивную теорию синтаксиса*, которая построена как раздел дискретного анализа и в которой *синтаксис* и его элементы (*словосочетание*, *выражение*, *предложение* и т. д.) имеют формальные определения [2].

Синтаксический анализ предложения на компьютере давно является предметом исследования различных инициативных групп, лабораторий и академических институтов. Исторически так сложилось, что под *синтаксическим анализом* предложения подразумевается его и *морфологический*, и *синтаксический*, и *семантический* анализ (и словоформ, и словосочетаний, и их значений). *Морфология с синтаксисом*, с одной стороны, и *семантика*, с другой стороны, изучают предметы несовместимой природы — мир материальных вещей (слова и их сочетания, план выражения) и мир представлений и мыслей (план содержания). Язык есть только *отображение (форма, по выражению Ф. де Соссюра) этих двух субстанций*. Попытка сходу объять их в рамках одной теории, а тем более в рамках одного алгоритма изначально была рискованным предприятием.

Хорошо известны и те, кто стоял у истоков «*синтаксического и глубинного анализа*» (Н. Хомский, Д. Слобин и др.), и труды российских ученых по созданию *синтаксических анализаторов* предложений русского языка.

В трудах В. А. Тузова (СПбГУ), Р. Г. Пиотровского (СПбГПУ им. А. И. Герцена), сотрудников МГУ (А. С. Старостин, М. Г. Мальковский и др.) и РАН (И. А. Мельчук, Ю. Д. Апресян и др.) и особенно в работах ежегодного сборника «Диалог-21» отражены достижения и тенденции развития компьютерных наук. А в работах А. О. Казенникова, И. П. Кузнецова и Н. В. Сомина, а также А. С. Старостина, М. Г. Мальковского и Н. В. Арефьева действительно используются методы дискретного анализа при построении анализаторов. И задача «извлечения из текста понятий» определенного типа [3] также известна давно и решается различными эвристическими методами.

На особенностях и имеющихся серьезных достижениях в построении синтаксических анализаторов не будем останавливаться по той простой причине, что в данной статье под *синтаксическим анализом* подразумевается нечто отличное от традиционного определения этого понятия.

На примере урезанного (упрощенного) синтаксиса русского языка приводится описание *формального синтаксиса* (без участия семантики в конструкциях *синтаксиса*). В рамках *формального синтаксиса* формулируется и решается немало

¹Дагестанский государственный университет, yash831@mail.ru

²Дагестанский государственный университет, sh_sh_b51@mail.ru

интересных задач лингвистики. На примере построения универсального алгоритма решения задачи *извлечения понятий из текста* показана одна из возможностей *формального синтаксиса*.

Формальный синтаксис — понятие дискретной математики. Как структура, определяемая лексикой и набором словосочетаний (упорядоченных пар словоформ), она открыта в том смысле, что может быть дополнена не только новыми словами и словосочетаниями, но и тем, что *элементы синтаксиса* могут быть инкапсулированы семантикой. Последнее обстоятельство дает основание и возможность для автономного изучения *синтаксиса* без услуг семантики и ощущение легкости формулировок задач и алгоритмов их решения.

Как известно, «из текстов извлекаются отдельные факты с помощью лексического анализа», а далее, на втором шаге, «за этим следует синтаксический разбор» [1]. Под *фактом*, видимо, следует подразумевать категорию семантики (например, *понятие* или *фрагмент знания*), выраженную одним словом или выражением из нескольких словоформ. Когда носителем языка, т. е. *слова* и его *значения*, является человек, тогда грамматика и языкознание опираются на *соответствие* между *словом* и его *значением*, имеющимся в сознании человека. В задачах компьютерной обработки текста разработчики соответствующих алгоритмов, как правило, пытаются в самом алгоритме реализовать свои интуитивные представления о *значении слова*.

Основной целью статьи является привлечение внимания читателя к следующим двум методологическим проблемам, имеющим место в компьютерной лингвистике, и путям их решения.

1. Возможность замены эвристических алгоритмов на точные алгоритмы решения задач выявления типа элементов синтаксиса (например, является ли предложением данная строка символов?).
2. Приемы построения *отображения* между двумя множествами: синтаксически правильных выражений и их значений (поэтому в статье в качестве первого множества рассматривается множество выражений с явной семантикой; если быть точнее, то выражений, отвечающих на вопрос *где?*).

Не вдаваясь в тонкости проблемы семантики языка, отметим: значение слова — это то, что просыпается в памяти человека при восприятии этого слова. А значение слова, описанное в толковом (семантическом) словаре, есть сопоставление одному слову некоторой другой группы слов и их сочетаний.

Даже тот, кто воспринимает слово *халва*, затрудняется сказать, чего больше для него в значении этого слова: ощущения сладости или пяти звуков, образующих это слово. Непосредственная связь между словом и его значением давно используется в решении задач лингвистики.

2 Формулировка задачи

В *синтаксическом анализе* предложения самым трудоемким является процесс определения *морфологических параметров* словоформы. Следующей по трудоемкости процедурой является *сегментация членов* предложения. Понятие *сегментации* нуждается в уточнении.

Предполагается, что предложение имеет древовидную структуру. Например, *предложению-последовательности* (словоформ)

**маленькая дочка соседа пошла на родник
за холодной водой** (1)

соответствует *предложение-дерево*, представленное на рис. 1.

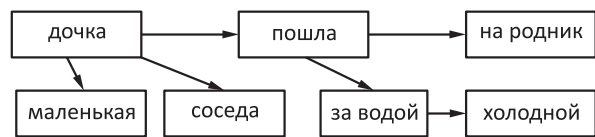


Рис. 1 Предложение-дерево T_1

Каждая ветвь *корневого дерева* [2, 4] на рис. 1 (обозначим его через T_1) также является корневым деревом. Например, ветвь сказуемого из предложения (1) — показана на рис. 2. У дерева несколько ветвей, вершины ветви предложения-дерева образуют *сегмент* в этом дереве.

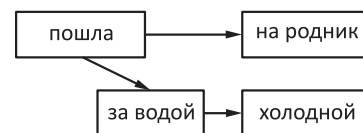


Рис. 2 Ветвь «пошла» дерева T_1

Запишем предложение-дерево T_1 в скобочном виде:

**дочка (маленькая, соседа,
пошла (на родник, за водой (холодной)))**. (2)

Трансформация двумерной структуры (2) в линейную цепь (1) интуитивно выполняется нашей языковой способностью. В конструктивной теории

языка, в частности в компьютерной лингвистике, этот процесс должен быть описан таким образом, чтобы иметь возможность найти соответствие между сегментами предложения-дерева (2) и *фрагментами* предложения-последовательности (1).

Фрагментом в последовательности (1) называется любой ее диапазон — непрерывная часть цепочки словоформ (1).

Фрагментация сегментов, т. е. трансформация *предложения-дерева* в *предложение-последовательность* таким образом, чтобы каждый сегмент первого был фрагментом во втором, — принципиальной важности задача в анализе предложения. Решается эта задача несложно — простым односторонним обходом дерева [4]. Однако правила трансформации, которыми пользуется наша языковая способность, частенько отступают от пути обхода дерева. Например, левосторонний обход дерева T_1 таков:

дочка маленькая соседа пошла

за водой холодной на родник

Рассматриваемая задача имеет непосредственное отношение к правилам трансформации вершин предложения-дерева во фрагменты предложения-последовательности.

Пусть (v, w) — некоторая дуга в предложении-дерева. Известно, что слово w доопределяет (уточняет) значение слова v . Следовательно, и вся ветвь с корнем в w присутствует в предложении исключительно для уточнения определенного семантического признака слова v . Признаком слова v может быть и *вес*, и *цвет*, и *местонахождение* и т. д. предмета с именем v . Такие сегменты назовем *описателями признаков* вещи с именем v , или *описателями семантических признаков* слова v .

Также известно, что *описатель семантического признака* всегда образует фрагмент в предложении. Например, в предложении «**Фермер отправил (на пять километров двести метров) (в двух контейнерах и в семи ящиках) (три тонны пятьсот килограмм) зерна**» в круглых скобках выделены три выражения, каждое из которых описывает один из трех признаков (глагола *отправил* и имени *зерна*). При любом нарушении фрагментарности этих сегментов их значения изменятся.

Используя это свойство предложения, можно разработать алгоритмы для выделения в предложении отдельных его фрагментов — описателей семантических признаков имен существительных (ИС) и глаголов, что существенно упростит анализ предложения. (Напомним, что неуверенность во *фрагментарности сегментов* существенно усложняет работу синтаксических анализаторов.)

Собирать дерево из нескольких его ветвей намного проще, нежели из многочисленных его вер-

шин. (Напомним также, анализ предложения означает: по заданному предложению-последовательности построить соответствующее ему предложение-дерево [4].)

Для решения поставленной задачи необходимы некоторые определения и обозначения из синтаксиса естественного языка (ЕЯ). В данном случае в качестве синтаксиса используется синтаксис русского языка.

3 Синтаксис

Задачи компьютерной лингвистики предполагают компьютерный морфологический анализ словоформ и синтаксический анализ предложения на *естественном языке*. Морфология и синтаксис письменного языка давно рассматриваются разработчиками синтаксических анализаторов и программ-переводчиков как дискретные системы с заданными множествами элементов и операций над ними.

Компьютерный синтаксический словарь (вернее было бы называть его морфологическим словарем) является обязательной частью базы данных в синтаксических анализаторах. У этих словарей такое же предназначение, как у обычных словообразовательных словарей: анализ (распознавание) и синтез (построение) словоформ. Построение компьютерного словаря, разработка алгоритмов анализа и синтеза словоформ по данному словарю является прикладной задачей дискретного анализа [2].

Под анализом словоформы, например «*ящичку*», подразумевается поиск исходной формы этого слова и *морфологических параметров* его: «*ящик 01 11 21 33*», где 01 — код ИС; 11 — код единственного числа; 33 — код дательного падежа.

Далее будем пользоваться следующей кодировкой частей речи и морфологических категорий: 01 — ИС; 02 — ИП (полная форма имени прилагательного); 10, 20 и 30 — *категории рода, числа и падежа* соответственно. Значения категорий также пронумерованы:

$$\begin{aligned} \text{род} &= (\text{мужской, средний, женский}) = \\ &= 10 = (11, 12, 13); \\ \text{число} &= (\text{единственное, множественное}) = \\ &= 20 = (21, 22); \\ \text{падеж} &= (\text{именительный, родительный,} \\ &\text{дательный, винительный, творительный,} \\ &\text{предложный}) = 30 = (31, 32, 33, 34, 35, 36). \end{aligned}$$

Нетрудно заметить, что значения категорий суть унарные операции, определенные на ИС и ИП. Действительно,

$$\begin{aligned} 22(\text{большой}) &= \text{большие}; \\ 35(\text{большой}) &= \text{большим}; \\ 33(\text{ящик}) &= \text{ящику}. \end{aligned}$$

Композиции этих операций образуют новые операции. Например:

$$\begin{aligned} 22 \cdot 35(\text{большой}) &= \text{большими}; \\ 22 \cdot 33(\text{ящик}) &= \text{ящикам}. \end{aligned}$$

Таким образом, на ИС и ИП определены по $3 \cdot 2 \cdot 6 = 36$ операций, в виде композиции трех операций типа 10, 20 и 30. (Операции 11, 12 и 13 не меняют ИС.) Каждая из них однозначно определяет новое значение своего операнда. Например:

$$\begin{aligned} 13 \cdot 21 \cdot 34(\text{большие}) &= \text{большую}; \\ 13 \cdot 22 \cdot 35(\text{ящик}) &= \text{ящиками}. \end{aligned}$$

Каждая из этих операций *образует* (строит, синтезирует) одну определенную словоформу.

Каждая из этих операций имеет обратную себе операцию: по заданной словоформе w определяется ее исходная форма w_1 и операция f такая, что $f(w_1) = w$. Например, для $w = \text{ящиками}$ имеем $w_1 = \text{ящик}$, $f = 13 \cdot 22 \cdot 35$.

Действие, обратное *синтезу* (построению) словоформы, называется *анализом* (распознаванием) словоформы.

Если $f(w_1) = w$, то три операции, образующие f , называются *морфологическими параметрами* словоформы w . Например, для словоформы *ящиками* морфологическими параметрами являются 13, 22 и 35. Иногда (для убедительности) за словоформой пишутся ее *параметры*. Например, **ящиками** 13 22 35.

Область значений операции f называется *лексической группой* f . Например, 01 13 22 35 — подмножество ИС в форме 13 22 35; ИС делится на 12 таких лексических групп.

Над лексическими группами можно выполнить множественные операции, например операцию объединения: $01\ 13\ 20\ 35 = 01\ 13\ 21\ 35 \cup 01\ 13\ 22\ 35$, представляющую собой множество ИС женского рода в творительном падеже в единственном и множественном числе, так как $20 = (21, 22)$.

Аналогично ИП делятся на 36 лексических групп.

Введенные обозначения можно использовать для формирования более сложных обозначений (формул), порождающих определенные группы элементов синтаксиса. Например, минимальная неделимая единица синтаксиса — *словосочетание*. Примером словосочетания является *согласованное определение ящиками нижними* или пара

(*ящиками, нижними*), которая, в свою очередь, является элементом прямого произведения $(01\ 13\ 22\ 35) \cdot (02\ 13\ 22\ 35)$. Таких сочетаний ИС с ИП в русском языке будут десятки тысяч. Резонно обозначить множество этих словосочетаний через $(01\ 13\ 22\ 35) \cdot (02\ 13\ 22\ 35)$ и называть его кодом *синтаксического отношения* (СО) между группами 01132235 и 02132235. Код лексической группы начинается с нуля, пробелы между кодами категорий и знак «·» необязательны: 0113223502132235.

Обозначив значения категорий буквами, получим все 36 СО:

$$01\ 1x\ 2y\ 3z\ 02\ 1x\ 2y\ 3z, \quad (3)$$

где $x = 1 \dots 3$; $y = 1 \dots 2$; $z = 1 \dots 6$, порождающих *согласованные определения*, имеющие место в русском языке.

Элементы множеств (3) обозначим через СО1. Все 36 записей (3) упорядочим лексикографически по возрастанию и обозначим через СО1001, СО1002, ..., СО1036.

Аналогично $3 \cdot 2 \cdot 6 \cdot 3 \cdot 2 = 216$ пар

$$01\ 1x\ 2y\ 3z\ 01\ 1r\ 2t\ 3z, \quad (4)$$

где $x, r = 1 \dots 3$; $y, t = 1 \dots 2$; $z = 1 \dots 6$, порождают словосочетания (*несогласованные определения* или СО2) вида: **ящиками фермеров, книгу библиотеки, заботы врачей** и т. д.

В (4) присутствуют пять параметров. Согласно грамматике русского языка все первые пять категорий, кроме последней категории «32», могут принимать всевозможные значения независимо друг от друга. Объединение всех 216 пар (4) удобно писать в виде:

$$01\ 10\ 20\ 30\ 01\ 10\ 20\ 3z. \quad (5)$$

Заменяя в (5) нули их допустимыми значениями и упорядочив полученные записи в лексикографическом порядке возрастания, получим 216 записей с кодами СО2001, СО2002, ..., СО2216.

В русском языке более 50 тыс. ИС и ИП. Слово-сочетаний из *согласованных определений* будет около миллиона. А словосочетаний из *согласованных и несогласованных определений* окажется десятки миллионов. На ЕЯ самым распространенным и многочисленным видом выражений являются сочетания из ИС и ИП, образованных по правилам СО1 и СО2.

Правила СО1 позволяют привязывать к ИС несколько ИП, например:

$$\text{хороший высокий умный студент.} \quad (6)$$

(Знаков препинания в рассматриваемом здесь синтаксисе нет.) Выражению (6) соответствует дерево в форме звезды (рис. 3).

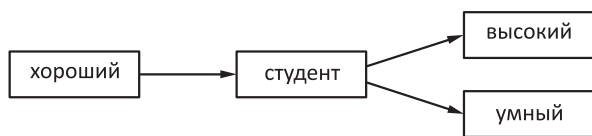


Рис. 3 Предложение-дерево. Звезда

Или в скобочной записи:

$$\text{студент (хороший, высокий, умный)}. \quad (7)$$

Шириной выражения с одним ИС называется число ИП в этом выражении. Выражение (7) имеет ширину три.

По правилам СО2 образуется линейная цепь связанных ИС, например:

$$\text{отряд студентов школы мужества}. \quad (8)$$

Выражению (8) соответствует цепь (которая на восточных языках называется *идафной цепью*; *идафа* (араб. إضافة) *дополнение*) — способ оформления несогласованного определения в арабском и ряде других восточных языков), представленная на рис. 4.



Рис. 4 Предложение-дерево. Цепь

Или в скобочной записи:

$$\text{отряд (студентов (школы (мужества)))}.$$

Длиной идафной цепи называется число ИС в этой цепи.

Далее для определенности и обозримости выражений вводятся ограничения на ширину *согласованного определения* — она не более одного, и на длину *несогласованного определения* — она не более трех.

В грамматике русского языка кроме 12 *беспредложных форм* ИС используются еще *предложные формы* ИС. Чтобы разнообразить рассматриваемый синтаксис, в лексику из ИС и ИП добавляются следующие предлоги: *над, под, на, в* — и соответствующие им *предложные формы* ИС: первые два предлога управляют творительным падежом, а последние два — предложным.

Если предположить, что эти предлоги имеют коды 21, 22, 23 и 34, то коды *предложных форм*, соответствующих этим предлогам, будут такими: 01 10 20 35 21, 01 10 20 35 22, 01 10 20 36 23 и 01 10 20 36 24. Например:

$$\begin{aligned} 0110\ 20\ 35\ 21(\text{отряд}) &= \text{над отрядом}; \\ 01\ 10\ 20\ 36\ 24(\text{отряд}) &= \text{в отряде}. \end{aligned}$$

Обозначим через S множество ИС и ИП в лексике русского языка и их формы, о которых было сказано выше. Обозначим через Q множество *согласованных и несогласованных определений*, определенных на S .

Пара (S, Q) представляет собой ориентированный граф, обозначим его через Sint . Имеются следующие основания называть этот граф *синтаксисом*.

Представим себе *корневое дерево* в этом графе:

- (а) у которого корнем является ИС;
- (б) каждое ИС в нем имеет ширину не более одного;
- (в) идафная цепь в нем имеет длину не более трех.

Деревья, удовлетворяющие требованиям (а)–(в), назовем *выражениями* в синтаксисе Sint . (Требования (б) и (в) здесь присутствуют для наглядности деревьев и *выражений*.) Примерами *выражений* являются: *дом; под домом; дом белый; под домом белым; в доме соседа; в доме высоком соседа хорошего врача зубного*.

Последнее выражение из шести слов является самым длинным в синтаксисе Sint . (Переставляя ИС и следующее за ним ИП, получим привычные для русского языка выражения.)

Данное определение *формального синтаксиса* никак не зависит от объема лексики и числа СО, на которых построен *синтаксис*.

Этот синтаксис не обременен семантикой. Вопрос инкапсуляции синтаксиса семантикой здесь не рассматривается. Хотя известна сила синтаксиса вдохнуть смысл даже в то, как «Глокая куздра штеко будланула бокра и курдючит бокренка». Поэтому выражения (деревья) в синтаксисе Sint , начинающиеся с *предложной формы* имен, обозначающих предметы, будут отвечать на вопрос «где?», и назовем их *указателями места*.

Элементами синтаксиса Sint являются обычные (синтаксически правильно построенные) выражения на русском языке, и задача заключается в поиске *указателей места* в этом синтаксисе.

Алгоритм извлечения *указателей места* из заданного текста очень прост. Имеет смысл рассмотреть работу алгоритма по шагам на простом примере:

$$\text{в высоком доме хорошего соседа зубного врача}. \quad (9)$$

Шаг 1. Выделяются словоформы в (9) и определяются коды лексических групп, к которым они принадлежат:

$$\begin{aligned} \text{в доме} &— 01213624; & \text{высоком} &— 022136; \\ \text{соседа} &— 012132; & \text{хорошего} &— 022132; \\ \text{врача} &— 012132; & \text{зубного} &— 022132. \end{aligned}$$

Шаг 2. Коды лексических групп позволяют группировать попарно словоформы из (9) таким образом, чтобы они были отношениями *согласованного* и *несогласованного определений*; полученные таким образом пары словоформ образуют дерево, представленное на рис. 5.

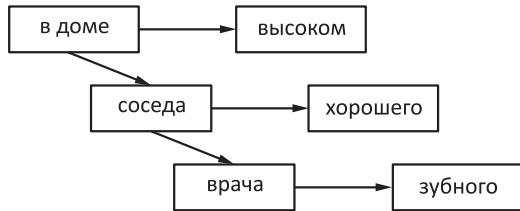


Рис. 5 Дерево, порожденное синтаксисом Sint на словоформах (9)

Шаг 3. После замены словоформ в дереве с рис. 5 на коды лексических групп, к которым они принадлежат, получится дерево, показанное на рис. 6.

Дерево на рис. 6 называется *синтаксической формой* (СФ) дерева с рис. 5, а выражения (9). Она является СФ для большего числа выражений; действительно, заменяя в этой СФ код каждой лексической группы некоторой словоформой из этой же группы, можно получить новые выражения. Так что СФ — *правило*, порождающее выражения определенной формы. Из сказанного следует шаг 4. (Сформулирован он в форме утверждения.)

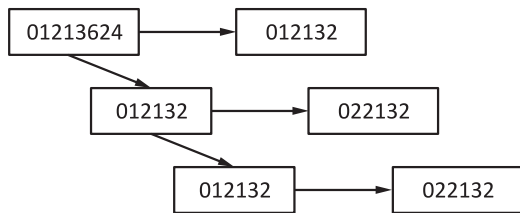


Рис. 6 Синтаксическая форма дерева, представленного на рис. 5

Шаг 4. Принадлежность выражения (9) к СФ с рис. 6 указывает на то, что (9) есть *указатель места*.

Рассмотрим еще несколько примеров *указателей места* из шести слов. Все они имеют одинаковую структуру: **в большом аule нижнего района нашей области; под нижним ящиком письменного стола большого начальника; на длинном хвосте серого слона бедного индуса.**

Анализируя эти выражения по тому же алгоритму из четырех шагов, можно убедиться в том, что и они — *указатели места*.

Обобщая СФ всевозможных *указателей места*, нетрудно заметить, что СФ, порождающие *указатели места*, имеют форму дерева, представленного на рис. 7.

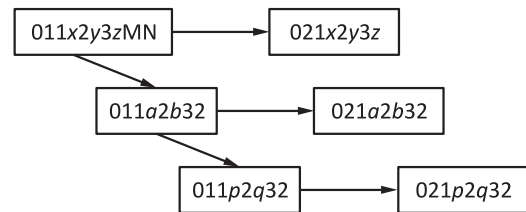


Рис. 7 Формы указателей места, объединенные в дереве T_{Space}

Корню этого дерева, обозначенного T_{Space} , соответствует код предложной формы ИС $011x2y3zMN$, где MN — код предлога, управляющего падежной формой $3z$.

Также нетрудно заметить, что любое поддерево дерева T_{Space} с корнем в $011x2y3zMN$ есть СФ, порождающая *указатели места* в синтаксисе Sint.

Более того, эти СФ порождают всевозможные *указатели места* в синтаксисе Sint.

4 Заключение

Авторы оперировали синтаксисом Sint — вырезкой из синтаксиса русского языка — для наглядной доступности задачи и алгоритма ее решения. Задача и алгоритм ее решения не претерпят существенных изменений, если их сформулировать для всего синтаксиса русского языка. Изменяются размеры словаря и синтаксиса (графа Sint). Для быстрого поиска слов в словаре (в упорядоченном массиве строк) можно использовать известные алгоритмы.

Приведенное выше описание алгоритма выявления *указателей места* в тексте, написанном в синтаксисе Sint, указывает на простоту и корректность алгоритма. Листинг программы, реализующей этот алгоритм, занимает не более одной страницы. Линейная сложность алгоритма обеспечивает ему мгновенное исполнение на компьютере.

Программа, анализирующая предложения в упрощенном синтаксисе и с лексикой из 16 слов (синтаксический анализатор), подготовлена авторами и реализована в IDE Delphi 7.

Литература

1. Барсегян А. А., Курприянов М. С., Степаненко В. В., Холод И. И. Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP. — СПб.: БХВ-Петербург, 2008. 384 с.
2. Шихиев Ф. Ш. Формализация и сетевая формулировка задачи синтаксического анализа. Дис. . . . канд. физ.-мат. наук. — СПб.: СПбГУ, 2006. 171 с.
3. Рубашкин В. Ш., Чуприн Б. Ю. Распознавание количественной информации в ЕЯ-текстах // Компьютерная лингвистика и интеллектуальные технологии: Тр. Междунар. конф. «Диалог-2006». — М.: РГГУ, 2006. С. 456–467.
4. Харари Ф. Теория графов / Пер. с англ. — М.: Едиториал УРСС, 2003. 296 с. (*Harary F. Graph theory.* — Boulder, CO, USA: Westview Press, 1994. 284 p.)

Поступила в редакцию 13.04.17

DISCRETE ANALYSIS IN PARSING

Ya. M. Mirzabekov and Sh. B. Shihiev

Dagestan State University, 43-a Gadzhiyev Str., Makhachkala 367000, Republic of Dagestan, Russian Federation

Abstract: An informal definition of syntax is given in terms of discrete mathematics and graph theory. The main difficulty for numerous attempts to formalize a natural language is the semantics of the language. It is shown how it is possible to classify expressions on a semantic basis, using their syntactic features. Classification of expressions by the kind of questions they answer is the simplest way of grouping expressions on the semantic basis. The present authors describe an algorithm that recognizes place pointers, that is, expressions which answer the question “where?”. On specific examples, the problem of analysis of expressions and the inverse problem of synthesis, more precisely, the applied problem of recognition of expressions are considered.

Keywords: natural language; discrete mathematics; graph theory; syntax; word forms; morphological parameters; consistent definitions; inconsistent definitions; vocabulary; semantics

DOI: 10.14357/19922264180214

References

1. Barsegyan, A. A., M. S. Kupriyanov, V. V. Stepanenko, and I. I. Cholod. 2007. *Tekhnologiya analiza dannykh: Data Mining, Visual Mining, Text Mining, OLAP* [Technology of data analysis: Data Mining, Visual Mining, Text Mining, OLAP]. St. Petersburg: BHV. 384 p.
2. Shikhiev, F. Sh. 2006. Formalizatsiya i setevaya formulirovka zadachi sintaksicheskogo analiza [Formalization and network formulation of the task of parsing]. St. Petersburg: SpbGU. PhD Diss. 171 p.
3. Rubashkin, V. Sh., and B. Y. Chuprin. 2006. Raspoznavanie kolichestvennoy informatsii v EYa-tekstakh [Quantitative data recognition at NLP]. Computational Linguistics and Intellectual Technologies: Conference (International) “Dialogue-2006” Proceedings. Moscow: RGGU. 456–467.
4. Harary, F. 1994. *Graph theory*. Boulder, CO: Westview Press. 284 p.

Received April 13, 2017

Contributors

Mirzabekov Yahya M. (b. 1983) — senior lecturer, Dagestan State University, 43-a Gadzhiyev Str., Makhachkala 367000, Republic of Dagestan, Russian Federation; yash831@mail.ru

Shihiev Shukur B. (b. 1951) — Candidate of Sciences (PhD) in physics and mathematics, associate professor, Dagestan State University, 43-a Gadzhiyev Str., Makhachkala 367000, Republic of Dagestan, Russian Federation; sh_sh_b51@mail.ru

MACHINE TRANSLATION OF RUSSIAN CONNECTIVES INTO FRENCH: ERRORS AND QUALITY FAILURES

V. Nuriev¹, N. Buntman², and O. Inkova³

Abstract: The paper shows what machine errors and quality failures may occur when translating connectives. To that end, a statistical machine translation (SMT) system has been used in order to generate translation samples. The opening part presents a brief retrospective on how machine translation has been developing for over 60 years; it sets out the necessary background and provides the context. Also, section 1 explains how an SMT system works. Then, the paper takes a closer look at the problem of evaluation of machine translation quality. Several approaches to classifying machine translation errors are considered to finally attempt a taxonomy that covers specifically the errors central to translation of connectives (from Russian into French). The closing section provides examples of these machine translation errors.

Keywords: statistical machine translation; corpus linguistics; machine translation errors; parallel texts

DOI: 10.14357/19922264180215

1 Introduction: Historical Background

Historically, in order to acquire the state-of-the-art accomplishments, machine translation has progressed through several stages, with this progress giving rise to a huge bulk of knowledge, which has quite naturally resulted in a large body of literature (see, for example, [1–7]). As the immediate purpose of the present paper is neither to produce an elaborate history of machine translation nor to inform readers on emerging trends in the field (i. e., neural machine translation), we will limit ourselves to outlining briefly main kinds of machine translation systems and their developmental steps.

The first significant attempts to coherently approach automatic translation date back to the early 1950s. They brought about Rule-Based Machine Translation (RBMT), the paradigm that remained prevalent for decades thereafter. Within it, there were three basic approaches, namely: (i) the ‘direct translation’ model; (ii) the ‘interlingua’ model; and (iii) the ‘transfer’ model.

In the ‘direct translation’ model, programming rules are developed and intended to translate “specifically from one source language (SL) into one particular target language (TL) with a minimal amount of analysis and syntactic reorganization” [8]. To put it differently, this model sees an SL sentence as a string of words, gets their

TL equivalents from accessible bilingual dictionaries, and reorganizes these equivalents into the TL sentence, maintaining the original word order as much as possible. Relevant linguistic knowledge is usually stored in bilingual dictionaries, it can also be described by algorithms and then coded. The translation potential of the model is constrained by both various problems of lexical ambiguity (e. g., homonyms) and syntactic limitations, as it is preferably applied to pairs of languages with similar syntactic features [9].

The ‘interlingua’ model favors the use of an artificial mediating language. The development of this model is grounded in a preconception that different natural languages can be reduced to a number of abstract language-neutral representations (codes independent of both SL and TL), and they would serve as a semantic-syntactic mediator when translating between languages. Hence, translation proceeds in two stages: the system converts an SL text into interlingua to further synthesize a TL text based on the interlingual representations. These representations are conceived to be clear-cut and express the content of an ST text to the fullest extent — its morphological, syntactic, and semantic information — since such RBMT “systems depend crucially on the common language into and out of which all sentences accepted by the system are mapped” [1]. One of the most serious shortcomings the ‘interlingua’ model is blamed for is the lack of language-specificity and an oversimplification through which natural language is perceived.

¹Institute of Linguistics of the Russian Academy of Sciences, 1 bld. 1 Bolshoy Kislovsky Lane, Moscow 125009, Russian Federation; Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation, nurieff.v@gmail.com

²M. V. Lomonosov Moscow State University, GSP-1, Leninskie Gory, Moscow 119991, Russian Federation, nabunt@hotmail.com

³University of Geneva, 24 du Général-Dufour Str., Genève 4 1211, Switzerland, Olga.Inkova@unige.ch

The third — ‘transfer’ — model converts the input text into a transfer structure that abstracts away from some grammatical details of the SL to produce a disambiguated representation of the source input. Then, this SL-dependent representation is transferred into a corresponding TL-dependent structure, from which a target output is generated. Thus, here, translation consists of three stages: analysis, transfer, and generation (synthesis). Therefore, at the very least, “transfer systems require monolingual modules to analyse and generate sentences, and transfer modules to relate translationally equivalent representations of those sentences” [1]. In contrast with the ‘direct translation’ approach, the ‘transfer’ model sees a sentence as a structure other than a linear string of words, which makes it more flexible and complex. Unlike interlingua systems, the SL and TL representations in this model are language-specific (and not language-neutral). It is noteworthy that for a transfer system, the number of transfer modules grows dramatically with the number of languages: for n languages, one needs at least $\lceil n(n-1) \rceil / 2$ transfer modules. “This is because for each of the n languages, there are $(n-1)$ possible TL (. . .) If these modules are reversible, then only half of this number of transfer modules are required” [1]. This is generally regarded as the foremost disadvantage of transfer systems, since they get more expensive the more languages they acquire.

Of the three RBMT approaches, the latter two persisted over the years. Only in 1981 Makoto Nagao¹ proposed a framework of machine translation that was aimed at overcoming weaknesses of RBMT when one translated between languages syntactic structures of which differed drastically (for more on this, see [10, 11]). His approach started from the analogy principle of translation, i. e., translation was perceived as a search for ‘analogues’ (similar in meaning and form) of SL sentences, phrases, or word strings in TL sentences stored in a database. Those databases were derived from parallel corpora where TL texts represented examples of translations previously produced by professional translators, and thus, the approach was dubbed ‘example-based machine translation’ (EBMT). On the one hand, since EBMT systems made use of human translations, they succeeded more in tackling such problems as idiomaticity. On the other hand, one of their major issues was the recombination of selected TL examples (most commonly short strings of words or phrases) in order to produce fluent and grammatical output [8].

Another corpus-based approach, of more recent origin, is SMT. It began to score its first successes in the 1990s. First, SMT systems had provided word-based translation only, but over time, they became phrase- and syntax-based. Statistical machine translation is proba-

bilistic by nature focusing on how to generate multiple hypothetical translations for the input string of words, and then how to work out which one of those is most likely. It employs two processes: training and decoding. In the course of training, a statistical model of translation is extracted from a parallel corpus, and a statistical model of the TL from a monolingual corpus. The translation model consists of a bilingual dictionary where to each possible translation for a given SL word or phrase, a probability is assigned. Yet, the model is not similar to a common dictionary, since its entries may contain not only the translations that are plausible and evident; there may be translations that are unlikely but not impossible, and the assigned probabilities show this. The language model comprises a database of TL word sequences, to each of which a probability is also assigned. During training, additional information can be extracted, such as models of relative sentence length, word reordering, etc. These models are then used in decoding, the process that actually produces an output translation. The decoding is per se a search process: a decoder searches over all possible translations for the one that will have the highest overall probability according to the translation and language models. In other words, SMT tries to solve the following decision problem: which translation is the most likely? There are two formulae to compute that score: the noisy channel model

$$\text{Translation} = \arg \max_T P(S|T)P(T) \quad (1)$$

and — more flexible and widespread — the log-linear model

$$\text{Translation} = \arg \max_T \sum_{m=1}^M \lambda_m h_m(T, S) \quad (2)$$

(for a full description of these formulae, see [12]).

As one can see, the leftmost parts of Eqs. (1) and (2) ($\text{Translation} = \arg \max_T \dots$) are the same. They are to be interpreted as follows: given that we have T candidate translations, we will let the output Translation be the T with the maximum ($\arg \max$) score. The rightmost parts of these equations differ as to how each candidate translation should be scored.

Model 1 has two component (or ‘feature’) scores $P(S|T)$ and $P(T)$ that are to be multiplied together. $P(S|T)$ gives the likelihood that the SL sentence S and the candidate translation T are translationally equivalent. This feature is referred to as the *translation model*. $P(T)$ gives the likelihood that the candidate translation T is actually a valid sentence in the TL and is referred to as the *language model*.

The right-hand-side of the model 2 equation shows that this model comprises a set of log feature scores to

¹The idea was presented by Makoto Nagao at a 1981 conference, and his paper on this was published only three years later.

be added together. The $\sum_{m=1}^M$ notation indicates that there are a total of M features to be scored and that their individual scores are to be added up. “These individual scores are to be computed by multiplying together two feature-specific values, λ_m and $h_m(T, S)$ where λ_m is simply the weight indicating the importance of that feature relative to the other features, and $h_m(T, S)$ is the log probability assigned to the source–candidate pair by that feature” [12].

In the present study, an SMT system¹ has been employed. It works with the major European languages and all translation directions are reversible (e.g., English–Spanish and vice versa). As well as other SMT systems, this one has three key components: a translation model, a language model, and a decoder. The translation model is expected to construct a graph that would encompass every possible way to translate an input sentence and estimate the probability of each translation. To that end, the translation model learns from bilingual parallel corpora. On the contrary, the language model gathers its data from single-language corpora and stores all TL frequent n -word combinations, along with information about frequency of use, n may be from 1 to 7 (usually, 5). The 3rd component, the decoder, performs the output translation. For every sentence of the SL text, it chooses all possible translation options, combining phrases from the translation model and sorting them out in terms of probability. The decoder then matches all the variations against the language model. In so doing, it searches for the sentence with the highest probability — according to the translation model, and the highest frequency of use — according to the language model.

2 Evaluation of Machine Translation Quality

In the present study, the ultimate purpose of which is to contrast human (reference) translations of Russian connectives into French against automated translations, an SMT system is used to generate samples of machine translations (for examples, see section 3). With such a purpose, one has to build a taxonomy of possible quality failures or errors occurring when a machine translator processes sentences with the linguistic units in question.

While in the literature there are a number of examples of those taxonomies, we will only pay attention to some especially interesting cases². The first case is described in [14], where the authors argue that automated quality metrics (BLEU, NIST, etc.), however useful and important, cannot be used all alone to identify machine translation errors; manual error analysis appears to be

also necessary. The typology of errors proposed in [14] has a hierarchical structure, with five classes of errors at the highest level: “Missing Words,” “Word Order,” “Incorrect Words,” “Unknown Words,” and “Punctuation” errors.

Errors of the first type occur when in the generated output, a word is missing. This class divides into two subclasses of errors arising (*i*) when the missing word is central to the meaning of the SL sentence; and (*ii*) when the missing word ensures the well-formedness of the output, but does not alter the meaning.

“Word Order” errors are classified according to the reordering principle, i.e., to what extent the generated output needs to be reordered so as to get a correct translation out of the generated hypothesis. The authors distinguish between word- and phrase-based reorderings and, further, between local and long-range reorderings. A word-based reordering involves moving separated words only, whereas a phrase-based reordering moves whole strings of words. The distinction between local and long-range is somewhat less obvious, it is stated that the former takes place when the words have to be reordered only in a local context (within the same syntactic chunk), while the latter results from moving the words into another chunk.

In this classification, “Incorrect Words” errors represent the largest class of machine translation errors. They occur once the system is unable to find a correct translation of a TL word or phrase. There are five types of errors falling into this category. Incorrect words are: (*i*) TL equivalents distorting the meaning of the sentence; (*ii*) TL equivalents in the wrong grammatical form (for inflected languages); (*iii*) redundant words (in case of spoken language translation); (*iv*) words that do not fit in stylistically with the rest of the TL sentence or neighboring context; and (*v*) idiomatic expressions unknown to the system and translated word-for-word (incorrectly).

“Unknown words” errors occur if the system fails to identify the SL word. This may happen when either unknown words (and stems) or unseen forms of the known stems are processed.

Also, the authors mention “Punctuation” errors, without paying much attention to them.

It is emphasized that the error types are not mutually exclusive and errors rarely occur in isolation, i.e., one may bring about another.

One more example of taxonomy of machine translation errors referred to here is taken from [13]. This taxonomy classifies errors in terms of the linguistic item affected by the error, taking into account different language levels where the error is located³.

¹Here, it is used as a web service.

²For a concise description of how the ideas behind the analysis of machine translation errors developed, see section 2 in [13].

³A somewhat similar approach is used in [15].

Orthography level errors result from misuse of punctuation and misspelling of words. They are divided into three types: punctuation, capitalization (inappropriate use of capital letters), and spelling errors.

Lexis level errors include omission, addition, and untranslated. Omission and addition errors are further analyzed considering the nature of words they affect: (a) content words (or lexical words), i. e., words that carry the meaning of a sentence; and (b) function (or grammatical) words, i. e., words that may have little lexical meaning, but convey grammatical relationships within a sentence. Omission errors occur when the translation of a word present in the SL text is missing in the TL output; an addition error represents the opposite. Untranslated errors happen when the system fails to find any translation candidate for a given SL word and leaves it as it is, without translation.

Grammar level errors cover morphological and syntactic content. They are grouped into two categories: misselection and misordering errors. Misselection errors are morphological, encompassing problems at word class-level (e. g., part-of-speech confusion) and at verbal level (tense and person), errors of agreement (gender, number, person) and those in contractions (between prepositions and articles). Misordering errors are related with syntactic problems that the TL output may demonstrate.

Semantic errors arise from problems related to the meaning of the words and wrong word selection. The authors propose to individuate three types of errors: confusion of senses, wrong choice, and collocational and idiomatic errors. Confusion of senses takes place when an SL word is translated by a TL equivalent representing one of its possible meanings, but, in the given context, the chosen translation does not fit. Wrong choice errors are produced if the chosen translation candidate is totally incorrect (e. g., in the case of homonyms). Collocational and idiomatic errors happen when the translation breaks the rules of compositional semantics.

Discourse-level errors result from failure to choose the most natural discursive option. At the discourse level, the errors are considered in terms of style, variety, and linguistic items that should not be translated. Style errors concern stylistically poor decision-making. Variety errors occur in cases where the TL is marked by dialect variation and the machine translator loosely combines lexical or grammatical structures, choosing from a mixture of dialects. Under the ‘should not be translated’ category, the authors consider SL sequences of words that should not be translated in the TL (e. g., book or film titles).

Based on these examples and some other insights, we have devised a taxonomy of machine translation errors,

identifying specifically the errors central to translation of connectives. To differentiate between the following types of errors, we propose:

- (1) AgramTotal (the output sentence is ungrammatical);
- (2) ErrorTotal (the SL connective is translated by a nonexisting linguistic item);
- (3) AgramPostCNT (there are grammatical errors in the chunk introduced by the connective);
- (4) AgramLocal (there are grammatical errors in the chunk that is not introduced by the connective);
- (5) AgramOrth (in the TL output, the connective is misspelled);
- (6) ErrorCNT (the choice of the TL connective is semantically ill-motivated);
- (7) Cyrillic (the TL sentence has nontranslated words left in Cyrillic); and
- (8) ErrorPart (a part of a multicomponent connective is mistranslated).

Machine translations, stored in a supracorpora database (SCDB) of connectives¹, are annotated according to this taxonomy. Also, while annotating, the translation is labelled as either ‘congruent’ (the Russian connective is translated by a French connective) or ‘divergent’ (the Russian connective is translated by a nonconnective). If there are no errors in the TL output, the translation is marked as NoError.

As in [14], we have not opted for automated quality metrics in order to identify errors of machine translation and estimate its quality. Instead, manual error analysis has been chosen, since connectives are complex linguistic units, having the ability to take both sentential and discursive scope. Manual error analysis proceeds in two steps: (i) the translation is annotated, i. e., an expert linguist identifies translation errors; and (ii) the annotation is double-checked for accuracy and reliability by another expert linguist (and translator).

In the SCDB of connectives, there is at least one reference — performed by a human — translation for each sample of machine translation, which opens up vast opportunities for users to compare human and automated translations.

3 Statistical Machine Translation of Connectives: Exemplifying Errors

The present section lists examples of SMT errors mentioned above. The translation direction is from Russian

¹For more on the SCDB of connectives, see [16–20].

Table 1 AgramTotal errors

SL fragment	Connective	Translation	TL equivalent and SMT errors
что ж тут, во всем этом, в самом деле, такого особенно преступного с моей стороны, то есть без предрассудков-то, а здраво судя?	то есть (CNT)	dans tout cela, en fait, ce sont particulièrement criminel de ma part, c'est à dire sans préjugés chose, et la santé pour en juger?	c'est-à-dire (AgramTotal)

Table 2 AgramTotal and Cyrillic errors

SL fragment	Connective	Translation	TL equivalent and SMT errors
заразительная игра воздуха в легких обойдет всех, причем иного прошибет слеза.	Причем (CNT)	заразительная le jeu de l'air dans les poumons évitera tous, et d'autre прошибет larme.	et (AgramTotal) (Cyrillic)

Table 3 AgramTotal and ErrorCNT errors

SL fragment	Connective	Translation	TL equivalent and SMT errors
Поверите ли, сударь, что собачонка не стоит восьми гривен, то есть я не дал бы за нее и восьми грошей. . .	то есть (CNT)	le Croirez —, monsieur, que le chien n'est pas la peine de huit hryvnia, c'est que je ne donnerais pour elle, et huit deniers. . .	c'est que (AgramTotal) (ErrorCNT)
Очень многие, и притом важнейшие, идеи, темы и образы его творчества — и предшествующего и последующего — появляются здесь в предельно острой и обнаженной форме. . .	и притом (CNT)	Très nombreux, et bien que les critiques, les idées, les thèmes et les images de sa créativité et précédant et suivant — apparaissent ici très aiguë et nue sous la forme de. . .	et bien que (AgramTotal) (ErrorCNT)

into French. All examples are collected from the SCDB of connectives¹.

AgramTotal errors make the TL output agrammatical, while the connective itself may be translated correctly (Table 1).

The degree of agrammaticality is hard to define, and the study is not aimed at providing such data. Nevertheless, it has been established that AgramTotal errors may occur in translations that contain Cyrillic errors (Table 2). As in the previous example, one can see that the connective is translated correctly by one of its regular equivalents.

AgramTotal errors may be encountered in the TL output where there are semantic errors in choice of connectives (ErrorCNT). This is to say that, in the TL sentence, the connective conveys the type of logical-semantic relations, drastically different from that expressed by the SL connective and inappropriate in the given context, which also contributes to agrammaticality

of the output sentence (Table 3). Here, one can see that in the output sentence *to est'*, a reformulation marker, is rendered by the causative connective *c'est que*, and *i protom* (the concomitance relation) is translated by *et bien que* (the concessive relation).

The kind of errors to pay special attention to is ErrorTotal. A nonexisting linguistic item generated as a TL equivalent of the SL connective may consist of elements each of which occurs naturally in the TL, yet combining these elements does not produce a fluent translation in the given context (Table 4).

It is noteworthy that the examples given in Table 4 demonstrate how the SMT system fails to consider a multiword connective as a whole — inseparable linguistic unit.

In the case of AgramPostCNT errors, agrammaticality of the chunk introduced by the connective also comes in different guises. It does not necessarily result from the wrong use of connectives, it may be caused by

¹As of December 17, 2017, the SCDB contains parallel texts (mostly fiction) in Russian and French of 3.5 million tokens, 15,328 annotations of translations from Russian into French (3,606 of which are annotations of machine translations).

Table 4 ErrorTotal and AgramLocal errors

SL fragment	Connective	Translation	TL equivalent and SMT errors
— Я отвык совсем ездить; с непривычки, да еще зимой, признаюсь, мне бы трудно было, не хотелось бы. . .	да еще <CNT>	Je suis tout à fait perdu l'habitude de conduire; de l'habitude, et oui, même en hiver, je l'avoue, il m'aurait été difficile, je ne voudrais pas. . .	et oui, même <ErrorTotal> <AgramLocal>
Он кинулся к комоду, с грохотом вытащил ящик, а из него портфель, бессвязно при этом выкрикивая. . .	при этом <CNT>	Il s'est jeté à la poitrine, avec le fracas tiré de lettres, de portefeuille, de façon incohérente lors de cette criant. . .	lors de cette <AgramTotal> <ErrorTotal>

Table 5 AgramPostCNT errors

SL fragment	Connective	Translation	TL equivalent and SMT errors
— Но, видно, это было неизбежно: зато как я покоен теперь и. . . как счастлив!	зато <CNT>	— Mais, on le voit, c'était inévitable, mais comme je l'est décédé maintenant, et. . . heureux!	mais <AgramPostCNT>

Table 6 AgramLocal and Cyrillic errors

SL fragment	Connective	Translation	TL equivalent and SMT errors
Шлиппенбах возбужденно жестикулировал. Зато водитель оставался совершенно невозмутимым.	зато <CNT>	Шлиппенбах excité жестикулировал. Mais le conducteur est resté totalement impassible.	mais <Cyrillic> <AgramLocal>

Table 7 AgramOrth and AgramLocal errors

SL fragment	Connective	Translation	TL equivalent and SMT errors
Он, как только проснулся, тотчас же вознамерился встать, умыться и, напившись чаю, подумать хорошенько, кое-что сообразить, записать и вообще заняться этим делом как следует.	и вообще <CNT>	Il est, dès que s'est réveillé, il a immédiatement décidé de se lever, de se laver et de peur qu'en buvant du thé, penser bien, quelque chose de comprendre, d'écrire et de s'occuper en gnral cette affaire comme il se doit.	éet en gnral <AgramLocal> <AgramOrth>

Table 8 ErrorPart and AgramPostCNT errors

SL fragment	Connective	Translation	TL equivalent and SMT errors
— Не могу сказать; но меня интересует при этом другое обстоятельство, так сказать. . .	но при этом <CNT>	Je ne puis le dire; mais ce qui m'intéresse lorsque cette autre circonstance. . .	mais lorsque <ErrorPart> <AgramPostCNT>

Table 9 NoError case

SL fragment	Connective	Translation	TL equivalent and SMT errors
И однако ж , одеваясь, он осмотрел свой костюм тщательнее обыкновенного.	и однако ж <CNT>	Et cependant , en s'habillant, il a regardé son costume plus soigneusement que d'habitude.	et cependant <NoError>

problems of grammatical agreement in the TL output (Table 5).

Unlikely AgramPostCNT errors, AgramLocal ones occur in the fragment that is not introduced by the connective (Table 6).

AgramOrth errors are relatively rare; however, there are a few examples (Table 7).

The ErrorPart label is used to annotate cases where the SMT system fails to translate a part of a multicomponent connective (Table 8).

In the database of connectives, there are cases annotated as NoError where the SMT system does not make any errors at all (Table 9).

4 Concluding Remarks

In the present paper, the authors have focused on the SMT approach and showed its evolution throughout the years. Statistical machine translation is not the leading paradigm any more, with neural machine translation systems being actively developed. Yet, many publicly available automated translation systems are still statistical, and we employ one of these systems in our study of connectives, at least at the current stage. That is why, it was necessary to show how SMT systems worked.

Also, the evaluation of machine translation quality has been discussed. Arguing that the role of automatic evaluation metrics is overly important nowadays, we suggested the necessity of manual error analysis made by real users and experts. To perform such analysis, one has to build a taxonomy of possible quality failures or errors occurring when a machine translation system processes the SL input. We considered several different cases of those taxonomies in order to propose our classification that would cover specifically the errors central to translation of connectives (eight types of errors).

Finally, we have provided specific examples to show errors that the SMT system can make when processing sentences with connectives.

This study opens up some perspectives of further research. First, it would be interesting to make it more grounded in empirical evidence, i. e., to gather reliable statistics on every type of SMT errors mentioned above. Second, it would be fruitful to see what errors neural machine translation systems make when translating connectives. Comparing their errors with those made by SMT systems could be insightful.

Acknowledgments

The study has been conducted at the Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of

Sciences with financial aid from the Russian Foundation for Basic Research (Grant No. 16-24-41002) and the Swiss National Science Foundation (Grant No. IZLRZ1_164059/1).

References

1. Trujillo, A. 1999. *Translation engines: Techniques for machine translation*. Berlin – Heidelberg – New York: Springer. 306 p.
2. Hutchins, W.J. 2000. *Early years in machine translation: Memoirs and biographies of pioneers*. Amsterdam–Philadelphia: John Benjamins. 400 p.
3. Nirenburg, S., H. Somers, and Y. Wilks, eds. 2003. *Readings in machine translation*. Cambridge–London: The MIT Press. 430 p.
4. Wilks, Y. 2009. *Machine translation. Its scope and limits*. New York, NY: Springer. 246 p.
5. Goutte, C., N. Cancedda, M. Dymetman, and G. Foster, eds. 2009. *Learning machine translation*. Cambridge–London: The MIT Press. 329 p.
6. Koehn, Ph. 2010. *Statistical machine translation*. New York, NY: Cambridge University Press. 447 p.
7. Du, J., and A. Way. 2017. Neural pre-translation for hybrid machine translation. *MT Summit XVI, the 16th Machine Translation Summit Proceedings*. Eds. S. Kurohashi and P. Fung. Nagoya. 27–40. Available at: http://aamt.info/app-def/S-102/mtsummit/2017/wp-content/uploads/sites/2/2017/09/MTSummitXVI_ResearchTrack.pdf (accessed December 12, 2017).
8. Hutchins, W.J. 2010. Machine translation: A concise history. *Transl. Stud.* 13(1-2):29–70.
9. Shiwen, Y., and X. Bai. 2015. Rule-based machine translation. *The Routledge encyclopedia of translation technology*. Ed. Ch. Sin-wai. Abingdon – New York: Routledge. 186–200.
10. Nagao, M. 1984. A framework of a mechanical translation between Japanese and English by analogy principle. *Artificial and human intelligence*. Eds. A. Elithorn and R. Banerji. Amsterdam: Elsevier. 173–180.
11. Hutchins, W.J. 2005. Towards a definition of example-based machine translation. *MT Summit X Second Workshop on Example-Based Machine Translation Proceedings*. Phuket. 63–70. Available at: <http://www.mt-archive.info/MTS-2005-Hutchins.pdf> (accessed December 12, 2017).
12. Hearne, M., and A. Way. 2011. Statistical machine translation: A guide for linguists and translators. *Lang. Linguist. Compass* 5(5):205–226.
13. Costa, Â., W. Ling, T. Luís, R. Correia, and L. Coheur. 2015. A linguistically motivated taxonomy for machine translation error analysis. *Machine Translation* 29:127–161.
14. Vilar, D., J. Xu, L. D’Haro, and H. Ney. 2006. Error analysis of statistical machine translation output. *5th Conference (International) on Language Resources and Evaluation Proceedings*. Genoa, Italy: European Language Resources

- Association (ELRA). Available at: http://www.lrec-conf.org/proceedings/lrec2006/pdf/413_pdf.pdf (accessed December 12, 2017).
15. Buntman, N., J.-L. Minel, D. Le Pesant, and I. Zatsman. 2010. Typology and computer modelling of translation difficulties. *Informatika i ee Primeneniya — Inform. Appl.* 3(4):77–83.
 16. Popkova, N. A., O. Yu. In'kova, I. M. Zatsman, and M. G. Kruzhev. 2015. Metodika postroyeniya monoekvivalentsiy v nadkorpusnoy baze dannykh konnektorov [Methodology of constructing monoequivalences in the supracorpora database of connectors]. *Tr. 2-y nauchn. konf. "Zadachi sovremennoy informatiki"* [2nd Scientific Conference "Problems of Modern Informatics" Proceedings]. Moscow: FRC CSC RAS. 143–153.
 17. Zatsman, I. M., O. Yu. In'kova, M. G. Kruzhev, and N. A. Popkova. 2016. Predstavlenie krossyazykovykh znaniy o konnektorakh v nadkorpusnykh bazakh dannykh [Representation of cross-lingual knowledge about connectors in supracorpora databases]. *Informatika i ee Primeneniya — Inform. Appl.* 10(1):106–118.
 18. Zaliznyak, Anna A., I. M. Zatsman, and O. Yu. In'kova. 2017. Nadkorpusnaya baza dannykh konnektorov: postroyeniye sistemy terminov [Supracorpora database on connectives: Term system development]. *Informatika i ee Primeneniya — Inform. Appl.* 11(1):100–108.
 19. Zatsman, I. M., O. S. Mamonova, and A. Yu. Shchurova. 2017. Obratimost' i al'ternativnost' generalizatsii modely perevoda konnektorov v parallel'nykh tekstakh [Reversibility and alternativeness of generalization of connectives translations models in parallel texts]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 27(2):125–142.
 20. In'kova, O. Yu., and N. A. Popkova. 2017. Statistical data as information source for linguistic analysis of Russian connectors. *Informatika i ee Primeneniya — Inform. Appl.* 11(3):123–131.

Received March 15, 2018

Contributors

Nuriev Vitaly A. (b. 1980) — Candidate of Science (PhD) in philology, senior scientist, Institute of Linguistics of the Russian Academy of Sciences, 1 bld. 1 Bolshoy Kislovsky Lane, Moscow 125009, 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; nurieff.v@gmail.com

Buntman Nadezhda V. (b. 1957) — Candidate of Science (PhD) in philology, associate professor, M. V. Lomonosov Moscow State University, GSP-1, Leninskie Gory, Moscow 119991, Russian Federation; nabunt@hotmail.com

Inkova Olga Yu. (b. 1965) — Doctor of Science (PhD) in philology, associate professor, Faculty of Arts, University of Geneva, 24 du Général-Dufour Str., Genève 4 1211, Switzerland; Olga.Inkova@unige.ch

ОШИБКИ И НЕТОЧНОСТИ МАШИННОГО ПЕРЕВОДА РУССКИХ КОННЕКТОРОВ НА ФРАНЦУЗСКИЙ ЯЗЫК*

В. А. Нуриев^{1,2}, Н. В. Бунтман³, О. Ю. Инькова⁴

¹Институт языкознания Российской академии наук

²Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

³Московский государственный университет им. М. В. Ломоносова

⁴Филологический факультет Женевского университета

Аннотация: Описываются ошибки и неточности, возникающие при машинном переводе коннекторов. Для генерирования образцов перевода использовался статистический машинный переводчик. В первом разделе рассматривается история развития машинного перевода и объясняются принципы работы статистического машинного переводчика. Во втором разделе поднимается проблема оценки качества машинного перевода. Здесь представлены несколько подходов к классификации ошибок машинного перевода, а также предлагается таксономия ошибок, разработанная применительно к переводу коннекторов (с русского языка на французский). В заключительной части демонстрируются примеры этих ошибок.

Ключевые слова: статистический машинный перевод; корпусная лингвистика; ошибки машинного перевода; параллельные тексты

DOI: 10.14357/19922264180215

* Исследование выполнено при финансовой поддержке РФФИ (грант № 16-24-41002) и Швейцарского национального научного фонда (грант № IZLRZ1.164059/1).

Литература

1. *Trujillo A.* Translation engines: Techniques for machine translation. — Berlin — Heidelberg — New York: Springer, 1999. 306 p.
2. *Hutchins W.J.* Early years in machine translation: Memoirs and biographies of pioneers. — Amsterdam—Philadelphia: John Benjamins, 2000. 400 p.
3. Readings in machine translation / Eds. S. Nirenburg, H. Somers, Y. Wilks. — Cambridge—London: The MIT Press, 2003. 430 p.
4. *Wilks Y.* Machine translation. Its scope and limits. — New York, NY, USA: Springer, 2009. 246 p.
5. Learning machine translation / Eds. C. Goutte, N. Cancedda, M. Dymetman, G. Foster. — Cambridge—London: The MIT Press, 2009. 329 p.
6. *Koehn Ph.* Statistical machine translation. — New York, NY, USA: Cambridge University Press, 2010. 447 p.
7. *Du J., Way A.* Neural pre-translation for hybrid machine translation // MT Summit XVI, the 16th Machine Translation Summit Proceedings / Eds. S. Kurohashi, P. Fung. — Nagoya, 2017. P. 27–40. http://aamt.info/app-def/S-102/mtsummit/2017/wp-content/uploads/sites/2/2017/09/MTSummitXVI_ResearchTrack.pdf.
8. *Hutchins W.J.* Machine translation: A concise history // *Transl. Stud.*, 2010. Vol. 13. No. 1-2. P. 29–70.
9. *Shiwen Y., Bai X.* Rule-based machine translation // *The Routledge encyclopedia of translation technology* / Ed. Ch. Sin-wai. — Abingdon—New York: Routledge, 2015. P. 186–200.
10. *Nagao M.* A framework of a mechanical translation between Japanese and English by analogy principle // *Artificial and human intelligence* / Eds. A. Elithorn, R. Banerji. — Amsterdam: Elsevier, 1984. P. 173–180.
11. *Hutchins W.J.* Towards a definition of example-based machine translation // MT Summit X Second Workshop on Example-Based Machine Translation Proceedings. — Phuket, 2005. P. 63–70. <http://www.mt-archive.info/MTS-2005-Hutchins.pdf>.
12. *Hearne M., Way A.* Statistical machine translation: A guide for linguists and translators // *Lang. Linguist. Compass*, 2011. No. 5(5). P. 205–226.
13. *Costa Â., Ling W., Luís T., Correia R., Coheur L.* A linguistically motivated taxonomy for machine translation error analysis // *Machine Translation*, 2015. Vol. 29. Iss. 2. P. 127–161.
14. *Vilar D., Xu J., D'Haro L., Ney H.* Error analysis of statistical machine translation output // 5th Conference (International) on Language Resources and Evaluation Proceedings. — Genoa, Italy: European Language Resources Association (ELRA), 2006. <http://www.lrec-conf.org/proceedings/lrec2006/pdf/413.pdf.pdf>.
15. *Buntman N., Minel J.-L., Le Pesant D., Zatsman I.* Typology and computer modelling of translation difficulties // *Информатика и её применения*, 2010. Т. 4. Вып. 3. С. 77–83.
16. *Попкова Н. А., Инькова О. Ю., Зацман И. М., Кружков М. Г.* Методика построения моноэквиваленций в надкорпусной базе данных коннекторов // *Задачи современной информатики: Тр. 2-й конф.* — М.: ФИЦ ИУ РАН, 2015. С. 143–153.
17. *Зацман И. М., Инькова О. Ю., Кружков М. Г., Попкова Н. А.* Представление кроссязыковых знаний о коннекторах в надкорпусных базах данных // *Информатика и её применения*, 2016. Т. 10. Вып. 1. С. 106–118.
18. *Зализняк Анна А., Зацман И. М., Инькова О. Ю.* Надкорпусная база данных коннекторов: построение системы терминов // *Информатика и её применения*, 2017. Т. 11. Вып. 1. С. 100–108.
19. *Зацман И. М., Мамонова О. С., Щурова А. Ю.* Обратимость и альтернативность генерализации моделей перевода коннекторов в параллельных текстах // *Системы и средства информатики*, 2017. Т. 27. № 2. С. 125–142.
20. *In'kova Yu. O., Popkova N. A.* Statistical data as information source for linguistic analysis of Russian connectors // *Информатика и её применения*, 2017. Т. 11. Вып. 3. С. 123–131.

Поступила в редакцию 15.03.2018

Батенков Александр Александрович (р. 1957) — инженер-исследователь Орловского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

Бахтеев Олег Юрьевич (р. 1993) — аспирант Московского физико-технического института; исследователь ЗАО «Антиплагиат»

Бунтман Надежда Валентиновна (р. 1957) — кандидат филологических наук, доцент Московского государственного университета имени М. В. Ломоносова

Быковец Евгений Владимирович (р. 1994) — студент магистратуры факультета вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова

Васильева София Николаевна (р. 1993) — аспирант Московского авиационного института (национального исследовательского университета)

Гасилов Артур Владимирович (р. 1992) — младший научный сотрудник Орловского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

Грушо Александр Александрович (р. 1946) — доктор физико-математических наук, профессор, главный научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Егоров Артём Юрьевич (р. 1992) — студент магистратуры Московского института электроники и математики Национального исследовательского университета «Высшая школа экономики»

Жуков Дмитрий Олегович (р. 1965) — доктор технических наук, профессор, заведующий кафедрой Московского технологического университета МИРЭА

Зальцман Анастасия Дмитриевна (р. 1989) — преподаватель Московского технологического университета МИРЭА

Инькова Ольга Юрьевна (р. 1965) — доктор филологических наук, доцент кафедры русского языка и литературы филологического факультета Женевского университета

Кан Юрий Сергеевич (р. 1960) — доктор физико-математических наук, профессор Московского авиационного института (национального исследовательского университета)

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

Лаврентьев Виктор Владимирович (р. 1955) — кандидат физико-математических наук, научный сотрудник факультета вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова

Лебедев Алексей Викторович (р. 1971) — доктор физико-математических наук, доцент кафедры теории вероятностей механико-математического факультета Московского государственного университета им. М. В. Ломоносова

Лери Марина Муксумовна (р. 1969) — кандидат технических наук, научный сотрудник Института прикладных математических исследований КарНЦ РАН, ФИЦ «Карельский научный центр РАН»

Лесько Сергей Александрович (р. 1983) — кандидат технических наук, доцент Московского технологического университета МИРЭА

Маньяков Юрий Анатольевич (р. 1984) — старший научный сотрудник Орловского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

Мирзабеков Яхья Мирзабекович (р. 1983) — старший преподаватель Дагестанского государственного университета

Мистрюков Андрей Вадимович (р. 1988) — аспирант кафедры математической статистики факультета вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова

Назаров Леонид Владимирович (р. 1957) — кандидат физико-математических наук, старший научный сотрудник факультета вычислительной математики

ки и кибернетики Московского государственного университета им. М. В. Ломоносова

Нуриев Виталий Александрович (р. 1980) — кандидат филологических наук, старший научный сотрудник Института языкознания Российской академии наук; старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Огальцов Александр Владимирович (р. 1993) — студент Национального исследовательского университета «Высшая школа экономики»; исследователь ЗАО «Антиплагиат»

Павлов Юрий Леонидович (р. 1949) — доктор физико-математических наук, профессор, главный научный сотрудник Института прикладных математических исследований КарНЦ РАН, ФИЦ «Карельский научный центр РАН»

Сурина Альфия Адгамовна (р. 1987) — аспирант Института естественных и точных наук Южно-Уральского государственного университета

Тимонина Елена Евгеньевна (р. 1952) — доктор технических наук, профессор, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Тырсин Александр Николаевич (р. 1961) — доктор технических наук, заведующий кафедрой прикладной математики Уральского энергетического института, Уральский федеральный университет имени первого Президента России Б. Н. Ельцина, старший научный сотрудник Института экономики Уральского отделения Российской академии наук

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

Хватова Татьяна Юрьевна (р. 1971) — доктор экономических наук, профессор Санкт-Петербургского политехнического университета Петра Великого

Шестаков Олег Владимирович (р. 1976) — доктор физико-математических наук, 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

Шихиев Шукур Бабаевич (р. 1951) — кандидат физико-математических наук, доцент Дагестанского государственного университета

Шнурков Петр Викторович (р. 1953) — кандидат физико-математических наук, доцент Московского института электроники и математики Национального исследовательского университета «Высшая школа экономики»

Шоргин Сергей Яковлевич (р. 1952) — доктор физико-математических наук, профессор, главный научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Яковлев Олег Альбертович (р. 1992) — младший научный сотрудник 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://www.translit.net/ru/bgn/> бесплатной программой транслитерации русского текста в латиницу.

Список литературы “References” приводится полностью отдельным блоком, повторяя все позиции из списка литературы к русскоязычной части, независимо от того, имеются или нет в нем иностранные источники. Если в списке литературы к русскоязычной части есть ссылки на иностранные публикации, набранные латиницей, они полностью повторяются в списке “References”.

Ниже приведены примеры ссылок на различные виды публикаций в списке “References”.

Описание статьи из журнала:

Zagurenko, A. G., V. A. Korotovskikh, A. A. Kolesnikov, A. V. Timonov, and D. V. Kardymon. 2008. Tekhniko-ekonomicheskaya optimizatsiya dizayna gidrorazryva plasta [Technical and economic optimization of the design of hydraulic fracturing]. *Neftyanoe hozyaystvo [Oil Industry]* 11:54–57.

Zhang, Z., and D. Zhu. 2008. Experimental research on the localized electrochemical micromachining. *Rus. J. Electrochem.* 44(8):926–930. doi:10.1134/S1023193508080077.

Описание статьи из электронного журнала:

Swaminathan, V., E. Lepkoswka-White, and B. P. Rao. 1999. Browsers or buyers in cyberspace? An investigation of electronic factors influencing electronic exchange. *JCMC* 5(2). Available at: <http://www.ascusc.org/jcmc/vol5/issue2/> (accessed April 28, 2011).

Описание 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.net/ru/bgn/>).
- Please take into account the following examples of Russian references appearance:
- Article in journal:**
Zhang, Z., and D. Zhu. 2008. Experimental research on the localized electrochemical micromachining. *Rus. J. Electrochem.* 44(8):926–930. doi:10.1134/S1023193508080077.
- Journal article in electronic format:**
Swaminathan, V., E. Lepkoswka-White, and B. P. Rao. 1999. Browsers or buyers in cyberspace? An investigation of electronic factors influencing electronic exchange. *JCMC* 5(2). Available at: <http://www.ascusc.org/jcmc/vol5/issue2/> (accessed April 28, 2011).
- Article from the continuing publication (collection of works, proceedings):**
Astakhov, M. V., and T. V. Tagantsev. 2006. Eksperimental’noe issledovanie prochnosti soedineniy “stal’–kompozit” [Experimental study of the strength of joints “steel–composite”]. *Trudy MGTU “Matematicheskoe modelirovanie slozhnykh tekhnicheskikh sistem” [Bauman MSTU “Mathematical Modeling of Complex Technical Systems” Proceedings]*. 593:125–130.
- Conference proceedings:**
Usmanov, T. S., A. A. Gusmanov, I. Z. Mullagalin, R. Ju. Muhametshina, A. N. Chervyakova, and A. V. Sveshnikov. 2007. Osobennosti proektirovaniya razrabotki mestorozhdeniy s primeneniem gidrorazryva plasta [Features of the design of field development with the use of hydraulic fracturing]. *Trudy 6-go Mezhdunarodnogo Simpoziuma “Novye resursosberegayushchie tekhnologii nedropol’zovaniya i povysheniya neftegazoidachi” [6th Symposium (International) “New Energy Saving Subsoil Technologies and the Increasing of the Oil and Gas Impact” Proceedings]*. Moscow. 267–272.
- Books and other monographs:**
Lindorf, L. S., and L. G. Mamikonians, eds. 1972. *Ekspluatatsiya turbogeneratorov s neposredstvennym okhlazhdeniem [Operation of turbine generators with direct cooling]*. Moscow: Energy Publs. 352 p.
- Dissertation and Thesis:**
Kozhunova, O. S. 2009. Tekhnologiya razrabotki semanticheskogo slovary informatsionnogo monitoringa [Technology of development of semantic dictionary of information monitoring system]. PhD Thesis. Moscow: IPI RAN. 23 p.
- State standards and patents:**
GOST 8.586.5-2005. 2007. Metodika vypolneniya izmereniy. Izmerenie raskhoda i kolichestva zhidkostey i gazov s pomoshch’yu standartnykh suzhayushchikh ustroystv [Method of measurement. Measurement of flow rate and volume of liquids and gases by means of orifice devices]. M.: Standardinform Publs. 10 p.
Bolshakov, M. V., A. V. Kulakov, A. N. Lavrenov, and M. V. Palkin. 2006. Sposob orientirovaniya po krenu letatel’nogo apparata s opticheskoy golovkoy samonavedeniya [The way to orient on the roll of aircraft with optical homing head]. Patent RF No. 2280590.
- References in Latin transcription are presented in the original language.
References in the text are numbered according to the order of their first appearance; the number is placed in square brackets.
All items from the reference list should be cited.
10. Manuscripts and additional materials are not returned to Authors by the Editorial Board.
11. Submissions of files by e-mail must include:
- the journal title and author’s name in the “Subject” field;
 - an article and additional materials have to be attached using the “attach” function;
 - an electronic version of the article should contain the file with the text and a separate file with figures.
12. “Informatics and Applications” journal is not a profit publication. There are no charges for the authors as well as there are no royalties.

Editorial Board address:

FRC CSC RAS, 44, block 2, Vavilov Str., Moscow 119333, Russia
Ph.: +7 (499) 135 86 92, Fax: +7 (495) 930 45 05
e-mail: rust@ipiran.ru (to Prof. Rustem Seyful-Mulyukov)
<http://www.ipiran.ru/english/journal.asp>