

*Intellectual Technologies
on Transport
No 3*



*Интеллектуальные технологии
на транспорте
№ 3*

*Санкт-Петербург
St. Petersburg
2015*

Интеллектуальные технологии на транспорте № 3, 2015

Сетевой электронный научный журнал, свободно распространяемый через Интернет.
Публикует статьи на русском и английском языках с результатами исследований и практических достижений
в области интеллектуальных технологий и сопутствующих им научных исследований

Журнал основан в 2015 году

Учредитель и издатель

Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования
«Петербургский государственный университет путей сообщения Императора Александра I» (ФГБОУ ВПО ПГУПС)

Сопредседатели редакционного совета

Илларионов А. В., вице-президент ОАО «РЖД», Москва, РФ
Панычев А. Ю., ректор ПГУПС, С-Петербург, РФ

Главный редактор

Хомоненко А. Д., проф., С-Петербург, РФ

Редакционный совет

Глухов А. П., зам. нач. Деп. без. ОАО «РЖД», Москва, РФ
Дудин А. Н., д. т. н., проф., БГУ, Минск, Белоруссия
Корниенко А. А., проф., ПГУПС, С-Петербург, РФ
Ковалец П., проф., Тех. Унив-тет, Варшава, Польша
Лыков Р. Ю., нач. ГВЦ ОАО «РЖД», Москва, РФ
Меркурьев Ю. А., проф., РТУ, Рига, Латвия

Нестеров В. М., проф., ген. дир. ЦР EMC2, С-Петербург
Пустарнаков В. Ф., ген. дир. «Газинформсервис»,
С-Петербург, РФ
Титова Т. С., проф., прорект. ПГУПС, С-Петербург, РФ
Федоров А. Р., ген. дир. «ДигДез», С-Петербург, РФ
Юсупов Р. М., проф., чл.-корр. РАН, С-Петербург, РФ

Редакционная коллегия

Бубнов В. П., проф., С-Петербург, РФ – зам. гл. ред.
Ададунов С. Е., проф., С-Петербург, РФ
Атилла Элчи, проф., университет Аксарай, Турция
Безродный Б. Ф., проф., МАДИ, Москва, РФ
Благовещенская Е. А., проф., С-Петербург, РФ
Булавский П. Е., д. т. н., доц., С-Петербург, РФ
Василенко М. Н., проф., С-Петербург, РФ
Гуда А. Н., проф., Ростов-на-Дону, РФ
Железняк В. К., проф., ПГУ, Белоруссия
Заборовский В. С., проф., С-Петербург, РФ
Зегжда П. Д., проф., С-Петербург, РФ
Канаев А. К., д. т. н., доц., С-Петербург, РФ
Когут А. Т., проф., Омск, РФ
Котенко А. Г., д. т. н., доц., С-Петербург, РФ
Куренков П. В., проф., Москва, РФ
Лецкий Э. К., проф., Москва, РФ

Мирзоев Т. асс. проф., Джорджия, США
Наседкин О. А., доц., С-Петербург, РФ
Никитин А. Б., проф., С-Петербург, РФ
Охтилев М. Ю., проф., С-Петербург, РФ
Соколов Б. В., проф., С-Петербург, РФ
Таранцев А. А., проф., С-Петербург, РФ
Утепбергенов И. Т., проф., Алма-Аты, Казахстан
Филипченко С. А., доц., Москва, РФ
Фозилов Ш. Х., проф., Ташкент, Узбекистан
Фу-Ниан Ху, проф, Джангсу, Китай
Хабаров В. И., проф., Новосибирск, РФ
Ходаковский В. А., проф., С-Петербург, РФ
Чехонин К. А., проф., Хабаровск, РФ
Яковлев В. В., проф., С-Петербург, РФ
Ялышев Ю. И., проф., Екатеринбург, РФ

Адрес редакции

190031 Санкт-Петербург, Московский пр., 9, ПГУПС
email: itt-pgups@yandex.ru, сайт: <http://itt-pgups.ru/>

ISSN 2413-2527

Журнал зарегистрирован Федеральной службой по надзору в сфере связи и массовых коммуникаций,
свидетельство Эл №ФС77-61707 от 07 мая 2015 г.

Журнал зарегистрирован в Российском индексе научного цитирования (РИНЦ)

Все выпуски журнала представлены в электронной библиотеке «CyberLeninka»
<http://cyberleninka.ru/journal/n/intellektualnye-tehnologii-na-transporte>

© Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования
«Петербургский государственный университет путей сообщения Императора Александра I», 2015.
Разрешается воспроизведение в прессе, а также сообщение в эфир или по кабелю опубликованных в составе периодиче-
ского издания-журнала «Интеллектуальные технологии на транспорте» статей по текущим экономическим, политическим,
социальным и религиозным вопросам с обязательным указанием автора статьи и сетевого электронного научного
периодического издания журнала «Интеллектуальные технологии на транспорте»

До 25.11.2015 г. идёт приём статей для 4-го номера журнала

Intellectual Technologies on Transport

Issue № 3, 2015

Network electronic scientific journal, open access. It publishes articles in Russian and English with the results of research and practical achievements in the field of intelligent technologies and associated research

Founded in 2015

Founder and Publisher

Federal State Educational Institution of Higher Professional Education
«Petersburg State Transport University»

Co-chairs of the Editorial Council

Illarionov A. V., Vice-President of JSC «Russian Railways», Moscow, Russia
Panychev A. Yu., rector of PSTU, St. Petersburg, Russia

Editor-in-Chief

Khomonenko A. D., Prof., St. Petersburg, Russia

Editorial Council Members

Glukhov A. P., deputy. beg. Dep. no. JSC «Russian Railways», Moscow, Russia	Nesterov V. M., Prof., director general at Russian EMC2 development center, St. Petersburg
Dudin A. N., Prof., BSU, Minsk, Belarus	Pustarnakov V. F., CEO at «Gazinformservice» LTD., St. Petersburg, Russia.
Kornienko A. A., Prof., PSTU, St. Petersburg, Russia	Titova T. S., Prof., PSTU, St. Petersburg, Russia
Kovalets P., Prof., Tech. University, Warsaw, Poland	Fedorov, CEO at «Digital Design» LTD., St. Petersburg, Russia
Lykov R. Yu., director at MCC – branch of JSC «Russian Railways», Russia	Yusupov R. M., Prof., Corr. Member of RAS, St. Petersburg, Russia
Merkuryev Yu. A., Prof., Academician of the Latvian Academy of Sciences, Riga, Latvia	

Editorial Board Members

Bubnov V. P., Prof., St. Petersburg, Russia – Deputy Editor-in-Chief	Letsky Ad. K., Prof., Moscow, Russia
Adadurov S. E., Prof., St. Petersburg, Russia	Mirzoev T. Ass. Prof., Georgia, USA
Attila Elci, Prof., Aksaray, Turkey	Nasedkin O. A., Ass. Prof., St. Petersburg, Russia
Bezrodny B. F., Prof., Moscow, Russia Blagoveshenskaya E. A., Prof., St. Petersburg, Russia	Nikitin A. B., St. Petersburg, Russia
Bulavsky P. E., Dr. Sc., Ass. Prof., St. Petersburg, Russia	Okhtilev M. Yu., Prof., St. Petersburg, Russia
Vasilenko M. N., Prof., St. Petersburg, Russia	Sokolov B. V., Prof., Dr. Sci., St. Petersburg, Russia
Guda A. N., Prof., Rostov-on-Don, Russia	Tarantsev A. A., Prof., St. Petersburg, Russia
Geleznyak V. K., Prof., ПГУ, Белоруссия	Utepbergenov I. T., Prof., Alma-Ata, Khazakhstan
Zaborovsky V. S., Prof., St. Petersburg, Russia	Filipchenko S. A., Ass. Prof., Moscow, Russia
Zegzda P. D., Prof., St. Petersburg, Russia	Fozilov S. Kh., Prof., Tashkent, Uzbekistan
Kanayev A. K., Ass. Prof., St. Petersburg, Russia	Fu-Nian Hu, Prof., Jiangsu, China
Kogut A. T., Prof., Omsk, Russia	Khabarov V. I., Prof., Novosibirsk, Russia
Kotenko A. G., Dr. Sc., Ass. Prof., St. Petersburg, Russia	Khodakosky V. A., Prof., St. Petersburg, Russia
Kurenkov P. V., Prof., Moscow, Russia	Chekxonin K. A., Prof., Khabarovsk, Russia
	Jakovlev V. V., Prof., St. Petersburg, Russia
	Jalyshev Yu. I., Prof., Ekaterinburg, Russia

Adress

190031, St. Petersburg, Moskovskiy pr., 9, 2–108
email: itt-pgups@yandex.ru, <http://itt-pgups.ru/>

ISSN 2413-2527

The journal is registered by the Federal Service for Supervision of Communications and Mass Media, EL №FS77-61707 testimony from May 7, 2015

The journal is registered in the Russian Science Citation Index (RSCI)

All issues of the journal are presented in electronic library «CyberLeninka»
<http://cyberleninka.ru/journal/n/intellektualnye-tehnologii-na-transporte>

© Federal State Educational Institution of Higher Professional Education «Petersburg State Transport University», 2015.
The reproduction in the press, as well as a message broadcast or cable published as part of the periodical – journal «Intellectual Technologies on Transport» articles on current economic, political, social and religious issues with the obligatory indication of the author, and the network of electronic scientific periodical journal «Intellectual Technologies on Transport»

Articles for the 4th issue of the journal are received until 25.11.2015.

Содержание

<i>Баушев А. Н.</i> Об информационном регулировании потоков автомобильного транспорта	5
<i>Рыжиков Ю. И., Хомоненко А. Д.</i> Расчет многоканальных систем обслуживания с абсолютным и относительным приоритетами на основе инвариантов отношения	11
<i>Пак А. А., Нарынов С. С. Касымова Д. Т., Утепбергенов И. Т.</i> Статистический метод выделения временных паттернов из естественных языков	17
<i>Захаров А. И., Загайнов А. И., Ходаковский В. А.</i> Мультифрактальный анализ: выявление границ применения при исследовании временных рядов	24
<i>Канаев А. К., Сахарова М. А., Бенета Э. В.</i> Определение вероятностно-временных характеристик процесса функционирования системы управления сетью передачи данных при поддержании показателей QoS	30
<i>Дорожко И. В., Тарасов А. Г., Барановский А. М.</i> Оценка надежности структурно сложных технических комплексов с помощью моделей байесовских сетей доверия в среде GeNIe	36
<i>Благовещенская Е. А., Зуев Д. В.</i> Модификация метода RPROP для решения задач распознавания образов	46

Contents

<i>Baushev A. N.</i> On the Information Regulation of Road Traffic Flows	5
<i>Ryzhikov Y.I., Khomonenko A. D.</i> Calculation of Multi-Channel Queueing Systems with Absolute and Relative Priorities on the Basis of Invariants Relationship	11
<i>Pak A. A., Narynov S. S., Kasymova D. T., Utepbergenov I. T.</i> Statistical Method for Isolating Temporal Patterns of Natural Languages	17
<i>Zakharov A. I., Zagaynov A. I., Khodakovsky V. A.</i> Multifractal Analysis: Identifying the Boundaries Application in the Study of Time Series	24
<i>Kanaev A. K., Sakharova M. A., Beneta E. V.</i> Determination Probability-Time Characteristics of the Process Functioning Management System Network While Maintaining QoS Performance	30
<i>Dorozhko I. V., Tarasov A. G., Baranovsky A. M.</i> Estimation to Reliability of Structural Complex Technical Systems by Using Bayesian Networks Belief Models in the Environment of GeNIe	36
<i>Blagoveshchenskaya E. A., Zuev D. V.</i> RPROP Method Modification for Pattern Recognition	46

Об информационном регулировании потоков автомобильного транспорта

Баушев А. Н.

Петербургский государственный университет
пути сообщения Императора Александра I
Санкт-Петербург, Россия
banban2008@yandex.ru

Аннотация. Предложен метод решения задачи о построении потока минимальной стоимости с нелинейной функцией стоимости. Метод заключается в сведении к задаче с линейной функцией стоимости в некоторой вспомогательной мультисети, структура которой определяется видом целевой функции в исходной задаче.

Ключевые слова: потоки в сетях, построение потока, минимальная стоимость, эквивалентность экстремальных задач.

ВВЕДЕНИЕ

Интенсивное развитие коммуникационных и информационных технологий породило новые способы регулирования автотранспортных потоков. Появление в автомобилях навигационного оборудования и использование интернета позволяет водителю оценить текущее состояние сети дорог и выбрать маршрут движения с учётом этого состояния. Более того, воспользовавшись, например, услугами компании Яндекс, водитель может в готовом виде получить оптимальный по времени маршрут следования и ввести его в свой навигатор.

Однако алгоритм последовательного поиска скорейшего маршрута может привести к далеко не оптимальному потоку, если в качестве критерия оптимальности выбрано суммарное (или среднее по потоку) время движения от начальной до конечной точки маршрута. Среднее время перемещения потока по участкам транспортной сети очевидно зависит от величины потока на них. Эту зависимость необходимо учитывать при оптимальном распределении потока по маршрутам следования.

Сказанное иллюстрирует искусственный пример (рис. 1), где первое число на дуге обозначает среднее время перемещения потока величины $\theta = 1$, а второе – среднее время перемещения потока величины $\theta \geq 2$.

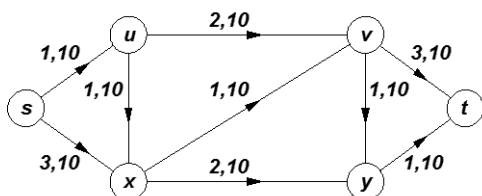


Рис. 1. Двухполюсная транспортная сеть

Рассмотрим задачу перемещения потока величины $\theta = 2$ из узла s в узел t по сети (рис. 1). Применяя алгоритм последовательного поиска кратчайшего по времени пути, находим:

$$\mu_1 = s \rightarrow u \rightarrow x \rightarrow v \rightarrow y \rightarrow t; \tau(\mu_1) = 5;$$

$$\mu_2 = s \rightarrow x \rightarrow y \rightarrow t; \tau(\mu_2) = 3 + 2 + 10 = 15.$$

Суммарное время перемещения составит 20, а среднее по потоку – 10.

Однако перемещение единичных потоков по путям $\mu_3 = s \rightarrow x \rightarrow y \rightarrow t$ и $\mu_4 = s \rightarrow u \rightarrow v \rightarrow t$ даёт суммарное время 12 и среднее по потоку – 6.

Приведённый пример наводит на мысль, что для интенсивного потока автотранспорта следует использовать «пакетный» подход к удовлетворению запросов водителей на оптимальный маршрут перемещения по сети дорог. В течение некоторого (малого) интервала времени собираются запросы от водителей, затем определяется величина θ перемещаемого потока. Далее находится оптимальное распределение этого потока по возможным маршрутам следования от пункта запроса до пункта назначения.

Математическая модель возникающей при этом оптимизационной задачи имеет вид

$$\tau(\varphi) = \sum_{a \in A} \tau_a(\varphi(a)) \rightarrow \min, \varphi \in \Phi_\theta, \quad (1)$$

где Φ_θ – множество всех потоков в двухполюсной транспортной сети $\Gamma = \langle X, A \rangle$, которые имеют величину θ ; $\varphi(a)$ – величина потока на дуге $a \in A$; τ_a – функция, описывающая зависимость времени перемещения по дуге a от величины потока.

Если

$$\tau_a(\varphi) = \gamma_a \varphi; \varphi \geq 0, \quad (2)$$

где γ_a – фиксированные числа, заданные на дугах сети, то задача (1) является классической задачей построения потока минимальной стоимости в транспортной сети, алгоритмы решения которой хорошо известны [5]. Однако для потоков автотранспорта функция (2), очевидно, неадекватно описывает фактическую зависимость. Более адекватной представляется функция

$$\tau_a(\varphi) = \begin{cases} 0, & \text{если } \varphi < 1, \\ \beta_a + \gamma_a(\varphi - 1), & \text{если } \varphi \geq 1. \end{cases} \quad (3)$$

Слагаемое β_a в (3) соответствует времени перемещения единичного потока по дуге при полностью свободном соответствующем участке дорожной сети. При этом $\beta_a > 0, \gamma_a \geq 0$.

Функция (3) является частным случаем более общей зависимости времени перемещения по участку дорожной сети

от величины потока на нём, графически представленной на рис. 2.

Настоящая статья посвящена описанию алгоритма решения задачи (1) с целевой функцией, слагаемые которой имеют вид (3). Предлагаемый подход к решению основан на сведении этой задачи к классической задаче построения потока минимальной стоимости в некоторой вспомогательной мультисети. Аналогичный метод можно применять и для решения задачи (1) с более сложными целевыми функциями, в частности, с целевой функцией, слагаемые которой имеют вид, представленный на рис. 2. Однако во избежание громоздкости описания соответствующей конструкции мы ограничимся изложением метода для указанного частного случая.

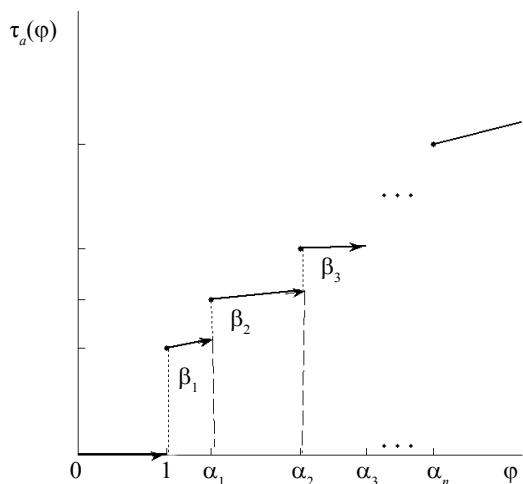


Рис. 2. Зависимость времени перемещения по участку сети

ПРЕДВАРИТЕЛЬНЫЕ СВЕДЕНИЯ

В этом разделе приведем сведения о потоках в сетях и об эквивалентности экстремальных задач, которые необходимы для понимания дальнейшего изложения [1, 2].

Пусть $G = \langle X, A \rangle$ – ориентированный граф. В этой записи X обозначает множество вершин графа G , а A – множество дуг. Для вершины $x \in X$ символ $A_-(x)$ будет обозначать множество всех дуг, для которых вершина x является конечной, а символ $A_+(x)$ – множество всех дуг, для которых вершина x является начальной.

Если каждой дуге $a \in A$ поставлено в соответствие некоторое число или специальный символ $c(a)$, то граф G называется ориентированной сетью, а вершины графа – узлами сети. Обычно в качестве специального символа в зависимости от контекста решаемой задачи используют символы $\pm\infty$ с естественной интерпретацией этих символов. Для обозначения сети применяется запись $\Gamma = \langle X, A, c \rangle$.

Ориентированная сеть $\Gamma = \langle X, A, c \rangle$ называется двухполюсной транспортной сетью, если:

- 1) существует единственный узел $s \in X$, называемый источником, для которого $A_-(s) = \emptyset$;
- 2) существует единственный узел $t \in X$, называемый стоком, для которого $A_+(t) = \emptyset$;
- 3) каждый узел сети лежит на некотором пути от источника к стоку;
- 4) для каждой дуги $a \in A$ задана величина $c(a) \geq 0$, называемая пропускной способностью дуги a .

Узлы s и t двухполюсной транспортной сети называют также полюсами сети, остальные узлы – промежуточные.

В практических ситуациях, не нарушая общности, всегда можно считать, что пропускная способность дуги либо выражается целым неотрицательным числом, либо равна $+\infty$.

Функция φ , заданная на дугах транспортной сети $\Gamma = \langle X, A, c \rangle$, называется (допустимым) потоком в этой сети, если для любой дуги $a \in A$ выполняются неравенства

$$0 \leq \varphi(a) \leq c(a) \quad (4)$$

и для любого промежуточного узла x выполняется условие баланса между общим потоком, «вытекающим» в узел, и общим потоком, «вытекающим» из узла:

$$\sum_{a \in A_-(x)} \varphi(a) = \sum_{a \in A_+(x)} \varphi(a). \quad (5)$$

Из соотношений (5) вытекает равенство

$$\sum_{a \in A_-(t)} \varphi(a) = \sum_{a \in A_+(s)} \varphi(a). \quad (6)$$

Общее значение сумм в равенстве (6) называется величиной потока.

Множество дуг $a \in A$, для которых $\varphi(a) > 0$, называется носителем потока φ . Целочисленный поток φ , имеющий единичную величину, называется единичным. Носитель единичного потока в двухполюсной транспортной сети представляет собой некоторый путь μ от источника к стоку в графе G . Единичный поток с носителем μ мы будем обозначать I_μ .

Любой поток в двухполюсной транспортной сети может быть представлен как

$$\varphi = \theta_1 \cdot I_{\mu_1} + \dots + \theta_k \cdot I_{\mu_k}, \quad (7)$$

где μ_1, \dots, μ_k – некоторые пути от источника к стоку, а величины $\theta_1, \dots, \theta_k$ положительны и их сумма равна величине потока θ .

Отметим, что целочисленные потоки, имеющие величину θ , можно представить выражением (7) либо

$$\varphi = I_{\pi_1} + \dots + I_{\pi_\theta} + I_{\delta_1} + \dots + I_{\delta_m}, \quad (8)$$

где π_1, \dots, π_θ – простые (т.е. не содержащие циклов) пути от источника к стоку, а $\delta_1, \dots, \delta_m$ – контуры (т.е. простые ориентированные циклы). При этом как в последовательности π_1, \dots, π_θ , так и в последовательности $\delta_1, \dots, \delta_m$, если последняя не пуста, возможны повторения [1].

Как отмечено во введении, предлагаемый метод решения задачи (1) заключается в её замене на эквивалентную классическую задачу построения потока минимальной стоимости в некоторой вспомогательной мультисети. Опишем кратко постановку и используемый нами метод решения классической задачи.

Предположим, что на множестве дуг транспортной сети $\Gamma = \langle X, A, c \rangle$ определена ещё одна функция γ , значение которой $\gamma(a)$ на дуге a интерпретируется как стоимость перемещения единичного потока по этой дуге.

Пусть Φ_θ обозначает множество допустимых потоков в сети Γ , имеющих величину θ . Величина

$$\gamma(\varphi) = \sum_{a \in A} \gamma(a)\varphi(a) \quad (9)$$

называется стоимостью потока φ , а экстремальная задача

$$\gamma(\varphi) \rightarrow \min, \varphi \in \Phi_\theta \quad (10)$$

называется задачей о построении потока величины θ , имеющего минимальную стоимость.

К задаче (10) могут быть сведены многие экстремальные задачи транспортного типа. В настоящее время известно несколько алгоритмов ее решения [5, 6]. Мы используем для этого классический алгоритм Басакера – Гоуэна [4].

В алгоритме Басакера – Гоуэна поток последовательно увеличивается (с нулевого потока) таким образом, что по завершении каждой итерации получается поток минимальной стоимости. Если в какой-то момент осуществление следующей итерации становится невозможным и при этом величина построенного потока меньше заданной величины θ , то это означает, что задача (10) не имеет решения, т. е. величина θ больше пропускной способности рассматриваемой сети. В противном случае итерации продолжают до тех пор, пока нужное количество θ не будет переправлено.

В процессе алгоритма исходная сеть дополняется вспомогательными дугами, ориентация которых противоположна ориентации соответствующих дуг исходной сети. Дуги исходной сети при этом называются прямыми, а вспомогательные дуги – обратными. Значение пропускной способности на обратной дуге полагается равным значению потока на соответствующей прямой дуге, а значение пропускной способности прямой дуги заменяется значением её остаточной пропускной способности.

Значение стоимости транспортировки единичного потока по обратной дуге полагается равным значению этой величины на прямой дуге, взятому с противоположным знаком. Если на каком-то шаге происходит насыщение прямой дуги, то значение стоимости транспортировки единичного потока по ней полагается равным $+\infty$, а если происходит насыщение обратной дуги (т. е. поток на прямой дуге становится нулевым), то эта дуга удаляется из сети. Построенная таким образом сеть называется модификацией исходной сети относительно протекающего в ней потока или, по терминологии [1], графом приращений.

Итерационная процедура алгоритма представляет собой последовательность таких шагов:

- 1) найти путь минимальной стоимости от источника к стоку в модифицированной сети;
- 2) определить пропускную способность пути, найденного на предыдущем шаге, т. е. остаточную пропускную способность дуг, входящих в рассматриваемый путь, минимальной величины;
- 3) определить минимальную величину пропускной способности пути и разность между заданным значением θ и величиной построенного до этого потока;
- 4) увеличить поток на прямых дугах рассматриваемого пути на величину, найденную на шаге 3, а поток на прямых дугах исходной сети, соответствующих обратным дугам пути, уменьшить на эту величину;
- 5) пересчитать пропускную способность и стоимость транспортировки единичного потока на дугах, соответствующих рассматриваемому пути, согласно описанным ранее правилам;
- 6) если на предыдущих шагах поток требуемой величины построен, то алгоритм останавливается, в противном случае процесс возвращается к шагу 1 с новым значением потока и с новой модификацией исходной сети.

Можно показать [1], что в результате применения итерационной процедуры, состоящей из шагов 1–6, будет либо построено решение задачи (10), либо обнаружено, что задача не имеет решения. Отметим, что отсутствие решения у задачи обнаруживается на шаге 1 в ситуации, когда величина построенного до этого потока меньше заданной, и при этом отсутствуют пути от источника к стоку с конечным значением стоимости транспортировки единичного потока по ним.

В настоящее время известны более быстрые алгоритмы построения потока минимальной стоимости в классическом случае [3], однако алгоритм Басакера – Гоуэна обладает рядом несомненных достоинств, из которых мы отметим два свойства, важных для излагаемого нами метода:

1) алгоритм применим в случае, когда вместо транспортных сетей рассматриваются транспортные мультисети, для которых аналогично задаче (10) ставится соответствующая задача;

2) поток, построенный по завершении каждой итерации алгоритма, состоящей из шагов 1–6, имеет минимальную стоимость среди стоимости всех допустимых потоков, величина которых равна величине построенного потока.

Возможность применять различные математические модели для решения задач оптимизации дает понятие эквивалентности экстремальных задач.

Пусть X и Y – произвольные непустые множества, f и g – числовые функции, заданные на множествах X и Y , соответственно. Экстремальные задачи

$$f(x) \rightarrow \min, x \in X \quad (11)$$

и

$$g(y) \rightarrow \min, y \in Y \quad (12)$$

называются эквивалентными, если существуют отображения $p: X \rightarrow Y$ и $q: Y \rightarrow X$ такие, что:

- 1) $g(p(x)) = f(x)$ для любого $x \in X$;
- 2) $f(q(y)) = g(y)$ для любого $y \in Y$.

Экстремальные задачи с пустыми множествами ограничений по определению считаются эквивалентными.

Если отображения p и q указаны явно, то задачи (11) и (12) также называются (p,q) -эквивалентными.

Ясно, что эквивалентные задачи одновременно разрешимы или неразрешимы. При этом если задачи (11) и (12) (p,q) -эквивалентны и x^* является решением задачи (11), то $y^* = p(x^*)$ – решение задачи (12), и наоборот: если y^* – решение задачи (12), то $x^* = q(y^*)$ – решение задачи (11).

ОПИСАНИЕ МЕТОДА

Пусть $\Gamma = \langle X, A \rangle$ – двухполосная транспортная сеть в задаче (1); s – источник сети; t – сток, а пропускные способности дуг неограниченны (равны $+\infty$). Построим двухполосную транспортную мультисеть $\Delta = \langle Y, B, c \rangle$ следующим образом. В качестве множества узлов мультисети возьмём то же множество, что и в сети Γ , т. е. положим $Y = X$, а каждую дугу $a \in A$ сети Γ представим в мультисети Δ парой дуг $(b^{(1)}, b^{(2)})$, начальный и конечный узлы которых совпадают, соответственно, с начальным и конечным узлом дуги a . Для каждой дуги a сети Γ в качестве пропускных способностей соответ-

ствующих дуг мультисети положим $c(b^{(1)}) = 1$, $c(b^{(2)}) = +\infty$, а в качестве стоимости транспортировки единичных потоков – $\delta(b^{(1)}) = \beta(a)$, $\delta(b^{(2)}) = \gamma_a$, где числа $\beta(a)$ и $\gamma(a)$ фигурируют в формуле (3).

Пусть $T = (\tau_a)_{a \in A}$ – семейство функций, каждая из которых имеет вид (3).

Поток ψ в мультисети $\Delta = \langle Y, B, c \rangle$ будем называть T-ранжированным, если из условия $\psi(b^{(1)}) \leq 1$ следует $\psi(b^{(2)}) = 0$, а из условия $\psi(b^{(2)}) > 0$ следует $\psi(b^{(1)}) = 1$.

Пусть Ψ_θ^T – множество всех T-ранжированных потоков величины θ в мультисети Δ . Рассмотрим оптимизационную задачу

$$\delta(\psi) = \sum_{b \in B} \delta(b) \psi(b) \rightarrow \min, \psi \in \Psi_\theta^T. \quad (13)$$

Покажем, что задачи (1) и (13) эквивалентны.

Пусть φ – поток величины θ в задаче (1). Построим поток ψ в мультисети Δ , определив его по соответствию между парами $(b^{(1)}, b^{(2)})$ дуг мультисети и дугами $a \in A$ исходной сети, следующим образом:

$$\psi(b^{(i)}) = \begin{cases} 0, & \text{если } \varphi(a) = 0, i = 1, 2; \\ 1, & \text{если } \varphi(a) > 0, i = 1; \\ \varphi(a) - 1, & \text{если } \varphi(a) > 0, i = 2. \end{cases}$$

Мы видим, что ψ – поток величины θ и что $\psi \in \Psi_\theta^T$. Положим $p(\varphi) = \psi$.

Пусть теперь ψ – T-ранжированный поток величины θ в мультисети Δ , $(b^{(1)}, b^{(2)})$ – пара дуг мультисети, соответствующая дуге $a \in A$. Определим поток φ в сети Γ равенством

$$\varphi(a) = \psi(b^{(1)}) + \psi(b^{(2)}).$$

Таким образом, φ – поток величины θ в сети Γ . Положим $q(\psi) = \varphi$.

Если $p(\varphi) = \psi$, то

$$\begin{aligned} \delta(\psi) &= \sum_{b^{(1)} \in B} \delta(b^{(1)}) \psi(b^{(1)}) + \sum_{b^{(2)} \in B} \delta(b^{(2)}) \psi(b^{(2)}) = \\ &= \sum_{\substack{a \in A, \\ \varphi(a) > 0}} \beta(a) \cdot 1 + \sum_{\substack{a \in A, \\ \varphi(a) > 0}} \gamma(a)(\varphi(a) - 1) = \sum_{a \in A} \tau_a(\varphi) = \tau(\varphi). \end{aligned}$$

Если $q(\psi) = \varphi$, то аналогично $\tau(\varphi) = \delta(\psi)$. Таким образом, задачи (1) и (13) эквивалентны.

Для решения задачи (13) можно применить алгоритм Басакера – Гоуэна, если его модифицировать таким образом, чтобы потоки, полученные по завершении каждой итерации, были T-ранжированными. Модификация заключается в том, что для каждой пары дуг $(b^{(1)}, b^{(2)})$ мультисети обратная дуга проводится только к одной дуге из этой пары.

При этом если $\psi(b^{(2)}) > 0$, то проводится дуга, обратная к дуге $b^{(2)}$, а дуга, обратная к дуге $b^{(1)}$, удаляется из сети. Дуга, обратная к дуге $b^{(1)}$, существует в том и только в том случае, когда она насыщена, а поток по дуге $b^{(2)}$ равен нулю. Значения пропускных способностей и стоимости на обратных дугах определяются так же, как в первоначальном варианте алгоритма Басакера – Гоуэна. Несложно убедиться, что

после такой модификации потоки, полученные по завершении каждой итерации, будут T-ранжированными.

Отметим, что если для поиска кратчайшего пути в ориентированной сети с весами дуг произвольных знаков применяется алгоритм Беллмана – Мура [7], то время работы описанного алгоритма оценивается сверху величиной $\theta \cdot O(mn)$, где m – число дуг в исходной сети; n – число узлов; θ – величина потока.

ПРИМЕР

Проиллюстрируем работу алгоритма построением оптимального потока величины $\theta = 2$ в транспортной сети, изображённой на рис. 1. Зависимость времени перемещения по дугам сети от величины перемещаемого по ним потока имеет вид, представленный на рис. 3.

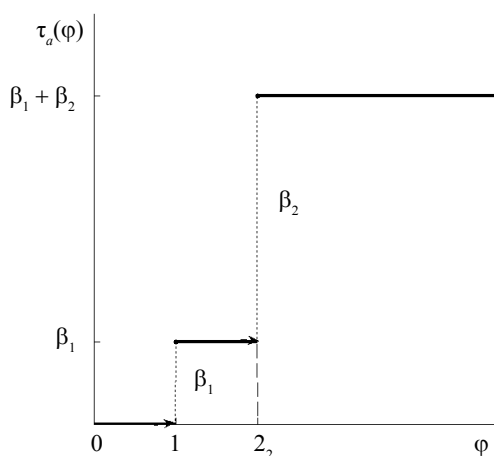


Рис. 3. Зависимость времени перемещения по дуге от величины потока

Несмотря на то, что зависимость $\tau_a(\varphi)$ в рассматриваемом случае имеет вид, отличный от (3), описанный нами метод применим, как утверждалось во введении, и в этом случае.

Каждую дугу a исходной сети представим в мультисети тремя дугами: $b^{(1)}, b^{(2)}, b^{(3)}$. Дуга $b^{(1)}$ отвечает скачку в точке 1, её пропускная способность равна 1, а стоимость транспортировки единичного потока – β_1 . Дуга $b^{(2)}$ отвечает скачку в точке 2, её пропускная способность равна 1, а стоимость транспортировки единичного потока – β_2 . Дуга $b^{(3)}$ отвечает линейному участку графика на множестве $\{\varphi \geq 2\}$, её пропускная способность равна $+\infty$, а стоимость транспортировки единичного потока равна нулю. Поскольку рассматриваются только целочисленные потоки, линейный участок графика на множестве $\{1 \leq \varphi < 2\}$ игнорируется.

Индексы дуг мультисети в соответствующей триаде будем называть рангами в представлении дуги исходной сети. Поток в мультисети называется ранжированным, если положительный поток по дугам более высокого ранга возможен только в том случае, когда дуги меньшего ранга насыщены.

Отметим, что при программной реализации алгоритма операция удаления дуги происходит через замену стоимости транспортировки единичного потока по ней на $+\infty$ так, что поиск кратчайшего пути (шаг 1) фактически осуществляется в подсети, образованной дугами, для которых указанные

величины не равны $+\infty$. Эту подсеть мы будем называть несущей подсетью мультисети. Несущая подсеть обновляется после каждой итерации.

На рис. 4–6 показаны несущие подсети, возникающие в процессе применения алгоритма для построения оптимального потока величины $\theta = 2$ в транспортной сети, изображённой на рис. 1. На рис. 4 первое число на дугах – стоимость транспортировки единичного потока, второе – пропускная способность, третье (в скобках) – ранг дуги в мультисети. Путь $\mu_1^* = s \rightarrow u \rightarrow x \rightarrow v \rightarrow y \rightarrow t$ – кратчайший. Время транспортировки единичного потока по этому пути равно $\tau(\mu_1^*) = 5$. На рис. 5 кратчайший путь $\mu_2^* = s \rightarrow x \rightarrow u \rightarrow v \rightarrow x \rightarrow y \rightarrow v \rightarrow t$. Время транспортировки единичного потока по этому пути равно $\tau(\mu_2^*) = 3 - 1 + 2 - 1 + 2 - 1 + 3 = 7$. Суммарное время движения оптимального потока равно $\tau(\mu_1^*) + \tau(\mu_2^*) = 5 + 7 = 12$. На рис. 6 по прямым дугам, для которых имеются обратные дуги, течёт единичный поток.

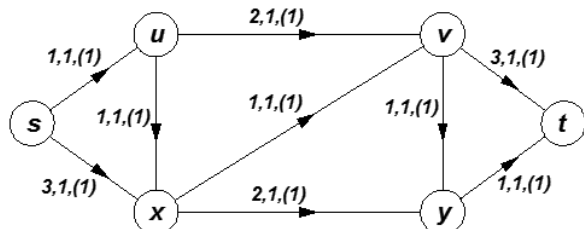


Рис. 4. Несущая подсеть перед началом 1-й итерации

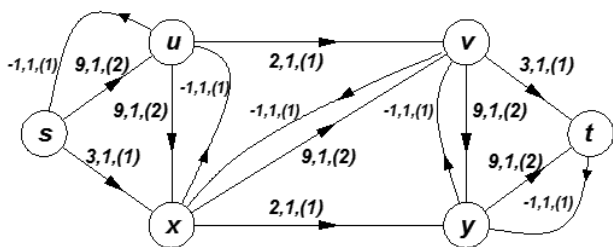


Рис. 5. Несущая подсеть перед началом 2-й итерации

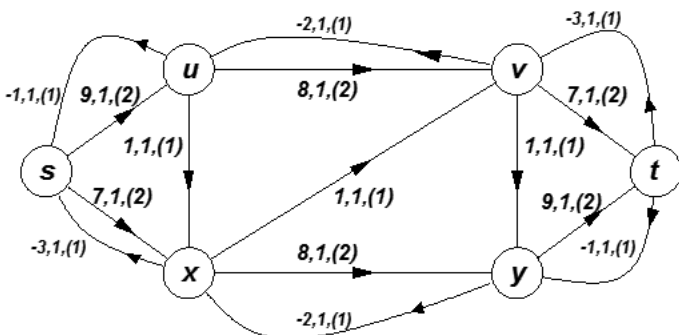


Рис. 6. Несущая подсеть после завершения 2-й итерации

ЗАКЛЮЧЕНИЕ

Задача построения потока минимальной стоимости в транспортной сети играет важную роль в решении проблемы оптимизации транспортных потоков. В интересных для практики случаях целевые функции в соответствующей задаче оказываются нелинейными, что не позволяет непосредственно использовать классические алгоритмы для её решения. Однако если целевая функция имеет аддитивную структуру, то задача построения потока минимальной стоимости оказывается эквивалентной классической задаче для некоторой специальным образом сконструированной транспортной мультисети.

ЛИТЕРАТУРА

1. Басакер Р. Конечные графы и сети / Р. Басакер, Т. Саати ; пер. с англ. – М.: Наука, 1974. – 368 с.
2. Баушев А. Н. Математическая модель многофазных железнодорожных грузоперевозок / А. Н. Баушев, А. Т. Осьминин, Л. А. Осьминин // Математическое моделирование. – 2013 – Т. 25, № 10. – С. 108-122.
3. Becker R. A simple efficient interior point method for min-cost flow / R. Becker, A. Kanenbauer // Lecture notes in computer sci. – 8889. – Pp. 753-765.
4. Busucker R. A Procedure for Determining a Family of Minimal-Cost Network Flow Patterns / R. Busucker, J. Gowen / Johns Hopkins Univ. O.R.O. Tech. – 1961. – P. 15.
5. Edmonds J. Theoretical improvements in algorithmic efficiency for network flow problems / J. Edmonds, R. M. Karp // Combinatorial Structures and Their Applications. – NY: Gordon and Breach, 1970. – Pp. 93-96.
6. Goldberg A. V. Finding minimum-cost circulations by successive approximation / A. V. Goldberg, R. E. Tarjan // Math. Oper. Res. – 1990. – Vol. 15, no. 3. – P. 430-466.
7. Moore E. F. The shortest path through a maze / E. F. Moore // Proc. Internat. Symp. Switching Theory, Part II. – Cambridge: Mass.: Harvard Univ. Press, 1957. – Pp. 285-292.

On the Information Regulation of Road Traffic Flows

Baushev A. N.

Petersburg State Transport University

St. Petersburg, Russia

banban2008@yandex.ru

Abstract. A method of solving of minimal-cost flow problem with nonlinear objective function is proposed. The method consists in the reduction of the problem to a problem with linear cost function on some subsidiary multi-net, the structure of which is determined by a target function in the source problem.

Keywords: flows in networks, minimal-cost flow problem, equivalence of optimization problems.

REFERENCES

1. Basaker R. R., Saati T. Finite Graphs and Networks. An Introduction with Application, McGraw-Hill Book Company, 1965. 294 p.

2. Baushev A. N., Os'minin A. T., Os'minin L. A. Mathematical Model of Multiphase Rail Freight [Matematicheskaya model' mnogofaznyh zheleznodorozhnyh gruzoperevozok]. *Matemat-*

icheskoe modelirovanie [Mathematical modeling], 2013, vol. 25, no. 10, pp. 108-122.

3. Becker R., Kanenbauer A. A Simple Efficient Interior Point Method for Min-Cost Flow. *Lecture Notes in Computer Sci.*, 8889, pp. 753-765.

4. Busucker R., Gowen J. A Procedure for Determining a Family of Minimal-Cost Network Flow Patterns. Johns Hopkins Univ. O. R. O. Tech., 1961. P. 15

5. Edmonds J., Karp R. M. Theoretical Improvements in Algorithmic Efficiency for Network Flow Problems. *Combinatorial Structures and Their Applications*. NY, Gordon and Breach, 1970. Pp. 93-96.

6. Goldberg A. V., Tarjan R. E. Finding minimum-cost circulations by successive approximation. *Math. Oper. Res.* 1990, vol. 15, no. 3, pp. 430-466.

7. Moore E. F. The Shortest Path Through a Maze. *Proc. Internat. Symp. Switching Theory, Part II*. Cambridge, Mass., Harvard Univ. Press, 1957. Pp. 285-292.

Расчет многоканальных систем обслуживания с абсолютным и относительным приоритетами на основе инвариантов отношения

Рыжиков Ю. И.

Санкт-Петербургский Институт информатики
и автоматизации РАН (СПИИРАН)
Санкт-Петербург, Россия
ryzhbox@yandex.ru

Хомоненко А. Д.

Петербургский государственный университет
путей сообщения Императора Александра I
Санкт-Петербург, Россия
khomon@mail.ru

Аннотация. Предлагается способ расчета среднего времени ожидания заявок разных типов в многоканальных системах обслуживания с абсолютным и относительным приоритетами, основанный на применении инвариантов отношения. Результаты сопоставляются с данными имитационного моделирования многоканальных приоритетных систем с произвольным распределением длительности обслуживания, аппроксимируемым распределениями Эрланга и гиперэкспоненциальным.

Ключевые слова: многоканальная система массового обслуживания, приоритет, инварианты отношения.

ВВЕДЕНИЕ

Системы обслуживания любого рода (технические и организационные) и назначения (производственные и информационные системы на транспорте, передача данных, здравоохранение, бытовое обслуживание, охрана общественного порядка, военное дело) время от времени нуждаются в профилактическом осмотре, проведении регламентных и ремонтных работ, а их персонал – в отдыхе.

Необходимость непрерывно выполнять ответственные функции, хотя бы и с уменьшенной производительностью, определяет применение многоканальных систем обслуживания. Такие системы, особенно в период функционирования в неполном составе, эксплуатируются в режиме, близком к насыщению, поэтому для достижения приемлемой оперативности обслуживания хотя бы по наиболее важным заявкам приходится вводить приоритеты.

Теория многоканальных систем обслуживания с приоритетами до настоящего времени практически разработана слабо. В статьях на эту тему задаются серьезные ограничения, как правило, предполагается экспоненциальное и к тому же одинаковое распределение времени обслуживания для всех типов заявок, см. например [9, 10, 12–14, 16, 18]).

Важный вопрос при исследовании приоритетных систем – установление характеристик распределения периода занятости, см. например [20]. На этой основе удастся рассчитать средние характеристики времени ожидания и пребывания заявок в многоканальных приоритетных системах.

В частности, в работах [6, 15] предложены численные решения для расчета среднего времени ожидания и пребы-

вания заявок в многоканальных системах с абсолютным и относительным приоритетом и с неоднородными классами заявок.

В [5] намечена свободная от названного выше ограничения схема решения задачи для средних значений ожидания (пребывания) в системе заявок каждого вида и демонстрируются удовлетворительные результаты ее применения. Однако существенными элементами этой схемы являются эвристики, полученные по результатам имитационного моделирования, не гарантирующие достаточной точности.

Ниже предлагается более надежная методика приближенного численного расчета средних значений в n -канальной системе обслуживания с абсолютным (прерыванием с дообслуживанием) и с относительным приоритетом. Методика основана на применении инвариантов отношения [1] и верифицируется с помощью имитационных моделей. Типы заявок нумеруются в порядке убывания приоритетов.

ИНВАРИАНТЫ ТЕОРИИ ОЧЕРЕДЕЙ

Согласно определению из [1], под инвариантами теории очередей понимают соотношения, определяемые не распределениями в целом, а некоторым количеством их начальных моментов. К примеру, инвариантом является формула Полячека – Хинчина для среднего времени ожидания заявок в системе $M/G/1$

$$w = \frac{\lambda b_2}{2(1 - \lambda b_1)},$$

где λ – интенсивность входящего потока; $\{b_i\}$ – начальные моменты распределения длительности обслуживания, а также формулы Литтла, связывающие средние длину очереди и число заявок в системе со средним временем ожидания и пребывания в системе, соответственно. С другой стороны, параметр ω , определяющий геометрическое распределение числа заявок перед прибытием очередной заявки в систему $GI/M/1$

$$\pi_k = (1 - \omega)\omega^k,$$

задается уравнением

$$\omega = \int_0^{\infty} e^{-\mu(1-\omega)t} dA(t), \quad (1)$$

где $A(t)$ – распределение интервалов между смежными заявками; μ – параметр экспоненциального распределения обслуживания. Следовательно, ω зависит от распределения $A(t)$ в целом, а не от конечного числа его моментов, и уравнение (1) инвариантом в указанном выше смысле не является.

Можно отметить некоторые современные работы, посвященные рассмотрению инвариантов отношения теории массового обслуживания. В [7] исследуются инвариантные характеристики хвостов стационарного распределения времени ожидания в системе массового обслуживания (СМО) $M/M/1/\infty$, $G/G/1/\infty$, определяемых субэкспоненциальными распределениями.

В [16] устанавливаются аналогии с большими отклонениями квазиобратимости класса процессов поступления заявок для систем $\bullet/M/1/\infty$. В работе [11] сравниваются свойства дуальных процессов Рамасвами и Брайта, показано, что последние обладают положительными свойствами, которые могут использоваться при разработке алгоритмов для анализа $GI/M/1$ и $M/G/1$.

Приведем еще несколько инвариантов из теории одноканальных приоритетных систем обслуживания, которые будут использованы в дальнейшем. В случае относительного (без права прерывания) приоритета среднее время ожидания начала обслуживания заявки j -го приоритета

$$w_j = \frac{\sum_{i=1}^k \lambda_i b_{i,2}}{2(1-R_{j-1})(1-R_j)}, \quad j = \overline{1, k}. \quad (2)$$

Здесь кумулянтный коэффициент загрузки $R_j = \sum_{i=1}^j \lambda_i b_{i,1}$.

Аналогичной формулой с ограничением суммирования в числителе верхним индексом j дается среднее время ожидания j -заявки в той же системе $\overline{M}_k / \overline{G}_k / 1$ с абсолютным приоритетом и дообслуживанием прерванной заявки. В этом случае ко времени ожидания начала обслуживания должна быть добавлена средняя длительность прерываний $b_{j,1} R_{j-1} / (1-R_{j-1})$, а для получения среднего времени пребывания – еще и средняя длительность обслуживания $b_{j,1}$. Итак, среднее время пребывания заявки в системе $\overline{M}_k / \overline{G}_k / 1$ с абсолютным приоритетом считается как

$$w_j = \frac{\sum_{i=1}^j \lambda_i b_{i,2}}{2(1-R_{j-1})(1-R_j)} + \frac{b_{j,1}}{1-R_{j-1}}, \quad j = \overline{1, k}. \quad (3)$$

Кратко охарактеризуем известные инварианты отношения многоканальных СМО.

ИНВАРИАНТЫ МНОГОКАНАЛЬНЫХ СИСТЕМ

К сожалению, перечень инвариантов для многоканальных систем обслуживания весьма ограничен. В частности, в [1] отмечено, что точная формула для среднего времени ожидания в системе $M/G/n$ не была получена в течение 40 лет, а с тех пор прошло еще 39. Еще острее (еще дальше от решения) этот вопрос стоит для многоканальных систем с приоритетами. В таких случаях для приближенного решения задачи можно применить инварианты отношения [1], основанные на

символических пропорциях для искомых средних показателей. В интересующем нас случае это будет

$$\frac{M/G/n}{M/G/1} \approx \frac{\overline{M}_k / \overline{G}_k / n}{\overline{M}_k / \overline{G}_k / 1}. \quad (4)$$

Заметим, что в [1] обсуждается более простой инвариант (с обязательным использованием показательных распределений обслуживания), поскольку к моменту выхода упомянутой книги ее авторы не располагали методами расчета немарковских многоканальных систем, исключая метод Кроммелена для обслуживания с постоянной длительностью.

Из (4) выводим базовое соотношение

$$\overline{M}_k / \overline{G}_k / n \approx \overline{M}_k / \overline{G}_k / 1 \cdot \frac{M/G/n}{M/G/1}.$$

Инварианты отношения не слабее прямых инвариантов, рассмотренных выше в том смысле, что если верны прямые инварианты, то верны и соответствующие инварианты отношения.

РЕАЛИЗАЦИЯ ИНВАРИАНТОВ ОТНОШЕНИЯ

Прежде всего отметим, что символическое равенство (4) правдоподобно, но строго не доказано, и точность построенного на его основе алгоритма должна подтверждаться при $n = 1$ согласованием с результатами счета по формулам типа (2) и (3), а для $n > 1$ – с результатами имитационного моделирования.

Ошибка в использовании инвариантов отношения будет тем меньше, чем ближе к интересующей нас модели $\overline{M}_k / \overline{G}_k / n$ будут условия обсчета моделей $M/G/1$ и $M/G/n$. Возможности численного расчета последней (с применением фазовых аппроксимаций) появились только в конце 1970-х годов и, разумеется, должны быть в полной мере использованы применительно к суммарной интенсивности потока заявок и к средневзвешенным моментам распределения длительности обслуживания. Для сопоставимости результатов среднее время пребывания в одноканальной системе считается для уменьшенной в n раз интенсивности потоков по формулам (2) и (3).

Для дисциплины с прерываниями наличие заявок с индексами $i > j$ (типы заявок упорядочены по убыванию приоритетов) никак не скажется на обслуживании заявок j -го типа, поэтому расчет упомянутых моделей делается при суммарном потоке интенсивности $\Lambda_j = \sum_{i=1}^j \lambda_i$ средневзвешенными моментами распределения обслуживания $\overline{b}_{j,m} = \Lambda_j^{-1} \sum_{i=1}^j \lambda_i b_{i,m}$, $m = 1, 2, \dots$

Практически расчет выполняется для модели $M/H_2/n$ с гиперэкспоненциальным 2-го порядка распределением обслуживания, параметры которого подбираются по трем средневзвешенным моментам. Отметим, что параметры аппроксимации в зависимости от коэффициента вариации полученного распределения могут оказаться вещественными, комплексными или парадоксальными (одна из вероятностей параллельных фаз отрицательна, а другая превышает единицу). Как показали численные эксперименты, упомянутые «псевдопатологии» не влияют на осмысленность конечных результатов.

После упомянутой аппроксимации расчет стационарных вероятностей состояний выполняется на основе итерационного метода Такахаси – Таками или методом матрично-геометрической прогрессии. Оба этих метода со ссылками на первоисточники и предложенными многочисленными модификациями расчетной схемы описаны в [2].

Через стационарные вероятности состояний можно вычислить среднее число заявок в системе (при его расчете можно ввести поправку на неограниченность очереди, дополнив реально вычисленные вероятности бесконечной убывающей геометрической прогрессией). Наконец, среднее время пребывания в системе определяется по формуле Литтла.

При обслуживании без прерывания расчетная схема оказывается несколько сложнее. Здесь для системы $M/G/n$ указанным выше способом по интенсивности входящего потока Λ_k и средневзвешенным моментам обслуживания $b_m = \Lambda_k^{-1} \sum_{i=1}^k \lambda_i b_{i,m}$ определяют стационарные вероятности и через них – среднюю длину очереди. Далее по формуле Литтла определяют общее для всех типов среднее время ожидания и добавлением к нему средней длительности обслуживания заявки каждого типа – среднее время их пребывания в системе.

Расчет среднего времени пребывания заявок каждого типа в одноканальной системе выполняется аналогично по идее, но более просто – с учетом того, что среднее время ожидания определяется суммарным потоком и средневзвешенными моментами обслуживания через формулу Полячка – Хинчина. Именно этими соображениями определялся отбор формул, включенных в разделы выше.

ИМИТАЦИОННЫЙ ЭТАЛОН

Для тестирования предложенных алгоритмов на современном Фортране разработаны программы соответствующих имитационных моделей. В модели системы с прерываниями на каждую заявку заводили паспорт, отражающий тип за-

явки, момент ее входа в систему, случайную трудоемкость, кратность прерываний, начало последнего прерывания данной заявки и суммарную длительность ее прерываний.

Этот набор сведений ориентирован на дальнейшее совершенствование аналитического метода расчета приоритетных систем. Содержание паспортов заносили в каналы (при наличии свободных) или в очереди – отдельно по типам заявок по возрастанию моментов прибытия. После каждого события каналы переупорядочивали, чтобы «кандидат на прерывание» всегда оказывался последним.

Логика модели без прерываний оказалась значительно проще и в дополнительных комментариях не нуждается.

ЧИСЛЕННЫЕ РЕЗУЛЬТАТЫ

При проведении численных экспериментов базовая интенсивность входящих простейших потоков трех типов принималась равной $\{0.222, 0.333, 0.445\}$, а средняя длительность обслуживания – $\{0.45, 0.90, 1.35\}$. Для получения высокого суммарного коэффициента загрузки 0,9, необходимого для иллюстрации вложенных прерываний, интенсивность потоков умножалась на этот коэффициент. Длительность обслуживания предполагалась подчиненной гамма-распределению с параметрами формы $\alpha = 3$ и $\alpha = 0,25$, что соответствовало коэффициентам вариации 0,577 и 2,0. Высшие моменты $\{b_i\}$ распределений обслуживания вычислялись через первые согласно

$$b_i = b_{i-1} \cdot b_1 \cdot [1 + (i-1) / \alpha], \quad i = 2, 3.$$

В процессе расчета по этим моментам подбирались H_2 -аппроксимация (для первого α ее параметры оказывались комплексными, а для второго – вещественными).

В табл. 1 и 2 приведены результаты расчета среднего времени пребывания заявок типов 1–3 в системе $M_k / G_k / n$ при относительном и абсолютном приоритетах в сопоставлении с результатами имитационных экспериментов. Модели прогоняли до обработки 300 тыс. заявок первого типа.

Относительный приоритет

Таблица 1

Тип	Метод	n = 1			n = 3			n = 5		
		1	2	3	1	2	3	1	2	3
E_3	Имит.	1,195	2,057	11,992	0,716	1,326	4,520	0,607	1,159	3,082
	Инвар.	1,192	2,059	11,931	0,680	1,260	4,634	0,579	1,102	3,195
H_2	Имит.	3,203	5,180	37,517	1,093	2,099	12,540	0,757	1,518	7,672
	Инвар.	3,233	5,246	41,028	1,276	2,189	13,121	0,888	1,584	7,599

Абсолютный приоритет

Таблица 2

Тип	Метод	n = 1			n = 3			n = 5		
		1	2	3	1	2	3	1	2	3
E_3	Имит.	0,488	1,313	12,743	0,450	0,948	4,865	0,450	0,911	3,296
	Инвар.	0,480	1,313	12,689	0,450	0,826	4,846	0,450	0,799	3,407
H_2	Имит.	0,560	2,186	35,197	0,449	1,022	11,652	0,450	0,924	7,163
	Инвар.	0,561	2,204	41,786	0,451	0,894	13,332	0,450	0,811	7,791

Из приведенных таблиц можно сделать следующие выводы.

1. Расхождение результатов «инвариантного» счета и имитационного моделирования обычно не превышает 15%, что свидетельствует о корректности как расчетных схем обоих сравниваемых подходов, так и реализующих их компьютерных программ.

2. Упомянутая корректность дополнительно подтверждается качественным соответствием результатов разумным ожиданиям – уменьшением по числу каналов среднего времени пребывания и его разброса по типам заявок, а также заметным увеличением этого разброса при переходе от относительных к абсолютным приоритетам. При увеличении числа каналов (и сохранении коэффициента загрузки) вследствие уменьшения ожидания среднее время пребывания в системе более приоритетных заявок стремится к средней длительности обслуживания (в случае абсолютных приоритетов – заметно быстрее).

3. Расхождение результатов допустимо в очень широком диапазоне коэффициентов вариации распределений длительности обслуживания (в рассмотренных примерах – от 0,577 до 2,0) и не обнаруживает тенденции к росту при увеличении числа каналов.

ЗАКЛЮЧЕНИЕ

Разработанный метод опирается на максимальную приближенность опорных вариантов к обчислываемой ситуации и дифференцированный учет влияния вида приоритетов. Так, для относительного приоритета по методу инвариантов вычислялось среднее время ожидания, к которому затем прибавлялось достоверно известное среднее время обслуживания. Поэтому результаты счета для относительного приоритета согласуются с имитационным моделированием заметно лучше, чем для абсолютного. Таким образом, инварианты отношения следует применять для оценки только той части результирующей величины, которая не может быть вычислена точно.

Основная доля трудоемкости метода приходится на расчет модели $M/H_2/n$, повторяемый для абсолютного приоритета по числу типов заявок, а для относительного приоритета – однократно. Выполненные автором модификации итерационного метода Такахаси – Таками и метода матрично-геометрической прогрессии [2] резко снизили эту трудоемкость.

Предложенная в настоящей статье схема использования инвариантов отношения, в отличие от [1], опирается не на приближенные оценки средней длины очереди, а на надежные численные алгоритмы. Она позволяет с достаточной для практических целей точностью рассчитывать среднее время пребывания заявок в многоканальной приоритетной системе с произвольными распределениями длительности обслуживания.

Схема избавляет от необходимости имитационного моделирования, трудоемкого как на этапе программирования, так и в процессе счета (обсуждаемая в статье имитационная модель использовалась только для верификации расчетных методик). Попутно отметим, что популярные системы моделирования типа GPSS для одноканальных систем не допускают вложенных прерываний, а прерывания многоканальных систем не позволяют моделировать вообще.

Дальнейшие исследования по разработке инвариантов отношения, на наш взгляд, целесообразно продолжить в следующих направлениях:

- для приоритетных систем массового обслуживания повышенной сложности, например, с вероятностным выталкивающим механизмом [21];
- многоканальных систем массового обслуживания повышенной сложности, например, с ограниченным временем пребывания заявок [3];
- нестационарных систем массового обслуживания с конечным источником заявок и распределениями фазового типа [8];
- разомкнутых и замкнутых немарковских моделей сетей массового обслуживания [4].

ЛИТЕРАТУРА

1. Бронштейн О.И. Модели приоритетного обслуживания в информационно-вычислительных системах / О. И. Бронштейн, И.М. Духовный. – М.: Наука, 1976. – 220 с.
2. Рыжиков Ю.И. Компьютерное моделирование систем с очередями: Курс лекций / Ю.И. Рыжиков. – СПб.: ВКА им. А.Ф. Можайского, 2007. – 164 с.
3. Рыжиков Ю.И. Расчёт гиперэкспоненциальной системы обслуживания $M/H_2/N-H_2$ с заявками, нетерпеливыми в очереди / Ю.И. Рыжиков, А.В. Уланов // Вестн. Томск. гос. ун-та. Управление, вычислительная техника и информатика. – 2014. – № 2 (27). – С. 47-53.
4. Рыжиков Ю.И. Расчет разомкнутых немарковских цепей с преобразованием потоков / Ю.И. Рыжиков, А.Д. Хомоненко // Автоматика и вычислительная техника. – 1989. – № 3. С. 15-24.
5. Рыжиков Ю.И. Средние времена ожидания и пребывания в многоканальных приоритетных системах / Ю.И. Рыжиков // Информационно-управляющие системы. – 2006. – № 6(25). – С. 43-49.
6. Хомоненко А.Д. Вероятностный анализ приоритетного обслуживания с прерываниями в многопроцессорных системах / А.Д. Хомоненко // Автоматика и вычислительная техника. – 1990. – № 2. – С. 55-61.
7. Цициашвили Г.Ш. Асимптотические инварианты в одноканальной системе массового обслуживания $G|G|1|_{\infty}$ / Г. Ш. Цициашвили, Н.В. Маркова // Дальневост. матем. журн. – 2002. – Т. 3, № 1. – С. 52-57.
8. Bubnov V.P. Model of Reliability of the Software with Coxian Distribution of Length of Intervals between the Moments of Detection of Errors / V.P. Bubnov, A.V. Tyrva, A.D. Khomonenko // Proc. of 34th Annual IEEE Computer Software and Appl. Conf. (COMPSAC 2010). 2010. – P. 238-243.
9. Davis R.H. Waiting-Time Distribution of Multiserver Priority Queueing System / R.H. Davis // Oper. Res. – 1966. – Vol. 14. – P. 133-136.
10. Gail H.R. Analysis of a Nonpreemptive Priority Multiserver Queue / H.R. Gail, S.L. Hantler, B.A. Taylor // Adv. Appl. Probab. – 1988. – Vol. 20, N 4. – P. 852-879.
11. Ganesh A. Invariant Rate Functions for Discrete-Time Queues / A. Ganesh, N. O'Connell, B. Prabhakar // Ann. Appl. Probab. – 2003. – Vol. 13, N 2. – P. 446-474.
12. Kao E. Analysis of Nonpreemptive Priority Queues with Multiple Servers and two Priority Classes / E. Kao, S. Wilson // Eur. J. Oper. Res. – 1999 – Vol. 118. – P. 181-193.

13. Kapadia A.S. Analysis of a Finite Capacity non Preemptive Priority Queue / A.S. Kapadia, M.P. Kazumi, A.C. Mitchell // *Comput. Oper. Res.* – 1984. – Vol. 11, N 3. – P. 337-343

14. Kella O. Waiting Times in Non-Preemptive Priority M/M/c queue / O. Kella, U. Yechialy // *Stochastic Models.* – 1985. – Vol. 1. – P. 257-262.

15. Khomonenko A.D. Performance Analysis of Multiprocessor Systems in the Priority Service of Heterogeneous Request Streams / A.D. Khomonenko // *Autom. Control and Comput. Sci.* – 1991. – Vol. 25, N 4. – P. 53-61 (Trans. from *Avtomatika i vychislitel'naya tekhnika.* – 1991. – N 4. – P. 55-64).

16. Korshunov D.A. Moments for Stationary Markov Chains with Asymptotically Zero Drift / D.A. Korshunov // *Siberian Math. J.* – 2011. – Vol. 52, N 4. – P. 655-664 (Trans. from *Sibirskiĭ Matematicheskiĭ Zhurnal.* – 2011. – Vol. 52, N 4. – P. 829-840).

17. Miller D.R. Steady-State Algorithmic Analysis of M/M/c Two Priority Queues with Heterogeneous Rates / D.R. Miller // *Appl. Probab. – Comput. Sci.: the interface.* – 1982. – Vol. 2. Boston: Birkhauser. P. 207-222.

18. Mitrani I. Multiprocessor systems with preemptive priorities / I. Mitrani, P.J.B. King // *Perform. Eval.* – 1981. – Vol. 1. – P. 118-125.

19. Taylor P.G. ON the Dual Relationship Between Markov Chains of GI/M/1 and M/G/1 Type / P.G. Taylor, Van Houdt B. // *Appl. Probab. Trust (24 Sept. 2009).* – P. 1-25.

20. Wiens P.D. On the Busy Period Distribution of the M/G/2 Queueing Systems / P.D. Wiens // *J. Appl. Probab.* – 1989. – Vol. 26, N 4. – P. 858-865.

21. Zaborovsky V. Preemptive Priority Queueing System with Finite Buffer Size and Randomized Push-Out Mechanism / V. Zaborovsky, V. Mulyukha, A. Ilyashenko, O. Zayats // *Mod. Traffic and Transp. Eng. Res.* – 2012. – Vol. 1, N 2. – P. 46-53.

Calculation of Multi-Channel Queueing Systems with Absolute and Relative Priorities on the Basis of Invariants Relationship

Ryzhikov Y. I.

St. Petersburg Institute for Informatics and Automation RAS (SPIIRAS)
Saint-Petersburg, Russia
ryzhbox@yandex.ru

Khomonenko A. D.

Petersburg State Transport University
Saint-Petersburg, Russia
khomon@mail.ru

Abstract. A method of calculating the average waiting time for various types of applications in multi-channel queueing systems with absolute and relative priorities, based on the application of the invariants of relations are proposed. The results are compared with data simulation of multi-channel priority systems with arbitrary distribution of the duration of service approximated by Erlang distribution and Hyperexponential.

Keywords: multichannel queueing systems, priorities, relationships invariants.

REFERENCES

1. Bronshtein O.I., Dukhovnyi I.M. *Modeli prioritetnogo obsluzhivaniia v informatsionno-vychislitel'nykh sistemakh*. [Models priority service in information and computing systems]. Moscow, Nauka, 1976. 220 p.
2. Ryzhikov Y.I. *Komp'iuternoe modelirovanie sistem s ocherediami: Kurs lektzii* [Computer simulation of queueing systems: Lectures]. SPb., VKA im. A. F. Mozhaiskogo, 2007. 164 p.
3. Ryzhikov Y.I., Ulanov A.V. Calculation Hyperexponential service system $M/H_2/N-H_2$ with applications, impatient queue [Raschet giperekspontsial'noi sistemy obsluzhivaniia $M/N_2/N-N_2$ s zaiavkami, neterpelivymi v ocheredi]. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naia tekhnika i informatika [Bull. Tomsk State Univ.]. Manage., Comp. Sci. and Inf.*, 2014, no. 2 (27), pp. 47-53.
4. Ryzhikov Y.I., Khomonenko A.D. Calculation of open Markov Chains with the Transformation Flows [Raschet razomknutykh nemarkovskikh tsepi s preobrazovaniem potokov]. *Avtomatika i vychislitel'naia tekhnika [Autom. Control and Comp. Sci.]*, 1989, no. 3, pp. 15-24.
5. Ryzhikov Y.I. Mean Waiting and Sojourn Times in Multi-channel Priority Systems [Srednie vremena ozhidaniia i prebyvaniia v mnogokanal'nykh prioritetnykh sistemakh]. *Informatsionno-upravliaiushchie sistemy [Inf. and Control Syst.]*, 2006, no. 6(25), pp. 43-49.
6. Khomonenko A.D. Probabilistic Analysis of Priority Service Interrupts in a Multiprocessor System [Veroiatnostnyi analiz prioritetnogo obsluzhivaniia s preryvaniiami v mnogoprotsektornykh sistemakh]. *Avtomatika i vychislitel'naia tekhnika [Autom. Control and Comput. Sci.]*, 1990, no. 2, pp. 55-61.
7. Ciciashvili G. Sh., Markova N.V. Asymptotic Invariants in Single-Channel Queueing System $G|G|1|\infty$ [Asimptoticheskie invarianty v odnokanal'noi sisteme massovogo obsluzhivaniia $G|G|1|\infty$]. *Dal'nevost. matem. zhurn. [Far East. Mat. Zh.]*, 2002, vol. 3, no. 1, pp. 52-57.
8. Bubnov V.P., Tyrva A.V., Khomonenko A.D. Model of Reliability of the Software with Coxian Distribution of Length of Intervals between the Moments of Detection of Errors. Proceedings of 34th Annual IEEE Comp. Software and Appl. Conf. (COMPSAC 2010), 2010, pp. 238-243.
9. Davis R.H. Waiting-Time Distribution of Multiserver Priority Queueing System. *Oper. Res.*, 1966, vol. 14, pp. 133-136.
10. Gail H.R., Hantler S.L., Taylor B.A. Analysis of a Non-preemptive Priority Multiserver Queue. *Adv. Appl. Probab.*, 1988, vol. 20, no. 4, pp. 852-879.
11. Ganesh A., O'Connell N., Prabhakar B. Invariant Rate Functions for Discrete-Time Queues. *Ann. Appl. Probab.*, 2003, vol. 13, no. 2, pp. 446-474.
12. Kao E., Wilson S. Analysis of Nonpreemptive Priority Queues with Multiple Servers and Two Priority Classes. *Eur. J. Oper. Res.*, 1999, vol. 118, pp. 181-193.
13. Kapadia A.S., Kazumi M.P., Mitchell A.C. Analysis of a Finite Capacity non Preemptive Priority Queue. *Comput. Oper. Res.*, 1984, vol. 11, no. 3, pp. 337-343.
14. Kella O., Yechialy U. Waiting Times in Non-Preemptive Priority $M/M/c$ Queue. *Stochastic Models*, 1985, vol. 1, pp. 257-262.
15. Khomonenko A.D. Performance Analysis of Multiprocessor Systems in the Priority Service of Heterogeneous Request Streams. *Autom. Control and Comput. Sci.*, 1991, vol. 25, no. 4, p. 53-61 (Trans. from *Avtomatika i vychislitel'naia tekhnika*, 1991, no. 4, pp. 55-64).
16. Korshunov D.A. Moments for Stationary Markov Chains with Asymptotically Zero Drift. *Siberian Math. J.*, 2011, vol. 52, no. 4, p. 655-664 (Trans. from *Sibirski'i Matematicheski'i Zhurnal*, 2011, vol. 52, no. 4, pp. 829-840).
17. Miller D.R. Steady-state algorithmic analysis of $M/M/c$ two priority queues with heterogeneous rates. *Appl. Probab. – Comput. Sci.: the Interface*, 1982, vol. 2. Boston: Birkhauser, pp. 207-222.
18. Mitrani I., King P.J.B. Multiprocessor Systems with Preemptive Priorities. *Perform. Eval.*, 1981, vol. 1, pp. 118-125.
19. Taylor P.G., Van Houdt B. ON the Dual Relationship Between Markov Chains of $GI/M/1$ and $M/G/1$ Type. *Appl. Probab. Trust* (24 Sept. 2009), pp. 1-25.
20. Wiens P.D. On the busy period distribution of the $M/G/2$ queueing systems. *J. Appl. Probab.*, 1989, vol. 26, no. 4, pp. 858-865.
21. Zaborovsky V., Mulyukha V., Ilyashenko A., Zayats O. Preemptive priority queueing system with finite buffer size and randomized push-out mechanism. *Mod. Traffic and Transp. Eng. Res.*, 2012, vol. 1, no. 2, pp. 46-53.

Статистический метод выделения временных паттернов из естественных языков

Пак А. А., Нарынов С. С.
ТОО «Alem Research»
Алматы, Казахстан
p.alexander@alem.kz

Касымова Д. Т., Утепбергенов И. Т.
Институт проблем информатики и управления
Алматы, Казахстан
i.utepbergenov@gmail.com

Аннотация. Рассматривается алгоритм рекурсивной самоорганизующейся карты (Recursive Self Organizing Map, RSOM) для выделения статистически значимых временных паттернов из символического потока контекстно зависимых грамматик. Главной идеей метода является предложенное Thomas Voegtlin в 2002 г. объединение классического алгоритма Teuvo Kohonen и неявного представления времени в виде комбинации текущего входа и контекстного отклика сети. Приведены результаты экспериментов над текстами на казахском и английском языках.

Ключевые слова: обработка естественных языков, нейронные сети, саморегуляция.

ВВЕДЕНИЕ

Индексация накопленного человечеством знания – задача нетривиальная. Размер текстов, накопленный в сети, составляет приблизительно 8000 терабайт. Среднегодовой объем электронной почты превышает 10000 терабайт. Очевидно, что в таком большом потоке данных выборка документов, полученная с помощью методов полнотекстового поиска, может содержать много нерелевантных или слабо релевантных результатов из-за синонимического представления запроса, опечаток, грамматических неточностей, сленга и др. [2].

Предлагаемый метод, позволяющий автоматически выделять темпоральные паттерны и конструкции естественного языка, основан на нейросетевом подходе самоорганизации и обучении без учителя. Сегодня разработано несколько способов в нейрокомпьютерике для представления времени, большинство из них связано с автоматическим распознаванием речи или с обработкой языка.

Условно можно разделить представление времени на явное, где время дополняет размерность пространства и коэффициент временной задержки используется для оценки мер процесса внутри окна наблюдения [14]; и неявное, когда время выражено опосредовано через логику представления изучаемого феномена [7], что характерно для рекуррентных сетей или так называемых «Leaky integrator»-процессов.

В общем случае явное представление времени сверхчувствительно и ограничено к временным деформациям, и напротив, неявное более устойчиво и универсально. Действительным ограничением квазиинтегративных нейронов является тот факт, что информация с долговременной зависимостью обычно затухает по экспоненциальному закону. Однако рекуррентное представление не обязательно должно

иметь эти ограничения [3], по крайней мере, до тех пор, пока не будет использован градиентный спуск.

В ряде научных работ предложены различные подходы интегрирования представления времени в самоорганизующуюся карту Кохонена (Self-Organizing Map, SOM). Явное представление рассмотрено в работах [11, 20], методы, основанные на латеральных и рекуррентных связях, предложены в [8, 9], техники, базирующиеся на квазиинтеграторе, – в [5, 18], также были изучены различные комбинации вышеозначенных принципов в работах [10, 12, 16]. Лейтмотивом перечисленных работ является идея обобщения процесса самоорганизации на время. Однако не совсем ясно, каким образом самоорганизация и время могут быть совмещены. Ошибка квантования или искажения в алгоритме SOM с учетом времени не может быть использована напрямую.

Предлагаемый метод основан на алгоритме рекуррентного модифицированного SOM, в котором обратная связь использована для представления времени. Представление времени неявно и самореферентно в том смысле, что карта учится классифицировать собственные выходы. При этом основная идея заключается в добавлении в классический алгоритм SOM рекуррентных связей при сохранении оригинального принципа самоорганизации. Чтобы оценить сходимость сети, предложено обобщить ошибки квантования временных рядов [21].

САМООРГАНИЗАЦИЯ И САМОРЕФЕРЕНТНОСТЬ

Широко известным примером рекуррентной нейронной сети является архитектура Simple Recurrent Network (SRN) предложенная Jeff Elman в 1990 г. SRN – это расширение классического Multi Layered Perceptron (MLP): добавлен еще один скрытый слой, который хранит отклики сети с заданным временным лагом в качестве дополнения к входному вектору. Алгоритм обратного распространения применяется к входному вектору и к его копии с лагом. Поскольку SRN обучается на своих прошлых активностях, его представление на скрытом слое является самореферентным. Самореферентные связи прямого и обратного распространения гомогенны. Нет никакой разницы между ними в уравнениях, описывающих активность сети, или в правилах обучения.

Покажем подход, основанный на данном принципе применительно к SOM с добавлением рекуррентных связей к оригинальной архитектуре SOM. Следует отметить, что уже

есть работы [8, 9, 12] в которых рекуррентные связи были интегрированы в карту Кохонена, но в этих работах прямая и обратная связь не гомогенны.

Архитектура, представленная в данной статье, использует активность сети с задержкой и гомогенность прямых обратных связей (см. рис. 1: оригинальный алгоритм карты рекурсивно применяется к входному вектору $x(t)$ и к отклику сети, полученному на предыдущем шаге $y(t-1)$. Пунктирные стрелки обозначают связи обучения. Сплошная стрелка означает связь «один к одному». Сеть учится ассоциировать текущий вход к предыдущему контексту [21]. Текущий входной вектор и предыдущий отклик сети рассматриваются как один входной вектор классического SOM-алгоритма. Иначе говоря, сеть учится кластеризовывать пару: входной вектор x и $y(t-1)$ контекст. Длинные последовательности сеть заучивает итеративно, используя более короткие последовательности в качестве описательных признаков.

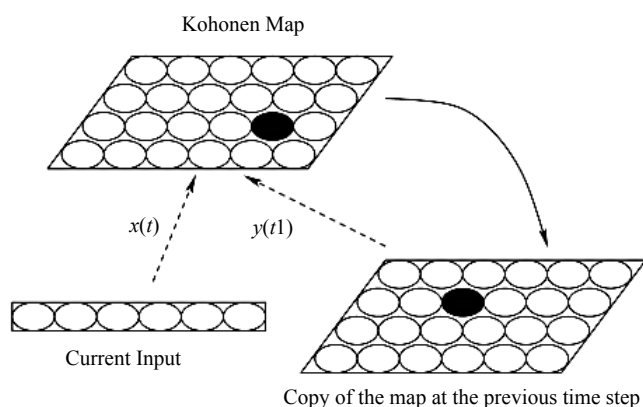


Рис. 1. Архитектура рекурсивного СОКа

В классическом алгоритме SOM каждый i -нейрон сравнивает свой вектор весов w_i , ($1 < i < N$) с входным вектором $x(t)$, где t – это переменная времени.

Обычно используется евклидова норма для определения победителя. Победитель определяется соревнованием среди N нейронов по наименьшей ошибке $E_i = \|w_i - x(t)\|$. По правилу обучения, вес нейрона-победителя должен быть модифицирован согласно

$$\Delta w_i = \gamma h_{ik} (x(t) - w_i), \quad (1)$$

где i – индекс-нейрона победителя; γ – функция скорости обучения; h_{ik} – функция соседства, традиционно используют так называемую «мексиканскую шляпу» или гауссиану (рис. 2) (в качестве функции используются соседства между нейронами SOM, см. [1, 15]). Существенным различием между ними является тот факт, что «мексиканская шляпа» позволяет больше специализировать отклик нейрона-победителя.

Мы в исследовании используем такую же нотацию. Дополнительно, пусть $y(t)$ – вектор активности карты в момент времени t . Каждый i -нейрон карты ($1 < i < N$) имеет пару весовых векторов w_i^x и w_i^y , которые сопоставляются с входным вектором $x(t)$ и вектором активности карты за предыдущую итерацию $y(t-1)$, соответственно. Чтобы определить нейрон-победитель, предлагается объединить

ошибки квантования, соответствующие $x(t)$ и $y(t-1)$. Поскольку прямые и обратные связи являются гомогенными, суммарная ошибка квантования выглядит следующим образом:

$$E_i = \alpha x(t) - w_i^{x^2} + \beta y(t-1) - w_i^{y^2}, \quad (2)$$

где $\alpha > 0$ – вес ошибки текущего входа; $\beta > 0$ – вес ошибки контекста.

Индекс нейрона-победителя определен по формуле $k = \arg \min \{E_i\}_{1 < i < N}$, тогда правила эволюции прямых и рекуррентных весов будут иметь вид

$$\Delta w_i^x = \gamma h_{ik} (x(t) - w_i^x), \quad (3)$$

$$\Delta w_i^y = \gamma h_{ik} (y(t-1) - w_i^y), \quad (4)$$

где γ – коэффициент скорости обучения; h_{ik} – функция соседства. Следует отметить, что это правила обучения оригинального SOM, применённые к $x(t)$ и $y(t-1)$.

Таким образом, в модели Recurrent SOM осуществляется обратная связь. Однако для того, чтобы завершить определение модели, необходимо привязать активность сети $y(t)$ ко входу и к синаптическим весам. Пусть компонента y_i вектора y – это активность i -нейрона, зависящая от квадрата ошибки квантования. Однако использование ошибки квантования напрямую $y_i = E_i$ может привести к нестабильности обучения. В таком случае предлагается использовать

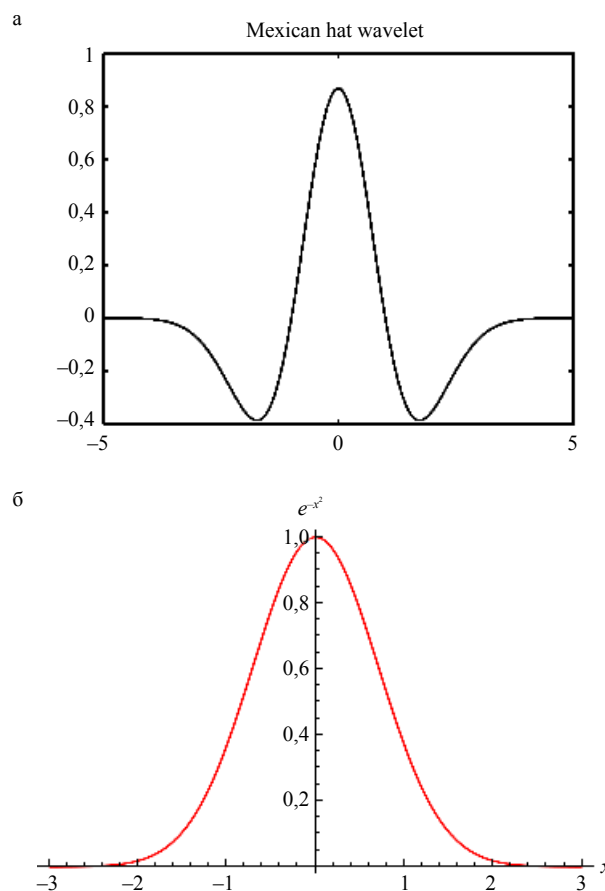


Рис. 2. Вейвлет «мексиканская шляпа» (а) и гауссиана (б)

функцию активации $y_i = F(E_i)$. Принцип выбора функции активации разьясним в следующем разделе.

Следует отметить, что параметры α, β могут быть исключены из формулы:

$$E_i^* = \|x(t) - w_i^x\|^2 \|y(t-1) - w_i^y\|^2 \text{ и } y_i = \sqrt{\frac{\beta}{\alpha}} F(\alpha E_i^*),$$

поэтому прямые и обратные связи являются гомогенными. Мы включили параметры α, β в E_i для удобства анализа сходимости и робастности алгоритма. В любом случае α, β могут быть рассмотрены как часть функции активации.

СТАБИЛЬНОСТЬ И ФУНКЦИЯ АКТИВАЦИИ

Идея использовать в традиционном SOM-алгоритме обратные связи уже была предложена в [4, 19]. Однако серьезной проблемой была нестабильность в обучении [19]. Иными словами, при изменении многомерного представления обучающих примеров на карте вектор обратных связей не успевал подстроиться, что препятствовало сходимости и стабильности обучения. Это подтверждено в работе [4]. Далее представлен аппроксимационный анализ стабильности рекуррентного SOM. Несмотря на приближительный характер анализа, будет дано представление о выборе трансферной функции.

Рассмотрим малое возмущение отклика сети $\delta y(t)$ при том, что все остальные переменные и параметры модели не изменяются. Возмущение $\delta y(t)$ будет иметь два эффекта. Во-первых, все последующие отклики $y(t+n), n > 0$ будут искажены рекуррентным распространением возмущения. Поставим в соответствие «будущим» искажениям вектор $\delta y(t+n)$.

Во-вторых, возмущение приведет к изменению распределения вектора $y(t)$, что в свою очередь индуцирует долгосрочную адаптацию синаптических весов. В первом приближении будем рассматривать только кратковременную эволюцию. При условии достаточно малой величины коэффициента обучения можно пренебречь адаптацией весов.

Стабильность на кратковременной динамике зависит от эволюции $\delta y(t+n)$. Если $\delta y(t+n) \rightarrow 0$ при $n \rightarrow \infty$, то возмущение затухнет

$$\sum_{n>0} \delta y(t+n) < +\infty.$$

Более того, если распределение y не изменится, вес не адаптируется. Также покажем, что эффект возмущения мал, если мало само возмущение. Можно выразить это утверждение, используя нотацию Ландау:

$$\sum_{n>0} \delta y(t+n) = O(\delta y(t)). \quad (5)$$

Рассмотрим воздействие адаптации весов на стабильность обучения. Строго говоря, стабильность весов должна быть рассмотрена только тогда, когда коэффициент обучения стремится к нулю. Однако мы рассматриваем стабильность в более широком смысле. Положим, что коэффициент обучения – малая константа и ожидаемые значения синаптических весов и их вариации можно сделать достаточно малыми правильным выбором коэффициента обучения. Для упрощения представления положим, что веса сходятся, когда

используется традиционный алгоритм SOM и распределение входного вектора является константой [12].

Положим, $W^y = (w_i^y), 1 < i < N$ – матрица весов рекуррентных связей. Рассмотрим малые изменения δW^y во время обучения. Как было сказано, мы рассматриваем матрицу W^y , элементы которой являются ожиданием рекуррентных весов во время обучения, δW^y – возмущение этих весов. δW^y индуцирует прямое изменение отклика $y(t)$ из-за того, что в уравнении (2) изменится w_i^y . Кроме того, к непрямоу изменению $y(t)$ может привести распространение возмущения отклика сети, как показано выше.

Возмущение отклика $\delta y(t)$ – результат суперпозиции двух компонент: прямой и косвенной. Прямая компонента может быть получена из уравнения (2) и будет иметь тот же порядок энергии, что и δW^y при непрерывной трансферной функции. Косвенная компонента – сумма величин, возникших в результате распространения возмущения на шагах $t-1, t-2$ и т. д. Однако если условие (5) удовлетворено, то энергия косвенной компоненты будет того же порядка, что и прямой компоненты, поэтому возмущения отклика сети $\delta y(t)$ и δW^y синаптических весов будут одного порядка:

$$\delta y(t) = O(\delta W^y). \quad (6)$$

В обучении подобной модели возникает так называемая проблема «движущейся цели», где цель – это набор синапсов, стабильных в период обучения. Под движущейся целью подразумевается изменение облака данных, на котором обучается RSOM из-за изменения рекуррентных связей. Иными словами, сеть учится классифицировать собственный отклик, который меняется во время обучения.

Пусть $T = (t_i), 1 < i < N$ соответствует матрице «движущихся целей», где t_i – «цель», к которой должен сойтись текущий вектор w_i^y по совместному стационарному распределению $(x(t), y(t-1))$. Следовательно, вектор «цель» в любой определенный момент времени является результатом суперпозиции векторов синапсов и статистического ожидания примеров обучения. Предполагаем, что для совместного стационарного распределения $(x(t), y(t-1))$ ряды $x(t)$ эргодичны. Однако любое изменение синаптических весов также изменит распределение $(x(t), y(t-1))$, и вектор t_i изменится. Чтобы гарантировать успешание модели за вектором цели, мы должны наложить дополнительное условие

$$\delta T < \delta W^y. \quad (7)$$

Из-за локальных латеральных связей на карте каждый вектор цели t_i является барицентром ожидания $y(t-1)$, когда нейрон i становится победителем в момент времени t . Хотя отношения соседства усложняют анализ возмущений в векторе целей, разумно предположить следующее неравенство:

$$\delta T \leq \sqrt{N} E[\delta y], \quad (8)$$

где δT – вариация матрицы целей из-за возмущения распределения y ; $E[\cdot]$ – статистическое ожидание.

Из неравенства (8) с учетом (7) получаем:

$$\sqrt{N} E[\delta y] \leq \delta W^y. \quad (9)$$

Таким образом, мы получили два условия стабильности. Условие (5) гарантирует стабильность нейронной динамики, а условие (9) – стабильность весов в период обучения.

Если оба условия удовлетворяются, то ожидаемые значения весов сходятся во время обучения. Выбор функции активации представлен следующим выражением:

$$y_i(t) = F(E_i) = \exp(-\alpha x(t) - w_i x^2 - \beta y(t-1) - w_i y^2). \quad (10)$$

Функция является непрерывной и ограниченной на отрезке (0,1]. Отклик нейрона-победителя находится ближе к 1, отклик проигравшего – ближе к 0. Кажется, что активность нейрона-победителя должна быть стабильной в случае возмущения перцептивных или синаптических полей. Гауссовская функция активации сглаживает выбросы. Таким же свойством обладает и отклик проигравших нейронов. В [21] показано, что (10) удовлетворяет условиям стабильности (5) и (9).

На рис. 3 (по [14]) представлены численные результаты экспериментов по определению области стабильности для рекурсивной карты Кохонена с функцией активации (10). Показаны области стабильности рекуррентного SOM, обученного на бинарных последовательностях для различных параметров α и β . При этом приведена средняя длина рецептивного поля нейрона-победителя. Для каждого условия данные усреднялись по восьми экспериментам. Низкие значения указывают на плохую стабильность, большие – на хорошую. В верхней правой области наблюдается стабильность при плохой сходимости.

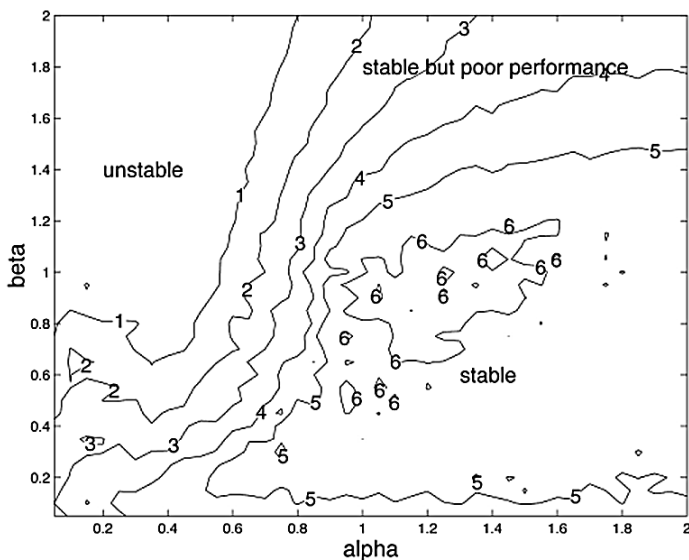


Рис. 3. Области стабильности рекуррентного SOM

ЭКСПЕРИМЕНТ И ПАРАМЕТРЫ ОБУЧЕНИЯ

Для тестирования и отладки проведены эксперименты, использовалась классическая функция соседства

$$h_{ik} = \exp\left(-\frac{d(i,k)^2}{\sigma^2}\right).$$

Размер карты составлял 28×28 для английского языка, 36×36 – для казахского. Тексты, на которых обучалась сеть: «Brave New World» Aldous Huxley, сборник исторических эссе о Казахстане, находящийся в сети в открытом доступе. Буквы были закодированы бинарным вектором, для казахского языка длина вектора составляла 6, для английского – 5 бит. В предобработке пунктуация убрана из текста, исключение составляла тильда в английском языке в таких словах как don't, haven't и т.д. Сам текст разбит на слова, разделителем служил пробел. Между словами контекст сети обнулялся.

Коэффициент обучения так же, как и в оригинальной статье, был константой $\gamma = 1$. Функция соседства представляла собой гауссиану с шириной $\sigma = 1,8$ и масштабом $s = 5$. Параметры $\alpha = 3, \beta = 0,7$. По ходу обучения нейроны втягивали в свое перцептивное поле образы различных последовательностей букв. Объем текстов для казахского и английского языков составил приблизительно 400 Кб. Длительность обучения для английского языка – около 8 ч, для казахского – около 15 ч, что объясняется большой размерностью карты и кодирования букв. Реализация алгоритма была сделана на языке JavaSE2. Использовались математические библиотеки линейной алгебры и аналитической геометрии Cern Colt 1.2.0 [6]. Для визуализации экспериментальных данных использовалась программа Pajek643.11 [17].

РЕЗУЛЬТАТЫ

На рис. 4 представлены временные паттерны, заученные нейронной сетью на текстах Aldous Huxley. Сеть была настроена для выявления характерных дифтонгов в тексте. Однако в силу рекуррентной природы последняя буква первого дифтонга является также первой буквой второго дифтонга. На рисунке видны статистически значимые комбинации дифтонгов, образующие слова английского языка:

- 1) was
- 2) with
- 3) whip
- 4) and
- 5) had
- 6) his,

а также характерные окончания:

- 1) -ing
- 2) -ent
- 3) -n't

Данные, представленные на рис. 4, подвергались постобработке. Прежде всего, были удалены дубликаты с общим корнем и одинаковым образом, что снизило количество узлов с 784 до 348. Также связи между узлами были прорежены по уровню 0.3.

На рис. 5 представлены временные паттерны казахского языка, характерные для исторических эссе, находящиеся в свободном доступе. Видно, что сеть выявила характерные сочетания, образующие целые слова:

- 1) философ
- 2) философия
- 3) бар
- 4) іні
- 5) қар
- 6) көр
- 7) бол,

а также характерные части слов:

- 1) дык
- 2) дер
- 3) ған
- 4) тар
- 5) сыр

Визуализация для казахского языка также была подвергнута фильтрации, редуцированию количества узлов с 1296 до 300 и прореживанию связей с уровнем 0.4.

ЗАКЛЮЧЕНИЕ

Сейчас наблюдается большой ажиотаж вокруг обработки естественных языков. Это связано, в первую очередь, с глубокой востребованностью автоматических методов обработки больших объемов текстовых данных. Естественно, подобное возможно лишь при помощи методов машинного обучения. Здесь наблюдается тенденция к более реалистичным когнитивным моделям, что связано с недавно появившимися эффективными методами «глубокого обучения», подразумевающими выявление в потоках данных иерархии более сложно устроенных признаков. Алгоритм рекуррентной карты является хорошим кандидатом для автоматического выявления сложных иерархий в потоке символов. Для его развития требуется решить проблемы мертвых нейронов, дубликатов и скорости обучения.

ЛИТЕРАТУРА

1. Калимолдаев. – URL : <http://www5a.wolframalpha.com/Calculate/MSP/MSP6701h6a8290ga35bb860003afghd1hgga79088?MSPStoreType=image/gif&s=61&w=304.&h=124.&cdf=RangeControl>.
2. Шумский С. А. Самоорганизующиеся семантические сети С. А. Шумский // *Нейроинформатика-2001 III Всерос. науч. конф. «Лекции по нейроинформатике»*. Москва, 2001, С. 182-212.
3. Bengio Y. Learning long-term dependencies with gradient descent is difficult / Y. Bengio, P. Simard, P. Frasconi // *IEEE Trans. Neural Net.* – 1994. – N 5(2). – P. 157-166.
4. Briscoe G. Learning temporal sequences in recurrent self-organising neural nets / G. Briscoe, T. Caelli // *Adv. topics in artificial intel.* : Proc. 10th Australian Joint Conf. on Artificial Intell., AI'97 / ed. A. Sattar. – Berlin : Springer, 1997. – P. 427-435.
5. Chappell G.J. The temporal Kohonen map / G.J. Chappell, J.G. Taylor // *Neural Net.* – 1993. – N 6. – P. 441-445.

6. Colt. – <http://acs.lbl.gov/software/colt>.
7. Elman J.L. Finding structure in time / J.L. Elman // *Cogn. Sci.* – 1990. – N 14. – P. 179-211.
8. Euliano N. Spatio-temporal self-organizing feature maps / N. Euliano, J. Principe // *Proc. Int. Conf. on Neural Net.* – 1996. – P. 1900-1905.
9. Hoekstra A. An extended Kohonen feature map for sentence recognition / A. Hoekstra, M. Drossaers // *Proc. Int. Conf. on Artificial Neural Net.* – 1993. – P. 404-407.
10. James D.L. SARDNET: A self-organizing feature map for sequences / D.L. James, R. Miikkulainen // *Adv. in Neural Inf. Proc. Syst.* – 1995. – N 7. – P. 577-584.
11. Kangas J. On the analysis of pattern sequences by self-organizing maps / J. Kangas. PhD Thesis. – Helsinki Univ. of Technology, 1994.
12. Kohonen T. *Self-Organizing Maps.* – Springer Verlag, 2001.
13. Koskela T. Time series prediction using recurrent SOM with local linear models / T. Koskela, M. Varsta, J. Heikkonen, K. Kaski // *Int. J. Knowledge-Based Intell. Eng. Syst.* – 1998. – N 2 (1). – P. 60-68.
14. Lang K.J. A time-delay neural network architecture for isolated word recognition / K.J. Lang, A.H. Waibel, G.E. Hinton // *Neural Net.* – 1990. – N 3. – P. 23-43.
15. Mexihat. – URL : <http://radio.feld.cvut.cz/matlab/toolbox/wavelet/mexihat.html>.
16. Mozayyani N., Alanou V., Dreyfus J., Vaucher G. A spatiotemporal data coding applied to Kohonen maps. *Proceedings of the International Conference on Artificial Neural Networks.* – 1995. P. 75–79.
17. Pajek. – URL : <http://pajek.imfm.si/doku.php>.
18. Privitera C.M. The analysis of continuous temporal sequences by a map of sequential leaky integrators / C.M. Privitera, P. Morasso // *Proc. Int. Conf. on Neural Net.* – 1994. – P. 3127-3130.
19. Scholtes J.C. Kohonen feature maps in natural language processing. Tech. Rep. CL-1991-01, Institute for Language, Logic and Information, Univ. of Amsterdam. Varsta M., Heikkonen J., Millan J.D.R. (1997). Context-learning with the self-organizing map // *Proc. Workshop on Self-Organizing Maps'97 Espoo.* – Finland: Helsinki Univ. of Technology. – 1991. – P. 197-202.
20. Vesanto J. Using the SOM and local models in time-series prediction / J. Vesanto // *Proc. Workshop on Self-Organizing Maps'97 Espoo.* – Finland: Helsinki Univ. of Technology. – 1997. – P. 209-214.
21. Voegtlin T. Recursive self-organizing maps / T. Voegtlin // *Neural Net.* – 2002. – Vol. 15, N 8-9. – P. 979-992.

Statistical Method for Isolating Temporal Patterns of Natural Languages

Pak A. A., Narynov S. S.
LLP „Alem Research“
Almaty, Kazakhstan
p.alexander@alem.kz

Kasymova D. T., Utepbergenov I. T.
Institute of information and computing technologies
Almaty, Kazakhstan
i.utepbergenov@gmail.com

Abstract. An algorithm for the recursive self-organizing map (Recursive Self Organizing Map, SOM) to highlight the statistically significant temporal patterns of the symbol stream of context-sensitive grammars. The main idea of this method is proposed Thomas Voegtlin in 2002, combining classical algorithm Teuvo Kohonen, and implicit representation of time as a combination of the current input and contextual response network. The results of experiments on the texts of the Kazakh and English languages.

Keywords: natural language processing, neural networks, self-organization.

REFERENCES

1. <http://www5a.wolframalpha.com/Calculate/MSP/MSP6701h6a8290ga35bb8600003afghd1hgga79088?MSPStoreType=image/gif&s=61&w=304.&h=124.&cdf=RangeControl>.
2. Shumsky S. Selforganizing Semantic Networks [Samoorganizuiushchiesia semanticheskije seti]. *Neuroinformatika-2001 III Vserpssiiskaia nauchnaia konferentiia "Lectsii po neuroinformatike" [Neuroinformatics-2001 III All-Rus. Sci. Conf. "Lectures on Neuroinformatics]*. Moscow, 2001. Pp. 182-212.
3. Bengio Y., Simard P., Frasconi P. Learning long-term dependencies with gradient descent is difficult. *IEEE Trans. Neural Net.*, 1994, no. 5 (2), pp. 157-166.
4. Briscoe G., Caelli T. Learning Temporal Sequences in Recurrent Self-Organising Neural Nets. *Adv. Topics in Artificial Intell. Proc. 10th Australian Joint Conf. on Artificial Intell., AI'97*; ed. A. Sattar. Berlin, Springer, 1997, pp. 427-435.
5. Chappell G. J., Taylor J. G. The Temporal Kohonen Map. *Neural Net.*, 1993, no. 6, pp. 441-445.
6. Colt (available at: <http://acs.lbl.gov/software/colt>).
7. Elman J. L. Finding Structure in Time. *Cogn. Sci.*, 1990, no. 14, pp. 179-211.
8. Euliano N., Principe J. Spatio-Temporal Self-Organizing Feature Maps. *Proc. Int. Conf. on Neural Net.*, 1996, pp. 1900-1905.
9. Hoekstra A., Drossaers M. An Extended Kohonen Feature Map for Sentence Recognition. *Proc. Int. Conf. Artificial Neural Net.*, 1993, pp. 404-407.
10. James D. L., Miikkulainen R. SARDNET: A Self-Organizing Feature Map for Sequences. *Adv. Neural Inf. Proc. Syst.*, 1995, no. 7, pp. 577-584.
11. Kangas J. On the Analysis of Pattern Sequences by Self-Organizing Maps. PhD Thesis. Helsinki Univ. of Technology, 1994.
12. Kohonen T. Self-Organizing Maps. Springer Verlag, 2001.
13. Koskela T., Varsta M., Heikkonen J., Kaski K. Time Series Prediction Using Recurrent SOM with Local Linear Models. *Int. J. Knowledge-Based Intell. Eng. Syst.*, 1998, no. 2 (1), pp. 60-68.
14. Lang K. J., Waibel A. H., Hinton G. E. A Time-Delay Neural Network Architecture for Isolated Word Recognition. *Neural Net.*, 1990, no. 3, pp. 23-43.
15. Mexihat (available at: <http://radio.feld.cvut.cz/matlab/toolbox/wavelet/mexihat.html>).
16. Mozayyani N., Alanou V., Dreyfus J., Vaucher G. A Spatiotemporal Data Coding Applied to Kohonen Maps. *Proc. Int. Conf. Artificial Neural Net.*, 1995, pp. 75-79.
17. Pajek (available at: <http://pajek.imfm.si/doku.php>).
18. Privitera C. M., Morasso P. The Analysis of Continuous Temporal Sequences by a Map of Sequential Leaky Integrators. *Proc. Int. Conf. Neural Net.*, 1994, pp. 3127-3130.
19. Scholtes J. C. Kohonen Feature Maps in Natural Language Processing. Tech. Rep. CL-1991-01, Inst. for Language, Logic and Information, Univ. of Amsterdam. Varsta M., Heikkonen J., Millan J. D. R. (1997). Context-Learning with the Self-Organizing Map. *Proc. Workshop on Self-Organizing Maps '97 Espoo*. Finland: Helsinki Univ. of Technology, 1991, pp. 197-202.
20. Vesanto J. Using the SOM and Local Models in Time-Series Prediction. *Proc. Workshop on Self-Organizing Maps '97 Espoo*. Finland: Helsinki Univ. of Technology, 1997, pp. 209-214.
21. Voegtlin T. Recursive Self-Organizing Maps. *Neural Net.*, 2002, Vol. 15, no. 8-9, pp. 979-992.

Multifractal Analysis: Identifying the Boundaries Application in the Study of Time Series

Zakharov A. I., Zagaynov A. I.

Military Space Academy named Mozhaiskiy
Saint-Petersburg, Russia
ana63916157@yandex.ru

Khodakovsky V. A.

Petersburg State Transport University
Saint-Petersburg, Russian
hva1104@mail.ru

Abstract. The work is devoted to revelation of possibilities from multifractal numerical procedures and similar handling techniques of time series developed until now. Application Frac-Lab 2.1 from the mathematic packet MatLab R2008b was used for this purpose, it processes test data (generated, for example, by the Henon map, Lorenz system of equation sand etc.), bases of heart rate variability records from the PhysioNet server and own electrocardiograms (ECG). Three-dimensional graphics of received wavelet transformation, graphs of isolines, multifractal characteristics in the form of Hölder exponent and scaling exponent were constructed. All possible data provided in the FracLab (for example, Mhat-wavelet, DoG-wavelet and etc.) were used as a wavelet-forming function in the course of calculating a spectrum. Disadvantages of used realization, received first of all for test data calculating the Hölder exponent, were determined. The conception of multifractal methodology widening was proposed for research of time series and creation of own software on its basis.

Keywords: time series, method of wavelet transform modulus maxima.

INTRODUCTION

The problem of searching for connection indicators of regulating system – is a perspective direction in the modern fundamental science, whose basic tendency is focused now on existence criterion determination (and detection) of determined chaos in the accentuated give time series of corresponding dynamical system.

As an example of such time series we mention recommendations for usage of physiological interpretation of consecutive intervals between QRS-complexes of electrocardiogram (heart rate variability), executed in 1996 by the European Cardiology Society, North American Electrophysiological Community [6] and Committee of Clinical Diagnostic Devices and Instruments.

Committee of New Medical Technology of the Russian Federation Ministry of Public Health [3] make these time lengths the most perspective in the course of noninvasive diagnostics in the modern practice of scientific investigations. In spite of the recommended (and accepted) range of variability indexes (for example, indexes of variational pulsometry (Mo, AMo, MxDMn, MxRMn, triangular index) [3, 6], indexes of correlation rhythmography (L, w, EllAs, EllSq) [3], indexes of power in various frequency intervals, low and high frequency spectral components, their relations (LF, HF, LF/HF, VLF, ULF, IC) [3, 6] and etc.) the results, received using them have specific disadvantages.

The problem here is in the statement (postulate) about stationarity of impacts of adaptive processes on the cardiac rhythm

regulating systems, such as impact of central and vegetative nervous system, pathologies of circulatory system (resulting in rhythm disturbance), thyroid gland and etc. Certainly, such solution makes a deep impression on possibilities of received results of investigations. At the same time, (including because of the above mentioned reason) the majority of calculated components use traditional (linear) methods of time series analysis (for example, statistic, spectral, correlation and etc.). It, in its turn, simplifies the form evaluation of conformance to considered processes, distorting more their real physiological interpretation.

For example, analyzing the same spectral concentration in the course of routine examination of variability in the various frequency bands, the values of power don't have four evident peaks (to which the frequency components are related). The power coefficients are calculated only thanks to connection with the specific preset interval, and not due to highlighting the specified peculiarity.

This situation is explained by nonlinearity, discontinuity, instability and etc. of the cardiac rhythm reference signal, where it is impossible to exclude both contained phenomena of its internal regulation and procedural error of registering abbreviations itself. To solve the existing difficulty, several researchers (together with our research team) propose usage of non-linear mathematical methods, the majority of which is based on the fractal analysis.

The fractal methods are based, first of all, on examination of scaling invariance (scaling) of the researched process conditions. Conventionally, they may be divided into methods which directly use the idea of fractal (as a geometric entity in the multidimensional phase space) and transferred to it by manipulating initial time series for setting coordinates and methods examining the scale invariance of the initial process direct features. The well known feature of scaling – fractal (correlation) dimensionality may be related to the first ones. But here, there are a number of difficulties connected first of all with convergency of such feature in the finite-dimensional space of enclosure.

Absence of convergence curve saturation of correlation dimensionality are demonstrated in several works appeared in recent years [1, 5, 7]. At the same time, the determined deterministic rhythms – for example, fetal rhythm for periods of gestation 38–40 weeks and ventricular fibrillation, it suggests that the direction of fundamental research vector is correct.

Development of the last one may take place both considering various improvements of the mentioned methodology and principally different conception, using only foundations of the previous one. The methods of forming the scaling of the exam-

ined process direct features on the basis of multifractal formalism conception, replacing the fractal exponent (restored attractor) by the directly built specialized spectrum, analyzed in the article, may be related to the last one.

MATERIALS AND METHODS

The scaling or scaling invariance is a main feature which is searched in the time-series in he analyzed methods of researches. At the same time, there are alternative approaches for its plotting, except the analyzed approach to the n-dimensional mapping. For example, the DFA method (detrended fluctuation analysis) is based on forming the automodeling process using the following sums:

$$b_k = \sum_{i=1}^k (a_i - \bar{a}) .$$

Afterwards, the formed row is divided in the areas with similar length n, where it is approximated (in the simplest variant) by the simple dependence using the method of the smallest squares. Setting the mean square error, the scaling is built by its comparison with the exponential function of the length value n:

$$F(n) = \sqrt{\frac{1}{N} \sum_{k=1}^N [b_k - \tilde{b}_k(n)]^2} ,$$

where $\tilde{b}_k(n)$ – is approximation b_k at the length n areas.

This approach is not perspective due to the beam analysis of scaling. The more advanced method WTMM (wavelet transform modulus maxima) uses plotting of the whole variety of the local maximum line of wavelet transform:

$$W_\psi(t, a) = \frac{1}{a} \int_{-\infty}^{\infty} f(x) \psi\left(\frac{x-t}{a}\right) dx,$$

where the initial signal $f(x)$ is divided using the function $\psi(x)$ generated from the soliton-like one with special features by its scale measurements and shifts [2]. In the simplest variant (Hölder’s exponent h) the scaling of one of the lines (for example, the maximum one) is researched:

$$|W_\psi(t_i, s)| \sim s^{h(t_i)} .$$

The more complicated approach to plotting the scaling is based on analyzing all lines by introducing the partial function with the weight degree of all wavelet transform maximums:

$$P_q(s) = \sum_i [W_\psi(t_i, s)]^q ,$$

and plotting the scaling using the scaling function $k(q)$:

$$P_q(s) \sim s^{k(q)} .$$

Above approach has a theoretically justified (and perfect) build capacity scaling, ranging from construction of the expansion of the original time series, setting the scale invariance of a certain line and ending with the introduction of the special function lines for finding local maxima of the generalized scaling. This is caused by frequent failure resulting in the construction of scaling a single line.

RESULTS AND DISCUSSION

To demonstrate usage of multifractal algorithms and determine their description possibilities we used standard time series in the theory of determined chaos – generated by nonlinear mapping (including solutions of nonlinear differential equations). Let’s remind, that the main advantage of such data involves a possibility to describe them using (mono)fractal methods, whose confirmation is convergence of fractal (correlation) dimensionality.

The application FracLab 2.1 of mathematic packet MatLab R2008b [4] was used for these purposes (at the first stage of researches. Results of analyzed methodology usage for Lorenz’s system of equations:

$$\begin{cases} \frac{dx}{dt} = -\sigma x + \sigma y, \\ \frac{dy}{dt} = xz + rx - y, \\ \frac{dz}{dt} = xy - bz. \end{cases}$$

with the values of parameters $\sigma = 10, b = 8/3, r = 28$ are given in fig. 1, where the graphs for Hölder exponent and scaling exponent using Mhat-wavelet are presented.

We would like to note that the solutions of mentioned system are received as a result of numerical integration using the methods Runge-Kutta of the 4th order. Even here we demonstrate unsuitability of the first (simplified) approach in the course of researching the scaling. It may be described by insufficiency of numerical structure of one line of local maximum (the result in the form of nondeterministic influence is possible in case of the described base transition), and inadequacy of the existing numeric base of algorithms. At that, the approach with introducing partial (generalized) function provides valid results – its approach to zero in the negative area of scaling and conation to straight line in the positive one.

Researching the time series (including for the considered system) we used various kinds of wavelet –forming function (Mhat, DoG, Wave, Morlet, Haar and etc.) The given results as well as usage of DoG and Wave wavelet are the most successful for the mentioned mapping.

Usage of discussed methodology for the two-dimensional mapping of Henon:

$$\begin{cases} x_{n+1} = 1 - ax_n^2 + y_n, \\ y_{n+1} = bx_n, \end{cases}$$

described with the parameters $a=1.4, b=0.3$ provide less demonstrative results (fig. 2). Similarly to attractor of Lorenz, the Hölder exponent is useless for practical interpretation of results. The scaling exponent behaves a little bit differently (including depending on the type of wavelet-forming basis). Here, there is no obvious conation to the direct correlation in the positive area of scaling. This (together with performed remarks) may be explained by direct forming of the researched attractor.

To this effect it is enough to examine its formation, for example, in the real time mode (realized in one of our projects [1]) and tracking the trajectory to make sure in the disruptive method of forming the next coordinate. In addition to attractors analyzed using multifractal methods we analyzed other attractors received using generators of determined chaos (Rossler equation system, Ikeda mapping and etc.) one.

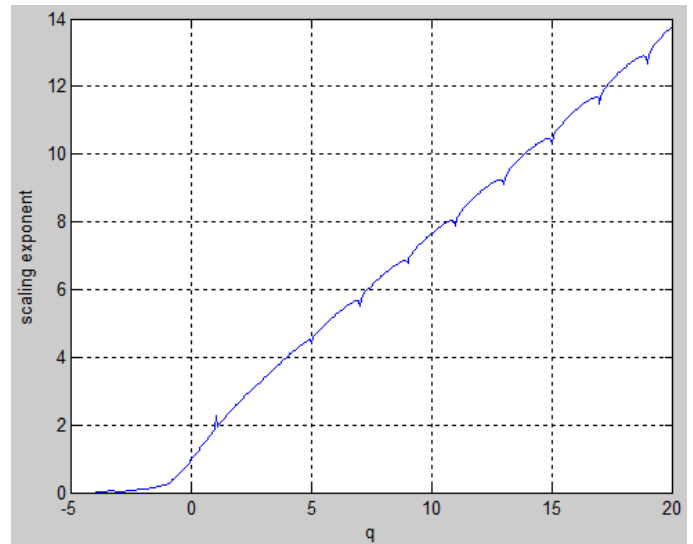
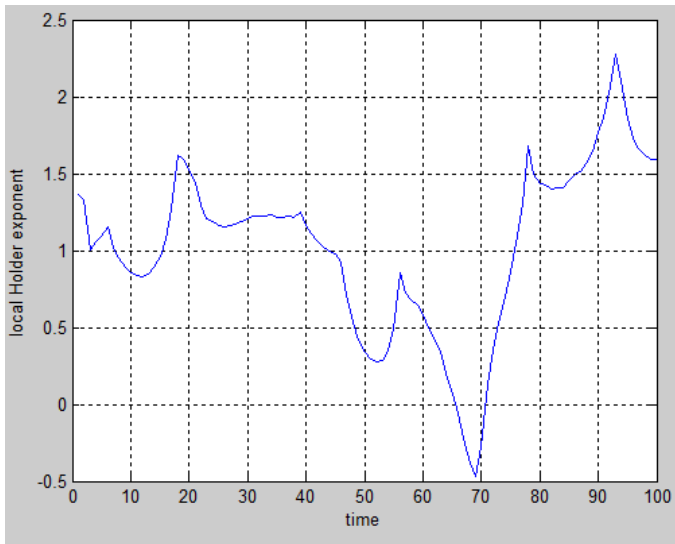


Fig. 1. Hölder exponent and scaling exponent developed for the Lorenz's system of equations

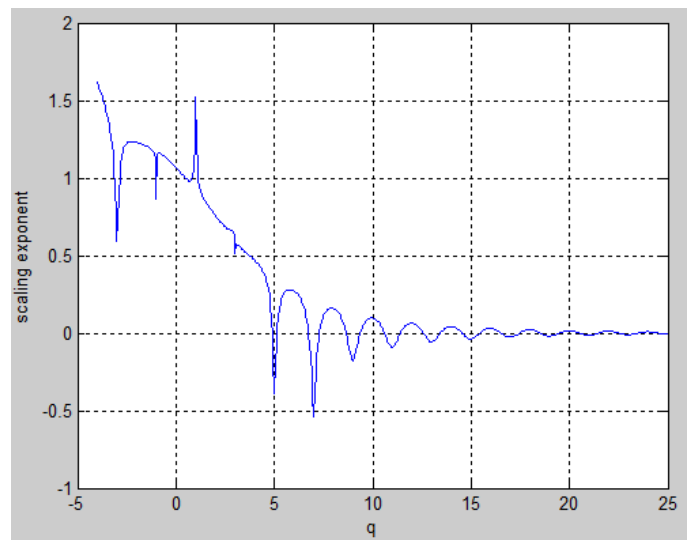
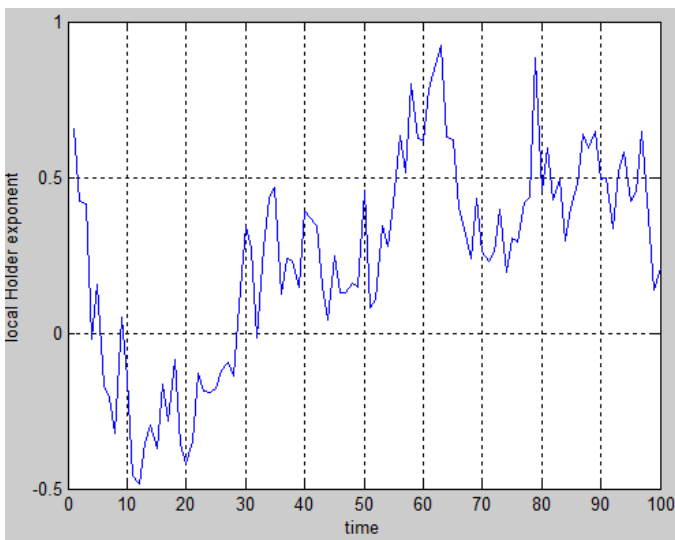


Fig. 2. Hölderexponent and scaling exponent formed for Henon mapping

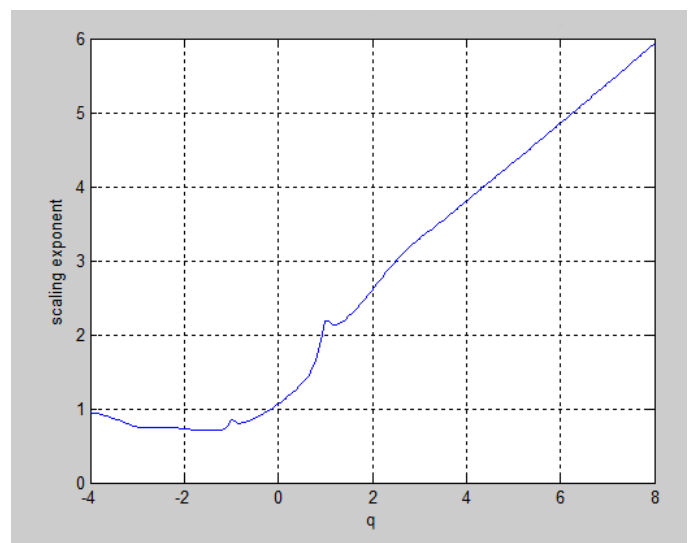
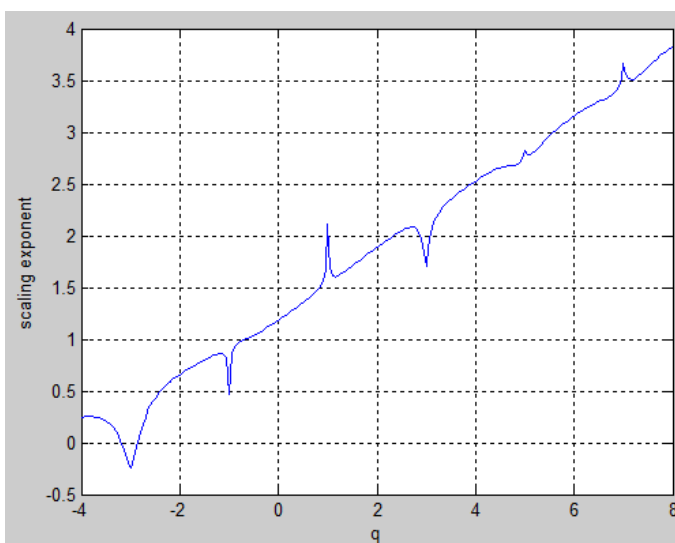


Fig. 3. Kind of scaling exponent for processed time series HRV

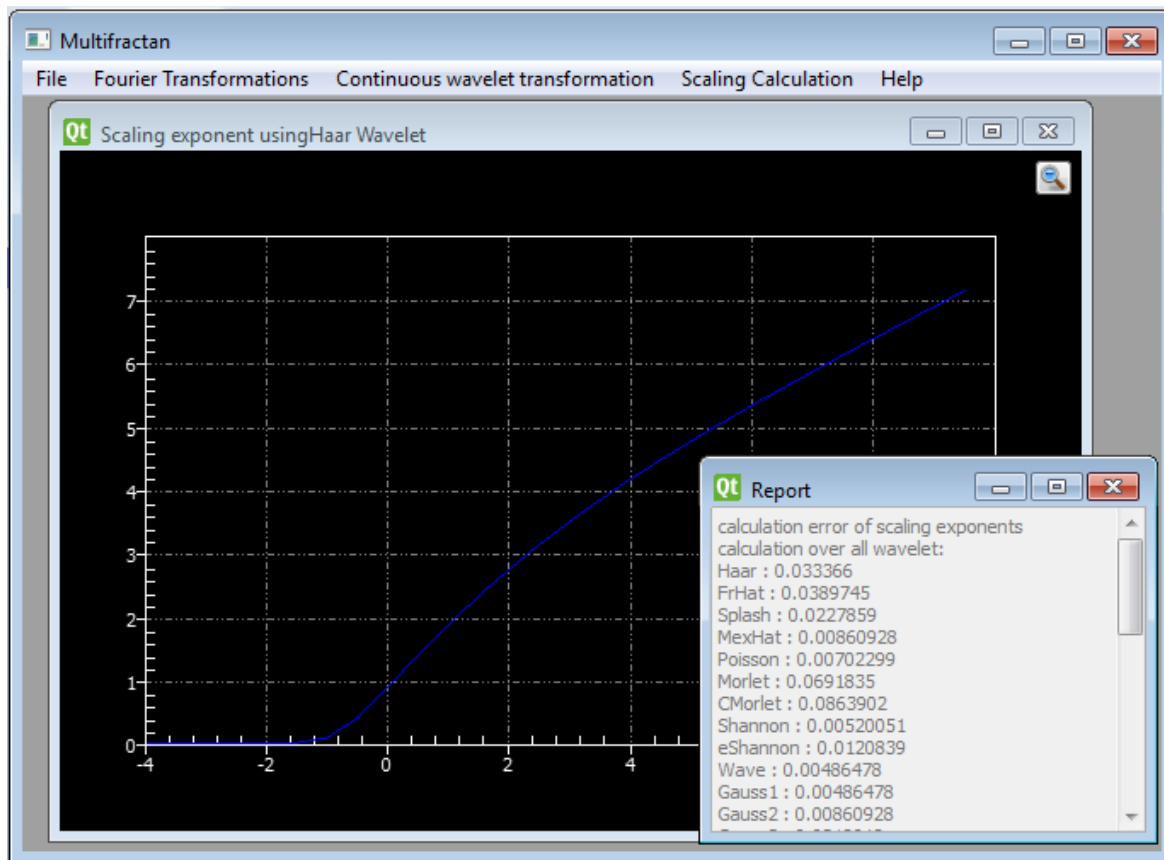


Fig. 4. Demonstration of deviation minimization functional in relation to the kind of wavelet-forming function

The second stage of researches consisted of revelation of the scaling function nonlinear dependencies for various groups of cardiac rhythm variability time series. Pursuant to peculiarities mentioned above we interpreted the results received using only the scaling exponent. In addition, here, we determined more detailed potential of used methodology.

We were interested in the problem of possibility to approximate the straight line in various ranges of scaling researching. We used the server bases [8] for these purposes. Let's demonstrate the most interesting results (from the point of view of the mentioned approach) of their processing (fig. 3).

Let's highlight the prospectivity of analyzed method in the course of comparing the results with model (test) data – in some cases we observed an expressed peculiarity of possibility to highlight the straight line in the specific areas of scaling analysis, it provides a perspective base for researching in the area of creation of (multifractal) deflection coefficients of scaling exponent from asymptotes, to determine the kind of non-deterministic chaos comparing them with (mono)fractal mapping.

Here, it is necessary to highlight that determination of the asymptotic dependence absence (in the course of analyzed approximation) doesn't imply (multi)fractality of initial time series. At the same time, indication on possibility of approximation by straight line provides adjustment for revelation of mono (fractal) by well known methods, and for permissibility to receive them by separating time series, accountable only to direct ranges of scaling change.

Fig. 4 proposes an example of optimized scaling exponent together with indication of its deviations from the direct relation

for various types of wavelet (the initial time series is generated by the Lorenz's equation system).

CONCLUSIONS

According to given arguments our research team proposes widening of basic multifraction methodology, connected mainly with optimization of scaling exponent in relation to wavelet-forming function. For these purposes we create special software, which owns a whole range of functional possibilities. We can declare that it has a special functional for calculation of scaling exponent deviations from the direct relation in relation to the kind of wavelet-forming functions. The mentioned realization is fulfilled under the auspices of RFBR and it will be published in our next article.

ACKNOWLEDGMENT

The reported study was supported by RFBR, research project № 12-08-31108-mol_a.

REFERENCES

- Antonov V.I., Zagaynov A. I., Vu. van Kuang. From the Basics of Numerical Multifractal Research to Create Automated Software Ensure [Ot osnov chislennykh multifraktalnykh issledovaniy k sozdaniyu avtomatizirovannogo programmnoy obeshcheniya]. *Nauchno-technicheskie vedomosti Sankt-Peterburgskogo gosudarstvennogo politehnicheskogo universiteta [Sci.*

and Tech. Bull. St. Petersburg state Polytechnical Univ.]. St. Petersburg, Izdatelstvo SPbGPU, 2013, no. 2 (169), pp. 71-77.

2. Astafeva N. M. Wavelet Analysis: Basic Theory and Examples of Application [Veyvlet-analiz: osnovy teorii i primery primeneniya]. *Uspekhi fizicheskikh nauk [Adv. Phys. Sci.]*, 1996, Vol. 166, no. 11, pp. 1145-1170.

3. Baevskiy R. M., Ivanov G. G., Gavrilushkin A. P., Dovgalevskiy P. Ya., Kukushkin Yu. A., Mironova T. F., Prilutskiy D. A., Semenov A. V., Fedorov V. F., Fleyshman A. N., Medvedev M. M., Chireykin L. V. Analysis of Variability of Heart Rhythm using Different Electrocardiographic Systems (part 1) [Analiz Variabelnosti Serdechnogo Ritma pri Ispolzovanii Razlichnykh Elektrokardiograficheskikh Sistem (chast 1)]. *Vestnik aritmologii [Bull. Arrhythmol.]*, 2002, no. 24, pp. 65-86.

4. FracLab: A Fractal Analysis Toolbox for Signal and Image Processing (available at: <http://fraclab.saclay.inria.fr>).

5. Gudkov G. V. The Role of Deterministic Chaos in the Structure of Variability of the Fetal Heart [Rol determinirovannogo khaosa v structure variabelnosti ritma serdtsa

ploda]. *Sovremennye problemy nauki i obrazovaniya [Mod. Prob. Sci. Educ.]*, Moscow, Krasnodar, 2008, Suppl. no. 1, pp. 413-423.

6. Malik M., Bigger D. T., Shvarts A. D., Kleyder R. E., Maliani A. M., Moss A. D., Shvarts P. D. Heart Rate Variability. Standards of Measurement, Physiological Interpretation of Clinical use [Variabelnosti serdechnogo ritma. Standarty izmereniya, fiziologicheskaya interpretatsiya klinicheskoe ispolzovanie]. *Evropeyskiy zhurnal po issledovaniyu serdtsa [Eur. J. Heart Res.]*, 1996, no. 17, pp. 354-381.

7. Mashin V. A. The Relationship of the Slope of the Regression Line of the Heart Rate Graph with Periodic and Non-Linear Dynamics of Heart in Stationary Short-Time Series [Svyaz tangensa ugla naklona linii regressii grafa serdechnogo ritma s periodicheskoy i nelineynoy dinamikoy serdtsa na korotkikh stacionarnykh otrezkach]. *Biofizika [Biophys.]*, 2006, Vol. 51, is. 3, pp. 534-538.

8. PhysioNet: The Research Resources for Complex Physiologic Signals (available at: <http://www.physionet.org>).

Мультифрактальный анализ: выявление границ применения при исследовании временных рядов

Захаров А. И., Загайнов А. И.
ВКАим. Можайского
Санкт-Петербург, Россия
zagainov239@gmail.com,
ana63916157@yandex.ru

Ходаковский В. А.
Петербургский государственный университет
путей сообщения Императора Александра I
Санкт-Петербург, Россия
hva1104@mail.ru

Аннотация. Работа посвящена выявлению возможностей разработанных к настоящему моменту мультифрактальных численных методов и средств подобной обработки временных рядов. Для этих целей использовано приложение FracLab 2.1 математического пакета MatLabR2008b, с помощью которого обработаны тестовые данные (сгенерированные, например, отображением Хенона, системой уравнений Лоренца и др.), базы записей variability сердечного ритма сервера PhysioNet и собственные ЭКГ. Построены трехмерные графики полученного вейвлет-преобразования, графики изолиний, мультифрактальные характеристики в виде экспоненты Хёлдера и скейлинговой экспоненты. В качестве вейвлет-образующей функции при вычислении спектра были использованы все возможные, предложенные в FracLab (напр. Mhat-вейвлет, DoG-вейвлет и др.). Установлены недостатки используемой реализации, полученные, прежде всего, для тестовых данных при расчетах экспоненты Хёлдера. Предложена концепция расширения мультифрактальной методологии исследования временных рядов и создание на ее основе собственного программного обеспечения.

Ключевые слова: временные ряды, метод модулей максимумов вейвлет-преобразования.

ЛИТЕРАТУРА

1. Антонов В. И. От основ численных мультифрактальных исследований к созданию автоматизированного программного обеспечения / В. И. Антонов, А. И. Загайнов, Вуван Куанг // Науч.-технич. ведомости Санкт-Петербургского гос. поли-

технич. ун-та. – СПб.: Изд-во СПбГПУ, 2013. – № 2 (169). – С. 71-77.

2. Астафьева Н. М. Вейвлет-анализ: основы теории и примеры применения / Н. М. Астафьева // Успехи физ. наук. – 1996. Т. 166, № 11. – С. 1145-1170.

3. Баевский Р. М. Анализ variability сердечного ритма при использовании различных электрокардиографических систем (ч. 1) / Р. М. Баевский, Г. Г. Иванов, А. П. Гаврилушкин и др. // Вестн. аритмологии. – 2002. – № 24. – С. 65-86.

4. FracLab: A fractal analysis toolbox for signal and image processing. – URL: <http://fraclab.saclay.inria.fr>.

5. Гудков Г. В. Роль детерминированного хаоса в структуре variability ритма сердца плода / Г. В. Гудков // Современ. проблемы науки и образования. – М.; Краснодар, 2008. – Прил. № 1. – С. 413-423.

6. Малик М. Variability сердечного ритма. Стандарты измерения, физиологическая интерпретации клиническое использования / М. Малик, Д. Т. Биггер, А. Д. Шварц и др. // Европейский журнал по исследованию сердца. – 1996. – № 17. – Р. 354-381.

7. Машин В.А. Связь тангенса угла наклона линии регрессии графа сердечного ритма с периодической и нелинейной динамикой сердца на коротких стационарных отрезках / В. А. Машин // Биофизика. – 2006. – Т. 51, вып. 3. – С. 534-538.

8. PhysioNet: The research resources for complex physiologic signals. – URL: <http://www.physionet.org>.

Определение вероятностно-временных характеристик процесса функционирования системы управления сетью передачи данных при поддержании показателей QoS

Канаев А. К., Сахарова М. А., Бенета Э. В.
Петербургский государственный университет
путей сообщения Императора Александра I
Санкт-Петербург, Россия
kanaevak@mail.ru, zuvakamariya@mail.ru, elinabeneta@yandex.ru

Аннотация. Предложена математическая модель для определения вероятностно-временных характеристик процесса функционирования системы управления сетью передачи данных с использованием систем поддержки принятия решений, ориентированной на поддержание заданных показателей качества обслуживания QoS. Применение систем поддержки принятия решений при управлении сетями передачи данных позволяет выполнять задачи по своевременному обнаружению отклонений параметров QoS в условиях ограниченных временных и вычислительных ресурсов.

Ключевые слова: управление сетями, сети передачи данных, системы поддержки принятия решений, топологическое преобразование стохастических сетей, нейронные сети, качество обслуживания.

ВВЕДЕНИЕ

Для обеспечения бесперебойной работы всех компонентов сети передачи данных (СПД) и предоставления пользователям услуг заданного качества требуется такая система управления (СУ), которая будет своевременно обнаруживать отклонения параметров качества обслуживания (QoS) СПД от требуемых значений и затем оценивать техническое состояние сети. При ограниченных временных и вычислительных ресурсах это можно сделать, используя специальные системы поддержки принятия решений (СППР). Однако для использования СППР в управлении СПД при оценке технического состояния сети и формировании плана восстановления ее работоспособности необходимо применять средства интеллектуального анализа данных.

В настоящей статье предложена блок-схема процесса функционирования СУ СПД с использованием СППР, для которого реализована математическая модель определения вероятностно-временных характеристик процесса, ориентированная на поддержание заданных показателей QoS.

ПОКАЗАТЕЛИ КАЧЕСТВА ОБСЛУЖИВАНИЯ В СЕТЯХ ПЕРЕДАЧИ ДАННЫХ

Под качеством обслуживания (QoS) понимается полезный эффект от обслуживания, который определяется сте-

пенью удовлетворения пользователя как от полученной услуги, так и от самой системы обслуживания [6, 8].

Для формирования решений в СУ СПД по оценке технического состояния СПД установим следующие группы диагностических показателей QoS [3, 8]:

- сетевые ориентированные параметры QoS, которые непосредственно связаны с определением уровня сетевого совершенства: учитываются ресурсные возможности СПД, определяется готовность ресурсов СПД к использованию, оценивается качество передачи информации.

Например [11, 12], IPLRmax (показатели потерь IP-пакетов во время их транспортировки – вероятность потери пакетов), IPSLBRmax (показатели потерь последовательностей IP-пакетов во время их транспортировки – коэффициент потерь IP-пакетов), IPERmax (показатели некорректной транспортировки IP-пакетов – количество пакетов с ошибками), K_г min (показатели эксплуатационной надежности сетевого оборудования – коэффициент готовности оборудования);

- сервисориентированные параметры качества обслуживания QoS, которые непосредственно связаны с определением качества предоставления услуги, имеющей такие потребительские свойства: обеспеченность, удобство пользования, действенность (доступность, целостность, непрерывность), безопасность использования.

Например [11, 12], IPTDmax (показатели задержки в передаче IP-пакетов – граница непревышения среднего значения величины задержек пакетов), IPDVmax (показатели вариаций в задержке IP-пакетов – граница непревышения среднего значения вариации задержек пакетов), Pmax (IPTDmax) (показатели задержки в передаче IP-пакетов – порог допустимой вероятности превышения значения IPTDmax), PIPSAmin (PIA-показатели доступности сетевой услуги – нижняя граница допустимого процента времени доступности услуги), TPUmax (TIU – показатели доступности сетевой услуги – верхняя граница допустимого времени недоступности услуги).

Также выделяют группу показателей сервис/сетезависимых, которые не связаны с определением качества услуги или уровня сетевого совершенства. Однако в статье они рассматриваться не будут.

ПРОЦЕСС ФУНКЦИОНИРОВАНИЯ СИСТЕМЫ УПРАВЛЕНИЯ СЕТЬЮ ПЕРЕДАЧИ ДАННЫХ, ОБЕСПЕЧИВАЮЩЕЙ ПОДДЕРЖАНИЕ ЗАДАННЫХ ПОКАЗАТЕЛЕЙ КАЧЕСТВА ОБСЛУЖИВАНИЯ

Процесс функционирования СУ СПД при периодических запросах на оценку технического состояния СПД на уровне управления сетью (NMS) классической концепции управления сетями связи (TMN) [2, 9, 10] осуществляется через равные интервалы времени, при этом должны поддерживаться непрерывный контроль параметров QoS СПД и постоянное пополнение новыми знаниями базы данных СУ СПД.

Блок-схема процесса функционирования СУ СПД представлена на рис. 1.

В качестве примера рассмотрим пуассоновский закон распределения случайных величин. Однако можно исполь-

зовать и другие законы распространения случайных величин. В данном случае случайными величинами являются запросы на оценку состояния по данным показателей QoS СПД [5].

Пусть на вход СУ СПД поступает пуассоновский поток запросов на оценку состояния СПД с функцией распределения (ФР) $W(t)$ и реализуется за время t_w .

Оценка технического состояния СПД выполняется с ФР $Z(t)$ за время t_z , далее измеряются параметры QoS, которые включают измерение и определение сетевых показателей QoS СПД за время t_b с ФР $B(t)$ и сервисориентированных показателей QoS за время t_m с ФР $M(t)$, результаты которых обрабатываются за время t_l с ФР $L(t)$ в СУ СПД с помощью СППР или самостоятельно лицом, принимающим решение (ЛПР).

После анализа диагностических данных параметров QoS СПД принимается решение о необходимости восстановления СПД. Если все показатели параметров QoS соответствуют нормативным значениям исправного состояния СПД, то с вероятностью $(1-Pv)$ за время t_{11} с ФР $H(t)$ цикл управления СУ СПД завершается.

Если хотя бы один показатель параметра QoS не соответствует нормативным значениям исправного состояния СПД, то необходимо провести ряд мероприятий по восстановлению работоспособного состояния СПД, а именно: определить причину изменения показателей QoS СПД за время t_x с ФР $X(t)$, изменить параметры СПД за время t_k с ФР $K(t)$. Далее с помощью системы технической диагностики (СТД) проводят восстановление СПД за время t_v с ФР $V(t)$.

По окончании восстановительных работ проводят контрольные измерения показателей QoS СПД, чтобы убедиться в достоверности восстановления требуемого состояния СПД.

ОПРЕДЕЛЕНИЕ ВЕРОЯТНОСТНО-ВРЕМЕННЫХ ХАРАКТЕРИСТИК ПРОЦЕССА ФУНКЦИОНИРОВАНИЯ СУ СПД

Для оценки оперативности функционирования СУ СПД можно использовать различные модели (графо-аналитические, марковские цепи, топологическое преобразование стохастических сетей (ТПСС) и другие) [4, 5, 13, 14].

Модель ТПСС позволяет декомпозировать сложный процесс на элементарные процессы, каждый из которых будет характеризоваться функцией распределения или средним временем выполнения процесса [13, 14]. Модель ТПСС будем использовать для решения поставленной задачи.

Представим этапы разработки математической модели процесса функционирования СУ СПД при помощи ТПСС [13]:

1 этап – качественно описывается процесс функционирования СУ СПД, предусматривается пошаговая декомпозиция процесса на множество элементарных процессов, определяются параметры процесса, условия функционирования СУ СПД, внутренние и внешние взаимосвязи;

2 этап – определяются исходные данные для описания каждого элементарного процесса, функция плотности, вероятности времени реализации элементарных процессов и соответствующие им преобразования;

3 этап – составляется стохастическая сеть. При описании процесса функционирования СУ СПД согласно блок-схеме

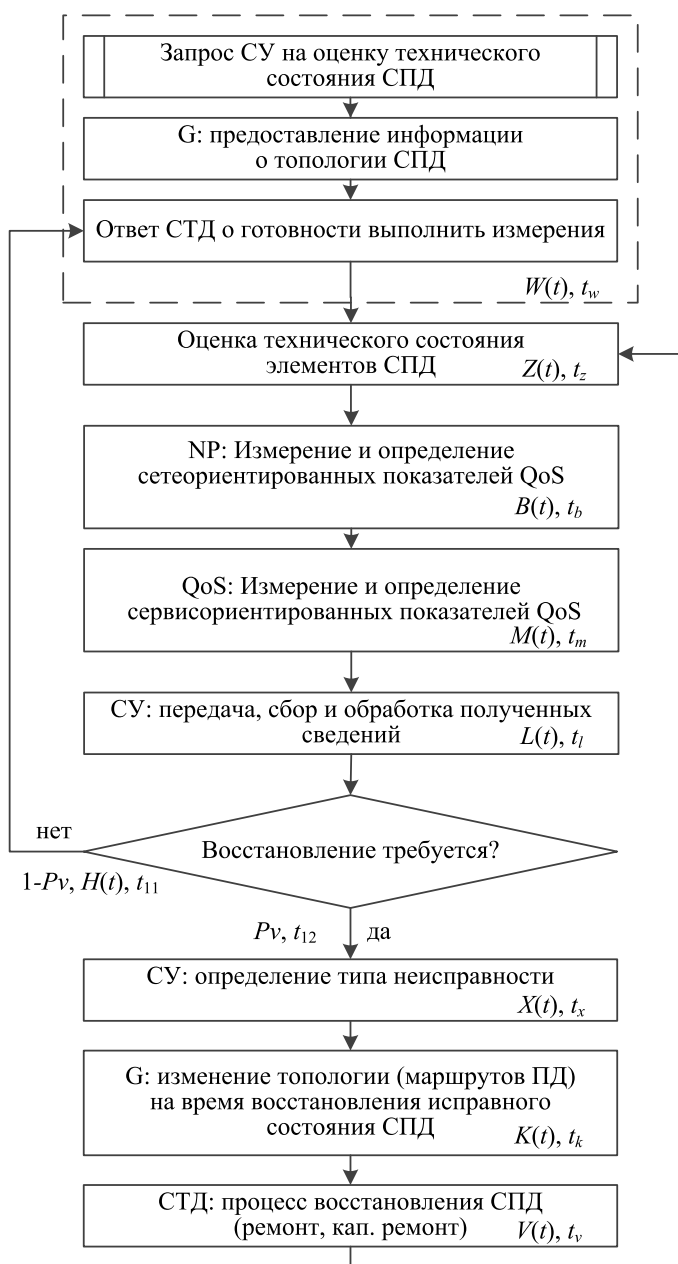


Рис. 1. Блок-схема процесса функционирования СУ СПД при измерении параметров QoS СПД

(рис. 1) решается основная задача СУ – восстановление требуемого состояния СПД;

4 этап – производится топологическое приведение стохастической сети к эквивалентной, при этом эквивалентная функция определяется из топологического уравнения Мейсона согласно выражению [13, 14]:

$$H = 1 + \sum_{k=1}^K (-1)^k Q_3(s) = 0. \quad (1)$$

Эквивалентная функция $Q_3(s)$ для сети произвольной структуры имеет вид

$$Q_3(s) = \frac{\prod_{j=1}^k f_j(s)}{1 + \sum_{l=1}^M (-1)^l W_l(s)}, \quad (2)$$

где k – количество функций $f_j(s)$ в прямой цепи стохастической сети;

5 этап – определяются вероятностно-временные характеристики СС при помощи метода двухмоментной аппроксимации.

Вероятность своевременной передачи заявок является характеристической функцией, позволяющей определить k -е начальные моменты случайного времени реализации стохастической сети. Первый начальный момент, или математическое ожидание случайного процесса, можно представить в виде выражения

$$M_k = (-1)^k \frac{d^k}{ds^k} \left[\frac{Q(s)}{Q(0)} \right]_{s=0}. \quad (3)$$

Отсюда среднее время передачи заявки на измерение, определяемое как начальный момент первого порядка,

$$\bar{t}_{\Pi} = -\frac{d}{ds} \left[\frac{Q(s)}{Q(0)} \right]_{s=0}. \quad (4)$$

Дисперсия времени передачи заявки на измерение $D[t_{\Pi}]$, определяемая как второй центральный момент, представлена выражением

$$D[t_{\Pi}] = \frac{d^2}{ds^2} \left[\frac{Q(s)}{Q(0)} \right]_{s=0} - \left\{ -\frac{d}{ds} \left[\frac{Q(s)}{Q(0)} \right]_{s=0} \right\}^2; \quad (5)$$

6 этап – оценивается эффективность исследуемой системы.

Вычисление математического ожидания и дисперсии позволяет приближенно определить функцию распределения времени передачи как неполную гамма-функцию [7, 8]:

$$F(t) = \begin{cases} 0, & t < 0; \\ \int_0^t \frac{\mu^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp[-\mu x] dx, & t \geq 0, \end{cases} \quad (6)$$

где $\alpha = \frac{\bar{t}_{\Pi}^{-2}}{D[t_{\Pi}]}$ – параметры формы; $\mu = \frac{\bar{t}_{\Pi}}{D[t_{\Pi}]}$ – параметры

масштаба.

Процесс функционирования СУ СПД при периодических запросах на оценку технического состояния СПД представлен на рис. 2 в виде стохастической сети, под которой понимается совокупность взаимосвязанных узлов (вершин) и ветвей [13], соединение которых соответствует блок-схеме процесса функционирования СУ СПД (см. рис. 1).

Для определения вероятностно-временных характеристик процесса функционирования СУ СПД с помощью метода ТПСС введем следующие ограничения и допущения:

- 1) функции распределения случайных величин относятся к классу экспоненциальных;
- 2) вероятности, соответствующие ветвям стохастической сети, определяются статистическими методами;
- 3) время реализации отдельных операций искомого процесса имеют экспоненциальное распределение;
- 4) модель предполагает отсутствие новых заявок до окончания обработки предыдущей;
- 5) потоки заявок являются неконкурирующими.

Необходимо определить функцию распределения $Q_k(t)$ и среднее время реализации цикла управления СУ СПД.

Для исследуемой стохастической сети процесса функционирования СУ СПД (рис. 2) эквивалентная функция $Q_k(t)$, имеет следующий вид в соответствии с выражением (2):

$$Q_k(s) = \frac{w \cdot z \cdot b \cdot m \cdot l \cdot P_v \cdot h}{w+s \cdot z+s \cdot b+s \cdot m+s \cdot l+s \cdot P_v \cdot h+s} \cdot \frac{1}{1 - \frac{z \cdot b \cdot m \cdot l \cdot (1-P_v) \cdot x \cdot k \cdot v}{z+s \cdot b+s \cdot m+s \cdot l+s \cdot (1-P_v) \cdot x+s \cdot k+s \cdot v+s}}. \quad (7)$$

Таким образом, для исследуемой стохастической сети (рис. 2) начальный момент случайного времени реализации СС модели функционирования СУ СПД при $s = 0$ будет $M1(0) = 189,1$. Дисперсия времени передачи заявки на измерение, определяемая как второй центральный момент по выражению (5), для $s = 0$ составит $D = 3,903 \cdot 10^4$.

Для определения функции распределения как неполной гамма-функции параметры формы при заданных исходных данных $\alpha = 0,916$ и масштаба $\mu = 4,845 \cdot 10^{-3}$.

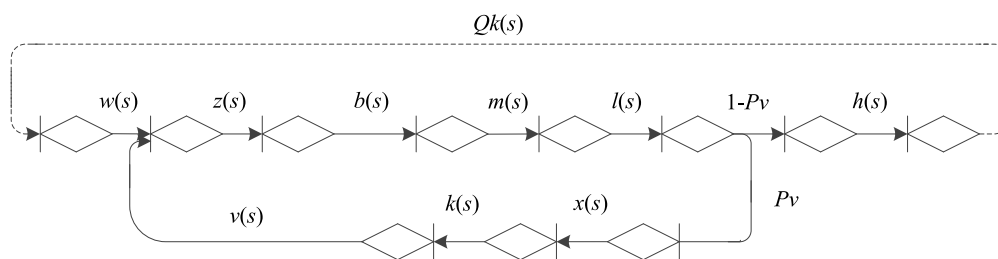


Рис. 2. Стохастическая сеть процесса функционирования СУ СПД при измерении параметров QoS

На рис. 3 представлена функция распределения времени реализации цикла управления СУ СПД при изменении вероятности P_V , при котором определяется вероятность появления мероприятий по восстановлению СПД. При увеличении времени реализации цикла управления растет значение вероятности завершения цикла управления за отведенное время. На рис. 3 показано, что с ростом значения вероятности

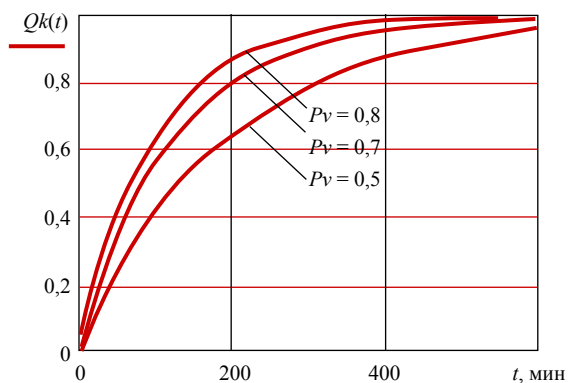


Рис. 3. Функция распределения времени реализации цикла управления СУ СПД

сти P_V значительно увеличивается время реализации цикла управления СУ СПД.

С помощью представленной математической модели можно определить среднее время $T(P_V)$ реализации цикла управления СУ СПД (рис. 4), который включает ряд меро-

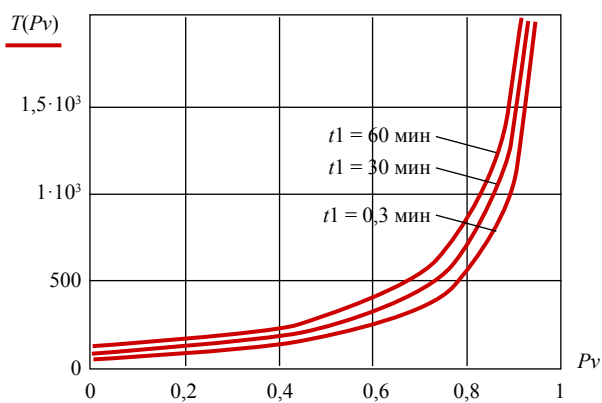


Рис. 4. Среднее время реализации цикла управления СУ СПД

приятий по восстановлению СПД, обеспечивающей выполнение требований к QoS.

Выявлено, что значение среднего времени реализации цикла управления СУ СПД при увеличении вероятности появления события P_V зависит от времени обработки диагностических показателей QoS СПД. Причем существенная разница значений среднего времени $T(P_V)$ наблюдается при значениях вероятности P_V более 0,4 для разных значений времени обработки показателей QoS t_1 .

Например, при использовании нейронных сетей для анализа данных затрачивается $t_y = 0,3$ мин, при этом $T(P_V)$ существенно меньше, чем в случае, если обработку полученных диагностических показателей проводит администратор сети

без использования СППР. Таким образом, при уменьшении времени на реализацию процесса передачи, сбора и обработки диагностической информации СПД t_1 уменьшается и среднее время $T(P_V)$ реализации цикла управления СУ СПД.

ЗАКЛЮЧЕНИЕ

Предложенная модель для определения вероятностно-временных характеристик процесса функционирования СУ СПД, обеспечивающей поддержание показателей QoS, работоспособна и чувствительна к изменению исходных данных, а результаты моделирования согласуются со статистическими данными по затрачиваемому времени на восстановление СПД.

Предложенная модель может применяться при различных законах распределения случайных величин, что позволяет сделать вывод о ее гибкости.

Расширить математическую модель процесса функционирования СУ СПД можно за счет добавления подпроцессов или декомпозиции отдельных подпроцессов в отдельную стохастическую сеть.

Результаты представленной математической модели можно учитывать при разработке методики функционирования СППР в структуре СУ СПД.

ЛИТЕРАТУРА

1. Буренин А. Н. Теоретические основы управления современными телекоммуникационными сетями: монография / А. Н. Буренин, В. И. Курносов. – М.: Наука, 2011. – 464 с.
2. Гребешков А. Ю. Управление сетями электросвязи по стандарту TMN: учеб. пособие / А. Ю. Гребешков. – М.: Радио и связь, 2004. – 155 с.
3. Игнациук П. Контроль перегрузки в сетях передачи данных [Congestion control in data transmission networks] / П. Игнациук, А. Бартожевич. – Лодзь: Спрингер, 2013. – 379 с.
4. Канаев А. К. Математическая модель процесса функционирования системы управления сетью передачи данных при запросах на определение технического состояния СПД / А. К. Канаев, М. А. Сахарова, Е. В. Скуднева // Изв. ПГУПС. – 2015. – № 1 (42). – С. 91-98.
5. Канаев А. К. Модель процесса функционирования системы управления сетью передачи данных при определении технического состояния элементов сети / А. К. Канаев, М. А. Сахарова // Транспорт Урала. – 2015. – № 2 (45). – С. 48-51.
6. Канаев А. К. Поддержание заданных показателей качества обслуживания в условиях отсутствия информации о видах аномалий в СПД / А. К. Канаев, М. А. Камынина // Геополитические факторы устойчивого развития Арктики и инновационные технологии прогнозирования и предотвращения чрезвычайных ситуаций: тезисы докл. VI Междунар. конгресса науч.-практич. конф., СПб., 15–16 ноя. 2013 г. – СПб., 2013. – С. 158-163.
7. Канаев А. К. Формирование элементов системы управления сетью передачи данных с применением аппарата нейронных сетей / А. К. Канаев, М. А. Камынин, Е. В. Опарин // Бюл. результатов науч. исследований. – 2012. – № 3 (2). – С. 47-55.
8. Коханович Г. Ф. Сети передачи данных / Г. Ф. Коханович, В. М. Чуприн. – Киев: МК-Пресс, 2006. – 272 с.

9. М.3000 Обзор рекомендаций TMN: Рекомендация сектора ИТУ по телекоммуникациям. Сер. М. – 2000.

10. М.3400 Обзор рекомендаций TMN: Рекомендация сектора ИТУ по телекоммуникациям. Сер. М. – 2000.

11. Необходимость производительности сети для услуг IP: Рекомендация Междунар. союза электросвязи. Сер. Т, У.1451. – 2002.

12. Передача IP-пакетов и предоставление параметров производительности: Рекомендация Междунар. союза электросвязи. Сер. Т, У. 1540. – 2002.

13. Привалов А. А. Метод топологического преобразования стохастических сетей и его использование для анализа систем связи ВМФ / А. А. Привалов. – СПб.: ВМА, 2000. – 166 с.

14. Привалов А. А. Моделирование процесса управления в единой системе мониторинга и администрирования связи ОАО «РЖД» / А. А. Привалов, А. П. Вандич, Е. В. Опарин // Цели развития тысячелетия» и инновационные принципы устойчивого развития арктических регионов: тезисы докл. IV Междунар. конгресса, СПб., 24–25 ноя. 2011 г. – СПб., 2011. – С. 130-133.

Determination Probability-Time Characteristics of the Process Functioning Management System Network While Maintaining QoS Performance

Kanaev A. K., Sakharova M. A., Beneta E. V.

Petersburg State Transport University

St. Petersburg, Russia

kanaevak@mail.ru zuvakamariya@mail.ru elinabeneta@yandex.ru

Abstract. In the article propose a mathematical model the probability-time characteristics of the process functioning system data network employment decision support systems, focused on maintaining the quality of service parameters specified QoS. The use of decision support systems in the management of data networks allows you to perform tasks for the timely detection of deviations QoS parameters in a limited time and computing resources.

Keywords: management of networks, data networks, decision support systems, topological transformation of stochastic networks, neural networks, quality of service.

REFERENCES

1. Burenin A. N., Kurnosov V. I. Teoreticheskiye osnovy upravleniya sovremennymi telekommunikatsionnymi setyami [Theoretical Basis of Modern Telecommunication Networks Management]. Moscow, Nauka, 2011. 221 p.
2. Grebeshkov A. Yu. Upravleniye setyami electrosvyazi po standartu TMN: uchebnoe posobie [Management of Telecommunication Networks on Standard TMN: Study Guide]. Moscow, Radio i svyaz, 2004. 155 p.
3. Ignaciuk P., Bartoszewicz A. Congestion control in data transmission networks. Lodz, Springer, 2013. 379 p.
4. Kanaev A. K., Sakharova M. A., Skudneva E. V. A Mathematical Model of the Process Management System Data Network in the Request for the Technical Condition of the SAP [Matematicheskaya model processa funktsionirovaniya sistemy upravleniya setyu peredachi dannykh pri zaprosakh na opredelenie tekhnicheskogo sostoyaniya SPD]. *Izvestiya PGUPS [News PGUPS]*, 2015, no. 1 (42), pp. 91–98.
5. Kanaev A. K., Sakharova M. A. Model of Management System Data Network in Determining the Technical Condition of the Network Elements [Model protsessa funktsionirovaniya sistemy upravleniya setyu peredachi dannykh pri opredelenii tekhnicheskogo sostoyaniya elementov seti]. *Transport Urala [Transp. Urals]*, 2015, no. 2 (45), pp. 48–51.
6. Kanaev A. K., Kamynina M. A. Maintaining a Set of Indicators of Quality of Service in the Absence of Information about the Kinds of Anomalies in the SPD [Podderzhaniye zadannykh pokazateley kachestva obsluzhivaniya v usloviyakh otsutstviya informatsii o vidakh anomalii v SPD]. *Trudy "Nauchno-prakticheskaya konferentsiya" (Proc. "Geopolitical factors of sustainable development of the Arctic and innovative technologies to predict and prevent emergencies")*. St. Petersburg, 2013, pp. 158–163.
7. Kanaev A. K., Kamynina M. A., Oparin E. V. Formation of Components of Network Management Data Using the Neural Network Apparatus [Formirovaniye elementov sistemy upravleniya setyu peredachi dannykh s primeneniem apparata neyronnykh setey]. *Byulleten rezultatov nauchnykh issledovaniy [Bulletin of research results]*, 2012, no. 3 (2), pp. 47–55.
8. Kokhanovich G. F., Chuprin V. M. Data Transfer Networks [Seti peredachi dannykh]. Kiev, MK-Press, 2006. 272 p.
9. TMN and Network Maintenance: International Transmission Systems, Telephone Circuits, Telegraphy, Facsimile and Leased Circuits. Telecommunications Management Network. ITU-T M.3000 Telecommunication standardization sector of ITU, Ser. M: Overview of TMN Recommendations, 2000.
10. TMN and Network Maintenance: International Transmission Systems, Telephone Circuits, Telegraphy, Facsimile and Leased Circuits. Telecommunications Management Network TMN Management Functions. ITU-T M.3400 Telecommunication standardization sector of ITU, Ser. M: Overview of TMN Recommendations, 2000.
11. Network Performance Objectives for IP Based Services. ITU-T Recommendation Y.1541. 2002.
12. IP Packet Transfer and Availability Performance parameters. ITU-T Recommendation Y.1540. 2002.
13. Privalov A. A. The Method of the Topological Transformation of Stochastic Networks and its use for the Analysis of Communication Systems of the Navy [Metod topologicheskogo preobrazovaniya stokhasticheskikh setey i ego ispolzovaniye dlya analiza sistem svyazi VMF]. St. Petersburg, VMA, 2000. 166 p.
14. Privalov A. A., Vandich A. P., Oparin E. V. Simulation Process Management in a Single System for Monitoring and Communication Manager of JSC "Russian Railways" [Modelirovaniye protsessa upravleniya v edinoi sisteme monitoringa i administrirovaniya svyazi OAO "RZHD"]. *Trudy "IV mezhdunarodnyy kongress" (Proc. "Millennium Development Goals and the innovative principles of sustainable development of the Arctic regions")*. St. Petersburg, 2011. pp. 130–133.

Оценка надежности структурно сложных технических комплексов с помощью моделей байесовских сетей доверия в среде GeNIe

Дорожко И. В., Тарасов А. Г.
Военно-космическая академия
им. А. Ф. Можайского
Санкт-Петербург, Россия
Doroghko-Igor@yandex.ru

Барановский А. М.
Петербургский государственный университет путей
сообщения Императора Александра I
Санкт-Петербург, Россия
bamvka@mail.ru

Аннотация. Предлагается новый подход к оценке показателей надежности сложных иерархических комплексов. Подход базируется на использовании математического аппарата динамических байесовских сетей доверия. Разработанная модель позволяет учитывать поступление новой информации о надежности элементов комплекса, что дает возможность прогнозировать значение вероятности безотказной работы технического комплекса, корректировать периодичность технического обслуживания и стратегию восстановления комплекса.

Ключевые слова: надежность, вероятность безотказной работы, байесовская сеть доверия.

ВВЕДЕНИЕ

В настоящее время оценка показателей надежности сложных технических комплексов является одной из важнейших задач [1–3, 5, 9, 10, 12]. С этапа формирования тактико-технических требований рассчитываются показатели надежности. Весь процесс разработки сложных технических комплексов сопровождается испытаниями на надежность. Конечной целью расчета показателей надежности технических комплексов является оптимизация конструктивных решений, параметров комплексов и режимов эксплуатации, организация технического обслуживания и ремонтов.

В технических заданиях (ТЗ) на выполнение опытно-конструкторских работ (ОКР) по созданию (модернизации) сложных технических комплексов обязательно присутствует раздел «Требования надежности», где, как правило, приводятся требуемые значения вероятности безотказной работы (ВБР) комплекса, коэффициента готовности, максимального времени восстановления и т. д. [7–9].

В соответствии с принятым вариантом схемно-конструкторского решения рассчитываются показатели надежности комплекса, которые позволяют сделать вывод о принципиальной возможности достичь требуемого уровня надежности в соответствии с ТЗ и перейти к этапу изготовления и заводских испытаний комплекса. При этом очевидна необходимость уточнения показателей надежности по результатам экспериментальной отработки на последующих этапах жизненного цикла изделия, что указывается в программе обеспечения надежности. В процессе работы комплексов могут отказывать отдельные элементы, поэтому персоналу необходимо иметь актуальную информацию о значении ВБР комплекса и его составных частей в текущий момент или на некоторый последующий период времени для принятия

решения о дальнейшей эксплуатации или ее приостановке, для планирования технического обслуживания и т. д.

В статье предлагается модель, основанная на байесовских сетях доверия (БСД) [4, 6, 11, 13–17], которая позволяет:

- оценивать ВБР комплекса по априорным данным о ВБР составных элементов и при поступлении новой информации о ВБР составных элементов;
- прогнозировать значение ВБР комплекса;
- поддерживать принятие решений при поиске дефектов в отказавшей системе.

Разработанная математическая модель оценки ВБР в виде БСД такова:

$$\langle \mathbf{G}, \mathbf{PAR} \rangle, \quad (1)$$

где \mathbf{G} – ациклический направленный граф; \mathbf{PAR} – множество параметров, определяющих БСД.

Вершинами ациклического направленного графа являются блоки комплекса с учетом иерархии:

$$\mathbf{B} = \{b_i^{<j>} \mid i = \overline{1, n}, j = \overline{1, m}\}, \quad (2)$$

где b – название блока; i – номер блока; n – число блоков; j – номер уровня; m – число уровней.

По расположению в байесовской сети доверия выделяют следующие типы переменных – вершин [11, 13–17]:

- родители ($b_i^{<j>} = \text{parent}(b_i^{<j+1>})$) – множество вершин $b_i^{<j>}$ таких, что существует вершина $b_i^{<j+1>}$ и направленное ребро от $b_i^{<j>}$ к $b_i^{<j+1>}$;
- дети ($b_i^{<j+1>} = \text{children}(b_i^{<j>})$) – множество вершин $b_i^{<j+1>}$ таких, что существует вершина $b_i^{<j>}$ и направленное ребро от $b_i^{<j>}$ к $b_i^{<j+1>}$.

Каждый блок $b_i^{<j>}$ может находиться в двух состояниях:

- 1) $\text{Rabot}_i^{<j>}$ – работоспособное состояние i -го блока j -го уровня;
- 2) $\text{Otkaz}_i^{<j>}$ – отказ i -го блока j -го уровня.

Параметрами сети \mathbf{PAR} являются:

- 1) для блоков, не имеющих родительских переменных, ($\forall b_i^{<j>} : \neg(\exists b_i^{<j-1>} = \text{parent}(b_i^{<j>}))$):

- ВБР и вероятности отказов в начальный момент времени (безусловные вероятности) – $P(\text{Rabot}_i^{<j>})_{t=0} = P_i^{<j>}(t=0)$ и $P(\text{Otkaz}_i^{<j>})_{t=0} = Q_i^{<j>}(t=0)$ (обозначения буквами «P» и «Q» вероятности безотказной работы и вероятности отказа приведены в [7–9] и широко используются в теории надеж-

ности, однако в теории вероятностей принято обозначение вероятности буквой «P», поэтому в статье приводится данное пояснение);

- вероятность того, что блок работоспособен в текущий момент времени, при условии, что блок был работоспособен в предыдущий момент (условная вероятность) – $P((Rabot_i^{<j>})_t / (Rabot_i^{<j>})_{t-1})$;
- вероятность отказа блока в текущий момент времени при условии, что блок был работоспособен в предыдущий момент (условная вероятность) – $P((Otkaz_i^{<j>})_t / (Rabot_i^{<j>})_{t-1})$;
- вероятность того, что блок работоспособен в текущий момент времени, при условии, что блок отказал в предыдущий момент (условная вероятность) – $P((Rabot_i^{<j>})_t / (Otkaz_i^{<j>})_{t-1})$;
- вероятность отказа блока в текущий момент времени при условии, что блок отказал в предыдущий момент (условная вероятность) – $P((Otkaz_i^{<j>})_t / (Otkaz_i^{<j>})_{t-1})$;

2) для блоков, имеющих родительские переменные ($\forall b_i^{<j>} : \exists b_i^{<j-1>} = \text{parent}(b_i^{<j>})$), задаются условные вероятности в зависимости от структурно-логической схемы надежности, т. е. оценивается влияние элементов более низкого иерархического уровня на блоки более высокого иерархического уровня.

Например, если блоки более низкого иерархического уровня $b_i^{<j-1>}$ соединены последовательно с точки зрения надежности, то условная вероятность того, что блок более высокого иерархического уровня $b_i^{<j>}$ работоспособен при условии работоспособности всех блоков более низкого уровня $b_i^{<j-1>}$, равна 1, а условная вероятность того, что блок более высокого иерархического уровня $b_i^{<j>}$ работоспособен при условии отказа хотя бы одного из блоков более низкого уровня $b_i^{<j-1>}$, равна 0.

На рис. 1 показан пример фрагмента структурно сложного комплекса с 4-уровневой структурой: на первом (нижнем) уровне блоки $b_1^{<1>}, b_2^{<1>}, b_3^{<1>}, b_4^{<1>}, b_5^{<1>}$ соединены в виде «моста» с точки зрения надежности; на втором – блоки $b_1^{<2>}, b_2^{<2>}, b_3^{<2>}$ соединены по мажоритарной схеме «2 из 3», а $b_4^{<2>}, b_5^{<2>}$ соединены параллельно; на третьем уровне блоки $b_3^{<3>}, b_2^{<3>}$ соединены последовательно в блоке четвертого уровня $b_1^{<4>}$.

На рис. 2 показана структура байесовской сети доверия для фрагмента технического комплекса, изображенного на

рис. 1, в открытой для программирования среде построения байесовских сетей GeNIe.

В табл. 1–4 приведены примеры задания параметров (условных вероятностей) в БСД для блоков, имеющих в своем составе блоки более низкого уровня, соединенные в виде «моста», по мажоритарной схеме «2 из 3», параллельно и последовательно. В табл. 1 представлен набор условных вероятностей для блока 2-го уровня $b_1^{<2>}$, состоящего из блоков $b_1^{<1>}, b_2^{<1>}, b_3^{<1>}, b_4^{<1>}, b_5^{<1>}$, соединенных в виде «моста». В табл. 2 представлен набор условных вероятностей для блока 3-го уровня $b_1^{<3>}$, состоящего из блоков $b_1^{<2>}, b_2^{<2>}, b_3^{<2>}$, соединенных по мажоритарной схеме «2 из 3». В табл. 3 представлен набор условных вероятностей для блока 3-го уровня $b_2^{<3>}$, состоящего из блоков $b_4^{<2>}, b_5^{<2>}$, соединенных параллельно. В табл. 4 представлен набор условных вероятностей для блока 4-го уровня $b_1^{<4>}$, состоящего из блоков $b_3^{<3>}, b_2^{<3>}$, соединенных последовательно.

На рис. 3 представлена динамическая байесовская сеть доверия фрагмента структурно-сложного технического комплекса, построенная по априорной информации (данным о ВБР элементов и о схемах их соединения). В среде GeNIe предоставляется возможность посмотреть графики ВБР всего комплекса или отдельных блоков.

Главное назначение байесовских сетей доверия – апостериорный вывод. Суть апостериорного вывода заключается в следующем: при поступлении свидетельства (новой информации) необходимо приравнять к нулю несовместимые со свидетельством вероятности исходов и нормировать оставшиеся вероятности, учитывая априорные данные с тем, чтобы вероятности оставшихся исходов в сумме снова давали единицу [11, 13–17]. Априорные данные пересчитываются, в результате получается апостериорная оценка, которая, в свою очередь, будет являться априорной информацией в случае поступления новых свидетельств.

Апостериорный вывод в БСД основан на процедуре опроса БСД. Под опросом БСД понимается вычисление безусловного распределения переменной на основе условных распределений БСД и поступивших данных (свидетельств) [11, 13–17]. Под свидетельством в данном случае понимается информация об отказе или работоспособности блоков.

На рис. 4 приведены графики изменения ВБР отдельных блоков по априорным данным и при поступлении новой

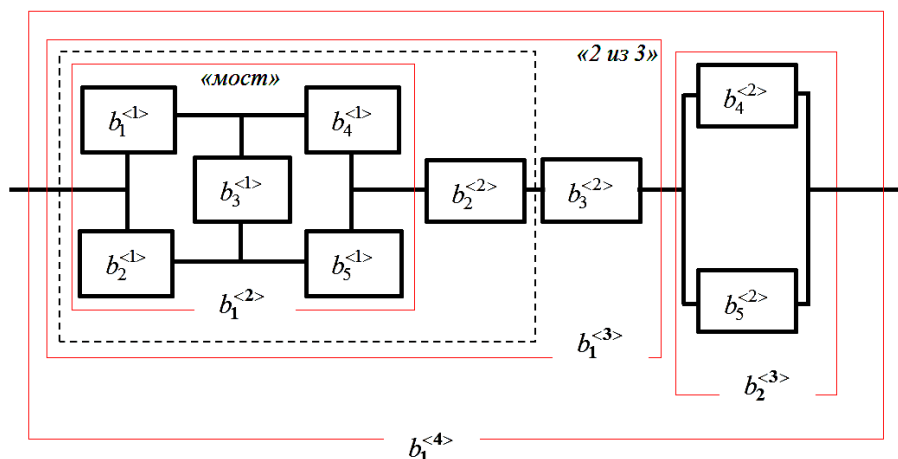


Рис. 1. Пример фрагмента структурно сложного технического комплекса

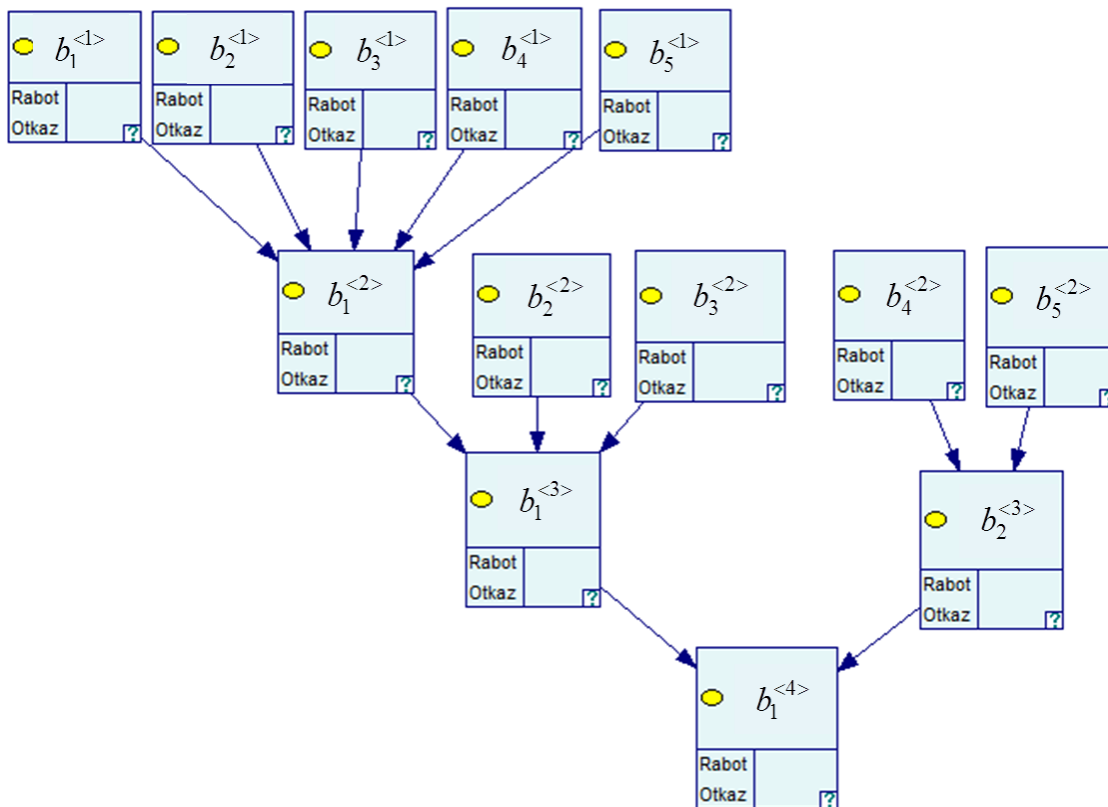


Рис. 2. Байесовская сеть доверия фрагмента структурно сложного технического комплекса

Таблица 1

Условные вероятности для блока 2-го уровня $b_1^{<2>}$

$b_1^{<1>}$	Rabot								Otkaz																								
$b_2^{<1>}$	Rabot				Otkaz				Rabot				Otkaz																				
$b_3^{<1>}$	Rabot		Otkaz		Rabot		Otkaz		Rabot		Otkaz		Rabot		Otkaz																		
$b_4^{<1>}$	Rabot	Otkaz	Rabot	Otkaz	Rabot	Otkaz	Rabot	Otkaz	Rabot	Otkaz	Rabot	Otkaz	Rabot	Otkaz	Rabot	Otkaz																	
$b_5^{<1>}$	Rabot	Otkaz	Rabot	Otkaz	Rabot	Otkaz	Rabot	Otkaz	Rabot	Otkaz	Rabot	Otkaz	Rabot	Otkaz	Rabot	Otkaz	Rabot	Otkaz															
$b_1^{<2>}$	Rabot	1	1	1	0	1	1	1	0	1	1	1	0	1	1	0	0	1	1	1	0	1	0	1	0	0	0	0	0	0	0	0	0
	Otkaz	0	0	0	1	0	0	0	1	0	0	0	1	0	0	1	1	0	0	1	0	1	1	0	1	1	1	1	1	1	1	1	1

Таблица 2

Условные вероятности для блока 3-го уровня $b_1^{<3>}$

$b_1^{<2>}$	Rabot								Otkaz								
$b_2^{<2>}$	Rabot				Otkaz				Rabot				Otkaz				
$b_3^{<2>}$	Rabot		Otkaz		Rabot		Otkaz		Rabot		Otkaz		Rabot		Otkaz		
$b_1^{<3>}$	Rabot	1		1		1		0		1		0		0		0	
	Otkaz	0		0		0		1		0		1		1		1	

Таблица 4

Условные вероятности для блока 4-го уровня $b_1^{<4>}$

$b_1^{<3>}$	Rabot				Otkaz				
$b_2^{<3>}$	Rabot		Otkaz		Rabot		Otkaz		
$b_1^{<4>}$	Rabot	1		0		0		0	
	Otkaz	0		1		1		1	

Таблица 3

Условные вероятности для блока 3-го уровня $b_2^{<3>}$

$b_4^{<2>}$	Rabot				Otkaz				
$b_5^{<2>}$	Rabot		Otkaz		Rabot		Otkaz		
$b_2^{<3>}$	Rabot	1		1		1		0	
	Otkaz	0		0		0		1	

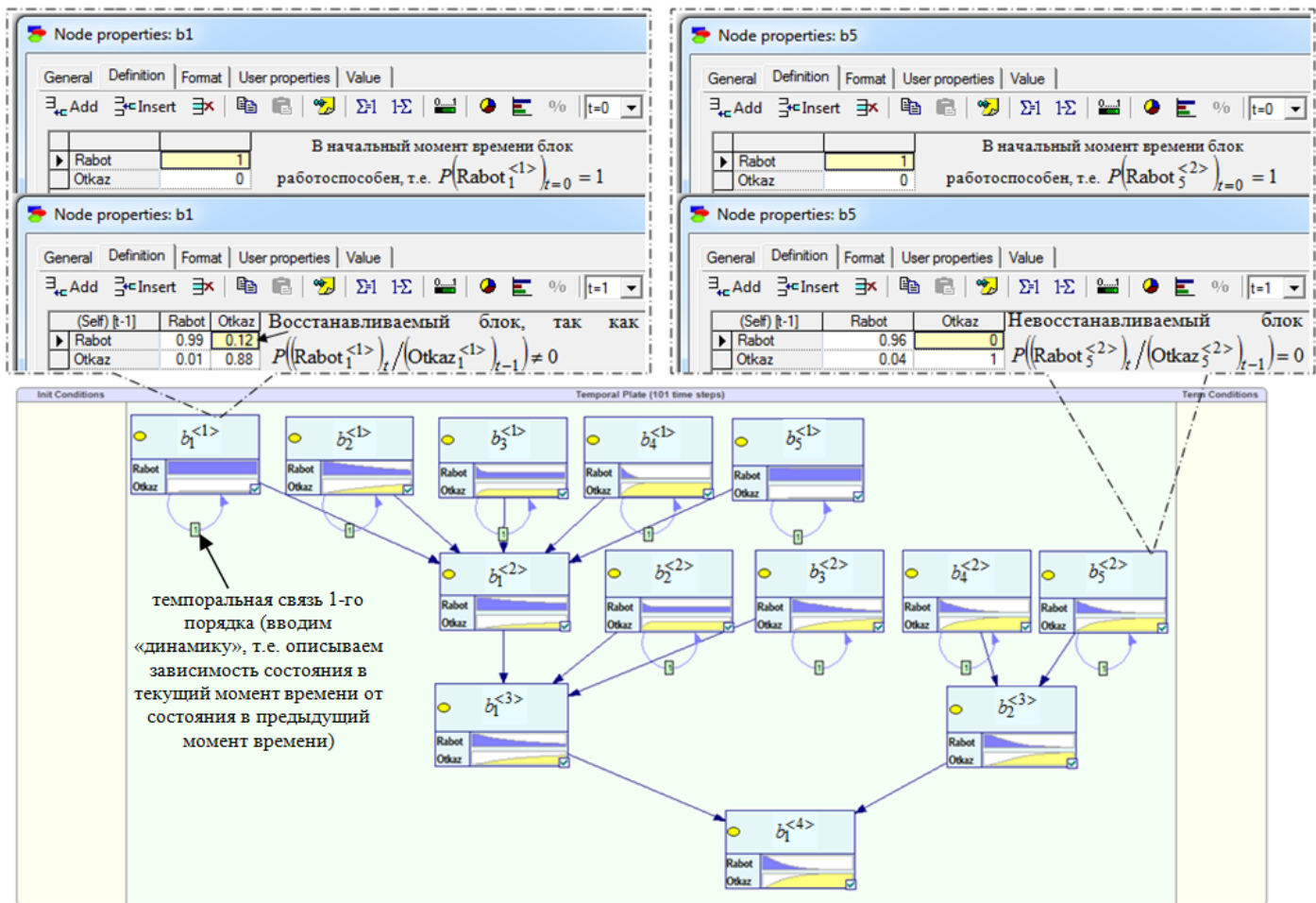


Рис. 3. Динамическая байесовская сеть в среде GeNIe

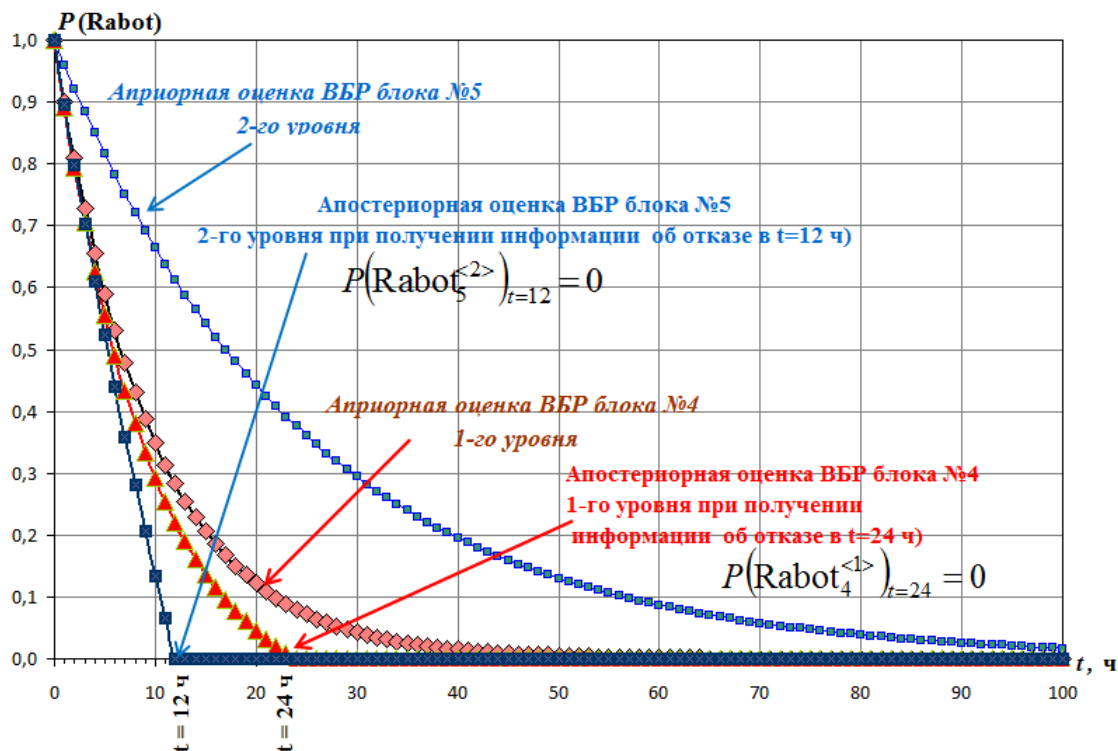


Рис. 4. Графики ВБР блоков по априорной информации и при поступлении информации об отказах

информации об отказах блоков в различные моменты времени: блок $b_5^{<2>}$ отказал в момент времени $t = 12$, а блок $b_4^{<1>}$ – $t = 24$. Аналогичные графики изменения ВБР могут быть получены для комплекса в целом по априорным данным и при поступлении новой информации.

Рассмотрим практическое применение разработанной модели на примере подсистемы управления руководителя работ (ПСУ РР), входящей в состав автоматизированной системы управления технологическими процессами (АСУ ТП). ПСУ РР выполняет функции координации и диспетчеризации работ.

Исходными данными для построения модели ПСУ РР, основанной на динамической байесовской сети доверия, являются:

- структурные схемы ПСУ РР;
- описание логики функционирования и принципов работы ПСУ РР;
- критерий отказа системы (задается в ТЗ или разработчиком);
- значения среднего времени безотказной работы или интенсивности потока отказов для блоков нижнего уровня детализации (элементов, модулей, плат), полученные из сертификатов и справочников.

Допущения и ограничения при моделировании:

- блоки системы могут находиться только в двух состояниях: отказа или работоспособности;
- случайные моменты отказов блоков нижнего уровня детализации имеют экспоненциальный закон распределения;
- при расчетах предполагается, что значения вероятностей безотказной работы кабельных соединений, соединительных внешних разъемов равны единице;
- к началу работы система проходит полный контроль, и ВБР системы (и ее элементов) равна единице в начальный момент времени;
- время полагается дискретным с шагом дискретизации один час (при необходимости можно изменить шаг дискретизации).

Построим модель для расчета ВБР ПСУ РР в среде GeNIe. Для блоков нижнего уровня детализации ПСУ РР $\forall b_i^{<j>} : \neg(\exists b_i^{<j-1>} = \text{parent}(b_i^{<j>}))$ (т. е. не имеющих родительских переменных) указывается (рис. 5):

- ВБР в начальный момент времени –

$$P(\text{Rabot}_i^{<j>})_{t=0} = P_i^{<j>} (t = 0) = 1;$$

- вероятность отказа в начальный момент времени –

$$P(\text{Otkaz}_i^{<j>})_{t=0} = Q_i^{<j>} (t = 0) = 1 - P_i^{<j>} (t = 0) = 0;$$

- вероятность того, что блок работоспособен в текущий момент времени t при условии, что блок был работоспособен в предыдущий момент $(t - \Delta t)$ –

$$P((\text{Rabot}_i^{<j>})_t / (\text{Rabot}_i^{<j>})_{t-\Delta t}) = \frac{e^{-\lambda_i^{<j>} t}}{e^{-\lambda_i^{<j>} (t-\Delta t)}} = e^{-\lambda_i^{<j>} \cdot \Delta t}$$

(в предположении экспоненциального закона распределения наработки на отказ). Так как время принято дискретным с шагом дискретизации один час, т. е. $\Delta t = 1$ ч,

$$P((\text{Rabot}_i^{<j>})_t / (\text{Rabot}_i^{<j>})_{t-1}) = \frac{e^{-\lambda_i^{<j>} t}}{e^{-\lambda_i^{<j>} (t-1)}} = e^{-\lambda_i^{<j>}}$$

(при необходимости шаг дискретизации можно изменить, если потребуется более высокая/низкая частота контроля ВБР комплекса);

- вероятность отказа блока в текущий момент времени при условии, что блок был работоспособен в предыдущий момент –

$$P((\text{Otkaz}_i^{<j>})_t / (\text{Rabot}_i^{<j>})_{t-1}) = 1 - e^{-\lambda_i^{<j>}};$$

- вероятность того, что блок работоспособен в текущий момент времени, при условии, что блок отказал в предыдущий момент –

$$P((\text{Rabot}_i^{<j>})_t / (\text{Otkaz}_i^{<j>})_{t-1}) = 0,$$

если блок не восстанавливается, или с учетом восстановления –

$$P((\text{Rabot}_i^{<j>})_t / (\text{Otkaz}_i^{<j>})_{t-1}) = 1 - e^{-\mu_i^{<j>}},$$

где $\mu_i^{<j>}$ – интенсивность восстановления i -го блока j -го иерархического уровня;

- вероятность отказа блока в текущий момент времени при условии, что блок отказал в предыдущий момент –

$$P((\text{Otkaz}_i^{<j>})_t / (\text{Otkaz}_i^{<j>})_{t-1}) = 1,$$

если блок не восстанавливается, или с учетом восстановления –

$$P((\text{Otkaz}_i^{<j>})_t / (\text{Otkaz}_i^{<j>})_{t-1}) = e^{-\mu_i^{<j>}}.$$

Для блоков, имеющих в своем составе блоки более низкого уровня детализации, указываются условные вероятности с учетом структурно-логической схемы надежности. Из структурно-логической схемы надежности «Канала № 1 питания контроллера» (рис. 6 а) видно, что «Источник питания № 1» и «Канал № 1 устройства ввода питания (УВП)» соединены последовательно с точки зрения надежности, т. е. отказ любого из них приведет к отказу «Канала № 1 питания контроллера». Параметры «Канала № 1 питания контроллера» представлены на рис. 6 б.

Для более компактного представления модели ПСУ РР были использованы конструкции среды GeNIe – субмодели («submodels»), которые позволяют скрыть блоки более низкого уровня для облегчения восприятия модели в целом и навигации по составным частям, но при необходимости можно раскрыть более подробный состав интересующей части модели. Использование субмоделей позволяет улучшить пользовательский интерфейс. На рис. 7 в свернутом виде представлена динамическая байесовская сеть доверия ПСУ РР.

Динамическая байесовская сеть доверия ПСУ РР состоит из:

- 347 блоков с 694 состояниями;
- 660 связей;
- 117 880 параметров;
- 78 субмоделей.

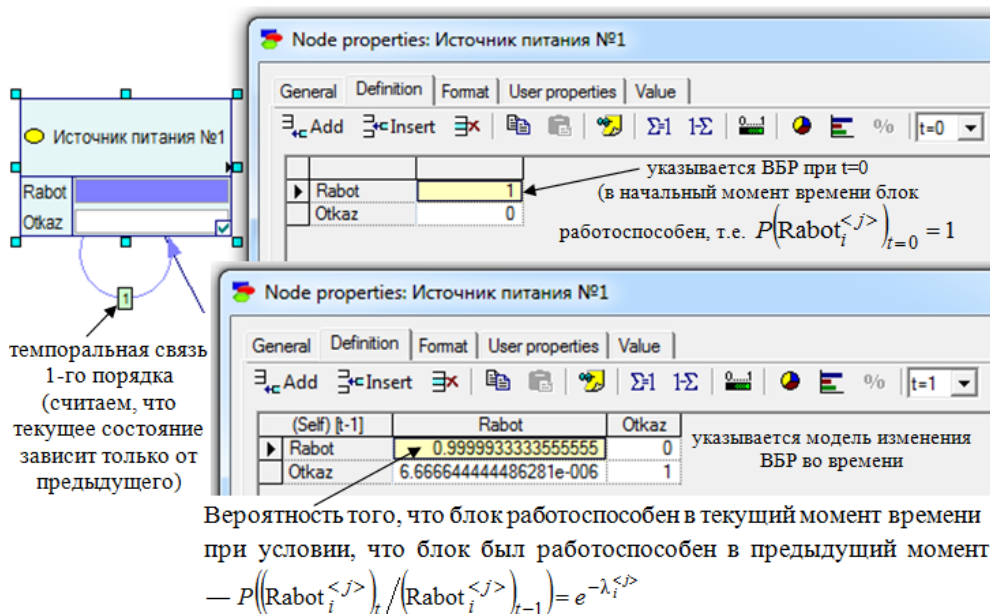


Рис. 5. Параметры блока нижнего уровня детализации

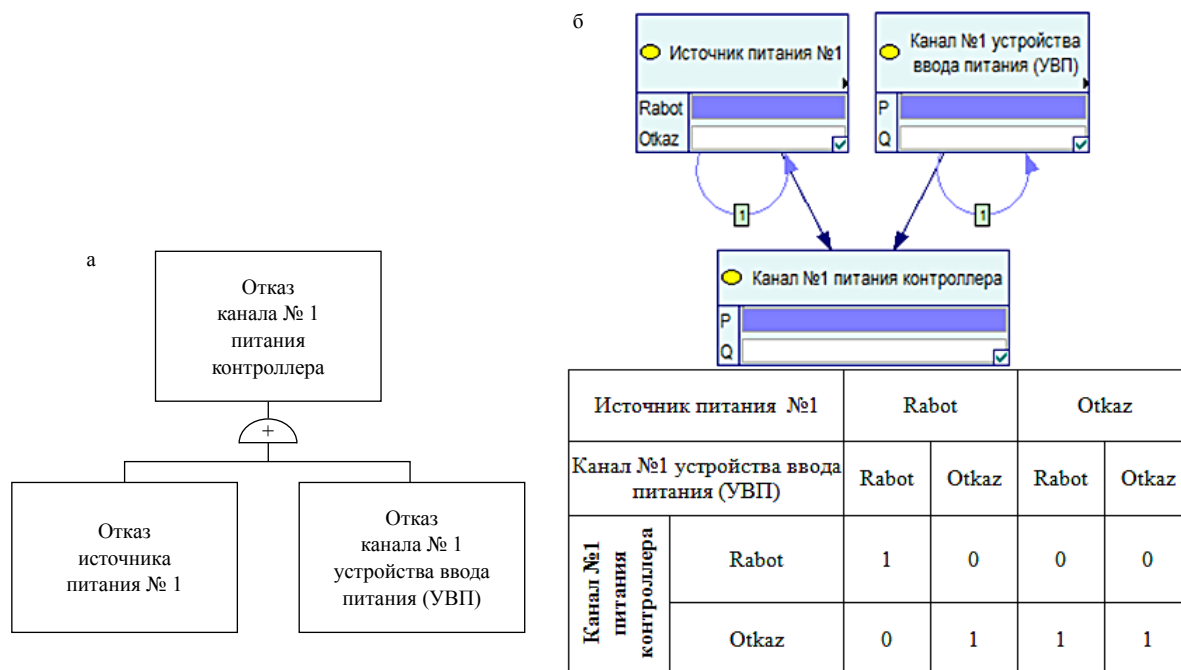


Рис. 6. Дерево отказов и БСД блока «Канал питания контроллера»: а) дерево отказов канала питания контроллера; б) параметры блока «Канал питания контроллера»

ПРИМЕР РАБОТЫ ДИНАМИЧЕСКОЙ БАЙЕСОВСКОЙ СЕТИ ДОВЕРИЯ

Рассмотрим пример, когда в различные моменты времени поступает информация об отказах блоков ПСУ РР (табл. 5):

С учетом данной информации сделан апостериорный вывод в динамической байесовской сети доверия и получена ВБР ПСУ РР. На рис. 8 изображены графики изменения ВБР по априорным данным и при поступлении новой информации об отказах блоков в различные моменты времени.

Согласно ТЗ, при выполнении работ одного расчетного цикла нормативное значение вероятности безотказной работы АСУ ТП и ПСУ РР должно быть не ниже 0,995.

Анализ графиков на рис. 8 позволяет сделать следующие выводы:

- значение ВБР ПСУ РР, полученное по априорной информации о надежности элементов, на момент времени $t = 51$ ч равно 0,995 и соответствует требованиям ТЗ;
- значение ВБР ПСУ РР, полученное с учетом информации об отказах блоков ПСУ РР, на момент времени $t = 20$ ч (табл. 5) равно 0,998. Прогнозное значение ВБР на момент времени $t = 51$ ч равно 0,9954, что соответствует требованиям ТЗ;
- при дальнейшем поступлении информации об отказах элементов ПСУ РР (табл. 5) происходит обновление вероят-

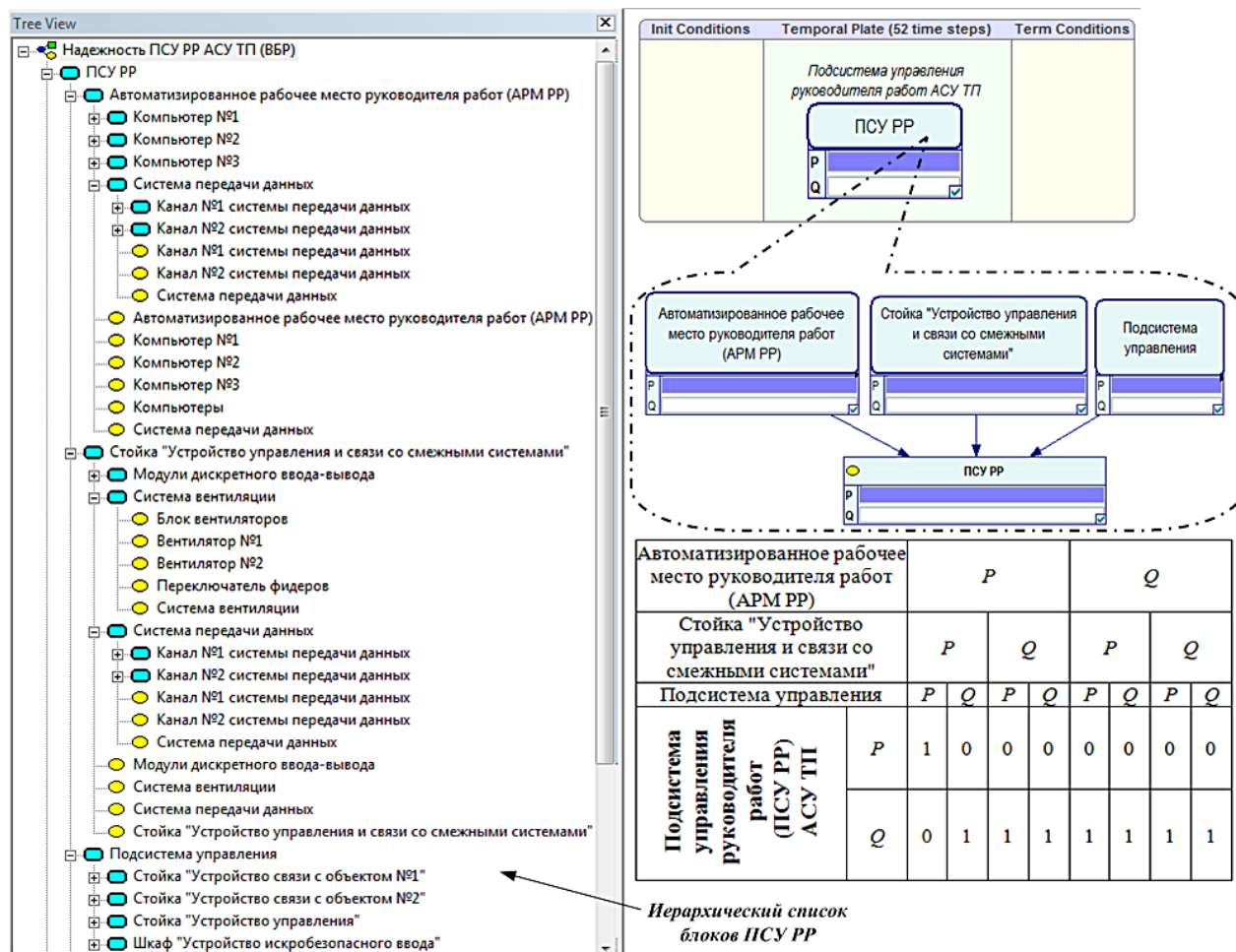


Рис. 7. Динамическая байесовская сеть доверия ПСУ PP

Таблица 5
Поступающая информация об отказах блоков ПСУ PP

Блок	Момент отказа, ч
Блок № 2 комплекта № 1 стойки «Устройство связи с объектом № 2»	5
Компьютер № 1 автоматизированного рабочего места руководителя работ (АРМ PP)	8
Вентилятор № 1 системы вентиляции стойки «Устройство управления»	9
Канал № 1 питания контроллера № 2 стойки «Устройство управления»	10
Блок № 1 комплекта № 1 стойки «Устройство связи с объектом № 1»	10
Модуль № 1 дискретного ввода-вывода	12
Вентилятор № 2 системы вентиляции	14
Платы интерфейсные № 1, № 3	16
Канал № 2 системы передачи данных	17
Блок № 2 комплекта № 2 стойки «Устройство связи с объектом № 2»	17
Питание системы передачи данных	20
Модуль № 2 дискретного ввода-вывода	22
Компьютер № 3 АРМ PP	25
Контроллеры № 1, № 3	29

ностей в модели, и значение ВБР на момент времени $t = 29$ ч, будет составляет 0,993. Прогнозное значение ВБР на момент времени $t = 51$ ч равно 0,987, что ниже значения ВБР, установленного требованиями ТЗ. Следовательно, необходимо восстановить отказавшую часть оборудования.

Таким образом, разработанная модель позволяет оценить ВБР ПСУ PP при поступлении новой информации о ВБР составных элементов в текущий и будущий моменты времени. Прогнозные значения ВБР целесообразно использовать для выбора стратегии восстановления системы.

Отметим, что при поступлении новой информации корректируется значение ВБР не только в текущий и будущий, но и в прошлый момент времени (рис. 8), т. е. решается задача ретроспективного анализа. Ретроспективный анализ обеспечивает лучшую оценку информации о состоянии, которая была доступна к тому времени (поскольку включает больше свидетельств).

РЕТРОСПЕКТИВНЫЙ АНАЛИЗ

Рассмотрим пример использования разработанной модели для ретроспективного анализа.

При расследовании аварийных ситуаций, анализе происшествий на сложных технических комплексах главной целью является определение причины аварии (происшествия). Особую актуальность это принимает, когда анализ произошедшего события, выявление первопричины аварии

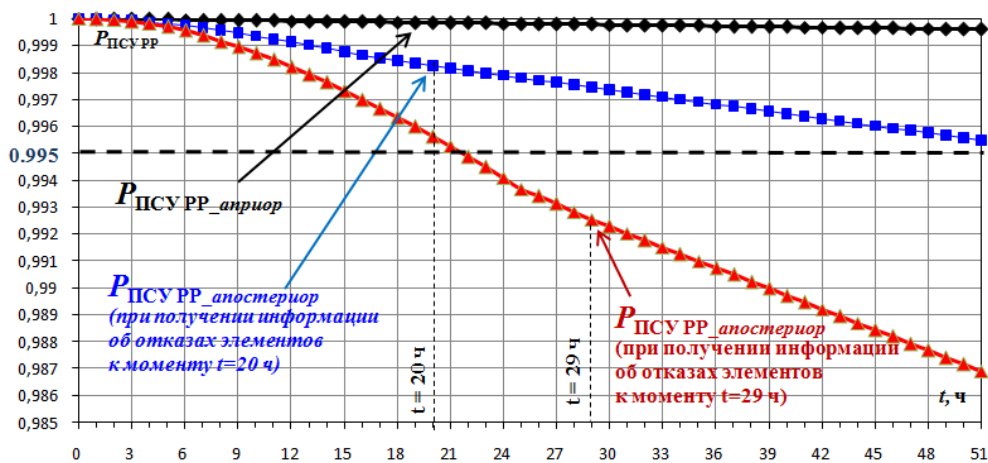


Рис. 8. Вероятность безотказной работы ПСУ РР с учетом информации об отказах

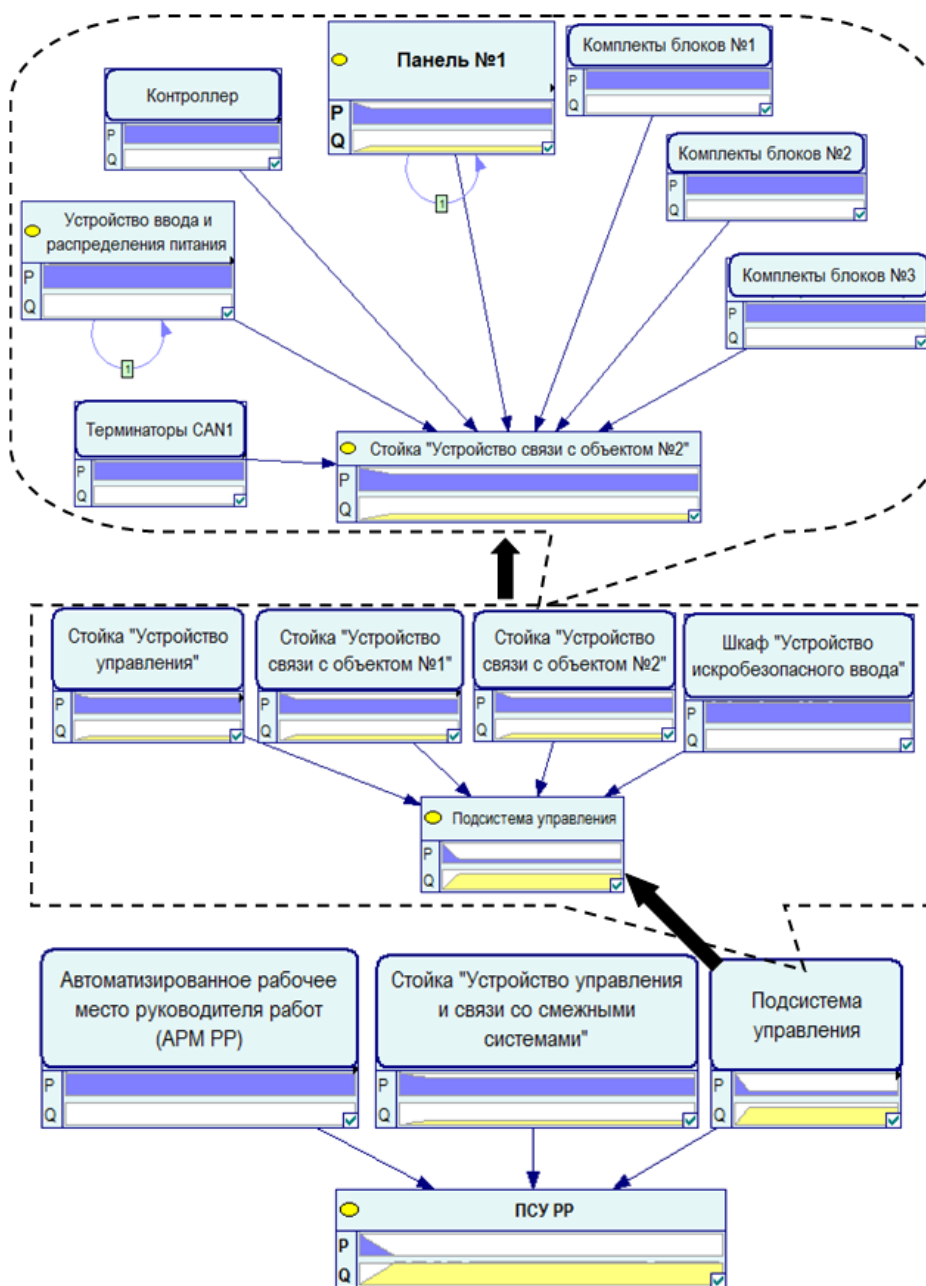


Рис. 9. Схема поиска причины отказа ПСУ РР

и принятие соответствующих мер позволяет исключить или уменьшить вероятность повторения неблагоприятных событий, а значит, выполнить поставленную задачу, уменьшить экономические потери и т. д.

Допустим, что в момент времени $t = 5$ ч произошел отказ ПСУ РР. Раскрывая последовательно блоки более низкого уровня, можно определить блок, вероятность отказа которого наибольшая. Из рис. 9 видно, что блок «Панель № 1» стойки «Устройство связи с объектом № 2» имеет наибольшую вероятность отказа по сравнению с другими блоками, следовательно, необходимо проверить блок «Панель № 1» в первую очередь и при необходимости заменить его.

Далее допустим, что блок «Панель № 1» проверен и сделан вывод, что он работоспособен, тогда в динамическую байесовскую сеть поступает данная информация, происходит обновление вероятностей, и поиск следующего блока с наибольшей вероятностью отказа продолжается до того момента, пока не будет найден отказавший блок. Описанный алгоритм поиска причины отказа позволяет оптимизировать время поиска неисправности, а следовательно, уменьшить время восстановления работоспособного состояния объекта, что крайне важно в условиях жестких временных ограничений выполнения технологического графика.

ЗАКЛЮЧЕНИЕ

Таким образом, в статье предлагается модель динамической байесовской сети доверия для оценки ВБР сложных многоуровневых комплексов. Данная модель предоставляет возможность руководителю работ (или оператору):

- контролировать ВБР комплекса при поступлении новой информации об отказах элементов. В БСД происходит обновление значения ВБР комплекса, учитывая поступающую новую информацию о надежности блоков;
- прогнозировать значение ВБР с целью выбора стратегии восстановления комплекса;
- поддерживать принятие решений при поиске причины отказа комплекса.

ЛИТЕРАТУРА

1. Аверьянов А. В. Определение пределов аппаратной избыточности информационных управляющих систем / А. В. Аверьянов, А. М. Барановский, К. А. Эсаулов // Изв. вузов. Приборостроение. – 2014. – Т. 57, № 3. – С. 23-25.
2. Гуменюк В. М. Надежность и диагностика электротехнических систем: учеб. пособие для вузов / В. М. Гуменюк. – Владивосток: Изд-во Дальневост. гос. техн. ун-та, 2010. – 218 с.
3. Гусеница Я. Н. О некоторых особенностях надежности программного обеспечения автоматизированных систем управления войсками / Я. Н. Гусеница, Ю. Л. Кругляк, Д. О. Пе-

трич // Тр. Воен.-космич. акад. им. А.Ф. Можайского. – СПб.: ВКА им. А.Ф. Можайского. – 2013. – Вып. 638. – С. 31-36.

4. Дорожко И. В. Методика синтеза оптимальных стратегий диагностирования автоматизированных систем управления сложными техническими объектами с использованием априорной информации / И. В. Дорожко, Н. А. Осипов // Тр. СПИИРАН. – 2012. – Вып. 1 (20). – С. 165-185.

5. Кочанов И. А. Прогнозирование интенсивности появления ошибок в программном обеспечении вычислительных систем реального времени на период эксплуатации / И. А. Кочанов, А. В. Кубуша // Тр. Воен.-космич. акад. им. А. Ф. Можайского. – СПб.: ВКА им. А.Ф. Можайского. – 2012. – Вып. 637-2. – С. 92-97.

6. Кузнецов А. Б. Методика диагностирования автоматизированных систем управления сложными техническими объектами с использованием априорной информации / А. Б. Кузнецов, Н. А. Осипов, И. В. Дорожко // Изв. вузов. Приборостроение. – 2013. – Т. 56, № 1. – С. 18-26.

7. Надежность в технике. Основные понятия. Термины и определения: ГОСТ 27.002-89. – Введ. 01.07.1990. – М.: Изд-во стандартов, 1990. – 39 с.

8. Надежность в технике. Термины и определения: ГОСТ Р 27.002-2009. – Введ. 01.01.2011. – М.: Стандартинформ, 2011. – 32 с.

9. Половко А. М. Основы теории надежности / А. М. Половко, С. В. Гуров. – СПб.: БХВ-Петербург, 2006. – 704 с.

10. Привалов А. Е. Онтологический подход к формированию ситуационной осведомленности при управлении эксплуатацией ракетно-космических комплексов / А. Е. Привалов, И. А. Кочанов, П. Ю. Бугайченко // Тр. Воен.-космич. акад. им. А.Ф. Можайского. – СПб.: ВКА им. А.Ф. Можайского. – 2013. – Вып. 639. – С. 93-98.

11. Рассел С. Искусственный интеллект: современный подход / С. Рассел, П. Норвиг ; пер. с англ. – 2-е изд. – М.: Вильямс, 2006. – 1408 с.

12. Рябинин И. А. Автоматизированное моделирование надежности структурно-сложных систем из элементов с тремя несовместными состояниями / И. А. Рябинин, А. В. Струков // Тр. СПИИРАН. – 2014. – Вып. 3 (34). – С. 89-111.

13. Тулупьев А. Л. Байесовские сети доверия: логико-вероятностный вывод в ациклических направленных графах / А. Л. Тулупьев, А. В. Сироткин, С. И. Николенко. – СПб.: Изд-во С.-Петерб. ун-та, 2009. – 400 с.

14. Тулупьев А. Л. Байесовские сети: логико-вероятностный подход / А. Л. Тулупьев, С. И. Николенко, А. В. Сироткин. – СПб.: Наука, 2006. – 607 с.

15. Jensen F.V., Nielsen T.D. Bayesian Networks and Decision Graphs. Berlin, Springer, 2007. 457 p.

16. Nilsson N.J. Probabilistic Logic Revisited. Artif. Intell. Amsterdam, Elsevier Sci. Publ. B.V., 1993, Vol. 59, pp. 31-36.

17. Pearl J. Probabilistic Reasoning in Intelligent Systems. NY, Morgan Kaufmann, 1988. 552 p.

Estimation to Reliability of Structural Complex Technical Systems by Using Bayesian Networks Belief Models in the Environment of GeNIe

Dorozhko I. V., Tarasov A. G.

Military Space academy named after A. F. Mozhaisky
St. Petersburg, Russia
Doroghko-Igor@yandex.ru

Baranovsky A. M.

Petersburg State Transport University
St. Petersburg, Russia
bamvka@mail.ru

Abstract. The article proposes a new approach to the estimation of parameters of reliability of complex hierarchical systems. Approach based on dynamic Bayesian Networks. The developed model allows to take into account the receipt of new information on the reliability of complex elements, which gives the opportunity to predict the probability of non-failure operation technical system, adjust the maintenance frequency and complex recovery strategy.

Keywords: reliability, survival function, Bayesian network.

REFERENCES

1. Averyanov A. V., Baranovsky A. M., Esaulov K. A. Determination of Hardware Redundancy Limits of Information Control Systems [Opredelenie predelov apparatnoy izbytochnosti informatsionnykh upravlyayushchikh sistem]. *Izvestiya vysshikh uchebnykh zavedeniy. Priborostroenie [Proc. Higher Educ. Establishments. Instrumentation]*, 2014, vol. 57, no. 3, pp. 23-25.
2. Gumenyuk V. M. Nadezhnost i diagnostika elektrotehnicheskikh system: uchebnoye posobiye dlya vuzov [Reliability and Diagnostics of Electrical Systems: Study Guide for Higher Educational Establishments]. Vladivostok, Publ. Far Eastern Univ., 2010. 218 p.
3. Gusenitsa Ya. N., Kruglyak Yu. L., Petrich D. O. Some Features of the Software Reliability of Automated Systems of Command and Control [O nekotorykh osobennostyakh nadezhnosti programmogo obespecheniya avtomatizirovannykh system upravleniya voyskami]. *Trudy Voenno-kosmicheskoy akademii imeni A. F. Mozhayskogo [Proc. Mil. Space acad. named after A. F. Mozhaisky]*, 2013, no. 638, pp. 31-36.
4. Dorozhko I. V., Osipov N. A. Technique of Synthesis of Optimal Strategies for Diagnosis of the Automated Control Systems of Complex Technical Objects with the use of Aprioristic Information [Metodika sinteza optimalnykh strategiy diagnostirovaniya avtomatizirovannykh system upravleniya slozhnymi tekhnicheskimi obektami s ispolzovaniem apriornoy informatsii]. *Trudy SPIIRAN [Proc. SPIIRAS]*, 2012, no. 1 (20), pp. 165-185.
5. Kochanov I. A., Kubusha A. V. Forecasting to Intensities of the Appearance Mistake in Software Ensuring the Computing Systems of Real Time for Period of the Usages [Prognozirovanie intensivnosti poyavleniya oshibok v programnom obespechenii vychislitelnykh system realnogo vremeni na period ekspluatatsii]. *Trudy Voenno-kosmicheskoy akademii imeni A. F. Mozhayskogo [Proc. Mil. Space acad. named after A. F. Mozhaisky]*, 2012, no. 637-2, pp. 92-97.
6. Kuznetsov A. B., Osipov N. A., Dorozhko I. V. Technique of Diagnosing of the Automated Control Systems of Complex Technical Objects with use Aprioristic Information [Metodika diagnostirovaniya avtomatizirovannykh system upravleniya slozhnymi tekhnicheskimi obektami s ispolzovaniem apriornoy informatsii]. *Izvestiya vysshikh uchebnykh zavedeniy. Priborostroenie [Proc. Higher Educ. Establishments. Instrumentation]*, 2013, vol. 56, no. 1, pp. 18-26.
7. Nadezhnost v tekhnike. Osnovnye ponyatiya. Terminy i opredeleniya: GOST 27.002-89 [Reliability in Technics. Basic Concepts. Terms and Definitions: GOST 27.002-89]. Moscow, Izdatelstvo standartov, 1990. 39 p.
8. Nadezhnost v tekhnike. Terminy i opredeleniya: GOST R 27.002-2009 [Reliability in Technics. Terms and Definitions: GOST 27.002-2009]. Moscow, Standartinform, 2011. 32 p.
9. Polovko A. M., Gurov S. V. Osnovy teorii nadezhnosti [Fundamentals of theory of reliability]. St. Petersburg, BVKh-Petersburg, 2006. 704 p.
10. Ptivalov A. E., Kochanov I. A., Bugaychenko P. Yu. The ontological approach to forming situation awareness during the management of the operation RKK [Ontologicheskii podkhod k formirovaniyu situatsionnoy osvedomlennosti pri upravlenii ekspluatatsiyey raketno-kosmicheskikh kompleksov]. *Trudy Voenno-kosmicheskoy akademii imeni A. F. Mozhayskogo [Proc. Mil. Space acad. named after A. F. Mozhaisky]*, 2013, no. 639, pp. 93-98.
11. Russel S., Norvig P. Iskusstvennyy intellekt: sovremennyy podkhod [Artificial Intelligence. A Modern Approach]. Moscow, Willams, 2006. 1408 p.
12. Ryabinin I. A., Strukov A. V. Automated Modeling of Reliability of Structurally Complex Systems Consisting of Elements with Three Disjoint States [Avtomatizirovannoe modelirovanie nadezhnosti strukturno-slozhnykh system iz elementov s tremya nesovmestnymi sostoyaniyami]. *Trudy SPIIRAN [Proc. SPIIRAS]*, 2014, no. 3 (34), pp. 89-111.
13. Tulupyev A. L., Sirotkin A. V., Nikolenko S. I. *Bayesovskie set idoveriya: logiko-veroyatnostnyy vyvod v atsiklicheskikh napravlennykh grafakh* [Bayesian Networks: logico-probabilistic inference in acyclic directed graphs], SPb, Publ. St. Petersburg Univ., 2009, 400 p.
14. Tulupyev A. L., Nikolenko S. I., Sirotkin A. V. *Bayesovskieseti: logiko-veroyatnostnyy podkhod* [Bayesian Networks: Logico-Probabilistic Approach]. St. Petersburg, Nauka, 2006. 607 p.
15. Jensen F. V., Nielsen T. D. Bayesian Networks and Decision Graphs. Berlin, Springer, 2007. 457 p.
16. Nilsson N. J. Probabilistic Logic Revisited. *Artif. Intell.* Amsterdam, Elsevier Sci. Publ. B. V., 1993, Vol. 59, pp. 31-36.
17. Pearl J. Probabilistic Reasoning in Intelligent Systems. NY, Morgan Kaufmann, 1988. 552 p.

Модификация метода RPROP для решения задач распознавания образов

Благовещенская Е. А., Зуев Д. В.
Петербургский государственный университет
путей сообщения Императора Александра I
Санкт-Петербург, Россия
zuevdv@gmail.com, kblag2002@yahoo.com

Аннотация. Алгоритмы, применяемые для обучения персептронов и для обучения специальных архитектур нейронных сетей часто содержат параметры, которые пользователь выбирает на основе опыта и известных рекомендаций других авторов («вручную»). От выбора этих параметров существенно зависит скорость достижения алгоритмом результата. Сегодня нет общей методики автоматизированного определения таких параметров. В работе описывается алгоритм модификации известного метода RPROP, позволяющий повысить скорость его сходимости при распознавании образов.

Ключевые слова: нейронные сети, RPROP, распознавание образов.

Одним из ключевых критериев оценки эффективности любого метода обработки данных является скорость его работы [1]. В данной статье предложен эффективный метод повышения скорости работы эвристических алгоритмов оптимизации [2] для задач минимизации функции ошибки $E(\{w_i\}): R^m \rightarrow R_+$ нейронных сетей [3, 9] при использовании их в решении задачи распознавания образов.

Рассмотрим сначала общую ситуацию. Особенностью большинства эвристических алгоритмов оптимизации является наличие параметров, выбираемых вручную. Такими параметрами часто являются константы или функции, характеризующие изменение «скорости шага» [3, 9]. В описаниях алгоритмов, как правило, даны либо конкретные рекомендуемые значения параметров, либо общие ограничения на их выбор.

В настоящей работе предложен общий подход к выбору параметров алгоритмов. В качестве объекта для применения предлагаемого метода выбран алгоритм оптимизации первого порядка RPROP [2, 10]. Результаты статистических исследований [5] показывают высокую скорость сходимости алгоритма RPROP в сравнении с другими градиентными алгоритмами. При этом отметим, что данный метод может быть перенесен на другие алгоритмы оптимизации без существенных изменений.

Приведем основные шаги алгоритма RPROP [11] для минимизации функции $E(\{w_i\})$ ошибки нейронной сети, заданной на множестве весовых коэффициентов $\{w_i\}$.

Шаг 1

Для каждого весового коэффициента w_i произвольно выбираем число $\Delta_i(0) > 0$, формулируем условие остановки алгоритма (например, максимальное количество итераций

или максимальное значение функции ошибки) и выбираем параметры алгоритма $\eta^+ > 1$ и $0 < \eta^- < 1$.

Шаг 2

На t -й итерации для всех i вычисляем значения

$$\Delta_i(t) = \begin{cases} (\eta^+) * \Delta_i(t-1), & \text{if } \frac{\partial E(t-1)}{\partial w_i} \frac{\partial E(t)}{\partial w_i} > 0; \\ (\eta^-) * \Delta_i(t-1), & \text{if } \frac{\partial E(t-1)}{\partial w_i} \frac{\partial E(t)}{\partial w_i} < 0; \\ \Delta_i(t-1), & \text{if } \frac{\partial E(t-1)}{\partial w_i} \frac{\partial E(t)}{\partial w_i} = 0, \end{cases} \quad (1)$$

а затем

$$\Delta w_i(t) = \begin{cases} -\Delta_i(t), & \text{if } \frac{\partial E(t)}{\partial w_i} > 0; \\ +\Delta_i(t), & \text{if } \frac{\partial E(t)}{\partial w_i} < 0; \\ 0, & \text{if } \frac{\partial E(t)}{\partial w_i} = 0. \end{cases} \quad (2)$$

Корректируем значения весовых коэффициентов по формуле

$$w_i(t+1) = w_i(t) + \Delta w_i(t).$$

Шаг 3

Проверяем условие остановки алгоритма. Если условие остановки не выполнено, переходим к шагу 2, в противном случае заканчиваем действия по алгоритму.

Особенностью RPROP является наличие константных величин η^+ и η^- . Конкретные значения этих величин существенно влияют как на сходимость, так и на скорость сходимости алгоритма.

Введем в рассмотрение вещественную непрерывную функцию $F(t, a_0, a_1, \dots, a_{n-1})$ от переменной t , зависящую от n параметров a_0, a_1, \dots, a_{n-1} . Если функция определена на всех возможных наборах вещественных параметров a_0, a_1, \dots, a_{n-1} , будем называть ее подстановкой.

ОПИСАНИЕ МЕТОДА

Заменим в (1) величину η^+ подстановкой $F(t, a_0, a_1, \dots, a_{n-1})$, где t – номер итерации алгоритма. Введём в рассмотрение отображение

$$A(X) : R^n \supset G \rightarrow N,$$

ставящее в соответствие набору параметров a_0, a_1, \dots, a_{n-1} количество итераций, за которые RPROP сходится при данных параметрах. Совокупность параметров a_0, a_1, \dots, a_{n-1} будем интерпретировать как точку $X = (a_0, a_1, \dots, a_{n-1}) \in R^n$.

Пусть множество $G \subset R^n$ состоит из тех точек X , на которых RPROP сходится. Назовём его пространством сходимости отображения A . Задача повышения скорости сходимости RPROP сводится к задаче минимизации отображения A на своём пространстве сходимости.

Для минимизации A предлагается следующая модификация алгоритма многогранника [4]:

Шаг 1. Выбираем точку $p(0)$ из пространства сходимости и некоторое значение параметра r . Формулируем условие остановки алгоритма. В качестве условия остановки может выступать, например, максимальное число итераций.

Шаг 2. На t -й итерации выбираем несколько точек, отстоящих на расстояние r от $p(t-1)$ (в смысле обычного расстояния в R^n). Вычисляем для каждой из получившихся точек значение отображения A , принимаем в качестве $p(t)$ ту точку, на которой значение отображения A меньше, чем $A(p(t-1))$.

Шаг 3. Проверяем условие остановки. Если оно выполнено, тогда заканчиваем действия по алгоритму. В противном случае переходим к шагу 2.

Геометрическая интерпретация алгоритма приведена на рис. 1.

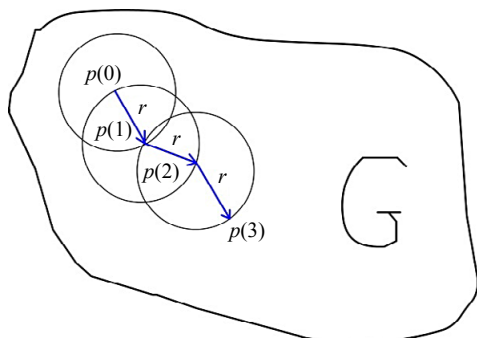


Рис. 1. Геометрическая интерпретация алгоритма

Стрелками иллюстрируется движение точки в «оптимальном» направлении.

Замечание. Результат применения алгоритма зависит от выбора начальной точки $p(0)$ и параметра r . Априори, особенно на практике, довольно сложно сделать выводы о пространстве сходимости отображения A и гарантировать, что хотя бы одна точка на окружности попадает в пространство сходимости. Для устранения этого недостатка, а также для повышения эффективности алгоритма предлагается рассматривать его вариации с динамически меняющимися радиусами, несколькими начальными точками и пр.

СХЕМА ПРИМЕНЕНИЯ МЕТОДА ДЛЯ МИНИМИЗАЦИИ ФУНКЦИИ ОШИБКИ $E(\{w_i\}) : R^m \rightarrow R_+$ НЕЙРОННОЙ СЕТИ

Часто для решения различных практических задач нейросетевым способом (распознавания образов, предсказания курса валют и др.) пользуются следующей схемой:

- 1) выбрать структуру нейронной сети;
- 2) выбрать алгоритм обучения и зафиксировать его параметры;

3) для множества «подобных» обучающих выборок применить выбранную нейронную сеть, выбранный алгоритм обучения и фиксированные параметры алгоритма.

Теперь подробнее поясним применение описанного алгоритма к решению задач распознавания образов. Отметим, что понятие подобности обучающих выборок определяется конкретной решаемой задачей. Для задачи распознавания рукописных букв подобными целесообразно считать обучающие выборки из одних и тех же букв, написанных разными людьми с похожим почерком.

Предложенный здесь метод показал свою эффективность именно в ситуации применения сети фиксированной структуры с фиксированным алгоритмом обучения и фиксированными параметрами алгоритма к множеству обучающих выборок.

Схема применения метода следующая:

- 1) выбрать структуру нейронной сети;
- 2) выбрать алгоритм обучения;
- 3) выбрать одну обучающую выборку из множества подобных обучающих выборок;
- 4) оптимизировать алгоритм обучения описанным методом для выбранной обучающей выборки;
- 5) зафиксировать оптимизированный алгоритм;
- 6) применить оптимизированный алгоритм для остальных обучающих выборок.

ПРИМЕРЫ ПОДСТАНОВОК И РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ

Полиномиальная подстановка

Довольно эффективной оказывается подстановка

$$F(t, a_0, a_1, \dots, a_{n-1}) = a_0 + a_1 t + \dots + a_{n-1} t^{n-1}.$$

Такую подстановку назовем полиномиальной. Число n для полиномиальной подстановки определяется исходя из условий конкретной задачи или класса задач. Результаты статических исследований использования полиномиальной подстановки показали, что для различных конфигураций нейронных сетей предложенный метод находит модификации алгоритма RPROP, которые примерно в 3 раза превосходят по скорости сходимости «стандартный» RPROP, причем с возрастанием n результат меняется незначительно. Часть абсолютных значений результата эксперимента представлена в таблице.

Результаты эксперимента

Количество элементов обучающей выборки	Количество эпох метода Градиентного спуска	Количество эпох метода RPROP	Количество эпох модифицированного метода RPROP
1	1252 239	5	4
2	5097 374	8	7
3	5135 375	22	12
4	5225 938	254 035	122 041
4	6963 032	21	18
4	55 070 050	1388 763	583 096

Выбор первого приближения для полиномиальной подстановки может быть проведен следующим образом. Пусть для заданной обучающей выборки стандартный алгоритм RPROP с заданным параметром η^+ сходится. Если в качестве первого приближения в описанной модификации метода многогранника выбрать точку $p(0) = (\eta^+, 0, \dots, 0)$, то по построению алгоритма видно, что число итераций сходимости модифицированного RPROP, по крайней мере, не увеличится. На практике число итераций значительно уменьшается.

НЕЙРОСЕТЕВАЯ ПОДСТАНОВКА

Утверждение теоремы Цыбенко [8] о нейронных сетях заключается в следующем: любая непрерывная функция может быть сколь угодно точно приближена однослойной нейронной сетью с произвольной нелинейной функцией активации. Это утверждение дает мотивацию к выбору нейронной сети в качестве подстановки. Рассмотрим полносвязную нейронную сеть с одним скрытым слоем (рис. 2). Пусть сеть содержит один нейрон во входном слое, один нейрон в выходном слое и такое количество нейронов в скрытом слое, чтобы суммарное количество весовых коэффициентов сети было n .

Пусть функции активации нелинейны.

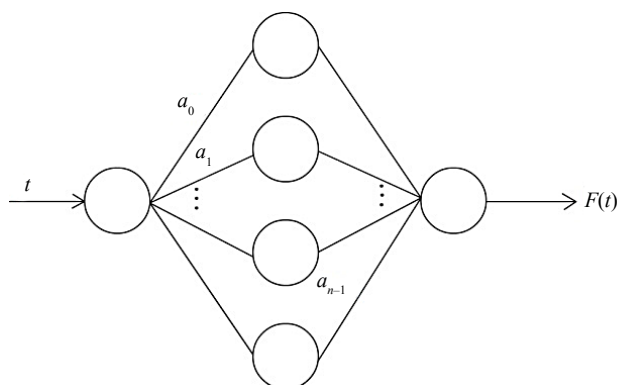


Рис. 2. Нейронная сеть с одним скрытым слоем

Такую сеть можно взять в качестве подстановки $F(t, a_0, a_1, \dots, a_{n-1})$, где набор $(a_0, a_1, \dots, a_{n-1})$ является набором весовых коэффициентов сети, а t – входной переменной. Для нейросетевой подстановки справедливо такое же замечание о выборе первого приближения, как и для полиномиальной подстановки, а именно: диапазон изменения значений подстановки определяется активационными функциями сети, таким образом, можно выбрать сеть с условием $F(t, 0, \dots, 0) = \eta^+$.

ЗАКЛЮЧЕНИЕ

Авторы разработали на языке C# для платформы Microsoft .NET программный комплекс, реализующий нейросетевую объектную модель и описанные алгоритмы. Прикладной задачей разработанного программного комплекса является задача распознавания образов. В качестве обучающей выборки образов для проведения эксперимента по определению эффективности предложенного в работе метода выступала база MNIST [6, 7]. Результаты работы программного комплекса показали, что применение полученных алгоритмов позволяет существенно экономить процессорное время вычислительной системы, что в несколько раз повышает эффективность её использования по сравнению с системами, использующими стандартный алгоритм RPROP для обучения нейронных сетей.

ЛИТЕРАТУРА

1. Благовещенская Е. А. О сходимости методов обучения нейронных сетей / Е. А. Благовещенская, Д. В. Зуев // Материалы XIX Междунар. конф. по вычислительной механике и современным прикладным программным системам (ВМСППС'2015), 24–31 мая 2015, Алушта. – М.: Изд-во МАИ, 2015. – С. 130-131.
2. Васильев А. Н. Нейросетевое моделирование. Принципы. Алгоритмы. Приложения / А. Н. Васильев, Д. А. Тархов. – СПб.: Изд-во Политехн. ун-та, 2009. – 528 с.
3. Круглов В. В. Нечеткая логика и искусственные нейронные сети / В. В. Круглов, М. И. Дли, Р. Ю. Голунов. – М.: Физматлит, 2000. – 224 с.
4. Метод Нелдера – Мида. – URL: http://ru.wikipedia.org/wiki/Метод_Нелдера_-_Мида (дата обращения 14.04.2013).
5. Описание алгоритма RPROP. – URL: <http://www.basegroup.ru/library/analysis/neural/rprop> (дата обращения 14.04.2013).
6. Описание базы данных MNIST. – URL: <http://yann.lecun.com/exdb/mnist> (дата обращения 14.04.2013).
7. Описание работы с базой данных MNIST и результаты экспериментов. – URL: <http://yann.lecun.com/exdb/publis/pdf/lecun-01a.pdf> (дата обращения 14.04.2013).
8. Теорема Цыбенко. – URL: http://actcomm.dartmouth.edu/gvc/papers/approx_by_superposition.pdf (дата обращения 14.04.2013).
9. Хайкин С. Нейронные сети. Полный курс / С. Хайкин. – М.: Вильямс, 2006. – 1104 с.
10. Riedmiller M. A direct adaptive method for faster back-propagation learning: The RPROP algorithm / M. Riedmiller, H. Braun. – San Francisco, 1993.
11. Riedmiller M. RPROP – A Fast Adaptive Learning Algorithm / M. Riedmiller, H. Braun. – Proc. Int. Symp. Computer and Inform. Sci. 1992. VII.

RPROP Method Modification for Pattern Recognition

Blagoveshchenskaya E. A., Zuev D. V.
Petersburg State Transport University
St. Petersburg, Russia
kblag2002@yahoo.com, zuevdv@gmail.com

Abstract. For perceptron learning and also for special architecture neural networks the learning algorithms used normally have some parameters chosen on the basis of the previous experience of people involved or according to some external recommendations, that is “manually”, as we say. The parameter choice can determine the speed of the algorithm result achievement. Today there is no universal method of automatic determination of such parameters. The article describes an algorithm of the well-known RPROP method modification leading to the increase of convergence speed in pattern recognition.

Keywords: neural networks, RPROP, pattern recognition.

REFERENCES

1. Blagoveshchenskaia E. A., Zuev D. V. O skhodimosti metodov obu-cheniia neuronnykh setei [On the Convergence of Methods of Training Neural Networks]. *Materialy XIX Mezhdunarodnoi konferentsii po Vychislitel'noi mekhanike i sovremennym prikladnym programmnyim sistemam (VMSPPS'2015) [Articles XIX Int. Conf. on Computer Mechanics and Modern Applied Codes]*. Moscow, 2015. Pp. 130–131.
2. Vasil'ev A. N., Tarkhov D. A. Neirosetevoe modelirovanie. Printsipy. Algoritmy. Prilozheniia [Neural network modeling. Principles. Algorithms. Applications]. St. Petersburg, Izdatelstvo Politekhnikheskogo Universiteta, 2009. 528 p.
3. Kruglov V. V., Dli M. I., Golunov R. Iu. Nechetkaia logika i iskusstvennye neuronnye seti [Fuzzy Logic and Artificial Neural Networks]. Moscow, Fizmatlit, 2000. 224 p.
4. Metod Neldera – Mida [Method Neldera – Mida]. Available at: https://ru.wikipedia.org/wiki/%D0%9C%D0%B5%D1%82%D0%BE%D0%B4_%D0%9D%D0%B5%D0%BB%D0%B4%D0%B5%D1%80%D0%B0_%E2%80%94%D0%9C%D0%B8%D0%B4%D0%B0 (accessed 14.04.2013)
5. Opisanie algoritma RPROP [Description of the Algorithm RPROP]. Available at: Rezhim dostupa: <http://www.basegroup.ru/library/analysis/neural/rprop> (accessed 14.04.2013).
6. Opisanie bazy dannykh MNIST [Database Description MNIST]. Available at: <http://yann.lecun.com/exdb/mnist/> (accessed 14.04.2013).
7. Opisanie raboty s bazoi dannykh MNIST i rezul'taty eksperimentov [How to Work with the Database and the Results of Experiments MNIST]. Available at: <http://yann.lecun.com/exdb/publis/pdf/lecun-01a.pdf> (accessed 14.04.2013).
8. Teorema Tsybenko [Theorem Tsybenko]. Available at: <http://deeplearning.cs.cmu.edu/pdfs/Cybenko.pdf> (accessed 06.10.2015).
9. Khaikin. S. Neironnye seti. Polnyi kurs [Neural networks. Full course]. Moscow, Vil'iams, 2006. 1104 p.
10. Riedmiller M., Braun H. A Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm. San Francisco, 1993.
11. Riedmiller M., Braun H. RPROP – A Fast Adaptive Learning Algorithm. Proc. Int. Symp. Computer and Inform. Sci. 1992. VII.