

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Том 21
2015
№ 1

ТЕОРЕТИЧЕСКИЙ И ПРИКЛАДНОЙ НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

Издается с ноября 1995 г.

УЧРЕДИТЕЛЬ
Издательство "Новые технологии"

СОДЕРЖАНИЕ

ОБЩИЕ ВОПРОСЫ

Цветков В. Я. Информационная неопределенность и информационная определенность в науках об информации 3

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ

Катруца А. М., Стрижов В. В. Проблема мультиколлинеарности при выборе признаков в регрессионных задачах 8

Имамвердиев Я. Н., Сухостат Л. В. Разработка робастного метода извлечения речевых признаков на основе эмпирического вейвлет-преобразования. 19

МОДЕЛИРОВАНИЕ И ОПТИМИЗАЦИЯ

Чеканин В. А., Чеканин А. В. Повышение эффективности конструирования ортогональной упаковки объектов. 24

Шелудько А. С., Ширяев В. И. Алгоритм гарантированного оценивания параметра одномерного хаотического отображения 30

Бойко А. А. Способ стратифицированного аналитического описания процесса функционирования информационно-технических средств. 35

Струченков В. И. Комбинированные методы решения вариационных задач с заданной структурой экстремали и ограничениями 43

МЕТОДЫ И ТЕХНОЛОГИИ ОБРАБОТКИ ДАННЫХ

Печенко И. С., Венгер О. В. Способ автоматизированного создания диаграмм последовательности операций для сценариев поведения системы, описанных в виде текста. 50

Бочков М. В., Бойков П. Н. Инфодинамическая модель поиска пользователя в социальной сети 56

Костин В. Н., Данышин Д. В. Метод оценки глубины прогноза развития (эволюции) характеристик сложных систем на основе энтропийного подхода. 62

БЕЗОПАСНОСТЬ ИНФОРМАЦИИ

Глушенко С. А., Долженко А. И. Система поддержки принятия решений нечеткого моделирования рисков информационной безопасности организации 68

ПРИКЛАДНЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Жусов Д. Л., Козленко А. В., Шульга Р. В. Алгоритм анализа глубины гипертекстовых переходов официальных сайтов федеральных органов исполнительной власти Российской Федерации 75

Главный редактор:
СТЕМПКОВСКИЙ А. Л.,
акад. РАН, д. т. н., проф.

Зам. главного редактора:
ИВАННИКОВ А. Д., д. т. н., проф.
ФИЛИМОНОВ Н. Б., д. т. н., с.н.с.

Редакционный совет:
БЫЧКОВ И. В., акад. РАН, д. т. н.
ЖУРАВЛЕВ Ю. И.,
акад. РАН, д. ф.-м. н., проф.
КУЛЕШОВ А. П.,
акад. РАН, д. т. н., проф.
ПОПКОВ Ю. С.,
чл.-корр. РАН, д. т. н., проф.
РУСАКОВ С. Г.,
чл.-корр. РАН, д. т. н., проф.
СОЙФЕР В. А.,
чл.-корр. РАН, д. т. н., проф.
СОКОЛОВ И. А., акад.
РАН, д. т. н., проф.
СУЕТИН Н. В., д. ф.-м. н., проф.
ЧАПЛЫГИН Ю. А.,
чл.-корр. РАН, д. т. н., проф.
ШАХНОВ В. А.,
чл.-корр. РАН, д. т. н., проф.
ШОКИН Ю. И.,
акад. РАН, д. т. н., проф.
ЮСУПОВ Р. М.,
чл.-корр. РАН, д. т. н., проф.

Редакционная коллегия:
АВДОШИН С. М., к. т. н., доц.
АНТОНОВ Б. И.
БАРСКИЙ А. Б., д. т. н., проф.
ВАСЕНИН В. А., д. ф.-м. н., проф.
ГАЛУШКИН А. И., д. т. н., проф.
ДИМИТРИЕНКО Ю. И., д. ф.-м. н., проф.
ДОМРАЧЕВ В. Г., д. т. н., проф.
ЗАГИДУЛЛИН Р. Ш., к. т. н., доц.
ЗАРУБИН В. С., д. т. н., проф.
ИСАЕНКО Р. О., к. т. н., с.н.с.
КАРПЕНКО А. П., д. ф.-м. н., проф.
КОЛИН К. К., д. т. н., проф.
КУЛАГИН В. П., д. т. н., проф.
КУРЕЙЧИК В. М., д. т. н., проф.
КУХАРЕНКО Б. Г., к. ф.-м. н., доц.
ЛЬВОВИЧ Я. Е., д. т. н., проф.
МИХАЙЛОВ Б. М., д. т. н., проф.
НЕЧАЕВ В. В., к. т. н., проф.
РЯБОВ Г. Г., чл.-корр. РАН, д. т. н., проф.
СОКОЛОВ Б. В., д. т. н., проф.
УСКОВ В. Л., к. т. н. (США)
ФОМИЧЕВ В. А., д. т. н., проф.
ЧЕРМОШЕНЦЕВ С. Ф., д. т. н., проф.
ШИЛОВ В. В., к. т. н., доц.

Редакция:
БЕЗМЕНОВА М. Ю.
ГРИГОРИН-РЯБОВА Е. В.
ЛЫСЕНКО А. В.
ЧУГУНОВА А. В.

Информация о журнале доступна по сети Internet по адресу <http://novtex.ru/IT>.
Журнал включен в систему Российского индекса научного цитирования.
Журнал входит в Перечень научных журналов, в которых по рекомендации ВАК РФ должны быть опубликованы научные результаты диссертаций на соискание ученой степени доктора и кандидата наук.

INFORMATION TECHNOLOGIES

INFORMACIONNYYE TEHNOLOGII

Vol. 21
2015
No. 1

THEORETICAL AND APPLIED SCIENTIFIC AND TECHNICAL JOURNAL

Published since November 1995

ISSN 1684-6400

CONTENTS

GENERAL QUESTIONS

Tsvetkov V. Ya. Information Uncertainty and Information Certainty in Information Science. 3

INTELLIGENT SYSTEMS AND TECHNOLOGIES

Katrutsa A. M., Strijov V. V. The Multicollinearity Problem for Feature Selection Methods in Regression. 8

Imamverdiyev Y. N., Sukhostat L. V. Development of Robust Speech Feature Extraction Method Based on Empirical Wavelet Transform 19

MODELING AND OPTIMIZATION

Chekanin V. A., Chekanin A. V. Improving the Efficiency of Construction of the Orthogonal Packing 24

Sheludko A. S., Shiryaev V. I. The Algorithm of Guaranteed Parameter Estimation for One-Dimensional Chaotic Map. 30

Boyko A. A. Method of Stratified Analytical Description of the Process of Functioning of Information-Technical Tools. 35

Srtuchukov V. I. Combined Algorithms for Solving Variational Problems with a Given Structure of the Extremal and Constraints 43

METHODS AND TECHNOLOGIES OF DATA PROCESSING

Pechenko I. S., Venger O. V. Generation of Flow Diagram from Textual Description of the Flow 50

Bochkov M. V., Boykov P. N. Infodynamic Model of Search of the User on a Social Network 56

Kostin V. N., Danshin D. V. Method of an Assessment of Depth of the Forecast of Development (Evolution) Difficult Systems on the Basis of Entropy Approach. 62

CRYPTOSAFETY INFORMATION

Glushenko S. A., Doljenko A. I. Decision Support Fuzzy Modeling System of Information Security Risks 68

APPLIED INFORMATION TECHNOLOGIES

Zhusov D. L., Kozlenko A. V., Shulga R. V. Hypertext Navigation Depth Analysis Algorithm for Federal Executive Authorities Official Websites in the Russian Federation. 75

Editor-in-Chief:

Stempkovsky A. L., Member of RAS,
Dr. Sci. (Tech.), Prof.

Deputy Editor-in-Chief:

Ivannikov A. D., Dr. Sci. (Tech.), Prof.
Filimonov N. B., Dr. Sci. (Tech.), Prof.

Chairman:

Bychkov I. V., Member of RAS,
Dr. Sci. (Tech.), Prof.

Zhuravljov Yu. I., Member of RAS,
Dr. Sci. (Phys.-Math.), Prof.

Kuleshov A. P., Member of RAS,
Dr. Sci. (Tech.), Prof.

Popkov Yu. S., Corresp. Member of RAS,
Dr. Sci. (Tech.), Prof.

Rusakov S. G., Corresp. Member of RAS,
Dr. Sci. (Tech.), Prof.

Soifer V. A., Corresp. Member of RAS,
Dr. Sci. (Tech.), Prof.

Sokolov I. A., Member of RAS,
Dr. Sci. (Phys.-Math.), Prof.

Suetin N. V.,
Dr. Sci. (Phys.-Math.), Prof.

Chaplygin Yu. A., Corresp. Member of RAS,
Dr. Sci. (Tech.), Prof.

Shakhnov V. A., Corresp. Member of RAS,
Dr. Sci. (Tech.), Prof.

Shokin Yu. I., Member of RAS,
Dr. Sci. (Tech.), Prof.

Yusupov R. M., Corresp. Member of RAS,
Dr. Sci. (Tech.), Prof.

Editorial Board Members:

Avdoshin S. M., Cand. Sci. (Tech.), Ass. Prof.
Antonov B. I.

Barsky A. B., Dr. Sci. (Tech.), Prof.

Vasenin V. A., Dr. Sci. (Phys.-Math.), Prof.

Galushkin A. I., Dr. Sci. (Tech.), Prof.

Dimitrienko Yu. I., Dr. Sci. (Phys.-Math.), Prof.

Domrachev V. G., Dr. Sci. (Tech.), Prof.

Zagidullin R. Sh., Cand. Sci. (Tech.), Ass. Prof.

Zarubin V. S., Dr. Sci. (Tech.), Prof.

Isaenko R. O., Cand. Sci. (Tech.)

Karpenko A. P., Dr. Sci. (Phys.-Math.), Prof.

Kolin K. K., Dr. Sci. (Tech.)

Kulagin V. P., Dr. Sci. (Tech.), Prof.

Kurechik V. M., Dr. Sci. (Tech.), Prof.

Kukhareenko B. G., Cand. Sci. (Phys.-Math.)

Ljvovich Ya. E., Dr. Sci. (Tech.), Prof.

Mikhailov B. M., Dr. Sci. (Tech.), Prof.

Nechaev V. V., Cand. Sci. (Tech.), Ass. Prof.

Ryabov G. G., Corresp. Member of RAS,
Dr. Sci. (Tech.), Prof.

Sokolov B. V., Dr. Sci. (Tech.)

Uskov V. L. (USA), Dr. Sci. (Tech.)

Fomichev V. A., Dr. Sci. (Tech.), Prof.

Chermoshentsev S. F., Dr. Sci. (Tech.), Prof.

Shilov V. V., Cand. Sci. (Tech.), Ass. Prof.

Editors:

Bezmenova M. Yu.

Grigorin-Ryabova E. V.

Lysenko A. V.

Chugunova A. V.

Complete Internet version of the journal at site: <http://novtex.ru/IT>.

According to the decision of the Higher Certifying Commission of the Ministry of Education of Russian Federation, the journal is inscribed in "The List of the Leading Scientific Journals and Editions wherein Main Scientific Results of Theses for Doctor's or Candidate's Degrees Should Be Published"

ОБЩИЕ ВОПРОСЫ GENERAL QUESTIONS

УДК 004.6; 528; 004.8

В. Я. Цветков, д-р техн. наук, проф., советник, e-mail: cvj2@mail.ru,
Московский государственный технический университет радиотехники,
электроники и автоматики (МГТУ МИРЭА)

Информационная неопределенность и информационная определенность в науках об информации

Статья описывает информационную неопределенность и информационную определенность с позиций теории информации К. Э. Шеннона и Н. Винера. Раскрывается сущность информационной неопределенности как разновидности "не знания". Дается различие между понятиями "не знание" и "незнание". Показано, что неопределенность и "не знание" описываются интервальным рядом. Показано условие эквивалентности сообщений по Шеннону сообщениям по Винеру—Флориди. Вводится понятие информационной конструкции как альтернатива понятию информационное сообщение. Раскрывается понятие измеримости информационной конструкции.

Ключевые слова: науки об информации, теория информации, знание, не знание, информационная определенность, информационная неопределенность, информационные конструкции, информационная измеримость, информационные модели, содержательность сообщения, информационные единицы, информация

Введение

Термин "науки об информации" является более общим, чем термин "теория информации". С одной стороны, он включает информатику и различные науки об информационных технологиях. С другой стороны, теория информации до настоящего времени не сформирована и существует два разных ее направления.

Первое направление основано на работе К. Э. Шеннона "математическая теория связи" [1]. Под "информацией" в этой теории понимают "нечто", что уменьшает неинформированность и неопределенность, а именно, уменьшает "незнание". Количество информации в этой теории называют информационной емкостью носителя информации, сообщения или иной информационной конструкции безотносительно к содержанию. Ее измеряют в байтах и битах безотносительно к тому, что в этих байтах содержится или не содержится ничего.

Винер называет теорию К. Э. Шеннона "статистической теорией количества информации" [2, 3]. "В этой теории за единицу количества информации принимается количество информации, *передаваемое при одном выборе между двумя равновероятными альтернативами*. Такая мысль возникла одновременно у Р. А. Фишера, у доктора Шеннона из *Bell Labs* и у автора настоящей книги" [2]. "При этом Фишер исходил из статистической теории, а Шеннон — из проблемы кодирования информации" [2].

Первоначально К. Э. Шеннон не говорил о своей работе как о теории информации и отмечал что "*семантические аспекты информации не релевантны техническим проблемам связи*" [1]. Однако позже, в 1964 г., появилась совместная работа Шеннона и Уивера [4], в которой точка зрения "теории информации на основе математической теории связи" развивалась как основная. В ней рассматривались "три уровня" проблемы коммуникации: техническая проблема передачи сообщений, семантическая проблема, проблема эффективности передачи сообщений. Следует подчеркнуть, что речь идет, как и в первой работе, о передаче сообщений, а не информации. Проблема семантики сводится к вопросу "How precisely do the transmitted symbols convey the desired meaning?" "Как точно передаваемые символы передают желаемое значение?". Однако ответ на этот вопрос не был полным и касался точности полученных символов в сравнении с исходными.

Такая точка зрения на теорию информации не была однозначно принята научным сообществом, и стали появляться работы, отражающие другие точки зрения. Вторым подходом развит в работах Н. Виннера [2, 3], Р. Карнапа [5], Л. Флориди [6], А. Н. Колмогорова [7], Ю. А. Шрейдера [8], А. А. Харкевича [9] и др. В этом подходе под информацией понимают содержание информационной конструкции безотносительно к ее объему и часто рассматривают это в аспекте семантики. Эту теорию часто называют

семантической теорией информации. Под "информацией" в этой теории понимают "нечто", что содержит знание и увеличивает в итоге знание получателя.

Оценка по Шеннону и по Винеру

По Шеннону за единицу количества информации принимается количество информации, передаваемое при одном выборе между двумя равновероятными альтернативами. По Винеру "Информация — обозначение содержания, полученного от внешнего мира в процессе приспособления к нему" [3]. Для пояснения различия в оценке количества информации по Шеннону и по Винеру возьмем пример из работы [10]. Рассмотрим три фразы:

"казнить нельзя, помиловать";
"казнить, нельзя помиловать";
"нить, ватнепомльзя казило".

По Шеннону все три фразы содержат одинаковое количество информации, поскольку имеют одинаковое число символов, содержащих равное число битов.

По Виннеру и Флориди первая и вторая фразы противоположны по смыслу, а третья бессмысленна, поскольку не содержит полезной информации. Критерий полезности можно использовать по Харкевичу [9].

По Шеннону информация уменьшает неопределенность, по Виннеру и Флориди — информация несет полезное содержание или содержит знание.

Дихотомия пары "не знание — знание"

Дихотомический анализ является одним из основных подходов [11], применяемых в системном анализе, в проектировании и при декомпозиции проектов. Он применим и при анализе содержания информационных сообщений.

По Шеннону информация уменьшает "наше незнание", по Виннеру и Флориди — информация увеличивает "наше знание". Увеличение знания и уменьшение незнания — не всегда одно и то же, что будет показано ниже. При этом термин "не знание" отличается от "незнания". Незнание обозначает одну сущность или элемент множества. "Не знание" обозначает совокупность сущностей или множество.

Дихотомическая пара "не знание — знание" представляет собой оппозицию [12, 13] и задает шкалу анализа для исследования содержательности информационных сообщений или других информационных конструкций.

Говоря о "не знании", следует остановиться на работе Нариньяни о "НЕ-факторах" [14], которая берет корни в его более ранних работах 1982 г. При этом мы упрощаем постановку задачи, поскольку Нариньяни говорит о знаниях в аспекте области

искусственного интеллекта, а мы переносим его рассуждения в область теории информации.

"НЕ-факторы отнюдь не нечто диковинное, — они "встроены" в нашу Картину мира как основная составляющая знаний, являясь их материей и строительным материалом" [14]. Следует отметить, что "картина мира" или "модель картины мира" может служить критерием оценки результата научных исследований или образования [15, 16].

Исследование НЕ-факторов Нариньяни связывает с идеей *недоопределенных множеств*, которую чаще называли теорией нечетких множеств (*fuzzy sets*).

В ряде работ Нариньяни формирует идею *недоопределенности* как общего явления. При переносе этого подхода в область теории информации мы приходим к информационной неопределенности. "Выяснилось, что интервальное представление может служить примером как недоопределенного числа, так и неточного с тем принципиальным различием, что первое может уточняться, а второе нет" [17]. Это отражает различие между "не знанием" и "незнанием", а также факт того, что информационная неопределенность может уменьшаться.

Здесь следует напомнить, что в "четкой логике" и в "четкой" физике или математике одно качественное явление (фактор) характеризуется одним количественным значением. Например говорят, что поезд прошел 400 км; в текущий момент времени он идет со скоростью 60 км/ч; он тормозит с ускорением $0,1 \text{ м/с}^2$. Это означает парадигму "один фактор — одно числовое значение", т. е. человек с вероятностью 1 определяет это значение.

При отсутствии уверенности человек задает интервал значений. Например, говорят, что поезд прошел не менее 250 км и не более 400 км; в текущий момент времени он идет со скоростью от 45 до 65 км/ч; он тормозит с ускорением $0,1 \pm 0,01 \text{ м/с}^2$. Это соответствует парадигме "один фактор — ряд числовых значений" или "один фактор — множество числовых значений". Такая ситуация характерна для "НЕ-факторов".

Работы в области "НЕ-факторов" относят к области плохо формализуемых явлений. Это направление в последнее время подвергается исследованию и изучению. Однако в неявной форме в науке придерживаются неформального правила "равенства категорий", которое восходит еще к Аристотелю. Согласно этому правилу сопоставлять можно величины, относящиеся к одной категориальной группе. Пространственные точечные величины сопоставляют с точечными и на этой основе решают задачи. Площадные величины сопоставляют с площадными, линейные с линейными и т. д.

При введении "не-фактора" или "не знания" оказывается, что оппозиционные факторы и "не знания" могут относиться к разным категориям. Например,

"человек — не человек". К области "не человек" можно отнести камень, автомобиль, бактерию, дерево и даже компьютерную программу, т. е. совокупности из разных несопоставимых или мало сопоставимых между собой категориальных групп. Это еще более усугубляет "малоформализуемость" таких явлений.

Модели неопределенности

Для пояснения различия между подходами Шеннона и Винера рассмотрим еще один пример. Допустим, точка M лежит в плоскости и содержится в первом квадранте.

По Шеннону все сообщения, уменьшающие неопределенность, содержат информацию, т. е. информационно измеримы. Информационно измеримой назовем такую информационную конструкцию, которая кроме качественной оценки может иметь одну или ряд количественных оценок, позволяющих сравнивать ее с другими информационными конструкциями.

Это сообщения (информационные конструкции) следующего характера: "Точка M не содержится во втором квадранте"; "Точка M не содержится в третьем квадранте"; "Точка M не содержится в четвертом квадранте". Каждое из перечисленных сообщений информирует получателя сообщения и уменьшает его информационную неопределенность. Однако получив только три сообщения, получатель методом исключения (так как квадрантов всего четыре) определит, что точка находится в первом квадранте.

По Винеру правильным и содержательным будет одно сообщение "Точка M содержится в первом квадранте". Все остальные сообщения типа: "точка может содержаться в первом или во втором", "точка не содержится..." являются бессодержательными. Но следует отметить, что в этом примере три сообщения по Шеннону, уменьшающие неопределенность, равнозначны одному содержательному сообщению по Винеру. Это означает, что интервал (интервальное значение) сообщений по Шеннону эквивалентен одному (точечное значение) сообщению по Винеру.

Можно усложнить эксперимент. Разобьем плоскость на 360 секторов. Пусть точка находится условно в 360-м секторе. В этом случае 359 сообщений по Шеннону "точка не находится в секторе I ", $I = 1...359$, будут равнозначны одному сообщению по Винеру "точка находится в секторе 360". Этим подчеркивается разница между "не знанием" и знанием. То есть для N состояний интервал из $N - 1$ сообщений, уменьшающих информационную неопределенность "не знание", эквивалентны одному сообщению, содержащему знание. Интервал из $N - 1$ сообщений назовем полным, так как он снимает информационную неопределенность.

Это дает возможность ввести условие эквивалентности сообщений по Шеннону сообщению по Винеру. Совокупность сообщений по Шеннону будет эквивалентна содержательному сообщению по Винеру, если они образуют полный интервальный ряд сообщений. Термин "сообщение" можно заменить на термин "информационная конструкция".

Следующий пример. Можно от дискретной модели перейти к непрерывной. Пусть точка имеет координаты $X = 3$, $Y = 4$. По Винеру одно высказывание, назовем его условно (A), будет истинным, т. е. информативным " $X = 3$, $Y = 4$ ", все остальные будут неинформативными.

По Шеннону бесконечное число высказываний будут содержать количество информации, уменьшающее неопределенность, т. е. быть информативными:

Первая группа $X = 3$, $Y =$ любое, кроме 4;

Вторая группа $X =$ любое, кроме 3, $Y = 4$;

Третья группа $X =$ любое, кроме 3, $Y =$ любое, кроме 4.

Но в этих группах не будет ни одного высказывания, равнозначного высказыванию (A).

Информационные единицы как составляющие сообщений

Общим для подходов Шеннона и Винера является использование информационных единиц [18, 19], хотя в разных аспектах рассмотрения. Поэтому исследование информационных единиц представляет интерес как для наук об информации, так и для наук, в которых понятие информации применяется [20, 21].

Простейшее информационное сообщение может включать одно слово. Такое слово можно рассматривать также как информационную модель [22] или как информационную единицу [23]. Слово есть обозначение. Смысл слова или его значение раскрывается через интерпретацию данного слова, через совокупность интерпретирующих его предложений [24].

Слово (как информационная единица) по Шеннону содержит такое количество информации, которое определяется количеством информации в символах, входящих в это слово. Если переставить символы в слове, то смысл его исчезнет, но по Шеннону количество информации не изменится. Информационный объем символов определяет количество информации. Следовательно, мера Шеннона определяет информационный объем сообщения.

По Винеру информация, которую содержит слово, определяется числом интерпретируемых его предложений. Если переставить символы в слове, то смысл его исчезнет, и это влечет исчезновение интерпретируемых предложений, т. е. количество информации становится равным нулю. Следовательно, мера Винера и Флориди определяет семантику сообщения.

Заключение

Информационная неопределенность описывается интервальным рядом сообщений и определяет "не знание". Информационная определенность описывается одним содержательным сообщением или полным рядом, снимающим неопределенность, и определяет "знание".

Подход Шеннона решает две задачи. Он исследует информационную неопределенность и позволяет оценить информационную емкость сообщения (информационной конструкции). Подход Шеннона задает интервальный ряд информационных конструкций для описания некоего события. Подход Шеннона задает информационные конструкции, определяющие "не знание". Подход Шеннона можно использовать для получения истинного знания, но при этом надо использовать полный интервальный ряд сообщений, уменьшающих неопределенность. Теория информации, задаваемая подходом Шеннона, является "статистической" и не является "семантической".

Подход Виннера — Флориди исследует задачу передачи смысла в информационной конструкции. Подход Виннера — Флориди задает информационные конструкции, определяющие "знание" и фактически является основой семантической теории информации.

Список литературы

1. Shannon C. E. A Mathematical Theory of Communication // Bell System Technical Journal. 1948. July & October. Vol. 27. P. 379—423, 623—656.
2. Winner N. Cybernetics or Control and Communication in the Animal and the Machine. New York — Herman et Cie, Paris: The technology Press and John Wiley & Sons Inc. 1948. 194 p.
3. Винер Н. Кибернетика, или управление и связь в животном и машине. М.: Сов. радио, 1958.
4. Shannon C. E. and Weaver W. The Mathematical Theory of Communication. The University of Illinois Press URBANA, 1964. 132 p.

5. Carnap R. et al. An outline of a theory of semantic information. — Research Laboratory of Electronics, Technical Report. N. 247. Massachusetts: MIT, 1952. С. 221—274.
6. Floridi L. Semantic Conceptions of Information. URL: <http://plato.stanford.edu/entries/information-semantic>.
7. Колмогоров А. Н. Три подхода к определению понятия "количество информации" // Проблемы передачи информации. 1965. Т. 1, № 1. С. 3—11.
8. Шрейдер Ю. А. О семантических аспектах теории информации // Информация и кибернетика. М.: Сов. радио. 1967. С. 15—47.
9. Харкевич А. А. Избранные труды. В 3-х т. Том 3. Теория информации. Оpozнание образов. М.: Наука, 1973. 524 с.
10. Иванников А. Д., Тихонов А. Н., Цветков В. Я. Основы теории информации. М.: МаксПресс, 2007. 356 с.
11. Tsvetkov V. Y. Dichotomous Systemic Analysis // Life Science Journal 2014. N. 11 (6). P. 586—590.
12. Tsvetkov V. Y. Opposition Variables as a Tool of Qualitative Analysis // World Applied Sciences Journal. 2014. N. 30 (11). P. 1703—1706.
13. Tsvetkov V. Y. Correlative analysis and opposition variables // European Journal Of Natural History. 2014. N. 1. P. 48—52.
14. Нариньяни А. С. НЕ-факторы: краткое введение // Новости искусственного интеллекта. М.: КОМКНИГА, 2006. Вып. 2. С. 52—63.
15. Цветков В. Я. Картина мира как образовательная парадигма // European Social Science Journal. 2013. N. 10-1 (37). P. 28—34.
16. Tsvetkov V. Y. Worldview Model as the Result of Education // World Applied Sciences Journal. 2014. N. 31 (2). P. 211—215.
17. Нариньяни А. С. Недоопределенные модели и операции с недоопределенными значениями. Препринт ВЦ СО АН СССР, № 400, 1982.
18. Цветков В. Я. Семантика информационных единиц // Успехи современного естествознания. 2007. № 10. С. 103—104.
19. Цветков В. Я. Язык информатики // Успехи современного естествознания. 2014. № 7. С. 129—133.
20. Соловьёв И. В. Идеальное, формальное, материальное — в информационных сообщениях // Перспективы науки и образования. 2014. № 1. С. 51—55.
21. Лакофф Д. Когнитивная семантика // Язык и интеллект. М.: Прогресс. 1996. С. 143—184.
22. Цветков В. Я. Модели в информационных технологиях. М.: Макс Пресс, 2006. 104 с.
23. Tsvetkov V. Ya. Information objects and information Units // European Journal of Natural History. 2009. N. 2. P. 99.
24. Фреге Г. Логика и логическая семантика. М.: Аспект пресс, 2000.

V. Ya. Tsvetkov, Advisor to the Rectorat, e-mail: cvj2@mail.ru, MSTU MIREA

Information Uncertainty and Information Certainty in Information Science

This article describes the information uncertainty and certainty of information. These concepts are revealed from the standpoint of information theory KE Shannon and Wiener. The article reveals the essence of information uncertainty as a kind of "not knowing." article gives the difference between the concepts of "not knowing" and "ignorance." It is shown that the uncertainty and "not knowing" describes interval number of information structures. Displaying posts equivalence condition Shannon reportedly N. Wiener — L. Floridi. Article introduces the concept of information design as an alternative to the concept of an information message. The notion of measurability of information construction

Keywords: information science, information theory, knowledge, not knowledge, information certainty, information uncertainty, information construction, information measurability, information model, the content of messages, information units, information

References

1. Shannon C. E. A Mathematical Theory of Communication". *Bell. System Technical Journal*. 1948. July & October. V. 27. P. 379—423 & 623—656.
2. Winner N. *Cybernetics or Control and Communication in the Animal and the Mashine*. New York — Herman et Cie, Paris: The Technology Press and John Wiley & Soris Inc., 1948. 194 p.
3. Winner N. *Cibernetika, ili upravlenie i svyaz v givotnom i mashine*. M.: Sovetskoe radio, 1958.
4. Shannon C. E., Weaver W. *The Mathematical Theory of Communication*. The University of Illinois Press — URBANA. 1964. 132 p.
5. Carnap R. et al. *An outline of a theory of semantic information*. Research Laboratory of Electronics, Technical Report. N. 247. Massachusetts: MIT, 1952. P. 221—274.
6. Floridi L. *Semantic Conceptions of Information*. URL: <http://plato.stanford.edu/entries/information-semantic>.
7. Kolmogorov A. N. Tri podhoda k opredeleniyu ponyatiya "kolichestvo informatsii". *Problemy peredachi informatsii*. 1965. V. 1, N. 1. P. 3—11.
8. Shreyder Yu. A. *O semanticheskikh aspektakh teorii informatsii*. M.: Sovetskoe radio. 1967. P. 15—47.
9. Kharkevich A. A. *Izbranye trudy v 3-h t. T. 3. Teoriya informatsii. Opoznanie obrazov*. M.: Nauka, 1973. 524 p.
10. Ivannikov A. P., Tichonov A. N., Tsvetkov V. Ya. *Osnovy teorii informatsii*. M.: MaksPress, 2007. 356 p.
11. Tsvetkov V. Y. Dichotomous Systemic Analysis. *Life Science Journal*. 2014. N. 11 (6). P. 586—590.
12. Tsvetkov V. Y. Opposition Variables as a Tool of Qualitative Analysis. *World Applied Sciences Journal*. 2014. N. 30 (11). P. 1703—1706.
13. Tsvetkov V. Y. Correlative analysis and opposition variables. *European Journal of Natural History*. 2014. N. 1. P. 48—52.
14. Narinyani A. S. Ne-factory: kratkoe vvedenie. *Novosti iskusstvennogo intellekta*. Vyp. 2. M.: KOMKNIGA, 2006. P. 52—63.
15. Tsvetkov V. Y. Kartina mira kak obrazovatel'naya paradigma. *European Social Science Journal*. 2013. N. 10-1 (37). P. 28—34.
16. Tsvetkov V. Y. Worldview Model as the Result of Education. *World Applied Sciences Journal*. 2014. N. 31 (2). P. 211—215.
17. Narinyani A. S. *Nedoopredelennyye modeli i operatsii s nedoopredelennymi znateniyami*. Preprint VTs SO AN SSSR, N. 400, 1982.
18. Tsvetkov V. Y. Semantika informatsionnykh edinits. *Uspechi sovremennogo estestvoznaniya*. 2007. N. 10. P. 103—104.
19. Tsvetkov V. Y. Yazyk informatiki. *Uspechi sovremennogo estestvoznaniya*. 2014. N. 7. P. 129—133.
20. Soloviev I. V. Idejnoe, formalnoe, materialnoe — v informatsionnykh soobcheniyakh. *Perspektivy nauki i obrazovaniya*. 2014. N. 1. P. 51—55.
21. Lakoff D. *Cognitivnaya semantika. Yazyk i intellekt*. M.: Progress. 1996. P. 143—184.
22. Tsvetkov V. Y. *Modeli v informatsionnykh technologiyyakh*. M.: MaksPress. 2006. 104 p.
23. Tsvetkov V. Y. Information objects and information Units. *European Journal of Natural History*. 2009. N. 2. P. 99.
24. Frege G. *Logika i logicheskaya semantika*. M.: Aspekt hress. 2000.

ИНФОРМАЦИЯ



V Международная научно-техническая конференция
"Открытые семантические технологии проектирования
интеллектуальных систем"
Open Semantic Technologies for Intelligent Systems
OSTIS-2015

19 – 21 февраля 2015 г., Белорусский государственный университет
информатики и радиоэлектроники, Минск, Республика Беларусь

НАПРАВЛЕНИЯ РАБОТЫ КОНФЕРЕНЦИИ

- Принципы, лежащие в основе семантического представления знаний, и их унификация. Типология знаний и особенности семантического представления различного вида знаний и метазнаний. Связи между знаниями и отношения, заданные на множестве знаний. Семантическая структура глобальной базы знаний, интегрирующей различные накапливаемые знания
- Языки программирования, ориентированные на параллельную обработку семантического представления баз знаний
- Модели решения задач, в основе которых лежит обработка знаний, осуществляемая непосредственно на уровне семантического представления обрабатываемых знаний. Семантические модели информационного поиска, интеграции знаний, анализа корректности и качества баз знаний, сборки информационного мусора, оптимизации баз знаний, дедуктивного и индуктивного вывода в базах знаний, правдоподобных рассуждений, распознавания образов, интеллектуального управления. Интеграция различных моделей решения задач
- Семантические модели восприятия информации о внешней среде и отображения этой информации в базу знаний
- Семантические модели мультимодальных пользовательских интерфейсов интеллектуальных систем, в основе которых лежит семантическое представление используемых ими знаний, и унификация этих моделей
- Семантические модели естественно-языковых пользовательских интерфейсов интеллектуальных систем. Структура семантического представления лингвистических баз знаний, описывающих естественные языки и обеспечивающих решение задач понимания естественно-языковых текстов и речевых сообщений, а также задач синтеза естественно-языковых текстов и речевых сообщений, семантически эквивалентных заданным фрагментам баз знаний
- Интегрированные комплексные логико-семантические модели интеллектуальных систем, основанные на семантическом представлении знаний, и их унификация
- Различные технические платформы и варианты реализации интерпретаторов унифицированных логико-семантических моделей интеллектуальных систем, основанных на семантическом представлении знаний
- Средства и методы, основанные на семантическом представлении знаний и ориентированные на проектирование различных типовых компонентов интеллектуальных систем (баз знаний, программ, решателей задач, интерфейсов)
- Средства и методы, основанные на семантическом представлении знаний и ориентированные на комплексное проектирование различных классов интеллектуальных систем (интеллектуальных справочных систем, интеллектуальных обучающих систем, интеллектуальных систем управления, интеллектуальных робототехнических систем, интеллектуальных систем поддержки проектирования и др.)
- Прикладные интеллектуальные системы, основанные на семантическом представлении используемых ими знаний

Рабочие языки конференции: русский, белорусский, английский.

Вся необходимая информация о предстоящей и предыдущих конференциях OSTIS находится на сайте конференции <http://conf.ostis.net>

УДК 519.584

А. М. Катруца, студент, Московский физико-технический институт,
В. В. Стрижов, доктор физ.-мат. наук, науч. сотр.,
Вычислительный центр РАН имени А. А. Дородницына, Москва, e-mail: strijov@gmail.com

Проблема мультиколлинеарности при выборе признаков в регрессионных задачах

Исследуется проблема мультиколлинеарности и ее влияние на эффективность методов выбора признаков. Предлагается процедура тестирования методов выбора признаков и методика порождения тестовых выборок с различными типами мультиколлинеарности между признаками. Рассматриваемые методы выбора признаков тестируются на порожденных выборках. Процедура тестирования заключается в применении методов выбора признаков к выборкам с различным типом мультиколлинеарности и оценивании количества мультиколлинеарных признаков в множестве отобранных признаков. В работе приводится критерий сравнения методов выбора признаков. Методы выбора признаков сравниваются согласно различным функционалам качества. Проведено сравнение методов выбора признаков для случая наличия в данных определенного типа мультиколлинеарности. Сделан вывод о качестве работы рассматриваемых методов на определенных типах данных.

Ключевые слова: регрессионный анализ, выбор признаков, мультиколлинеарность, тестовые выборки, критерий качества

Введение

Работа посвящена тестированию методов выбора признаков. Предполагается, что исследуемая выборка содержит значительное число мультиколлинеарных признаков. *Мультиколлинеарность* — это сильная корреляционная связь между отбираемыми для анализа признаками, совместно воздействующими на целевой вектор, которая затрудняет оценивание регрессионных параметров и выявление зависимости между признаками и целевым вектором. Проблема мультиколлинеарности, возможные способы ее обнаружения и устранения описаны в работах [1–3]. Также мультиколлинеарность приводит к уменьшению устойчивости оценок вектора параметров. Оценка вектора параметров называется устойчивой, если малое изменение некоторой компоненты этого вектора приводит к малому изменению соответствующей компоненты оценки целевого вектора.

В задачах анализа данных для уменьшения размерности [4, 5], упрощения использования стандартных алгоритмов машинного обучения [6], удаления нерелевантных признаков [7] и повышения обобщающей способности применяемого алгоритма [8] применяют методы выбора признаков. Также методы выбора признаков используют для решения проблемы мультиколлинеарности в задачах регрессии [9].

Задача выбора оптимального подмножества признаков является одной из основных задач предварительной обработки данных. Методы выбора признаков основаны на минимизации некоторого функционала, который отражает качество рассматриваемого подмножества признаков. В работах [10–12] сделан обзор существующих методов выбора признаков, проведена классификация методов выбора признаков по используемым функционалам качества и стратегии поиска оптимального подмножества признаков.

При наличии мультиколлинеарности в регрессионных задачах применение методов выбора признаков приводит к повышению устойчивости оценок параметров и уменьшению их дисперсии. Для этого применяют методы отбора признаков с различными регуляризаторами или стратегиями добавления и удаления признаков с использованием статистических тестов для проверки значимости добавляемого признака. Примерами методов, использующих регуляризаторы, являются гребневая регрессия [13], где регуляризатор — взвешенная евклидова норма вектора параметров; Lasso [14] и LARS [15], где регуляризатор — взвешенная сумма модулей параметров; Elastic net [16], где регуляризатор — линейная комбинация предыдущих двух регуляризаторов. Методом, использующим проверку значимости добавляемого или удаляемого признака, является шаговая регрессия [17] с различными ком-

бинациями процедур добавления или удаления признаков.

Для тестирования методов выбора признаков в работе [9] предложен метод генерации выборок и функционал, позволяющий оценить качество процедуры выбора признаков. Однако предложенный способ не позволяет оценить изменение критерия качества при непрерывном изменении параметров выборок и структурного параметра мультиколлинеарности.

В нашей работе предложена другая процедура генерации тестовых выборок, основанная на задании свойств признаков. Рассматриваются следующие свойства признаков: мультиколлинеарность между признаками; коррелированность целевому вектору; ортогональность между признаками; ортогональность признаков целевому вектору. Задание количества признаков, обладающих каждым из этих свойств, позволяет генерировать выборки с различным взаимным расположением признаков и целевого вектора. Такой метод генерации тестовых выборок дает возможность исследовать зависимость эффективности методов выбора признаков при непрерывном изменении параметра мультиколлинеарности.

В работе предложен критерий ранжирования методов выбора признаков и методики их тестирования. Критерием ранжирования является число мультиколлинеарных признаков в множестве отобранных признаков, удаление которых приводит к росту ошибки не больше некоторого заданного значения. Методика тестирования заключается в последовательном применении различных методов выбора признаков к тестовым выборкам, каждая из которых отражает некоторый тип мультиколлинеарности, и оценке качества полученного подмножества признаков для каждой пары, включающей метод выбора признаков и тестовую выборку.

1. Постановка задачи выбора признаков

Задана выборка $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}$, $i \in I = \{1, \dots, m\}$, множество свободных переменных — вектор $\mathbf{x} = [x_1, \dots, x_j, \dots, x_n]$, где $j \in J = \{1, \dots, n\}$. Предполагается, что эти переменные принадлежат множеству действительных чисел, либо его подмножеству: $x_i \in \mathbb{X} \subseteq \mathbb{R}^n$ и $y_i \in \mathbb{Y} \subseteq \mathbb{R}^1$. Введем обозначения: $\mathbf{y} = [y_1, \dots, y_m]^T$ — вектор значений зависимой переменной, целевой вектор; $\boldsymbol{\chi}_j = [x_{1j}, \dots, x_{mj}]^T$ — реализация j -й свободной переменной; j -й признак и $\mathbf{X} = [\mathbf{x}_1^T, \dots, \mathbf{x}_m^T]^T = [\boldsymbol{\chi}_1, \dots, \boldsymbol{\chi}_m]$ — матрица плана. Предполагается, что вектор \mathbf{x}_i и число y_i связаны соотношением

$$y_i = f(\mathbf{w}, \mathbf{x}_i) + \varepsilon(\mathbf{x}_i), \quad (1)$$

где $f: \mathbb{W} \times \mathbb{X} \rightarrow \mathbb{Y}$ отображение декартова произведения пространства допустимых параметров \mathbb{W} и пространства значений \mathbb{X} свободной переменной в об-

ласть значений \mathbb{Y} зависимой переменной, а $\varepsilon(\mathbf{x}_i)$ — i -й компонент вектора регрессионных остатков $\boldsymbol{\varepsilon} = \mathbf{f} - \mathbf{y}$. Обозначим вектор-функцию

$$\mathbf{f} = \mathbf{f}(\mathbf{w}, \mathbf{X}) = [f(\mathbf{w}, \mathbf{x}_1), \dots, f(\mathbf{w}, \mathbf{x}_m)]^T \in \mathbb{Y}^m.$$

Определим функцию ошибки

$$S: \mathbb{X} \times \mathbb{W} \times \mathbb{Y} \rightarrow \mathbb{R}_+$$

и представление множества индексов элементов выборки в виде

$$I = L \cup C.$$

Далее в качестве функции ошибки S зададим квадрат нормы вектора регрессионных остатков $\boldsymbol{\varepsilon}$:

$$S = \sum_{i=1}^m \varepsilon^2(\mathbf{x}_i) = \text{RSS}; \text{TSS} = \sum_{i=1}^m (y_i - \bar{y})^2, \quad (2)$$

где $\bar{y} = \frac{1}{m} \sum_{i=1}^m y_i$; RSS (residual sum of squares) — сум-

ма квадратов регрессионных остатков; TSS (total sum of squares) — полная сумма квадратов.

Требуется найти такой оптимальный вектор параметров $\mathbf{w}^* \in \mathbb{W}$, при котором функция приближает целевой вектор \mathbf{y} наилучшим образом в смысле функции ошибки S .

Назовем моделью пару (\mathbf{f}, A) , где $A \subset J$ — подмножество индексов признаков, используемое для вычисления вектор-функции \mathbf{f} . Ниже фиксирована функция $\mathbf{f} = \mathbf{X}\mathbf{w}$, после этого модель зависит только от множества A , поэтому вместо (\mathbf{f}, A) для обозначения применяемой модели будем использовать A . Таким образом, выбор модели сводится к нахождению оптимального множества индексов A^* в смысле функции ошибки S , вычисляемой на элементах выборки D_C :

$$A^* = \arg \min_{A \subset J} S(A|\mathbf{w}^*, D_C). \quad (3)$$

Для решения задачи (3) необходимо найти вектор параметров \mathbf{w}^* как решение задачи минимизации функции ошибки на элементах выборки D_L с индексами из множества L :

$$\mathbf{w}^* = \arg \min_{\mathbf{w} \in \mathbb{W}} S(\mathbf{w}|A, D_L). \quad (4)$$

Задача (3) является задачей выбора признаков и заключается в нахождении подмножества индексов признаков $A^* \subset J$, минимизирующего функцию ошибки S .

2. Анализ мультиколлинеарности при выборе признаков

В дальнейшем будем считать, что векторы признаков $\boldsymbol{\chi}_j$ и целевой вектор \mathbf{y} нормированы. Рассмотрим некоторое подмножество $B \subset J$ индексов признаков. Назовем признаки мультиколлинеарными,

если найдутся такие коэффициенты a_k , $k \in B$ и достаточно малое $\delta > 0$, что

$$\|\mathbf{x}_j - \sum_{k \in B} a_k \mathbf{x}_k\| < \delta, \quad (5)$$

где j — индекс признака и $j \notin B$. Чем меньше δ , тем выше степень мультиколлинеарности.

Назовем признаки с индексами i, j *коррелирующими*, если найдется достаточно малое δ_{ij} такое, что

$$\|\mathbf{x}_j - \mathbf{x}_i\| < \delta_{ij}. \quad (6)$$

Из определения следует, что и формула (6) есть частный случай формулы (5) при $\delta_{ji} = \delta_{ij}$ и $a_k = 1$, $k = j$.

Назовем признак \mathbf{x}_j *коррелированным с целевым вектором*, если найдется достаточно малое δ_{yj} , такое что

$$\|\mathbf{y} - \mathbf{x}_j\| < \delta_{yj}.$$

2.1. Фактор инфляции дисперсии

Широкоизвестным критерием анализа мультиколлинеарности авторы считают фактор инфляции дисперсии [18]. Фактор инфляции дисперсии VIF_j определяется для j -го признака и является показателем наличия линейной зависимости между j -м и остальными признаками. Для нахождения VIF_j необходимо определить оценку $\hat{\mathbf{w}}$ для вектора коэффициентов \mathbf{w} в задаче (1) при $y_i = x_{ij}$, $i \in I$ и $J = J \setminus j$. Аналогично (2) определяются RSS и TSS . Величина VIF_j определяется следующим выражением:

$$VIF_j = \frac{1}{1 - R_j^2}, \quad (7)$$

где $R_j^2 = 1 - \frac{RSS}{TSS}$ — коэффициент детерминации.

Согласно [18] значение $VIF_j \gtrsim 5$ означает наличие зависимости между j -м и всеми остальными признаками. Недостатками этого критерия мультиколлинеарности является то, что он может принимать большие значения сразу для нескольких признаков, что мешает определить какой из признаков необходимо удалить.

Другим критерием наличия мультиколлинеарности между признаками является число обусловленности к матрицы $\mathbf{X}^T \mathbf{X}$, которое равно отношению максимального и минимального по модулю собственных чисел λ_{\max} и λ_{\min} :

$$\kappa = \frac{\lambda_{\max}}{\lambda_{\min}}. \quad (8)$$

Оно показывает насколько матрица близка к вырожденной. Чем больше κ , тем ближе матрица к вырожденной.

2.2. Метод Белсли

Для обнаружения и исключения мультиколлинеарных признаков в наборе отобранных признаков предлагается явно поставить оптимизационную задачу, используя метод Белсли. Критерием сравнения методов выбора признаков в данной работе является критерий, основанный на исключении признака, мультиколлинеарного некоторым другим признакам из набора выбранных признаков. Исключение проводится методом Белсли. Предлагаемый критерий сравнения методов выбора признаков в дальнейшем называется критерием наличия мультиколлинеарных признаков среди отобранных признаков. Будем считать, что на множестве параметров \mathbb{W} задано нормальное распределение

$$\mathbf{w} \sim N(\mathbf{w}_{ML}, \mathbf{A}^{-1})$$

с матожиданием \mathbf{w}_{ML} и ковариационной матрицей \mathbf{A}^{-1} . Оценка $\hat{\mathbf{A}}^{-1}$ ковариационной матрицы в случае линейной модели будет

$$\hat{\mathbf{A}}^{-1} = (\mathbf{X}^T \mathbf{X})^{-1}.$$

Используя сингулярное разложение матрицы \mathbf{X}

$$\mathbf{X} = \mathbf{U} \mathbf{\Lambda} \mathbf{V}^T,$$

где \mathbf{U} и \mathbf{V} — ортогональные матрицы, а $\mathbf{\Lambda}$ — диагональная с собственными значениями λ_i на диагонали, такими что

$$\lambda_1 > \lambda_2 > \dots > \lambda_n,$$

получим выражение для $(\mathbf{X}^T \mathbf{X})^{-1}$:

$$(\mathbf{X}^T \mathbf{X})^{-1} = \mathbf{V} \mathbf{\Lambda}^{-2} \mathbf{V}^{-1}.$$

Столбцы матрицы \mathbf{V} — собственные векторы, а квадраты сингулярных чисел — собственные значения матрицы $\mathbf{X}^T \mathbf{X}$:

$$\mathbf{X}^T \mathbf{X} = \mathbf{V} \mathbf{\Lambda}^2 \mathbf{V}^T = \mathbf{V} \mathbf{\Lambda}^2 \mathbf{V}^T,$$

$$\mathbf{X}^T \mathbf{X} \mathbf{V} = \mathbf{V} \mathbf{\Lambda}^2.$$

Отношение максимального собственного значения λ_{\max} к i -му собственному значению λ_i назовем индексом обусловленности η_i :

$$\eta_i = \frac{\lambda_{\max}}{\lambda_i}.$$

Большое значение η_i указывает на зависимость, близкую к линейной, между признаками и чем больше η_i , тем сильнее зависимость. Поэтому на этапе удаления нужно найти такой индекс i^* , что

$$i^* = \arg \max_{i \in F_{k-1}} \eta_i$$

где F_{k-1} текущее подмножество признаков. Оценками дисперсий параметров будут диагональные элементы матрицы $X^T X$:

$$\text{Var}(w_i) = \sum_{j=1}^n \frac{v_{ij}^2}{\lambda_j^2}.$$

Далее определим дисперсионную долю q_{ij} как вклад j -го признака в дисперсию i -го элемента вектора параметров \mathbf{w} :

$$q_{ij} = \frac{v_{ij}^2 / \lambda_j^2}{\sum_{j=1}^n v_{ij}^2 / \lambda_j^2},$$

где $[v_{ij}] = \mathbf{V}$, а λ_j — собственное значение. Большие значения дисперсионных долей означают наличие зависимостей между признаками, это следует из способа их получения.

Следовательно, по найденному максимальному индексу обусловленности i^* находим признак j^* :

$$j^* = \arg \min_{i \in F_{k-1}} q_{i^*j}, \quad (9)$$

который вносит наибольший вклад в дисперсию i -го элемента вектора \mathbf{w} , т. е. является коллинеарным некоторому другому признаку.

3. Методы построения тестовых выборок

Для тестирования методов выбора признаков предлагается использовать синтетические выборки, которые определяются с помощью множеств P_f , P_y , C_f , C_y и R , индексирующих соответственно ортогональные признаки, признаки ортогональные целевому вектору, мультиколлинеарные признаки, признаки, коррелирующие с целевым вектором и случайно расположенные признаки. Определим следующие множества, задающие структуру выборки:

- 1) множество ортогональных признаков χ_j с индексами j из множества P_f ;
- 2) множество признаков χ_j , ортогональных целевому вектору \mathbf{y} , с индексами j из множества P_y ;
- 3) множество мультиколлинеарных признаков χ_j с индексами j из множества C_f ;
- 4) множество признаков χ_j , коррелирующих с целевым вектором, с индексами j из множества C_y ;
- 5) множество случайных признаков χ_j с индексами из множества R .

Для регулирования степени мультиколлинеарности используется параметр мультиколлинеарности k : при $k = 1$ признаки коллинеарны, при $k = 0$ — ортогональны.

При этом параметр k используется как для определения степени мультиколлинеарности признаков, так и для определения степени коррелированности признаков и целевого вектора.

Рассмотрим базовые варианты взаимного расположения мультиколлинеарных признаков и целевого вектора, из которых варьированием параметров можно генерировать различные выборки для тестирования методов выбора признаков.

1. Признаки χ_j с индексами как из множества мультиколлинеарных между собой признаков $j \in C_f$, так и из множества ортогональных целевому вектору \mathbf{y} , $j \in P_y$:

$$\langle \mathbf{y}, \chi_j \rangle = 0, j \in J; \|\chi_i - \sum_{l \in B} a_l \chi_l\| < \delta, i \in J, i \notin B \subset J; \quad (10)$$

$$J = P_y \cap C_f.$$

Схематично взаимное расположение признаков и целевого вектора изображено на рис. 1. Выборки с такой структурой будем называть выборками первого типа.

2. Все признаки χ_j порождены случайно из n -мерной случайной величины, $j \in R$. Эта случайная величина взята из равномерного распределения на единичном кубе размерности r . При этом найдется некоторый признак χ_i , $i \in R$, приближающийся целевой вектор \mathbf{y} :

$$J = R, |R| = r, \chi_1, \dots, \chi_r \sim U[0, 1]^r, \|\mathbf{y} - \chi_i\| < \delta. \quad (11)$$

Схематично взаимное расположение признаков и целевого вектора изображено на рис. 2. Выборки с такой структурой будем называть выборками второго типа.

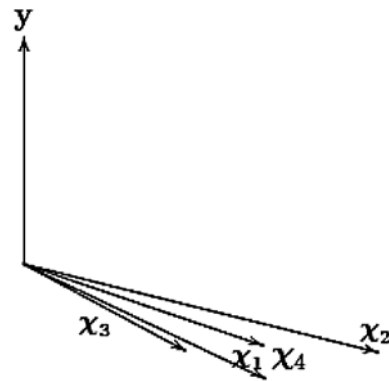


Рис. 1. Неадекватная коррелирующая выборка

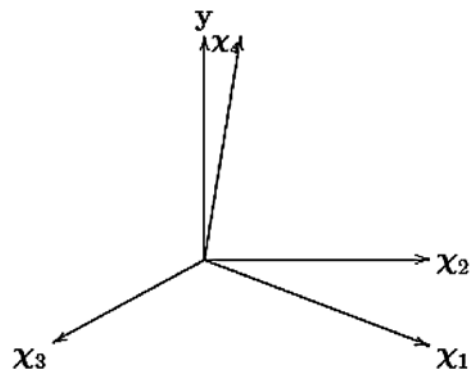


Рис. 2. Адекватная случайная выборка

3. Все признаки χ_j коррелируют и хорошо приближают целевой вектор \mathbf{y} :

$$\|\chi_j - \chi_i\| < \delta_{ij}, i, j \in J; \|\mathbf{y} - \chi_j\| < \delta, j \in J, J = C_y. \quad (12)$$

Схематично расположение признаков и целевого вектора изображено на рис. 3. Выборки с такой структурой будем называть выборками третьего типа.

4. Множество признаков χ_j с индексами из множества $j \in J$ состоит из объединения двух множеств: множества ортогональных признаков с индексами из множества P_f и множества признаков χ_c , коррелированных с некоторыми из них. Индексы c лежат в множестве C_f . При этом целевой вектор \mathbf{y} хорошо приближается линейной комбинацией ортогональных признаков $\chi_j, j \in J$:

$$\langle \chi_i, \chi_j \rangle = 0, i, j \in P_f; \mathbf{y} = \sum_{j \in P_f} a_j \chi_j; \|\chi_j - \chi_i\| < \delta_{ij}, i \in P_f, j \in C_f; J = P_f \cup C_f. \quad (13)$$

Схематично взаимное расположение признаков и целевого вектора изображено на рис. 4. Выборки с такой структурой будем называть выборками четвертого типа.

Комбинируя описанные выше варианты взаимного расположения признаков и целевого вектора, варьируя параметр мультиколлинеарности, а также изменяя мощности p_f, p_y, c_f, c_y и r множеств P_f, P_y, C_f, C_y и R , индексирующих множества признаков со свойствами, определенными в выражениях (10)–(13), можно генерировать выборки для тестирования методов выбора признаков.

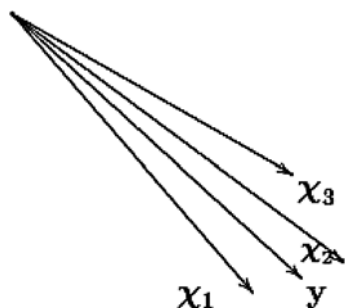


Рис. 3. Адекватная избыточная выборка

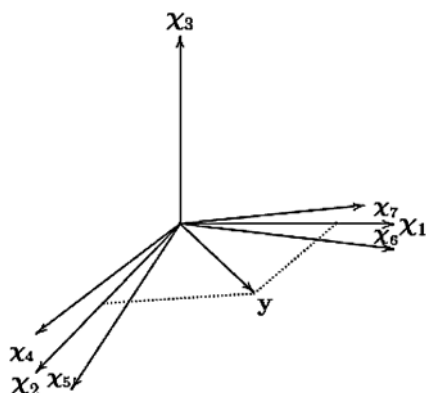


Рис. 4. Адекватная коррелирующая выборка

4. Критерии сравнения методов выбора признаков

Для анализа методов выбора признаков определим следующий критерий, позволяющий оценить сколько мультиколлинеарных признаков есть в множестве отобранных признаков. Зададим некоторое предельное значение s_0 функции ошибки S . Результатом работы метода выбора признаков является набор признаков с индексами из множества $A \subset J, p = |A|$. Для найденного множества признаков получен оптимальный вектор параметров \mathbf{w}_A^* . Назовем h максимальную мощность множества индексов признаков $J_h \subset A$, при удалении которого значение функции ошибки S не превосходит s_0 :

$$h = \arg \max_{S(J_h, \mathbf{w}_h, D) < s_0} |J_h|, \quad (14)$$

где $S(J_h, \mathbf{w}_h, D)$ — функция ошибки, в которой первый аргумент — это матрица \mathbf{X} со столбцами, индексы которых лежат в множестве J_h , второй аргумент — вектор параметров \mathbf{w}_h , составленный из элементов \mathbf{w}_A^* с индексами из множества J_h , и третий аргумент — выборка. Ниже в разделе "Вычислительный эксперимент" определялась величина d , значение которой равно числу признаков, удаление их приводит к ошибке, не превышающей s_0 :

$$d = |A| - h. \quad (15)$$

Определение индексов удаляемых признаков проводилось методом Белсли, задача (9). Методы выбора признаков ранжируются по возрастанию величины d : большие значения d показывают, что выбранное подмножество признаков (решение задачи (3)) содержит избыточные признаки, удаление которых приводит к росту функции ошибки вплоть до s_0 .

Ранее авторами [18, 19] были предложены следующие критерии сравнения линейных регрессионных моделей.

1. Скорректированный (adjusted) коэффициент детерминации R_{adj}^2 учитывает добавление избыточных признаков и выражается как

$$R_{\text{adj}}^2 = 1 - \frac{\text{RSS}/(m-p)}{\text{TSS}/(m-1)}, \quad (16)$$

где $p = |A|$, m — число строк в матрице \mathbf{X} , а RSS и TSS определяются из (2). Чем ближе значение к единице, тем лучше модель описывает целевой вектор.

2. Критерий C_p позволяет достичь компромисса между значением RSS и числом используемых переменных $p = |A|$, а также ликвидировать возможную коллинеарность признаков. Величина C_p определяется следующим образом:

$$C_p = \frac{\text{RSS}_A}{\text{RSS}} - m + 2p, \quad (17)$$

где RSS_A — это величина, аналогичная RSS , но найденная при использовании признаков с индексами из множества A . Меньшие значения соответствуют лучшему набору признаков.

3. Информационный критерий BIC вычисляется по следующей формуле:

$$BIC = RSS + p \log m, \quad (18)$$

где $p = |A|$ и m — это количество строк в матрице X . Чем меньше величина BIC, тем лучше модель описывает целевой вектор.

4. F -тест используется в случае линейной модели для проверки отсутствия релевантных признаков. Если ни один из признаков не приближает целевой вектор, то величина

$$\frac{(TSS - RSS)/p}{RSS/(n - p - 1)} \sim F_{p, n-p-1} \quad (19)$$

имеет распределение Фишера с p , $n - p - 1$ степенями свободы.

5. Вычислительный эксперимент

В вычислительном эксперименте проведено сравнение методов выбора признаков по различным функционалам качества при фиксированном значении предельной функции ошибки $s_0 = 0,5$ и при двух значениях параметра мультиколлинеарности $k = 0,2$ и $k = 0,8$. Для каждой выборки и для каждого метода выбора признаков были получены зависимости между предельным значением функции ошибки s_0 и максимальным числом d , а также между фактором инфляции дисперсии VIF и параметром мультиколлинеарности k . При этом VIF определялся для признаков из множества A , что показывает наличие мультиколлинеарных признаков в множестве отобранных признаков A . Эксперименты проводи-

ли на выборках при $k = 0,2$ и $k = 0,8$. Для выборок второго типа график зависимости VIF от параметра мультиколлинеарности k и числа избыточных признаков d в множестве отобранных признаков от предельного значения функции ошибки s_0 не строили, так как в этом типе выборок нет мультиколлинеарных признаков.

Результаты экспериментов сведены в табл. 1—7.

В экспериментах генерировались выборки четырех типов, определяемых формулами (10)—(13) для двух значений параметра мультиколлинеарности $k = 0,2$ и $k = 0,8$. Перед проведением экспериментов векторы признаков и целевой вектор были отнормированы, так что евклидова норма векторов признаков и целевого вектора равна единице. Измеряемые значения критериев усреднены по пяти повторениям. Значения элементов вектора w , меньшие 10^{-6} , считались незначительными и равными нулю. Значения p -value соответствуют проверке нулевой гипотезы о том, что вектор параметров w — нулевой, т. е. отсутствуют признаки, с помощью которых можно приблизить целевой вектор y , против альтернативы, что среди столбцов матрицы X есть подходящие для описания целевого вектора y , при уровне значимости 0,05. Если значение p -value меньше 0,05, то нулевая гипотеза отвергается. Это означает, что среди признаков есть такие, которые хорошо приближают целевой вектор y . Проверка выполняется с помощью F -теста (19). В таблицах, где отсутствует столбец со значениями p -value, они пренебрежимо малы. Значение предельной функции ошибки $s_0 = 0,5$.

Сравнивали методы LARS, Lasso, ElasticNet, Ridge и Stepwise. Все, кроме последнего, являются методами, которые одновременно решают задачи (4) и (3). Отбор признаков проводится обнулением незначительных коэффициентов в оптимальном векторе пара-

Таблица 1

Значения функционалов качества для выборок первого типа при $k = 0,2$

Метод	d	C_p	RSS	κ	VIF	R_{adj}^2	BIC	p -value
Lasso	0	−997	1	3,84	1,05	−3,32	314,62	0,11
Ridge	0	−997	1	4,13	1,05	−3,31	346,39	0,1
LARS	—	−997	—	—	—	—	—	—
Stepwise	0	−997	1	4,13	1,05	−3,41	346,41	$5,28 \cdot 10^{-4}$
Elastic Net	0	−997	1	3,84	1,05	−3,32	314,32	0,11

Таблица 2

Значения функционалов качества для выборок первого типа при $k = 0,8$

Метод	d	C_p	RSS	κ	VIF	R_{adj}^2	BIC	p -value
Lasso	0	−997	1	717,8	16,6	−3,32	310,48	0,06
Ridge	0	−997	1	801	16,6	−3,31	346,39	0,05
LARS	—	−997	—	—	—	—	—	—
Stepwise	0	−997	1,68	801	16,6	−6,22	347,01	10^{-10}
Elastic Net	0	−997	1	717,8	16,6	−3,32	310,48	0,06

Значения функционалов качества для выборок второго типа

Метод	d	C_p	RSS	κ	VIF	R_{adj}^2	BIC
Lasso	0	$7 \cdot 10^6$	$8,50 \cdot 10^{-4}$	1	0,25	1	6,9
Elastic Net	0	$8,76 \cdot 10^{-4}$	$8,76 \cdot 10^{-4}$	1	0,25	1	6,9
Ridge	0	$7,97 \cdot 10^9$	0,97	1	0,25	-3	7,88
LARS	0,2	-997	$1,30 \cdot 10^{-10}$	2,19	0,32	1	8,29
Stepwise	4,6	-997	$1,33 \cdot 10^{-10}$	28,86	0,89	1	53,88

Значения функционалов качества для выборок третьего типа при $k = 0,2$

Метод	d	C_p	RSS	$\kappa \cdot 10^8$	$VIF \cdot 10^7$	R_{adj}^2	BIC
Ridge	0	$2,3 \cdot 10^9$	0,97	24	1,14	-3,17	346,36
Lasso	1	$2 \cdot 10^6$	$8,50 \cdot 10^{-4}$	0,95	0,58	1	13,82
Elastic Net	3,2	$2 \cdot 10^6$	$8,50 \cdot 10^{-4}$	2,8	0,97	1	41,45
LARS	36	-997	$4,22 \cdot 10^{-10}$	24	1,14	1	345,39
Stepwise	36	-997	$4,22 \cdot 10^{-10}$	24	1,14	1	345,39

Значения функционалов качества для выборок третьего типа при $k = 0,8$

Метод	d	C_p	RSS	κ	VIF	R_{adj}^2	BIC
Lasso	0	$5,16 \cdot 10^8$	$8,50 \cdot 10^{-4}$	1	0,24	1	6,9
Ridge	0	$5,9 \cdot 10^{11}$	0,97	$6,07 \cdot 10^{11}$	$2,9 \cdot 10^9$	-3,17	346,36
Elastic Net	3,2	$5,16 \cdot 10^8$	$8,50 \cdot 10^{-4}$	$7,3 \cdot 10^{10}$	$2,5 \cdot 10^9$	1	41,45
LARS	36	-997	$1,65 \cdot 10^{-12}$	$6,07 \cdot 10^{11}$	$2,9 \cdot 10^9$	1	345,39
Stepwise	36	-997	$1,73 \cdot 10^{-12}$	$6,07 \cdot 10^{11}$	$2,9 \cdot 10^9$	1	345,39

Значения функционалов качества для выборок четвёртого типа при $k = 0,2$

Метод	d	C_p	RSS	κ	VIF	R_{adj}^2	BIC
Ridge	0	$6 \cdot 10^{30}$	0,95	$8,42 \cdot 10^{15}$	$1,15 \cdot 10^{23}$	-3	210,95
Stepwise	1	-868,95	$5,45 \cdot 10^{-29}$	1	0,63	1	13,82
LARS	1,8	$5,38 \cdot 10^{29}$	0,38	$2,1 \cdot 10^{16}$	$3,3 \cdot 10^{30}$	-0,62	102,62
Elastic Net	17,6	$5,84 \cdot 10^{27}$	$9,18 \cdot 10^{-4}$	$1,4 \cdot 10^{16}$	$5,32 \cdot 10^{20}$	1	150,59
Lasso	18	$5,84 \cdot 10^{27}$	$9,18 \cdot 10^{-4}$	$1,4 \cdot 10^{16}$	$5,32 \cdot 10^{20}$	1	150,60

Значения функционалов качества для выборок четвёртого типа при $k = 0,8$

Метод	d	C_p	RSS	κ	VIF	R_{adj}^2	BIC
Ridge	0	$1,8 \cdot 10^{30}$	0,95	10^{16}	$8,65 \cdot 10^{16}$	-2,97	152,92
Stepwise	1	$9,4 \cdot 10^5$	$8,8 \cdot 10^{-25}$	1	0,63	1	13,82
LARS	1,2	10^{30}	0,38	$3 \cdot 10^{29}$	10^{20}	-0,57	108,15
Lasso	14,8	$1,73 \cdot 10^{27}$	$9,2 \cdot 10^{-4}$	$9,92 \cdot 10^{15}$	10^{17}	1	150,59
Elastic Net	15,2	$1,70 \cdot 10^{27}$	$9,2 \cdot 10^{-4}$	$9,92 \cdot 10^{15}$	10^{17}	1	150,59

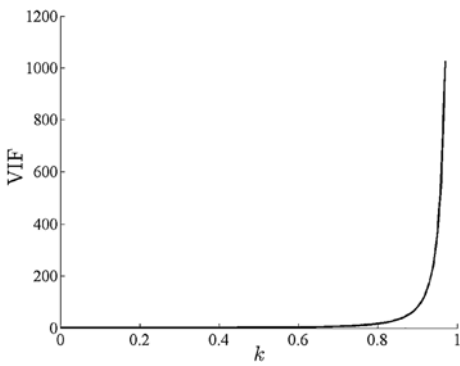


Рис. 5. Зависимость фактора инфляции дисперсии VIF от параметра мультиколлинеарности k для первого типа выборки

метров \mathbf{w}^* . Метод Stepwise последовательно решает задачи (3) и (4), добавляя и удаляя признаки в соответствии с их значимостью, определяемой статистическим тестом. В алгоритме ElasticNet используется взвешенная сумма регуляризаторов из алгоритмов Lasso и Ridge, веса у обоих регуляризаторов равны 0,5. Прочерк в таблице означает, что метод выбора признаков не отбирает ни один признак и получаемый вектор \mathbf{w}^* нулевой.

В столбцах таблиц стоят ранее введенные критерии качества модели: число мультиколлинеарных признаков d в множестве отобранных признаков (15); критерий C_p (17); остаточная сумма квадратов RSS (2); число обусловленности к матрицы $\mathbf{X}^T\mathbf{X}$ (8); значение VIF (7); скорректированный коэффициент детерминации R_{adj}^2 (16); информационный критерий BIC (18) и p -value для F -теста (19).

Для выборок первого типа (10) $n = p_y = 50$, $m = 1000$ результаты приведены в табл. 1 и 2 при $k = 0,2$ и $k = 0,8$ соответственно.

Для выборок второго типа (11) $n = r = 50$, $m = 1000$ результаты приведены в табл. 3.

Для выборок третьего типа (12) $n = c_y = 50$, $m = 1000$ результаты приведены в табл. 4 и 5 при $k = 0,2$ и $k = 0,8$ соответственно.

Для выборок четвертого типа (13) $p_f = 10$, $c_f = 40$, $m = 1000$ результаты приведены в табл. 6 и 7 при $k = 0,2$ и $k = 0,8$ соответственно.

На рис. 5–8, представлена зависимость VIF от параметра мультиколлинеарности k для каждого типа выборок, где эта зависимость имеет место.

На рис. 5 показана зависимость VIF от параметра мультиколлинеарности k для первого типа выборок при работе

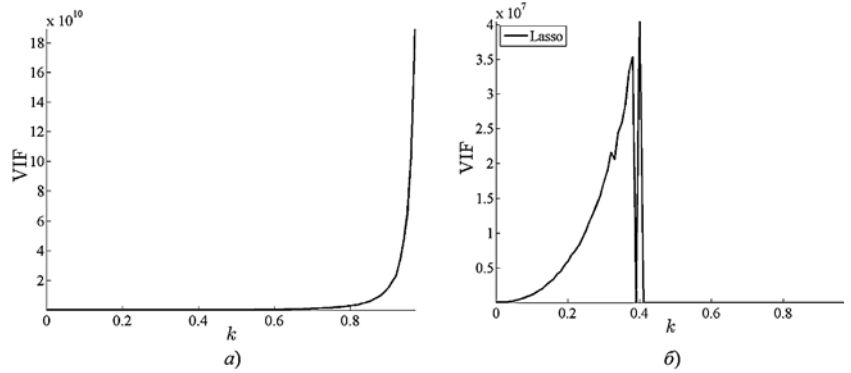


Рис. 6. Зависимость фактора инфляции дисперсии VIF от параметра мультиколлинеарности k для третьего типа выборок: а — при работе всех рассматриваемых методов отбора кроме Lasso; б — при работе метода Lasso

различных алгоритмов. На этом рисунке видно, что все алгоритмы показывают одинаковые результаты, и ни один из рассматриваемых методов выбора признаков не решает проблему мультиколлинеарности в случае ортогональности всех признаков целевому вектору и взаимной коррелированности.

На рис. 6 изображена зависимость VIF от параметра мультиколлинеарности k для третьего типа выборок. Видно, что все методы показывают одинаковый вид зависимости, кроме метода Lasso. Для него при росте параметра мультиколлинеарности наблю-

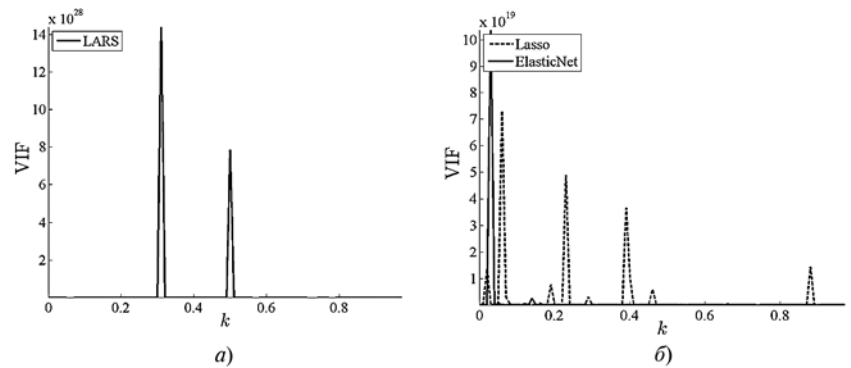


Рис. 7. Зависимость фактора инфляции дисперсии VIF от параметра мультиколлинеарности k для четвертого типа выборок: а — при работе метода LARS; б — при работе методов Lasso и Elastic Net

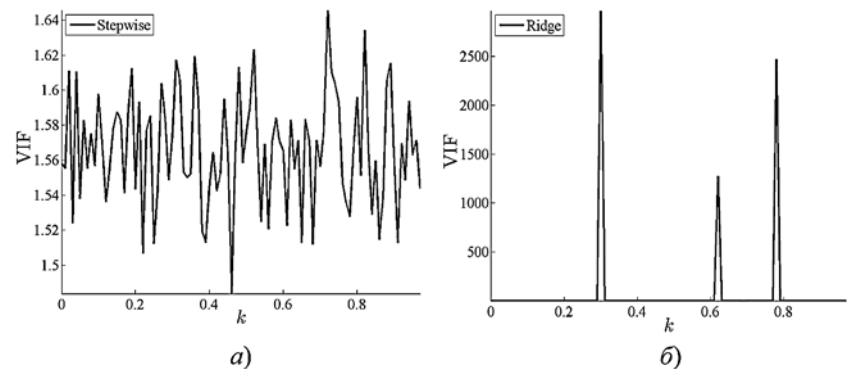
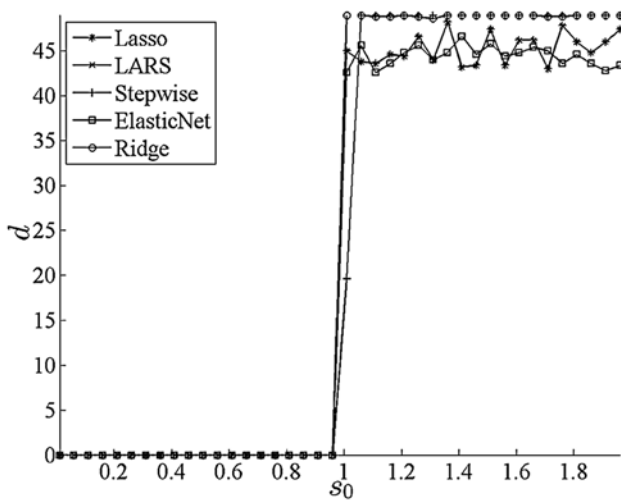


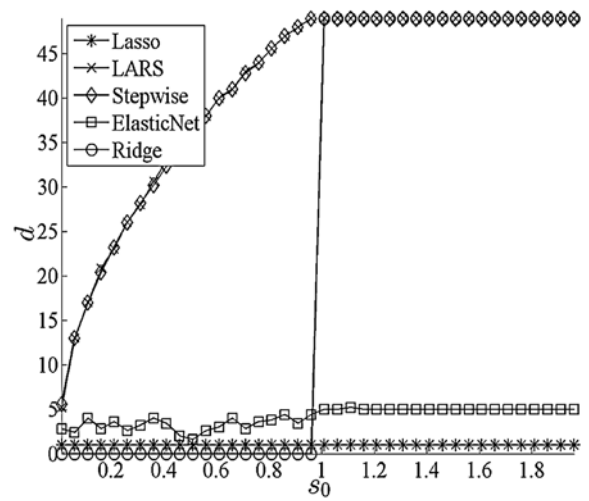
Рис. 8. Зависимость фактора инфляции дисперсии VIF от параметра мультиколлинеарности k для четвертого типа выборок: а — при работе метода Stepwise; б — при работе метода Ridge

дается резкое уменьшение значения величины VIF. Это говорит об отсутствии линейной зависимости между выбранными признаками в выборках, сгенерированных при $k \geq 0,4$.

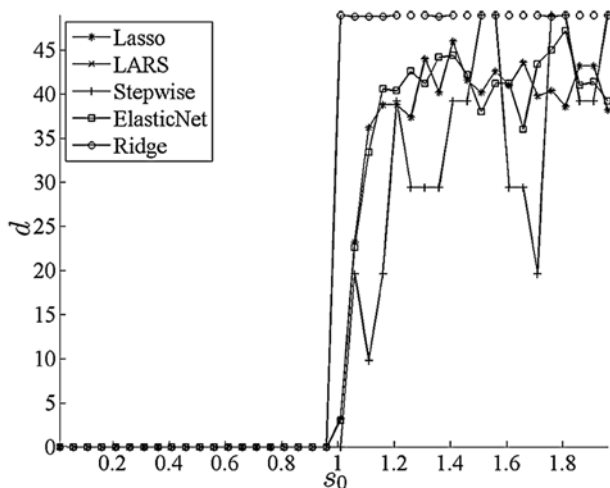
На рис. 7, 8 показана зависимость VIF от параметра мультиколлинеарности k для четвертого типа выборок при работе различных методов. Метод LARS показывает резкие скачки значений VIF, как и метод Ridge, но у метода Ridge амплитуда скачков ниже. Методы Lasso и ElasticNet демонстрируют скачки, схожие со скачками у LARS, но меньшей амплитуды и более высокой частоты. Для выборок четвертого типа после применения метода Stepwise значения VIF не превышают двух при росте коэффициента k . Это означает, что метод Stepwise для выборок четвертого типа дает набор линейно независимых признаков.



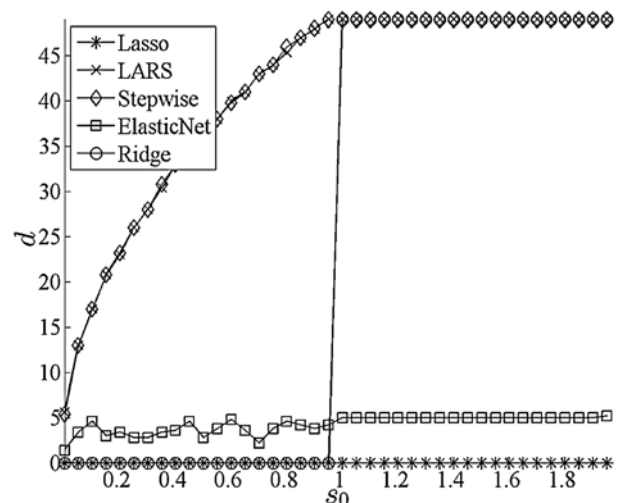
а)



а)



б)



б)

Рис. 9. Зависимость числа мультиколлинеарных признаков d в множестве отобранных признаков от предельного значения функции ошибки s_0 для первого типа выборок:

а — при $k = 0,2$; б — при $k = 0,8$

Рассмотрим зависимость числа мультиколлинеарных признаков d в множестве отобранных признаков от значений предельной ошибки s_0 для ранее рассмотренных типов выборок на рис. 9—11.

На рис. 9 показана зависимость числа лишних признаков d в множестве отобранных признаков от предельного значения функции ошибки s_0 для первого типа выборок при значениях $k = 0,2$ и $k = 0,8$. Значение d стабильно равно нулю вследствие ортогональности целевого вектора и всех признаков вплоть до значений, близких к единице. Далее идет резкий скачок d , так как предельное значение функции ошибки выросло достаточно, чтобы удалить сразу почти все признаки.

На рис. 10 показана зависимость величины d от параметра s_0 для третьего типа выборок при значениях $k = 0,2$ и $k = 0,8$. Метод Lasso отбирает один

Рис. 10. Зависимость числа мультиколлинеарных признаков d в множестве отобранных признаков от предельного значения функции ошибки s_0 для третьего типа выборок:

а — при $k = 0,2$; б — при $k = 0,8$

или два признака, наилучшим образом приближающие целевой вектор, поэтому значение d для этого метода равно нулю или единице. Аналогично, но чуть хуже, работает метод Elastic Net, он отбирает чуть больше лишних признаков нежели метод Lasso. Зависимость d от s_0 для метода Ridge схожа с зависимостью для первого типа выборок по той же причине: сначала она достаточно велика, чтобы удалить хоть один признак, но как только приближается к единице становится возможным удалить сразу почти все признаки. Для методов LARS и Stepwise наблюдается постепенный рост величины d с ростом предельного значения функции ошибки s_0 с выходом на константу при достижении s_0 значения, близкого к единице.

На рис. 11 показана зависимость d от параметра s_0 для четвертого типа выборок при $k = 0,2$ и $k = 0,8$.

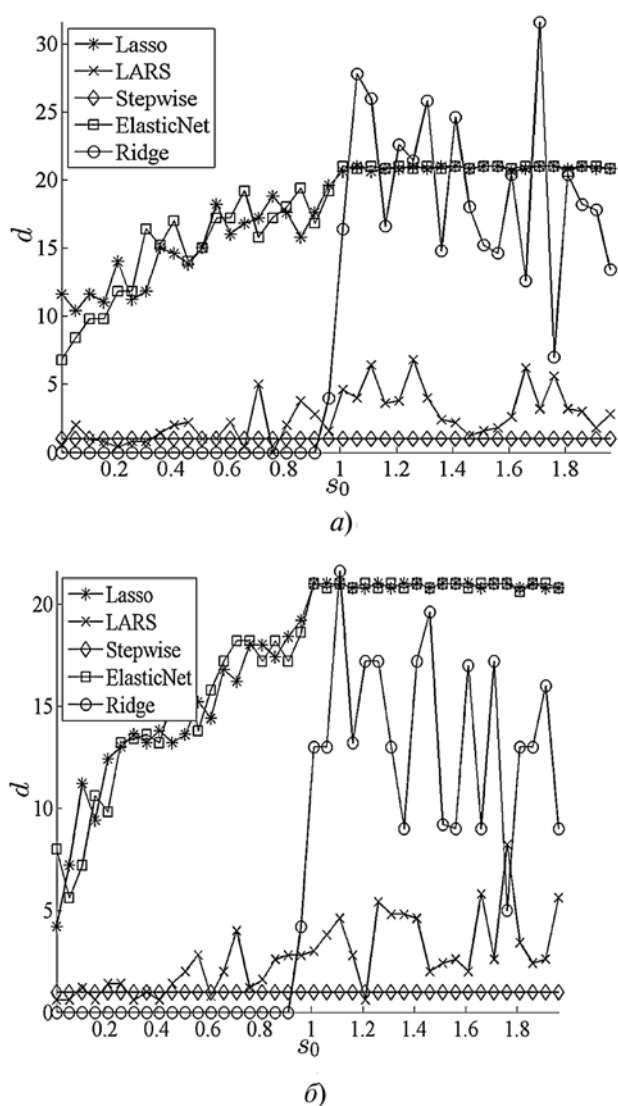


Рис. 11. Зависимость числа мультиколлинеарных признаков d в множестве отобранных признаков от предельного значения функции ошибки s_0 для четвертого типа выборок:

a — при $k = 0,2$; b — при $k = 0,8$

Наиболее стабильные решения дает метод Stepwise, у которого в среднем обнаруживается только один признак, удаление которого приводит к ошибке, не превышающей s_0 . Чуть хуже работает метод LARS: число лишних признаков d среди отобранных им не превышает пяти при росте предельного значения функции ошибки s_0 . Для методов Lasso и ElasticNet наблюдается рост d при росте s_0 до единицы, а затем стабилизация на уровне $d \approx 20$. Для метода Ridge вид зависимости схож с предыдущими типами выборок, только для четвертого типа после преодоления s_0 значения, равного единице, величина d начала сильно колебаться. Это показывает неустойчивость набора признаков, получаемого методом Ridge для четвертого типа выборок.

Заключение

В работе проведено исследование эффективности методов выбора признаков в случае выборок с мультиколлинеарными признаками. Эксперименты показали, что из рассмотренных методов проблему мультиколлинеарности при отборе признаков решают методы Lasso (для выборок третьего типа) и Stepwise (для выборок четвертого типа). Для выборок первого типа все рассмотренные методы показывают одинаковые результаты: ни один из рассматриваемых методов выбора признаков не решает проблему мультиколлинеарности в случае ортогональности всех признаков целевому вектору. Предложенный критерий показывает, что как при малых, так и при больших значениях k устойчивые решения дают одинаковые методы. Также вид зависимости между величинами s_0 и d практически одинаков в рамках одной выборки для больших и маленьких значений k . Для выборок первого типа все рассматриваемые методы показывают одинаковый результат, для выборок третьего типа наиболее устойчивый результат дает метод Lasso, для выборок четвертого типа — методы LARS и Stepwise.

Работа выполнена при поддержке РФФИ, проект 14-07-31046.

Список литературы

1. **Askin R. G.** Multicollinearity in regression: Review and examples // Journal of Forecasting. 1982. V. 1, N. 3. P. 281—292.
2. **Learner E. E.** Multicollinearity: A Bayesian Interpretation // The Review of Economics and Statistics. 1973. V. 55, N. 3. P. 371—380.
3. **Belsley D. A., Kuh E., Welsch R. E.** Regression diagnostics: Identifying influential data and sources of collinearity. New York: John Wiley & Sons, 2005.
4. **Yu Lei, Liu Huan.** Feature selection for high-dimensional data: A fast correlation-based filter solution // ICML. Washington D. C. 2003. V. 3. P. 856—863.
5. **Стрижов В. В., Кузнецов М. П., Рудаков К. В.** Метрическая кластеризация последовательностей аминокислотных остатков в ранговых шкалах // Математическая биология и биоинформатика. 2012. Т. 7. № 1. С. 345—359.
6. **Chen Yi-Wei, Lin Chih-Jen.** Combining SVMs with various feature selection strategies // Feature Extraction. Foundations and Applications. Berlin: Springer, 2006. P. 315—324.

7. **George H. J., Kohavi R., Pfleger K.** et al. Irrelevant Features and the Subset Selection Problem // *ICML*. New Brunswick. 1994. V. 94. P. 121–129.
8. **Vorontsov K.** Combinatorial probability and the tightness of generalization bounds // *Pattern Recognition and Image Analysis*. 2008. V. 18, N. 2. P. 243–259.
9. **Chong Il-Gyo, Jun Chi-Hyuck.** Performance of some variable selection methods when multicollinearity is present // *Chemometrics and Intelligent Laboratory Systems*. 2005. V. 78, N. 1–2. P. 103–112.
10. **Guyon I., Elisseeff A.** An Introduction to Variable and Feature Selection // *The Journal of Machine Learning Research*. 2003. V. 3. P. 1157–1182.
11. **Bolón-Canedo V., Sánchez-Marroño N., Alonso-Betanzos A.** A review of feature selection methods on synthetic data // *Knowledge and information systems*. 2013. V. 34, N. 3. P. 483–519.
12. **Ladha L., Deepa T.** Feature Selection Methods and Algorithms // *International Journal on Computer Science & Engineering*. 2011. V. 3, N. 5.
13. **El-Dereeny M., Rashwan N. I.** Solving multicollinearity problem using ridge regression models // *International Journal of Contemporary Mathematical Sciences*. 2011. V. 6. P. 585–600.
14. **Tibshirani R.** Regression Shrinkage and Selection Via the Lasso // *Journal of the Royal Statistical Society, Series B*. — 1994. — V. 58. — P. 267–288.
15. **Efron B., Hastie T., Tibshirani R.** Least angle regression // *The Annals of statistics*. 2004. V. 32, N. 2. P. 407–499.
16. **Zou Hui, Hastie T.** Regularization and variable selection via the Elastic Net // *Journal of the Royal Statistical Society, Series B*. — 2005. — V. 67. — P. 301–320.
17. **Harrell F. E.** *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. — Berlin: Springer, 2001.
18. **Paul R. K.** Multicollinearity: causes, effects and remedies // Accessed. Apr. 23, 2013. URL: <http://bit.ly/lqDHObV>.
19. **Strijov V., Krymova E., Weber G.-W.** Evidence optimization for consequently generated models. // *Mathematical and Computer Modeling*. 2013. V. 57, N. 1–2. P. 50–56.

A. M. Katrutsa, Student,

Moscow Institute of Physics and Technology, Moscow, amkatrutsa@yandex.ru,

V. V. Strijov, Rechecher,

Dorodnicyn Computing Center of Russian Academy of Sciences, Moscow, strijov@ccas.com

The Multicollinearity Problem for Feature Selection Methods in Regression

The paper investigates the multicollinearity problem in regression analysis and its influence on the performance of feature selection methods. The authors propose a procedure to test feature selection methods. A criteria is proposed to compare the feature selection methods, according to their performance when the multicollinearity is present. The feature selection methods are compared according to the other well-known evaluation measures. Methods to generate data sets of different multicollinearity types were proposed. The authors investigate performance of feature selection methods. The feature selection methods were tested on the data sets of different multicollinearity types.

Keywords: regression analysis, feature selection, multicollinearity, test data sets

References

1. **Askin R. G.** Multicollinearity in regression: Review and examples. *Journal of Forecasting*. 1982. V. 1, N. 3. P. 281–292.
2. **Learner E. E.** Multicollinearity: A Bayesian Interpretation. *The Review of Economics and Statistics*. 1973. V. 55, N. 3. P. 371–380.
3. **Belsley D. A., Kuh E., Welsch R. E.** *Regression diagnostics: Identifying influential data and sources of collinearity*. New York: John Wiley & Sons, 2005.
4. **Yu Lei, Liu Huan.** Feature selection for high-dimensional data: A fast correlation-based filter solution. *ICML*. Washington D. C., 2003. V. 3. P. 856–863.
5. **Strijov V. V., Kuznetsov M. P., Rudakov K. V.** Metricheskaya klasterizatsiya posledovatel'nostey aminokislotnykh ostatkov v rangovykh shkalakh. *Matematicheskaya biologiya i bioinformatika*. 2012. V. 7, N. 1. P. 345–359.
6. **Chen Yi-Wei, Lin Chih-Jen.** Combining SVMs with various feature selection strategies. *Feature Extraction. Foundations and Applications*. Berlin: Springer, 2006. P. 315–324.
7. **George H. J., Kohavi R., Pfleger K.** et al. Irrelevant Features and the Subset Selection Problem. *ICML*. New Brunswick, 1994. V. 94. P. 121–129.
8. **Vorontsov K.** Combinatorial probability and the tightness of generalization bounds. *Pattern Recognition and Image Analysis*. 2008. V. 18, N. 2. P. 243–259.
9. **Chong Il-Gyo, Jun Chi-Hyuck.** Performance of some variable selection methods when multicollinearity is present. *Chemometrics and Intelligent Laboratory Systems*. 2005. V. 78, N. 1–2. P. 103–112.
10. **Guyon I., Elisseeff A.** An Introduction to Variable and Feature Selection. *The Journal of Machine Learning Research*. 2003. V. 3. P. 1157–1182.
11. **Bolón-Canedo V., Sánchez-Marroño N., Alonso-Betanzos A.** A review of feature selection methods on synthetic data. *Knowledge and information systems*. 2013. V. 34, N. 3. P. 483–519.
12. **Ladha L., Deepa T.** Feature Selection Methods and Algorithms. *International Journal on Computer Science & Engineering*. V. 3, N. 5. 2011.
13. **El-Dereeny M., Rashwan N. I.** Solving multicollinearity problem using ridge regression models. *International Journal of Contemporary Mathematical Sciences*. 2011. V. 6. P. 585–600.
14. **Tibshirani R.** Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society, Series B*. 1994. V. 58. P. 267–288.
15. **Efron B., Hastie T., Tibshirani R.** Least angle regression. *The Annals of Statistics*. 2004. V. 32, N. 2. P. 407–499.
16. **Zou Hui, Hastie T.** Regularization and variable selection via the Elastic Net. *Journal of the Royal Statistical Society, Series B*. 2005. V. 67. P. 301–320.
17. **Harrell F. E.** *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. Berlin: Springer, 2001.
18. **Paul R. K.** Multicollinearity: Causes, Effects and Remedies. — Accessed Apr. 23, 2013, URL: <http://bit.ly/lqDHObV>.
19. **Strijov V., Krymova E., Weber G.-W.** Evidence optimization for consequently generated models. *Mathematical and Computer Modeling*. 2013. V. 57, N. 1–2. P. 50–56.

Я. Н. Имамвердиев¹, канд. техн. наук, науч. сотр., зав. отд., e-mail: yadigar@lan.ab.az,

Л. В. Сухостат², e-mail: lsuhostat@hotmail.com,

Институт информационных технологий НАНА, Баку, Азербайджан

Разработка робастного метода извлечения речевых признаков на основе эмпирического вейвлет-преобразования

Извлечение векторов признаков речевого сигнала является важным этапом для систем распознавания диктора. В настоящее время остаются актуальными работы по поиску информативных признаков речевого сигнала, обеспечивающих его адекватное описание и низкий процент ошибок при распознавании. В данной работе представлен подход для извлечения речевых признаков на основе эмпирического вейвлет-преобразования, повышающего точность распознавания, сохраняя при этом приемлемые показатели по вычислительной трудоемкости.

Ключевые слова: распознавание диктора, эмпирическое вейвлет-преобразование, дискретный алгоритм деления энергии, мгновенная амплитуда, мгновенная частота

Введение

Существует множество голосовых признаков, характеризующих диктора. Речь является сложным сигналом, возникающим в результате нескольких преобразований, происходящих на различных уровнях: семантическом, лингвистическом, артикуляционном и акустическом [1]. Различия, связанные с диктором, являются результатом сочетания анатомических различий, присущих вокальному тракту и манере разговора разных людей. Все эти различия могут быть использованы при распознавании диктора.

В настоящее время не существует формальной процедуры получения системы информативных признаков речевого сигнала, обеспечивающих качественное распознавание диктора. Обычно их выбирают исключительно на основе опыта и интуиции специалиста. Затем из полученной таким образом исходной системы признаков тем или иным формальным способом выбирают более экономичную и наиболее информативную подсистему описания речевого сигнала.

Исследования физики голосового аппарата [2], периферической слуховой системы, опытов по чтению динамических спектрограмм речевого сигнала, называемых видимой речью, и различных психофизических экспериментов показывают, что передача информации в речевом сигнале реализуется изменениями его кратковременного амплитудного спектра.

Цель выделения признаков заключается в преобразовании сигнала речи к некоторому типу параметрического представления для дальнейшего анализа и обработки. Кратковременные спектральные признаки наиболее часто применяют в задачах распознавания диктора и речи. В отличие от признаков высокого уровня, требующих более сложной пред-

варительной обработки [2, 3], их легче вычислить и получить хорошие результаты [4]. Кепстральные признаки тесно связаны с лингвистическим содержанием речи. Помимо кепстральных особенностей, речь имеет и источник возбуждения, который, как полагают, содержит полезные свойства для распознавания диктора. Кроме того, в реальных ситуациях существуют большие различия между этапами разработки и практического применения системы распознавания диктора. Как следствие, кепстральные признаки недостаточны, чтобы обеспечить удовлетворительную и надежную точность распознавания диктора. Они также не учитывают нестационарность и нелинейность человеческой речи.

Данная работа направлена на исследование новых и эффективных параметров для робастного распознавания диктора и предлагает метод извлечения признаков речевого сигнала для задачи распознавания диктора на основе эмпирического вейвлет-преобразования (Empirical Wavelet Transform, EWT).

1. Краткий обзор методов извлечения признаков речевого сигнала

Многие исследования были посвящены разработке различных схем извлечения характерных для диктора акустических признаков из речевых высказываний. J. Wolf в работе [5] среди наиболее существенных параметров выделяет частоту основного тона, спектральные признаки гласных и назальных согласных, оценку голосового источника, продолжительность слова и время начала "озвончения" (*voice onset time*). В работе [6] авторы сделали обзор и обобщили основные особенности речи, которые были использованы для системы распознавания диктора. Наряду с классическими и ведущими признаками были приведены некоторые недавно полученные наборы параметров.

Свойства идеальных речевых признаков, применяемых в системах распознавания диктора [5, 7]:

- большая вариабельность между дикторами и небольшая изменчивость у каждого диктора;
- устойчивость к фоновому шуму и искажениям;
- частое использование в обычной речи;
- простота в измерении;
- стабильность во времени и независимость от здоровья/настроения говорящего;
- трудно имитируемые.

Признаки, как правило, классифицируются на пять групп с точки зрения их физической интерпретации [8]: спектральные, спектрально-временные признаки голосового источника, просодические и признаки высокого уровня.

Имеется множество спектральных признаков, характеризующих речевой сигнал в задачах распознавания речи и диктора. Среди них можно выделить коэффициенты линейного предсказания (*linear prediction coefficients*, LPC) [9], кепстральные коэффициенты линейного предсказания (*linear prediction cepstral coefficients*, LPCC), кепстральные коэффициенты по шкале мел (*mel-frequency cepstral coefficients*, MFCC), которые были впервые применены к распознаванию диктора [10] и другие.

MFCC-признаки являются наиболее известными и популярными спектральными признаками. MFCC- и LPCC-признаки первоначально были разработаны для распознавания речи и основаны на линейной модели источник — фильтр для системы речеобразования.

В последние годы при описании и анализе свойств речи был использован подход на основе АМ — FM (Amplitude Modulation — Frequency Modulation) моделирования.

Монокомпонент АМ — FM сигнала описывается уравнением

$$x(n) = A(n)\cos[\Theta_n], \quad (1)$$

где $A(n)$ — мгновенная амплитуда монокомпонентного сигнала; Θ_n — мгновенная фаза. При этом многокомпонентный сигнал сначала разлагается, каждый его компонент описывается мгновенной огибающей и мгновенной частотой. Данный подход показывает значительные улучшения показателей распознавания диктора [11].

2. Эмпирическое вейвлет-преобразование

Мы предлагаем метод построения семейства вейвлетов, адаптированных к обрабатываемым сигналам. Один из путей в достижении адаптивности состоит в предположении, что фильтры зависят от расположения информации в спектре анализируемого сигнала. Для четкости рассмотрим реальные сигналы (где спектр симметричен относительно частоты $\omega = 0$), но подобные рассуждения могут

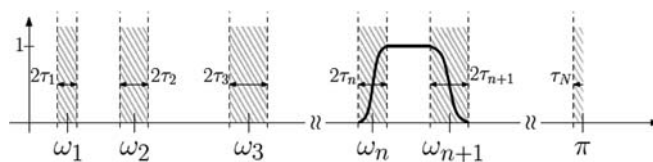


Рис. 1. Разбивка оси Фурье

быть также расширены для комплексных сигналов путем построения различных фильтров на положительных и отрицательных частотах. Мы также рассмотрим нормализованную ось Фурье, которая имеет периодичность для того, чтобы удовлетворить критерию Шеннона, и ограничим наше обсуждение отрезком $\omega \in [0, \pi]$.

Начнем с предположения, что отрезок $[0, \pi]$ делится на N смежных сегментов. Обозначим границы между сегментами ω_n (где $\omega_0 = 0$ и $\omega_N = \pi$), как показано на рис. 1. Каждый сегмент $\Lambda_n = [\omega_{n-1}, \omega_n]$. Вокруг ω_n определяем переходную фазу T_n шириной $2\tau_n$.

Эмпирические вейвлеты [12] определяются как полосовые фильтры на каждом Λ_n . Для этого используем идею, применяемую при построении вейвлетов Littlewood-Paley и Meyer. Тогда для всех $n > 0$ определяем эмпирическую масштабируемую функцию $\hat{\phi}_n$ и эмпирические вейвлеты согласно следующим выражениям:

$$\hat{\phi}_n(\omega) = \begin{cases} 1, & \text{если } |\omega| \leq \omega_n - \tau_n; \\ \cos \left[\frac{\pi}{2} \beta \left(\frac{1}{2\tau_n} (|\omega| - \omega_n + \tau_n) \right) \right], & \text{если } \omega_n - \tau_n \leq |\omega| \leq \omega_n + \tau_n; \\ 0, & \text{в противном случае,} \end{cases} \quad (2)$$

$$\hat{\psi}_n(\omega) = \begin{cases} 1, & \text{если } \omega_n + \tau_n \leq |\omega| \leq \omega_{n+1} - \tau_{n+1}; \\ \cos \left[\frac{\pi}{2} \beta \left(\frac{1}{2\tau_{n+1}} (|\omega| - \omega_{n+1} + \tau_{n+1}) \right) \right], & \text{если } \omega_{n+1} - \tau_{n+1} \leq |\omega| \leq \omega_{n+1} + \tau_{n+1}; \\ \sin \left[\frac{\pi}{2} \beta \left(\frac{1}{2\tau_n} (|\omega| - \omega_n + \tau_n) \right) \right], & \text{если } \omega_n - \tau_n \leq |\omega| \leq \omega_n + \tau_n; \\ 0, & \text{в противном случае.} \end{cases} \quad (3)$$

Функция $\beta(x)$ из $C^k([0, 1])$ (пространство k раз дифференцируемых функций на интервале $[0, 1]$) удовлетворяет условию

$$\beta(x) = \begin{cases} 0, & \text{если } x \leq 0; \\ 1, & \text{если } x \geq 1. \end{cases} \quad (4)$$

и $\beta(x) + \beta(1-x) = 1 \quad \forall x \in [0, 1]$.

Что касается выбора τ_n , возможны несколько вариантов. Самый простой состоит в выборе τ_n пропорционально ω_n : $\tau_n = \gamma \omega_n$, где $0 < \gamma < 1$. Следова-

тельно, для всех $n > 0$ уравнения (2) и (3) принимают вид

$$\hat{\phi}_n(\omega) = \begin{cases} 1, & \text{если } |\omega| \leq (1 - \gamma)\omega_n; \\ \cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\gamma\omega_n}(|\omega| - (1 - \gamma)\omega_n)\right)\right], & \text{если } (1 - \gamma)\omega_n \leq |\omega| \leq (1 + \gamma)\omega_n; \\ 0, & \text{в противном случае} \end{cases} \quad (5)$$

и

$$\hat{\psi}_n(\omega) = \begin{cases} 1, & \text{если } (1 + \gamma)\omega_n \leq |\omega| \leq (1 - \gamma)\omega_{n+1}; \\ \cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\gamma\omega_{n+1}}(|\omega| - (1 - \gamma)\omega_{n+1})\right)\right], & \text{если } (1 - \gamma)\omega_{n+1} \leq |\omega| \leq (1 + \gamma)\omega_{n+1}; \\ \sin\left[\frac{\pi}{2}\beta\left(\frac{1}{2\gamma\omega_n}(|\omega| - (1 - \gamma)\omega_n)\right)\right], & \text{если } (1 - \gamma)\omega_n \leq |\omega| \leq (1 + \gamma)\omega_n; \\ 0, & \text{в противном случае.} \end{cases} \quad (6)$$

Теперь можем определить эмпирическое вейвлет-преобразование $W_f^e(n, t)$ так же как и в случае классического вейвлет-преобразования.

Результаты представлены в виде скалярных произведений с эмпирическими вейвлетами

$$\begin{aligned} W_f^e(n, t) &= \langle f, \psi_n \rangle = \int f(\tau) \overline{\psi_n(\tau - t)} d\tau = \\ &= \hat{f}(\omega) \overline{\hat{\psi}_n(\omega)}, \end{aligned} \quad (7)$$

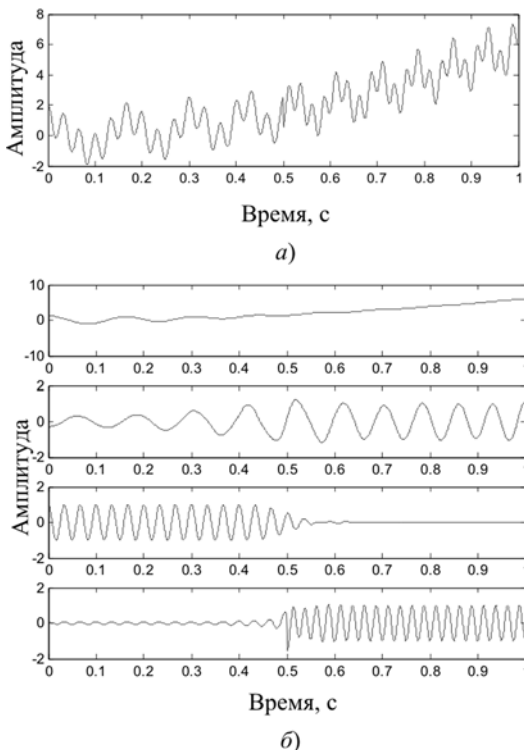


Рис. 2. Входной сигнал $f_{sig1}(t)$ (а) и эмпирические моды, полученные путем EWT (б)

а аппроксимирующие коэффициенты — в виде скалярных произведений с масштабируемой функцией

$$\begin{aligned} W_f^e(0, t) &= \langle f, \phi_1 \rangle = \int f(\tau) \cdot \overline{\phi_1(\tau - t)} d\tau = \\ &= \hat{f}(\omega) \cdot \overline{\hat{\phi}_1(\omega)}, \end{aligned} \quad (8)$$

где $\hat{\psi}_n(\omega)$ и $\hat{\phi}_1(\omega)$ определяются из уравнений (5) и (6) соответственно. Обратное преобразование принимает вид

$$\begin{aligned} f(t) &= W_f^e(0, t) * \phi_1(t) + \sum_{n=1}^N W_f^e(n, t) * \psi_n(t) = \\ &= \hat{W}_f^e(0, \omega) * \hat{\phi}_1(\omega) + \sum_{n=1}^N \hat{W}_f^e(n, \omega) * \hat{\psi}_n(\omega). \end{aligned} \quad (9)$$

Эмпирическая мода (*Intrinsic Mode Function, IMF*) f_k определяется следующим образом:

$$f_0(t) = W_f^e(0, t) * \phi_1(t), \quad (10)$$

$$f_k(t) = W_f^e(k, t) * \psi_k(t). \quad (11)$$

Примеры получения IMF-компонент с помощью EWT показаны на рис. 2 и 3. Первый тестовый сигнал $f_{sig1}(t)$ (рис. 2) получен путем суммирования трех компонент (для $t \in [0, 1]$)

$$f_{s1}(t) = 6t^2; \quad (12)$$

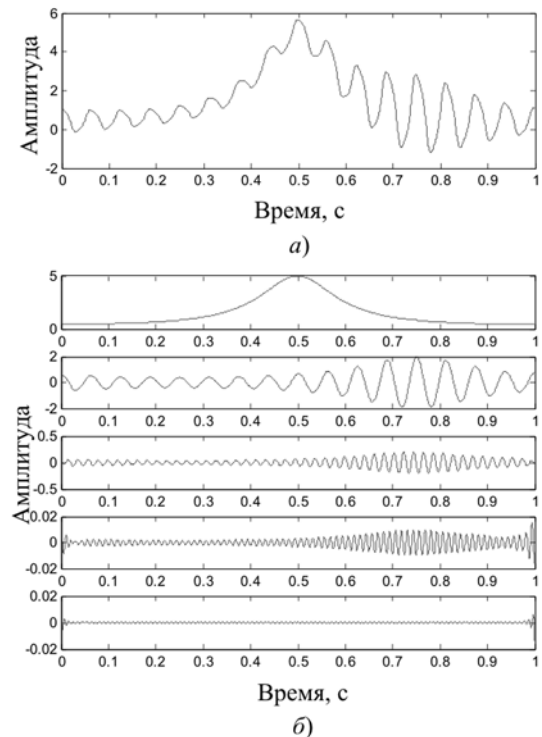


Рис. 3. Входной сигнал $f_{sig2}(t)$ (а) и эмпирические моды, полученные путем EWT (б)

$$f_{s2}(t) = \cos(10\pi t + 10\pi t^2); \quad (13)$$

$$f_{s3}(t) = \begin{cases} \cos(80\pi t - 15\pi), & \text{если } t > 0,5; \\ \cos(60\pi t), & \text{в противном случае.} \end{cases} \quad (14)$$

$$f_{sig1}(t) = f_{s1}(t) + f_{s2}(t) + f_{s3}(t). \quad (15)$$

Компоненты второго сигнала $f_{sig2}(t)$ принимают вид

$$f_{s1}(t) = \frac{1}{1,2 + \cos(2\pi t)}; \quad (16)$$

$$f_{s2}(t) = \frac{1}{1,5 + \sin(2\pi t)}; \quad (17)$$

$$f_{s3}(t) = \cos(32\pi t + \cos(64\pi t)) \quad (18)$$

и

$$f_{sig2}(t) = f_{s1}(t) + f_{s2}(t)f_{s3}(t). \quad (19)$$

3. Дискретный алгоритм разделения энергии

После получения IMF необходимо выбрать метод для выделения мгновенной амплитуды и частоты. Обычно применяется преобразование Гильберта [13], однако дискретный алгоритм разделения энергии (*Discrete Energy Separation Algorithm*, DESA) [14] превосходит его по вычислительной сложности и скорости на реальных сигналах. А также рассматривает энергию, необходимую для генерации каждого монокомпонента АМ — FM сигнала. Это помогает при исследовании вибраций голосовых складок для выделения отличительных особенностей каждого диктора.

Пусть $d^m(n)$ — значение IMF для каждого фрейма при $n = 1, \dots, N$ и $m = 1, \dots, M_x$, где M_x обозначает число мод, на которые $x(t)$ разбивается.

Затем мы можем применить дискретный оператор Тигера

$$\Psi[d^m(n)] = \frac{(d^m(n))^2 - d^m(n-1)d^m(n+1)}{n}, \quad n = 2, \dots, N-1. \quad (20)$$

Если $d^m(n)$ — дискретный косинусный с постоянной амплитудой A и частотой ω , $d^m(n) = A\cos(\Omega n + \theta)$ при $\Omega = \omega T$ и T — период дискретизации, то

$$\Psi[d^m(n)] = A^2\omega^2 \left(\frac{\sin\Omega}{\Omega} \right)^2. \quad (21)$$

Затем применим алгоритм DESA для АМ — FM разделения. Он оценивает мгновенную частоту $\Omega(n)$ и мгновенную огибающую $a(n)$ следующим образом:

$$\Omega(n) = \arccos\left(1 - \frac{\Psi[y(n)] + \Psi[y(n+1)]}{4\Psi[d^m(n)]}\right), \quad (22)$$

$$|a(n)| = \sqrt{\frac{\Psi[d^m(n)]}{1 - \left(1 - \frac{\Psi[y(n)] + \Psi[y(n+1)]}{4\Psi[d^m(n)]}\right)^2}}, \quad (23)$$

где $y(n) = d^m(n) - d^m(n-1)$ для $n = 2, \dots, N$.

4. Извлечение речевых признаков на основе предлагаемого подхода

На рис. 4 показана блок-схема получения признаков, характерных для каждого диктора, путем модуляции источника возбуждения. Приводятся компоненты мгновенной частоты и мгновенной амплитуды, а также вектор, объединяющий все эти компоненты.

Процесс вычисления вектора признаков следующий:

1. *Выделение вокализованных/невокализованных участков.* Вектор признаков извлекается только из вокализованных участков.

2. *Получение IMF-компонент на основе EWT:*

а. Применение преобразования Фурье.

б. Вычисление локального максимума на отрезке $[0, \pi]$ и нахождение множества $\{\omega_n\}$.

в. Выбор параметра согласно $\gamma < \min_n \left(\frac{\omega_{n+1} - \omega_n}{\omega_{n+1} + \omega_n} \right)$.

д. Построение банка фильтров.

е. Фильтрация сигнала для получения каждого компонента IMF.

3. *Применение DESA.* Нахождение мгновенной частоты и мгновенной амплитуды на основе алгоритма выделения энергии Тигера.

4. *Объединение векторов признаков.* Полученные вектора мгновенных частот и мгновенных амплитуд далее объединяются для получения нового вектора признаков.

Заключение

В работе описан подход для распознавания диктора, основанный на эмпирическом вейвлет-преобразовании. Он дает физически

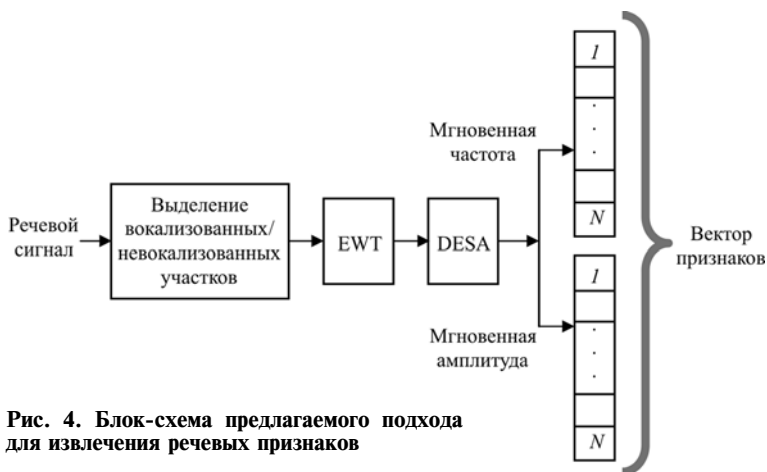


Рис. 4. Блок-схема предлагаемого подхода для извлечения речевых признаков

значимые результаты в режиме реального времени. Генерирует IMF через адаптивный алгоритм из набора данных. Применяя EWT, мы получаем IMF, которые являются уникальными особенностями для каждого диктора. Для выделения мгновенной частоты и мгновенной амплитуды был применен алгоритм DESA. Он превосходит преобразование Гильберта, сохраняя внутренние свойства данных, не ограничиваясь принципом неопределенности.

Данная работа выполнена при финансовой поддержке Фонда развития науки при Президенте Азербайджанской Республики — грант № EIF-RITN-MQM-2/IKT-2-2013-7(13)-29/18/1.

Список литературы

1. **Quatieri T. F.** Discrete-time speech signal processing: principles and practice, ser. NJ: Prentice Hall, 2001. 819 p.
2. **Benesty J., Sondhi M., Huang Y.** Springer handbook of speech processing. Springer, 2008. 1159 p.
3. **Doddington G.** Speaker recognition based on idiolectal differences between speakers // Proc. of Eurospeech. 2001. V. 4. P. 2521—2524.
4. **Reynolds D.** Channel robust speaker verification via feature mapping // Proc. of ICASSP. 2003. V. 2. P. 53—56.
5. **Wolf J. J.** Efficient acoustic parameters for speaker recognition // J. Acoustical Society of America. 1982. V. 51, N. 6 (Part 2). P. 2044—2056.
6. **Kinnunen T., Li H.** An overview of text-independent speaker recognition: from features to supervectors // Speech Communication. 2010. V. 52, N. 1. P. 12—40.
7. **Rose P.** Forensic speaker identification / Taylor & Francis forensic science series. New York: Taylor & Francis, 2002. 380 p.
8. **Kinnunen T.** Spectral features for automatic text-independent speaker recognition: Licentiate thesis. Department of Computer Science. University of Joensuu. Finland, 2003.
9. **Маркел Дж., Грей А. Х.** Линейное предсказание речи. М.: Связь, 1980. 308 с.
10. **Furui S.** Cepstral analysis techniques for automatic speaker verification // IEEE tran. acoust., speech, signal processing. 1981. V. 27. P. 254—272.
11. **Holambe R. S., Deshpande M. S.** Noise Robust Speaker Identification: Using Nonlinear Modeling // Forensic Speaker Recognition. 2012. P. 153—182.
12. **Gilles J.** Empirical Wavelet Transform // IEEE Transactions on Signal Processing. 2013. V. 61, N. 16. P. 3999—4010.
13. **Huang N. E.** Hilbert-Huang Transform and its applications. World Scientific Publishing, 2005. 311 p.
14. **Schlotthauer G., Torres M. E., Rufiner H. L.** A new algorithm for instantaneous F0 speech extraction based on Ensemble Empirical Mode Decomposition // 17th European Signal Processing Conf. 2009. P. 2347—2351.

Y. N. Imamverdiyev, Head of Department, e-mail: yadigar@lan.ab.az,

L. V. Sukhostat, Researcher, e-mail: lsuhostat@hotmail.com,

Institute for Information Technologies, Azerbaijan National Academy of Sciences, Baku, Azerbaijan

Development of Robust Speech Feature Extraction Method Based on Empirical Wavelet Transform

Speech feature vectors extraction is an important step for speaker recognition systems. Currently, state-of-art works remain relevant to find informative features of speech signals, ensuring its appropriate description and low error rate during recognition. In this paper we present an approach for speech feature extraction based on empirical wavelet transform. To calculate the instantaneous frequency and instantaneous amplitude of IMFs Discrete Energy Separation Algorithm is used, which overcomes the disadvantages of Hilbert transform. The proposed method increases the recognition accuracy, while maintaining an acceptable level of computational complexity.

Keywords: speaker recognition, empirical wavelet transform, discrete energy separation algorithm, instantaneous amplitude, instantaneous frequency

References

1. **Quatieri T. F.** Discrete-time speech signal processing: principles and practice, ser. NJ: Prentice Hall, 2001. 819 p.
2. **Benesty J., Sondhi M., Huang Y.** Springer handbook of speech processing. Springer, 2008. 1159 p.
3. **Doddington G.** Speaker recognition based on idiolectal differences between speakers // Proc. of Eurospeech. 2001. V. 4. P. 2521—2524.
4. **Reynolds D.** Channel robust speaker verification via feature mapping // Proc. of ICASSP. 2003. V. 2. P. 53—56.
5. **Wolf J. J.** Efficient acoustic parameters for speaker recognition // J. Acoustical Society of America. 1982. V. 51, N. 6 (Part 2). P. 2044—2056.
6. **Kinnunen T., Li H.** An overview of text-independent speaker recognition: from features to supervectors // Speech Communication. 2010. V. 52, N. 1. P. 12—40.
7. **Rose P.** Forensic speaker identification. Taylor & Francis forensic science series. New York: Taylor & Francis. 2002. 380 p.
8. **Kinnunen T.** Spectral features for automatic text-independent speaker recognition. Licentiate thesis. Department of Computer Science. University of Joensuu. Finland. 2003.
9. **Маркел Дж., Грей А. Х.** Линейное предсказание речи. М.: Связь, 1980. 308 с.
10. **Furui S.** Cepstral analysis techniques for automatic speaker verification // IEEE tran. acoust., speech, signal processing. 1981. V. 27. P. 254—272.
11. **Holambe R. S., Deshpande M. S.** Noise Robust Speaker Identification: Using Nonlinear Modeling // Forensic Speaker Recognition. 2012. P. 153—182.
12. **Gilles J.** Empirical Wavelet Transform // IEEE Transactions on Signal Processing. 2013. V. 61, N. 16. P. 3999—4010.
13. **Huang N. E.** Hilbert-Huang Transform and its applications. World Scientific Publishing, 2005. 311 p.
14. **Schlotthauer G., Torres M. E., Rufiner H. L.** A new algorithm for instantaneous F0 speech extraction based on Ensemble Empirical Mode Decomposition // Proc. 17th European Signal Processing Conf. 2009. P. 2347—2351.

МОДЕЛИРОВАНИЕ И ОПТИМИЗАЦИЯ MODELING AND OPTIMIZATION

УДК 004.021, 519.854.2

В. А. Чеканин, канд. техн. наук, доц., vladchekanin@rambler.ru,

А. В. Чеканин, д-р техн. наук, проф., зав. каф., avchekanin@rambler.ru

Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования "Московский государственный технологический университет "СТАНКИН", г. Москва

Повышение эффективности конструирования ортогональной упаковки объектов¹

Рассматривается NP-полная оптимизационная задача ортогональной упаковки объектов, актуальная при решении большого числа практических задач автоматизации и управления. Предлагается многоуровневая связанная структура данных, обеспечивающая высокую скорость доступа к содержимому контейнеров при конструировании ортогональной упаковки. Проведен анализ эффективности новой структуры данных на тестовых задачах двухмерной и трехмерной ортогональной упаковки.

Ключевые слова: упаковка, задача ортогональной упаковки, структура данных, многоуровневая связанная структура данных, оптимизация, распределение ресурсов, вычислительный эксперимент

Введение

Задача ортогональной упаковки объектов представляет собой NP-полную задачу оптимального размещения набора ортогональных объектов в ортогональных контейнерах [1]. Эта задача является актуальной, к ее решению сводится множество практических оптимизационных задач в автоматизации и управлении. Она возникает, в частности, при решении задач складирования и перевозки грузов, раскроя материалов, распределения ресурсов в вычислительных сетях, календарного планирования, компоновки элементов интегральных схем и многих других оптимизационных задач [2–5]. Разработке методов оптимального решения задачи ортогональной упаковки посвящены работы большого числа как отечественных (А. С. Филиппова, А. Ф. Валеева, Ю. Г. Стоян, В. М. Картак, Ю. И. Валиахметова, В. В. Бухвалова и др.), так и зарубежных исследователей (S. Martello, E. Hopper, A. Lodi, D. Vigo, A. Bortfeldt, G. Wascher, H. Haubner, H. Schumann, S. Fekete, J. Schepres, T. Crainic, G. Perboli и др.). Значительный вклад в развитие методов оптимального решения задачи ортогональной упаковки объектов внесла отечественная научная школа под руководством Э. А. Мухачевой.

¹ Данная работа финансировалась Министерством образования и науки РФ в рамках государственного задания в сфере научной деятельности.

В общем виде постановка D -мерной задачи ортогональной упаковки объектов подразумевает наличие двух наборов элементов:

1) набор N ортогональных контейнеров (D -мерных параллелепипедов) с габаритными размерами $\{W_j^1, W_j^2, \dots, W_j^D\}$, $j \in \{1, \dots, N\}$, которые служат для размещения в них ортогональных объектов;

2) набор n ортогональных объектов (D -мерных параллелепипедов) с габаритными размерами $\{w_i^1, w_i^2, \dots, w_i^D\}$, $i \in \{1, \dots, n\}$, которые необходимо разместить в минимальном числе контейнеров.

Обозначим положение D -мерного ортогонального объекта i в контейнере j через $(x_{ij}^1; x_{ij}^2; \dots; x_{ij}^D)$. При решении задачи ортогональной упаковки объектов должны быть выполнены следующие условия корректности размещения [6–7]:

1) ребра размещенных в контейнере ортогональных объектов параллельны ребрам этого контейнера;

2) размещенные объекты не перекрывают друг друга, т. е.

$$(x_{ij}^d \geq x_{kj}^d + w_k^d) \vee (x_{kj}^d \geq x_{ij}^d + w_i^d),$$

$\forall j \in \{1, \dots, N\}, \forall d \in \{1, \dots, D\}, \forall i, k \in \{1, \dots, n\}, i \neq k;$

3) размещенные объекты не выходят за границы контейнеров, т. е.

$$(x_{ij}^d \geq 0) \wedge (x_{ij}^d + w_i^d \leq W_j^d) \quad \forall j \in \{1, \dots, N\}, \\ \forall d \in \{1, \dots, D\}, \forall i \in \{1, \dots, n\}.$$

Перед размещением объекты классифицируются по различным типам. Объекты одного и того же типа имеют одинаковые геометрические и физические характеристики, т. е. имеют одни и те же габаритные размеры и выполнены из одного материала.

Решение задачи ортогональной упаковки n объектов произвольной размерности может быть закодировано строкой размещения $s = \{A_1, A_2, \dots, A_n\}$, содержащей последовательность выбора объектов для их размещения в контейнерах, в которой число A_i представляет собой номер типа i -го размещаемого объекта.

Конструирование ортогональной упаковки

Процесс решения любой оптимизационной задачи упаковки объектов можно представить в виде схемы, приведенной на рис. 1.

Для быстрого поиска удовлетворительного решения задачи упаковки применяют метаэвристические алгоритмы [8–14], осуществляющие рациональный поиск такой строки размещения, декодирование которой обеспечивает получение упаковки требуемого качества.

Эффективность поиска удовлетворительного решения NP -полной задачи ортогональной упаковки объектов определяется в большей мере используемым оптимизационным алгоритмом. В процессе поиска решения с использованием многопроходных эвристических алгоритмов формируются промежуточные строки решения, эффективность декодирования которых определяется используемой моделью конструирования упаковки. Среди наиболее распространенных моделей конструирования упаковки можно выделить блочную [15], матричную [16] и узловую модели [17, 18].

В работе [19] показана эффективность применения предложенной ранее авторами настоящей статьи модели потенциальных контейнеров для конструирования ортогональной упаковки, которая построена на базе узловой модели. При использовании модели потенциальных контейнеров свободное пространство

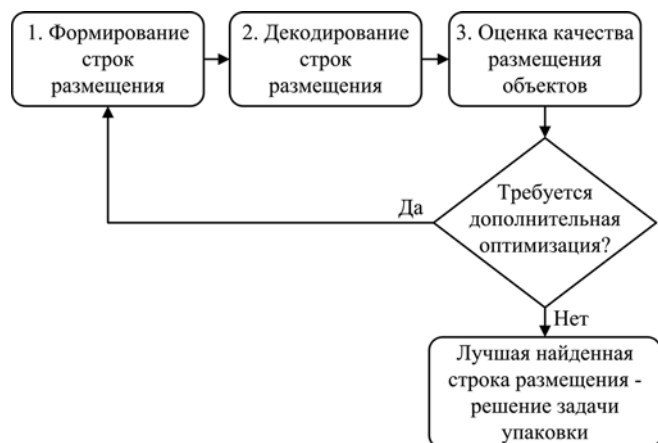


Рис. 1. Процесс решения задачи упаковки

каждого контейнера представляется в виде набора ортогональных объектов — потенциальных контейнеров (ПК), описывающих всевозможные свободные пространства контейнера.

Каждый ПК k описывается вектором $\{p_k^1; p_k^2; \dots; p_k^D\}$, содержащим его габаритные размеры, а также вектором $\{x_k^1; x_k^2; \dots; x_k^D\}$, содержащим координаты точки ПК, ближайшей к началу координат контейнера, в котором размещен этот ПК.

В модели ПК при перекрытии ортогонального объекта i с габаритными размерами $\{w_i^1; w_i^2; \dots; w_i^D\}$, размещенного в точке $\{x_i^1; x_i^2; \dots; x_i^D\}$ D -мерного ортогонального контейнера объекта, потенциальным контейнером k с габаритными размерами $\{p_k^1; p_k^2; \dots; p_k^D\}$, расположенном в точке $\{x_k^1; x_k^2; \dots; x_k^D\}$ контейнера, в пространстве размещаемого контейнера образуются не более $2D$ новых ПК из двух наборов:

1) набор ПК с габаритными размерами $\{p_k^1; p_k^2; \dots; p_k^{d-1}; x_i^d - x_k^d; p_k^{d+1}; \dots; p_k^D\}$, расположенных в точке $\{x_k^1; x_k^2; \dots; x_k^d; \dots; x_k^D\}$ и создаваемых при выполнении условий перекрытия $x_i^d > x_k^d$ и $x_i^d < x_k^d + p_k^d \forall d \in \{1, \dots, D\}$;

2) набор ПК с габаритными размерами $\{p_k^1; p_k^2; \dots; p_k^{d-1}; x_k^d + p_k^d - x_i^d - w_i^d; p_k^{d+1}; \dots; p_k^D\}$, расположенных в точках $\{x_k^1; x_k^2; \dots; x_k^{d-1}; x_i^d + w_i^d; x_k^{d+1}; \dots; x_k^D\}$ и создаваемых при выполнении условий перекрытия $x_i^d + w_i^d > x_k^d$ и $x_i^d + w_i^d < x_k^d + p_k^d \forall d \in \{1, \dots, D\}$.

В качестве примера рассмотрим двухмерный прямоугольный контейнер с габаритными размерами $\{L; H\}$, содержащий единственный ПК с габаритными размерами $\{l; h\}$, расположенный в точке $\{0; 0\}$ размещаемого контейнера. При размещении прямоугольного объекта с габаритными размерами $\{l; h\}$ в точке $\{x; y\}$ контейнера образуются следующие новые ПК (рис. 2):

1) в точке $\{0; 0\}$ — ПК с габаритными размерами $\{L; y\}$ и $\{x; H\}$;

2) в точке $\{x + l; 0\}$ — ПК с габаритными размерами $\{W - x - l; H\}$ и в точке $\{0; y + h\}$ — ПК с габаритными размерами $\{L; H - y - h\}$.

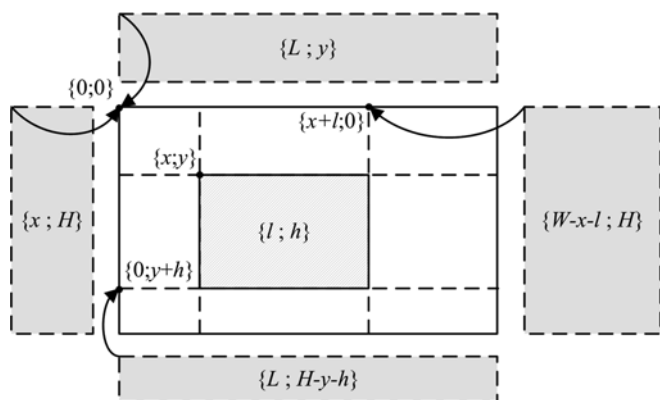


Рис. 2. Потенциальные контейнеры двумерного прямоугольного контейнера

Постановка задачи D -мерной ортогональной упаковки включает также задание направления загрузки контейнера в виде приоритетного списка выбора его координатных осей $L_P = \{P_1; P_2; \dots; P_D\}$, где $P_d \in [1; D] \forall d \in [1; D]$. Для загрузки контейнера в заданном направлении выбор ПК при размещении каждого объекта осуществляется в порядке, определенном в приоритетном списке L_P .

Алгоритм декодирования строки размещения для модели потенциальных контейнеров.

1. Выбрать очередной объект i из строки размещения. Если выбраны все объекты, выполнить переход к п. 4. Выбрать первый контейнер j , содержащий, как минимум, один ПК.

2. В текущем контейнере j провести последовательный поиск ближайшего к его началу координат ПК k , к которому возможно присоединить текущий объект i (проверка выполнения условия $w_i^d \leq p_k^d \forall d \in \{1, \dots, D\}$). Если искомым ПК не найден, выполнить переход к следующему контейнеру $j := j + 1$ и повторить п. 2. Если среди всех контей-

неров не найден искомым узел, выполнить переход к п. 1.

3. Разместить объект i в найденном ПК k контейнера j . Образовать новые ПК и выполнить процедуру поиска и удаления вложенных ПК. Упорядочить все ПК в порядке убывания приоритета присоединения к ним объектов, т. е. для любого ПК k контейнера j должно выполняться следующее неравенство:

$$\sum_{h=1}^D \left(x_k^{P_h} \prod_{d=P_{h+1}}^D W_j^d \right) \leq \sum_{h=1}^D \left(x_{k+1}^{P_h} \prod_{d=P_{h+1}}^D W_j^d \right).$$

Выполнить переход к п. 1.

4. Завершить декодирование.

Наиболее трудоемким этапом алгоритма декодирования строки размещения является п. 3, в котором выполняется упорядочение всех ПК. Например, для направления загрузки $L_P = \{1; 2; 3\}$ набор потенциальных контейнеров (табл. 1) должен быть упорядочен по неубыванию координат сначала вдоль координатной оси 1, затем — вдоль оси 2, и наконец, — вдоль оси 3 (табл. 2).

Для повышения эффективности декодирования при решении задач ортогональной упаковки объектов с использованием модели потенциальных контейнеров предлагается новая структура данных — многоуровневая связная структура, которая снимает необходимость сортировки всех координат потенциальных контейнеров.

Многоуровневая связная структура данных

В основу разработанной многоуровневой связной структуры данных положена идея представления набора координат потенциальных контейнеров в виде рекурсивно вложенных линейных связных списков.

Набор K потенциальных контейнеров, расположенных в точках $\{x_k^1; x_k^2; \dots; x_k^D\}$, $k \in K$, контейнера, представляется в виде D -уровневых рекурсивно вложенных друг в друга упорядоченных по возрастанию линейных связных списков (рис. 3). Каждый элемент j списка i на уровне вложенности P_d содержит координату $s_{i,j}^{P_d} = x_k^{P_d}$ такого потенциального контейнера k , что внутри каждого связного списка выполняется неравенство $s_{i,j}^{P_d} < s_{i,j+1}^{P_d} \forall P_d \in L_P$.

Каждый элемент j списка i на уровне вложенности P_d содержит координату $s_{i,j}^{P_d} = x_k^{P_d}$ такого потенциального контейнера k , что внутри каждого связного списка выполняется неравенство $s_{i,j}^{P_d} < s_{i,j+1}^{P_d} \forall P_d \in L_P$.

Таблица 1

Исходные координаты потенциальных контейнеров

Порядковый номер ПК k	1	2	3	4	5	6	7	8	9	10
Координата x_k^1	0	2	2	2	4	0	4	0	4	2
Координата x_k^2	1	3	7	9	1	2	1	2	1	7
Координата x_k^3	0	1	5	6	2	1	3	3	1	2

Таблица 2

Координаты потенциальных контейнеров после упорядочения

Порядковый номер ПК k	1	2	3	4	5	6	7	8	9	10
Координата x_k^1	0	0	0	2	2	2	2	4	4	4
Координата x_k^2	1	2	2	3	7	7	9	1	1	1
Координата x_k^3	0	1	3	1	2	5	6	1	2	3

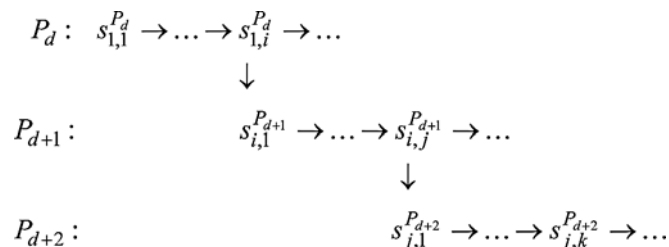


Рис. 3. Многоуровневая связная структура данных

В качестве примера рассмотрим ортогональный трехмерный контейнер (параллелепипед), содержащий набор K потенциальных контейнеров, вектора $X_k = \{x_k^1; x_k^2; x_k^3\}$, $k \in K$ которых приведены в табл. 1.

Представление набора координат потенциальных контейнеров при использовании многоуровневой связной структуры данных для направления загрузки $L_p = \{1; 2; 3\}$ приведено на рис. 4.

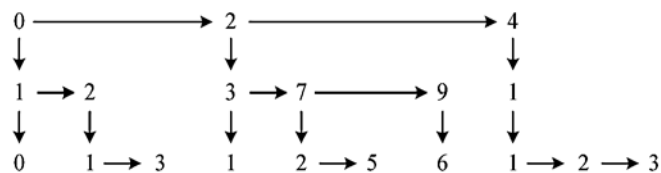


Рис. 4. Многоуровневая связная структура данных трехмерного контейнера для направления загрузки $L_p = \{1; 2; 3\}$

Многоуровневая связная структура данных при размещении n объектов позволяет по сравнению с простым линейным связным списком как минимум в $\frac{2n}{n+1}$ раз повысить скорость доступа к ПК по их координатам.

При размещении n объектов в D -мерном ортогональном контейнере формируется не более $2Dn$ ПК [19]. При организации набора ПК в виде линейного связного списка среднее время доступа к любому ПК равно

$$T_L = \frac{2Dnt}{2} = Dnt, \quad (1)$$

где t — время перехода к следующему элементу списка.

Любой линейный связный список в многоуровневой связной структуре состоит не более чем из $n+1$ элементов, так как при условии плотного размещения объектов два соседних размещенных объекта имеют одну общую координату, следовательно, не более $n+1$ ПК будут иметь различные координаты вдоль каждой из координатных осей контейнера. Поэтому при организации набора ПК в виде многоуровневой связной структуры среднее время доступа к D -мерному ПК равно

$$T_M = \frac{D(n+1)t}{2}. \quad (2)$$

Анализируя формулы (1) и (2), делаем вывод, что скорость доступа к ПК при использовании многоуровневой структуры данных выше, чем при использовании линейного связного списка, как минимум в $\frac{2n}{n+1}$ раз. Эта оценка проведена для самого алго-

ритмически сложного случая, когда каждый ПК для каждой координатной оси имеет уникальную координату, не совпадающую с координатами других ПК.

Вычислительные эксперименты

Анализ эффективности работы предложенной многоуровневой связной структуры данных проводился при решении тестовых задач двухмерной контейнерной упаковки из библиотеки OR-library [20] для наборов двухмерных прямоугольных объектов, взятых из задач 2DBPP (2D Bin Packing Problem), сформулированных *S. P. Fekete* и *J. Schepers* [21].

Задачи решали с использованием разработанного прикладного программного обеспечения *Packer* [22], предназначенного для решения задач одномерной упаковки, двухмерной упаковки прямоугольников (включая задачи прямоугольного раскроя), а также задач трехмерной упаковки параллелепипедов с использованием метаэвристических методов. Эксперименты проводили на персональной ЭВМ (ЦП — AMD 1,79 ГГц; ОЗУ — 1,12 Гбайт).

В ходе каждого вычислительного эксперимента решали задачи двухмерной ортогональной упаковки трех различных типов с различным процентным соотношением классов размещаемых объектов. Объемы выборки тестовых задач равны $m = 40, 50, 100, 150, 250, 500$ и 1000 объектов. Полные параметры тестовых задач приведены в работе [10].

На основе усредненных полученных результатов тестирования построена диаграмма, приведенная на рис. 5. Относительная временная эффективность многоуровневой связной структуры данных рассчитывается по формуле $T = t_L/t_M$, где t_L и t_M — время решения задачи упаковки при использовании линейного списка и многоуровневого связного списка соответственно. Из диаграммы видно, что многоуровневая связная структура, используемая для представления набора ПК, обеспечивает более быстрое размещение объектов по сравнению с обычным линейным списком, упорядочение в котором проводится с помощью одного из наиболее быстрых алгоритмов сортировки Quicksort [23].

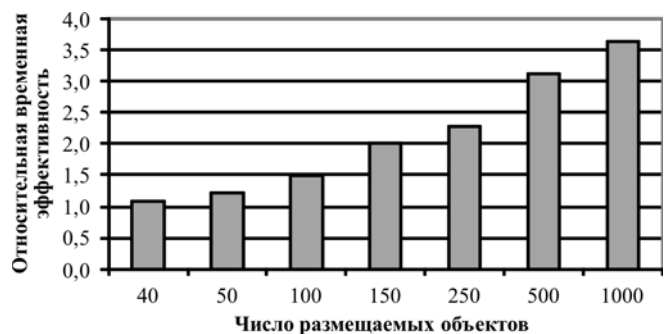


Рис. 5. Временная эффективность многоуровневой связной структуры при решении тестовых задач двухмерной ортогональной упаковки

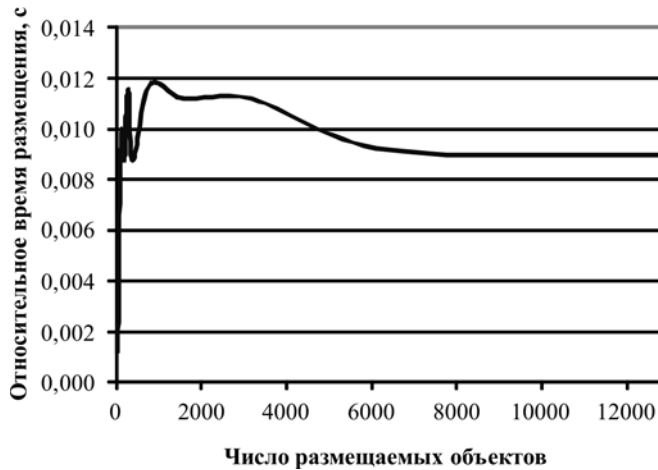


Рис. 6. Относительное время размещения трехмерных ортогональных объектов

Проведенные вычислительные эксперименты показали, что эффективность многоуровневой связанной структуры данных при решении задач двухмерной ортогональной упаковки объектов растет с увеличением числа размещаемых объектов.

Особенность многоуровневой связанной структуры данных заключается в том, что среднее время размещения, затрачиваемое на один объект, не зависит от общего числа размещаемых объектов. На рис. 6 представлен график зависимости относительного времени размещения одного и того же набора трехмерных объектов пяти различных типов ($10 \times 10 \times 10$, $20 \times 20 \times 20$, $10 \times 40 \times 20$, $30 \times 5 \times 10$, $5 \times 10 \times 15$) в трехмерный контейнер с габаритными размерами $200 \times 1000 \times 1000$. Относительное время размещения характеризует время, затрачиваемое на размещение одного объекта, и определяется как отношение общего времени размещения всех объектов к их числу n .

Заключение

Для получения быстрого доступа к потенциальным контейнерам при конструировании ортогональной упаковки предложена новая структура данных — многоуровневая связанная структура, в основу которой положена идея представления набора координат потенциальных контейнеров в виде рекурсивно вложенных линейных связанных списков.

Проведенные вычислительные эксперименты на тестовых задачах ортогональной упаковки объектов показали высокую временную эффективность предложенной структуры в сравнении с обычным линейным списком, требующим сортировки всех элементов после размещения каждого очередного объекта в контейнере. Использование многоуровневой связанной структуры позволяет более чем в 2 раза повысить скорость доступа к потенциальным контейнерам, что в конечном итоге обеспечивает более быстрое формирование ортогональной упаковки.

1. Garey M., Johnson D. Computers intractability: a guide to the theory of NP-completeness. San Francisco: W. H. Freeman, 1979. 338 p.
2. Lodi A., Martello S., Monaci M. Two-dimensional packing problems: A survey // European Journal of Operational Research. 2002. V. 141, N. 2. P. 241–252.
3. Wu Y., Li W.-K., Goh M., De Souza R. Three dimensional bin packing problem with variable bin height // European Journal of Operational Research. 2010. V. 202, N. 2. P. 347–355.
4. Egeblad J., Garavelli C., Lisi S., Pisinger D. Heuristics for container loading of furniture // European Journal of Operational Research. 2010. V. 200, N. 3. P. 881–892.
5. Liao C. S., Hsu C. H. New lower bounds for the three-dimensional orthogonal bin packing problem // European Journal of Operational Research. 2013. Vol. 225, N. 2. P. 244–252.
6. Wascher G., Haubner H., Schumann H. An improved typology of cutting and packing problems // European Journal of Operational Research. 2007. V. 183, N. 3. P. 1109–1130.
7. Bortfeldt A., Wascher G. Constraints in container loading — A state-of-the-art review // European Journal of Operational Research. 2013. V. 229, N. 1. P. 1–20.
8. Чеканин А. В., Чеканин В. А. Алгоритмы эффективного решения задачи ортогональной упаковки объектов // Журнал вычислительной математики и математической физики. 2013. Т. 53, № 10. С. 1639–1648.
9. Валеева А. Ф. Применение метаэвристики муравьиной колонии к задачам двумерной упаковки // Информационные технологии. 2005. № 10. С. 36–43.
10. Чеканин В. А., Чеканин А. В. Алгоритм решения задач ортогональной упаковки объектов на основе мультиметодной технологии // Информационные технологии. 2013. № 7. С. 17–21.
11. Чеканин В. А., Ковшов Е. Е. Моделирование и оптимизация технологических операций в промышленном производстве на основе эволюционных алгоритмов // Технология машиностроения. 2010. № 3. С. 53–57.
12. Чеканин В. А., Ковшов Е. Е. Систематизация и анализ структур данных при автоматизации управления складом на основе генетических алгоритмов // Проблемы полиграфии и издательского дела. 2008. № 5. С. 42–51.
13. Чеканин В. А., Чеканин А. В. Исследование генетических методов оптимизации распределения прямоугольных ресурсов // Материалы 2-й международной научно-практической конференции "Современное машиностроение. Наука и образование" (14–15 июня 2012 г.). СПб.: Изд-во Политехн. ун-та, 2012. С. 798–804.
14. Валиахметова Ю. И., Филиппова А. С. Мультиметодный генетический алгоритм для решения задач ортогональной упаковки // Информационные технологии. 2007. № 12. С. 50–56.
15. Филиппова А. С. Моделирование эволюционных алгоритмов решения задач прямоугольной упаковки на базе технологии блочных структур // Информационные технологии. 2006. № 6. Приложение. 32 с.
16. Картак В. М. Матричный алгоритм поиска оптимального решения для задачи упаковки прямоугольников в полубесконечную полосу // Информационные технологии. 2008. № 2. С. 24–30.
17. Crainic T. G., Perboli G., Tadei R. Extreme point-based heuristics for three-dimensional bin packing // INFORMS. Journal on Computing. 2008. V. 20, N. 3. P. 368–384.
18. Чеканин В. А., Чеканин А. В. Эффективные модели представления ортогональных ресурсов при решении задачи упаковки // Информационно-управляющие системы. 2012. № 5. С. 29–32.
19. Chekanin A. V., Chekanin V. A. Improved packing representation model for the orthogonal packing problem // Applied Mechanics and Materials. 2013. V. 390. P. 591–595.
20. Библиотека OR-library наборов объектов из задач S. P. Fekete и J. Schepers. URL: <http://people.brunel.ac.uk/~mastjib/jeb/info.html> (дата обращения: 22 мая 2014).
21. Fekete S. P., Schepers J. New classes of lower bounds for bin packing problems // Integer Programming and Computational Optimization. Lecture Notes in Computer. 1998. V. 1412. P. 257–270.
22. Чеканин В. А., Чеканин А. В. Оптимизация решения задачи ортогональной упаковки объектов // Прикладная информатика. 2012. № 4 (40). С. 55–62.
23. Weiss M. A. Data Structures and Algorithm Analysis in C++. Boston: Pearson Education, 2014. 656 p.

Improving the Efficiency of Construction of the Orthogonal Packing

In this paper is considered the NP-completed optimization orthogonal packing problem that is actual in solving of many practical problems of automation and control. The solution of the multidimensional orthogonal packing problem can be represented as a placement string which contains a sequence of objects to be packed into containers. Constructing of a pack for a given sequence of objects from the placement string is performed by decoding block. To increase the effectiveness decoding is offered a new data structure — multilevel linked data structure that is based on a recursively embedded each to other linear queues. The offered data structure provides high-speed access to the packing during its formation. The effectiveness of the new data structure is investigated on the standard two- and threedimensional test orthogonal packing problems. The carried out computational experiments demonstrate high time efficiency of the proposed data structure compared to the ordered simple linked list. The proposed data structure is applicable for any dimensional orthogonal bin packing problems.

Keywords: packing; packing problem; orthogonal packing problem; data structure; multilevel linked data structure; optimization; discrete optimization; resources allocation; waste minimization; computational experiment

References

1. Garey M., Johnson D. *Computers intractability: a guide to the theory of NP-completeness*. San Francisco: W. H. Freeman, 1979. 338 p.
2. Lodi A., Martello S., Monaci M. Two-dimensional packing problems: A survey. *European Journal of Operational Research*. 2002. V. 141, N. 2. P. 241—252.
3. Wu Y., Li W.-K., Goh M., De Souza R. Three dimensional bin packing problem with variable bin height. *European Journal of Operational Research*. 2010. V. 202, N. 2. P. 347—355.
4. Egeblad J., Garavelli C., Lisi S., Pisinger D. Heuristics for container loading of furniture. *European Journal of Operational Research*. 2010. V. 200, N. 3. P. 881—892.
5. Liao C. S., Hsu C. H. New lower bounds for the three-dimensional orthogonal bin packing problem. *European Journal of Operational Research*. 2013. V. 225, N. 2. P. 244—252.
6. Wascher G., Haubner H., Schumann H. An improved typology of cutting and packing problems. *European Journal of Operational Research*. 2007. V. 183, N. 3. P. 1109—1130.
7. Bortfeldt A., Wascher G. Constraints in container loading — A state-of-the-art review. *European Journal of Operational Research*. 2013. V. 229, N. 1. P. 1—20.
8. Chekanin A. V., Chekanin V. A. Algoritmy effektivnogo resheniya zadachi ortogonal'noy upakovki ob'yektov. *Zhurnal vychislitel'noy matematiki i matematicheskoy fiziki*. 2013. V. 53, N. 10. P. 1639—1648.
9. Valeyeva A. F. Primeneniye metaevristiki murav'inoy kolonii k zadacham dvumernoy upakovki. *Informatsionnyye tekhnologii*. 2005. V. 10. P. 36—43. (rus.).
10. Chekanin V. A., Chekanin A. V. Algoritm resheniya zadach ortogonal'noy upakovki ob'yektov na osnove mul'timetodnoy tekhnologii. *Informatsionnyye tekhnologii*. 2013. V. 7. P. 17—21. (rus.).
11. Chekanin V. A., Kovshov E. E. Modelirovaniye i optimizatsiya tekhnologicheskikh operatsiy v promyshlennom proizvodstve na osnove evolyutsionnykh algoritmov. *Tekhnologiya Mashinostroeniya*. 2010. V. 3. P. 53—57. (rus.).
12. Chekanin V. A., Kovshov E. E. Sistematizatsiya i analiz struktur dannykh pri avtomatizatsii upravleniya skladom na osnove geneticheskikh algoritmov. *Izvestiya vysshikh uchebnykh zavedeniy. Problemy poligrafii i izdatel'skogo dela*. 2008. N. 5. P. 42—51. (rus.).
13. Chekanin V. A., Chekanin A. V. Issledovaniye geneticheskikh metodov optimizatsii raspredeleniya pryamougol'nykh resursov. *Materialy 2-y mezhdunarodnoy nauchno-prakticheskoy konferentsii "Sovremennoye mashinostroeniye. Nauka i obrazovaniye"*. SPb.: Izd-vo Politekh. un-ta, 2012. P. 798—804. (rus.).
14. Valiahmetova Yu. I., Filippova A. S. Mul'timetodnyy geneticheskiiy algoritm dlya resheniya zadach ortogonal'noy upakovki. *Informatsionnyye tekhnologii*. 2007. N. 12. P. 50—56. (rus.).
15. Filippova A. S. Modelirovaniye evolyutsionnykh algoritmov resheniya zadach pryamougol'noy upakovki na baze tekhnologii blochnykh struktur. *Informatsionnyye tekhnologii*. 2006. N. 6. Prilozheniye. 32 p. (rus.).
16. Kartak V. M. Matrichnyy algoritm poiska optimal'nogo resheniya dlya resheniya zadachi upakovki pryamougol'nikov v polubeskonechnuyu polosyu. *Informatsionnyye tekhnologii*. 2008. N. 2. P. 24—30. (rus.).
17. Crainic T. G., Perboli G., Tadei R. Extreme point-based heuristics for three-dimensional bin packing. *INFORMS, Journal on Computing*. 2008. V. 20, N. 3. P. 368—384.
18. Chekanin V. A., Chekanin A. V. Effektivnyye modeli predstavleniya ortogonal'nykh resursov pri reshenii zadachi upakovki. *Informatsionno-Upravlyayushchiye systemy*. 2012. V. 5. P. 29—32. (rus.).
19. Chekanin A. V., Chekanin V. A. Improved packing representation model for the orthogonal packing problem. *Applied Mechanics and Materials*. 2013. V. 390. P. 591—595.

А. С. Шелудько, инженер, e-mail: sheludkoas@susu.ac.ru,
 В. И. Ширяев, д-р техн. наук, проф., зав. каф., e-mail: vis@susu.ac.ru,
 Южно-Уральский государственный университет, г. Челябинск

Алгоритм гарантированного оценивания параметра одномерного хаотического отображения

Рассматривается применение гарантированного подхода для нахождения множественной оценки параметра квадратичного отображения по единственной зашумленной реализации измерений. Представлены результаты работы алгоритма для различных моделей ошибок измерений. Исследовано, в каких случаях удастся уточнить априорную оценку, а также получить точное значение параметра.

Ключевые слова: хаотическое отображение, задача идентификации, гарантированный подход

Введение и постановка задачи

Модели и методы хаотической динамики применяются в различных областях исследований: динамическая обработка и защита информации [1], моделирование биологических [2] и экономических [3] систем. Распространенным приложением является реконструкция математической модели по результатам измерений [4–6]. Одним из подходов к решению задачи реконструкции является разложение измерений y_k по системе базисных процессов [7–9]:

$$y_k = a_0 + \sum_{i=1}^n a_i x_k^{(i)} + \zeta_k, \quad k = 1, 2, \dots, N,$$

где a_i — коэффициенты разложения (i — номер базисного процесса); ζ_k — ошибки аппроксимации.

Для задания моделей процессов $x_k^{(1)}, \dots, x_k^{(n)}$ могут быть использованы одномерные хаотические отображения [10, 11]:

$$x_{k+1} = f(x_k, \lambda). \quad (1)$$

Применение моделей вида (1) требует разработки соответствующих алгоритмов идентификации для хаотических процессов (см., например, [5, 12]), в том числе в реальном времени. В данной работе исследуется применение гарантированного подхода [13–15] в задаче оценивания параметра λ хаотического отображения (1) по единственной зашумленной реализации измерений

$$y_k = x_k + v_k, \quad k = 1, 2, \dots, N. \quad (2)$$

При этом априорная информация об ошибках v_k представляется только в виде множественных оценок V_k : $v_k \in V_k$. Алгоритм гарантированного оценивания предполагает рекуррентное нахождение множественных оценок (информационных множеств) Λ_k и X_k для параметра λ и переменной состояния x_k отображения (1): $\lambda \in \Lambda_k, x_k \in X_k$. В работе [16] рассмотрен алгоритм оценивания переменной состоя-

ния x_k в случае, когда параметр λ модели (1) является известным. Данная работа продолжает исследования, описанные в работах [11, 14, 16, 17].

Актуальность разрабатываемого подхода

Распространенным подходом к решению задачи апостериорного оценивания параметра λ хаотического отображения (1) по зашумленным измерениям (2) является использование метода наименьших квадратов [5, 18], что предполагает решение задачи оптимизации

$$\min_{\lambda \in \Lambda, x_0 \in X_0} \sum_{k=1}^N (y_k - f^k(x_0, \lambda))^2. \quad (3)$$

При таком подходе основной сложностью является многоэкстремальность [17] целевой функции в задаче (3), которая возникает вследствие ряда причин.

1. Модели вида (1) описывают временные процессы, которые по своим характеристикам близки к шумам, составляющим ошибки измерений.

2. Сильная чувствительность хаотических процессов к малым изменениям параметра λ и начального условия x_0 , в результате чего целевая функция является "изрезанной".

3. На практике число измерений N может быть небольшим, это может привести к тому, что локальные экстремумы целевой функции будут мало отличаться друг от друга.

В конечном счете это приводит к необходимости применения трудоемких алгоритмов поиска глобального экстремума [18, 19]. В этом случае эффективным подходом может стать предварительная обработка измерений (2) с помощью алгоритма гарантированного оценивания для уточнения множества поиска Δ параметра λ . Это позволит уменьшить число локальных экстремумов целевой функции в задаче (3), а также сократить время вычислений при использовании алгоритмов глобальной оптимизации.

Одним из приемов, который применяется для преодоления проблемы многоэкстремальности целевой функции, является итерирование в обратном времени [5], т. е. решение задачи

$$\min_{\lambda \in \Lambda, x_N \in X_N} \sum_{k=0}^{N-1} (y_{N-k} - f^{(-k)}(x_N, \lambda))^2.$$

В данном случае алгоритм гарантированного оценивания может быть использован как для уточнения множества Λ , так и для задания множества X_N по результатам предварительной обработки измерений (2).

Алгоритм гарантированного оценивания параметра

Рассмотрим алгоритм нахождения информационных множеств Λ_k и X_k для параметра λ и переменной состояния x_k хаотического отображения (1) по зашумленным измерениям (2). Исходными данными для алгоритма являются априорные множественные оценки Λ_0 , X_0 и V_k для параметра λ , начального значения x_0 переменной состояния и ошибок v_k соответственно: $\lambda \in \Lambda_0$, $x_0 \in X_0$, $v_k \in V_k$. Информационное множество X_k на шаге k определим следующим образом (см., например, [14, 16]):

$$X_k = X_{k/k-1} \cap Y_k, \quad (4)$$

где $X_{k/k-1}$ — множество прогнозов; Y_k — множество, совместное с измерениями. Для построения множества прогнозов $X_{k/k-1}$ используют информационные множества X_{k-1} и Λ_{k-1} , найденные на предыдущем шаге:

$$X_{k/k-1} = \bigcup_{\lambda \in \Lambda_{k-1}} S(X_{k-1}, \lambda), \quad (5)$$

где $S(X_{k-1}, \lambda)$ — множество прогнозов, построенное для конкретного значения параметра λ :

$$S(X_{k-1}, \lambda) = \{x | x = f(t, \lambda), t \in X_{k-1}\}. \quad (6)$$

Множество, совместное с измерениями, находим исходя из априорно заданного множества V_k :

$$Y_k = \{x | x = y_k - v, v \in V_k\}. \quad (7)$$

Если в результате операции пересечения (4) $X_k \neq X_{k/k-1}$, то для некоторых значений параметра $\lambda \in \Lambda_{k-1}$ может оказаться, что $S(X_{k-1}, \lambda) \cap X_k = \emptyset$. В этом случае за счет исключения таких значений λ уточняется множественная оценка Λ_k параметра λ :

$$\Lambda_k = \{\lambda \in \Lambda_{k-1} | S(X_{k-1}, \lambda) \cap X_k \neq \emptyset\}. \quad (8)$$

Результат работы алгоритма является гарантированным: если исходные данные Λ_0 , X_0 и V_k заданы корректно, то на каждом шаге k истинные значения параметра λ и переменной состояния x_k принадлежат соответствующим информационным множествам Λ_k и X_k , найденным с помощью уравнений (4)–(8).

Численные эксперименты показали, что если на некотором шаге k реализующаяся ошибка измерений v_k оказывается близкой к одной из границ множества V_k , то на этом шаге возможно уточнение множественной оценки Λ_k параметра λ независимо от модели ошибок. Это условие не является достаточным: результат зависит также от реализации исходного процесса x_k . Таким образом, эффективность рассмотренного алгоритма зависит от того, на сколько множественные оценки V_k адекватны реально реализующимся ошибкам v_k . Как правило, множественные оценки V_k задаются постоянными для некоторого интервала времени: $V_k = V$. Выбор множества V можно организовать с помощью параллельных вычислений, рассматривая результаты работы алгоритма для разных значений V :

1. Если множество V задано неверно, т. е. на некотором шаге $v_k \notin V$, то на последующих шагах информационное множество X_k в результате операции пересечения (4) становится пустым: $X_k = \emptyset$. Такая ситуация также возникает, если неверно задано множество X_0 .

2. Если множество V задано слишком большим, то множество прогнозов (5) регулярно оказывается внутри множества, совместного с измерениями (7): $X_{k/k-1} \subset Y_k$.

3. Если описанных выше ситуаций не возникает, то критерием выбора является точность найденных множественных оценок для переменной состояния x_k . В качестве такого критерия можно использовать средний размер информационных множеств X_k на рассматриваемом временном интервале.

Численные эксперименты

Исследуем возможности описанного алгоритма гарантированного оценивания на примере квадратичного (логистического) отображения

$$x_{k+1} = f(x_k, \lambda) = \lambda x_k(1 - x_k), \quad (9)$$

которое имеет хаотические решения при $x_0 \in (0; 1)$ и $\lambda \in (3,569945\dots; 4]$ (за исключением тех значений параметра λ , которым соответствуют циклы отображения f). Пусть истинное значение параметра $\lambda = 3,7$, начальное условие $x_0 = 0,15$, число измерений $N = 50$ (рис. 1). Рассмотрим результаты работы алгоритма для разных моделей ошибок измерений v_k . Во всех случаях задаются следующие исходные данные:

$$\Lambda_0 = [3,57; 4], X_0 = [0; 1].$$

1. В качестве ошибок измерений v_k возьмем реализацию белого гауссовского шума (рис. 2) с нулевым математическим ожиданием и средним квадратическим отклонением $\sigma = 0,05$ (отношение сигнал/шум С/Ш = 12,1 дБ). Интервал измерений был разбит на пять одинаковых промежутков, на кото-

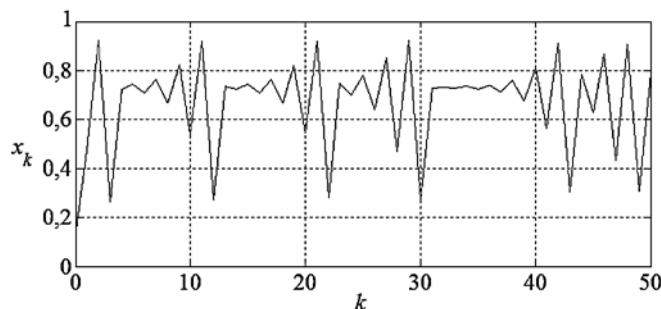


Рис. 1. Реализация квадратичного отображения

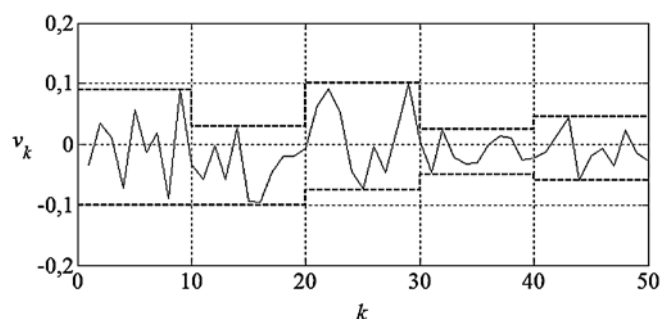


Рис. 2. Ошибки в виде реализации белого гауссовского шума

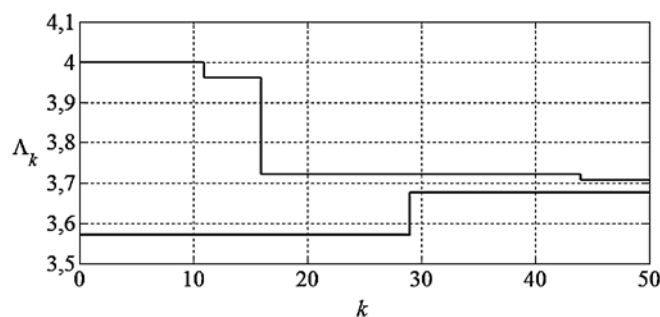


Рис. 3. Множественная оценка параметра

рых множественные оценки V_k для ошибок v_k выбирались следующим образом:

$$V_k = \begin{cases} [-0,100; 0,090], & k \in [1; 10]; \\ [-0,100; 0,030], & k \in [11; 20]; \\ [-0,075; 0,100], & k \in [21; 30]; \\ [-0,050; 0,025], & k \in [31; 40]; \\ [-0,060; 0,045], & k \in [41; 50]. \end{cases}$$

При обработке измерений множественная оценка Λ_k параметра λ уточняется при $k \in \{11, 16, 29, 44\}$ (рис. 3), конечная оценка

$$\Lambda_N = [3,6774; 3,7068].$$

Таким образом, начальную неопределенность (размер множества Λ_0) удается уменьшить в 14 раз.

2. Пусть ошибки измерений v_k являются хаотическим процессом, например, $v_k = 0,2x_k$, где x_k — реализация квадратичного отображения (9) при $\lambda = 3,65$ и $x_0 = 0,5$. Реализация ошибок показана на рис. 4 (С/Ш = 12,7 дБ). Зададим постоянную на всем интервале множественную оценку

$$V_k = [0,0583; 0,1825].$$

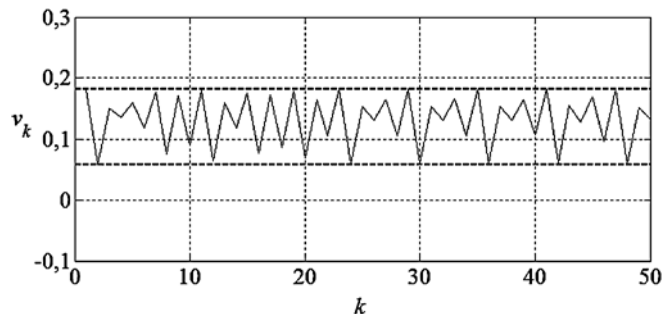


Рис. 4. Ошибки в виде реализации хаотического отображения

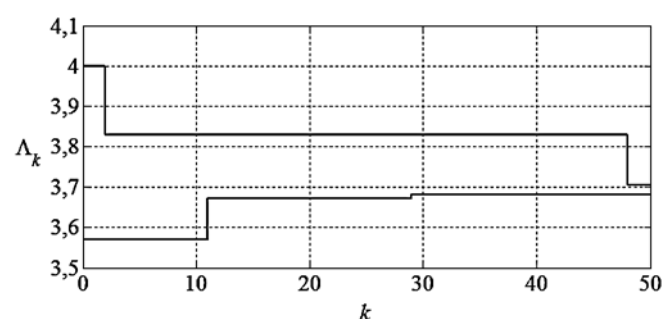


Рис. 5. Множественная оценка параметра

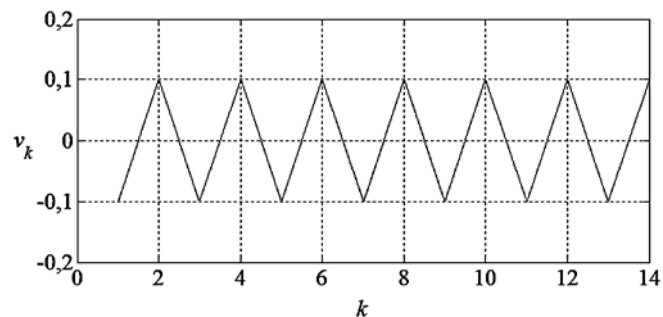


Рис. 6. Реализация ошибок измерений

В данном случае априорную оценку Λ_0 параметра λ удается уменьшить в 18 раз:

$$\Lambda_N = [3,6810; 3,7038].$$

При этом множественная оценка Λ_k параметра λ уточнялась при $k \in \{2, 11, 29, 48\}$ (рис. 5).

3. Полученные выше результаты подтверждают вывод о том, что множественную оценку Λ_k для параметра λ удается уточнить, если реализующаяся ошибка измерений v_k оказывается близкой к одной из границ множества V_k . Рассмотрим реализацию ошибок, значения которых выбираются только на границах множества $V_k = [-\alpha; \alpha]$, $\alpha > 0$, например, $v_k = \alpha \cdot \cos \pi k$ (рис. 6). На рис. 7 показаны множественные оценки Λ_k , полученные при $\alpha = 0,1$. За 13 шагов получена точная оценка параметра λ . При этом уже при $k = 2$ одна из границ множества Λ_k является истинным значением параметра λ . Расчеты показали, что для данного типа ошибок измерений параметр λ можно найти точно независимо от зна-

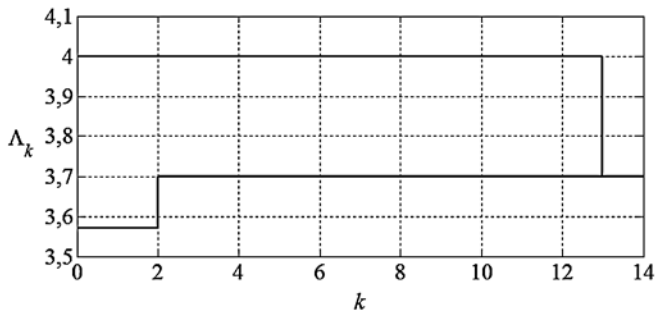


Рис. 7. Множественная оценка параметра

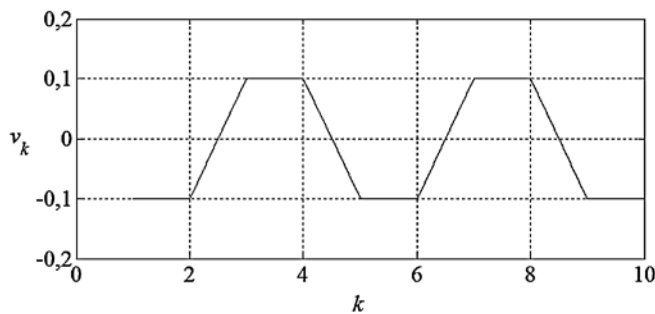


Рис. 8. Реализация ошибок измерений

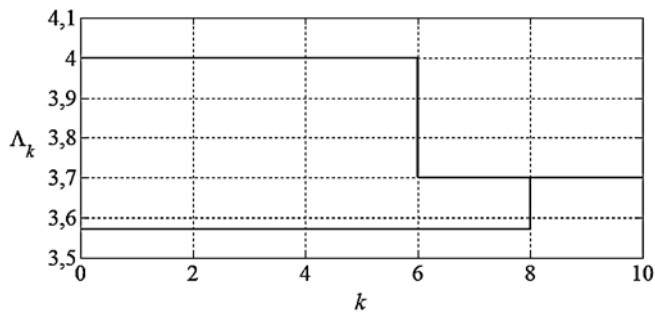


Рис. 9. Множественная оценка параметра

чения α , т. е. при значительно меньшем отношении $C/\text{Ш}$, чем в предыдущих случаях. Число шагов, за которое удается это сделать, зависит от конкретного правила, по которому выбирается значение ошибок на одной из границ множества V_k . Например, если $v_k = \alpha \cdot \cos\left[\frac{k}{2}\right]$ (рис. 8), то точная оценка будет найдена за 8 шагов (рис. 9).

Заключение

Рассмотренный в работе алгоритм гарантированного оценивания параметра λ хаотического отображения (1) по единственной зашумленной реализации измерений (2) имеет следующие особенности:

- не требуется выдвигать предположений о модели ошибок v_k ;
- алгоритм может быть использован при обработке измерений в реальном времени;
- результат работы алгоритма является гарантированным: при корректных исходных данных ис-

тинное значение параметра λ принадлежит информационному множеству Λ_k : $\lambda \in \Lambda_k$;

- алгоритм может быть использован для предварительной обработки измерений в целях уточнения множества возможных значений (множества поиска) параметра λ при применении метода наименьших квадратов, что позволяет уменьшить число локальных экстремумов целевой функции и сократить время вычислений при использовании алгоритмов глобальной оптимизации.

Проведенные эксперименты показали, что использование предложенного алгоритма гарантированного оценивания позволяет уменьшить априорное множество поиска параметра λ квадратичного отображения (9) в 14 раз, если ошибки измерений являются реализацией белого гауссовского шума ($C/\text{Ш} = 12,1$ дБ), и в 18 раз, если ошибки являются реализацией хаотического процесса ($C/\text{Ш} = 12,8$ дБ). Для некоторых моделей ошибок возможно за конечное число шагов получить точную оценку параметра λ при значительно меньшем отношении $C/\text{Ш}$.

Список литературы

1. Колесников А. А., Веселов Г. Е., Колесникова С. И. и др. Синерго-кибернетический подход к созданию хаосодинамических систем обработки и защиты информации // Информационные технологии. Приложение. 2013. № 10. С. 1–32.
2. Переварюха А. Ю. Новый метод компьютерного моделирования режима переходного хаоса // Информационные технологии. 2010. № 2. С. 18–25.
3. Буланчев В. А., Серков Л. А. Модельный подход к самоорганизующимся системам с детерминированным хаосом // Информационные технологии. 2006. № 7. С. 48–53.
4. Никульчев Е. В. Моделирование систем с нелинейной динамикой на основании экспериментальных данных // Мехатроника, автоматизация, управление. 2006. № 5. С. 6–14.
5. Смирнов Д. А., Власкин В. С., Пономаренко В. И. Метод оценки параметров одномерных отображений по хаотическим временным рядам // Письма в ЖТФ. 2005. Т. 31, № 3. С. 18–26.
6. Voss H. U., Timmer J., Kurths J. Nonlinear dynamical system identification from uncertain and indirect measurements // International Journal of Bifurcation and Chaos. 2004. V. 14, N. 6. P. 1905–1933.
7. Садовничий В. А., Козодеров В. В., Ушакова Л. А., Ушаков С. А. Предсказуемость глобальных и региональных явлений в природе и обществе // Вестник ОГГГН РАН. 2000. № 1. С. 84–113.
8. Huang N. E., Shen Z., Long S. R. et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis // Proc. of the Royal Society of London A. 1998. V. 454. P. 903–995.
9. Vitanov N. K., Sakai K., Dimitrova Z. I. SSA, PCA, TDPSC, ACFA: Useful combination of methods for analysis of short and nonstationary time series // Chaos, Solitons & Fractals. 2008. V. 37, N. 1. P. 187–202.
10. Чернов В. М. Арифметические методы синтеза быстрых алгоритмов дискретных ортогональных преобразований. М.: Физматлит, 2007. 264 с.
11. Шелудько А. С., Ширяев В. И. Совместное использование фильтра Калмана и минимаксного фильтра в задаче оценивания параметров модели хаотического процесса // Вестник ЮУрГУ. Серия "Компьютерные технологии, управление, радиоэлектроника". 2012. № 35. С. 59–64.
12. Тратас Ю. Г. Применение методов статистической теории связи к задачам приема хаотических колебаний // Успехи современной радиоэлектроники. 1998. № 11. С. 57–80.

13. Куржанский А. Б. Задача идентификации — теория гарантированных оценок // Автоматика и телемеханика. 1991. № 4. С. 3—26.

14. Ширяев В. И. Алгоритмы управления динамическими системами в условиях неопределенности // Мехатроника. 2001. № 8. С. 2—5.

15. Simon D., El-Sherief H. Hybrid Kalman/minimax filtering in phase-locked loops // Control Engineering Practice. 1996. Vol. 4, N. 5. P. 615—623.

16. Шелудко А. С., Ширяев В. И. Алгоритм минимаксной фильтрации для одномерного хаотического процесса // Мехатроника, автоматизация, управление. 2014. № 5. С. 8—12.

17. Елсаков С. М., Ширяев В. И. О многоэкстремальности в задачах оценивания систем детерминированного хаоса // Вестник ЮУрГУ. Серия "Компьютерные технологии, управление, радиоэлектроника". 2009. № 3. С. 37—41.

18. Banerjee A., Abu-Mahfouz I. A comparative analysis of particle swarm optimization and differential evolution algorithms for parameter estimation in nonlinear dynamic systems // Chaos, Solitons & Fractals. 2014. V. 58. P. 65—83.

19. Карпенко А. П. Популяционные алгоритмы глобальной поисковой оптимизации. Обзор новых и малоизвестных алгоритмов // Информационные технологии. Приложение. 2012. № 7. С. 1—32.

A. S. Sheludko, Engineer of Applied Mathematics Chair, sheludkoas@susu.ac.ru,
 V. I. Shiryaev, PhD, Professor, Head of the Chair, vis@susu.ac.ru
 South Ural State University, Chelyabinsk, Russia

The Algorithm of Guaranteed Parameter Estimation for One-Dimensional Chaotic Map

In this article, we consider the problem of parameter estimation from a single noisy realization of one-dimensional chaotic map. Based on the guaranteed approach the proposed algorithm finds the interval estimates of uncertain variables (parameter and state variable). The key features of the algorithm:

1. The a priori information about the measurement errors is presented only as an interval of possible values. It is not necessary to assume the model of measurement errors or its statistical properties.
2. The algorithm is recurring procedure that can be applied in real-time processing.
3. The result of computations is guaranteed: at every time step interval parameter estimate (information set) always contains the true value of the unknown parameter.
4. If the parameter estimation problem is solved by the least squares method, the algorithm can be used to specify the search set for the parameter. It decreases the number of local minimums of the multiextremal cost function.

We present numerical experiments for logistic map and different types of measurement errors.

Keywords: chaotic map, parameter estimation, guaranteed approach

References

1. Kolesnikov A. A., Veselov G. E., Kolesnikova S. I. et al. Синерго-кибеметический подход к созданию хаосодинамических систем обработки и защиты информации. *Информационные технологии. Приложение*. 2013. N. 10. P. 1—32.
2. Perevaryukha A. Yu. Novyy metod komp'yuternogo modelirovaniya rezhima perekhodnogo khaosa. *Информационные технологии*. 2010. N. 2. P. 18—25.
3. Bulanichev V. A., Serkov L. A. Model'nyy podkhod k samoorganizuyushchimsya sistemam s determinirovannym khaosom. *Информационные технологии*. 2006. № 7. P. 48—53.
4. Nikul'chev E. V. Modelirovanie sistem s nelineynoy dinamikoy na osnovanii eksperimental'nykh dannykh. *Mekhatronika, avtomatizatsiya, upravlenie*. 2006. N. 5. P. 6—14.
5. Smirnov D. A., Vlaskin V. S., Ponomarenko V. I. Metod otsenki parametrov odnomernykh otobrazheniy po khaoticheskim vremennym ryadam. *Pis'ma v ZhTF*. 2005. V. 31, N. 3. P. 18—26.
6. Voss H. U., Timmer J., Kurths J. Nonlinear dynamical system identification from uncertain and indirect measurements. *International Journal of Bifurcation and Chaos*. 2004. V. 14, N. 6. P. 1905—1933.
7. Sadovnichiy V. A., Kozoderov V. V., Ushakova L. A., Ushakov S. A. Predskazuemost' global'nykh i regional'nykh yavleniy v prirode i obshchestve. *Vestnik OGGGGN RAN*. 2000. N. 1. P. 84—113.
8. Huang N. E. The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis. Proc. of the Royal Society of London A. 1998. V. 454. P. 903—995.
9. Vitanov N. K., Sakai K., Dimitrova Z. I. SSA, PCA, TDPSC, ACFA: Useful combination of methods for analysis of short and nonstationary time series. *Chaos, Solitons & Fractals*. 2008. V. 37. N. 1. P. 187—202.
10. Chernov V. M. Arifmeticheskie metody sinteza bystrykh algoritmov diskretnykh ortogonal'nykh preobrazovaniy. M.: Fizmatlit, 2007. 264 p.
11. Sheludko A. S., Shiryaev V. I. Sovmestnoe ispol'zovanie fil'tra Kalmana i minimaksnogo fil'tra v zadache otsenivaniya parametrov modeli khaoticheskogo protsessa. *Vestnik YuUrGU. Seriya. "Komp'yuternye tekhnologii, upravlenie, radioelektronika"*. 2012. N. 35. P. 59—64.
12. Tratas Yu. G. Primenenie metodov statisticheskoy teorii svyazi k zadacham priema khaoticheskikh kolebaniy. *Uspekhi sovremennoy radioelektroniki*. 1998. N. 11. P. 57—80.
13. Kurzhanskiy A. B. Zadacha identifikatsii — teoriya garantirovannykh otsenok. *Avtomatika i telemekhanika*. 1991. N. 4. P. 3—26.
14. Shiryaev V. I. Algoritmy upravleniya dinamicheskimi sistemami v usloviyakh neopredelennosti. *Mekhatronika*. 2001. N. 8. P. 2—5.
15. Simon D., El-Sherief H. Hybrid Kalman/minimax filtering in phase-locked loops // Control Engineering Practice. 1996. V. 4, N. 5. P. 615—623.
16. Sheludko A. S., Shiryaev V. I. Algoritm minimaksnoy fil'tratsii dlya odnomernogo khaoticheskogo protsessa. *Mekhatronika, avtomatizatsiya, upravlenie*. 2014. N. 5. P. 8—12.
17. Elnakov S. M., Shiryaev V. I. O mnogoekestremal'nosti v zadachakh otsenivaniya sistem determinirovannogo khaosa. *Vestnik YuUrGU. Seriya "Komp'yuternye tekhnologii, upravlenie, radioelektronika"*. 2009. N. 3. P. 37—41.
18. Banerjee A., Abu-Mahfouz I. A comparative analysis of particle swarm optimization and differential evolution algorithms for parameter estimation in nonlinear dynamic systems. *Chaos, Solitons & Fractals*. 2014. V. 58. P. 65—83.
19. Karpenko A. P. Populyatsionnye algoritmy global'noy poiskovoy optimizatsii. Obzor novykh i maloizvestnykh algoritmov. *Информационные технологии. Приложение*. 2012. N. 7. P. 1—32.

А. А. Бойко, канд. техн. наук, доц., зам. нач. отдела, algeminy@mail.ru,

Военный учебно-научный центр Военно-воздушных сил

"Военно-воздушная академия имени профессора Н. Е. Жуковского и Ю. А. Гагарина" (г. Воронеж)

Способ стратифицированного аналитического описания процесса функционирования информационно-технических средств

Предложен способ математического описания процесса функционирования информационно-технических средств в виде совокупности взаимосвязанных страт протекающих в них частных процессов.

Ключевые слова: стратификация, информационно-техническое средство, полумарковский процесс

Введение

Информационно-технические средства являются основой современных организационно-технических систем. Термин "информационно-техническое средство" является собирательным. Он выделяет широкий спектр технических средств, участвующих в процессе создания и выполнения операций с данными, из которых может быть получена информация. Информационно-техническим средством (ИТС) может быть радиоэлектронное средство или средство вычислительной техники, комбинация таких средств друг с другом и (или) с другими видами технических средств, определенными в ГОСТ [1]. Персональная ЭВМ с встроенным Wi-Fi-адаптером может являться примером ИТС, комбинирующего в себе средство вычислительной техники и радиоэлектронное средство.

ИТС могут решать информационно-расчетные задачи (ИРЗ), формализующие информацию и реализующие специальные математические расчеты, коммуникационные задачи (КЗ), обеспечивающие информационное взаимодействие с другими ИТС, и задачи непосредственного управления исполнительными устройствами (УЗ). Для средства вычислительной техники, например, исполнительным устройством может быть принтер. Однако принтер в зависимости от уровня развития своего технического и программного обеспечения также может рассматриваться как ИТС. При этом будет иметь место не управление исполнительным устройством, а информационное взаимодействие двух ИТС.

В процессе конфликтного взаимодействия организационно-технических систем, включая стадию их антагонистического противоборства, ИТС играют двойственную роль [2]. С одной стороны, они обеспечивают повышение производительности своей организационно-технической системы, а с другой стороны, подвержены влиянию деструктивных факторов: информационно-техническим (программным и электромагнитным) воздействиям (ИТВ) [3], приводящим к нарушению конфиденциальности, целостности и доступности обрабатываемой ин-

формации и алгоритмов ее обработки, физическому износу и уничтожению. Поэтому при проектировании и эксплуатации ИТС необходимо обосновывать требования к параметрам их функционирования с учетом этой двойственности.

1. Постановка задачи

Одним из наиболее распространенных методов обоснования требований к параметрам функционирования сложных систем является математическое моделирование [4]. Однако на сегодняшний день применение данного метода связано для разработчиков ИТС со следующими ограничениями.

С практической точки зрения, известные модели функционирования ИТС [5–8] и аналогичные им либо не учитывают специфику выполняемых ИТС задач, либо представляют эти задачи однотипными. В данных моделях не принимаются во внимание особенности функционирования технического (ТО), общего программного и информационного обеспечения (операционной системы, баз данных и систем управления ими, базовой системы ввода–вывода и др., далее — ОПИО), а из деструктивных факторов учитываются только электромагнитные воздействия, приводящие к нарушению доступности информации. В результате осуществить адекватную оценку практически значимой номенклатуры показателей функционирования ИТС с учетом всего спектра выполняемых этими средствами задач, качества входящих в них компонентов и комплексного влияния деструктивных факторов с применением известных моделей невозможно.

С методической точки зрения ИТС относятся к таким сложным системам, многообразие ключевых процессов в которых, с одной стороны, не представляется возможным аналитически описать в виде причинно-следственных моделей параллельных действий, а с другой стороны, их недостаточно рассматривать только в статике или в установившихся режимах с использованием графо-сигнальных моделей [4]. Для аналитического описания функционирования систем такого класса в наибольшей степени

применим математический аппарат полумарковских процессов [5, 9, 10]. Однако число ключевых процессов функционирования в типовом ИТС настолько велико, что разработка для него аналитического описания в виде единой детальной полумарковской модели неизбежно приводит к "взрыву" пространства состояний.

С учетом изложенного актуальной является задача разработки способа аналитического описания совокупности процессов функционирования информационно-технических средств, базирующегося на применении математического аппарата полумарковских процессов и парирующего "взрыв" пространства состояний при учете всего спектра выполняемых этими средствами задач, качества входящих в их состав компонентов и комплексного влияния деструктивных факторов.

2. Способ решения задачи

Для решения поставленной задачи воспользуемся изложенным в работе [1] методом стратификации, заключающемся в рассмотрении сложной системы в различных слоях ее функционального пространства, называемых стратами, и анализе каждой страты без учета влияния остальных страт.

Предлагаемый способ стратифицированного аналитического описания процесса функционирования ИТС состоит в выполнении следующих шагов.

Шаг 1. Выделяют множество страт функционирования ИТС. Одна из страт может характеризовать специализированный суперпроцесс, координирующий некоторое подмножество страт.

Шаг 2. Разрабатывают аналитическое описание страт. Страты описывают в виде полумарковских процессов, а суперпроцесс в виде системы с очередями.

Ввиду того, что произвольная плотность распределения неотрицательной случайной величины с достаточной степенью точности аппроксимируется с помощью обобщенного закона Эрланга n -го порядка, плотность распределения времени нахождения полумарковского процесса в каждом состоянии целесообразно описывать с использованием этого закона:

$$E_{s,v}(t) = (-1)^{n-1} \prod_{i=0}^{n-1} \lambda_{s,v,i} \sum_{j=0}^{n-1} \frac{e^{-\lambda_{s,v,j}t}}{\prod_{k=0, k \neq j}^{n-1} (\lambda_{s,v,j} - \lambda_{s,v,k})}, \quad (1)$$

где $E_{s,v}(t)$ — плотность распределения времени нахождения системы в s -м состоянии до перехода в v -е состояние; $\lambda_{s,v,i}$ — параметр обобщенного закона Эрланга n -го порядка.

Следует отметить, что на практике 2-й порядок закона (1) является достаточным для исследования систем рассматриваемого класса.

Шаг 3. Формируют группы состояний. Число групп полагают равным числу исследуемых характеристик объекта. Состояния включают в группу по критерию их непосредственного влияния на некоторую характеристику исследуемого объекта. Каждая группа может включать состояния различных страт и учитывать влияние суперпроцесса.

Шаг 4. Определяют вероятностно-временные характеристики страт.

К числу наиболее практичных следует отнести два метода аналитического описания полумарковских процессов. Первый из них описан в работах [5, 9] и позволяет решить задачу определения вероятности нахождения исследуемой системы в каждом состоянии в любой момент времени, не прибегая к численным методам и используя в качестве исходных данных вероятности переходов из каждого состояния в последующие и плотности распределения вероятностей времен этих переходов. Второй метод изложен в работе [10] и позволяет решить указанную задачу только на основе плотностей распределения вероятностей времен переходов, но требует при этом применения численных методов. Поскольку вычислительные мощности современных ЭВМ позволяют обеспечить приемлемую точность полученных численными методами решений, а минимизация исходных данных имеет приоритетное значение, за основу при реализации данного шага представляется целесообразным взять второй метод.

Шаг 5. Вычисляют значения показателей групп состояний. Для этого применяют следующие правила: а) значения учитываемых в показателе вероятностно-временных характеристик состояний, принадлежащих одному полумарковскому процессу, складываются; б) суммарные значения вероятностно-временных характеристик разных полумарковских процессов перемножаются; в) значения характеристик суперпроцесса учитываются в множителях, описывающих состояния зависящих от него полумарковских процессов. Данные правила обусловлены тем, что в каждом полумарковском процессе совокупность всех его состояний образует полную группу событий, а частные процессы считаются независимыми. Поэтому уместно применение правил сложения и умножения вероятностей.

Рассмотрим результаты применения предложенного способа на примере ИТС, которое может одновременно выполнять ИРЗ, КЗ и УЗ.

3. Пример решения задачи

Относительно самостоятельные частные процессы функционирования ИТС представим на рис. 1 в виде стратифицированной структуры.

Данная структура включает модели ИРЗ, КЗ и УЗ, модель диспетчера, являющегося суперпроцессом и координирующего выполнение задач, а также модели ТО, ОПИО и информационного конфликта между реализующими программное воздействие специальными программными средствами (СПС) и подсистемой защиты информации (ПЗИ) ИТС. Характер взаимовлияния частных процессов в рассматриваемом ИТС таков, что его модели задач являются независимыми, а модели ТО, ОПИО и информационного конфликта СПС и ПЗИ одновременно влияют друг на друга и на модели задач.

Рассмотрим оригинальные описания показанных на рис. 1 моделей.

Модель ТО с учетом процессов, характеризующих надежность технических систем [12], и особенностей реализации потенциально возможных ИТВ представим на рис. 2 в виде графа состояний.

Граф динамики функционирования ТО включает шесть состояний: T_1 — нормальное функционирование; T_2 — ТО функционирует с нулевой производительностью (например, по причине и в течение воздействия мощного электромагнитного излучения на его электронные компоненты или схемы с обратимым эффектом); T_3 — произошел сбой (самовосстанавливающийся частичный отказ, обусловленный, например, некачественной пайкой или внутренними дефектами элементов устройств); T_4 — произошел отказ ТО (например, по причине перегорания электронного компонента в результате его износа или воздействия мощного электромагнитного излучения с необратимым эффектом); T_5 — проводится ремонт; T_6 — уничтожено.

Модель ОПИО с учетом процессов, характеризующих надежность информационных систем [12], в наиболее часто встречающихся на практике условиях, когда устранение ошибок в нем невозможно, и требования к функционированию в режиме реального времени не предъявляются, представим в виде графа со следующими состояниями: C_1 — нормальное функционирование; C_2 — зависание (самовосстанавливающийся частичный отказ, перезагрузка не требуется); C_3 — сбой (самовосстанавливающийся частичный отказ, требуется перезагрузка); C_4 — отказ (например, по причине ошибки в ее программном коде); C_5 — восстановление. Граф этого процесса показан на рис. 3.

Модель ИРЗ аналогична модели ОПИО с учетом того, что состояния именованы буквой "R", а вероятности соответствующих состояний имеют вид $P^r(t)$. Исходя из практики, надежность функционирования



Рис. 1. Стратифицированная структура процесса функционирования информационно-технического средства

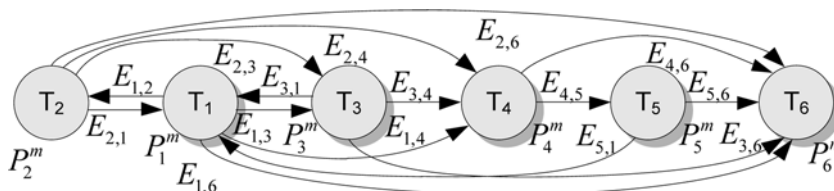


Рис. 2. Граф состояний, отражающий динамику функционирования технического обеспечения

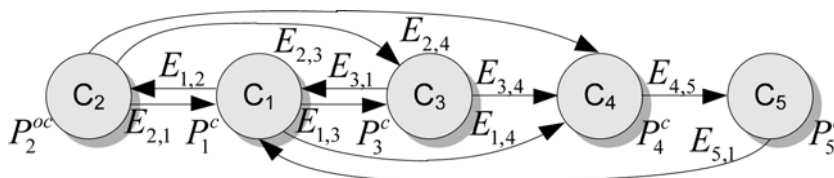


Рис. 3. Граф состояний, отражающий динамику функционирования общего программного и информационного обеспечения

ОПИО может существенно отличаться от надежности функционирования специального программного и информационного обеспечения (СПИО), реализующего ИРЗ.

Модель информационного конфликта СПС и ПЗИ, учитывающая известные из практики особенности программного воздействия, показана на рис. 4. Состояния графа этой модели имеют следующие описания: S_1 — исходное работоспособное состояние ИТС; S_2 — получен физический доступ к ИТС, анализируются механизмы технического доступа к ИТС на физическом уровне эталонной модели взаимодействия открытых систем (ЭМВОС); S_3 — ПЗИ нейтрализует возможность получения физического доступа к ИТС; S_4 — получен технический доступ к ИТС, проводится анализ преамбул протоколов канального уровня ЭМВОС; S_5 — ПЗИ нейтрализует возможность получения технического доступа к ИТС; S_6 — СПС получило возможность использовать в своих целях управляющую информацию протоколов канального уровня ЭМВОС; S_7 — ПЗИ нейтрализует возможность доступа к протоколам информационного взаимодействия ИТС на сетевом и вышестоящих уровнях ЭМВОС; S_8 — СПС получило доступ к протоколам информационного взаимодействия ИТС на сетевом и выше-

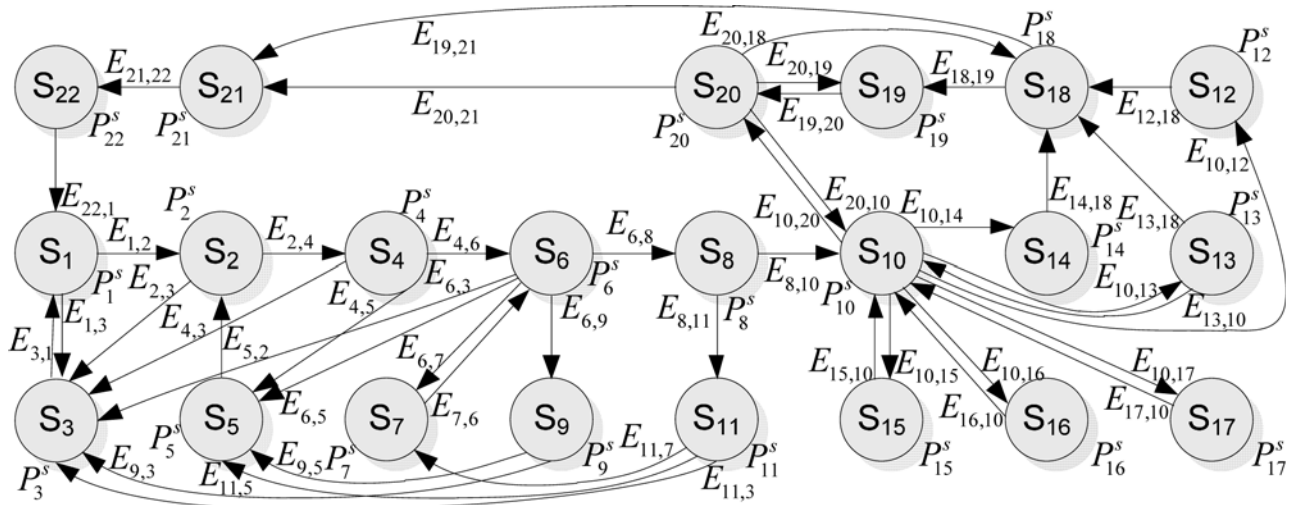


Рис. 4. Граф состояний, отражающий динамику конфликта специальных программных средств и подсистемы защиты информации

стоящих уровнях ЭМВОС; S_9 — СПС скрыто нарушает доступность информации ИТС на канальном уровне ЭМВОС; S_{10} — СПС внедрилось в программную среду ИТС и анализирует его текущее состояние; S_{11} — СПС скрыто нарушает доступность информации на сетевом, транспортном, сеансовом и представительском уровнях ЭМВОС при выполнении ИТС КЗ, не влияя на выполнение ИРЗ и УЗ; S_{12} — СПС открыто осуществляет отказ в обслуживании ИТС, приводящий к невозможности выполнения любых задач; S_{13} — СПС открыто применяет ИТС в своих целях; S_{14} — СПС открыто осуществляет вывод ИТС из строя; S_{15} — СПС размножается; S_{16} — СПС дезинформирует ИТС путем уничтожения или подмены программ, информации в базах данных, передаваемых сообщений, в том числе в интересах обеспечения бескомпроматности; S_{17} — СПС проводит разведку путем чтения и анализа программ, информации в базах данных, передаваемых сообщений; S_{18} — ПЗИ проводит принудительный поиск СПС; S_{19} — СПС осуществляет самомодификацию; S_{20} — СПС переходит в режим ожидания; S_{21} — выполняются регламентные работы по поддержанию работоспособности ИТС; S_{22} — проводится подготовка ИТС к работе.

Модель диспетчера, определяющего дисциплину обслуживания ИРЗ, КЗ и УЗ, представим в виде показанной на рис. 5 простейшей одноканальной системы массового обслуживания типа М/М/1 с множественным доступом и дисциплиной FIFO [13].

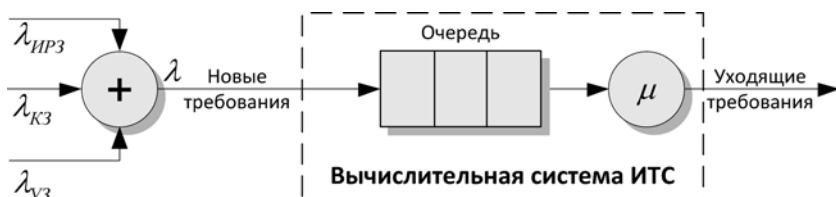


Рис. 5. Модель диспетчера задач информационно-технического средства

Среднее число выполняемых в ИТС задач определяется [13] формулой:

$$N = \frac{\lambda_{ИРЗ} + \lambda_{КЗ} + \lambda_{УЗ}}{\mu - \lambda_{ИРЗ} - \lambda_{КЗ} - \lambda_{УЗ}}, \quad (2)$$

где $\lambda_{ИРЗ}$, $\lambda_{КЗ}$, $\lambda_{УЗ}$ — интенсивности поступления на выполнение ИРЗ, КЗ и УЗ, соответственно; μ — интенсивность выполнения задач в ИТС. При этом предполагается, что ИТС эксплуатируется в режиме, не превышающем его пропускной способности, т. е. $0 \leq (\lambda_{ИРЗ} + \lambda_{КЗ} + \lambda_{УЗ}) \leq \mu$.

Оценить с учетом (2) среднее число выполняющихся в ИТС задач каждого типа возможно с использованием следующих выражений:

$$N_{ИРЗ} = N \frac{\lambda_{ИРЗ}}{\lambda_{ИРЗ} + \lambda_{КЗ} + \lambda_{УЗ}} = \frac{\lambda_{ИРЗ}}{\mu - \lambda_{ИРЗ} - \lambda_{КЗ} - \lambda_{УЗ}},$$

$$N_{УЗ} = \frac{\lambda_{УЗ}}{\mu - \lambda_{ИРЗ} - \lambda_{КЗ} - \lambda_{УЗ}}, \quad (3)$$

$$N_{КЗ} = \frac{\lambda_{КЗ}}{\mu - \lambda_{ИРЗ} - \lambda_{КЗ} - \lambda_{УЗ}},$$

где $N_{ИРЗ}$, $N_{УЗ}$ и $N_{КЗ}$ — среднее число задач, выполняющихся в ИТС ИРЗ, УЗ и КЗ, соответственно. ИТС может не выполнять задачи какого-либо типа. В таком случае интенсивность поступления на выполнение задачи соответствующего типа будет равна нулю.

Модель КЗ описывает передачу сообщений другим ИТС по различным доступным линиям связи в условиях ИТВ. За ее основу возьмем результаты работы [14], в которой предложена дискретно-событийная модель конфликта направления радиосвязи между двумя абонентами, состоящая из нескольких объединенных единой системой управ-

ления линий радиосвязи и системы радиопротиводействия. Данная модель позволяет оценить вероятность $K_0(t)$ передачи сообщения в любой момент времени с учетом адаптивного применения различных мер помехозащиты. Несмотря на ориентацию представленной в работе [14] модели на радиосвязь, ее без изменений можно применить ко всему спектру беспроводных и проводных коммуникационных каналов, используемых ИТС. С учетом того, что в работе [14] рассматривается преимущественно конфликт коммуникационных средств и средств подавления связи на физическом уровне ЭМВОС, в условиях дополнительного программного воздействия на вышестоящих уровнях этой модели вероятность передачи сообщения для ИТС будет отличаться от значения $K_0(t)$. Полагая, что ИТВ на физическом и вышестоящих уровнях ЭМВОС в интересах минимизации расходуемых ресурсов могут осуществляться одновременно по разным каналам одного ИТС, вероятность передачи сообщения КЗ в любой момент времени может быть определена с учетом описанной выше модели информационного конфликта СПС и ПЗИ на основе использования формулы вероятности произведения двух совместных событий:

$$P_{K3}(t) = K_0(t) \left(1 - \sum_i P_i^s(t) \right), \quad (4)$$

где $P_{K3}(t)$ — вероятность передачи сообщения КЗ; $P_i^s(t)$ — вероятность нахождения конфликта СПС и ПЗИ ИТС в i -м состоянии при $i = 9, 11, 12, 14, 21, 22$.

Модель УЗ имеет следующую особенность. При выполнении УЗ задействовано как соответствующее СПИО, так и управляемые ИТС исполнительные устройства. Поэтому модель выполнения этой задачи, по своей сути, интегрирует в себе состояния модели функционирования ОПИО (или ИРЗ) и состояния модели функционирования ТО. Граф состояний такой модели представлен на рис. 6, его состояния имеют следующее описание: U_1 — нормальное выполнение; U_2 — зависание СПИО; U_3 — зависание исполнительного устройства; U_4 — сбой СПИО; U_5 — сбой исполнительного устройства;

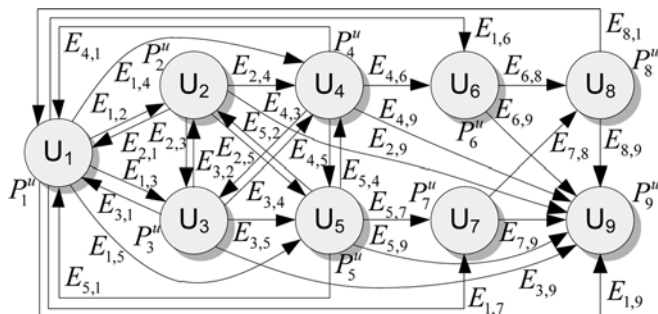


Рис. 6. Граф состояний, отражающий динамику функционирования управляющей задачи

U_6 — отказ СПИО; U_7 — отказ исполнительного устройства; U_8 — восстановление работоспособности; U_9 — исполнительное устройство уничтожено.

Изложенные частные модели позволяют с учетом (2)—(4) определить следующие важные для организационно-технической системы относительно самостоятельные показатели функционирования ИТС:

- эффективность $E(t)$, определяемая как вероятность успешного выполнения ИРЗ, КЗ и УЗ с использованием выражения

$$E(t) = P_1^t(t) P_1^c(t) \sum_j P_j^s(t) \frac{1}{N} [N_{ИРЗ} P_1^r(t) Q_{ИРЗ} + N_{КЗ} P_{КЗ}(t) Q_{КЗ} + N_{УЗ} P_1^u(t) Q_{УЗ}], \quad (5)$$

где j — индекс состояния конфликта СПС и ПЗИ ИТС, в котором работоспособность ИТС не нарушается ($j = 1...8, 10, 13, 15...20$); $Q_{ИРЗ}$, $Q_{КЗ}$, и $Q_{УЗ}$ — показатели качества выполнения задач ИТС, определяемые в интервале $[0, 1]$;

- способность к заражению других ИТС $Z(t)$, определяемая как вероятность нахождения конфликта СПС и ПЗИ ИТС в состоянии S_{15} с учетом работоспособности ТО, ОПИО и КЗ с использованием выражения:

$$Z(t) = P_1^t(t) \times P_1^c(t) \times P_{15}^s(t) \times P_{K3}(t) \times Q_{K3}; \quad (6)$$

- подверженность дезинформации $D(t)$, определяемая как вероятность нахождения конфликта СПС и ПЗИ ИТС в состоянии S_{16} с учетом работоспособности ТО, ОПИО и КЗ с использованием выражения

$$D(t) = P_1^t(t) \times P_1^c(t) \times P_{16}^s(t); \quad (7)$$

- способность быть разведанным $R(t)$, определяемая как вероятность нахождения конфликта СПС и ПЗИ ИТС в состоянии S_{17} с учетом работоспособности ТО, ОПИО и КЗ с использованием выражения

$$R(t) = P_1^t(t) \times P_1^c(t) \times P_{17}^s(t); \quad (8)$$

- способность принадлежать своей ОТС $A(t)$, определяемая как вероятность нахождения конфликта СПС и ПЗИ ИТС в состоянии S_{13} с учетом работоспособности ТО, ОПИО и КЗ с использованием выражения:

$$A(t) = 1 - P_1^t(t) \times P_1^c(t) \times P_{13}^s(t) \times P_1^u(t) \times Q_{УЗ}. \quad (9)$$

В качестве примера определения вероятностно-временных характеристик частных процессов функционирования ИТС рассмотрим ИТС, в котором постоянно выполняются 2 ИРЗ, 2 КЗ и 1 УЗ, $Q_{ИРЗ} = Q_{КЗ} = Q_{УЗ} = 1$, а $K_0(t)$ является неизменным во времени и равным 0,85. Полагая, что времена переходов частных моделей функционирования

Пример исходных данных

Частная модель	s	v	τ, мин	s	v	τ, мин	s	v	τ, мин	s	v	τ, мин	s	v	τ, мин
ТО	1	2	1440	1	6	4320	2	4	43200	3	4	43200	4	6	4320
	1	3	10080	2	1	1	2	6	4320	3	6	4320	5	1	180
	1	4	43200	2	3	10080	3	1	15	4	5	1	5	6	4320
ОПИО	1	2	1440	1	4	4320	2	3	10080	3	1	15	4	5	1
	1	3	10080	2	1	1	2	4	4320	3	4	4320	5	1	180
Конфликт СПС и ПЗИ ИТС	1	2	1	6	3	1440	9	5	30	11	5	30	18	19	1440
	1	3	1440	6	5	1440	10	12	15	11	7	1440	18	21	60
	2	3	1440	6	7	1440	10	13	15	12	18	30	19	20	0,017
	2	4	1	6	8	30	10	14	1440	13	10	15	20	10	15
	3	1	30	6	9	60	10	15	15	13	18	15	20	18	1440
	4	3	1440	7	6	15	10	16	15	14	18	15	20	19	1440
	4	5	1440	8	10	1	10	17	15	15	10	15	20	21	1440
	4	6	1	8	11	60	10	20	0,017	16	10	60	21	22	180
5	2	1	9	3	1440	11	3	1440	17	10	30	22	1	30	
ИРЗ	1	2	30	1	4	180	2	3	120	3	1	15	4	5	1
	1	3	120	2	1	1	2	4	180	3	4	180	5	1	60
УЗ	1	2	30	2	1	1	3	4	120	4	9	4320	6	9	4320
	1	3	1440	2	3	1440	3	5	10080	5	1	15	7	8	1
	1	4	120	2	4	120	3	9	4320	5	2	30	7	9	4320
	1	5	10080	2	5	10080	4	1	15	5	4	120	8	1	180
	1	6	180	2	9	4320	4	3	1440	5	7	43200	8	9	4320
	1	7	43200	3	1	1	4	5	10080	5	9	4320			
	1	9	4320	3	2	30	4	6	180	6	8	1			

Примечание: τ — среднее время до перехода из s-го в v-е состояние.

ния ИТС распределены по обобщенному закону Эрланга 2-го порядка со средними значениями, приведенными в таблице, можно получить графические зависимости, представленные на рис. 7—11.

Из рис. 7 видно, что в течение недели в заданных условиях функционирования ИТС вероятность нор-

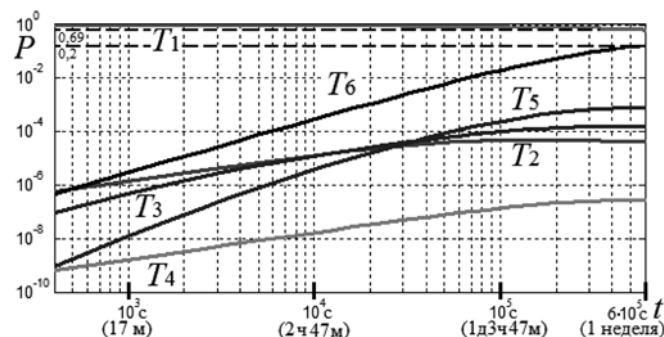


Рис. 7. Динамика состояний технического обеспечения

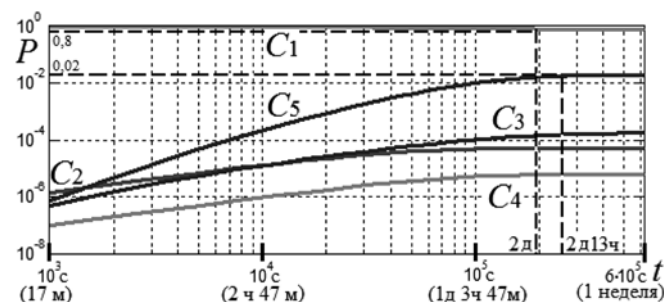


Рис. 8. Динамика состояний общего программного и информационного обеспечения

мального функционирования ТО снижается более чем на 30 %. При этом его вероятность быть уничтоженным к концу анализируемого периода увеличивается до 0,2.

На рис. 8 видно, что в течение первых 2 дней вероятность нормального функционирования ОПИО снижается до 0,8 и далее до конца недели остается неизменной. При этом уже через 2 дня 13 ч вероятность нахождения ОПИО в состоянии восстановления установится на уровне 0,02.

Зависимости на рис. 9, а и б показывают, что для указанных в таблице характеристик информационного конфликта СПС и ПЗИ в течение первых 2 ч ИТС практически не будет подвергаться воздействию СПС (кривые S₉, S₁₁—S₁₄, S₁₆ и S₁₇). После этого СПС будет активно воздействовать на ИТС со среднестатистической вероятностью 0,04 и размножаться с вероятностью 0,018 (кривая S₁₅ на рис. 9, б).

На рис. 10 показано, что в течение первых 3 ч вероятность нормального функционирования ИРЗ устанавливается на уровне 0,4, что в 2 раза меньше, чем у ОПИО. Через 8 ч вероятность нахождения ИРЗ в состоянии восстановления устанавливается на уровне 0,08, что в 4 раза больше, чем у ОПИО.

На рис. 11 видно, что уже через 1,5 часа вероятность нормального выполнения УЗ будет на уровне 0,1. При этом вероятность уничтожения исполнительного устройства с этого момента до окончания временного интервала моделирования монотонно возрастает с 0,05 до 0,95, что свидетельствует

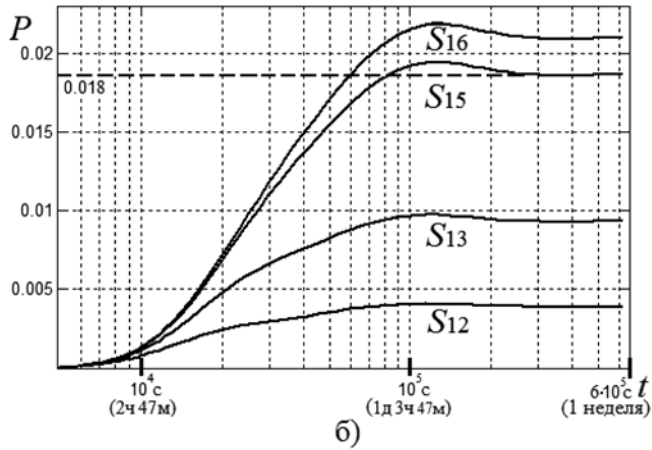
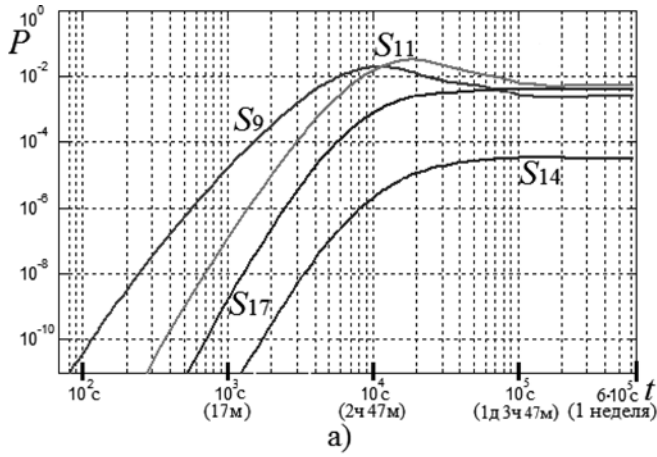


Рис. 9. Динамика изменения вероятностей нахождения информационного конфликта специальных программных средств и подсистемы защиты информации в состояниях $S_9, S_{11}-S_{17}$

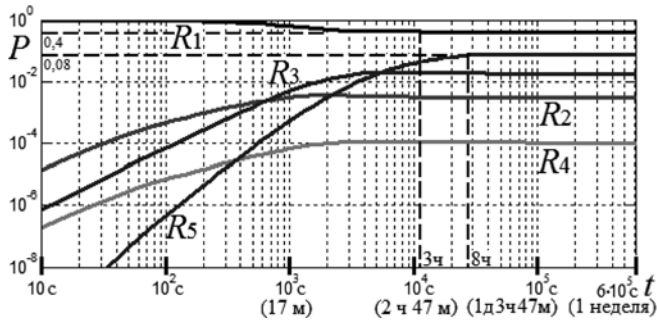


Рис. 10. Динамика состояний информационно-расчетной задачи

о необходимости принятия дополнительных мер по повышению его живучести.

Определение графической зависимости для показателя каждой сформированной группы состояний частных процессов функционирования ИТС продемонстрируем на рис. 12.

Из рис. 12 видно, что в заданных условиях эффективность ИТС в течение 2 дней резко снижается на два порядка, а к концу анализируемого периода возрастет на порядок. Данный "провал" обусловлен тем, что до внедрения СПС в ИТС злоумышлен-

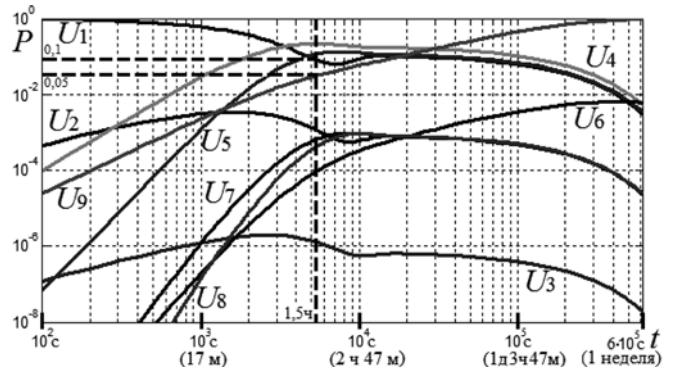


Рис. 11. Динамика состояний задачи непосредственного управления исполнительным устройством

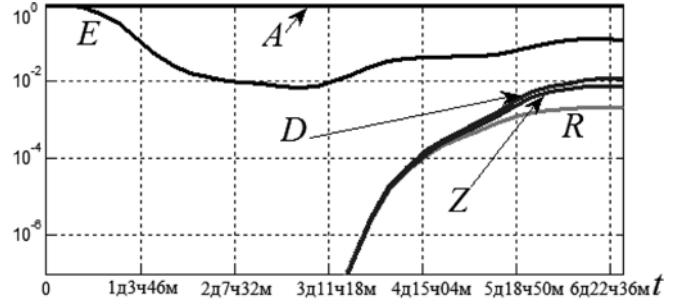


Рис. 12. Динамика показателей функционирования информационно-технического средства

ником активно применяются меры по выводу ИТС из строя, а после внедрения приоритет отдается скрытым ИТВ, которые с точки зрения эффективности организационно-технической системы наиболее целесообразны.

Заключение

Таким образом, предложен способ стратифицированного аналитического описания процесса функционирования информационно-технических средств, позволяющий при учете всего спектра выполняемых этими средствами задач, качества входящих в их состав компонентов и комплексного влияния деструктивных факторов парировать "взрыв" пространства состояний путем введения частных относительно самостоятельных полумарковских процессов и координирующего их суперпроцесса. С применением способа получены аналитические выражения для оценки показателей эффективности информационно-технических средств, их способности к заражению аналогичных средств, подверженности дезинформации, способности быть разведанными и принадлежать своей организационно-технической системе.

Результаты работы могут быть применимы в ходе проектирования и эксплуатации информационно-технических средств в интересах обоснования требований к параметрам их функционирования.

Список литературы

1. ГОСТ Р 50397—92. Совместимость технических средств электромагнитная. Термины и определения. М.: Изд-во стандартов, 1992. 15 с.
2. **Бойко А. А., Храмов В. Ю.** Модель информационного конфликта информационно-технических и специальных программных средств в вооруженном противоборстве группировок со статическими характеристиками // Радиотехника. 2013. № 7. С. 5—10.
3. **Бойко А. А., Дьякова А. В.** Способ разработки тестовых удаленных информационно-технических воздействий на пространственно распределенные системы информационно-технических средств // Информационно-управляющие системы. 2014. № 3. С. 84—92.
4. **Резников Б. А.** Системный анализ и методы системотехники. Часть 1. Методология системных исследований. Моделирование сложных систем. М.: Изд-во Министерство обороны СССР, 1990. 522 с.
5. **Модели информационного конфликта средств поиска и обнаружения.** / Под ред. Ю. Л. Козирацкого. М.: Радиотехника, 2013. 232 с.
6. **Владимиров В. И., Владимиров И. В.** Основы оценки конфликтно-устойчивых состояний организационно-технических систем (в информационных конфликтах). Воронеж: ВАИУ, 2008. 231 с.
7. **Колесниченко В. И., Корниенко В. В., Семенов С. А.** Оценка эффективности автоматизированной информационной системы с ограниченными ресурсами // Вопросы радиоэлектроники. Системы отображения информации и управления спецтехникой. 2008. Вып. 1. С. 16—23.
8. **Борисов В. В., Сысков В. В.** Мультиагентное моделирование сложных организационно-технических систем в условиях противоборства // Информационные технологии. 2012. № 4. С. 7—14.
9. **Гнеденко Б. В., Коваленко И. Н.** Введение в теорию массового обслуживания. М.: Наука, 1987. 336 с.
10. **Чикин М. Г.** Метод аналитического описания процессов с дискретным множеством состояний и не показательными распределениями времен переходов // Информационно-измерительные и управляющие системы. 2004. № 5. С. 8—11.
11. **Кузнецов В. И.** Системное проектирование радиосвязи: Методы и обеспечение. Ч. 1. Системотехника. Воронеж: ВНИИС, 1994. 387 с.
12. **Половко А. М., Гуров С. В.** Основы теории надежности. СПб.: БХВ-Петербург, 2006. 704 с.
13. **Клейнрок Л.** Вычислительные системы с очередями. М.: Мир, 1979. 600 с.
14. **Чикин М. Г.** Особенности использования аппарата полумарковских процессов для моделирования направлений радиосвязи в интересах оценки эффективности радиоподавления // Радиотехника. 2005. № 9. С. 35—39.

A. A. Boyko, PhD, Tech., Associate Professor, algeminy@mail.ru,
Military Education-Science Center of Military Air Forces "Professor N. E. Zhukovsky
and Yu. A. Gagarin Military Air Academy", Voronezh, Russian Federation

Method of Stratified Analytical Description of the Process of Functioning of Information-Technical Tools

The method of stratified analytical description of the process of functioning of information-technical tools, allows in view all spectrum of these tools tasks, quality they components and complex influence of destructive factors parry "explosion" of the space of states through the introduction of private relatively independent semi-Markov processes and coordinating their superprocess. Method allows to get an analytical expressions to evaluate efficiency of functioning of information-technical tool, its ability to infect other similar tools, exposure misinformation, ability to be explored and to belong to their organizational-technical system. The results can be applied during the process of design and operation of information-technical tools in the interests of substantiation requirements of their functioning parameters.

Keywords: stratification, information-technical tool, semi-Markov process

References

1. GOST R 50397—92. *Sovmestimost' tehniceskikh sredstv jelektromagnitnaja*. Terminy i opredelenija. M.: Izd-vo standartov, 1992. 15 p.
2. **Boyko A. A., Hramov V. Ju.** Model informacionnogo konflikta informacionno-tehniceskikh i special'nyh programnyh sredstv v vooruzhennom protivoborstve gruppirovok so statichnymi harakteristikami. *Radiotehnika*. 2013. N. 7. P. 5—10.
3. **Boyko A. A., Djakova A. V.** Sposob razrabotki testovyh udalennyh informacionno-tehniceskikh vozdeistvij na prostranstvenno raspredelennye sistemy informacionno-tehniceskikh sredstv. *Informatsionno-upravliaiushchie sistemy*. 2014. N. 3. P. 84—92.
4. **Reznikov B. A.** *Sistemnyj analiz i metody sistemotehniki. Chast' 1. Metodologija sistemnyh issledovanij. Modelirovanie slozhnyh sistem*. M.: Izdatel'stvo Ministerstvo oborony SSSR, 1990. 522 p.
5. **Modeli informacionnogo konflikta sredstv poiska i obnaruzhenija**. [red. Ju. L. Kozirackogo]. M.: Radiotehnika, 2013. 232 p.
6. **Vladimirov V. I., Vladimirov I. V.** *Osnovy ocenki konfliktno-ustojchivyh sostojanij organizacionno-tehniceskikh sistem (v informacionnyh konfliktah)*. Voronezh: VAIU, 2008. 231 p.
7. **Kolesnichenko V. I., Kornienko V. V., Semenov S. A.** Ocenka jeffektivnosti avtomatizirovannoj informacionnoj sistemy s ograni-
- chennymi resursami. *Voprosy radiojelektroniki. Sistemy otobrazhenija informacii i upravlenija spectehnikoj*. 2008. V. 1. P. 16—23.
8. **Borisov V. V., Syskov V. V.** Mul'tiagentnoe modelirovanie slozhnyh organizacionno-tehniceskikh sistem v uslovijah protivoborstva. *Informacionnyye tehnologii*. 2012. N. 4. P. 7—14.
9. **Gnedenko B. V., Kovalenko I. N.** *Vvedenie v teoriju massovogo obsluzhivaniya*. M.: Nauka, 1987. 336 p.
10. **Chikin M. G.** Metod analiticheskogo opisaniya processov s diskretnym mnozhestvom sostojanij i ne pokazatel'nymi raspredelenijami vremen perehodov. *Informacionno-izmeriternye i upravljajushhie sistemy*. 2004. N. 5. P. 8—11.
11. **Kuznecov V. I.** *Sistemnoe proektirovanie radiosvjazi: Metody i obespechenie*. Ch. 1. Sistematika. Voronezh: VNIIS, 1994. 387 p.
12. **Polovko A. M., Gurov S. V.** *Osnovy teorii nadezhnosti*. SPb.: BHV-Peterburg, 2006. 704 p.
13. **Klejnrok L.** *Vychislitel'nye sistemy s ocheredjami*. M.: Mir, 1979. 600 p.
14. **Chikin M. G.** Osobennosti ispol'zovaniya apparata polumarkovskih processov dlja modelirovaniya napravlenij radiosvjazi v interesah ocenki jeffektivnosti radiopodavlenija. *Radiotehnika*. 2005. N. 9. P. 35—39.

В. И. Струченков, д-р техн. наук, проф., e-mail: str1942@mail.ru,
МГТУ МИРЭА

Комбинированные методы решения вариационных задач с заданной структурой экстремали и ограничениями

Рассматривается задача поиска экстремали заданного функционала (двух- или трехмерной кривой), которая должна удовлетворять целому ряду ограничений. Специфическая особенность искомой кривой заключается в том, что она должна состоять из элементов заданного вида, на параметры которых накладываются ограничения. Число элементов не известно и предстоит определению в процессе решения задачи. Такие задачи возникают, в частности, при проектировании трасс линейных сооружений. Предлагается многоэтапный подход с использованием методов нелинейного и динамического программирования.

Ключевые слова: функционал, целевая функция, нелинейное программирование, приведенный антиградиент, динамическое программирование

Введение

Многие прикладные задачи оптимизации из различных областей практики сводятся к поиску двух- или трехмерных кривых, которые должны удовлетворять условиям гладкости и другим ограничениям. При этом некоторый интегральный показатель, однозначно определяемый для каждой допустимой кривой, должен принимать минимальное значение. Другими словами, требуется найти экстремаль заданного функционала, удовлетворяющую системе ограничений.

Примером таких задач могут служить поиск оптимальных трасс (плана и продольного профиля) проектируемых дорог и других линейных сооружений как на этапе нового строительства, так и на этапе реконструкции. Функционал отражает строительные затраты или приведенные затраты на строительство и эксплуатацию сооружения.

Принципиальная особенность таких задач, заключается в том, что искомая кривая (экстремаль) должна состоять из элементов заданного вида. В зависимости от вида сооружения в качестве элементов могут использоваться отрезки прямых, дуги окружностей, парабол, клотоид и др.

Параметры элементов должны удовлетворять ограничениям на первые производные и кривизну. Длины элементов также ограничены. Число элементов искомой кривой неизвестно.

Целесообразность оптимизации проектов таких дорогостоящих объектов, как железные и автомобильные дороги, очевидна. В частности, в условиях пересеченного рельефа и сложной геологии затраты на строительство и последующую эксплуатацию могут быть существенно снижены при оптимальном расположении трассы проектируемой дороги на местности. Это было установлено еще 40 лет тому назад при проектировании отдельных участков БАМ на таких ЭВМ, как БЭСМ 4 [1].

Отмеченные особенности задачи, затрудняющие разработку алгоритмов и программ вместо утраченных в 90-х годах, а также отсутствие заинтересованности проектировщиков и строителей в снижении строительных затрат привели к тому, что математические методы оптимизации трасс практически не применяются. Считается достаточным использование разного рода эвристических алгоритмов в процессе интерактивного проектирования.

В настоящее время даже в наиболее совершенных современных системах автоматизированного проектирования (САПР), таких как CARD/1 [2], Bentley Rail Track [3], или их российские аналоги Topomaic Robur [4] и Geonics [5], компьютер используется для решения сопутствующих рутинных задач, но не как инструмент выработки оптимальных проектных решений.

Известно, что в одних и тех же условиях, располагая одной и той же информацией, различные специалисты предлагают различные варианты проектных решений. Рассмотрение ограниченного числа интуитивно назначаемых вариантов не гарантирует близость к оптимуму конечного результата такого процесса. В то же время известно, что относительно небольшие изменения положения трассы на местности могут приводить к существенным изменениям затрат на строительство и эксплуатацию дороги [6].

Следовательно, проблема разработки адекватных математических моделей и математически корректных алгоритмов оптимизации трасс новых дорог остается актуальной. Это является главным направлением совершенствования САПР линейных сооружений.

Цель настоящей статьи — показать, как исходная вариационная задача, не решаемая классическими методами, может быть решена с помощью комбинации методов нелинейного и динамического программирования.

В этой статье дана общая постановка исходной вариационной задачи и рассмотрены наиболее типичные частные случаи. Далее представлены математические модели. Случай поиска двумерной кривой рассмотрен подробно, начиная с определения числа элементов и кончая оптимизацией их параметров. При этом реализация методов нелинейного и динамического программирования в конкретных алгоритмах будет основываться на специфических особенностях системы ограничений.

1. Постановка задачи и математические модели

Найти трехмерную кривую $x(s)$, $y(s)$, $z(s)$, для которой достигается

$$\min \int_0^S F(x(s), y(s), z(s)) ds. \quad (1)$$

Здесь s — текущая длина кривой, а S — неизвестная суммарная длина кривой, соединяющей две заданные точки A и B . $F(x(s), y(s), z(s))$ — заданная функция, имеющая непрерывные первые частные производные по всем трем своим аргументам. В подынтегральной функции F в качестве аргументов могут присутствовать частные производные функций $x(s)$, $y(s)$, $z(s)$, на которые накладываются ограничения. Ограничения в виде неравенств могут накладываться и на координаты отдельных точек искомой кривой.

Применительно к проектированию оптимальных трасс линейных сооружений $x(s)$, $y(s)$ задают план трассы, а $z(s)$ — ее продольный профиль. При этом $z(s)$ — однозначная функция и $\left| \frac{dz}{ds} \right| \leq v$, где v (максимальный продольный уклон) имеет значение в несколько промилле, и потому отличие длины искомой кривой от длины ее проекции на горизонтальную плоскость практически несущественно. Ограничение на уклон может быть и двусторонним.

Ограничена не только первая производная в продольном профиле, но и кривизна в плане и продольном профиле, а также координаты отдельных точек (например, при пересечении водотоков и существующих коммуникаций).

Ограничена не только первая производная в продольном профиле, но и кривизна в плане и продольном профиле, а также координаты отдельных точек (например, при пересечении водотоков и существующих коммуникаций).

Рассматриваемая задача имеет существенные отличия от задач, рассматриваемых в вариационном исчислении [7]. Прежде всего это наличие ограничений в виде неравенств. Даже в случае отсутствия дополнительных условий, в соответствии с которыми искомая кривая должна состоять из элементов заданного вида, отмеченные особенности не позволяют использовать аппарат классического вариационного исчисления, в частности уравнения Эйлера [7], для решения рассматриваемой задачи.

Если в (1) функции $x(s)$, $y(s)$ заданы, то получаем задачу поиска плоской кривой $z(s)$, при этом $z(s)$ — однозначная функция.

Другая частная задача возникает при поиске кривой на плоскости XOY , т. е. при поиске функций $x(s), y(s)$. В этом случае искомая кривая не обязательно является графиком однозначной функции.

Применительно к проектированию трасс линейных сооружений поиск $z(s)$ соответствует проектированию оптимального продольного профиля при заданном положении трассы в плане. Эта задача может иметь самостоятельное значение при проектировании трасс в обжитых районах из-за отсутствия возможности варьирования в плане. В качестве примера достаточно вспомнить историю защиты Химкинского леса.

Поиск плоской кривой $x(s)$, $y(s)$ соответствует проектированию плана трассы в условиях, когда продольный профиль однозначно определяется этим планом или слабо зависит от него. Примером может служить прокладка траншей на заданной глубине, проектирование трасс новых железных дорог на участках напряженного хода [8] или проектирование плана трассы реконструируемой железной дороги (выправка пути) [8].

Рассмотрим подробнее двумерные задачи применительно к проектированию трасс железных дорог.

В случае проектирования продольного профиля искомая проектная линия $z(s)$ — это ломаная линия, соединяющая начальную и конечную точки, число элементов которой неизвестно.

Если обозначить профиль земли $H(s)$, то в первом приближении задача состоит в следующем. По заданной $H(s)$ найти такую ломаную $z(s)$, чтобы она удовлетворяла всем ограничениям, и был

$$\min \int_0^{S_0} F(z(s), H(s), s) ds, \quad (2)$$

где S_0 — заданная длина трассы в плане, а функция F моделирует затраты на элементе длины.

Реальные модели должны учитывать конструкции поперечных профилей земляного полотна, наличие водопропускных и других искусственных сооружений, распределение земляных масс и способы производства земляных работ и др. Эти модели подробно рассмотрены в работах [8–10].

Задача вариационного исчисления (2) сводится к задаче нелинейного программирования, обладающей интересными особенностями независимо от конкретного вида функции F .

Поскольку число элементов искомой ломаной неизвестно, то приходится считать, что переломы профиля земли и проектной линии (т. е. профиля трассы) имеют одни и те же абсциссы. Профиль земли всегда представлен в виде ломаной с неравномерным шагом, и такое допущение позволяет фиксировать число элементов n (размерность задачи) и длины s_i элементов (в плане). При этом получится ломаная с большим, чем нужно, числом элементов, но из-за многочисленных ограничений ее отклонения от окончательной $z(s)$ невелики [8]. Идея в

том, чтобы найти эту ломаную путем решения задачи оптимизации, затем преобразовать ее в ломаную с элементами, длины которых не менее допустимой длины, определив тем самым реальную размерность задачи и начальное приближение, и на последнем этапе выполнить оптимизацию при всех ограничениях и необходимых уточнениях целевой функции. Такой многоэтапный процесс с уточнением математической модели и ее параметров является обычным для решения сложных проектных задач творческого характера.

Зная число и длины элементов искомого ломаной, можно аналитически выразить все ограничения на $z(s)$, если принять в качестве неизвестных z_i ($i = 1, 2, \dots, n$) ее ординаты в точках перелома. Эти ограничения делятся на три группы:

- на ординаты в отдельных точках $z_i \leq z_i^{\max}$ или $z_i \geq z_i^{\min}$;

- на уклоны элементов профиля:

$$a_i \leq (z_{i+1} - z_i)/s_i \leq b_i \quad (i = 1, 2, \dots, n - 1),$$

где s_i — длины элементов. Эти ограничения являются дискретными аналогами ограничений на первую производную; b_i, c_i, d_i — заданные константы;

- на разности уклонов смежных элементов:

$$c_i \leq (z_{i+2} - z_{i+1})/s_{i+1} - (z_{i+1} - z_i)/s_i \leq d_i \quad (i = 1, 2, \dots, n - 1).$$

Эти ограничения являются дискретными аналогами ограничений на кривизну.

В силу малости проектных уклонов длина элемента и его проекции практически совпадают.

Интеграл (2) превращается в сумму [9, 10] и получается следующая задача нелинейного программирования с линейной системой ограничений.

Найти $\min \Phi(\mathbf{x}, \mathbf{c})$ при $\mathbf{Ax} \leq \mathbf{b}$, где \mathbf{x} — вектор неизвестных, \mathbf{c} — вектор параметров, матрица \mathbf{A} и вектор \mathbf{b} задают систему линейных ограничений, $\Phi(\mathbf{x}, \mathbf{c})$ — целевая функция.

В случае проектирования плана трассы различных линейных сооружений, т. е. при поиске функций $x(s), y(s)$, элементами искомого линии могут быть отрезки прямых, окружностей, парабол, клотоид и др. Число элементов неизвестно. Ограничены минимальные длины элементов, максимальная кривизна и координаты отдельных точек. Искомая кривая может быть приближенно представлена в виде ломаной линии с большим, но известным числом элементов. В качестве переменных могут быть приняты координаты вершин искомого ломаной, через которые аналитически выражаются все ограничения. Снова получается задача нелинейного программирования, но с нелинейной системой ограничений.

При проектировании трасс реконструируемых железных дорог в качестве переменных удобно

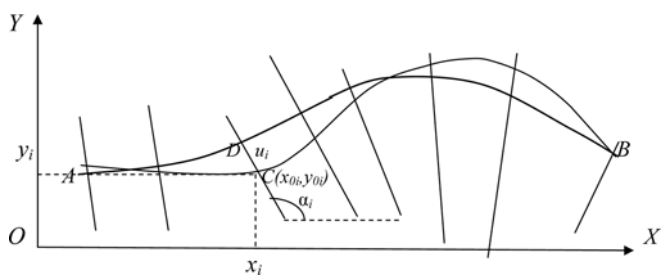


Рис. 1. Дискретное представление трассы

принять не координаты x_i, y_i вершин искомого ломаной, а сдвиги u_i по нормальям (поперечникам) в точках съемки плана трассы существующего пути (CD на рис. 1). Через эти сдвиги также можно выразить все ограничения на искомую кривую. При этом получим задачу нелинейного программирования с меньшим числом переменных.

Имеются следующие виды ограничений:

1. Кривизна окружности, проведенной через любые три смежные точки в плане, должна быть в заданных пределах $k_{\min i} \leq k(u_{i-1}, u_i, u_{i+1}) \leq k_{\max i}$, где

$$k(u_{i-1}, u_i, u_{i+1}) = \frac{\begin{vmatrix} x_{i-1} & y_{i-1} & 1 \\ x_i & y_i & 1 \\ x_{i+1} & y_{i+1} & 1 \end{vmatrix}}{l_{i-1} \cdot l_{i-1, i+1} \cdot l_{i, i+1}}. \quad (3)$$

В числителе (3) стоит определитель третьего порядка, в который входят декартовы координаты трех смежных точек, а в знаменателе — произведение расстояний l между этими точками. Определитель — это удвоенная площадь треугольника с вершинами в рассматриваемых точках. Это известная формула для радиуса описанной окружности $R = abc/(4S)$, где a, b, c — длины сторон, S — площадь треугольника.

Через переменные u_i ($i = 1, 2, \dots, n$) выражаем сначала декартовы координаты точек пересечения плана трассы с поперечниками

$$x_i = x_{0i} + u_i \cos \alpha_i; \quad y_i = y_{0i} + u_i \sin \alpha_i$$

затем через них вычисляем расстояния между точками и кривизну. Здесь x_{0i}, y_{0i} — координаты начальной точки на i -й нормали, α_i — угол нормали с осью X .

2. Ограничения на координаты отдельных точек, включая условия пересечения других коммуникаций и водотоков, теперь записываются просто:

$$c_k \leq u_k \leq d_k,$$

где k — номер нормали, соответствующей коммуникации или водотоку. Таких ограничений может быть много. Для строго фиксированных точек $c_k = d_k$.

Ограничения на план трассы существенно нелинейны.

2. Особенности возникающих задач нелинейного программирования и методы их решения

Задачи нелинейного программирования, получаемые при дискретном представлении искомой экстремали, обладают существенными особенностями, обусловленными спецификой рассматриваемых объектов. Эти особенности относятся прежде всего к системам ограничений на искомую экстремаль.

Так, ни одно из ограничений не связывает больше трех смежных переменных. Это означает, что в линейном случае (проектирование продольного профиля) матрица системы ограничений имеет блочный характер. Блоки, соответствующие отмеченным выше типам ограничений, содержат в каждой строке только один, два или три ненулевых элемента. Это позволяет построить итерационный процесс, на каждом шаге которого при любом наборе активных ограничений направление спуска определяется без решения систем линейных уравнений. В качестве такого направления принимается приведенный антиградиент [11, 12].

Если A_k — матрица системы активных ограничений на k -й итерации и $N(A_k)$ — ее нуль-пространство [12], то отмеченные свойства системы ограничений для любого активного набора позволяют легко построить базис в $N(A_k)$ [13]. Если C — базисная матрица и f — антиградиент целевой функции, то $CC^T f$, называемый приведенным антиградиентом, является направлением спуска. Действительно, его скалярное произведение на антиградиент неотрицательно, так как $(CC^T f, f) = (C^T f, C^T f) \geq 0$.

Система нелинейных ограничений, получаемая при дискретном представлении плана трассы, также в каждом из неравенств содержит не более трех смежных переменных. При ее линеаризации в текущей итерационной точке получается система линейных неравенств, обладающая отмеченными особенностями. Это позволяет построить базис в нуль-пространстве, соответствующем касательной плоскости. Но спуск из текущей итерационной точки целесообразно осуществлять не в касательной плоскости, а по кривой в граничной поверхности без нарушения активных нелинейных ограничений. Касательная к этой кривой совпадает с на-

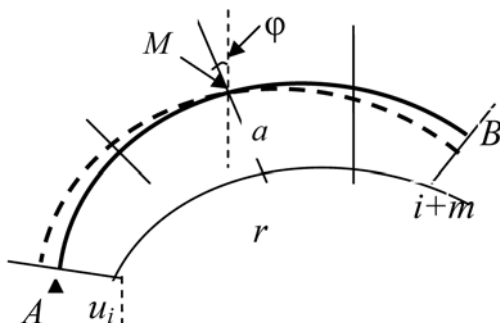


Рис. 2. Пример базисного смещения (сдвиг + поворот)

правлением спуска в касательной плоскости. В этом состоит смысл метода базисных смещений [14].

Пусть на некоторой итерации точки с i -й по $(i + m)$ -ю лежат на кривой предельного радиуса (рис. 2). Ее начальное положение AB . Необходимо найти такие изменения переменных $u_i, u_{i+1}, \dots, u_{i+m}$, при которых радиус остается предельным. Это означает, что окружность смещается как единое целое, и любое ее смещение можно представить как комбинацию двух смещений, например, сдвиг вдоль некоторой прямой и поворот в плоскости вокруг некоторого центра. В качестве прямой можно взять любую нормаль с номером $i \leq k \leq i + m$, а за центр вращения принять точку пересечения плана трассы и выбранной нормали. На рис. 2 это точка M на поперечнике с номером r . Пусть a — смещение окружности вдоль поперечника, а φ — угол поворота. Если величины a и φ известны, то приращения Δu_k ($i \leq k \leq i + m$) вычисляются через a, φ, u_i углы нормалей с осью OX и исходные значения u_k ($i \leq k \leq i + m$).

Если на окружности оказалась одна точка, у которой координата u_j имеет минимальное или максимальное значение, то ее следует принять за центр вращения, а соответствующий j -й поперечник — за линию сдвига. В этом случае $a = 0$, и все искомые смещения по поперечникам определяются углом поворота.

Если на окружности предельного радиуса оказалось две или более фиксированных точек, то искомые смещения или приращения координат Δu_k ($i \leq k \leq i + m$) равны нулю.

Набор линейно независимых базисных смещений является аналогом базиса в линейном случае. При линеаризации базисных смещений получается базис в соответствующем ему нуль-пространстве.

Важно отметить, что отмеченные особенности систем ограничений как в линейном, так и в нелинейном случае позволяют без громоздких вычислений осуществлять модификацию активного набора [13].

3. Определение числа элементов экстремали и построение начального приближения

Решение задачи нелинейного программирования дает экстремаль в виде ломаной линии. Ее число элементов определяется спецификой оптимизируемого объекта и требуемой точностью решения задачи. Однако это число существенно больше, чем требуется, если изначально были ограничения на минимальную длину элемента. Поэтому необходимо преобразовать полученную ломаную с минимальными отклонениями от нее в последовательность элементов заданного вида с соблюдением всех ограничений, включая ограничения по минимальной длине элемента. Поскольку экстремаль в виде ломаной удовлетворяет всем ограничениям (кроме ограничений по минимальной длине элемента), то ее отклонения от искомой последовательности эле-

ментов оказываются много меньше, чем ширина исходной области поиска оптимального решения [14]. Это позволяет использовать разработанные ранее алгоритмы динамического программирования для решения возникающих задач аппроксимации полученной ломаной плоской кривой, состоящей из элементов заданного вида при наличии ограничений, рассмотренных выше [15, 16].

Характерно, что ограничения в данном случае способствуют сокращению времени счета из-за сокращения числа допустимых вариантов.

При проектировании продольного профиля железных дорог ломаная линия, состоящая из коротких элементов, преобразуется в ломаную же линию, но с элементами, длина которых не менее заданной величины. Алгоритм преобразования подробно рассмотрен в работах [8, 15].

При проектировании продольного профиля автомобильных дорог преобразование может осуществляться в систему парабол второй степени (в частности, отрезков прямых) или в систему круговых кривых с или без прямых вставок. Соответствующие алгоритмы динамического программирования приведены в работах [8, 16, 17].

При проектировании плана трассы преобразование ломаной в последовательность круговых кривых с прямыми вставками особых сложностей не вызывает. Особенность только в том, что искомая экстремаль в общем случае не является графиком однозначной функции. Поэтому для использования динамического программирования сетка варьирования строится с использованием нормалей [14].

Преобразование ломаной в последовательность круговых и прямых, сопрягаемых клотоидами, потребовало разработки отдельного алгоритма, так как с помощью динамического программирования эту задачу решить не удалось.

4. Оптимизация начального приближения

Экстремаль, получаемая после преобразования ломаной линии, должна рассматриваться как начальное приближение для получения окончательного решения.

Смысл этапов расчета ломаной линии и ее преобразования состоит в определении размерности задачи (числа элементов) и построении начального приближения. Дело не только в том, что динамическое программирование не обеспечивает точности решения задачи из-за дискретности варьирования, но в таких задачах, как проектирование продольного профиля новых дорог, имеется дополнительная взаимосвязь элементов в насыпях и выемках, сооружаемых совместно. Целевая функция в этом случае не может вычисляться для каждого элемента отдельно, что не позволяет использовать динамическое программирование [14].

Если искомая экстремаль состоит только из прямолинейных элементов, как это имеет место при

проектировании продольного профиля железных дорог, то оптимизация начального приближения осуществляется с помощью того же алгоритма нелинейного программирования, что и на первом этапе при построении ломаной. В этом случае ордината любой точки искомой экстремали линейно зависит от ординат концов соответствующего элемента, поэтому добавлен только пересчет производных целевой функции по ординатам концов коротких элементов в производные по ординатам концов элементов искомой экстремали.

При использовании параболических элементов удобно взять в качестве неизвестных не ординаты, а уклоны концов элементов.

Рассмотрим случай фиксированных уклонов в начальной и конечной точках.

Уравнение каждой параболы в декартовой системе координат (l, H) , помещенной в начало элемента, имеет вид

$$H = al^2 + bl + c,$$

где c — ордината в начальной точке элемента; b — уклон в начале элемента; l — расстояние от начала элемента.

Если считать неизвестными все уклоны — параметры b_j ($j = 1, 2, \dots, n + 1$), где n — число элементов, то через эти уклоны можно найти ординату любой точки, считая заданной ординату начальной точки. Действительно, параметр a_j каждой параболы вычисляется через уклоны в ее начальной и конечной точках, так как длина элемента L_j известна:

$$a_j = (b_{j+1} - b_j)/(2L_j); \quad (4)$$

$$c_{j+1} = (b_{j+1} + b_j)L_j/2 + c_j. \quad (5)$$

Это означает, что при любой комбинации параметров b_j можно вычислить все ординаты, которые раньше считались независимыми переменными, а следовательно, и частные производные (градиент) целевой функции. Для этого нужно последовательно вычислять параметры a_j и c_{j+1} по формулам (4) и (5).

Ограничения по уклонам записываются совсем просто:

$$b_j^{\min} \leq b_j \leq b_j^{\max}.$$

Ограничения по радиусам:

$$L_j/R_j^{\min1} \leq b_{j+1} - b_j \leq L_j/R_j^{\min2}.$$

Здесь $R_j^{\min1} < 0$ и $R_j^{\min2} > 0$ — минимально допустимые радиусы выпуклой и вогнутой кривых соответственно.

Система ограничений чрезвычайно проста, поэтому при любом наборе активных ограничений по уклонам и радиусам для соответствующего участка сдвиг в пространстве уклонов невозможен при наличии хотя бы одного предельного уклона (нулевой

участок) или если активны только ограничения по радиусам, то все компоненты единственного базисного вектора равны единице (участок сдвига). Отметим, что в исходном координатном пространстве нулевой участок соответствует сдвигу по оси ординат, а участок сдвига соответствует повороту с центром в начальной точке участка.

Однако есть еще и ограничения на ординаты отдельных точек, в частности, всегда активное ограничение по выходу в заданную конечную точку. Это обстоятельство не позволяет строить направление спуска отдельно по участкам, если уклоны элементов между ними и кривизна не предельны.

Рассматривалось два способа преодоления возникающих трудностей [14]:

- использование штрафных функций за нарушение ограничений по ординатам;
- построение направления спуска с помощью решения вспомогательных систем линейных уравнений, число которых равно числу активных ограничений по ординатам.

Однако в дальнейшем удалось упростить алгоритм с помощью специального приема преобразования базиса. Для этого берется канонический базис в пространстве уклонов для всей искомого экстремали в целом. Ему соответствует единичная матрица E .

Если с k -го по $k + m$ -й уклон имеем нулевой участок, то столбцы с k -го по $k + m$ -й из матрицы вычеркиваются. При этом в матрице образуются нулевые строки.

Если с r -го по $r + p$ -й уклон имеем участок сдвига, то соответствующие столбцы заменяются их суммой.

Обозначим полученные вектор-столбцы e_1, e_2, \dots, e_q .

Активное ограничение по ординате некоторой точки требует, чтобы направление спуска p_1, p_2, \dots, p_n удовлетворяло дополнительному условию

$$\alpha_1 p_1 + \alpha_2 p_2 + \dots + \alpha_k p_k = 0. \quad (6)$$

Здесь α_i ($i = 1, 2, \dots, k - 1$) — длины элементов, предшествующих элементу с номером k , на котором есть точка с ограниченной ординатой; α_k — часть длины k -го элемента от его начала до точки с ограниченной ординатой. Далее будем рассматривать вектор α , у которого первые k компонент равны α_i ($i = 1, 2, \dots, k$) и $\alpha_i = 0$ при $i > k$, если $k < n$.

Условие (6) уменьшает размерность нуль-пространства и требует преобразования базиса. Это можно сделать различными способами. Например, исключим из базиса первый из векторов e_j , для которого $(\alpha, e_j) \neq 0$. Остальные векторы преобразуем по формуле

$$e'_k = e_k - \frac{(\alpha, e_k)}{(\alpha, e_j)} e_j. \quad (7)$$

Линейная независимость преобразованных базисных векторов гарантирована, так как любая их линейная комбинация является линейной комбинацией исходных базисных векторов. Очевидна и ортогональность любого из них вектору α .

Формулу (7) можно использовать повторно при наличии любых других линейных ограничений.

Более того, эта формула позволяет построить базис в нуль-пространстве любой матрицы с линейно независимыми строками, начиная с канонического базиса.

При оптимизации начального приближения экстремали, не соответствующей однозначной функции, например, кривой, состоящей из отрезков прямых и круговых кривых или с дополнительным сопряжением прямых и круговых клотоидами, в качестве переменных принимались длины элементов и кривизна круговых кривых.

Задача была решена с помощью метода DFP-оптимизации [11].

Этот вопрос требует отдельного рассмотрения.

Заключение

Рассмотренные задачи из области проектирования трасс линейных сооружений — это частный случай решения вариационной задачи, которой присущи отмеченные выше особенности. Подобные задачи возникают и в других областях, например, в планировании и при управлении технологическими процессами.

Идея комбинированного применения методов нелинейного и динамического программирования может оказаться полезной и при решении вариационных задач, в которых подынтегральная функция содержит не только искомую экстремаль, но и ее производные.

К настоящему времени комбинированные методы используются в разработанных автором САПР, предназначенных для проектирования трасс новых и реконструируемых железных дорог.

В этих интерактивных системах в диалоге проектировщик — компьютер последний выступает как интеллектуальный партнер, способный вырабатывать оптимальные решения с использованием математических моделей и математически корректных алгоритмов оптимизации. В этом принципиальное отличие новых систем от систем, в которых проектировщик подбирает элементы проектного решения, "ползая по экрану компьютера", а компьютер увязывает их, реализуя творческий замысел проектировщика. Миф о том, что подобным образом достигаются оптимальные решения весьма устойчив, но, будем надеяться, не вечен.

Список литературы

1. Использование математических методов оптимизации и ЭВМ при проектировании продольного профиля железных до-

рог // Тр. Всесоюзного НИИ транспортного строительства. Вып. 101. М.: Транспорт, 1977.

2. **CARD/1**. URL: <http://www.card-1.com/en/home/>
3. **Bentley Rail Track**. URL: <http://www.bentley.com/>
4. **Topomatic Robur**. URL: <http://www.topomatic.ru> Visited: July 14, 2012.
5. **Курилко Ю., Чешева В.** Geonics ЖЕЛДОР — САПР // CADmaster 2007. № 1(36).
6. **Shafahi Y., Shahbazi M. J.** Optimum railway alignment. URL: http://www.uic.org/cdrom/2001/wcrr2001/pdf/sp/2_1_1/210.pdf
7. **Ожегова А. В., Насибуллин Р. Г.** Вариационное исчисление: задачи, алгоритмы, примеры. Казань: Казанский ун-т, 2013.
8. **Струченков В. И.** Методы оптимизации трасс в САПР линейных сооружений. М.: Солон — Пресс, 2014.
9. **Струченков В. И.** Математические модели и методы оптимизации в системах проектирования трасс новых железных дорог // Информационные технологии. 2013. № 7 (203).
10. **Struchenkov V.** Mathematical Models and Optimization in Line Structure Routing: Survey and Advanced Results // International Journal of Communications, Network and System Sciences. Special Issue Models & Algorithms for Applications. 2012. N. 5.

11. **Гилл Ф., Мюррей У., Райт М.** Практическая оптимизация: Пер. с англ. М.: Мир, 1985.

12. **Аоки М.** Введение в методы оптимизации: Пер. с англ. М.: Наука, 1977.

13. **Struchenkov V. I.** Nonlinear Programming Algorithms for CAD Systems of Line Structure Routing // World Journal of Computer Application and Technology. 2014. N. 2 (5).

14. **Струченков В. И.** Методы оптимизации в прикладных задачах. М.: Солон — Пресс, 2009.

15. **Struchenkov V. I.** Piecewise Linear Approximation of Plane Curves with Restrictions in Computer-Aided Design of Railway Routes // World Journal of Computer Application and Technology. 2014. N. 2 (1).

16. **Struchenkov V. I.** Per Element Approximation of Plane Curves with Restrictions in Computer-Aided Design of Road Routes // American Journal of Systems and Software, 2013. V. 1, N. 1.

17. **Struchenkov V. I.** Piecewise Parabolic Approximation of Plane Curves with Restrictions in Computer-Aided Design of Road Routes // Transaction on Machine Learning and Artificial Intelligence. Society Science and Education. United Kingdom. 2013. V. 1, Is. 1.

V. I. Struchenkov, Professor, e-mail: stl942@mail.ru,
Moscow State University of Radio Engineering, Electronics and Automation

Combined Algorithms for Solving Variational Problems with a Given Structure of the Extremal and Constraints

Under study is the problem of searching the extremal of given functional (two or three-dimensional curve), which must satisfy a number of given restrictions. A specific feature of the desired curve is that it should consist of a given types elements, whose parameters are limited. The number of elements is unknown and to be determined in the process of solving the problem. Such problems arise in particular when designing linear structures routes.

In the case searching of two-dimensional extremal piecewise linear and piecewise parabolic curves are considered. Such problems arise in the design of optimal longitudinal profile of railways and roads.

Multi-stage approach is proposed using the methods of nonlinear and dynamic programming.

In the first stage using nonlinear programming we define polyline consisting of elements of small length. In the second stage using dynamic programming we determine a number of the elements and the initial approach for the third stage. In the third stage using a special non-linear programming algorithm we find the optimal decision.

Keywords: functional, extremal, objective function, nonlinear programming, reduced antigradient, dynamic programming

References

1. **Ispol'zovanie** matematicheskikh metodov optimizacii i JeVm pri proektirovanii prodol'nogo profilya zheleznyh dorog. Tr. Vsesojuznogo NIИ transportnogo stroitel'stva. Vyp. 101. М.: Transport, 1977.
2. **CARD/1**. URL: <http://www.card-1.com/en/home/>
3. **Bentley Rail Track**. URL: <http://www.bentley.com/>
4. **Topomatic Robur**. URL: <http://www.topomatic.ru> Visited: July 14, 2012.
5. **Kurilko Ju., Chesheva V.** Geonics ZhELDOR — SAPR. CAD-master 2007. N 1 (36).
6. **Shafahi Y., Shahbazi M. J.** Optimum railway alignment. URL: http://www.uic.org/cdrom/2001/wcrr2001/pdf/sp/2_1_1/210.pdf
7. **Ozhegova A. V., Nasibullin R. G.** Variacionnoe ischislenie: zadachi, algoritmy, primery. Kazan': Kazanskiy universitet, 2013.
8. **Struchenkov V. I.** Metody optimizacii trass v SAPR linejnyh sooruzhenij. М.: Solon — Press, 2014.
9. **Struchenkov V. I.** Matematicheskie modeli i metody optimizacii v sistemah proektirovanija trass novyh zheleznyh dorog. Informacionnye tehnologii. 2013. N. 7 (203).
10. **Struchenkov V. I.** Mathematical Models and Optimization in Line Structure Routing: Survey and Advanced Results. International

Journal of Communications, Network and System Sciences. Special Issue Models & Algorithms for Applications. 2012. N. 5.

11. **Gill F., Mjurrej U., Rajt M.** Prakticheskaja optimizacija: Per. s angl. М.: Mir, 1985.

12. **Aoki M.** Vvedenie v metody optimizacii: Per. s angl. М.: Nauka, 1977.

13. **Struchenkov V. I.** Nonlinear Programming Algorithms for CAD Systems of Line Structure Routing. World Journal of Computer Application & Technology. 2014. V. 2, N. 5.

14. **Struchenkov V. I.** Metody optimizacii v prikladnyh zadachah. М.: Solon — Press, 2009.

15. **Struchenkov V. I.** Piecewise Linear Approximation of Plane Curves with Restrictions in Computer-Aided Design of Railway Routes. World Journal of Computer Application and Technology. 2014. V. 2, N. 1.

16. **Struchenkov V. I.** Per Element Approximation of Plane Curves with Restrictions in Computer-Aided Design of Road Routes. American Journal of Systems and Software. 2013. V. 1, N. 1.

17. **Struchenkov V.** Piecewise Parabolic Approximation of Plane Curves with Restrictions in Computer-Aided Design of Road Routes. Transaction on Machine Learning and Artificial Intelligence. Society Science and Education. United Kingdom. 2013. V. 1, N. 1.

МЕТОДЫ И ТЕХНОЛОГИИ ОБРАБОТКИ ДАННЫХ

METHODS AND TECHNOLOGIES OF DATA PROCESSING

УДК 004.91

И. С. Печенко, науч. сотр., e-mail: ivan.pechenko@intel.com,
О. В. Венгер, ст. науч. сотр., e-mail: oleg.v.venger@intel.con,
ЗАО "Интел А/О"

Способ автоматизированного создания диаграмм последовательности операций для сценариев поведения системы, описанных в виде текста

Одним из способов спецификации поведения вычислительных систем является описание набора сценариев, по которым должно определяться взаимодействие ее компонентов. 95 % такого рода спецификаций составляют документы, в которых сценарии описаны в виде текста. Такому описанию присущи неточности, неоднозначности, неполнота, невозможность машинной обработки. Предлагается способ преобразования текстовой спецификации к виду, пригодному для машинной обработки и более подходящему для обмена информацией между группами участников проекта. В текстовое описание добавляются тэги, позволяющие структурировать спецификацию. Затем, с использованием графической нотации BPMN, из структурированного текста автоматически создается диаграмма последовательности операций. Предлагаемый способ был реализован и протестирован на нескольких моделях сценариев поведения систем. Данный подход может быть применен на практике для сценариев, описывающих поведение сложных вычислительных систем и комплексов.

Ключевые слова: модель поведения, сценарий поведения, структурированный текст, машинная обработка спецификации, диаграмма последовательности операций, BPMN

Введение

Современные вычислительные системы состоят из большого числа различных программных и аппаратных компонентов. Аппаратная часть насчитывает десятки блоков, таких как процессоры, графические ядра, контроллеры памяти, контроллеры устройств ввода-вывода и т. д. Программная часть включает в себя прошивки для аппаратных блоков, драйверы, операционную систему и др. Одной из задач, возникающих при проектировании таких систем, является спецификация и описание того, каким образом компоненты должны взаимодействовать между собой и с окружением вычислительной системы. Качество такой спецификации является определяющим фактором для успеха других этапов проектирования, в частности, эффективности тестирования системы. Кроме того, эффективность процесса проектирования можно повысить путем использования результатов машинной обработки спецификации в дальнейшей разработке и тестировании.

Одним из способов спецификации поведения системы является описание набора сценариев, по которым должно осуществляться взаимодействие ее компонентов. В настоящее время не существует единого стандарта для описания сценариев. Среди используемых методов можно отметить диаграммы коммуникации и последовательности UML (Unified Modeling Language, стандарт графического описания

для объектного моделирования в области разработки программного обеспечения [15]), диаграммы последовательности сообщений, диаграммы последовательности операций и т. д. На практике значительную часть такого рода спецификации составляют документы, в которых сценарии описаны в виде обычного текста. Согласно [1] 95 % технических требований при разработке программного обеспечения представлены в виде обычного либо структурированного текста.

Текстовое описание сценариев является удобным начальным способом описания поведения системы. Это особенно справедливо для сложных сценариев с большим числом событий и взаимодействующих компонентов. Его легко редактировать и дополнять, в том числе вставлять новые блоки информации в произвольное место сценария. Однако описание в виде обычного текста имеет ряд недостатков. Во-первых, свободному текстовому описанию присущи неточности, неоднозначности, неполнота. Во-вторых, форма, удобная для создания спецификации, не всегда является удобной для понимания и использования ее другими участниками проекта. Для понимания больше подходит визуальное представление, например, такое как диаграмма последовательности операций. В-третьих, автоматическая обработка текстовых описаний невозможна без специальной подготовки.

Существующие работы, связанные с анализом и использованием текстовой спецификации, в боль-

шинстве случаев нацелены на улучшение процесса разработки программного обеспечения. Их можно разделить на две категории. Часть работ предлагает способы обнаружения проблем в текстовой спецификации [2, 7], оценки и улучшения качества текстовой спецификации [3–6]. Другая группа работ предлагает способы экстракции UML-диаграмм из текстовой спецификации [8–10, 13]. В обоих случаях исходное текстовое описание сначала структурируется, что обеспечивает возможность последующей машинной обработки. За счет такого структурирования достигается соответствие текста некой формализованной структуре, которую затем нетрудно перести в диаграммы UML. Такое структурирование достигается двумя основными путями: путем лингвистического анализа с последующим устранением выявленных неточностей и дефектов в текстовых спецификациях и путем написания спецификаций по заранее установленному структурированному шаблону. Интересным способом структурирования текстового описания является SBVR (Semantics of Business Vocabulary and Rules) [14], продвигаемый консорциумом OMG (Object Management Group) [12], который занимается разработкой и продвижением объектно-ориентированных технологий и стандартов. Членами консорциума являются сотни ведущих компаний-производителей программного обеспечения.

В данной работе предлагается способ преобразования спецификации сценариев в виде обычного текста к виду, пригодному для машинной обработки и более подходящему для обмена информацией между различными группами участников проекта, ответственными за проектирование, разработку, тестирование программно-аппаратных платформ. Суть предлагаемого решения заключается в следующем: сначала в исходное текстовое описание добавляются специальные тэги, позволяющие структурировать полученную спецификацию. Затем из текста с тэгами автоматически создается графическое представление — диаграмма последовательности операций. В отличие от методов структурирования текста спецификаций, основанных на лингвистическом анализе, описанный в данной статье метод не предполагает привязки к определенной области применения и использования словарей понятий и ключевых фраз, а преимущество описанного метода перед использованием заранее определенных шаблонов текста заключается в возможности работать с готовыми текстами спецификаций без приложения значительных усилий на их полную переработку. Предлагаемый способ был реализован и используется для преобразования реальных сценариев, состоящих из десятков шагов каждый.

Пример-иллюстрация: исходный вид

В качестве иллюстрации будем использовать описание процесса подготовки статьи для публи-

кации в материалах научной конференции. В текстовом виде оно имеет следующий вид:

Организаторы конференции рассылают информационное сообщение. Получив сообщение, авторы готовят начальную версию статьи и отправляют ее организаторам. Организаторы рассылают полученные работы рецензентам. Рецензенты изучают присланные работы и готовят рецензии. Получив результаты рецензирования, организационный комитет принимает решение о приеме работы к публикации и информирует авторов о своем решении. В случае положительного решения авторы с учетом замечаний готовят окончательную версию статьи и слайды доклада и отправляют организаторам для корректурной правки. В случае наличия замечаний авторы должны их устранить. Вместе с этим авторы должны предоставить дополнительные документы (акт приема-передачи, лицензионное соглашение). Организаторы осуществляют проверку дополнительных документов. Процесс заканчивается после устранения замечаний.

Данное описание (хотя и неформализованное) дает представление о том, как выглядит процесс подачи статьи для публикации. Мы видим, что в процессе принимают участие три объекта (организаторы, авторы, рецензенты), мы видим, с чего начинается и чем заканчивается процесс, а также получаем представление о последовательности действий. Мы также видим, что некоторые действия происходят всегда, в то время как другие происходят только при выполнении некоторых условий. Мы получаем представление о том, как взаимодействуют различные объекты: они обмениваются сообщениями, содержащими определенные данные. Такой способ взаимодействия как раз характерен для компонентов программно-аппаратных платформ, что и позволяет использовать этот абстрактный пример для описания предложенного способа.

Структурирование примера-иллюстрации

Человек, прочитавший представленное описание, может создать на ее основе эквивалентную блок-схему, однако автоматизация создания блок-схемы затруднительна. В данной главе описывается, каким образом можно с минимальными усилиями сделать подобное описание пригодным для машинной обработки.

Выявим сущности, используемые при описании нашего сценария. Прежде всего, мы имеем объекты, осуществляющие действия. Будем называть такие объекты *агентами*. Агентами в нашем сценарии являются "Организатор", "Автор", "Рецензент". *Действие* само по себе также является сущностью. Каждое предложение исходного описания описывает какое-либо действие. Как правило, сценарий предполагает определенную последовательность действий. Некоторые действия могут осуществляться только после того, как выполнены другие действия. В тексте последовательность возникает естествен-

ным образом: действия-предшественники упоминаются раньше, чем действия-следствия. *Сообщение* — еще одна сущность, характеризующая взаимодействие между агентами. Частью сообщения являются данные, необходимые для выполнения дальнейших действий. Наконец, мы видим, что переход от действия к действию характеризуется *условием перехода*.

Первым шагом структуризации является использование нумерованного списка для указания последовательности действий. Элемент списка соответствует атомарному действию. На усмотрение автора спецификации остается гранулярность атомарных действий. Одна и та же операция может быть представлена как одно действие или как последовательность нескольких действий. В зависимости от цели использования оптимальным может являться как один, так и другой вариант. Заметим, что последовательность элементов списка не дает полного представления о сценарии в случае наличия ветвлений или параллельного выполнения действий.

Вторым шагом является однозначное определение агента, осуществляющего действие. Как правило, идентификатор агента уже присутствует: это подлежащее предложения, которое описывает действие. Для сценариев с ветвлениями или параллельными действиями иногда необходимо идентифицировать отдельные действия (как это делать, будет описано ниже).

Третьим шагом является определение всех непосредственных переходов между действиями, включая определение условий перехода. Значительная их часть не требует дополнительных действий — переход от какого-либо элемента списка к следующему элементу является переходом по умолчанию.

Вот каким образом выглядит структурированное описание нашего сценария:

1. *[[flow:подготовка_публикации]].*
2. *[[start]].*
3. *[[Организатор]] Рассылает информационное сообщение.*
4. *[[Автор]] Готовит начальную версию статьи.*
5. *[[Организатор]] Рассылает полученные работы рецензентам.*
6. *[[Рецензент]] Изучает присланные работы и готовит рецензии.*
7. *[[Организатор]] Принимает решение о принятии работы к публикации.*
8. *[[Автор]] Работа принята? [[to работа_принята, подготовка_документов 'да']] [[to end 'нет']],*
9. *[[Автор: работа_принята]] Готовит окончательную версию статьи и слайды доклада с учетом замечаний.*
10. *[[Организатор]] Осуществляет корректурную правку.*
11. *[[Организатор]] Есть замечания? [[to Автор: работа_принята 'да']] ?[[to конец 'нет']].*
12. *[[Автор: подготовка_документов]] Предоставляет дополнительные документы. {{ИТ: сведения об*

авторах, экспертное заключение о возможности публикации}}.

13. *[[Организатор]] Осуществляет проверку документов.*

14. *[[Организатор]] Есть замечания? [[to Автор: подготовка_документов 'да']] [[to конец 'нет']].*

15. *[[Организатор:конец]] [[merge]].*

16. *[[end]].*

Структурированное описание представляет собой нумерованный список и содержит служебную информацию (тэги) внутри двойных квадратных скобок. В данном примере ключевые слова выделены жирным шрифтом для улучшения восприятия.

Текст спецификации начинается с элемента списка, содержащего тэг *[[flow:label]]*, где *label* — имя сценария, а заканчивается элементом списка, являющегося тэгом *[[end]]*. Каждый элемент списка представляет собой задачу, выполняемую на данном этапе, или блок ветвления. Каждый элемент списка должен начинаться с тэга *[[agent_name<:label>]]*, где *agent_name* — имя агента, выполняющего данное действие, а *label* — необязательная метка, которую можно использовать как указатель на действие или блок ветвления. Также в элементе списка могут присутствовать другие тэги:

- тэг *[[start]]* обозначает одну из возможных точек, с которых начинается выполнение сценария. Допустимо наличие нескольких стартовых точек;
- тэг *[[end]]* после спецификации задачи обозначает, что после данной задачи выполнение сценария заканчивается. В отличие от предыдущего, данный тэг должен входить в состав элемента списка, описывающего задачу;
- тэг *[[to <agent:>label <"condition">]]* означает переход, отличный от перехода по умолчанию. Здесь *agent* — имя агента, которому передается управление, *label* — метка, к элементу с которой осуществляется переход, *condition* — условие, при котором выполняется данный переход. В случае отсутствия идентификатора агента имеется в виду переход к другому действию текущего агента. Если тэга *[[to ...]]* в элементе списка нет, то по умолчанию считается, что после выполнения задачи выполняется следующая за ней в списке. Если же этот тэг присутствует, то возможны несколько вариантов. Если такой тэг один в элементе списка, тогда после выполнения задачи, соответствующей этому элементу, осуществляется переход к элементу с указанной меткой. Если данных тэгов в элементе списка несколько, то по умолчанию считается, что данный элемент соответствует блоку, в котором происходит разветвление сценария в зависимости от условий, заданных в *conditon*.

Описанные выше тэги позволяют описать, как выглядит граф сценария, где вершинами являются действия, а ребрами — переходы между ними. В дополнение к этой базовой информации структурированный текст содержит дополнительные аннота-

ции внутри блоков {...}. Простейший случай использования аннотаций — добавление комментариев, распознаваемых при машинной обработке. Однако внутри аннотации могут определяться существенные свойства, расширяющие базовую модель сценария в контексте конкретного применения. Для того чтобы различить аннотации, добавленные для разных приложений, текст внутри блока {...} начинается с метки. Так, хотя в нашем примере сценарий описывает процесс подачи статьи в общем, метка ИТ говорит нам о том, что информация о дополнительных документах относится к требованиям журнала "Информационные технологии".

Преобразование структурированного текста в диаграмму

Программа для преобразования структурированного текстового описания сценариев была реализована как расширение к среде Microsoft Office. Текстовый процессор Microsoft Word был использован для спецификации сценариев. Объектная модель MS Word представляет богатый API для работы со списками, встречающимися в документе. Для визуализации сценариев был использован графический редактор Microsoft Visio. Одной из причин такого решения явилось то, что MS Visio включает в себя библиотеку графической нотации BPMN.

Как и SBVR, нотация BPMN (Business Process Model and Notation) [11] была разработана консорциумом OMG. Целью BPMN является предоставление способа описания, легко понятного широкой аудитории — от бизнес-аналитиков, создающих начальное черновое описание процессов, до разработчиков реальных технологий, реализующих эти процессы, и пользователей этих технологий.

Нотация BPMN призвана служить связующим звеном между фазой проектирования бизнес-процесса и фазой его реализации. Как указано выше, существует потребность в аналогичном связующем звене между фазами проектирования, разработки и тестирования вычислительных систем. Для описания сценариев поведения вычислительных систем мы предлагаем использовать подмножество элементов BPMN.

Для иллюстрации использования BPMN рассмотрим диаграмму примера-иллюстрации (рис. 1). Сущности, описанные в предыдущей главе, отображаются на элементы BPMN следующим образом.

А. Агенты

Агент соответствует элементу BPMN "дорожка". Создание диаграммы сценария начинается с определения набора агентов. Объекты агентов служат контейнерами для других элементов диаграммы. В нашем примере в сценарии участвуют три агента ("Организатор", "Автор", "Рецензент"). Мы используем вертикальную ориентацию агентов, поскольку она является основной для диаграмм последова-

тельности сообщений, которые часто используются при проектировании вычислительных систем.

В. Действия

Действия описывают атомарные операции, выполняемые агентами. В нашем примере большинство действий совершает Организатор: он рассылает информационное сообщение, затем рассылает полученные работы для рецензии, принимает решение о принятии статьи к публикации и т. д. Действия соответствуют элементу BPMN "действие" и изображаются прямоугольниками со скругленными углами. Текст внутри прямоугольника содержит описание действия.

С. Блоки ветвления

Блоки ветвления описывают логические разветвления в сценарии, при которых то, какая задача будет выполняться следующей, зависит от некоего логического условия, определенного внутри него. Этот элемент соответствует элементу BPMN "шлюз" и изображается ромбом. В нашем примере сценарий будет иметь разный вид для принятых и отклоненных работ, что и указано наличием блока ветвления внутри агента "Автор".

Разновидностью блока ветвления является элемент объединения, соответствующий элементу "параллельный шлюз" BPMN. Он изображается ромбом, внутри которого помещен знак '+'. Наличие элемента объединения перед действием указывает на то, что агент начинает выполнять действие только после того, как завершатся все действия-предшественники. В случае отсутствия элемента объединения перед действием предполагается, что агент начинает действие, когда хотя бы одно из действий-предшественников завершено.

Д. Соединяющие объекты (коннекторы)

Коннекторы определяют последовательность выполнения задач. Коннектор типа "последовательность" соединяет задачи внутри одного объекта и соответствует элементу BPMN "поток управления" и изображается сплошной линией со стрелкой, указывающей направление перехода. Коннектор типа "сообщение" соединяет задачи в разных агентах и таким образом определяет, как происходит взаимодействие агентов. Он изображается штриховой линией. Сообщения могут представлять, например, команду одного агента другому, отклик на ранее полученную команду и т. п. Такое использование сообщений позволяет, помимо прочего, сократить структурированное текстовое описание. Свободный текст "Организаторы конференции рассылает информационное сообщение. Получив сообщение, авторы готовят начальную версию статьи" превращается в лаконичное:

1. [[Организатор]] рассылает информационное сообщение.

2. [[Автор]] готовит начальную версию статьи.

Наличие в нумерованном списке смежных элементов с разными идентификаторами агентов подразумевает пересылку сообщения от первого ко второму. С коннекторами может быть связано условие перехода от задачи к задаче, когда задача-следствие активируется только при выполнении определенного условия. Условие изображается как текст, ассоциированный с коннектором. "Сообщение" соответствует элементу BPMN "поток сообщений".

Е. Аннотации

Аннотации, добавляемые к объектам диаграммы, служат для уточнения значения элементов диаграммы и повышения ее информативности. Для изображения аннотаций на диаграмме-примере используются объекты-ссылки с текстом, содержащимся внутри блока {{...}}.

Ф. Вспомогательные элементы

Нотация BPMN также содержит элементы "начало" и "конец". Они используются для улучшения читаемости диаграммы. Для обозначения начала сценария используется символ окружности. Для обозначения конца сценария используется также символ

окружности жирной линией. В контексте задачи спецификации поведения вычислительных систем эти элементы применяются для обозначения действий, с которых может начинаться сценарий, и действий, после которых сценарий завершается. В общем случае сценарий может иметь несколько элементов "начало" и "конец".

Реализация генератора диаграмм

Диаграмма на рис. 1 была автоматически создана из приведенного выше структурированного текстового описания с помощью программы-конвертера. Блок-схема работы генератора диаграмм представлена на рис. 2.

Сначала входное текстовое описание проверяется на отсутствие ошибок. Ошибкой является, например, ссылка на неопределенную метку. В случае обнаружения ошибок пользователю сообщается диагностика и предлагается устранить ошибки. Затем определяется взаимное расположение агентов, а также размеры и координаты агентов, действий, блоков ветвления. В нашем случае агенты упорядочены в порядке первого упоминания во входной спецификации (это являлось требованием заказчика). В общем случае заслуживают внимания другие критерии упорядочивания агентов, например, минимизация суммарной длины соединяющих коннекторов. Затем в диаграмму добавляются коннекторы. Добавление коннектора сводится к определению его начальной и конечной точек. Трассировка между ними осуществляется встроенными средствами MS Visio.

Критерием выбора начальной и конечной точек коннектора является расстояние между ними. Мы стремимся минимизировать это расстояние при сохранении читаемости диаграммы. MS Visio позволяет помещать объекты диаграммы на слои. Мы используем это свойство при работе с аннотациями. Как было указано выше, частью аннотации может быть метка. При наличии метка интерпретируется как идентификатор слоя, на котором должна размещаться аннотация. Аннотации с разными метками помещаются на разные слои. Аннотации без метки считаются комментариями, которые автор спецификации желает видеть на диаграмме. Они помещаются на базовый слой.

Слой является удобным механизмом детализации базовой спецификации. В случае необходимости детализации для какого-то конкретного приложения название приложения может быть использовано в качестве иденти-

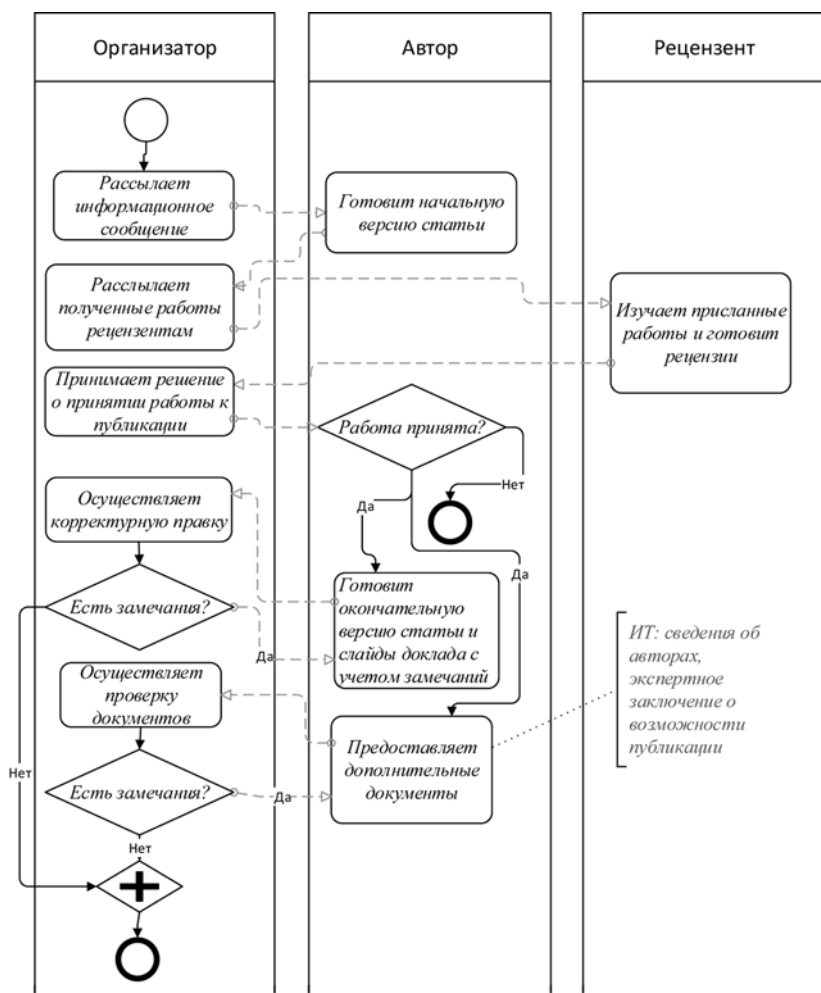


Рис. 1. Процесс подготовки статьи к публикации

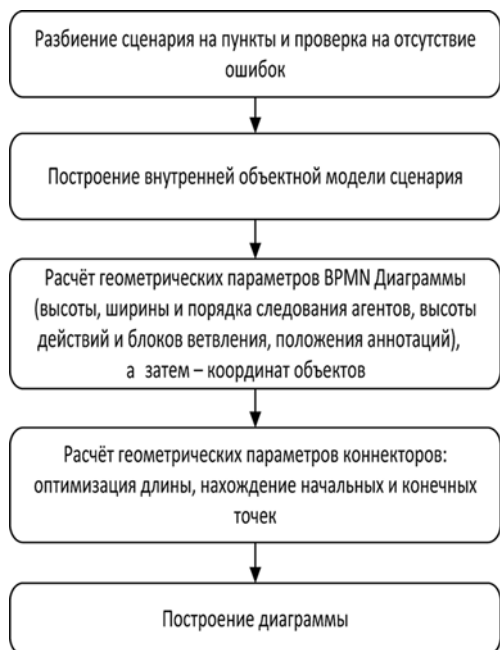


Рис. 2. Схема работы генератора диаграмм

фикатора слоя. Дополнительные данные могут быть определены как аннотации, принадлежащие этому слою. Таким образом, они могут быть отфильтрованы от других имеющихся аннотаций при машинной обработке.

Заключение

Описанный способ преобразования текстового описания в диаграмму последовательности операций реализован и успешно применен на практике для сценариев, описывающих поведение сложных вычислительных систем. Генератор диаграмм использовался не только для визуализации финальной версии сценария, но и как инструмент самоконтроля авторов спецификаций сценариев на промежуточных стадиях. В этом случае диаграмма служила инструментом визуального контроля соответствия написанного текста намерениям автора. Простота структурирования неформализованного текста и использование BPMN в качестве графической нотации способствовали широкому использованию предложенного способа.

Вместе с тем, существует целый ряд направлений для дальнейших улучшений. Одной из задач является анализ качества спецификации в виде структурированного текста, включающий в себя выявление неоднозначности, неполноты, противоречивости. Желательно иметь возможность автоматически определять проблемные места и информировать о них пользователя.

Другой задачей является повышение качества планировки диаграммы последовательности опе-

раций. Более компактное расположение элементов диаграммы позволит улучшить восприятие сценариев, состоящих из большого числа действий. Поддержка иерархической спецификации, когда диаграмма сценария верхнего уровня содержит ссылки на подсценарии, позволит создать удобную спецификацию сценариев с повторяющимися подсценариями.

Наконец, путем продуманного добавления аннотаций к базовому описанию сценария (или группы сценариев) можно довести спецификацию до состояния, в котором она может быть использована для автоматизированного создания моделей, пригодных для формального анализа корректности поведения вычислительной системы.

Авторы планируют продолжить исследования по этим направлениям.

Список литературы

1. Mich L., Franch M., Inverardi P. N. Market Research for Requirements Analysis Using Linguistic Tools // Requirement Engineering Journal. 2004. N. 9 (1). P. 40–56.
2. Kiyavitskaya N., Zeni N., Mich L., Berry D. M. Requirements for Tools for Ambiguity Identification and Measurement in Natural Language Requirements Specifications // Requirement Engineering Journal. 2008. N. 13. P. 207–240.
3. Hussain I., Ormandjieva O., Kosseim L. Automatic Quality Assessment of SRS Text by Means of a Decision-Tree-Based Text Classifier // In Proceedings of the 7th International Conference on Quality Software. 2007. P. 209–218.
4. Tjong S. F., Hallam N., Hartley M. Improving the Quality of Natural Language Requirements Specifications through Natural Language Requirements Patterns // In Proceedings of the 6th IEEE International Conference on Computer and Information Technology. 2006.
5. Kamalrudin M., Hosking J., Grundy J. Improving Requirements Quality Using Essential Use Case Interaction Patterns // In Proceedings of the 33rd International Conference on Software Engineering. 2011.
6. Denger C., Berry D. M., Kamsties E. Higher Quality Requirements Specifications through Natural Language Patterns // In Proceedings of the IEEE International Conference on Software — Science, Technology and Engineering — SwSTE'03. 2003.
7. Holtmann J., Meyer J., von Detten M. Automatic Validation and Correction of Formalized, Textual Requirements // In Proceedings of the IEEE 4th International Conference on Software Testing, Verification and Validation Workshops — ICSTW. 2011.
8. Gelhausen T., Tichy W. F. Thematic Role Based Generation of UML Models from Real Word Requirements // In Proceedings of the 1st IEEE International Conference on Semantic Computing — ICSC. 2007. P. 282–289.
9. Sharma V. S., Sarkar S., Verma K., Panayappan A., Kass A. Extracting High-level Functional Design from Software Requirements // In Proceedings of the 16th IEEE Asia-Pacific Software Engineering Conference — APSEC. 2009. P. 35–42.
10. Deeptimahanti D. K., Sanyal R. An Innovative Approach for Generating Static UML Models from Natural Language Requirements // Advances in Software Engineering. Communications in Computer and Information Science. 2009. V. 30. P. 147–163.
11. URL: <http://www.bpmn.org/> Дата обращения: 01.01.2014.
12. URL: <http://www.omg.org/> Дата обращения: 01.01.2014.
13. Afreen H., Bajwa I. S. Generating UML Class Models from SBVR Software Requirements Specifications // In 23rd Benelux Conference on Artificial Intelligence — BNAIC. 2011. P. 23–32.
14. URL: <http://www.omg.org/spec/SBVR/> Дата обращения: 01.01.2014.
15. URL: <http://www.omg.org/spec/> Дата обращения: 25.04.2014.

Generation of Flow Diagram from Textual Description of the Flow

Nowadays computer systems design is a very complex process, and the specification of components interaction is a serious part of the design process. The quality of the specification is the important factor for system design and test efficiency. Commonly used model of system specification is a set of scenarios (or flows) of components interaction. There is no common standard for flow description, and now 95 % of such specifications are textual scenarios descriptions. They are commonly inaccurate, inconsistent, incomplete and cannot be automatically processed. Here in this work we propose a method of converting textual flow specifications to a form with better readability, usable for automatic processing and more convenient for data sharing between project stakeholders. Special tags are added to the textual flow description to structure it, and then flow diagram is automatically generated from the text using BPMN notation. This method was implemented and successfully used for improving complex systems specifications.

Keywords: behavior model, behavior scenario, structured text, automatic specification processing, flow diagram, BPMN

References

1. Mich L., Franch M., Inverardi P. N. Market Research for Requirements Analysis Using Linguistic Tools. *Requirement Engineering Journal*. 2004. N. 9 (1). P. 40–56.
2. Kiyavitskaya N., Zeni N., Mich L., Berry D. M. Requirements for Tools for Ambiguity Identification and Measurement in Natural Language Requirements Specifications. *Requirement Engineering Journal*. 2008. N. 13. P. 207–240.
3. Hussain I., Ormandjieva O., Kosseim L. Automatic Quality Assessment of SRS Text by Means of a Decision-Tree-Based Text Classifier. In *Proceedings of the 7th International Conference on Quality Software*. 2007. P. 209–218.
4. Tjong S. F., Hallam N., Hartley M. Improving the Quality of Natural Language Requirements Specifications through Natural Language Requirements Patterns. In *Proceedings of the 6th IEEE International Conference on Computer and Information Technology*. 2006.
5. Kamalrudin M., Hosking J., Grundy J. Improving Requirements Quality Using Essential Use Case Interaction Patterns. In *Proceedings of the 33rd International Conference on Software Engineering*. 2011.
6. Denger C., Berry D. M., Kamsties E. Higher Quality Requirements Specifications through Natural Language Patterns. In *Proceedings of the IEEE International Conference on Software – Science, Technology and Engineering – SwSTE'03*. 2003.
7. Holtmann J., Meyer J., von Detten M. Automatic Validation and Correction of Formalized, Textual Requirements. In *Proceedings of the IEEE 4th International Conference on Software Testing, Verification and Validation Workshops – ICSTW*. 2011.
8. Gelhausen T., Tichy W. F. Thematic Role Based Generation of UML Models from Real Word Requirements. In *Proceedings of the 1st IEEE International Conference on Semantic Computing – ICSC*. 2007. P. 282–289.
9. Sharma V. S., Sarkar S., Verma K., Panayappan A., Kass A. Extracting High-level Functional Design from Software Requirements. In *Proceedings of the 16th IEEE Asia-Pacific Software Engineering Conference – APSEC*. 2009. P. 35–42.
10. Deepthimahanti D. K., Sanyal R. An Innovative Approach for Generating Static UML Models from Natural Language Requirements. *Advances in Software Engineering. Communications in Computer and Information Science*. 2009. V. 30. P. 147–163.
11. URL: <http://www.bpmn.org/> Дата обращения: 01.01.2014.
12. URL: <http://www.omg.org/> Дата обращения: 01.01.2014.
13. Afreen H., Bajwa I. S. Generating UML Class Models from SBVR Software Requirements Specifications. In *23rd Benelux Conference on Artificial Intelligence – BNAIC*. 2011. P. 23–32.
14. URL: <http://www.omg.org/spec/SBVR/> Data obrasheniya: 01.01.2014.
15. URL: <http://www.omg.org/spec/> Data obrasheniya: 25.04.2014.

УДК 004.912

М. В. Бочков, д-р техн. наук, проф.,
НОУ ДПО "Центр предпринимательских рисков", г. Санкт-Петербург,
П. Н. Бойков, вед. специалист,
ОАО "НИИ "Рубин", г. Санкт-Петербург, e-mail: boykova@yandex.ru

Инфодинамическая модель поиска пользователя в социальной сети

Исследованы закономерности представления пользователями соцсетей своих данных, на основе которых разработан алгоритм формирования оптимальной стратегии поиска в социальных сетях.

Ключевые слова: социальная сеть, поиск в социальной сети, алгоритм оптимизации поиска в социальных сетях, инфодинамическая модель

Социальная сеть как объект исследования

Среди ресурсов в сети Интернет все большую популярность приобретают онлайн-социальные сети (ОСС). К типовым возможностям их участников можно отнести:

- обмен информационными ресурсами с другими участниками ОСС;
- публикация и обсуждение идей;
- выбор и участие в социальных группах (сообществах);
- использование развлекательных и досуговых сервисов ОСС и др.

Очевидной тенденцией в развитии ОСС является рост числа пользователей и развитие их функциональных сервисов [1]. Динамика изменения числа пользователей наиболее популярных ОСС представлена на рис. 1 (см. вторую сторону обложки).

Информационную основу ОСС образуют персональные пользовательские страницы. Как правило, создатели ОСС стремятся получить от пользователя максимум информации. С этой целью регистрационная форма предлагает опубликовать максимум идентификационной и другой персональной информации. На рис. 2 (см. вторую сторону обложки) показаны набор регистрационных данных в ОСС "ВКонтакте".

Очевидно, что наиболее полное представление пользователями своих данных повышает точность и полноту результатов запроса и, следовательно, однозначность идентификации участников ОСС. Вместе с тем, среднестатистический пользователь подсознательно стремится представить минимум информации о себе, ограничить круг своего общения, обеспечив себе комфортное общение в ОСС. Таким образом, проявляется конфликт интересов владельцев и пользователей ОСС — одни хотят знать все, а другие хотят обойтись минимумом информации о себе [2].

Постановка задачи. Целью настоящего исследования является выявление закономерностей представления идентифицирующей пользователя информации в ОСС. Знание таких закономерностей и их описание в виде формальной модели позволит сформировать оптимальную стратегию поиска, при которой вероятность точного нахождения требуемого пользователя ОСС за минимальное число итераций поиска будет максимальна.

Модель представления информации пользователями ОСС

Исходные данные. В настоящем исследовании использован модельный фрагмент ОСС, сформированный путем обезличивания репрезентативного дампа общедоступных в сети Интернет пользовательских страниц. На основе полученной информа-

ции проведен расчет статистик, характеризующих атрибуты регистрационных данных пользователей [3].

Ранжирование пользовательских атрибутов и интерпретация полученных результатов. Для последующих исследований были выделены следующие пользовательские атрибуты:

- идентификатор пользователя в ОСС (a_0);
- фамилия (a_1);
- имя (a_2);
- город проживания (a_3);
- пол пользователя (a_4);
- дата рождения (a_5);
- наименование и год окончания вуза (a_6);
- наименование и год окончания школы (a_7);
- место работы (a_8);
- семейное положение (a_9).

Для исследований были введены следующие упрощения:

- параметр a_0 не использован при анализе, так как является уникальным для каждого пользователя и однозначно идентифицирует пользователя в социальной сети;
- параметры a_4 и a_9 исключены из анализа ввиду малого диапазона принимаемых значений, незначительного влияния на результаты поиска;
- аналогичные значения атрибута a_2 "Иван", "Ваня", "Ванька" считаются эквивалентными поисковому запросу "Иван";
- для атрибута a_3 значения, подобные "СПб", "Санкт-Петербург", "Питер", считались эквивалентными поисковому запросу "Санкт-Петербург".

На рис. 3 (см. вторую сторону обложки) представлены ранжированные значения вероятностей присутствия пользовательских атрибутов, отражающих закономерности отображения информации в ОСС. В дальнейшем данная закономерность используется в качестве модели представления информации пользователями ОСС.

Алгоритм формирования максимальной стратегии

Использование модели представления в свою очередь позволяет решить задачу формирования оптимального алгоритма поиска пользователя на основе математического аппарата инфодинамического моделирования [4], который основывается на анализе распределений значений результатов поиска от используемых атрибутов. Оригинальность предлагаемого подхода состоит в том, что основные формальные соотношения получены в дифференциальной форме (т. е. позволяют оценить влияние конкретных значений атрибутов на конкретные результаты поиска), а средние оценки (энтропия, условная и взаимная информация) представляют усреднение дифференциальных оценок. Такой подход позволяет более детально проанализировать информационные связи в системе отношений между информационными атрибутами и результатами поиска.

Алгоритм формирования максимальной стратегии на основе математического аппарата инфодинамического моделирования состоит из трех этапов.

1. Вначале определяем ряд информационных оценок (собственная информация, условная информация и взаимная информация) поисковой функции и информационных атрибутов [4].

Собственной информацией $I(a_j)$ (или количеством собственной информации) значения информационного атрибута a_j переменной \mathfrak{A}_j поисковой функции f называется величина

$$I(a_j) = -\log p(a_j),$$

где $p(a_j)$ — вероятность того, что информационный атрибут указан пользователем на странице ОСС.

Условной информацией $I(\beta|a_j)$ в значении β переменной \mathfrak{A}_j поисковой функции f при заданном значении информационного атрибута a_j называется величина

$$I(\beta|a_j) = I(\beta a_j) - I(a_j).$$

Взаимная информация между значениями β и a_j есть информация в значении информационного атрибута a_j о значении β поисковой функции f (и наоборот, информация в значении β о значении a_j) вычисляется следующим образом:

$$I(\beta; a_j) = I(a_j; \beta) = I(\beta) - I(\beta|a_j) = I(a_j) - I(a_j|\beta).$$

Взаимная информация между значениями информационного атрибута a_f и поисковой функцией f используется для оценки взаимного влияния атрибута и функции, т. е. данная информационная оценка указывает степень информационной связи присутствия атрибута с конкретным значением поисковой функции.

2. Проведем исследование пользовательских атрибутов для определения влияния атрибутов на результаты поиска с учетом их встречаемости на пользовательских страницах в ОСС. Для этого на основе выборки пользователей, полученной из ОСС, выполним поиск случайных пользователей, изменяя атрибуты поиска и сравнивая полученные результаты. Поиск будем осуществлять на выборке случайных пользователей, на страницах которых содержатся все информационные атрибуты.

Рассчитаем вероятность того, что по заданным атрибутам результаты поиска будут успешными, т. е. пользователь будет найден:

$$P(f(\beta = 1)) = 1 - \frac{n}{N},$$

где n — результат поиска для a_j ; N — общее число пользователей.

Результаты расчетов значений условной энтропии поисковой функции и взаимной информации между поисковой функцией и значением каждого атрибута a_j показаны в табл. 1.

Как видно из результатов, приведенных в табл. 1, использование одного пользовательского атрибута в поисковом запросе, даже с учетом того, что пользователь указал на своей странице все атрибуты, не всегда позволяет найти пользователя в ОСС. Таким образом, целесообразно исследовать значения поисковой функции при использовании в поисковом запросе попарно двух атрибутов между собой. При этом в дальнейших расчетах будем использовать среднее значение вероятностей успешного результата поиска, значение которых получены как среднее арифметическое вероятностей для каждого атрибута выбранных пользователей (табл. 2).

В табл. 3, приведены следующие значения: U — номер теста, $P(U_j)$ — вероятность совместного

Таблица 1

Id_user	a_1	a_2	a_3	a_5	a_6	a_7	a_8
36	0,9802	0,9037	0,7861	0,9992	0,9791	0,9996	0,9999
2482	0,9523	0,7504	0,7861	0,9993	0,9842	0,9997	1
227385	0,9792	0,7719	0,9993	0,9998	0,9999	1	1
125812	0,9965	0,9958	0,9937	0,9992	1	0,9998	1
178272	0,9988	0,8178	0,6201	0,9991	0,9977	0,9997	0,9999
93388	1	0,9848	0,9995	0,9994	0,9996	0,9999	0,9998

Таблица 2

Id_user	a_1	a_2	a_3	a_5	a_6	a_7	a_8
Среднее значение	0,9845	0,8707	0,8641	0,9993	0,9934	0,9997	0,9999

Таблица 3

U	$P(U_j)$	$a_j a_j$	f
u_1	0,02	0 0 0 0 1 1	1
u_2	0,03	0 0 0 1 0 0 1	1
u_3	0,06	0 0 0 1 0 1 0	1
u_4	0,09	0 0 0 0 1 0 1	1
u_5	0,2	0 0 0 0 1 1 0	1
u_6	0,24	0 0 0 1 1 0 0	1
u_7	0,11	1 0 0 0 0 0 1	1
u_8	0,24	1 0 0 0 0 1 0	1
u_9	0,28	1 0 0 1 0 0 0	1
u_{10}	0,86	1 0 0 0 1 0 0	0
u_{11}	0,11	0 1 0 0 0 0 1	1
u_{12}	0,24	0 1 0 0 0 1 0	0
u_{13}	0,86	0 1 0 0 1 0 0	0
u_{14}	0,03	0 0 1 0 0 0 1	0
u_{15}	0,06	0 0 1 0 0 1 0	0
u_{16}	0,28	0 1 0 1 0 0 0	0
u_{17}	0,08	0 0 1 1 0 0 0	0
u_{18}	0,23	0 0 1 0 1 0 0	0
u_{19}	1	1 1 0 0 0 0 0	0
u_{20}	0,27	1 0 1 0 0 0 0	0
u_{21}	0,27	0 1 1 0 0 0 0	0

присутствия атрибутов a_j на пользовательской странице (данное значение рассчитано для всей полученной выборки) и f — значение решающей функции (результат поиска).

Исходя из результатов, приведенных в табл. 3, вероятности значений решающей функции $p(\beta)$ (где β — значение поисковой функции: 0 — результат поиска неудачен, 1 — поиск успешен), присутствия атрибутов a_j и их комбинаций приведены в табл. 4.

Следующим шагом определяем средние оценки (энтропия, условная и взаимная информация), характеризующие ансамбли значений поисковой функции от значений атрибутов поиска. Для отражения "привлекательности" поискового атрибута выбраны показатели условной энтропии и взаимной энтропии.

Условная энтропия $H(f|a_j)$ функции поиска f при заданном значении атрибута a_j

$$H(f|a_j) = \sum_{\beta=0}^1 p(\beta a_j) I(\beta|a_j).$$

Взаимная информация $I(f; a_j)$ между поисковой функцией f и атрибутом a_j определяется выражением

$$I(f; a_j) = \sum_{\beta=0}^1 p(\beta a_j) I(\beta; a_j).$$

Вычисление информационных оценок (для табл. 1) и результаты выбора переменных при построении дерева решений приведены в табл. 5.

3. Дадим содержательную интерпретацию информационных оценок и полученных соотношений с позиции анализа информационных атрибутов процедуры поиска.

Для выработки логически обоснованных критериев выбора состава и порядка атрибута при разработке стратегии поиска необходимо более детально проанализировать содержательную сторону приведенных формальных оценок. Для этого представим связь атрибута и поисковой функции следующим образом:

$$H(f) - H(f|a_j) = I(f; a_j) = H(a_j) - H(a_j|f).$$

Рассмотрим содержательно слагаемые этого выражения.

Энтропия $H(f)$ поисковой функции — среднее количество информации, которое необходимо извлечь для определения значения функции.

Энтропия $H(a_j)$ атрибута — среднее количество информации, которое извлекается при добавлении атрибута поиска.

Взаимная информация $I(f; a_j)$ — среднее количество информации, которое несет атрибут поиска о результатах поиска.

Для решения поставленной задачи интерес представляет взаимная информация как индикатор того, насколько уменьшится диапазон результатов поиска

Таблица 4

Атрибут	Вероятности	Значения a_j, β		Взаимная вероятность $p(\beta a_j)$	Комбинация значений	
		a_j	β		a_j	β
a_1	$p(a_j)$	0	1	$p(\beta a_1)$	0	1
	$p(\beta)$	1	1		0,3	0,2
a_2	$p(a_j)$	0,3.....0,7	0,8.....0,2	$p(\beta a_2)$	0,1	0,7
	$p(\beta)$	0,1.....0,9				
a_3	$p(a_j)$	0,3.....0,7	1.....0	$p(\beta a_3)$	0,3	0
	$p(\beta)$	0,7.....0,3				
a_5	$p(a_j)$	0,3.....0,7	0,7.....0,3	$p(\beta a_5)$	0,3	0,3
	$p(\beta)$	0,8.....0,2				
a_6	$p(a_j)$	0,3.....0,7	0,8.....0,2	$p(\beta a_6)$	0,3	0,2
	$p(\beta)$	0,7.....0,3				
a_7	$p(a_j)$	0,3.....0,7	0,7.....0,3	$p(\beta a_7)$	0,3	0,3
	$p(\beta)$	0,9.....0,1				
a_8	$p(a_j)$	0,3.....0,7	0,9.....0,1	$p(\beta a_8)$	0,3	0,1
	$p(\beta)$					

Таблица 5

Атрибут	a_1	a_2	a_5	a_6	a_7	a_8
$H(f a_j)$	0,358	0,158	0,363	0,358	0,396	0,279
$I(f; a_j)$	0,802	0,698	0,681	0,802	0,936	1,049

при наличии того или иного атрибута. Другими словами, из информации $H(a_j)$ оценивается та ее часть, которая позволяет уменьшить энтропию $H(f)$ функции до значения $H(f|a_j)$.

Под стоимостью решения будем понимать время отклика на выполнение поискового запроса и прием его одинаковым для каждого атрибута поиска. Таким образом, критерием оптимизации при выборе следующего теста будет выступать выражение

$$a_j^* = \max_{a_j \in A} I(f, a_j),$$

где a_j^* — следующий атрибут для добавления в запрос.

В общем виде задача формирования оптимальной стратегии поиска в терминах инфодинамического моделирования соответствует задаче конструирования деревьев решений. Процедуру конструирования дерева решений представим в виде алгоритма.

Шаг 1. Задача состоит в выборе атрибута, который целесообразно использовать первым. Результаты вычислений (табл. 6) показывают, что критерию

Таблица 6

Уровень дерева решений	Условие (известные значения переменных)	Атрибуты, из которых осуществляется выбор	Взаимная информация $I(f, a_j)$	Выбор
1	—	a_1 a_2 a_5 a_6 a_7 a_8	0,802 0,698 0,681 0,802 0,936 1,049	a_8
2	$a_8 = 1$	a_1 a_2 a_5 a_6 a_7	0,802 0,698 0,681 0,802 0,936	a_7
3	$a_7 = 1$	a_1 a_2 a_5 a_6	0,802 0,698 0,681 0,802	a_1
4	$a_1 = 1$	a_2 a_5 a_6	0,698 0,681 0,802	a_6
5	$a_6 = 1$	a_2 a_5	0,698 0,681	a_2

оптимизации удовлетворяет атрибут a_8 . Построим первый уровень дерева решений при условии, что корневому узлу дерева соответствие этого атрибута не позволит идентифицировать пользователя, поэтому на основе вероятности совместной встречаемости атрибутов целесообразно в качестве второго атрибута использовать a_1 . Если атрибут a_1 отсутствует, то в качестве второго атрибута выбирается атрибут, следующий по значению совместной вероятности встречаемости атрибутов с a_8 .

Шаг 2. Если атрибут a_8 отсутствует, то выбирается следующий по значимости атрибут, а алгоритм добавления второго атрибута в параметры запроса аналогичен шагу 1.

Шаг 3. На этом шаге выбирается переменная $a_8 = a_7 = 0$. Из таблицы решений видно, что поисковая функция f принимает максимальное значение только в случае, если будут известны следующие пары атрибутов: $a_1 a_5$ и $a_5 a_6$. Таким образом, целесообразно проверить общий для этих пар атрибут a_5 . Если он отсутствует, то поиск целесообразно прекратить, так как оставшихся атрибутов недостаточно для идентификации пользователя в социальной сети.

Дерево решений как результат оптимизации представлено на рис. 4.

Экспериментальные исследования полученного алгоритма поиска пользователя социальной сети

На заключительном этапе исследования практическая пригодность полученного алгоритма для поиска пользователя в социальной сети, в частности, определялось, возможно ли конструирование множества деревьев решений за приемлемое время. Затем оценивалось преимущество предложенного алгоритма по сравнению с алгоритмом последовательного добавления атрибутов поиска в поисковый запрос.

Для получения априорных сведений о пользователях социальной сети использовались три наиболее популярные социальные сети "ВКонтакте", "Фейсбук" и "Одноклассники". Выборка осуществлялась путем случайного отображения информации о пользователе социальной сети (такая функциональная возможность предоставляется OSS). В результате эксперимент проводился на общей выборке из 4500 случайно выбранных пользователей социальных сетей.

В ходе эксперимента были получены следующие результаты. Во всех трех социальных сетях алгоритм поиска пользователя сконструировал дере-

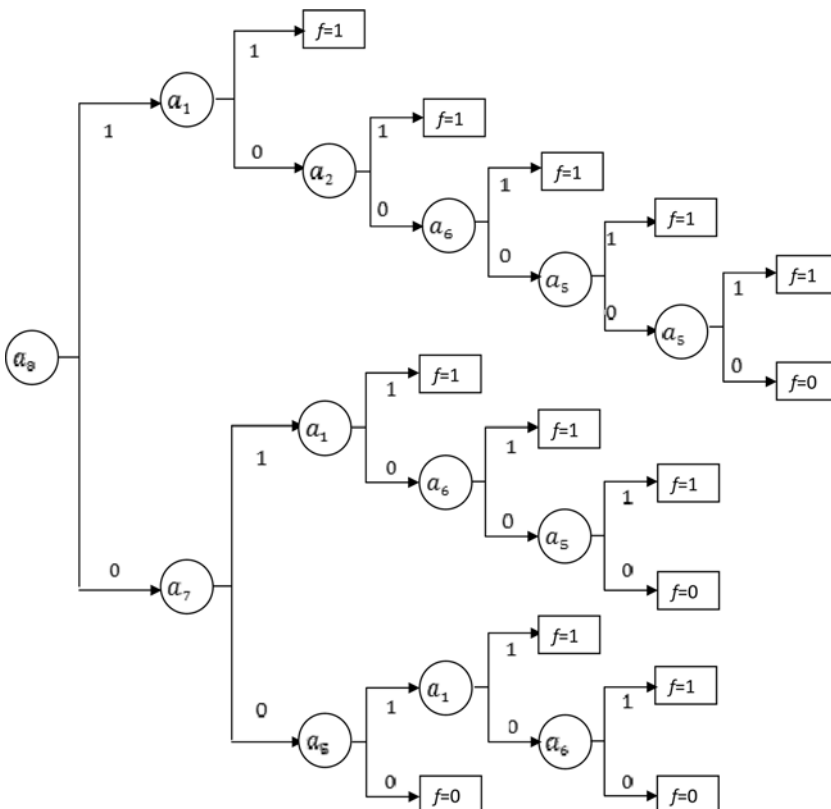


Рис. 4. Дерево решений на основе информационных оценок

Стоимость деревьев решений

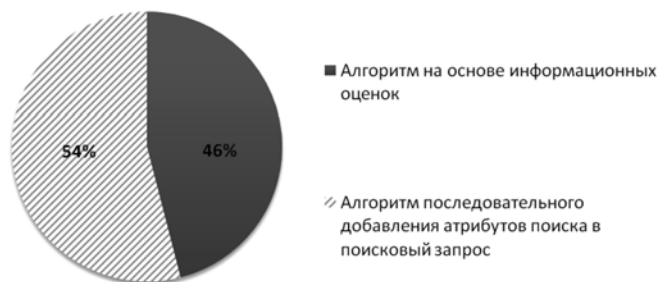


Рис. 5. Сравнение вычислительных затрат двух алгоритмов для успешного поиска пользователя в социальной сети

вья меньшей стоимостью, чем алгоритм последовательного добавления атрибутов поиска в поисковый запрос. В среднем стоимость деревьев решений на 8 % меньше (рис. 5). Также снизились и вычислительные затраты; так, в среднем для успешного поиска пользователя в социальных сетях необходимо использовать три атрибута, а не четыре.

Выводы

- ◆ Предложенный алгоритм конструирования дерева решений на основе информационных оценок

пользовательских атрибутов позволяет оптимизировать время поиска пользователей в ОСС при достаточно большом числе атрибутов поиска.

- ◆ Использование разработанного алгоритма позволяет сократить вычислительные затраты на обработку сервером социальной сети поискового запроса, тем самым давая возможность уменьшить время, затраченное аналитиком на поиск путем последовательного добавления пользовательских атрибутов в поисковый запрос.
- ◆ Полученный подход применим к любым базам данных, содержащим большое число объектов учета с множеством идентифицирующих и характеризующих объекты атрибутов, что позволяет говорить о формировании универсальной стратегии поисковой оптимизации.

Список литературы

1. Губанов Д. А., Новиков Д. А., Чхартишвили А. Г. Социальные сети: модели информационного влияния, управления и противоборства. МЦНМО, 2010.
2. Бочков М. В., Бойков П. Н., Яшин А. А. Социальные сети как основной источник утечки персональных данных // Inside # Защита информации. 2010. № 3.
3. Бочков М. В., Бойков П. Н. Способ автоматического рубрицирования неструктурированной информации в сети Интернет // Информационные технологии. 2012. № 2.
4. Курбацкий А. Н., Чеусhev В. А. Информационный метод анализа и оптимизации в системах поддержки принятия решений. Минск: Ин-т техн. кибернетики НАН Беларуси, 1999.

M. V. Bochkov, Professor, UKC "The Center of enterprise risks", SPb,
P. N. Boykov, Leading Specialist,
Public Corporation of Scientific Research Institute "Rubin", SPbf,
e-mail: boykovpn@yandex.ru

Infodinamicheskyy Model of Search of the User on a Social Network

Among resources in the Internet online social networks (OSN) are becoming more popular. Information basis OSN is formed personal user pages. Obviously, the most complete of the user's picture of their data increases the accuracy and completeness of query results, and is therefore uniquely of participants' identify OSN. On the other hand, the average user's subconsciously prefer to present a minimum information about themselves, limit their circles of contacts, securing a comfortable communication in OSN. The current article explores the patters of personal data presentation by social media users based on which an algorithm of an optimized social networks search strategy has been developed.

Keywords: social network, search in social network, the algorithm of an optimized social networks search strategy, infodinamicheskyy model

References

1. Gubanov D. A., Novikov D. A., Chxartishvili A. G. *Socialnir seti: mo informacionnogo vliyaniya, upravleniya i protivoborstva*. M.: MCNMO, 2010.
2. Bochkov M. V., Boykov P. N., Yashin A. A. *Socialnie seti kak osnovnoy istochnik utechki personalnich dannich. Inside. Zashita informacii*, Sant-Petersburg. 2010. N. 3.

3. Bochkov M. V., Boykov P. N. *Sposob avtomaticheskogo rubricirovaniya nestrukturirovannoy informacii v seti Internet. Informacionnie tehnologii*. 2012. N. 2.
4. Kurbackiy A. N., Cheushev V. A. *Informacionniy metod analiza i optimizacii v sistemach podderzhki prinyatiya resheniy*. Minsk: Institut tech. kibemetiki NAN Belarus, 1999.

В. Н. Костин, канд. тех. наук, доц., e-mail: vladimirkostin57@mail.ru, **Д. В. Даньшин**, студент, Оренбургский государственный университет

Метод оценки глубины прогноза развития (эволюции) характеристик сложных систем на основе энтропийного подхода

Предлагается на основе заданных соотношений энтропии (ошибок первого и второго рода) определять допустимое время прогнозирования развития характеристик любых сложных систем. Результаты прогнозирования могут использоваться для задания характеристик перспективных систем при их проектировании. Апробация метода проведена для оценки глубины прогнозирования развития социальной ситуации в целях формирования параметров модели нарушителя при проектировании систем физической защиты.

Ключевые слова: оценка глубины прогнозирования развития характеристик сложных систем, энтропия, ошибки первого и второго рода

Введение

При разработке любой новой системы (технической, экономической, экологической, социальной и т. д.), рассматривая процесс совершенствования как эволюционное развитие системы, необходимо задать перспективные (оптимальные) характеристики будущей системы на определенный момент времени. В качестве меры эволюции системы выступает информационная энтропия. Очевидно, что чем дальше по времени осуществляется прогноз развития, тем больше неопределенность и тем меньше достоверность и надежность получаемых характеристик, т. е. имеется оптимально приемлемое время прогнозирования.

Существует множество математических методов, которые позволяют прогнозировать поведение параметров системы во времени [1]. Однако методов оценки глубины прогнозирования развития систем крайне мало. Среди них марковские модели, которые позволяют определять время наступления предельных вероятностей переходов состояний в системе для однородной марковской модели.

Особенность информационного метода заключается в определении изменения прогнозируемых параметров развития системы в виде порции приращения соотношений энтропии. При достижении заданного значения соотношений энтропии (ошибок прогнозирования) проводится оценка характеристик системы и глубины прогнозирования по времени. Полученные характеристики являются оптимальными, так как система обладает наилучшим соотношением энтропии как меры степени упорядоченности и организованности структуры системы. Если порция энтропии от предыдущей системы будет велика, то система будет недоразвитой, т. е. это модернизация старой системы (не новая система). Если порция энтропии от предыдущей системы мала, то возникает опасность, что новая система будет не приспособлена (не адаптирована) к существующей

среде, т. е. новая система обладает малой преемственностью от предыдущей системы. При построении метода оценки глубины прогнозирования использовался принцип максимума энтропии.

Данные вопросы рассматривались в зарубежных источниках [2], а также в работах [3, 4] только с позиции теории принятия решений как выбор наилучшей системы из совокупности существующих систем. Результатом же применения данного метода являются прогнозируемые характеристики развития перспективной системы на определенный момент времени, которые являются входными данными для проектирования.

Постановка задачи

Необходимо путем прогноза развития параметров сложной системы определить ее характеристики и глубину прогнозирования по времени через заданные ошибки первого и второго рода в виде соотношений энтропии.

Известно, что степень упорядоченности и организованности структуры любой системы можно описать с помощью энтропии [3, 4], поэтому каждый период развития системы определялся в виде оценочного потенциала — энтропии. Математический аппарат метода опирается на материал, изложенный в работе [3].

Формализацию задачи осуществим на основе концепции метода комитетов, рассматривая ее с позиции правомерного существования схемы нахождения сложного предпочтения на множестве временных интервалов развития системы. Комитетом оценки эффективности принимаемых решений назовем такое конечное множество векторов класса R^n : $K = (x_1, x_2, \dots, x_n) \in R^n$, для которого принятие нулевой гипотезы H_0 при заданном уровне значимости обеспечивает определенное число компонент вектора X_m (характеристик системы).

Решение сформулированной задачи связано, во-первых, с определением вектора X_m , количественно отображающего потенциал системы; во-вторых, с оценкой уровня значимости $\alpha_0^{\text{эф}}$ и мощности критерия $\beta_i^{\text{эф}}$ стратегического прогноза. Информационное отображение конкретной ситуации при этом укладывается в следующую схему: имеется n сравниваемых между собой периодов развития системы; каждому периоду существования системы поставлена в соответствие совокупность m характеристик, определяющих потенциал системы в данное время. В этом случае ситуация прогнозирования в развернутой форме характеризуется табл. 1, у которой один вход образован множеством временной шкалой развития системы $\{A_j\}$, а другой — множеством ее признаков.

Вес параметров в формировании оценочного потенциала характеризуется количественной мерой степени уверенности в ситуации объективно существующей неопределенности и отождествляется с распределением вероятностей p_{ji} . Компоненты $\{X_{ji}\}$ задаются в единых или различных физических шкалах. Поэтому для приведения компонент $\{X_{ji}\}$ к единой общей шкале воспользуемся естественной нормализацией, осуществляемой относительно экстремальных значений компонент $\{X_{ji}\}$ как без смены ингредиента на противоположный

$$r_{ji} = \frac{x_{ji}}{x_{\max j}}, \quad (1)$$

так и со сменой ингредиента на противоположный

$$r_{ji} = \frac{x_{\min j}}{x_{ji}} \quad (2)$$

с отображением в $x_{ji} \rightarrow r [0, 1]$. Зависимости (1) и (2) обеспечили отображение выборочного пространства, приведенного в табл. 1, в другое, имеющее мощность континуума, приведенного в табл. 2.

Элементы r_{ji} (пространства мощности континуума) в единой шкале будем идентифицировать с элементарными событиями. При этом определенная на r_{ji} нормированная мера соответствует вероятности $p(r)$, которая отождествляется с понятием интегрального потенциала заданного комплекса элементарных событий. Смысл данной меры состоит в том, чтобы соответствующим образом интерпретировать понятие вероятность. При этом вероятность как категория диалектики совмещает в себе и меру объективной возможности события и степень субъективной уверенности в появлении событий. В целях формализации задачи выбора решения идентифицируем множества систем с пространством событий $\{A\}$, а множество признаков — с со-

Морфологическая матрица

Характеристика системы (признак)	Временная шкала развития системы				
	$\{A_1\}$...	$\{A_j\}$...	$\{A_n\}$
X_1	X_{11}	...	X_{1j}	...	X_{1n}
X_j	X_{j1}	...	X_{ji}	...	X_{jn}
...
X_m	X_{m1}	...	X_{mi}	...	X_{mn}

Таблица 2

Матрица принятия решений

Характеристика системы (признак)	Временная шкала развития системы				
	$\{A_1\}$...	$\{A_j\}$...	$\{A_n\}$
X_1	r_{11}	...	r_{1j}	...	r_{1n}
X_j	r_{j1}	...	r_{ji}	...	r_{jn}
...
X_m	r_{m1}	...	r_{mi}	...	r_{mn}

бытиями $\{x\}$. Тогда связь между всеми компонентами, формирующими оценочный потенциал ситуации, осуществляется через определенную на этих компонентах нормированную меру, которая отождествляется с вероятностью $p(a)$. Распределение вероятностей $p(a)$ обеспечивает оценку нулевой гипотезы H_0 при заданном уровне значимости $\alpha_0^{\text{эф}}$ и мощности критерия $\beta_0^{\text{эф}}$ прогноза развития системы.

Информация в вероятностно-статистической теории выступает в качестве снимаемой уменьшаемой неопределенности, а ее количество измеряется с помощью энтропии через вероятность.

Поэтому дальнейшее построение метода связано с исследованием законов преобразования информации поля декартова произведения двух множеств (систем и признаков) в количественные составляющие информации. С этой целью в логическую схему введем такие понятия, как априорные, апостериорные и условные вероятности, применим теорему Байеса и формулу полной вероятности, а также введем понятие условной вероятности p проявления j -й характеристики в формировании оценочного потенциала при условии, что события, формирующие оценочный потенциал, произошли.

Для получения зависимости определения величины $p(r)$, являющейся нормированной мерой на элементарных событиях $\{r\}$, воспользуемся тем, что понятие оценочного потенциала заданного комплекса элементарных событий можно отождествить с функцией принадлежности, которая ставит в соответствие каждому r действительное число в интервале $[0, 1]$. При этом, не нарушая общности рас-

суждений, искомая зависимость функции принадлежности представляется в форме

$$p_{ji}(r) = r_{ji} / \sum_{i=1}^n r_{ji}. \quad (3)$$

Один из методов расчета вероятности проявления j -го признака сравниваемых интервалов времени на формирование оценочного потенциала основан на введенном В. В. Хоменюком понятии потенциального распределения вероятности [3, стр. 106]. Формализованный расчет этих вероятностных оценок связан с принципом максимума неопределенности, сами оценки получаются в результате решения следующей задачи на условный экстремум:

$$H(P) = - \sum_{j=1}^m p_j \log p_j \rightarrow \max; \quad (4)$$

$$\sum_{j=1}^m p_j = 1; \quad (5)$$

$$\prod_{j=1}^m r^{p_j} = \text{const}. \quad (6)$$

Формула (4) — это информационная энтропия Больцмана—Шеннона, выступающая в качестве меры неопределенности, где p_j — вероятность проявления j -го признака системы. Формула (5) является условием нормировки, а (6) постулирует постоянство среднегеометрического показателя:

$$r_{ji} = \sum_{i=1}^n r_{ji}^{p_j}. \quad (7)$$

Отыскивая условный экстремум (4)—(6) по правилу неопределенного множителя Лагранжа, получаем зависимость для определения оценок вектора потенциального распределения вероятностей, доставляющей максимум целевой функции (4):

$$\hat{p}_j(r) = \sum_{i=1}^n r_{ji} / \sum_{j=1}^m \sum_{i=1}^n r_{ji}. \quad (8)$$

Принцип потенциального распределения вероятностей основан на том, что предпочитается выбор с большей вероятностью тех признаков системы, свойства которых имеют больший вклад в суммарное значение оценочного потенциала. При этом отметим, что для принципа потенциального распределения вероятностей (8) априорная информация о состоянии характеристик основана на принципе недостаточности знаний.

Однако не вызывает сомнения тот факт, что вес различных признаков в формировании оценочного потенциала различен. Получение оценок априорного распределения p_j связано с отношением порядка предпочтения, которое подробно исследовано в тру-

дах Фишборна [5]. Для простого линейного отношения порядка оценки Фишборна априорных вероятностей образуют убывающую арифметическую прогрессию:

$$\check{p}_j = 2(m - j + 1)/(m(m + 1)). \quad (9)$$

Вводя на основе оценок Фишборна априорную вероятность, мы, задавая "вход" в модель, учитываем различный вес характеристики в формировании оценочного потенциала системы. Затем, используя принцип потенциального распределения (8) и положения теоремы Байеса, получаем логически обоснованный "выход" из модели в виде апостериорных условных вероятностей в форме

$$p_j = \sum_{i=1}^n r_{ji} \check{p}_j / \sum_{j=1}^m \sum_{i=1}^n r_{ji} \check{p}_j. \quad (10)$$

Таким образом, польза от введения априорной вероятности, с одной стороны, состоит в том, что это дает присоединение необходимой для анализа информации, а с другой стороны, это выигрыш с точки зрения логической ясности.

После введения априорной вероятности в информационно-вероятностный метод и вычисления апостериорного значения условной вероятности p_j перейдем к следующему этапу моделирования, связанному с получением вероятностных оценок проявления j -го признака i -го варианта на формирование оценочного потенциала. С этой целью воспользуемся теоремой Байеса, в которой речь идет об обращении порядка утверждений в условной вероятности, т. е. в принятых нами обозначениях связываются $p_{ji}(r)$ и p_j . Тогда вероятность $p(a)$ в рассматриваемой информационной ситуации определяется зависимостью

$$p_{ji}(a) = \frac{p_{ji}(r)p_j}{\sum_{i=1}^n \sum_{j=1}^m p_{ji}(r)p_j}. \quad (11)$$

Для построения решающего алгоритма воспользуемся аналогией статистических понятий, а именно понятиями ошибок первого и второго рода. *Ошибка первого рода* заключается в непринятии проверяемой гипотезы H_0 , когда она верна. Вероятность совершения такой ошибки $\alpha_0^{\text{эф}}$ называется уровнем значимости. Уровень значимости $\alpha_0^{\text{эф}}$ характеризует риск разработчика новой системы. При этом очевидно, что риск разработчика тем выше, чем меньше степень упорядоченности и организации рассматриваемой системы, характерной для того или иного периода времени. В нашем случае это вероятность оказаться в старой системе, т. е. принять старую систему (параметры новой системы незна-

чимо отличаются от существующей системы — модернизация).

Ошибка второго рода — это принятие проверяемой гипотезы H_0 , когда она неверна. Вероятность совершения ошибки второго рода $\beta_1^{\text{эф}}$ характеризует риск заказчика (потребителя). Отметим, что заказчик рискует не только тогда, когда альтернативный вариант обладает низкой степенью упорядоченности и организации, но и когда он обладает низкой степенью приспособляемости к изменению различных внешних условий. Ошибка второго рода характеризует степень порции энтропии преемственности от предыдущей ситуации. Таким образом, это вероятность того, что прогнозируемые параметры не будут обладать преемственностью от исходных данных начальной системы, т. е. прогнозируемые параметры не будут адаптированы к новым условиям.

В нашем случае ошибка второго рода означает, что дальнейшее прогнозирование не целесообразно, так как возникает большая неопределенность использования системы, т. е. характеристики прогнозируемой системы обладают низкой приспособленностью к существующей среде (не связаны с реальностью).

Измерение степени упорядоченности и степени организации системы осуществляется через количество энтропии. Тогда уровень значимости оценим через функцию неопределенности:

$$H_i(p) = -\sum p_{ji}(a) \log p_{ji}(a).$$

При этом не опровергается та система, которой соответствует меньшее значение величины:

$$\alpha_i^{\text{эф}} = H_{\text{best}}(p) - H_i(p),$$

где $H_{\text{best}}(p)$ — значение энтропии гипотетической прогнозируемой системы, обладающей наилучшими характеристиками для данной информационной ситуации; $H_i(p)$ — значение энтропии для i -го варианта существующей системы.

Известно, что количество накопленной и сохраняемой в структуре систем информации I_i в точности равно уменьшению их энтропии $\Delta H_i(p)$. При этом разность между энтропией $H_i(p)$, объективно существующей в оценке влияния каждой конкретной характеристики на формирование оценочного потенциала, и максимальной энтропией H_{max} в рассмотренной информационной ситуации и есть количество информации I_i , накопленной в данной области исследований:

$$I_i = H_{\text{max}} - H_i(p). \quad (12)$$

Чтобы любая система в процессе своей эволюции не достигла предела "приспособленности" (в результате которого системы существуют только в определенных жестко детерминированных условиях), она должна сохранять в себе непредсказуемость, характеризующую определенной порцией энтро-

пии. Удельный вес этой порции определяется зависимостью

$$G_H^i = H_i(p)/I_i.$$

Следующий шаг конструктивистского метода связан с решением задачи минимизации меры неопределенности. Решение данной задачи обеспечит определение количественного значения удельного веса энтропии.

В результате анализа большого числа работ из различных научных областей доказано существование оптимального значения величины G_H , характеризующей удельный вес порции энтропии [3, 4]. Получено оптимальное значение порции энтропии $G_H^{\text{опт}} = 0,27$, это наилучшее соотношение непредсказуемости и детерминированности. Следовательно, чем больше расчетная величина G_H^i отличается от оптимального значения, тем выше вероятность принятия гипотезы G_0 , когда она не верна. Поэтому ошибка второго рода, определяемая мощностью критерия (удельный вес порции энтропии), представляется зависимостью в форме

$$\beta_i^{\text{эф}} = G_H^i - G_H^{\text{опт}}. \quad (13)$$

Величины $\alpha_0^{\text{эф}}$ и $\beta_0^{\text{эф}}$ количественно характеризуют систему договоренностей, т. е. для принятия гипотезы H_0 необходимо представителям заказчика и разработчика договориться о числовом значении ошибок первого и второго рода. При этом если выполняется неравенство вида

$$\alpha_i^{\text{эф}} \leq \alpha_0^{\text{эф}} \text{ и } \beta_i^{\text{эф}} \leq \beta_0^{\text{эф}},$$

то нулевая гипотеза H_0 (вариант i -й системы) принимается, и процесс прогнозирования заканчивается.

Пример. Применение метода рассмотрено для прогнозирования развития социальной ситуации при решении модельного примера. Базовых нарушителей выберем согласно руководящему документу. Приказом министра промышленности и энергетики РФ от 04.05.2007 № 150 определены следующие типы нарушителей:

- x_1 — внешний нарушитель 1-го типа: террористическая группа 5—12 человек. Целью нарушителя является совершение террористического акта (ТА). Последствия несанкционированного действия нарушителя будут выходить за пределы федеральной, региональной или территориальной зон чрезвычайной ситуации (ЧС);
- x_2 — внешний нарушитель 2-го типа: малочисленная группа (2—4 человека). Целью нарушителя является совершение ТА. Последствия действий нарушителя будут выходить за пределы санитарной зоны объекта;

Таблица 3

Статистика криминальной ситуации

Тип нарушителя	Временная шкала ситуаций (год)				
	2008	2009	2010	2011	2012
x_1	2	3	2	2	3
x_2	4	5	4	5	6
x_3	7	6	8	7	8
x_4	2	3	1	4	3
x_5	3	2	3	4	4
x_6	2	1	3	1	2

- x_3 — внешний нарушитель 3-го типа: одиночный подготовленный нарушитель, не имеющий санкционированного доступа на объект. Данный тип нарушителя действует под принуждением или воздействием психотропных препаратов. Целью нарушителя является ТА;
 - x_4 — внешний нарушитель 4-го типа: одиночный нарушитель, не имеющий санкционированного доступа на объект, имеющий целью хищение материальных ценностей;
 - x_5 — внутренний нарушитель 1-го типа: работник объекта, имеющий санкционированный доступ на объект. Целью нарушителя является хищение ради собственной наживы, однако не исключается возможность совершения ТА;
 - x_6 — внутренний нарушитель 2-го типа: работник охраны объекта. Нарушитель данного типа может осуществить хищение материальных ценностей, а также вступить в сговор с внешним нарушителем 1-го и 2-го типа в целях наживы.
- Имеется, например, статистика по каждому типу угрозы в регионе (федерации) за несколько лет,

которая представлена в табл. 3. Типы нарушителей отсортированы в порядке убывания опасности.

Прогнозирование поведения параметров ситуации на будущие года для табл. 3 осуществлялось на основе метода наименьших квадратов по критерию:

$$\min \sum_{i=0}^n [f(x_i) - P_3(x_i)]^2, \quad (14)$$

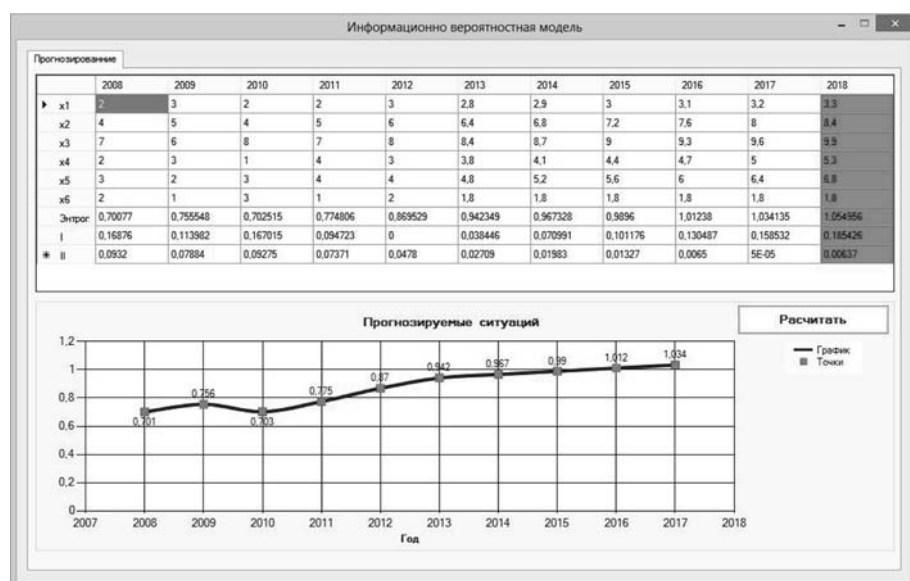
с использованием полинома третьей степени $P_3(x_i)$, так как данный полином описывает статистические результаты с наименьшей ошибкой согласования со статистическими данными $f(x_i)$ табл. 3 [1]. Последователно с помощью аппроксимирующей функции получали прогнозируемые результаты ситуаций и с использованием информационно вероятностного метода [формулы (1)–(3), (8)–(14)] оценивали ошибку первого и второго рода. Задача решается итерационно, пока ошибки не достигнут заданных значений. Процесс расчета автоматизирован при помощи разработанной программы на языке программирования C#, результаты приведены на рисунке.

Заключение

Полученные результаты показывают, что прогнозирование поведения ситуации по времени целесообразно осуществлять на шесть лет. При этом ошибка первого рода составит $\alpha_0 = 0,185$, а ошибка второго рода $\beta_0 = 0,0064$, т. е. G_H^i стремится к наилучшему соотношению непредсказуемости и детерминированности. Характеристики социальной ситуации на 2010 г. приведены на рисунке в последнем столбце и используются как входные данные в методике при задании требований к СФЗ при проектировании [6].

Список литературы

1. Вержбицкий В. М. Основы численных методов: учебник для вузов. М.: Высш. шк., 2002. 840 с.
2. Gutlin Libal Information Theory and the Living System. N. Y.: Columbia University Press, 1972.
3. Мушков А. Ю., Тихомиров В. А., Тихомиров А. В. Модели и методы стратегического управления сложными экономическими и технологическими системами. Тверь: ВУ ПВО, 2003. 244 с.
4. Седов Е. А. Эволюция и информация. М.: Наука, 1976.
5. Фишборн П. С. Теория полезности для принятия решений. М.: Наука 1978.
6. Костин В. Н., Шевченко С. Н. Методика формирования требований к системе физической защиты на основе концептуальной имитационной модели // Инфокоммуникационные технологии. 2013. № 2. С. 91–98.



Прогнозирование развития ситуации

Method of an Assessment of Depth of the Forecast of Development (Evolution) Difficult Systems on the Basis of Entropy Approach

In article it is offered on the basis of the set ratios of entropy (errors of the first and second sort) to define admissible time of forecasting of improvement (development) of parameters of any difficult systems. Results of forecasting can be used for a task of characteristics of perspective systems at their design. Approbation of a method is carried out for an assessment of depth of forecasting of development of a social situation for the purpose of formation of parameters of model of the violator at design of systems of physical protection of potentially dangerous objects.

Keywords: assessment of depth of forecasting of development of difficult systems, entropy, errors of the first and second sort

References

1. Verzhbitskiy V. M. *Osnovy chislennykh metodov: uchebnyk dlya vuzov*. M.: Vyssh. Shk., 2002. 840 p.
2. Gutlin *Libal Information Theory and the Living System*. N. Y.: Columbia University Press, 1972.
3. Mushkov A. Yu., Tikhomirov V. A., Tikhomirov A. V. *Modeli i metody strategicheskogo upravleniya slozhnyimi ekonomicheskimi i tekhnologicheskimi sistemami*. Tver': VU PVO, 2003. 244 p.

4. Sedov E. A. *Evolutsiya i informatsiya*. M.: Nauka, 1976.
5. Fishborn P. S. *Teoriya poleznosti dlya prinyatiya resheniy*. M.: Nauka, 1978.
6. Kostin V. N., Shevchenko S. N. Metodika formirovaniya trebovaniy k sisteme fizicheskoy zashchity na osnove kontseptual'noy imitatsionnoy modeli. *Infokommunikatsionnye tekhnologii*. 2013. N. 2. P. 91–98.

ИНФОРМАЦИЯ



IX Международная научная конференция "Параллельные вычислительные технологии" (ПаВТ 2015)

30 марта — 3 апреля 2015 г.
Уральский федеральный университет
имени первого Президента России Б.Н. Ельцина (г. Екатеринбург)

"Параллельные вычислительные технологии (ПаВТ) 2015" — международная научная конференция, девятая в серии ежегодных конференций, посвященных развитию и применению параллельных вычислительных технологий в различных областях науки и техники. **Главная цель конференции** — предоставить возможность для обсуждения перспектив развития параллельных вычислительных технологий и представления результатов, полученных ведущими научными группами в использовании суперкомпьютерных технологий для решения задач науки и техники в странах СНГ и всего Мира.

Организаторы конференции:

- Российская академия наук
- Суперкомпьютерный консорциум университетов России

Тематика конференции покрывает все аспекты применения высокопроизводительных вычислений в науке и технике, включая приложения, аппаратное и программное обеспечение, специализированные языки и пакеты.

В первый день работы конференции будет объявлена **22-я редакция списка Top50** самых мощных компьютеров СНГ.

Во все дни работы конференции будет действовать суперкомпьютерная выставка, на которой ведущие производители аппаратного и программного обеспечения представят свои новейшие разработки в области высокопроизводительных вычислений.

Языки конференции: русский, английский.

Официальный сайт конференции: <http://agora.guru.ru/pavt2015/>

БЕЗОПАСНОСТЬ ИНФОРМАЦИИ CRYPTOSAFETY INFORMATION

УДК 004.64

С. А. Глушенко, аспирант, e-mail: www.555.sergey@mail.ru,
А. И. Долженко, д-р эконом. наук, доц., проф., e-mail: doljenkoalex@gmail.ru,
Ростовский государственный экономический университет, г. Ростов-на-Дону

Система поддержки принятия решений нечеткого моделирования рисков информационной безопасности организации

Обосновывается целесообразность применения нечеткой логики для оценки риска информационной безопасности организации, а также предъявляются функциональные требования к системе поддержки принятия решений управления рисками. Проводится реализация процесса нечеткого моделирования базы правил посредством разработанной СППР ModelingFuzzySet. Выполнение нечеткого вывода реализуется на основе алгоритма Мамдани (Mamdani).

Ключевые слова: риск, нечеткое множество, терм-множество, нечеткая производственная модель, лингвистическая переменная, база правил, функция принадлежности, система поддержки принятия решений

Введение

Процесс внедрения информационных технологий и средств вычислительной техники в производство и управление современных организаций является эффективным инструментом повышения производительности труда. Однако существующие и проектируемые информационные инфраструктуры организаций и предприятий часто приобретают неструктурированный характер, что способствует неконтролируемому росту уязвимостей и риску информационной безопасности в целом.

Информационная безопасность (ИБ) организации — защищенность информации и поддерживающей инфраструктуры от случайных или преднамеренных воздействий естественного или искусственного характера, которые могут нанести неприемлемый ущерб [1].

В целях обеспечения ИБ организации строится Система обеспечения информационной безопасности (СОИБ). СОИБ предприятия представляет собой совокупность мер организационного и программно-технического уровня, направленных на защиту информационных ресурсов предприятия от потенциальных угроз. Меры защиты организационного уровня реализуются путем проведения соответствующих мероприятий, предусмотренных документированной политикой информационной безопасности. Меры защиты программно-технического уровня реализуются с помощью соответствующих средств и методов защиты информации [2].

Экономический эффект от внедрения СОИБ проявляется в виде снижения возможного матери-

ального, репутационного и иных видов ущерба, наносимого предприятию, за счет использования мер, направленных на формирование и поддержание режима ИБ.

Определить перечень необходимых мер защиты информации, выбрать стратегию развития информационной структуры организации и поддерживать на должном уровне безопасность организации возможно только по результатам аудита уязвимостей предприятия и анализа рисков.

В работе [3] рассматривается риск как фактор, сущность или элемент, представляющий опасность для ИБ организации, величина которой не определена.

Недостатками существующих подходов к анализу и управлению рисками является отсутствие методологической основы для интегрального анализа качественных и количественных факторов рисков [4].

Для устранения вышеперечисленных недостатков предлагается использовать нечеткие модели, которые являются эффективным инструментом, в тех случаях, когда [5]:

- имеется недостаточность или неопределенность знаний об исследуемой системе или процессе;
- получение требуемой информации сопряжено с различными трудностями или вообще невозможно;
- основная часть информации получена на основе экспертных данных или эмпирических описаний процессов;
- параметры и входные данные не являются точными и корректно представленными.

Данные доводы обуславливают целесообразность применения нечетких моделей [5, 6] при оценке состояния системы обеспечения информационной безопасности организации, которые позволяют использовать общеизвестный лингвистический подход. Оценка компонентов СОИБ проводится терминами *средний уровень программно-аппаратной защиты, высокий уровень организационной защиты, низкая рыночная ценность информационного ресурса* и др., т. е. в тех случаях, когда ИТ-менеджерам сложно придать им точную (объективную) количественную оценку и описать с помощью математического языка. При таком подходе характеристики системы целесообразно рассматривать с точки зрения теории нечетких множеств как лингвистические переменные. Применение методов получения оценок рисков на основе нечеткой логики позволяет использовать как количественные характеристики, которым объективно свойственна неопределенность, так и качественные, субъективные оценки экспертов, выраженные нечеткими понятиями, а также формализовать нечеткие описания с помощью нечетких чисел, множеств, лингвистических переменных и нечетких свидетельств [7].

В работах [8, 9] описана нечеткая производственная модель (НПМ), которая позволяет снять ограничения на число учитываемых входных переменных и интегрировать как качественные, так и количественные подходы к оценке рисков. В НПМ определены семь входных лингвистических переменных, характеризующих факторы риска, и четыре выходных лингвистических переменных, характеризующих риски различных областей ИБ. Модель содержит четыре базы правил и позволяет проводить лингвистический анализ рисков, которые несут потенциальные угрозы и ущерб организации, а также выявлять приоритеты рисков (очень высокий, высокий, средний, низкий, очень низкий), которые важны для ИТ-менеджмента информационной структуры предприятия.

На рис. 1 приведена диаграмма вариантов использования (Use Cases), которая охватывает функциональные требования к СППР управления рисками информационной безопасности организации, поддерживающей нечеткие модели.

Лицо принимающее решение (ЛПР), ИТ-менеджер, задействует вариант использования (ВИ) "Формирование модели анализа" при оценке риска, который включает ВИ "Ввод факторов риска" и "Формирование базы правил". Далее посылается сообщение на выполнение ВИ "Задание цели анализа". Цель анализа определяет необходимые входные факторы риска и базу правил, которые должен ввести ИТ-менеджер. База правил и база фактов (рабочая область) определяют возможность выполнения ВИ "Инициировать нечеткий вывод", который вычисляет функцию принадлежности лингвистической переменной, характеризующей показатель риска. Вариант использования "Выполнить оценку риска" выполняется по сигналу от ИТ-менеджера и использует ВИ "Инициировать нечеткий вывод".

На рис. 2 приведена общая схема взаимодействия компонентов СППР.

Система электронной обработки данных (СЭОД) имеет большой набор источников данных, включая внешние источники, об объекте управления, который приводится к единому формату и согласовыва-

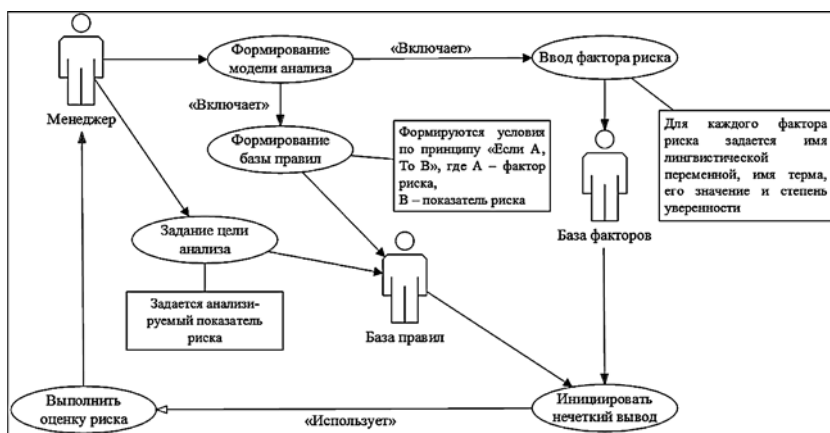


Рис. 1. Диаграмма Use Cases

Постановка задачи для инструментария управления рисками

Автоматизация процесса анализа рисков посредством применения систем поддержки принятия решений (СППР) повышает эффективность работы ИТ-менеджера. Однако существующие программные пакеты анализа и оценки рисков не обладают возможностями построения нечетких производственных сетей и интегрального анализа качественных и количественных факторов риска [3].



Рис. 2. Схема компонентов СППР

Обозначение лингвистических переменных

Обозначение в модели	Наименование лингвистической переменной
ЛП01	Программно-аппаратный уровень защиты
ЛП02	Уровень организационной защиты
ЛП03	Уровень правовой защиты
ЛП04	Риск снижения эффективности защиты
ЛП05	Мотивация источника угроз
ЛП06	Возможности источника угроз
ЛП07	Риск возникновения потенциальных угроз
ЛП08	Рыночная ценность информационного ресурса
ЛП09	Объем данных в информационном ресурсе организации
ЛП10	Риск материального ущерба
ЛП11	Риск ИБ организации

ется до минимально требуемого уровня обобщения для расчета поставленной задачи с помощью математических моделей в подсистеме данных СППР.

Подсистема моделей включает собрание математических и аналитических моделей, которые применяются для моделирования и анализа поставленной задачи.

Система управления интерфейсом обеспечивает простое взаимодействие между пользователем, подсистемой данных и моделями, а также обеспечивает проведение анализа и моделирования проблемы для помощи в принятии управленческих решений [11].

В работе [5] описывается процесс нечеткого моделирования базы правил НПМ оценки рисков ИБ организации посредством применения специализированного пакета Fuzzy Logic Toolbox программного средства MATLAB и реализацией нечеткого вывода на основе алгоритма Мамдани (Mamdani) [12].

Используемый пакет является достаточно универсальным, однако для реализации нечеткой продукционной сети, ориентированной на оценку рисков ИБ, не обеспечивает требуемой функциональности, так как имеет ограничения по числу входных переменных для баз правил и характеризуется ограниченными возможностями по организации проведения интерактивного имитационного эксперимента при оценке рисков по схеме "Если...То...".

Таким образом, вышеприведенные ограничения программных пакетов побудили выполнить собственную разработку системы поддержки принятия решений анализа рисков на базе нечетких продукционных сетей, которая позволит получить как качественные, так и количественные оценки.

Применение разработанной СППР для построения НПМ

Для реализации процесса нечеткого моделирования рисков ИБ организации посредством разработанной СППР ModelingFuzzySet [13] IT-менеджеру необходимо выполнить следующие действия.

Шаг 1. Сформировать нечеткую продукционную модель оценки рисков с помощью дизайнера модели. Формирование модели предполагает задание лингвистических переменных (ЛПхх), формализующих факторы и показатели рисков ИБ, а также базы нечетких продукционных правил (БПхх). IT-менеджер может использовать предложенную НПМ (рис. 3, см. третью сторону обложки), либо провести модификацию модели, определяя факторы и показатели риска, актуальные для конкретной организации, а также изменить правила нечеткого вывода.

Обозначения лингвистических переменных, используемых в сформированной НПМ, представлены в табл. 1.

Шаг 2. Фазификация — введение нечеткости. На этом шаге необходимо задать функции принадлежности для терм-множеств входных и выходных

лингвистических переменных. Программное средство ModelingFuzzySet обеспечивает формирование функций принадлежности терм-множеств различных видов: треугольные, трапециевидные, Z-, S- и П-образные.

Для входной переменной ЛП01 терм-множество состоит из трех термов $T = \{\text{Низкий (Н)}, \text{Средний (С)}, \text{Высокий (В)}\}$, которые характеризуют низкий, средний и высокий уровни программно-аппаратной защиты организации. Функции принадлежности для входной переменной ЛП01 являются трапециевидными. В общем случае трапециевидная функция принадлежности имеет следующий вид:

$$\mu_T(x, a, b, c, d) = \begin{cases} 0, & x \leq a; \\ \frac{x-a}{b-a}, & a \leq x \leq b; \\ 1, & b \leq x \leq c; \\ \frac{d-x}{d-c}, & c \leq x \leq d; \\ 0, & d \leq x, \end{cases} \quad (1)$$

где a, b, c, d — числовые параметры, характеризующие нижнее основание трапеции (a, d) и верхнее (b, c), причем должно выполняться условие $a \leq b \leq c \leq d$.

С учетом формулы (1) функции принадлежности нечетких терм-множеств лингвистической переменной "Программно-аппаратный уровень защиты" будут иметь следующий вид:

$$\mu_{\Delta}^{\text{Н}}(x; 0; 0; 0,15; 0,4), \quad \mu_{\Delta}^{\text{С}}(x; 0,2; 0,4; 0,6; 0,8), \\ \mu_{\Delta}^{\text{В}}(x; 0,6; 0,85; 1,0; 1,0).$$

На рис. 4 (см. третью сторону обложки) приведены графики функций принадлежности терм-множеств лингвистической переменной ЛП01 — "Программно-аппаратный уровень защиты".

Для входной переменной ЛПО2 терм-множество состоит из трех термов $T = \{H, C, B\}$, функции принадлежности которых являются треугольными. В общем случае треугольная функция принадлежности имеет следующий вид:

$$\mu_{\Delta}(x, a, b, c) = \begin{cases} 0, & x \leq a; \\ \frac{x-a}{b-a}, & a \leq x \leq b; \\ \frac{c-x}{c-b}, & b \leq x \leq c; \\ 0, & c \leq x, \end{cases}$$

где a, b, c — числовые параметры, характеризующие основание треугольника (a, c) и его вершину (b), причем должно выполняться условие $a \leq b \leq c$.

С учетом формулы (2) функции принадлежности нечетких терм-множеств лингвистической переменной "Уровень организационной защиты" будут иметь следующий вид:

$$\mu_{\Delta}^H(x; 0; 0; 0,4), \mu_{\Delta}^C(x; 0,2; 0,5; 0,8), \\ \mu_{\Delta}^B(x; 0,6; 1,0; 1,0).$$

На рис. 5 (см. третью сторону обложки) приведены графики функций принадлежности терм-множеств лингвистической переменной ЛПО2 — "Уровень организационной защиты".

Для входной переменной ЛПО3 терм-множество состоит из трех термов $T = \{H, C, B\}$, функции принадлежности которых являются трапециевидными.

С учетом формулы (1) функции принадлежности нечетких терм-множеств лингвистической переменной "Уровень правовой защиты" будут иметь следующий вид:

$$\mu_{\Delta}^H(x; 0; 0; 0,15; 0,45), \mu_{\Delta}^C(x; 0,1; 0,4; 0,6; 0,9), \\ \mu_{\Delta}^B(x; 0,55; 0,85; 1,0; 1,0).$$

На рис. 6 (см. четвертую сторону обложки) приведены графики функций принадлежности терм-множеств лингвистической переменной ЛПО3 — "Уровень правовой защиты".

Для выходной переменной ЛПО4 (лингвистическая переменная "Риск снижения эффективности защиты") терм-множество состоит из пяти термов: $T = \{\text{Очень низкая очевидность риска (ОНОР)}; \text{Низкая очевидность риска (НОР)}; \text{Средняя очевидность риска (СОР)}; \text{Высокая очевидность риска (ВОР)}; \text{Очень высокая очевидность риска (ОВОР)}\}$. Функции принадлежности лингвистических переменных являются трапециевидными.

С учетом формулы (1) функции принадлежности нечетких терм-множеств лингвистической пе-

ременной "Риск снижения эффективности защиты" будут иметь следующий вид:

$$\mu_{\Delta}^{\text{ОНОР}}(x; 0; 0; 0,1; 0,25), \\ \mu_{\Delta}^{\text{НОР}}(x; 0,1; 0,25; 0,35; 0,5), \\ \mu_{\Delta}^{\text{СОР}}(x; 0,3; 0,45; 0,55; 0,7), \\ \mu_{\Delta}^{\text{ВОР}}(x; 0,5; 0,65; 0,75; 0,9), \\ \mu_{\Delta}^{\text{ОВОР}}(x; 0,75; 0,9; 1,0; 1,0).$$

На рис. 7 (см. четвертую сторону обложки) приведены графики функций принадлежности терм-множеств лингвистической переменной ЛПО4 — "Риск снижения эффективности защиты".

Аналогичным образом задаются функции принадлежности для других лингвистических переменных.

Шаг 3. Задание нечетких правил. Правила модели формируются на основе общих закономерностей поведения исследуемой системы и позволяют "вложить" в механизм вывода логическую модель прикладного уровня. С учетом ограничений алгоритма Мамдани модифицируем базу правил БПО1 в вид, приведенный на рис. 8 (см. четвертую сторону обложки). Также в процессе формирования нечетких правил модели ИТ-менеджер задает уровень влияния показателя риска на систему информационной безопасности организации, который используется для оценки возможного ущерба при возникновении данного рискованного события.

В алгоритме Мамдани для агрегирования степени истинности предпосылок используем T -норму и \min -конъюнкцию:

$$T(\mu_A(x), \mu_B(x)) = \min(\mu_A(x), \mu_B(x)).$$

Определение степени истинности заключений по каждому правилу (импликация) основано на операции \min -активизации:

$$\mu_R(x, y) = \min\{\mu_A(x), \mu_B(y)\}.$$

Шаг 4. Аккумуляция заключения по всем правилам проведено с применением операции \max -дизъюнкции. При дефазификации использован метод центра тяжести для дискретного множества значений функций принадлежности:

$$y' = \frac{r=1}{Y_{\max}} \frac{\sum_{r=1}^{Y_{\max}} y_r \mu_{B'}(y_r)}{\sum_{r=1}^{Y_{\max}} \mu_{B'}(y_r)},$$

где Y_{\max} — число элементов y_r в дискретизированной для вычисления "центра тяжести" области Y .

Реализуя систему нечеткого вывода на этапе дефазификации, получим оценку приоритета риска. Приоритезация является основной целью анализа рисков и основополагающим фактором в процессе принятия решений по управлению рисками ИБ организации.

Использование модели оценки рисков ИБ организации

В процессе идентификации рисков ИТ-менеджерами были определены уровни программно-аппаратной, организационной и правовой защиты

Рис. 9. Форма ввода значения лингвистической переменной ЛП01

Рис. 10. Форма ввода значения лингвистической переменной ЛП02

Рис. 11. Форма ввода значения лингвистической переменной ЛП03

паратной, организационной и правовой защиты организации. Данная информация будет использоваться для моделирования показателя риска "Снижение эффективности защиты" в разработанной СППР ModelingFuzzySet.

Значение входной лингвистической переменной ЛП01 — "Программно-аппаратный уровень защиты" соответствует терму В — "полностью соответствует уровню конфиденциальности информации" с уровнем уверенности $\mu_{ЛП01}^В = 0,7\mu$ (рис. 9).

Значение входной лингвистической переменной ЛП02 — "Уровень организационной защиты" соответствует терму Н — "слабое планирование и отсутствие мониторинга уязвимостей" с уровнем уверенности $\mu_{ЛП02}^Н = 0,8\mu$ (рис. 10).

Значение входной лингвистической переменной ЛП03 — "Уровень правовой защиты" соответствует терму С — "документация имеется, но недостаточно детальная" с уровнем уверенности $\mu_{ЛП03}^С = 0,7\mu$ (рис. 11).

По заданным исходным условиям активизируется правило 12 (см. рис. 8), которое определяет значение выходной лингвистической переменной ЛП04 — "Риск снижения эффективности защиты", равное СОР — "Средняя очевидность риска" с уровнем уверенности $\mu_{ЛП04}^{СОР} = 0,7\mu$ и значением показателя риска 50 (рис. 12). Реализация нечеткого вывода определяется на основе алгоритма Мамдани (Mamdani) [14].

Используя результаты моделирования, ИТ-менеджер может определить сумму возможного ущерба по табл. 2.

Таким образом, снижение эффективности защиты предприятия может привести к возможному ущербу в размере \$500 тыс.

Для получения интегральной оценки риска ИБ организации необходимо провести моделирование всех выходных лингвистических переменных первого уровня, затем программное средство рассчитает значение выходной лингвистической переменной ЛП11 — "Риск ИБ организации", используя полученные ранее оценки в качестве входных данных. Результаты нечеткого моделирования выход-

Таблица 2

Шкала оценки последствий возникновения риска

Ранг показателя риска	Значение показателя риска				
	0...20	21...40	41...60	61...80	81...100
1	\$100	\$300	\$500	\$700	\$900
2	\$1 тыс.	\$3 тыс.	\$5 тыс.	\$7 тыс.	\$9 тыс.
3	\$10 тыс.	\$30 тыс.	\$50 тыс.	\$70 тыс.	\$90 тыс.
4	\$100 тыс.	\$300 тыс.	\$500 тыс.	\$700 тыс.	\$900 тыс.
5	\$1 млн	\$3 млн	\$5 млн	\$7 млн	\$10 млн



Рис. 12. Результат нечеткого моделирования ЛП04

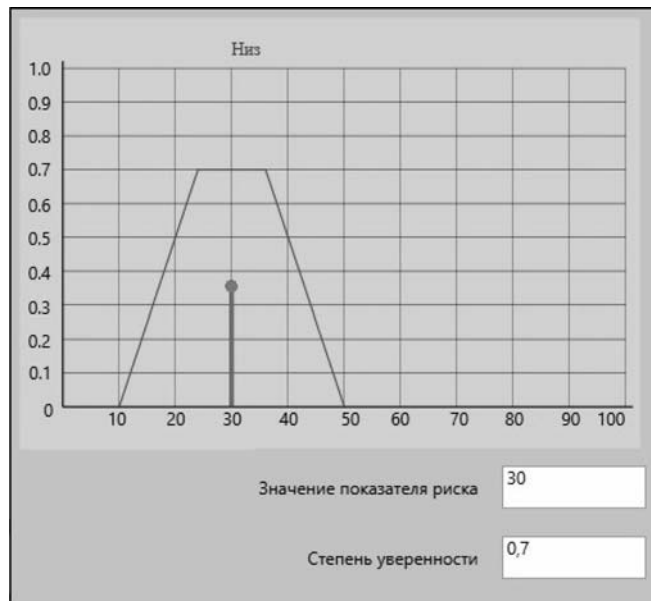


Рис. 13. Результат нечеткого моделирования интегральной ЛП11

ной лингвистической переменной ЛП07 — "Риск возникновения потенциальных угроз" показали НОР — "Высокая очевидность риска" с уровнем уверенности $\mu_{\text{ЛП07}} = 0,8\mu$ и значением показателя риска 70. Результаты нечеткого моделирования выходной лингвистической переменной ЛП10 — "Риск материального ущерба" показали НОР — "Низкая очевидность риска" с уровнем уверенности $\mu_{\text{ЛП10}}^{\text{НОР}} = 1,0\mu$ и значением показателя риска 30. Результаты нечеткого моделирования интегральной выходной лингвистической переменной ЛП11 — "Риск ИБ организации" показали НОР — "Низкая очевидность риска" с уровнем уверенности $\mu_{\text{ЛП11}}^{\text{НОР}} = 0,7\mu$ и значением показателя риска 30 (рис. 13).

Таким образом, оценка показывает, что уровень риска ИБ организации позволяет работать ИТ-менеджерам в нормальном режиме, но имеются предпосылки нарушения системы обеспечения информационной безопасности (СОИБ). Поэтому необходимо разработать и применить план по поиску и устранению уязвимостей СОИБ, иначе ущерб организации может составить около \$3 млн.

Заключение

Разработанная система поддержки принятия решений позволяет строить многоуровневые нечеткие производственные модели оценки рисков, а используемый механизм нечеткого вывода позволяет получить числовое значение риска, лингвистическое описание степени риска, а также степень уверенности эксперта в возникновении рисковом со-

бытия. Полученные данные помогают ИТ-менеджерам выявить приоритеты рисков (очень высокий, высокий, средний, низкий, очень низкий) и выработать план мероприятий по снижению влияния наиболее опасных угроз на ИБ организации.

Механизм получения оценок риска на основе нечеткой логики обладает широкими возможностями и позволяет адаптировать его к имеющимся на предприятии моделям управления рисками, а также модифицировать с учетом реальных условий политики информационной безопасности организации.

Список литературы

1. **Рекомендации** по стандартизации "Информационные технологии. Основные термины и определения в области технической защиты информации" (Р 50.1.053—2005).
2. **Национальный стандарт РФ** "Информационная технология. Практические правила управления информационной безопасностью" (ГОСТ Р ИСО/МЭК 17799—2005).
3. **Симонов С. В.** Анализ рисков, управление рисками // Информационный бюллетень "Jet Info". 1999. № 1 (68).
4. **Долженко А. И., Глушенко С. А.** Анализ предпосылок создания СППР управления рисками проектов информационных систем // Информационные системы, экономика, управление трудом и производством. Ученые записки. Вып. 12. Ростов-на-Дону: РГЭУ (РИНХ). 2011. № 12.
5. **Борисов В. В., Круглов А. С. и Федулов А. С.** Нечеткие модели и сети. М.: Горячая линия — Телеком, 2007. 284 с.
6. **Заде Л. А.** Понятие лингвистической переменной и его применение к принятию приближенных решений. М.: Мир, 1976. 168 с.
7. **Долженко А. И.** Модель анализа риска потребительского качества проектов экономических информационных систем // Вестник Северо-Кавказского государственного технического университета. 2009. № 1 (18). С. 129—134.
8. **Глушенко С. А.** Нечеткая производственная модель оценки рисков информационной безопасности организации // Вопросы экономики и права: Сборник статей аспирантов и соискателей.

телей ученой степени кандидата наук. Ростов-на-Дону: РГЭУ (РИНХ), 2013.

9. **Глушенко С. А.** Применение системы matlab для оценки рисков информационной безопасности организации // БИЗНЕС-ИНФОРМАТИКА. 2013. № 4 (24).

10. **Глушенко С. А.** Анализ функциональной полноты программных систем управления рисками // Вестник РГЭУ (РИНХ). 2012. № 2 (38).

11. **Трахтенгерц Э. А.** Компьютерная поддержка принятия решений: научно-практическое изд. Сер. "Информатизация России на пороге XXI века". М.: СИНТЕГ, 1998. 376 с.

12. **Леоненков А. В.** Нечеткое моделирование в среде MATLAB и fuzzyTECH. СПб.: БХВ-Петербург, 2005. 736 с.

13. **Долженко А. И.** Система моделирования производственной нечеткой сети (ПРОНЕС) № 2010612952 / Долженко А. И., Глушенко С. А., Калугян К. Х., Лозина Е. Н., Чередниченко А. С. Свидетельство о государственной регистрации программы для ЭВМ. М., 25.05.2010 г.

14. **Анисимова Г. Б., Долженко А. И.** Система оценки риска потребительского качества проектов информационных систем // Вестник ИНЖЭКОНА. Сер. "Экономика". 2010. № 1 (36). С. 179—191.

S. A. Glushenko, Graduate Student, e-mail: www.555.sergey@mail.ru,
A. I. Doljenko, Professor, e-mail: doljenkoalex@gmail.ru,
Rostov State Economic University (RINE), Russia, Rostov-on-Don

Decision Support Fuzzy Modeling System of Information Security Risks

The article explains the importance of the application of risk analysis in the management system of information security (IS) and explains the usefulness of fuzzy logic to evaluate the risk. Application of fuzzy models allows to take into account both quantitative and qualitative characteristics, as well as represent fuzzy description using fuzzy sets and linguistic variables.

In the article, the problem statement for risk management tools to support fuzzy models and expediency proprietary decision support system (DSS) risk analysis. The following describes the process of implementing the fuzzy rule base modeling developed by DSS ModelingFuzzySet. Mechanism for obtaining risk assessments based on the algorithm of Mamdani allows obtaining the numerical value of the risk, the linguistic description of the risk and the degree of confidence in expert risk event occurs.

IT-managers used the simulation results to identify priority risks (very high, high, medium, low, very low) and allowed to develop effective action plan to reduce the impact of the most dangerous threats to the information security organization.

Keywords: risk, fuzzy set, term set, fuzzy production model, the linguistic variable, rule base, the membership function, decision support system

References

1. **Рекомендации** по стандартизации "Информационные технологии. Основные термины и определения в области технической зашхиты информации" (R 50.1.053—2005).

2. **Национальный** стандарт RF "Информационная технология. Практические правила управления информационной безопасностью" (GOST R ISO/MJeK 17799—2005).

3. **Simonov S. V.** Анализ рисков, управление рисками. *Информационный бюллетень "Jet Info"*. 1999. N. 1 (68).

4. **Dolzhenko A. I., Glushenko S. A.** Анализ предпосылок создания СПР управления рисками проектов информационных систем. *Информационные системы, экономика, управление трудом и производством. Ученые записки*. Вып. 12. Ростов-на-Дону: РГЭУ (РИНХ), 2011. N. 12.

5. **Borisov V. V., Kruglov A. S. i Fedulov A. S.** *Nechetkie modeli i seti*. М.: Gorjachaja linija-Telekom, 2007. 284 p.

6. **Zade L. A.** *Ponjatije lingvisticheskoy peremennoy i ego primenenije k prinjatiju priblizhennyh reshenij*. М.: Mir, 1976. 168 p.

7. **Dolzhenko A. I.** Model' analiza riska potrebitel'skogo kachestva proektov jekonomicheskikh informacionnyh sistem. *Vestnik Severo-Kavkazskogo gosudarstvennogo tehničeskogo universiteta*. 2009. N. 1 (18). P. 129—134.

8. **Glushenko S. A.** Nechetkaja produkcionnaja model' ocenki riskov informacionnoj bezopasnosti organizacii. *Voprosy jekonomiki i prava: Sbornik statej aspirantov i soiskatelej uchenoj stepeni kandidata nauk*. Ростов-на-Дону: РГЭУ(РИНХ), 2013.

9. **Glushenko S. A.** Применение системы matlab длia оценки рисков информационной безопасности организации. *BIZNES—INFORMATIKA*. 2013. N. 4 (24).

10. **Glushenko S. A.** Анализ функциональной полноты программных систем управления рисками. *Vestnik RGEU (RINH)*. 2012. N. 2 (38).

11. **Trahtengerc Je. A.** *Komp'yuternaja podderzhka prinjatija reshenij: Nauchno-prakticheskoe izd. Ser. "Informatizacija Rossii na poroge XXI veka"*. М.: SINTEG, 1998. 376 p.

12. **Leonenkov A. V.** *Nechetkoe modelirovanie v srede MATLAB i fuzzyTECH*. СПб.: BHV-Peterburg, 2005. 736 p.

13. **Dolzhenko A. I.** *Sistema modelirovanija produkcionnoj nechetkoj seti (PRONES)* — № 2010612952. Dolzhenko A. I., Glushenko S. A., Kalugjan, K. H., Lozina, E. N., Cherednichenko A. S. Svidetel'stvo O gosudarstvennoj registracii programmy dlja JeVM. М., 25.05.2010 g.

14. **Anisimova G. B., Dolzhenko A. I.** Sistema ocenki riska potrebitel'skogo kachestva proektov informacionnyh sistem. *Vestnik INZhEKONA*. Ser. "Jekonomika". 2010. N. 1 (36). P. 179—191.

Д. Л. Жусов, канд. техн. наук, сотрудник, e-mail: d.zhusov@mail.ru,
А. В. Козленко, канд. техн. наук, сотрудник, Р. В. Шульга, сотрудник,
Академия ФСО России, г. Орел

Алгоритм анализа глубины гипертекстовых переходов официальных сайтов федеральных органов исполнительной власти Российской Федерации

Предложен алгоритм анализа глубины гипертекстовых переходов официальных сайтов федеральных органов исполнительной власти России, позволяющий повысить оперативность контроля их навигационных средств.

Ключевые слова: web-сайт, гипертекстовый переход, глубина гипертекстовых переходов

Введение

В настоящее время необходимость совершенствования механизмов государственного управления и бурный рост числа пользователей сети Интернет обусловили широкое применение современных информационных технологий для организации взаимодействия государства и общества.

В целях обеспечения доступа к информации о деятельности федеральных органов исполнительной власти (ФОИВ) РФ наиболее широкое применение получили официальные web-сайты. Важность

и необходимость их качественного функционирования подчеркивается разработкой нормативных документов [1, 2], регламентирующих их работу. В целях обеспечения быстрого доступа к информации на официальном сайте ФОИВ РФ его навигационные средства должны соответствовать ряду требований, одним из которых является ограничение на максимальное число последовательных гипертекстовых переходов, начиная с главной страницы сайта. При этом число таких переходов по кратчайшей последовательности должно быть не более пяти [2].

Данный факт определяет необходимость контроля соответствия официальных сайтов ФОИВ РФ требованиям руководящих документов.

Анализ соответствия официальных web-сайтов ФОИВ руководящим документам

С января по февраль 2013 г. Фонд Свободы Информации провел ежегодный мониторинг официальных сайтов ФОИВ РФ [3]. Его результаты (рис. 1) свидетельствуют об отсутствии единого подхода к формированию структуры официальных web-сайтов ФОИВ РФ, а также о не выполнении требований к ним в полном объеме. Так, глубина вложения информации на сайтах не всегда удовлетворяет требуемому ограничению, а существующие способы их проверки достаточно трудоемки.

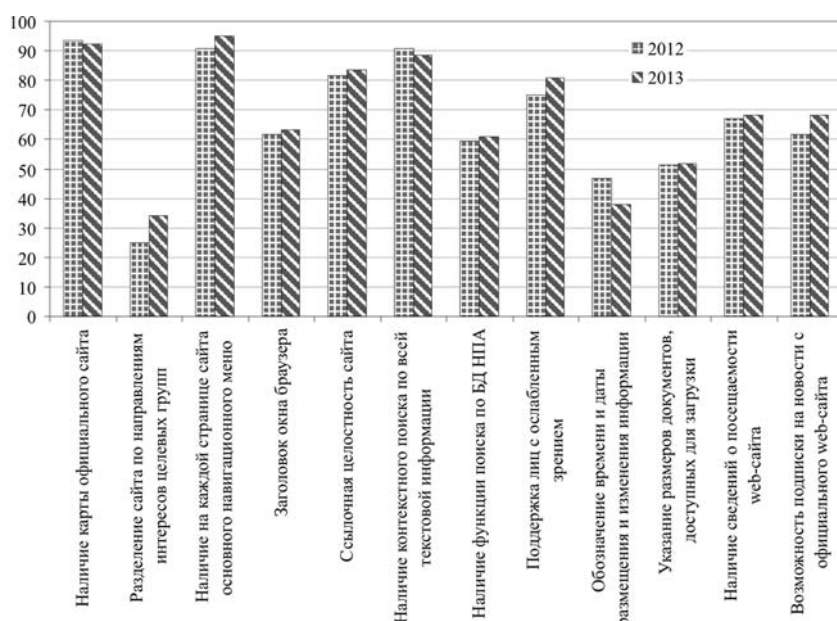


Рис. 1. Статистика соответствия официальных сайтов ФОИВ РФ требованиям руководящих документов

Следовательно, представляется необходимой разработка алгоритма анализа глубины гипертекстовых переходов web-сайтов.

Структурное представление web-сайтов

Структура каждого web-сайта может быть представлена деревом $G = (V, E)$, где V — множество вершин дерева (web-страниц), а E — множество его ребер (гипертекстовых переходов по сайту — гиперссылок) [4], а для расчета глубины гипертекстовых переходов целесообразно применять метод обхода вершин графа "поиск в ширину" [5].

Пусть имеется web-сайт, представленный в виде графа (рис. 2).

Последовательными переходами p_{01}, p_{02}, p_{03} по гиперссылкам со стартовой страницы осуществляется загрузка страниц web-сайта первого уровня: $www.test.ru/1, www.test.ru/2, www.test.ru/3$. Затем последовательно со страницы $www.test.ru/1$ переходами p_{14} и p_{15} по гиперссылкам загружаются страницы $www.test.ru/4, www.test.ru/5$; со страницы $www.test.ru/2$ переходом p_{28} загружается $www.test.ru/8$; а со страницы $www.test.ru/3$ переходами p_{36} и p_{37} соответственно $www.test.ru/6$ и $www.test.ru/7$. Загруженные страницы составляют второй уровень дерева web-сайта. Отсутствие гипертекстовых переходов со страниц второго уровня определяет остановку обхода вершин и фиксирует дерево web-сайта в виде рис. 3.

Рассмотренный метод обхода вершин дерева позволил разработать алгоритм анализа глубины гипертекстовых переходов web-сайтов (рис. 4). Стоит отметить, что объектами анализа могут являться любые web-сайты, по отношению к которым целесообразна формулировка требований по глубине гипертекстовых переходов. Возможным множеством сайтов являются официальные web-сайты ФОИВ РФ.

Основу разработанного алгоритма составляет набор рекурсивных процедур, представленный блоками 3—15. Его сложность определяется поиском в коде загруженной web-страницы тегов $\langle a \rangle$ с атрибутом *href* и числом страниц анализируемого web-сайта и вычисляется как $O(\eta) \cdot O(V + E)$, где η — число символов в коде web-страниц.

Рассмотренный алгоритм программно реализован на платформе JavaSE в виде отдельного модуля, взаимодействующего через JDBC-интерфейс с базой данных хранения результатов, функционирующей под управлением СУБД MySQL. С его помощью было проведено экспериментальное определение максимальной глубины гиперссылок как для локально расположенных сайтов, так и для размещенных в сети Интернет. Ввиду отсутствия доступа к локальной сети,

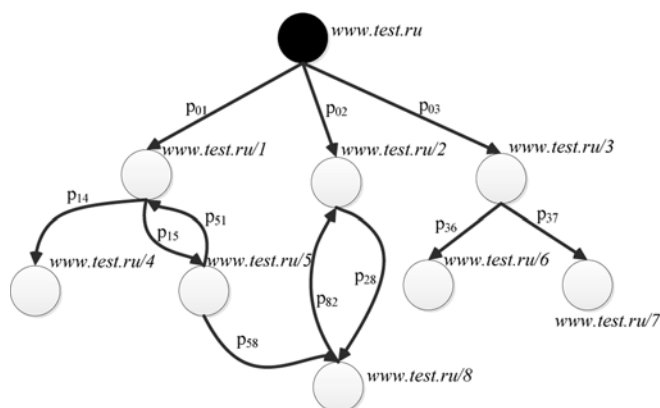


Рис. 2. Графовое представление web-сайта

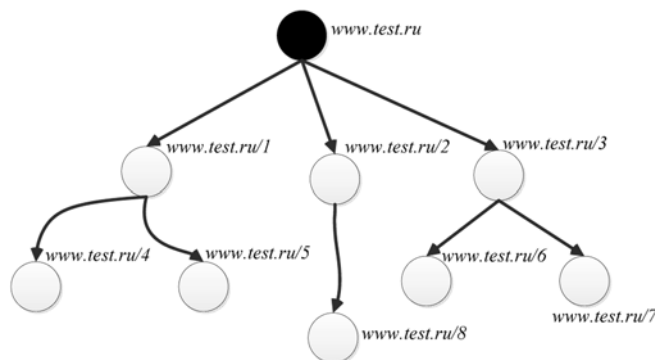


Рис. 3. Дерево web-сайта, построенное методом поиска "в ширину"

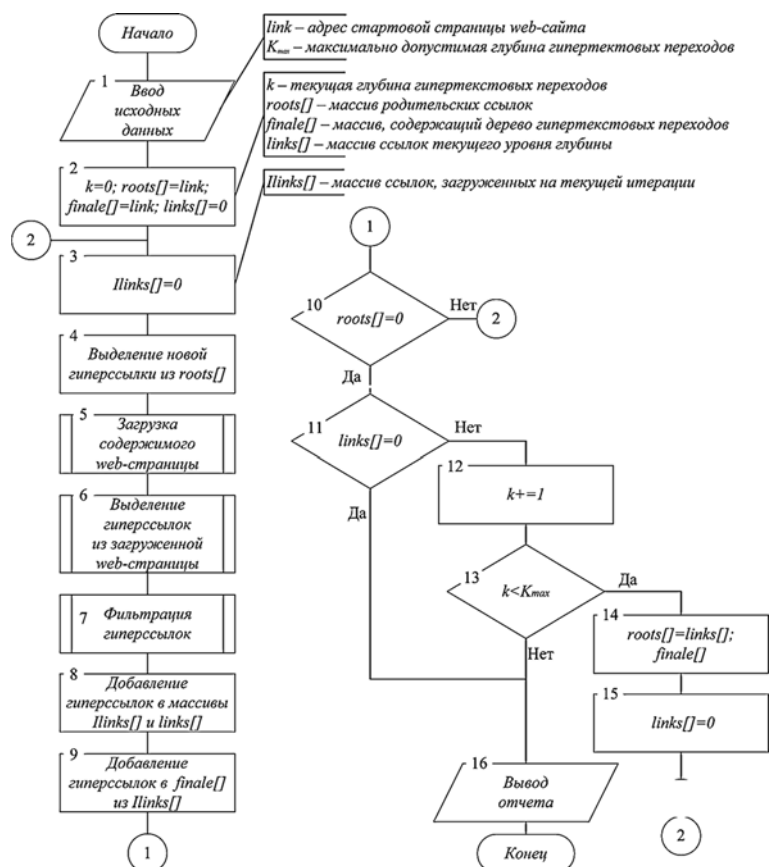


Рис. 4. Алгоритм анализа глубины гипертекстовых переходов официальных сайтов ФОИВ РФ

в которой размещены web-серверы с сайтами ФОИВ, для проведения фрагмента эксперимента web-сайты были предварительно полностью скачаны с помощью программного продукта Teleport Ultra 1.70. Для этих реализаций web-сайтов, размещенных на экспериментальной ПЭВМ с конфигурацией — AMD Athlon FX4000 8 ГБ ОЗУ — были получены значения времени анализа глубины гиперссылок (табл. 1).

Для реализации эксперимента на web-сайтах, размещенных в сети Интернет, использовалось подключение со скоростью 40 Мбит/с. Однако в процессе эксперимента скорость доступа к анализируемым web-страницам (их загрузки в программе) колебалась в пределах 0,1...1 Мбайт/с (табл. 2). Наиболее очевидными причинами этого можно считать предопределенная настройками web-серверов скорость обработки запросов и максимальное возможное число одновременных соединений с одного IP-адреса. Кроме того, при проведении эксперимента невозможно было ограничить нагрузку на web-серверы со стороны других пользователей сети Интернет.

Результаты свидетельствуют о том, что на время анализа глубины гипертекстовых переходов кроме объема web-сайта (число web-страниц) влияют скорость Интернет-подключения и настройки web-сервера, на котором размещен web-сайт ФОИВ (табл. 1, 2, рис. 5).

Обобщая данные таблиц 1, 2, можно представить график зависимости времени анализа от числа страниц web-сайта (см. рис. 5).

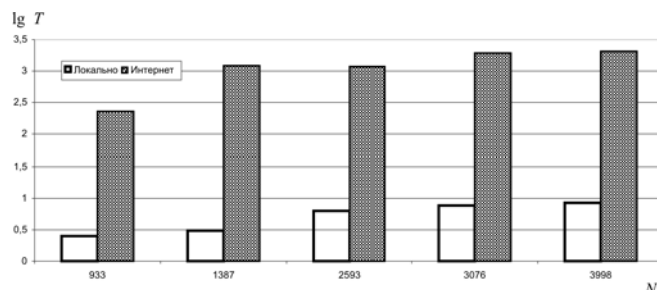


Рис. 5. Результаты экспериментальных исследований

Из рисунка видно, что время анализа глубины гиперссылок разработанным программным средством существенно меньше времени доступа к сети Интернет. Это позволяет говорить о возможности повышения оперативности контроля навигационных средств web-сайтов с помощью программной реализации представленного алгоритма при наличии высокоскоростного подключения и функционировании web-серверов в условиях непредельной загрузки.

Заключение

Программная реализация представленного в работе алгоритма анализа гипертекстовых переходов web-сайтов может эффективно использоваться для контроля требований нормативных документов. При этом объектами анализа могут являться не только официальные web-сайты ФОИВ РФ, но и сайты других государственных органов, организаций и учреждений.

Практическая значимость работы определяется возможностью применения представленного алгоритма для решения задачи контроля функционирования web-сайтов различного назначения.

Направлением дальнейших исследований является интеграция программной реализации алгоритма с модулем сбора и обработки данных о посещаемости web-сайтов с различных серверов сбора статистики, например, Liveinternet, Mail, Rambler и т. п.

Список литературы

1. **Федеральный закон** № 8-ФЗ "Об обеспечении доступа к информации о деятельности государственных органов и органов местного самоуправления" от 09 февраля 2009 г.
2. **Приказ** Министерства экономического развития РФ от 16 ноября 2009 г. № 470 "О требованиях к технологическим, программным и лингвистическим средствам обеспечения пользования официальными сайтами федеральных органов исполнительной власти".
3. **Результаты мониторинга официальных сайтов федеральных органов исполнительной власти**. 2013. URL: <http://svobodainfo.org/ru/node/2527>
4. **Райгородский А. М.** Модели случайных графов. М.: МЦНМО, 2011. 136 с.
5. **Кормен Т. Х., Лейзерсон Ч. И., Ривест Р. П., Штайн К.** Алгоритмы: построение и анализ. 2-е изд. М.: Вильямс, 2006. 1296 с.

Таблица 1

Время анализа глубины гиперссылок локально расположенных web-сайтов

web-сайт	Число web-страниц, N	Глубина гиперссылок, K	Время анализа, T, с	lg T
minsvyaz.ru	933	3	2,5	0,398
minfin.ru	2593	4	6,3	0,799
mintrans.ru	3998	6	8,4	0,924
rosminzdrav.ru	1387	3	3	0,477
minregion.ru	3076	5	7,6	0,88

Таблица 2

Время анализа глубины гиперссылок web-сайтов, расположенных в сети Интернет

web-сайт	Число web-страниц, N	Глубина гиперссылок, K	Время анализа, T, мин	lg T	Скорость доступа к web-сайту, Мбайт/с
minsvyaz.ru	933	3	3,87	2,366	1
minfin.ru	2593	4	19,5	3,068	1
mintrans.ru	3998	6	34,4	3,315	0,5
rosminzdrav.ru	1387	3	20,1	3,081	0,1
minregion.ru	3076	5	32,55	3,291	0,8

Hypertext Navigation Depth Analysis Algorithm for Federal Executive Authorities Official Websites in the Russian Federation

In the article the results of researches on scope of Russian Federation federal executive authorities official websites monitoring is considered. The purpose of the article was an increasing the hypertext navigation depth analysis efficiency for the websites. Achievement of the formulated purpose is realized by the solution of the hypertext navigation depth analysis algorithm development task, which considering features of websites creation. The method of graph peaks bypass "in width" and set of the recursive procedures on hyperlinks separation on each web page of the site is its cornerstone. The computing complexity assessment for developed algorithm and platform option for its program implementation are offered. The experimental results contains websites time analysis estimates depending on number of the web pages included in them and the maximum hypertext navigation depth on the shortest way. They can be an evidence of a possibility for websites monitoring efficiency increasing taking into account requirements of the guidelines. It is expedient to apply it to increase of developed algorithm application efficiency at the moments of Web-server loading limit absence in the presence of high speed connection to the Internet.

The further researches direction on number of websites visits data processing on the basis of these generally available servers of statistics is formulated.

Keywords: website, hypertext navigation, hypertext navigation depth

References

1. **Federal'nyj** zakon № 8-FZ "Ob obespechenii dostupa k informacii o dejatel'nosti gosudarstvennyh organov i organov mestnogo samoupravlenija" ot 09 fevralja 2009 g.
2. **Prikaz** Ministerstva jekonomicheskogo razvitija RF ot 16 nojabrja 2009 g. N. 470 "O trebovanijah k tehnologicheskim, programmym i lingvisticheskim sredstvam obespechenija pol'zovanija oficial'nymi sajтами federai'nyh organov ispolnitel'noj vlasti".

3. **Rezul'taty** monitoringa oficial'nyh sajtov federal'nyh organov ispolnitel'noj vlasti. 2013. URL: <http://svobodainfo.org/ru/node/2527>

4. **Rajgorodskij A. M.** Modeli sluchainyh grafov. M.: MCNMO, 2011. 136 p.

5. **Kormen T. H., Lejzerson C. I., Rivest R. P., Shtajn K.** *Algoritmy: postroenie i analiz.* 2-e izd. M.: Vil'jams, 2006. 1296 p.

Уважаемые авторы!

Сообщаем Вам, что в связи с подготовкой журнала "Информационные технологии" к включению в международные наукометрические базы данных редакция скорректировала требования к статьям и к сопровождающим их материалам, направляемым в адрес редакции. Просьба детально ознакомиться с новыми правилами и строго руководствоваться ими.

ПРАВИЛА ДЛЯ АВТОРОВ

Материалы статьи должны соответствовать тематике журнала и содержать новые результаты исследований, не опубликованные ранее и не предназначенные к публикации в других печатных или электронных изданиях.

Для опубликования статьи в редакцию журнала направляются следующие материалы:

- рукопись статьи в doc- и pdf-форматах;
- таблицы, иллюстрации и перечень подрисуночных подписей;
- сведения об авторах, содержащие фамилию, имя, отчество, ученые степень и звание, должность, место работы, служебный и домашний адреса, телефоны и e-mail;
- экспертное заключение о возможности публикации статьи в открытой печати;
- англоязычная информация, содержащая данные статьи, необходимые для индексирования журнала в международных наукометрических базах данных.

Если статья высылается почтовой бандеролью, то в нее необходимо вложить диск с электронной версией перечисленных материалов.

Все статьи проходят обязательное рецензирование. Просьба редакции о доработке рукописи не означает, что статья принята к печати, так как она вновь будет рассматриваться рецензентами, а затем редакцией. Редакция оставляет за собой право не возвращать автору отклоненную статью и не вести дискуссию по мотивам отклонения. Редакция оставляет за собой право при необходимости сокращать статьи, редактировать и отсылать авторам на доработку.

ТРЕБОВАНИЯ К СОДЕРЖАНИЮ СТАТЬИ

Статья должна быть структурирована, т. е. должна включать *разделы* с названиями, кратко и точно отражающими их содержание, в том числе:

- введение, содержащее обоснование актуальности и краткий обзор проблематики;
- четкую постановку задачи исследования;
- описание метода решения задачи исследования;
- прикладную интерпретацию и иллюстрацию полученных результатов исследования;
- заключение, включающее обобщение и указание области применения полученных результатов, не повторяющее аннотацию и не ограничивающееся простым перечислением того, что сделано в работе.

ПРАВИЛА ПОДГОТОВКИ РУКОПИСИ СТАТЬИ

Объем рукописи статьи, предлагаемой к публикации, должен быть не менее 10 и не более 20 страниц машинописного текста, напечатанного на одной стороне белого листа бумаги формата А4 с полями со всех сторон не менее 2 см, с абзацным отступом 1 см, с полуторным межстрочным интервалом, с использованием текстового редактора Microsoft Word (любая версия) с шрифтом Times New Roman размером 14 pt.

В указанный объем статьи входят: текст, приложения, иллюстрации, таблицы, список литературы. В отдельных случаях по решению редколлегии объем статьи может быть увеличен. Страницы рукописи должны быть пронумерованы, начиная с первой.

Материал статьи должен быть изложен в следующей последовательности:

- *шифр УДК* в левом верхнем углу;
- *инициалы и фамилия автора* (полужирным шрифтом по центру), ученое звание, должность, e-mail;
- *название организации* (без сокращений), город (если организаций больше, чем одна, то указываются все организации со сносками после фамилий соответствующих авторов; в случае совпадения места работы авторов нужно указывать его один раз для всей группы авторов);
- *название статьи* (полужирным шрифтом по центру) не более 12 слов;
- *аннотация* (объемом 5—7 строчек), в предельно сжатой форме излагающая суть работы и полученные результаты;
- *ключевые слова* (8—10 слов), несущие в тексте основную смысловую нагрузку;
- *текст статьи* с учетом указанных далее требований к его оформлению;
- *список литературы*, оформленный по ГОСТ Р 7.0.5—2008;
- *англоязычная информация по статье*.

В тексте статьи следует придерживаться общепринятой терминологии, а все используемые специальные термины, обозначения и аббревиатуры должны быть раскрыты и разъяснены.

Формулы должны быть набраны четко с соблюдением различий в написании прописных и строчных букв, а также букв, сходных по начертанию. Индексы и показатели степени следует писать строго ниже или выше строки. В формулах русские и греческие буквы следует набирать прямо, а латинские буквы, обо-

значающие скалярные переменные, — курсивом. Величины, обозначающие векторы и матрицы, должны быть выделены полужирным шрифтом и набраны прямо (допускается также набор всех величин, обозначенных латинскими буквами, в том числе матриц и векторов, светлым курсивом); знак транспонирования — буквой "t" строчной прямой; экспонента — буквой "e" полужирной прямой. Стандартные математические обозначения (например, \max , \log , \sin и т. д.) должны быть набраны прямо. Номера формул пишутся справа в круглых скобках. Нумерация формул сквозная и дается в порядке их появления, причем нумеруются только формулы, на которые имеются ссылки в тексте.

Список литературы имеет сквозную нумерацию в порядке упоминания в тексте и содержит следующие данные: фамилия и инициалы автора, название работы, название журнала, сборника (если это не монография), город и полное название издательства (для монографии), год издания, том, номер журнала и страницы. Ссылки на литературные источники даются в квадратных скобках. Допускаются ссылки на электронные носители.

Рисунки и таблицы должны быть выполнены качественно (графическая обработка рисунков в редакции не предполагается). В журнале все рисунки воспроизводятся в черно-белом варианте, за исключением цветных рисунков, размещаемых по усмотрению редакции на обложке журнала.

При подготовке рисунков в электронном виде просим соблюдать следующие требования к растровым изображениям:

TIFF (PhotoShop): — полутоновые цветные — 300 dpi, CMYK;
— полутоновые черно-белые — 300 dpi;
— штриховые иллюстрации — 600 dpi.

Толщина линий должна быть не менее 0,15 мм.

Масштаб надписей должен быть соразмерным с самим рисунком.

Рисунки и таблицы дублируются на отдельных листах, подрисовочные подписи даются отдельным списком. Обозначения на рисунках должны строго соответствовать обозначениям, принятым в тексте, и иметь идентичное написание. Надписка на рисунках должна быть достаточно крупной, чтобы сохранять четкость при уменьшении масштаба рисунка. Не допускаются особо мелкие надписки на рисунках, включая обозначения и оцифровку осей на графиках. Цифровые выноски на рисунках должны быть выполнены курсивом; единицы измерения на осях графиков пишутся через запятую и на русском языке.

Англоязычная информация должна включать (согласно требованиям зарубежных аналитических баз данных) следующие данные:

- инициалы и фамилии каждого автора, e-mail, место работы с указанием индекса и города (при совпадении места работы авторов нужно указывать его один раз для всей группы авторов);
- название статьи; аннотация статьи объемом не менее 220 слов, написанная качественным английским языком, отражающая основное содержание статьи с указанием цели, задачи, результатов исследования и кратких выводов;
- ключевые слова (10–12) слов;
- перевод списка литературы на английский язык.

Редакция не предоставляет авторских экземпляров журнала и может выслать авторам электронный вариант статьи в pdf-формате. Если авторы хотят гарантированно получить журнал со своей статьей, то им необходимо оформить подписку на этот номер.

Публикация в журнале бесплатная.

Адрес редакции:

107076, Москва, Стромынский пер., 4

Телефон редакции журнала (499) 269-5510

E-mail: it@novtex.ru

Технический редактор *Е. В. Конова*.

Корректор *Т. В. Пчелкина*.

Сдано в набор 10.11.2014. Подписано в печать 22.12.2014. Формат 60×88 1/8. Бумага офсетная.

Усл. печ. л. 8,86. Заказ ГТ115. Цена договорная.

Журнал зарегистрирован в Министерстве Российской Федерации по делам печати, телерадиовещания и средств массовых коммуникаций.

Свидетельство о регистрации ПИ № 77-15565 от 02 июня 2003 г.

Оригинал-макет ООО "Авансед солюшнз". Отпечатано в ООО "Авансед солюшнз".

119071, г. Москва, Ленинский пр-т, д. 19, стр. 1.