

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

2(210)
2014

ТЕОРЕТИЧЕСКИЙ И ПРИКЛАДНОЙ НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

Издается с ноября 1995 г.

УЧРЕДИТЕЛЬ
Издательство "Новые технологии"

СОДЕРЖАНИЕ

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ

- Кухаренко Б. Г., Солнцева М. О.** Использование методов сокращения фона при сегментировании телеметрических изображений для идентификации групп объектов 3
- Савченко Л. В.** Автоматическое распознавание изолированных слов на основе теории нечетких множеств и кластерной модели минимальных речевых единиц . 9
- Мочалов И. А., Хрисат М. С.** Оценивание параметров модели по нечетким случайным данным 14

СЕТИ И СИСТЕМЫ СВЯЗИ

- Балонин Н. А., Сергеев М. Б.** Современные сетевые технологии в научных исследованиях 23

МОДЕЛИРОВАНИЕ И ОПТИМИЗАЦИЯ

- Левин В. И.** Устойчивость решения задач оптимизации в условиях неопределенности 27
- Орехов Э. Ю.** Равновероятная генерация целочисленных задач прямоугольного раскроя-упаковки 33
- Гаврюшин С. С., Евгеньев Г. Б.** Многокритериальная оптимизация в жизненном цикле изделий 37
- Амосова Е. С.** Разработка онтологической модели предметной области "Инновационные разработки в научных исследованиях" 42

СТРУКТУРНЫЙ СИНТЕЗ И ПРИНЯТИЕ РЕШЕНИЙ

- Николаева М. А., Зотова О. Ф., Шолохова Н. В.** Система поддержки принятия решений при управлении депозитным портфелем физических лиц коммерческого банка 46

ГЕОИНФОРМАЦИОННЫЕ СИСТЕМЫ

- Потапов В. П., Попов С. Е., Замараев Р. Ю.** ГИС-сервис для расчета интенсивности землетрясений на базе технологии Google App Engine 52
- Агаев Ф. Г., Джавадов Н. Г., Асадов Х. Г., Халафов Р. В.** Новый способ оптимизации дистанционного зондирования неоднородных участков природных и техногенных сред 56
- Черняев А. В., Павлов А. А.** Географическая информационная система малого речного бассейна и особенности ее формирования в задачах моделирования нефтяных загрязнений 60

АВТОМАТИЗИРОВАННЫЕ СИСТЕМЫ УПРАВЛЕНИЯ

- Волынская А. В.** Разработка устройств формирования и оптимальной обработки сложных двоичных сигналов для информационных систем железнодорожного транспорта 65
- Приложение. Артамонов С. Е., Затуливетер Ю. С., Козлов В. А., Фищенко Е. А.** Элементы стратегии и архитектурные предпосылки опережения в области однокристальных многопроцессорных компьютеров с массовым параллелизмом .

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.

Published since November 1995

УЧРЕДИТЕЛЬ
Издательство "Новые технологии"

CONTENTS

INTELLIGENT SYSTEMS AND TECHNOLOGIES

- Kukhareno B. G., Solnceva M. O.** Use of Background Reduction Methods under Segmentation of Telemetric Images for Identification of Object Groups. 3
- Savchenko L. V.** Automatic Recognition of Isolated Words on the Basis of Fuzzy Set Theory and Cluster Model of Minimal Speech Units. 9
- Mochalov I. A., Khrisat M. S.** Estimation Parameter Model Using Fuzzy Random Data. 14

NETWORKS AND COMMUNICATION SYSTEMS

- Balonin N. A., Sergeev M. B.** Modern Network Technologies in Scientific Researches 23

MODELING AND OPTIMIZATION

- Levin V. I.** Solution Stability of Optimization Problems in Condition of Uncertainty . 27
- Orekhov E. Yu.** Generating Instances to Test Heuristic Algorithms for Rectangular Cutting-Packing 33
- Gavriushin S. S., Evgenev G. B.** Multicriteria Optimization in The Life Cycle of Products 37
- Amosova E. S.** Development of the Ontological Domain Model for Innovative Developments in Scientific Research. 42

STRUCTURAL SYNTHESIS AND DECISION MAKING

- Nikolaeva M. A., Zotova O. F., Sholokhova N. V.** DSS for the Commercial Bank Deposit Portfolio Management 46

GEOINFORMATION SYSTEMS

- Potapov V. P., Popov S. E., Zamaraev R. Yu.** Web Service for Calculating the Macro-seismic Field Based on Google App Engine 52
- Agayev F. G., Javadov N. H., Asadov H. H., Khalafov R. V.** New Technique of Optimization for Remote Sensing of Non-Homogenous Zones of Natural and Technogenic Medium 56
- Chernyaev A. V., Pavlov A. A.** Geographic Information System of the Small River Basin and Features of its Generation in Oil Pollution Modeling Problems 60

AUTOMATED CONTROL SYSTEMS

- Volynskaya A. V.** Development of Devices of Formation and Optimum Processing of Difficult Signals for the Information Systems of Railway Transport 65

- APPLICATION. Artamonov S. E., Zatuliveter Yu. S., Kozlov V. A., Fishchenko E. A.** Elements of the Strategy and Architectural Conditions in Advance of Single-Chip Multi-processor Computers with Massively Parallelism

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.
Kureichik V. M., Dr. Sci. (Tech.), Prof.
Kukhareno 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"

УДК 004.85, 519.724, 519.177

Б. Г. Кухаренко, канд. физ.-мат. наук, ст. науч. сотр., вед. науч. сотр.,
Институт машиноведения РАН, г. Москва, e-mail: kukharenko@imash.ru,
М. О. Солнцева, аспирант, Московский физико-технический институт (ГУ),
e-mail: solnceva.chalei@gmail.com

Использование методов сокращения фона при сегментировании телеметрических изображений для идентификации групп объектов

Рассматривается задача наблюдения за группой объектов при низком качестве телеметрического изображения. Идентификация объектов осуществляется посредством сегментирования телеметрического изображения в пространстве характеристик его пикселей. Показано, как для оценки мод распределения пикселей в пространстве характеристик телеметрического изображения используется метод сдвига среднего.

Избыточная инициализация этого метода всеми пикселями изображения сокращается по методу вырезания (удаления) швов, уменьшающего размер фона изображения. На примере телеметрического изображения показано, что метод сдвига среднего сглаживает его текстуру, усиливая визуальное восприятие объектов.

Ключевые слова: мобильные системы, децентрализованное управление, телеметрические изображения, сегментирование изображений, метод сдвига среднего, метод удаления швов

B. G. Kukharenko, M. O. Solnceva

Use of Background Reduction Methods under Segmentation of Telemetric Images for Identification of Object Groups

A problem object group observation is under study at a low quality of telemetric image. Object identification is performed by segmentation of telemetric image in its pixel feature space. As shown, for estimation of pixel distribution modes in the image pixel feature space, the Mean shift algorithm can be in use. An overabundant initialization of the algorithm by all image pixels is reduced by the Seam carving algorithm, which decreases the image background size. As shown by a telemetric image sample, the Mean shift algorithm smoothies its texture to visualize the object perception.

Keywords: mobile systems, decentralized control, telemetric images, image segmentation, Mean shift, Seam carving

Введение

Для мобильных систем (*mobile systems*) одним из ключевых технических вопросов является координация движений индивидуальных объектов (*vehicles*). Ввиду распределения в пространстве и ограниченного обмена информацией используется децентрализованное управление (*decentralized control*) [1]. Близкие по характеристикам объекты объединяются в группы для выполнения общей задачи. Если такими характеристиками являются координаты объектов в пространстве, задача группировки объектов сводится к определению кластеров объектов на основе их близости в пространстве [2, 3]. Однако специфическое децентрализованное управление

объектами и их группировка на основе относительного расположения в принципе не требуют явного знания координат объектов, а могут основываться на удаленных (телеметрических) изображениях реального времени для этих групп объектов. Мотивацией для разработки систем децентрализованного управления на основе телеметрических изображений является обеспечение мониторинга мультимедийных систем, состоящих из десятков объектов и разделенных сотнями метров расстояния [4]. При удаленном наблюдении (телеметрии) за объектами в режиме реального времени с высокой частотой кадров часто получаются изображения низкого качества, которые быстро теряют актуальность.

Задача анализа телеметрических изображений низкого качества достаточно сложна, так как качество результирующего изображения зависит от выбора параметров для его восстановления. Пиксели изображения описываются векторами пространства R^7 характеристик, состоящими из координат пикселя x и y на плоскости изображения, цвета пикселя по шкале RGB, относительной яркости и текстуры. Вследствие низкого качества телеметрического изображения пространство его характеристик является произвольно структурированным. Поэтому оно должно быть проанализировано быстрыми непараметрическими методами, не требующими предварительной информации [5].

Непараметрические методы иерархической кластеризации объединяют или разбивают векторы пространства характеристик на кластеры на основе некоторой меры близости [2]. Однако эти методы вычислительно трудоемки и, кроме того, критерий остановки для процедуры объединения или разбиения векторов определяется только по косвенным показателям. Непараметрические методы кластеризации на основе оценки мод плотности используют эмпирическую оценку функции плотности в пространстве характеристик изображения. Области высокой плотности в пространстве характеристик соответствуют локальным максимумам функции плотности, т. е. ее модам. Моде функции плотности соответствует кластер в пространстве характеристик, и, следовательно, сегмент изображения [6–8]. Подход, основанный на определении мод, и сегментирование изображения по методу сдвига среднего (*mean shift*) впервые предложены в работе [9]. Как показано в работах [10, 11], многократные (≈ 100) итерации процедуры сдвига среднего существенно сглаживают текстуру изображения, улучшая визуализацию объектов, распознаваемых в результате его сегментирования. Однако эффективность этой процедуры падает с ростом размеров изображения, так как при определении мод функции плотности процедура сдвига среднего последовательно инициализируется всеми пикселями изображения.

В настоящей работе задача анализа телеметрического изображения в режиме реального времени решается в два этапа, с использованием прореженного (уменьшенного) изображения, позволяющего однозначно идентифицировать наблюдаемые объекты. Суммарное время анализа изображения сопоставимо со временем, за которое происходит смена кадра. Изображение уменьшается за счет удаления неинформативной части фона с помощью метода швов (*seam carving*) [12]. Процедура удаления шва требует незначительного времени, так как шов строится исходя из значений градиента яркости пикселей. Многократное последовательное применение этой процедуры также требует незначительного времени. В результате последующее применение процедуры сдвига среднего к уменьшенному

изображению, во-первых, требует значительно меньше времени, во-вторых, достаточно одного применения этой процедуры для получения изображения однозначно идентифицируемых объектов. В работе представлены результаты двухэтапной обработки телеметрического изображения.

1. Метод швов, не искажающий границы объектов в панорамном кадре

При эффективном изменении размеров изображений необходимо принимать во внимание не только геометрические ограничения, налагаемые формой отдельных объектов (прямые линии должны выглядеть как прямые, окружность как окружность и т. п.), но также рассматривать содержимое изображения в целом. Для этого вводится специальный оператор, названный *seam carving* ("удаление шва"), позволяющий как уменьшать, так и увеличивать исходное изображение [12]. Шов — это непрерывная последовательность пикселей с "минимальной энергией", формула для которой приводится ниже. Шов может быть вертикальным и горизонтальным. При заданном направлении формирования шва в строке или в столбце матрицы, представляющей изображение, выбирается пиксель с минимальной энергией, а следующий пиксель шва выбирается среди восьми его ближайших соседей (*8-connected path*) таким образом, что на каждом шаге из строки или столбца матрицы удаляется только один пиксель. Применение оператора удаления шва в вертикальном и горизонтальном направлениях изменяет размеры изображения. Пиксели можно не только удалять, но и добавлять, не нарушая баланс между исходным изображением и искусственно вставленными пикселями.

Для процедуры удаления шва подходят несколько типов функций энергии: на основе градиента яркости, ее энтропии, визуального восприятия и др. Процедуры удаления и вставки не зависят от параметра, однако требуют непосредственного контроля. Простая функция энергии определяется градиентом

$$e_1(\mathbf{I}) = \left| \frac{\partial}{\partial x} \mathbf{I} \right| + \left| \frac{\partial}{\partial y} \mathbf{I} \right|. \quad (1)$$

Если при заданной функции энергии нужно уменьшить ширину изображения, то для изображения \mathbf{I} размера $m \times n$ вертикальный шов определяется следующей формулой:

$$s^x = \{s_i^x, i = \overline{1, n}\} = \{(x(i), i), i = \overline{1, n}\},$$

где $\forall i, |x(i) - x(i-1)| \leq 1$, x — отображение $x: [1, \dots, n] \rightarrow [1, \dots, m]$. Аналогично, если $y: [1, \dots, m] \rightarrow [1, \dots, n]$, тогда горизонтальный шов определяется как

$$s^y = \{s_j^y, j = \overline{1, m}\} = \{(j, y(j)), j = \overline{1, m}\},$$

где $\forall j, |y(j) - y(j-1)| \leq 1$. Пиксели вдоль пути шва \mathbf{s} (вертикального шва $\{s_i^x, i = \overline{1, n}\}$) описываются как

$$\mathbf{I}_s = \{\mathbf{I}(s_i), i = \overline{1, n}\} = \{\mathbf{I}(x(i), i), i = \overline{1, n}\}.$$

В противоположность удалению строки или столбца удаление пикселей шва изображения имеет только локальный эффект: все пиксели изображения сдвигаются влево (или вверх) для того, чтобы компенсировать недостающий пиксель пути шва. Визуальный эффект заметен только вдоль пути шва, при этом оставшаяся часть изображения не изменяется. Можно заменить ограничение $|x(i) - x(i-1)| \leq 1$ на $|x(i) - x(i-1)| \leq k$ и получить либо просто столбец (или строку) для $k = 0$, либо пространственно связанный, или даже полностью несвязанный набор пикселей для любого значения $1 \leq k \leq m$.

При заданной функции энергии e (1) стоимость шва определяется как

$$E(\mathbf{s}) = E(\mathbf{I}_s) = \sum_{i=1}^n e(\mathbf{I}(s_i)).$$

Далее ищется оптимальный шов \mathbf{s}^* , стоимость которого минимальна:

$$\mathbf{s}^* = \min_{\mathbf{s}} E(\mathbf{s}) = \min_{\mathbf{s}} \sum_{i=1}^n e(\mathbf{I}(s_i)).$$

Оптимальный шов определяется в результате рекурсии в матрице стоимости всех швов. На первом шаге, в случае вертикального шва, пиксели изображения анализируются, начиная со второй строки до последней. При этом вычисляется общий минимум энергии для всех возможных связанных путей швов, начинающихся с элемента (i, j) :

$$M(i, j) = e(i, j) + \min(M(i-1, j-1), M(i-1, j), M(i-1, j+1)).$$

Минимальное значение элемента в последней строке матрицы M показывает конец минимального связанного вертикального шва. На втором шаге возвращаются назад, начиная с минимального элемента последней строки для того, чтобы найти путь оптимального шва. Построение оптимального горизонтального шва выполняют аналогично.

2. Процедура сдвига среднего (Mean Shift)

Как отмечалось во введении, каждый пиксель изображения представляется вектором характеристик \mathbf{x}_i . В настоящем разделе учитывается положение пикселя (x_i, y_i) и цвет (r_i, g_i, b_i) , т. е. используются пятимерные векторы $\mathbf{x}_i = (x_i, y_i, r_i, g_i, b_i) \in \mathbb{R}^5$. Считается, что они представляют собой выборку для функции распределения плотности вероятности. Поэтому более плотные области в пространстве характеристик соответствуют локальным максимумам функции распределения плотности вероятности,

т. е. модам этого (неизвестного) распределения. Мода определяется на основе локальной структуры пространства состояний и в результате устанавливается центр связанного с ней кластера [6–9]. Используемый в настоящем разделе подход основан на процедуре сдвига среднего [10, 11].

Оценка распределения плотности вероятности $f^*(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^d$ на основе векторов $\{\mathbf{x}_i, i = \overline{1, n}\} \in \mathbb{R}^d$ с использованием ядра известна как метод окна Парзена (*Parzen window*). Эта оценка вычисляется с помощью многомерного ядра $K(\mathbf{x})$ и симметричной, положительно определенной матрицы ширины полосы (*bandwidth matrix*) \mathbf{H} размерности $d \times d$ по формуле

$$f^*(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n K_{\mathbf{H}}(\mathbf{x} - \mathbf{x}_i), \quad (2)$$

где

$$K_{\mathbf{H}}(\mathbf{x}) = |\mathbf{H}|^{-1/2} K(|\mathbf{H}|^{-1/2} \mathbf{x}) \quad (3)$$

является финитной функцией и удовлетворяет стандартным условиям

$$\int_{\mathbb{R}^d} K(\mathbf{x}) d\mathbf{x} = 1, \quad \lim_{\|\mathbf{x}\| \rightarrow \infty} \|\mathbf{x}\|^d K(\mathbf{x}) = 0,$$

$$\int_{\mathbb{R}^d} K(\mathbf{x}) d\mathbf{x} = 0, \quad \int_{\mathbb{R}^d} \mathbf{x} \mathbf{x}^T K(\mathbf{x}) d\mathbf{x} = c_K \mathbf{I},$$

где c_K — константа. На практике матрица \mathbf{H} выбирается либо диагональной $\mathbf{H} = \text{diag}(h_1^2, \dots, h_d^2)$, либо пропорциональной единичной матрице $\mathbf{H} = h^2 \mathbf{I}$, где $h > 0$ — параметр ширины полосы. Многомерное ядро генерируется из симметричного одномерного ядра $K_1(x)$ либо произведением d таких ядер, либо его вращением в \mathbb{R}^d , в результате чего получают радиально-симметричное ядро [10, 11]. Качество оценки распределения плотности вероятности с помощью многомерного ядра измеряется средней квадратичной ошибкой между распределением плотности вероятности и его оценкой, проинтегрированной по области определения. Однако на практике можно вычислить только асимптотическую аппроксимацию этой меры (сокращенно AMISE — *asymptotic approximation of the measure*). При асимптотике число точек наблюдения $n \rightarrow \infty$, в то время как ширина полосы $h \rightarrow 0$ со скоростью, меньше чем n^{-1} . Для обоих типов многомерных ядер мера AMISE минимизируется ядром Епанечникова (*Epanechnikov kernel*), имеющим профиль

$$k_E(x) = \begin{cases} 1 - x^2, & 0 \leq x \leq 1; \\ 0, & x > 1, \end{cases} \quad (4)$$

который приводит к радиально симметричному ядру

$$K_E(\mathbf{x}) = \begin{cases} \frac{1}{2} c_d^{-1} (d+2)(1 - \|\mathbf{x}\|^2), & \|\mathbf{x}\| \leq 1; \\ 0, & \|\mathbf{x}\| > 1, \end{cases} \quad (5)$$

где c_d — объем единичной d -мерной сферы. Профиль

$$k_N(x) = \exp\left(-\frac{1}{2}x\right), \quad x \geq 0 \quad (6)$$

задает многомерное нормальное ядро

$$K_N(\mathbf{x}) = (2\pi)^{-d/2} \exp\left(-\frac{1}{2}\|\mathbf{x}\|^2\right). \quad (7)$$

Оба ядра (5) и (7) используются в данной работе. Применяя обозначение профиля (4) и (6), оценку плотности распределения вероятности (2), (3) можно представить в следующем виде:

$$f_{h,K}^*(\mathbf{x}) = \frac{c_{k,d}}{nh^d} \sum_{i=1}^n k\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right). \quad (8)$$

Моды распределения $f(\mathbf{x})$ (8) соответствуют нулевым значениям градиента $\nabla f(\mathbf{x}) = 0$. Процедура сдвига среднего (*mean shift*) является способом локализации мод, не оценивая само распределение вероятности.

2.1. Оценка градиента распределения

Из оценки распределения (8) получаем оценку градиента распределения вероятности

$$\begin{aligned} \nabla^* f_{h,K}(\mathbf{x}) &\equiv \nabla f_{h,K}^*(\mathbf{x}) = \\ &= \frac{2c_{k,d}}{nh^{d+2}} \sum_{i=1}^n (\mathbf{x} - \mathbf{x}_i) k'\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right). \end{aligned} \quad (9)$$

Определим функцию

$$g(x) = -k'(x). \quad (10)$$

Обозначим профиль ядра $G(\mathbf{x})$ как $g(x)$, тогда ядро определяется как

$$G(\mathbf{x}) = c_{g,d} g(\|\mathbf{x}\|^2), \quad (11)$$

где $c_{g,d}$ — соответствующая константа нормировки. Формула (9) для оценки градиента распределения вероятности принимает вид

$$\begin{aligned} \nabla^* f_{h,K}(\mathbf{x}) &= \frac{2c_{k,d}}{nh^{d+2}} \sum_{i=1}^n (\mathbf{x}_i - \mathbf{x}) g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right) = \\ &= \frac{2c_{k,d}}{nh^{d+2}} \left[\sum_{i=1}^n g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right) \right] \times \\ &\times \frac{\left[\sum_{i=1}^n \mathbf{x}_i g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right) \right]}{\left[\sum_{i=1}^n g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right) \right]} - \mathbf{x}, \end{aligned} \quad (12)$$

где предполагается, что $\sum_{i=1}^n g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right)$ — положительное число. Это условие выполняется для профилей (4) и (6). Множители в формуле (12), заключенные в квадратные скобки, имеют специальное значение. Как следует из (11), первый из них пропорционален оценке распределения вероятности в точке \mathbf{x} , вычисленной с помощью ядра $G(\mathbf{x})$,

$$f_{h,G}^*(\mathbf{x}) = \frac{c_{g,d}}{nh^d} \sum_{i=1}^n g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right). \quad (13)$$

Второй множитель представляет отклонение среднего значения характеристик $\{\mathbf{x}_i, i = \overline{1, n}\} \in \mathbb{R}^d$ в точке \mathbf{x} от значения в точке \mathbf{x} , т. е. отражает сдвиг среднего (*mean shift*):

$$\mathbf{m}_{h,G}(\mathbf{x}) = \frac{\sum_{i=1}^n \mathbf{x}_i g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right)} - \mathbf{x}. \quad (14)$$

Формула (14) показывает разницу между взвешенным средним, полученным при использовании ядра $G(\mathbf{x})$, и центром ядра (окна) — точкой \mathbf{x} . С учетом (13) и (14) формула (12) принимает вид

$$\nabla^* f_{h,K}(\mathbf{x}) = f_{h,G}^*(\mathbf{x}) \frac{2c_{k,d}}{h^2 c_{g,d}} \mathbf{m}_{h,G}(\mathbf{x}). \quad (15)$$

Откуда следует

$$\mathbf{m}_{h,G}(\mathbf{x}) = \frac{1}{2} h^2 c \frac{\nabla^* f_{h,K}(\mathbf{x})}{f_{h,G}^*(\mathbf{x})}. \quad (16)$$

Таким образом, вектор сдвига среднего всегда направлен в сторону максимального увеличения плотности вероятности. Локальное среднее сдвигается в направлении области, в которой находится большинство наблюдаемых точек. Поскольку вектор сдвига среднего выравнивается соответственно с оценкой локального градиента, он может определить путь, ведущий к стационарной точке оцениваемого распределения. Моды распределения являются такими стационарными точками.

2.2. Достаточное условие сходимости

Пусть $\{\mathbf{y}_j\}_{j=1,2,\dots}$ — последовательность положений ядра $G(\cdot)$ (11). В соответствии с формулой (14)

$$\mathbf{y}_{j+1} = \frac{\sum_{i=1}^n \mathbf{x}_i g\left(\left\|\frac{\mathbf{y}_j - \mathbf{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{\mathbf{y}_j - \mathbf{x}_i}{h}\right\|^2\right)}, \quad j = 1, 2, \dots, \quad (17)$$

является взвешенным средним в точке \mathbf{y}_j , вычисленным с помощью ядра $G(\cdot)$ (11), и \mathbf{y}_1 — исходное положение этого ядра. Соответствующая последовательность $\{f_{h,K}^*(j)\}_{j=1,2,\dots}$ оценок распределения, вычисленных с помощью ядра K , задается следующим образом:

$$\{f_{h,K}^*(j) = \{f_{h,K}^*(\mathbf{y}_j), j = 1, 2, \dots \quad (18)$$

При определенных условиях ядро $K(\cdot)$ обеспечивает сходимость последовательностей $\{\mathbf{y}_j\}_{j=1,2,\dots}$ и $\{f_{h,K}^*(j)\}_{j=1,2,\dots}$ [10, 11].

2.3. Обнаружение моды, основанное на сдвиге среднего

Кластеризация набора векторов характеристик пикселей $\{\mathbf{x}_i, i = \overline{1, n}\}$ происходит последовательным выбором в качестве начальной точки \mathbf{y}_1 в (17) всех векторов $\mathbf{x}_i, i = \overline{1, n}$. Точки пространства характеристик объединяются в кластеры в соответствии с неподвижными (стационарными) точками (17), определяющими моды распределения. Согласно (14) и (15) в стационарной точке \mathbf{y}_C уравнения (17) сдвиг среднего $\mathbf{m}_{h,G}(\mathbf{y}_C) = 0$ и градиент распределения $\nabla f_{h,K}^*(\mathbf{y}_C) = 0$. Набор всех $\mathbf{x}_i, i = \overline{1, n}$, для которых итерации (17) сходятся к одной и той же моде, определяет область притяжения (*basin of attraction*) этой моды. Отсюда следует практический алгоритм для обнаружения моды: сначала процедурой сдвига среднего определяются стационарные точки (17), затем отбрасываются все точки, которые не являются локальными максимумами. Точки локальных максимумов определяются как уникальные стационарные точки в пределах некоторой малой открытой сферы. Они могут быть проверены возму-

щением стационарной точки случайным вектором с малой нормой и применением процедуры сдвига среднего. Если стационарная точка не меняется (в пределах некоторой допустимой погрешности), то она является локальным максимумом распределения плотности вероятности.

3. Численный эксперимент

В настоящей работе для улучшения распознавания объектов на телеметрических изображениях предлагается двухэтапный подход к обработке этих изображений. Этот подход заключается в многократном применении процедуры удаления шва с последующей обработкой промежуточного изображения с помощью однократной процедуры сдвига среднего. На рис. 1 (см. третью сторону обложки) рассматривается типичный пример телеметрического изображения. Размер изображения 965×832 пикселей. Это изображение невысокого качества с засвеченным верхним правым углом. В данном случае объекты наблюдения участвуют в соревнованиях по воздухоплаванию. На исходном изображении выделяются 23 объекта. Интерес представляет только та часть изображения, на которой присутствуют объекты наблюдения. В целях упрощения наблюдения за объектами неинформативная часть кадра сокращается методом вырезания швов.

Продемонстрируем результат выполнения процедуры выделения вертикального шва для изображения, приведенного на рис. 1. На рис. 2, а приведен черно-белый негатив этого изображения и полученный на его основе градиент яркости (рис. 2, б). Рис. 2, в показывает выделенный первый вертикальный шов в исходном изображении. Многократное повторение процедуры удаления шва значительно сокращает фон изображения, не затрагивая границы объектов.

С помощью метода вырезания швов группируются объекты, что способствует значительному со-

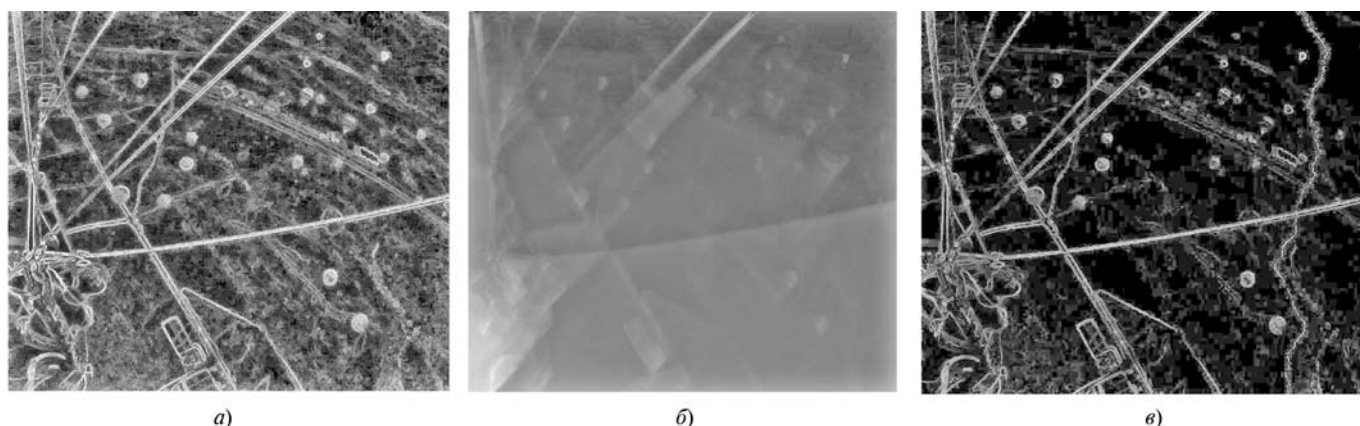


Рис. 2. Основные шаги процедуры вырезания шва для изображения на рис. 1: а — черно-белый негатив изображения; б — градиент яркости негатива; в — шов в изображении

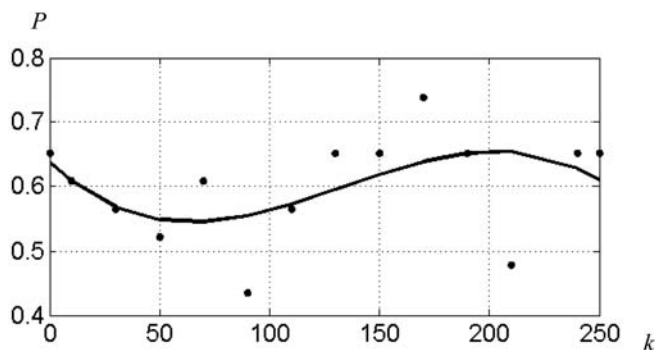


Рис. 4. Доля P идентифицированных объектов наблюдения после применения процедуры сдвига среднего в зависимости от числа швов k , вырезаемых на первом этапе обработки изображения на рис. 1. Сплошной линией показана линия тренда при аппроксимации полиномом 3-й степени

крашению времени последующей обработки изображения, удаляются пиксели фона и улучшается распознавание объектов. На рис. 3, а (см. третью сторону обложки) показан градиент яркости исходного цветного изображения (см. рис. 1). На рис. 3, б показано отображение доверия (*confidence map*). На рис. 3, в представлен результат однократного применения процедуры сдвига среднего к исходному изображению без предварительного вырезания швов. При его сравнении с исходным изображением уже можно отметить уменьшение фона в засвеченной части кадра. Нумерация объектов на рис. 3, в соответствует исходному изображению (см. рис. 1).

На рис. 4 показана зависимость доли идентифицированных объектов (от исходного числа объектов — 23) от числа вырезанных швов в исходном изображении (см. рис. 1). Исследование выполняется в два этапа: вырезание некоторого числа швов завершается обработкой изображения с помощью однократной процедуры сдвига среднего. Начиная с вырезания 130 швов, доля распознаваемых объектов практически не изменяется и соответствует начальному результату распознавания, однако время обработки изображения сокращается на порядок. Сплошной линией показана линия тренда при аппроксимации полиномом 3-й степени с коэффициентами при степенях аргумента в порядке убывания: $(-0,000000081668490, 0,000032854275516, -0,003218979870898, 0,637810565797082)$.

В изображении, приведенном на рис. 5 (см. четвертую сторону обложки) сначала удаляются 30 швов (рис. 5, а) и затем применяется процедура сдвига среднего (рис. 5, б). При сравнении рис. 5, б и рис. 3, в, полученного без вырезания швов в результате применения одной процедуры сдвига среднего к исходному изображению, заметно сокращение засвеченной части кадра.

На рис. 6 (см. четвертую сторону обложки) показаны результаты двухэтапной обработки исходного изображения при вырезании 150 швов. Этот пример соответствует количественному интервалу налагаемых швов (от 130 до 250 и более), в котором доля распознаваемых объектов в среднем такая же, как на исходном изображении.

На рис. 7 (см. четвертую сторону обложки) приведен вариант, соответствующий вырезанию 200 швов при обработке изображения рис. 1, он демонстрирует практически полное отсутствие дефектов фона (засвеченной области правого верхнего угла).

Заключение

Предварительное выполнение процедуры вырезания швов обеспечивает сокращение времени анализа этого изображения процедурой сдвига среднего без потери его качества и сохранение количества распознаваемых объектов.

Список литературы

1. **Siljak D.** Decentralized Control of Complex Systems. New York: Academic Press, 1991.
2. **Jain A. K., Dubes R. C.** Algorithms for Clustering Data. Englewood Cliffs, NJ: Prentice Hall, 1988.
3. **Кухаренко Б. Г., Солнцева М. О.** Принцип минимальной длины описания при анализе графов с разреженными матрицами смежности в задачах кластеризации их узлов // Информационные технологии. 2013. № 7. С. 37–42.
4. **Mahacek P., Mas I., Petrovic O., Acain J., Kitts C.** Cluster space control of autonomous surface vessels // Marine Technology Society Journal. 2009. V. 43, No. 1. P. 13–20.
5. **Деврой Л., Дьерфи Л.** Непараметрическое оценивание плотности. L_1 -подход. М.: Мир, 1988.
6. **Herbin M., Bonnet N., Vautrot P.** A clustering method based on the estimation of the probability density function and on the skeleton by influence zones // Pattern Recognition Letters. 1996. V. 17. P. 1141–1150.
7. **Touzani A., Postaire J. G.** Clustering by mode boundary detection // Pattern Recognition Letters. 1989. V. 9. P. 1–12.
8. **Wren C., Azarbayejani A., Darrell T., Pentland A.** Pfunder: Real-time tracking of the human body // Transactions on Pattern Analysis and Machine Intelligence. 1997. V. 19, No. 7. P. 780–785.
9. **Fukunaga K., Hostetler L. D.** The estimation of the gradient of a density function, with applications in pattern recognition // Transactions on Information Theory. 1975. V. 21. P. 32–40.
10. **Comaniciu D., Meer P.** Mean shift: A robust approach toward feature space analysis // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002. V. 24, No. 5. P. 603–619.
11. **Comaniciu D., Meer P.** An algorithm for data driven bandwidth selection // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2003. V. 25, No. 2. P. 281–288.
12. **Avidan S., Shamir A.** Seam carving for content aware image resizing // ACM Transactions on Graphics. 2007. V. 26, No. 3. P. 10–18.
13. **Кухаренко Б. Г.** Сегментирование изображений на основе сечения графов, Марковских случайных полей и алгоритмов разведывания данных // Информационные технологии. 2011. № 2. Приложение. С. 1–32.

Л. В. Савченко, аспирант, Нижегородский государственный
лингвистический университет им. Н. А. Добролюбова, e-mail: LyudmilaSavchenko@yandex.ru

Автоматическое распознавание изолированных слов на основе теории нечетких множеств и кластерной модели минимальных речевых единиц

На основе кластерной модели минимальных речевых единиц и теории нечетких множеств предложен алгоритм распознавания фонем в слухе, если для каждой фонемы доступно несколько эталонных реализаций. С использованием такого метода нечеткого фонетического кодирования—декодирования разработан новый алгоритм распознавания изолированных слов. Представлены результаты его экспериментального исследования для русского языка. Показано, что предложенный подход характеризуется повышением точности распознавания на 3–5,5 % по сравнению с известными аналогами.

Ключевые слова: распознавание изолированных слов, принцип минимума информационного рассогласования, кластерная модель минимальных речевых единиц, метод нечеткого фонетического кодирования—декодирования

L. V. Savchenko

Automatic Recognition of Isolated Words on the Basis of Fuzzy Set Theory and Cluster Model of Minimal Speech Units

A novel automatic phoneme recognition algorithm was proposed on the basis of cluster model of minimal speech units and fuzzy set theory, if several model phonemes are available. By using this fuzzy coding-decoding method the algorithm of isolated words recognition was developed. The results of its experimental study for Russian language are presented. It is shown that the accuracy of the proposed approach is 3–5,5 % higher than the accuracy of other well-known methods.

Keywords: automatic speech recognition of isolated words, minimum information discrimination principle, cluster model of minimal speech units, phonetic fuzzy coding-decoding method

Введение

Направление автоматического распознавания речи (АРР) и, в частности распознавания изолированных слов в задаче голосового управления, в настоящее время становится все более актуальным [1]. К сожалению, качество существующих коммерческих решений (таких как Microsoft Voice Command, Windows Speech Recognition, Nuance Dragon Naturally Speaking, Google Voice Search и т. п.), основанных на широко распространенном методе скрытых Марковских моделей [1, 2], оказывается для некоторых практически важных задач (в системах голосового управления с повышенными требованиями к надежности [3], при наличии акустических помех [4] и т. п.), особенно для русского языка [5], все еще недостаточным. Как известно [3], для повышения качества распознавания русской речи в системах голосового управления с повышенной помехозащищенностью можно использовать известный метод фонетического декодирования слов (ФДС) [6], в котором задача распознавания слов сводится к выделению границ слогов и распознаванию гласной фонемы в каждом слого. Его основными преимуществами являются быстрая настройка под конкретного диктора и автоматически перенастраиваемый

рабочий словарь [3]. К сожалению, точность ФДС иногда оказывается недостаточной, так как близкие по звучанию фонемы часто перепутываются между собой [7]. Для повышения качества распознавания фонем за счет использования дополнительной информации о попарных расстояниях между эталонными минимальными речевыми единицами (МРЕ) нами было предложено обобщение ФДС — метод нечеткого фонетического кодирования—декодирования (НФКД) [8]. В настоящей работе предложена модификация НФКД для задачи распознавания слов на случай, когда для каждой фонемы доступно несколько эталонных реализаций. Полученные результаты и сделанные по ним выводы рассчитаны на широкий круг специалистов в области АРР.

Метод фонетического декодирования в задаче распознавания гласной фонемы в слого

Рассмотрим задачу автоматического распознавания гласных фонем в слого: вновь поступающему на вход слого — речевому сигналу x с частотой дискретизации F (в герцах) — требуется поставить в соответствие одну из R гласных фонем. Каждая r -я фонема ($r = \overline{1, R}$) задается множеством эталонов

$X_r = \{\mathbf{x}_{r,j}\}$, где $\mathbf{x}_{r,j}, j = \overline{1, J_r}, J_r \gg 1$ — одноименные (однофонемные) реализации; J_r — число эталонных фонем r -го класса.

В методе ФДС [6] одноименные реализации в сознании человека группируются в соответствующие классы вокруг некоторого центра — эталонной метки. Для определения понятия "фонема" используется кластерная модель МРЕ [9]: речевая метка $\mathbf{x}_r^* \in X_r$ образует информационный центр-эталон (ИЦ) r -го класса, если в пределах множества X_r она характеризуется минимальной суммой информационных рассогласований по Кульбаку—Лейблеру [10] относительно всех других его меток-реализаций $\mathbf{x}_{r,j}, j = \overline{1, J_r}$, в пределах каждого r -го кластера:

$$\mathbf{x}_r^* = \arg \min_{\mathbf{x}_{r,k}, k \in \{1, \dots, J_r\}} \sum_{j=1}^{J_r} \rho_{KL}(\mathbf{x}_{r,k}/\mathbf{x}_{r,j}), \quad (1)$$

где $\rho_{KL}(\mathbf{x}_{r,k}/\mathbf{x}_{r,j})$ — информационное рассогласование Кульбака—Лейблера [10] между $\mathbf{x}_{r,k}$ и $\mathbf{x}_{r,j}$, совпадающее, в случае гауссовского распределения фонемы с нулевым средним и неизвестной автокорреляционной матрицей, с издавна применяющимся в автоматической обработке речи рассогласованием Итакуры—Саито [1, 11]. Для широко используемой в АРР [1, 3] авторегрессионной (АР) модели порядка p речевого сигнала рассогласование Кульбака—Лейблера может быть записано как [1, 5]

$$\rho_{KL}(\mathbf{x}(t)/\mathbf{x}_r^*) = \frac{1}{2} \left[\frac{\sigma_r^2(\mathbf{x})}{\sigma_0^2} - 1 \right]. \quad (2)$$

Здесь $\sigma_0^2 = \sigma_x^2$, где σ_x^2 — выборочная оценка дисперсии порождающего процесса $\mathbf{x}(t)$, а $\sigma_r^2(\mathbf{x})$ — выборочная оценка дисперсии отклика r -го обеляющего фильтра $y_{r,j}(t) = \|y_{r,1}(t) \dots y_{r,M-p}(t)\|$, где

$$y_{r,j}(t) = x_{j+p}(t) - \sum_{m=1}^p a_{r,m} x_{j+p-m}(t), \quad j = \overline{1, M-p}. \quad (3)$$

ИЦ-эталон \mathbf{x}_r^* задается своим вектором АР-коэффициентов $\mathbf{a}_r = \{a_{r,m}\}, m = \overline{1, p}$, полученным, например, с помощью алгоритма Берга и рекурсивной процедуры Левинсона—Дурбина [12].

Для решения поставленной задачи разобьем сигнал \mathbf{x} на ряд из T непересекающихся сегментов данных $\mathbf{x}(t), t = \overline{1, T}$ длиной $\tau = 0,01 \dots 0,015$ с. [1, 13]. После этого каждый парциальный сигнал $\mathbf{x}(t) = \|x_1(t) \dots x_M(t)\|$ (здесь $M = \tau F$ — число отсчетов в сегменте) рассматривается в пределах конечного списка фонем $\{\mathbf{x}_r^*\}$ и отождествляется с той $\mathbf{x}_{v(t)}$ из них, которая отвечает принципу минимума значения заданной исследователем меры близости между сигналом $\mathbf{x}(t)$ и ИЦ \mathbf{x}_r^* . Асимптотически оптимальное в байесовском смысле решение задачи классификации с оценкой автокорреляционной матрицы в предположении [9, 11] о независимости векторов

$\|x_1(t) \dots x_p(t)\|, \|x_2(t) \dots x_{p+1}(t)\|, \dots, \|x_{M-p+1}(t) \dots x_M(t)\|$ дает основанный на принципе минимума информационного рассогласования Кульбака—Лейблера [10] критерий

$$v(t) = \arg \min_{r = \overline{1, R}} \rho_{KL}(\mathbf{x}(t), \mathbf{x}_r^*), \quad t = \overline{1, T}. \quad (4)$$

Таким образом, для каждого сегмента $\mathbf{x}(t)$ имеем вектор расстояний

$$\rho(t) = \|\rho_{KL}(\mathbf{x}(t)/\mathbf{x}_1^*), \dots, \rho_{KL}(\mathbf{x}(t)/\mathbf{x}_R^*)\|. \quad (5)$$

Тогда для распознавания гласной фонемы требуется по набору векторов (5) для всех $t = \overline{1, T}$ принять решение в пользу принадлежности к одному из R классов $\{\mathbf{x}_r^*\}$. Согласно методу ФДС [6], каждой МРЕ \mathbf{x} ставится в соответствие некий числовой код $c(r)$, где $c(r) \in \{1 \dots C\}$, в общем случае $C \leq R$. Для каждого сегмента в момент времени t решение принимается по критерию (4). Итоговое решение принимается в пользу наиболее часто встречающегося кода c^* :

$$c^* = \arg \max_{c = \overline{1, C}} \sum_{t=1}^T \delta(c(v(t)) - c), \quad (6)$$

где $\delta(x)$ — дискретная дельта-функция; $v(t)$ определяется согласно (4). Таким образом, формулы (4)—(6) описывают алгоритм распознавания гласной фонемы в слого на основе метода ФДС.

Алгоритм распознавания гласной фонемы в слого на основе операций с нечеткими множествами

В работах [7, 8] для распознавания фонем был предложен метод НФКД как обобщение ФДС. В НФКД предложено рассматривать фонему как нечеткое множество эталонных МРЕ: j -й ($j = \overline{1, R}$) МРЕ ставится в соответствие не один ИЦ-эталон \mathbf{x}_r^* как в методе ФДС, а нечеткое множество вида $\{\{\mathbf{x}_r^*, \mu_j(\mathbf{x}_r^*)\}\}$, где $\mu_j(\mathbf{x}_r^*)$ — степень принадлежности эталона \mathbf{x}_r^* к j -й МРЕ. Известно [10], что рассогласование Кульбака—Лейблера между объектами одного класса имеет в асимптотике распределение χ_S^2 с числом степеней свободы s , определяемым числом независимых параметров классифицируемого объекта. В предположении о гауссовском распределении МРЕ и для указанных выше условий оптимальности критерия (4) имеем $s = M - p$ (см. работу [9]). Тогда степень принадлежности $\mu_j(\mathbf{x}_r^*)$ определяется [7] как

$$\mu_j(\mathbf{x}_r^*) = P(\mathbf{x}_j^*/\mathbf{x}_r^*) = P(\rho_{KL}(\mathbf{x}_j^*/\mathbf{x}_r^*) < \chi_{M-p}^2). \quad (7)$$

Здесь $P(\mathbf{x}_j^*/\mathbf{x}_r^*)$ — условная вероятность принадлежности \mathbf{x}_j^* к j -й МРЕ. Таким образом, чем больше рассогласование между эталонными МРЕ, тем меньше степень принадлежности $\mu_j(\mathbf{x}_r^*)$. Аналогично, входному сигналу $\mathbf{x}(t)$ ставится в соответствие нечеткое множество вида $\{\mathbf{x}_r^*, \mu(\mathbf{x}(t)/\mathbf{x}_r^*)\}$, где

$$\mu(\mathbf{x}(t)/\mathbf{x}_r^*) = P(\rho_{KL}(\mathbf{x}(t)/\mathbf{x}_r^*) < \chi_{M-p}^2). \quad (8)$$

Далее для каждого сегмента $\mathbf{x}(t)$, используя операцию нечеткого пересечения [14] множеств, получаем результирующее множество $\{(\mathbf{x}_r^*, \mu(r, t))\}$, где

$$\mu(r, t) = \min(\mu_{v(t)}(\mathbf{x}_r^*), \mu(\mathbf{x}(t)/\mathbf{x}_r^*)). \quad (9)$$

Здесь $v(t)$ определяется согласно (4).

На следующем шаге алгоритма на основе всех $\mu(r, t)$ (9) определяется итоговая степень принадлежности $\mu(r)$ сигнала к r -й МРЕ:

$$\mu(r) = \frac{1}{T} \sum_{t=1}^T \mu(r, t). \quad (10)$$

В результате на выходе алгоритма распознавания гласной фонемы в слоге получаем нечеткое множество $\{(\mathbf{x}_r^*, \mu(r))\}$. Таким образом, решение в НФКД принимается по формулам (7)–(10). Алгоритм (7)–(10) использует информацию только об ИЦ-эталонах $\{\mathbf{x}_r^*\}$ (8), (9).

В настоящей работе предлагается обобщение НФКД на случай с несколькими реализациями МРЕ для каждого эталонного класса. Вначале для каждого класса-эталона ищется его ИЦ-эталон \mathbf{x}_r^* (1). Далее, чтобы учесть информацию обо всех реализациях $\mathbf{x}_{j, k}$, $k = \overline{1, J_j}$, МРЕ, используя операцию нечеткого объединения [14], определим $\mu_j(\mathbf{x}_r^*)$ как

$$\mu_j(\mathbf{x}_r^*) = \max_{k \in \{1, \dots, J_j\}} P(\mathbf{x}_{j, k}/\mathbf{x}_r^*). \quad (11)$$

Итоговое решение в обобщенном НФКД принимается по формулам (7)–(11). Определение (9) позволяет не занижить степени принадлежности $\mu(r, t)$ при использовании операции (11), если хотя бы одна из эталонных МРЕ из множества $\{\mathbf{x}_{v(t), k}\}$, $k = \overline{1, J_{v(t)}}$ близка к ИЦ-эталону \mathbf{x}_r^* . Для остальных фонем $\mu(r, t)$ уменьшаются по сравнению с $\mu(\mathbf{x}(t)/\mathbf{x}_r^*)$ и таким образом, их вклад в результирующее решение (10) незначителен. Это особенно важно в последующей задаче распознавания слов, степень принадлежности которых может быть вычислена как среднее степеней принадлежности составляющих слово слогов.

Результаты экспериментальных исследований в задаче распознавания слогов

Исследование эффективности предложенного обобщения НФКД (7)–(11) проводили на десяти различных реализациях слогов для каждого гласного звука русского языка (а, я, у, ю, о, ё, э, е, ы, и) одного диктора, так как известно, что точность распознавания ударных гласных звуков наиболее высока [1, 5, 15]. Например, для звука "ё" распознавались следующие слоги: мё, лё, ёж, слё, дё, ёл, пёс, всё, тёс, вё. В качестве эталонов брали десять МРЕ этого же диктора по четыре реализации на каждый звук. Для записи сигнала применяли встроенный в ноутбук микрофон. Частота дискретизации F была установлена равной 8 кГц. Длина одного сегмента

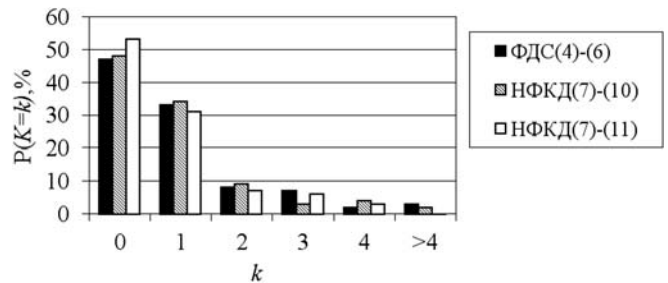


Рис. 1. Распределение случайной величины K

данных составляла $M = 120$ отсчетов ($\tau = 0,015$ с), порядок АР-модели $p = 20$. Далее представлены результаты сравнительного анализа трех методов: ФДС (4)–(6), оригинального НФКД (7)–(10) и предложенной модификации (7)–(11). В методе ФДС во избежание пропуска истинной фонемы близкие эталоны объединяли в один класс ($\{а, я\}$, $\{у, ю\}$, $\{о, ё\}$, $\{э, е\}$, $\{ы, и\}$). Для всех алгоритмов распознавания гласной фонемы в слоге оценивалось распределение случайной величины K — числа МРЕ, степень принадлежности которых больше, чем истинная $\mu(\gamma)$. Результаты, усредненные по всем тестовым реализациям, представлены на рис. 1. Здесь k — значения, которые может принимать случайная величина K .

Из рис. 1 видно, что при использовании рассогласования (2) для обобщенного метода НФКД, вероятность

$$p_1 = P(K = 0 \vee K = 1) = P(K = 0) + P(K = 1)$$

того, что не более одной МРЕ имели более высокую степень принадлежности, чем истинная МРЕ, \mathbf{x}_r^* равна $53 + 31 = 84$ %. Для сравнения аналогичная величина $p_1 = 82$ % для ранее предложенного НФКД и $p_1 = 80$ % для метода ФДС.

Таким образом, обобщение НФКД имеет большую точность распознавания по сравнению с оригинальным НФКД и методом ФДС. Это связано, в первую очередь, с использованием операции нечеткого объединения (11), в которой учитывается информация о каждой реализации эталонного класса.

Распознавание изолированных слов

Пусть задано множество из $L > 1$ эталонных слов $\{X_l\}$, где $l = \overline{1, L}$ — номер слова-эталона. Согласно популярному фонетическому подходу [16], каждое эталонное слово разбивается на последовательность фонем (транскрипцию) $X_l = \{c_{l, 1}, c_{l, 2}, \dots, c_{l, L_l}\}$. Здесь L_l — длительность команды (в фонемах), а числа $c_{l, j} \in \{1, \dots, R\}$ — номера фонем из фонетического алфавита $\{\mathbf{x}_r^*\}$, $r = \overline{1, R}$, где R — число фонем в алфавите. Задача состоит в том, чтобы поступившему на вход речевому сигналу X поставить в соответствие наиболее близкое к нему слово-эталон.

Известно [3], что задачу распознавания слов можно свести к задаче распознавания слогов, когда

степень принадлежности слова определяется как среднее арифметическое степеней принадлежности составляющих его слогов. Предположим, что входное слово X разбито на N слогов, причем границы каждого n -го слога ($n = \overline{1, N}$) определены с точностью до номера сегмента ($t_n^{(1)}, t_n^{(2)}$). Тогда для распознавания n -му слогу каждого эталона x_r^* ставится в соответствие последовательность относительных частот $\mu_n(r)$ сегментов слога, к которым x_r^* оказался ближе (в смысле (4)) остальных МРЕ:

$$\mu_n(r) = \frac{1}{t_n^{(2)} - t_n^{(1)} + 1} \sum_{t=t_n^{(1)}}^{t_n^{(2)}} \mu(r, t),$$

где $\mu(r, t)$ определяется согласно (9). Далее для каждого слова-эталона X_l оценивается его корреляция с распознаваемым речевым сигналом:

$$\mu_l = \begin{cases} \frac{1}{N} \sum_{n=1}^N \mu_n(c_l, n), & L_l = N, \\ 0, & L_l \neq N. \end{cases} \quad (12)$$

Итоговое решение в задаче распознавания слов принимается в пользу слова X^* по критерию максимума μ_l .

Результаты экспериментальных исследований в задаче распознавания слов

Проведем сравнение методов ФДС (4)—(6), (12) оригинального НФКД (7)—(10), (12) и его модификации (7)—(12). В качестве словаря эталонов использовали перечень лекарств (1900 наименований). Для тестирования диктор произносил 200 слов/словосочетаний из словаря с четким выделением слогов (каждое слово произносили по 3 раза). В качестве эталонов брали десять гласных звуков этого же диктора по четырем реализациям (для каждой МРЕ). Оценка вероятности ошибки распознавания (в %) в зависимости от числа слогов представлена на рис. 2. Слово считается распознанным верно, если не существует ни одного слова с большей степенью принадлежности (12), чем истинное. Как

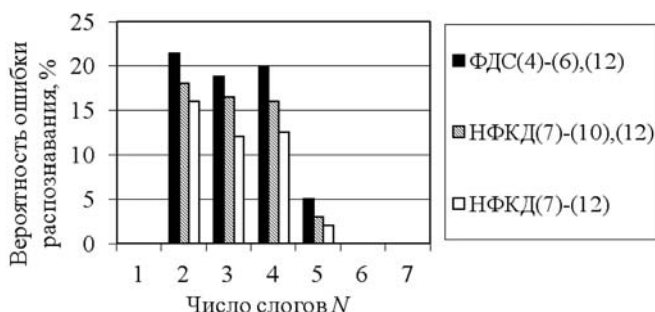


Рис. 2. Зависимость вероятности ошибки распознавания слов от числа слогов N

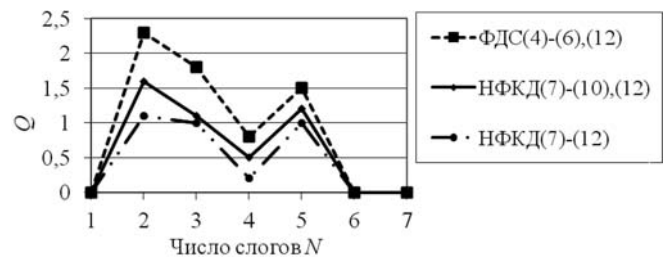


Рис. 3. Зависимость среднего числа слов Q , имеющих большую степень принадлежности, чем истинное слово, от числа слогов N

видно из этого рисунка, применение предложенного НФКД (7)—(12) позволяет увеличить точность распознавания. Например, вероятность ошибки распознавания для слов, состоящих из трех слогов, составляет 12 %, что на 4,5 % ниже ошибки распознавания для оригинального НФКД и на 6,75 % ниже ошибки распознавания для метода ФДС. Вероятность ошибки распознавания, усредненной по всем тестовым реализациям слов, для метода НФКД составляет 10,5 %, что на 3 и 5,5 % ниже аналогичного показателя для оригинального НФКД и ФДС соответственно.

На рис. 3 представлена зависимость среднего числа слов Q , имеющих большую степень принадлежности, чем истинное слово, от числа слогов. Из этого рисунка видно, что в среднем для предложенного НФКД число альтернативных слов на выходе алгоритма АРР на 0,2...1,2 слова ниже по сравнению с методами ФДС и оригинальным НФКД.

В заключительном эксперименте среднюю вероятность ошибки для метода НФКД сопоставляли с аналогичным показателем следующих известных систем АРР для русского языка:

- русскоязычной клиент-серверной версии системы Google Voice Search [17], качество дикторонезависимого распознавания которой считается в настоящее время очень высоким;
- системы фонемного распознавания Perpetuum M. (<http://jkit.narod.ru/index.htm>), в которой не требуется настройки на диктора, а словарь вводится в текстовом виде.

Указанные системы АРР работают с непрерывной речью, поэтому паузы между слогами дикторы не произносили. Запись речевого сигнала осуществлялась через внешний микрофон с функцией шумоподавления из гарнитуры A4Tech HS (отношение сигнал/шум 10 дБ). Для записи речевого сигнала использовался формат wav (моно, $F = 8$ кГц, 16 бит на отсчет), при этом из сигнала удаляли начальные и конечные паузы. Использовали два словаря: наименований лекарств (1900 фраз) и список городов с областями (1200 наименований). Для тестирования диктор произносил по 500 слов/словосочетаний из каждого словаря с четким выделением слогов, каждое слово он произносил по 2 раза. В качестве фонем-эталонных для методов ФДС и НФКД использовали фонетическую базу данных

Вероятность ошибки распознавания слов, %

Методы АРР	Словарь лекарств	Словарь городов с областями
Google Voice Search	65	18,9
Perpetuum M v. 3.01	48,6	53,1
ФДС (4)—(6), (12) (дикторонезависимый режим)	14,2	10,4
ФДС (4)—(6), (12) (дикторозависимый режим)	8,7	7,6
НФКД (7)—(12) (дикторонезависимый режим)	9,9	7,2
НФКД (7)—(12) (дикторозависимый режим)	5,2	4,9

из предыдущего эксперимента. Вероятность ошибки АРР приведена в таблице. Из этой таблицы хорошо видно, что точность распознавания для метода НФКД является наивысшей, что во многом обусловлено требованием к слоговому произношению и применением операций нечеткого пересечения (9) и объединения (11). Преимуществом метода НФКД по сравнению с рассмотренными аналогами является быстрая настройка под конкретного пользователя (пользователю достаточно произнести гласные звуки). Между тем качество популярной системы Google Voice Search для словаря лекарств оказалось неудовлетворительным, так как большая часть наименований лекарственных препаратов отсутствует в универсальном словаре системы.

Заключение

В последнее время теория нечетких множеств находит все большее применение в разнообразных задачах в области искусственного интеллекта [18]. В настоящей работе предложено обобщение метода НФКД [8] для автоматического распознавания гласной фонемы в слоге и изолированных слов в случае, если для каждой фонемы доступно несколько эталонных реализаций. Результаты экспериментальных исследований показывают, что использование предложенного алгоритма позволяет повысить качество распознавания по сравнению с методами ФДС и НФКД. Так, например, для предложенного алгоритма вероятность p_1 того, что не более одной МРЕ (при распознавании фонемы в слоге) имели более высокую степень принадлежности, чем истинная МРЕ равна 84 %, что на 2—4 % выше аналогичного показателя для методов ФДС и НФКД (см. рис. 1). Также для предложенного метода повысилась точность распознавания (см. рис. 2) и снизилось среднее число альтернативных слов на выходе алгоритма распознавания АРР (см. рис. 3) по сравнению с аналогичными показателями для оригинального метода НФКД и метода ФДС соответственно. Отмеченное повышение точности можно объяснить тем, что операция объединения (11) учитывает все эталонные реализации фонемы.

В дальнейшем при использовании операции пересечения нечетких множеств (9) степени принадлежности фонем, близких к истинной, остаются практически без изменения, а других фонем сильно занижаются. В таблице показано, что метод НФКД превосходит по точности распознавания известные системы АРР для русского языка, что связано с послоговым произношением и применением теории нечетких множеств. Таким образом, в работе предложен новый алгоритм распознавания слов (7)—(12), основанный на принципе минимума информационного рассогласования Кульбака—Лейблера и теории нечетких множеств. Этот алгоритм имеет высокие эксплуатационные характеристики.

Список литературы

1. Benesty J., Sondh M., Huang Y. (eds.). Springer Handbook of Speech Recognition. New York: Springer, 2008. P. 1159.
2. Rabiner L. R. A tutorial on hidden Markov models and selected applications in speech recognition // Proc. of the IEEE. Feb. 1989. Vol. 77, No. 2. P. 257—286.
3. Савченко А. В. Адаптивный алгоритм распознавания речи на основе метода фонетического декодирования слов в задаче голосового управления // Информационные технологии. 2013. № 4. С. 34—39.
4. Корсун О. Н., Яцко А. А., Финаев И. М., Чучупал В. Я. Экспериментальное исследование влияния акустических помех разных видов на результаты автоматического распознавания речевых команд // Наука и образование. 2013. № 1. С. 103—114. URL: <http://technomag.edu.ru/doc/508614.html>
5. Кипяткова И. С., Карпов А. А. Аналитический обзор систем распознавания русской речи с большим словарем // Труды СПИИРАН. 2010. № 12. С. 7—20.
6. Савченко В. В. Метод фонетического декодирования слов в задаче автоматического распознавания речи на основе принципа минимума информационного рассогласования // Известия вузов России. Радиоэлектроника. 2009. Вып. 5. С. 31—41.
7. Савченко Л. В., Савченко А. В. Алгоритм автоматического распознавания фонем на основе логики нечетких множеств в информационной метрике Кульбака—Лейблера // Вестник компьютерных и информационных технологий. 2013. № 3. С. 36—41.
8. Savchenko L. V., Savchenko A. V. Fuzzy Phonetic Decoding Method in a Phoneme Recognition Problem // Proc. of NOLISP—2013 Int. Conf., LNCS/LNAI. 2013. Vol. 7911. P. 176—183.
9. Савченко В. В. Фонема как элемент информационной теории восприятия речи // Известия вузов России. Радиоэлектроника. 2008. Вып. 4. С. 3—11.
10. Kullback S. Information Theory and statistics. New York: Dover Pub., 1997. P. 399.
11. Basseville M. Distance measures for signal processing and pattern recognition // Signal Processing. 1989. Vol. 18. P. 349—369.
12. Марпл С. Л.-мл. Цифровой спектральный анализ и его приложения. М.: Мир, 1990. С. 584.
13. Цыплихин А. И., Сорокин В. Н. Сегментация речи на кардинальные элементы // Информационные процессы. 2006. Т. 6, № 3. С. 177—207.
14. Zadeh L. A. Fuzzy Sets // Information Control. 1965. Vol. 8. P. 338—353.
15. Ronzhin A. L., Yusupov R. M., Li I. V., Leontieva A. B. Survey of Russian Speech Recognition Systems // Proc. of the International Conference SPECOM'2006, St. Petersburg, Russia. 2006. P. 54—60.
16. Козлов А. В., Саввина Г. В., Шелепов В. Ю. Система фонемного распознавания отдельно произносимых слов // Искусственный интеллект. 2003. № 1. С. 156—165.
17. Schuster M. Speech Recognition for Mobile Devices at Google. // Proc. of 11th Pacific Rim International Conference on Artificial Intelligence (Daegu, Korea). Lecture Notes and Computer Science. 2010. Vol. 6230. P. 8—10.
18. Jensen R., Cornelis C. Fuzzy-Rough Nearest Neighbour Classification // Transactions on Rough Sets. 2011. Vol. XIII. P. 56—72.

Оценивание параметров модели по нечетким случайным данным

Рассматриваются традиционные методы точечных оценок применительно к "гибридным" данным, которые являются одной из разновидностей нечетких случайных переменных. Методами теории нечетких линейных систем показано, что при обработке "гибридных" данных возникают "сильные—слабые" оценки точечных параметров. Приведены простейшие примеры обработки нечетких данных.

Ключевые слова: нечеткие системы, гибридные данные, метод наименьших квадратов, метод моментов, метод максимального правдоподобия

I. A. Mochalov, M. S. Khrisat

Estimation Parameter Model Using Fuzzy Random Data

The article deals with the traditional methods of point estimates with respect to the "hybrid" data, which are a type of fuzzy random variables. Using the theory of fuzzy linear systems shows that the treatment of "hybrid" data there are "strong—weak" point estimates of parameters. Is a simple example of processing of fuzzy data.

Keyword: fuzzy systems, hybrid data, the method of least squares, method of moments, maximum likelihood method

Введение

Теория нечетких множеств находит широкое применение в решении разнообразных прикладных задач управления. Одна из таких актуальных задач состоит в учете различного рода неопределенностей при обработке данных в целях идентификации параметров модели.

Целью статьи является применение традиционных методов математической статистики для обработки "гибридных" данных. Они появляются при интерпретации неопределенности в виде двух компонент: одна — случайная, а другая — нечеткая. Нечеткая обычно представляет собой параметры случайной компоненты в виде нечеткой переменной. Такое представление возникает, когда ошибка измерений трактуется случайной составляющей, а погрешность модели — как некая нечеткость.

Ниже будет показано, что при обработке "гибридных" данных возникают специфические проблемы. Можно полагать, что их решение составляют научную новизну данной статьи.

В качестве методов получения точечных оценок по "гибридным" данным будем использовать:

- метод моментов;
- метод максимального правдоподобия;
- метод наименьших квадратов.

1. Нечеткие вычисления [1]

1.1. Нечеткое число

Нечеткое число $x_H \in R^1$ определяется как отображение $r: R^1 \rightarrow [0; 1] \in R^1$, где $r(x)$ — функция принадлежности. Из-за отсутствия взаимной однозначности выделяются "левая" $\underline{r}(x)$ и "правая" $\bar{r}(x)$

ветви относительно $r(x) = 1$, каждая из которых определяет уже взаимно однозначное отображение. В теории нечетких множеств используется эквивалентная уровневая форма представления нечеткого числа, задаваемая в виде обратного отображения $r^{-1}(x): R^1 \supset [0; 1] \rightarrow R^1$. Для отображения $x(r)$, $r \in [0, 1]$ выделяются "нижняя" $\underline{x}(r)$ и "верхняя" $\bar{x}(r)$ ветви.

Таким образом, для нечеткого числа $x_H \in R^1$ используется цепочка эквивалентных представлений:

$$\begin{aligned} x_H \in R^1 &\Leftrightarrow r(x), r \in [0, 1] \Leftrightarrow \\ &\Leftrightarrow (\underline{r}(x); \bar{r}(x) | r \in R^1, \underline{r}, \bar{r}, \in [0, 1]) \Leftrightarrow \\ &\Leftrightarrow (\underline{x}(r); \bar{x}(r) | 0 \leq r \leq 1). \end{aligned}$$

Символ | (здесь и далее) означает "при условии". Относительно $r(x)$ должны выполняться следующие свойства:

- функция $r(x)$ полунепрерывна сверху;
- функция $\underline{r}(x)$ монотонно возрастает;
- функция $\bar{r}(x)$ монотонно убывает.

Кроме этого для $x(r)$ должно выполняться условие $\underline{x}(r) \leq \bar{x}(r)$. Если $x(r)$ имеет треугольную форму, то перечисленные свойства выполняются для остроугольного треугольника, однако не каждый тупоугольный треугольник может изображать нечеткое число (рис. 1).

Обычно для нечеткого числа x_H применяется обозначение

$$x_H \Leftrightarrow (\underline{x}(r), \bar{x}(r) | 0 \leq r \leq 1).$$

Арифметические операции * ("+", "-", "×", "÷") нечетких чисел x_H и y_H определяются соотношением

$$x_H * y_H = z_H \Leftrightarrow \max_{x_H * y_H = z_H} \min(r(x), r(y)).$$

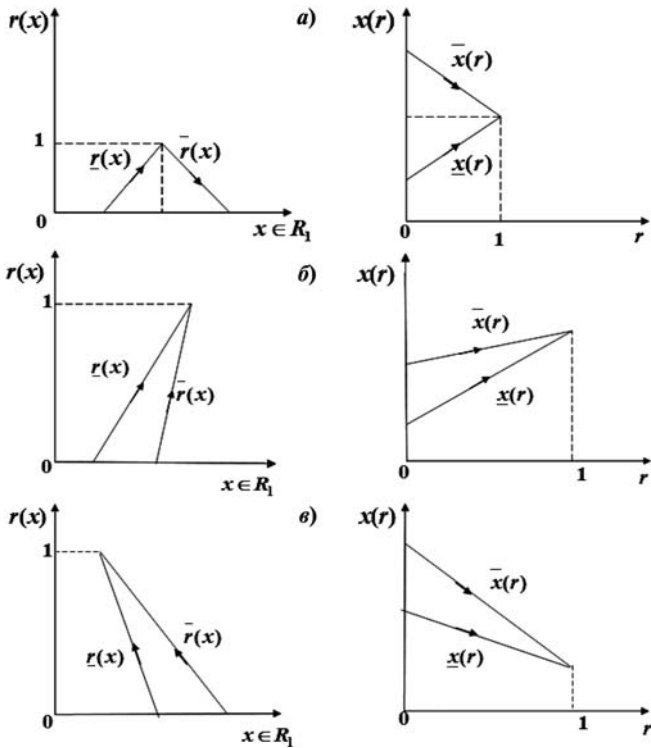


Рис. 1. Нечеткое число (а), не четкие числа (б, в)

Операции сравнения " \leq ", " \geq " следуют из определения [1]: имеем нечеткие числа x_H и y_H такие, что $x_H = (\underline{x}(r), \bar{x}(r)|0 \leq r \leq 1)$, $y_H = (\underline{y}(r), \bar{y}(r)|0 \leq r \leq 1)$, тогда $x_H \geq y_H$, если:

$$T(x_H) = \int_0^1 r[\underline{x}(r) + \bar{x}(r)]dr \geq \int_0^1 r[\underline{y}(r) + \bar{y}(r)]dr = T(y_H). \quad (1)$$

Рассмотрим совокупность $X_H = \{x_H\}$ нечетких чисел и зададим для ее элементов операции сложения (+), умножения (\times) на скаляр k , вычитания (-). Последняя операция эквивалентна заданию "нулевого" или обратного элемента.

Имеем:

- сложение (+) элементов x_{1H} и x_{2H} : $(x_{1H} + x_{2H}) = (\underline{x}_1(r) + \underline{x}_2(r); \bar{x}_1(r) + \bar{x}_2(r)|r \in [0; 1])$;
- умножение (\times) элемента x_H на скаляр $k \in R_1$:
 $k \times x_H = \begin{cases} k \times \underline{x}(r); k \times \bar{x}(r)|r \in [0; 1], k \geq 0; \\ k \times \bar{x}(r); k \times \underline{x}(r)|r \in [0; 1], k < 0; \end{cases}$
- существование элемента $r_2(x)$, обратного к $r_1(x)$, т. е. $r_2(x) = r_1(-x)$ или $r_1(x) + r_2(x)|_{r_2=r_1(-x)} \equiv r_1(x) + r_1(-x) = 0$ (рис. 2);
- ассоциативность (+): $(x_{1H} + x_{2H}) + x_{3H} = x_{1H} + (x_{2H} + x_{3H})$;
- дистрибутивность (\times): $(k_1 + k_2)x_{1H} = k_1x_{1H} + k_2x_{1H}$; $k(x_{1H} + x_{2H}) = kx_{1H} + kx_{2H}$.

В результате имеем совокупность X_H с введенными операциями над его элементами, удовлетворяющими перечисленным аксиомам, которая является

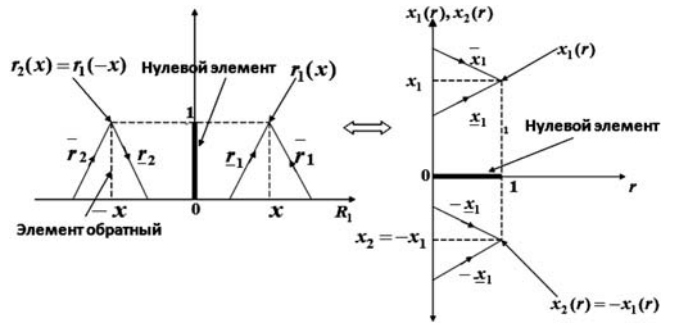


Рис. 2. Нечеткое число $r_2(x)$, обратное нечеткому числу $r_1(x)$

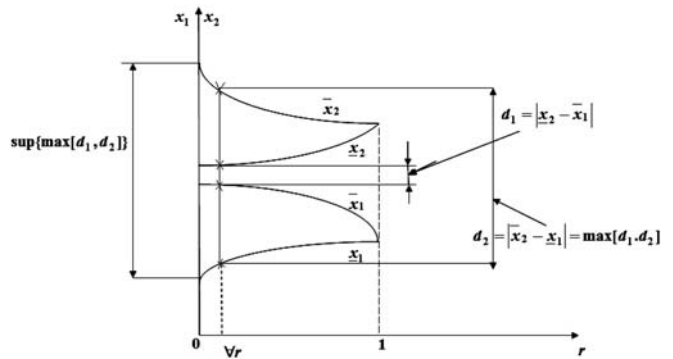


Рис. 3. Геометрическое представление метрики пространства X_H

векторным (линейным) пространством. В символической форме совокупность X_H записывается:

$$X_H = \underbrace{\{x_H\}}_{\text{совокупность нечетких чисел}} + \text{операции} \begin{pmatrix} "+" \\ "\times" \text{ на } k \\ "-" \end{pmatrix} \text{ над } x_{iH} +$$

$$+ \text{аксиомы операций} \begin{pmatrix} \text{коммутативность} \\ \text{ассоциативность} \\ \text{дистрибутивность} \end{pmatrix} =$$

= векторное (линейное) пространство.

В векторном пространстве X_H вводим метрику $D(\cdot)$ между любыми элементами x_{1H} , x_{2H} (рис. 3):

$$D(x_{1H}, x_{2H}) = \sup\{\max\{|\bar{x}_1 - \underline{x}_2|, |\underline{x}_1 - \bar{x}_2|\}\}.$$

Норму $\|\cdot\|_{X_H}$ пространства X_H определим в виде

$$\|(x_{1H} - x_{2H})\|_{X_H} = D(x_{1H}, x_{2H}).$$

Нетрудно показать, что последовательность x_{iH} будет сходиться в X_H относительно метрики D , поэтому векторное пространство X_H является полным и, соответственно, определяет банахово пространство подобно тому, как совокупность $C_0[a, b]$ непрерывных функций на $[a, b]$ с нормой $\sup_{t \in [a, b]} |x(t)|$, $C_1[a, b]$ с нормой $\sup_t |x(t)| + \sup_t |\dot{x}(t)|$, $C_2[a, b]$ и т. д. образуют банаховы пространства.

1.2 Нечеткая функция

Нечеткая функция $\varphi_H(x)$ определяется как отображение $\varphi_H: R^1 \rightarrow F = \{r(x)\}$, где F — совокупность функций принадлежности $r(x)$. Это отображение

параметризуется относительно $r \in [0, 1]$ и может быть представлено в виде

$$\varphi_H(x) = (\underline{\varphi}(x, r); \overline{\varphi}(x, r) | 0 \leq r \leq 1).$$

По аналогии с формулой (1) для нечеткой функции $\varphi_H(x)$ вводится критерий

$$T(\varphi_H(x)) = \int_0^1 r[\underline{\varphi}(x, r) + \overline{\varphi}(x, r)] dr.$$

Имеют место следующие утверждения:

- нечеткая функция $\varphi_H(x)$ монотонно возрастает (убывает), если для любых x_1 и x_2 выполняется $x_1 \leq x_2 \Rightarrow T(\varphi_H(x_1)) \leq T(\varphi_H(x_2))$ ($x_2 \leq x_1 \Rightarrow T(\varphi_H(x_2)) \leq T(\varphi_H(x_1))$);
- нечеткая функция $\varphi_H(x)$ непрерывна для $x \in [c, d] \subset R^1$, если $T(\varphi_H(x))$ непрерывна;
- нечеткая функция $\varphi_H(x)$ дифференцируема, если $T(\varphi_H(x))$ дифференцируема; производная от нечеткой функции $\varphi_H(x)$ равна $\dot{\varphi}_H(x) = \left(\frac{\partial}{\partial x} \underline{\varphi}(x, r); \right.$

$$\left. \frac{\partial}{\partial x} \overline{\varphi}(x, r) | 0 \leq x \leq 1 \right);$$

- нечеткая функция $\varphi_H(x)$ интегрируема по Риману, если $T(\varphi_H(x))$ интегрируема; интеграл от нечеткой функции $\varphi_H(x)$ равен

$$\int_c^d \varphi(x) dx = \left(\int_c^d \underline{\varphi}(x, r) dx; \int_c^d \overline{\varphi}(x, r) dx | 0 \leq x \leq 1 \right).$$

Приведенные утверждения показывают, что трудно сконструировать нечеткие аналоги основных структур классического математического анализа: максимум (минимум) нечеткой функции, нечеткие точки перегиба, нечеткие дифференциалы, касательные и т. д.

2. Нечеткие системы

линейных алгебраических уравнений (НСЛАУ)

Принято следующее определение НСЛАУ [2]: это уравнение относительно вектора X линейного типа:

$$AX = Y_H,$$

где $A = \{a_{ij}\}$, $i, j = \overline{1, n}$, — матрица с элементами a_{ij} из четких чисел, т. е. функция принадлежности $r_{ij}(a)$ элементов a_{ij} является одиночного (singleton) типа:

$$r_{ij}(a) = \text{singl}(a - a_{ij}) = \begin{cases} 1, & a = a_{ij}; \\ 0, & a \neq a_{ij}; \end{cases}$$

$Y_H^T = (y_{1H}, \dots, y_{nH})$ — вектор, имеющий нечеткие компоненты с заданными функциями принадлежности $r_k(y)$, $k = \overline{1, n}$.

В символической форме имеем: дано

$A, Y_H \xrightarrow{AX=Y_H} X$ — найти вектор с нечеткими компонентами.

Пример. НСЛАУ с размерностью (2×2) :

$$\begin{aligned} \begin{cases} 1 \cdot x_1 - 1 \cdot x_2 = y_{1H} \\ 1 \cdot x_1 + 2 \cdot x_2 = y_{2H} \end{cases} &\Leftrightarrow \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_{1H} \\ y_{2H} \end{pmatrix} \Rightarrow \\ &\Rightarrow X_H^T = (x_{1H}, x_{2H}) = \\ &= (\underline{x}_1(r), \overline{x}_1(r); \underline{x}_2(r), \overline{x}_2(r) | r \in [0; 1])_{(1 \times 4)}. \end{aligned}$$

Необходимо определить X_H^T . Очевидно, имеет место несогласованность размерностей переменных. Для ее устранения рассмотрим алгоритмы решения нечеткой системы в простейшем случае (1×1) :

$$ax = y_H,$$

a — четкое число, y_H — нечеткое число.

Случай 1. Пусть $a > 0$, тогда

$$ax = y_H \Rightarrow A = a_{(1 \times 1)} > 0,$$

поэтому при $y_H = (\underline{y}(r), \overline{y}(r) | r \in [0; 1])$ имеем:

$$\begin{cases} a \cdot \underline{x} = \underline{y}; \\ a \cdot \overline{x} = \overline{y}. \end{cases}$$

Если $\underline{x}, \overline{x}$ — решение системы, то $\underline{x} + \overline{x}$ тоже есть решение системы как соответствующая линейная комбинация, поэтому

$$\begin{aligned} \begin{cases} a \cdot \underline{x} + 0 \cdot \overline{x} = \underline{y} \\ 0 \cdot \underline{x} + a \cdot \overline{x} = \overline{y} \end{cases} &\Leftrightarrow \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} \underline{x} \\ \overline{x} \end{pmatrix} = \begin{pmatrix} \underline{y} \\ \overline{y} \end{pmatrix} \Rightarrow \\ &\Rightarrow |S| = a^2 \neq 0 \Rightarrow X = S^{-1} \cdot Y_H \Rightarrow \\ &\Rightarrow \begin{pmatrix} \underline{x} \\ \overline{x} \end{pmatrix} = \begin{pmatrix} a^{-1} & 0 \\ 0 & a^{-1} \end{pmatrix} \begin{pmatrix} \underline{y} \\ \overline{y} \end{pmatrix} \Rightarrow \\ &\Rightarrow x(a^{-1} \cdot \underline{y}(r), a^{-1} \cdot \overline{y}(r) | r \in [0; 1]), a > 0. \end{aligned}$$

Случай 2. Пусть $a < 0$, для простоты $a = -1$, тогда

$$a \cdot x |_{a=-1} = y_H \Rightarrow -x = y \Leftrightarrow x = -y.$$

Из геометрических представлений (рис. 4) получаем

$$\begin{aligned} -y_H = \begin{cases} \overline{x}^r + \underline{y} \equiv 0 \\ \underline{x} + \overline{y}_r \equiv 0 \end{cases} &\Leftrightarrow \begin{cases} (-1) \cdot \overline{x} = \underline{y} \\ 1 \cdot \underline{x} = \overline{y} \end{cases} \Leftrightarrow \\ \begin{cases} 0 \cdot \underline{x} + 1(-\overline{x}) = \underline{y} \\ 1 \cdot \underline{x} + 0(-\overline{x}) = -\overline{y} \end{cases} &\Leftrightarrow \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}_S \underbrace{\begin{pmatrix} \underline{x} \\ -\overline{x} \end{pmatrix}}_X = \underbrace{\begin{pmatrix} \underline{y} \\ -\overline{y} \end{pmatrix}}_{Y_H}. \end{aligned}$$

Если теперь $a < 0$ и $a \neq -1$, то

$$\begin{aligned} a \cdot x = y_H &\Leftrightarrow a(-x) = y_H \Leftrightarrow ax = -y_H \Leftrightarrow \\ \Leftrightarrow \begin{cases} 0 \cdot \underline{x} + a(-\overline{x}) = \underline{y} \\ a \cdot \underline{x} + 0(-\overline{x}) = -\overline{y} \end{cases} &\Leftrightarrow \underbrace{\begin{pmatrix} 0 & a \\ a & 0 \end{pmatrix}}_S \underbrace{\begin{pmatrix} \underline{x} \\ -\overline{x} \end{pmatrix}}_X = \underbrace{\begin{pmatrix} \underline{y} \\ -\overline{y} \end{pmatrix}}_{Y_H} \Rightarrow \\ &\Rightarrow |S| = -a^2 \neq 0 \Rightarrow X = S^{-1} \cdot Y_H \Rightarrow \\ &\Rightarrow \begin{pmatrix} \underline{x} \\ -\overline{x} \end{pmatrix} = \begin{pmatrix} 0 & a^{-1} \\ a^{-1} & 0 \end{pmatrix} \begin{pmatrix} \underline{y} \\ -\overline{y} \end{pmatrix} \Rightarrow \\ &\Rightarrow x = (a^{-1} \cdot (-\overline{y}(r)), a^{-1} \cdot (\underline{y}(r)) | r \in [0; 1]), a < 0. \end{aligned}$$

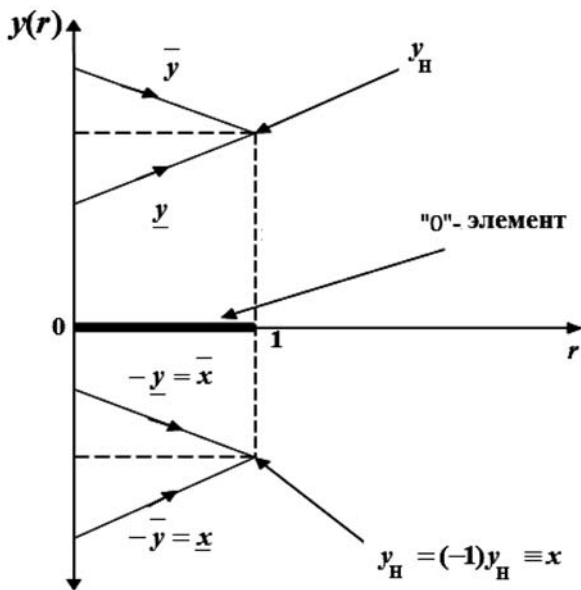


Рис. 4. Геометрическое представление "0"-го элемента

После объединения результатов случаев 1, 2 будем иметь следующий алгоритм решения НСЛАУ. По заданной НСЛАУ:

$$A_{(n \times n)} \cdot X_{(n \times 1)} = Y_{H(n \times 1)}$$

с заданной матрицей $A_{(n \times n)}$ составляется расширенная система

$$S_{(2n \times 2n)} \cdot X_{(2n \times 1)} = Y_{H(2n \times 1)}$$

с матрицей $S_{(2n \times 2n)}$; $X^T = (\underline{X} \ ; \ -\bar{X})$; $Y_H^T = (\underline{Y} \ ; \ -\bar{Y})$. Здесь S — блочная матрица

$$S = \begin{pmatrix} B & \vdots & C \\ \vdots & \ddots & \vdots \\ C & \vdots & B \end{pmatrix}_{(2n \times 2n)},$$

где блоки B и C находятся по матрице A : B состоит из положительных элементов A , а отрицательные элементы A заменяются соответственно нулями; C состоит из модулей отрицательных элементов A , а положительные элементы A заменяются соответственно нулями; $A = B - C \Rightarrow C = B - A$.

При решении расширенной системы возникают следующие проблемы:

1. При каких условиях $|S| \neq 0$?
2. Могут ли компоненты вектора X иметь функцию принадлежности в уровневой форме в виде тупоугольного треугольника, т. е. отличаться от нечеткого числа? Если такая ситуация возможна, то что тогда необходимо понимать под решением расширенной системы?

Ответ на вопрос по п. 1 дается с помощью следующей теоремы.

Теорема 1. $|S| \neq 0 \Leftrightarrow |B - C| \neq 0, |B + C| \neq 0$.

Проблема: если S не вырождена ($|S| \neq 0$), то, как вычислить S^{-1} ?

Теорема 2. Если S^{-1} существует, то она должна иметь такую же структуру, что S :

$$S^{-1} = \begin{pmatrix} D & \vdots & E \\ \vdots & \ddots & \vdots \\ E & \vdots & D \end{pmatrix}.$$

Теорема 3. В условиях теоремы 2 имеем:

$$D = 0,5[(B + C)^{-1} + (B - C)^{-1}];$$

$$E = 0,5[(B + C)^{-1} - (B - C)^{-1}].$$

Таким образом, если S не вырождена ($(B + C)$, $(B - C)$ — не вырождены), то X — единственный вектор и он равен

$$X = S^{-1} \cdot Y_H.$$

Однако X может быть неподходящим нечетким вектором в том смысле, что функции принадлежности треугольного типа некоторых компонент вектора X могут иметь в основании треугольника углы больше 90° (см. рис. 1, б, в), т. е. не являются нечеткими числами. Например, пусть для фиксированного i получено:

$$x_i = (\underline{x}_i(r), \bar{x}_i(r) | r \in [0; 1]),$$

которое не является нечетким числом, тогда проводится замена (рис. 5):

$$x_i \rightarrow u_i,$$

$$u_i = (\underline{u}_i(r) = \min\{\underline{x}_i, \bar{x}_i, \underline{x}_i(r=1)\};$$

$$\bar{u}_i(r) = \max\{\underline{x}_i, \bar{x}_i, \underline{x}_i(r=1)\} | r \in [0; 1]).$$

После такой замены

$$X \rightarrow U.$$

Решение U называют "слабым" (weak [week]-англ.) нечетким решением НСЛАУ. Если же все компоненты вектора X являются нечеткими числами, т. е. углы при основании треугольной функции принадлежности меньше 90° (см. рис. 1, а), тогда найденный вектор X называют "сильным" (strong-англ.) нечетким решением НСЛАУ.

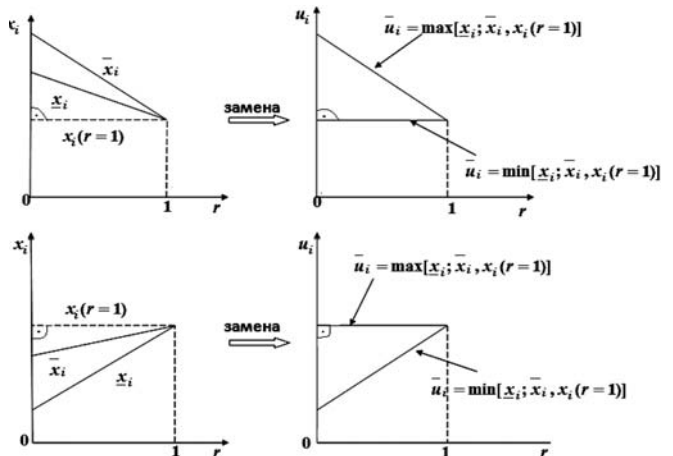


Рис. 5. Замена нечеткой компоненты (числа) x_i на его "слабую" компоненту u_i

Теорема 4. Для того чтобы решение X НСЛАУ

$$X = S^{-1} \cdot Y_H$$

имело "сильное" нечеткое решение, необходимо и достаточно, чтобы элементы $(S^{-1})_{ij} \geq 0, \forall i, j$.

Имеет место следующий факт. Матрица A НСЛАУ не вырождена, т. е. $|A| \neq 0$, однако матрица S расширенной НСЛАУ может быть вырождена, т. е. $|S| = 0$. В символической форме

$$|A| \neq 0 \Rightarrow |S| \neq 0.$$

Пример. Имеем НСЛАУ

$$\begin{cases} 1 \cdot x_1 - 1 \cdot x_2 = y_{1H} \\ 1 \cdot x_1 + 1 \cdot x_2 = y_{2H} \end{cases}, |A| = \begin{vmatrix} 1 & -1 \\ 1 & 1 \end{vmatrix} = 1 - (-1) = 2 \neq 0.$$

Расширенная НСЛАУ имеет матрицу S :

$$S = \begin{pmatrix} \underbrace{B} & & \underbrace{C} \\ 1 & 0 & \vdots & 0 & 1 \\ 1 & 1 & \vdots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 1 & \vdots & 1 & 0 \\ 0 & 0 & \vdots & 1 & 1 \\ \underbrace{C} & & & \underbrace{B} & \end{pmatrix}.$$

Из теории блочных матриц ($|B| \neq 0$) имеем:

$$|S| = \begin{vmatrix} \underbrace{B} & \vdots & \underbrace{C} \\ \vdots & \dots & \vdots \\ \underbrace{C} & \vdots & \underbrace{B} \end{vmatrix} = |B| \cdot |B - C \cdot B^{-1} \cdot C|,$$

где $|B| = \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} = 1 \neq 0$; $|B - C \cdot B^{-1} \cdot C| =$
 $= \begin{vmatrix} 1 & 0 \\ \frac{1}{B} & \frac{1}{C} \end{vmatrix} - \begin{vmatrix} 0 & 1 \\ \frac{0}{0} & \frac{0}{0} \end{vmatrix} \cdot \begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} \cdot \begin{vmatrix} 0 & 1 \\ \frac{0}{0} & \frac{0}{0} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} = 0.$

Таким образом, $|A| \neq 0$, однако ($|S| = 0$), т. е. расширенная НСЛАУ не имеет единственного решения. В этом случае НСЛАУ решается в соответствии с общей теорией решения систем по методу Гаусса путем приведения S к ступенчатому виду:

$$\begin{pmatrix} 1 & 0 & 0 & 1 & Y_1 \\ 0 & 1 & 0 & 0 & Y_2 \\ 0 & 1 & 1 & 0 & Y_3 \\ 0 & 0 & 1 & 1 & Y_4 \end{pmatrix} \sim \begin{pmatrix} \dots & \dots & \dots & Y_1 \\ 0 & -1 & 0 & 1 & Y_1 - Y_2 \\ \dots & \dots & \dots & Y_3 \\ \dots & \dots & \dots & Y_4 \end{pmatrix} \sim$$

$$\begin{pmatrix} \dots & \dots & \dots & Y_1 \\ \dots & \dots & \dots & Y_1 - Y_2 \\ 0 & 0 & 1 & 1 & Y_1 - Y_2 + Y_3 \\ 0 & 0 & 1 & 1 & Y_4 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 1 & Y_1 \\ 1 & -1 & 0 & 0 & Y_1 - Y_2 \\ 0 & 1 & 1 & 0 & Y_1 - Y_2 + Y_3 \\ 0 & 0 & 1 & 1 & Y_1 - Y_2 + Y_3 - Y_4 \end{pmatrix}.$$

Здесь для простоты обозначений вектор $Y_H^T = (\underline{y}_1, \underline{y}_2, -\bar{y}_1, -\bar{y}_2)$ заменен на вектор $Y^T = (Y_1, Y_2, Y_3, Y_4)$.

Матрица S приведена к ступенчатому виду S_3 , поэтому возможны следующие варианты решения НСЛАУ.

Вариант 1. Пусть имеем:

$$Y_1 - Y_2 + Y_3 - Y_4 \neq 0,$$

т. е. векторы линейно независимые, но тогда из последней строки матрицы S_3 имеем: $0|Y_1 - Y_2 + Y_3 - Y_4 \neq 0 \Rightarrow 0 \neq 0 \Rightarrow$ противоречие, поэтому НСЛАУ несовместна, т. е. не имеет нечеткого решения. Пусть для определенности имеем:

$$\begin{aligned} y_{1H} &= (\underline{y}_1 = r; \bar{y}_1 = 2 - r | r \in [0; 1]); \\ y_{2H} &= (\underline{y}_2 = 3 + r; \bar{y}_2 = 4 | r \in [0; 1]). \end{aligned}$$

Тогда в терминах компонент $Y_i, i = \overline{1, 4}$, вектора Y получим:

$$Y^T = (\underbrace{y_1 = r}_{Y_1}; \underbrace{y_2 = 3 + r}_{Y_2}; \underbrace{-\bar{y}_1 = r - 2}_{Y_3}; \underbrace{\bar{y}_2 = -4}_{Y_4}).$$

Из линейной независимости компонент $Y_i, i = \overline{1, 4}$, имеем:

$$Y_1 - Y_2 + Y_3 - Y_4 \neq 0 \Rightarrow Y_1 - Y_2 + Y_3 \neq Y_4.$$

После подстановки соответствующих значений Y_i получим:

$$\begin{aligned} Y_1|_{y_1=r} - Y_2|_{y_2=3+r} + Y_3|_{-\bar{y}_1=r-2} \neq Y_4|_{-\bar{y}_2=-4} \Rightarrow \\ \Rightarrow r - (3 + r) + r - 2 \neq -4 \Rightarrow \frac{r-5}{Y_1 - Y_2 + Y_3} \neq \frac{4}{Y_4}. \end{aligned}$$

Это означает, что заданная НСЛАУ с $|A| \neq 0, |S| = 0$ при заданных y_{1H}, y_{2H} не имеет нечеткого решения.

Вариант 2. Пусть имеем:

$$Y_1 - Y_2 + Y_3 - Y_4 = 0,$$

т. е. векторы линейно зависимы, тогда из последней строки матрицы S_3 имеем:

$$0|Y_1 - Y_2 + Y_3 - Y_4 = 0 \Rightarrow Y_1 - Y_2 + Y_3 = Y_4.$$

Пусть теперь

$$\begin{aligned} y_{1H}^* &= (\underline{y}_1 = r; \bar{y}_1 = 2 - r | r \in [0; 1]); \\ y_{2H}^* &= (\underline{y}_2 = 3 + r; \bar{y}_2 = r - 5 | r \in [0; 1]), \end{aligned}$$

тогда вектор Y равен

$$Y = (\underbrace{y_1 = r}_{Y_1}; \underbrace{y_2 = 3 + r}_{Y_2}; \underbrace{-\bar{y}_1 = r - 2}_{Y_3}; \underbrace{\bar{y}_2 = 5 - r}_{Y_4}).$$

Из условия линейной зависимости имеем:

$$\begin{aligned} Y_1|_{y_1=r} - Y_2|_{y_2=3+r} + Y_3|_{-\bar{y}_1=r-2} = Y_4|_{\bar{y}_2=5-r} \Rightarrow \\ \Rightarrow r - (3 + r) + (r - 2) = \frac{r-5}{Y_1 - Y_2 + Y_3} = \frac{r-5}{Y_4}. \end{aligned}$$

Это означает, что НСЛАУ с $|A| \neq 0, |S| = 0$ при заданных y_{1H}^*, y_{2H}^* имеет бесконечное число решений.

Найдем одно из возможных решений. Пусть задано:

$$\underline{x}_1 = \alpha + \beta \cdot r | r \in [0; 1], \alpha, \beta \in R_1,$$

тогда находим $\bar{x}_1, \underline{x}_2, \bar{x}_2$ из системы:

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}}_S \underbrace{\begin{pmatrix} \underline{x}_1 = \alpha + \beta \cdot r \\ \underline{x}_2 = u_1 + u_2 \cdot r \\ -\bar{x}_1 = u_3 + u_4 \cdot r \\ -\bar{x}_2 = u_5 + u_6 \cdot r \end{pmatrix}}_X \underbrace{\begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{pmatrix}}_{Y_H} = \underbrace{\begin{pmatrix} y_1 = r \\ y_2 = 3 + r \\ -y_3 = r - 2 \\ -y_4 = -r + 5 \end{pmatrix}}_{Y_H} \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{pmatrix}.$$

При заданных α, β неизвестные u_1, \dots, u_6 находятся из соответствующей системы. Например, одно из уравнений этой системы имеет вид:

$$1 \cdot X_1 + 0 \cdot X_2 + 0 \cdot X_3 + 1 \cdot X_4 \equiv Y_1 \Leftrightarrow \Leftrightarrow \underbrace{\alpha + \beta \cdot r}_{\bar{x}_1} + \underbrace{u_5 + u_6 \cdot r}_{\bar{x}_4} \equiv 0 + 1 \cdot r \Rightarrow$$

$$\Rightarrow \begin{cases} \alpha + u_5 = 0 \\ \beta + u_6 = 1 \end{cases} \Rightarrow u_5 = -\alpha; u_6 = 1 - \beta \Rightarrow -\bar{x}_2 = -\alpha + (1 - \beta) \cdot r \Rightarrow \bar{x}_2 = \alpha + (1 - \beta) \cdot r.$$

После аналогичных вычислений получим:

$$\underline{x}_2 = (3 - \alpha) + (1 - \beta) \cdot r; \bar{x}_1 = (5 - \alpha) - \beta \cdot r.$$

Таким образом, одно из возможных решений из бесконечного их числа равно:

$$\begin{aligned} x_{1H} &= (\underline{x}_1 = \alpha + \beta \cdot r; \\ \bar{x}_1 &= (5 - \alpha) - \beta \cdot r | r \in [0; 1]; \\ x_{2H} &= (\underline{x}_2 = (3 - \alpha) + (1 - \beta) \cdot r; \\ \bar{x}_2 &= \alpha + (\beta - 1) \cdot r | r \in [0; 1], \end{aligned}$$

однако при этом должны выполняться условия:

$$\underline{x}_1(r=1) = \bar{x}_1(r=1) \text{ и } \underline{x}_2(r=1) = \bar{x}_2(r=1).$$

Из первого из них имеем:

$$\frac{\alpha + \beta \cdot r}{\underline{x}_1} |_{r=1} = \frac{(5 - \alpha) - \beta \cdot r}{\bar{x}_1} |_{r=1} \Rightarrow \alpha + \beta = (5 - \alpha) - \beta \Rightarrow 2(\alpha + \beta) = 5 \Rightarrow \alpha + \beta = 2,5 \Rightarrow \Rightarrow \beta = 2,5 - \alpha.$$

Из этого следует:

$$\begin{aligned} x_{1H} |_{\beta = 2,5 - \alpha} &= (\underline{x}_1 = \alpha + (2,5 - \alpha)r; \\ \bar{x}_1 &= (5 - \alpha) - (2,5 - \alpha)r | r \in [0; 1]; \end{aligned}$$

$$\begin{aligned} x_{2H} |_{\beta = 2,5 - \alpha} &= (\underline{x}_2 = (3 - \alpha) + (\alpha + 1,5)r; \\ \bar{x}_2 &= \alpha + (1,5 - \alpha)r | r \in [0; 1]). \end{aligned}$$



Рис. 6. Диапазон изменения параметров α, β правой части НСЛАУ для получения "сильных" решений

Типы решений НСЛАУ

НСЛАУ $A_{(n \times n)} \cdot X_{(n \times 1)} = Y_{(n \times 1)}$	Расширенная НСЛАУ $S_{(2n \times 2n)} \cdot X_{(2n \times 1)} = Y_{(2n \times 1)}$	Тип решения НСЛАУ
$ A \neq 0$	$ S \neq 0$	"Сильное" или "слабое" решение
$ A \neq 0$	$ S = 0$	Не имеет решения или имеет "слабое/сильное" решение

Пусть ищем "сильное" нечеткое решение, тогда:

$$\begin{cases} \underline{x}_1(r=0) \leq \bar{x}_1(r=0) \Leftrightarrow \alpha \leq 5 - \alpha \Rightarrow 2\alpha \leq 5 \Rightarrow \alpha \leq 2,5; \\ \underline{x}_2(r=0) \leq \bar{x}_2(r=0) \Leftrightarrow 3 - \alpha \leq \alpha \Rightarrow \alpha \geq 1,5, \end{cases}$$

откуда

$$1,5 \leq \alpha \leq 2,5$$

и с учетом $\beta = 2,5 - \alpha$ получаем $0 \leq \beta \leq 1$.

Найденный диапазон изменения параметров α, β обеспечивает получение "сильных" нечетких решений среди их бесконечного числа (рис. 6). Результаты решения НСЛАУ сведены в таблицу.

3. Постановка задачи оценивания для гибридных данных

Одной из задач традиционной математической статистики является оценка неизвестных четких параметров выбранной параметрической модели. В этом случае полагается, что закон распределения $f(x, \theta)$ четкого случайного вектора "x" генеральной совокупности задан, а четкий вектор "theta" параметров является неизвестным. Задача оценивания в этом случае формулируется следующим образом. Необходимо найти четкий вектор оценки "theta" по выборке четких случайных данных $x = (x_1; x_2; \dots; x_n)$, полученной из генеральной совокупности случайным образом.

Подобная задача оценивания имеет место в нечетком случае, когда x — нечеткий случайный вектор с θ нечетким вектором параметров. Такие нечеткие случайные данные принято называть гибридными данными (ГД) [3]. В терминах вероятностных пространств полагается, что метрика в этом пространстве задается в виде евклидовой метрики.

Далее будем рассматривать задачу оценивания для ГД. В этом случае для оценивания применяются традиционные методы математической статистики, а для нахождения функции принадлежности полученной оценки используются приемы и терминология [2]. В результате этого появляются "сильные—слабые" оценки.

Имеется многомерная нечеткая плотность $f(x_H, \theta_H)$ генеральной совокупности, где x_H, θ_H — нечеткие векторы. Необходимо по ГД $x_{1H}, x_{2H}, \dots, x_{nH}$ случайной выборки из генеральной совокупности найти вектор нечеткой оценки $\hat{\theta}_H$ для нечеткого вектора параметров θ_H . Для решения сформулированной выше задачи будем использовать методы: моментов; максимального правдоподобия; наименьших квадратов [4].

4. Статистические методы обработки гибридных данных

4.1. Метод моментов

Имеется нечеткая случайная выборка $x_H = (x_{1H}, x_{2H}, \dots, x_{nH})$ из генеральной совокупности, распределение которой $f(x_H, \theta_H)$ известно с точностью до вектора нечетких параметров $\theta_H = (\theta_{1H}, \theta_{2H}, \dots, \theta_{rH})$ в виде нечетких чисел с заданными функциями принадлежности треугольного типа, которые в уровневой форме имеют вид:

$$\theta_{iH}(\theta) = (\underline{\theta}_i(r), \overline{\theta}_i(r) | r \in [0; 1]), i = \overline{1, r}.$$

Необходимо найти оценку нечеткого параметра θ_H по случайной выборке. Полагается, что у нечеткой случайной величины x существуют первые r нечетких начальных моментов $m_{iH}(\theta)$, зависящих от нечеткого вектора θ_H :

$$m_{iH}(\theta) = \int_{-\infty}^{\infty} x^i \cdot f(x_H, \theta_H) dx, i = \overline{1, r}.$$

Вместе с тем, имеются выборочные нечеткие начальные моменты:

$$\hat{m}_{iH} = n^{-1} \sum_{k=1}^n x_{kH}^i, i = \overline{1, n}.$$

Аналогичные переменные могут быть введены для нечетких центральных моментов $\mu_{iH}(\theta)$ и их выборочных нечетких аналогов $\hat{\mu}_{iH}$.

В соответствии с концепцией метода моментов нечеткие параметры находятся из системы нечетких уравнений:

$$m_{iH}(\theta) = \hat{m}_{iH}, i = \overline{1, r},$$

которая в общем случае является нелинейной относительно θ .

Пример. Пусть нечеткая случайная величина X имеет экспоненциальное распределение с неизвестным нечетким параметром λ_H :

$$f(x, \lambda_H) = \lambda_H e^{-\lambda_H x}, x \geq 0.$$

Задана нечеткая случайная выборка: $x_{1H}, x_{2H}, \dots, x_{nH}$.

Необходимо найти нечеткую оценку $\hat{\lambda}_H$ неизвестного параметра λ_H по методу моментов. Для экспоненциального распределения имеем: $m_{iH} = \lambda_H^{-1}$ [4].

Первый нечеткий выборочный начальный момент \hat{m}_{iH} равен $\hat{m}_{1H} = n^{-1} \sum_{i=1}^n x_{iH}$. Откуда имеем:

$$\hat{\lambda}_H = (n^{-1} \sum_{i=1}^n x_{iH})^{-1}.$$

Для упрощения дальнейших вычислений переходим к обратной величине $\hat{T}_H = \lambda_H^{-1} = n^{-1} \sum_{i=1}^n x_{iH}$.

Положим:

$$x_{iH}(r) = (\underline{x}_i(r) = (x_i - \beta) + \beta r; \overline{x}_i(r) = (x_i + \alpha) - \alpha r | r \in [0; 1]), \alpha > 0, \beta > 0, \alpha > \beta, \alpha, \beta -$$

параметры, тогда:

$$\hat{T}_H = (\hat{T}(r) = n^{-1} \sum_{i=1}^n [(x_i - \beta) + \beta r],$$

$$\overline{\hat{T}}(r) = n^{-1} \sum_{i=1}^n [(x_i + \alpha) - \alpha r] | r \in [0; 1]),$$

откуда:

$$\hat{T}_H = (\hat{T}(r) = (T_{cp} - \beta) + \beta r,$$

$$\overline{\hat{T}}(r) = (T_{cp} + \alpha) - \alpha r | r \in [0; 1]), T_{cp} = \sum_{i=1}^n x_i$$

Для того чтобы \hat{T}_H было нечетким числом, должны выполняться неравенства [2]

$$\begin{cases} (T_{cp} + \alpha) - \alpha r \leq 0; \\ (T_{cp} - \beta) + \beta r \geq 0, (T_{cp} + \alpha) - \alpha r \geq (T_{cp} - \beta) + \beta r, \\ \forall r \in [0; 1], \end{cases}$$

которые справедливы всегда, так как $\alpha > 0, \beta > 0$.

Это означает, что нечеткая оценка \hat{T}_H является

"сильной". Расчеты показывают, что $\hat{\lambda}_H$ также является "сильной" нечеткой оценкой.

4.2. Метод максимального правдоподобия

Вектор нечеткой оценки $\hat{\theta}_H$ находится из условия:

$$\max_{\theta_H} \ln L(x_H, \theta_H),$$

где $L(x_H, \theta_H) = \prod_{i=1}^n f(x_{iH}, \theta_H)$ — функция правдоподобия. Это приводит к нечеткой системе уравнений относительно нечетких компонент θ_{kH} вектора θ_H :

$$\frac{\partial \ln L(x_H, \theta_H)}{\partial \theta_H} = 0, \quad k = \overline{1, r}.$$

Пример. В теории управления достаточно часто используют нечувствительные, или робастные, нечеткие системы [5], которые мало чувствительны к внешним возмущениям. Аналогичные системы существуют при нечетком оценивании, когда возникает задача фильтрации аномальных измерений. Одним из способов их описания состоит в использовании для этих целей нечеткого распределения Лапласа [4].

Пусть задана нечеткая случайная выборка $x = (x_{1H}, x_{2H}, \dots, x_{nH})$ с независимыми компонентами, которые имеют плотность

$$f(x_{iH}, \theta_H) = 2^{-1/2} \cdot \sigma \exp \left\{ -\frac{2^{1/2} |x_{iH} - \theta_H|}{\sigma} \right\}, \quad i = \overline{1, n}.$$

Здесь $\theta_H = Ex$ — нечеткое математическое ожидание, подлежащее оцениванию; σ — заданная четкая константа.

Имеем:

$$\begin{aligned} f(x_H, \theta_H) &= \prod_{i=1}^n f(x_{iH}, \theta_H) = \\ &= \sigma^{-n} 2^{-n/2} \exp \left\{ -\frac{\sqrt{2} \sum_{i=1}^n |x_{iH} - \theta_H|}{\sigma} \right\} \end{aligned}$$

— функция правдоподобия нечетких переменных.

Логарифмическая функция правдоподобия имеет вид

$$\begin{aligned} \ln f(x_H, \theta_H) &= -n \ln \sigma - 0,5 n \ln 2 - 2^{-1/2} \sigma^{-1} \sum_{i=1}^n |x_{iH} - \theta_H| = \\ &= a + b \left(\sum_1 + \sum_2 \right), \end{aligned}$$

где

$$\sum_1 = \sum_{i=1}^{n_1} |x_{iH} - \theta_H|, \quad x_{iH} > \theta_H;$$

$$\sum_2 = \sum_{i=1}^{n_2} |x_{iH} - \theta_H|, \quad x_{iH} < \theta_H;$$

$$a = -n \ln \sigma - 0,5 n \ln 2; \quad b = -2^{-1/2} \sigma^{-1}; \quad n_1 + n_2 = n.$$

Нечеткое уравнение для нахождения $\hat{\theta}_H$ имеет вид

$$\begin{aligned} \max_{\theta_H} \ln f(x_H, \theta_H) &\Leftrightarrow \frac{\partial}{\partial \theta_H} \left[a + b \left(\sum_1 \theta_H + \sum_2 \theta_H \right) \right] = \\ &= 0 \Leftrightarrow b \cdot (-n_1 + n_2) = 0 \Leftrightarrow n_1 = n_2. \end{aligned}$$

Это означает, что

$$\hat{\theta}_H = \text{Me} \{ x_{iH} \}_{i=1}^n,$$

где Me — символ медианы совокупности нечетких переменных x_{iH} , $i = \overline{1, n}$.

Медиана нечеткой совокупности находится аналогичным способом, существующим для четких переменных. В этом случае отношение порядка ">", "<", "=" для нечетких переменных, определенные в работе [1], используются для построения нечеткого вариационного ряда:

$$x_{(1)H}, x_{(2)H}, \dots, x_{(n)H}.$$

В результате получим:

$$\begin{aligned} \hat{\theta}_H &= \text{Me} \{ x_{iH} \}_{i=1}^n = \\ &= \begin{cases} 0,5 [x_{(2m)H} + x_{(2m+1)H}], & n - \text{четное}; \\ x_{(2m+1)H}, & n - \text{нечетное}. \end{cases} \end{aligned}$$

Пусть n — нечетное, тогда $\hat{\theta}_H$ является "сильной" оценкой, так как по условию задачи полагается, что все элементы $\{ x_{iH} \}_{i=1}^n$ имеют функции принадлежности в виде нечетких чисел.

Очевидно, что при n четном также получается "сильная" оценка.

4.3. Метод наименьших квадратов (МНК)

В соответствии с МНК нечеткие случайные данные y_{iH} , $i = \overline{1, m}$, связаны между собой линейной моделью:

$$y_H(t) = \sum_{i=1}^n x_{iH} f_i(t) + e_H(t), \quad t \in [0, T] \in R_1,$$

где e_H — нечеткая случайная переменная с симметричной плотностью и параметрами:

$$Ee_H(0) = 0; \quad De_H(t) = \sigma^2 \cdot I,$$

I — единичная матрица, σ^2 — заданная четкая константа; нечеткость $e_H(t)$ задается с помощью функции принадлежности $r(e_H)$ треугольного типа; $f_i(t)$, $i = \overline{1, n}$, — заданные базисные функции модели; x_{iH} , $i = \overline{1, n}$, — неизвестные нечеткие параметры модели, подлежащие определению по m нечетким случайным измерениям:

$$Y_H^T = (y_H(t_1), y_H(t_2), \dots, y_H(t_m)) = (y_{1H}, y_{2H}, \dots, y_{mH}),$$

которые получены в моменты времени t_1, t_2, \dots, t_m ; $m > n$ (число измерений m больше числа n неизвестных параметров модели).

Нечеткий вектор оценок находится из условия

$$\min_x e_H^T \cdot e_H = \min_x \|Y_H - FX\|_{E_H}^2,$$

где $F = (f_i(t_j))_{m \times n}$ — прямоугольная матрица.

Это приводит к необходимости решения нечеткой линейной системы [2]: $AX = U_H$, где $A = (f_i, f_j)$,

$i, j = n$, — квадратная матрица $\dim A = (n \times n)$; $U_H^T = (u_{1H} = (f_1, Y_H); u_{2H} = (f_2, Y_H), \dots, u_{nH} = (f_n, Y_H))$ — вектор нечетких переменных, $u_{iH} = (f_i, Y_H)$, $i = \overline{1, n}$, — скалярное произведение векторов f_i, Y_H в E_n . В результате находится вектор нечетких оценок $\hat{X}_H^T =$

$= (\hat{x}_{1H}, \hat{x}_{2H}, \dots, \hat{x}_{nH})$, компоненты которого согласно [2] могут быть "сильными" или "слабыми".

Пример. Имеем модель: $y_H(t) = x_{1H} \cdot 1 + x_{2H} \cdot t + e_H(t)$. Здесь $f_1(t) = 1, f_2(t) = t$ — базисные функции ($n = 2$). Получены нечеткие случайные данные ($m = 3$):

точка 0_1 : ($t_1 = -1; y_{1H} = (y_1 = r; \bar{y}_1 = 2 - r | r \in [0; 1])$);

точка 0_2 : ($t_2 = 0; y_{2H} = (y_2 = r; \bar{y}_2 = 0 | r \in [0; 1])$);

точка 0_3 : ($t_3 = 1; y_{3H} = (y_3 = r; \bar{y}_3 = (\beta - 1) - \beta \cdot r | r \in [0; 1])$), $\beta > 0$.

Вычисления дают:

$$(f_1, f_1) = \sum_{i=1}^3 1 \cdot 1 = 3; (f_1, f_2) = (f_2, f_1) = \sum_{i=1}^3 1 \cdot t_i = 0;$$

$$(f_2, f_2) = \sum_{i=1}^3 t_i^2 = 2;$$

$$(f_1, Y_H) = \sum_{i=1}^3 1 \cdot y_{iH} = \underbrace{(y_{1H} + y_{3H})}_{u_{1H}} = u_{1H} =$$

$$= (\underline{u}_1(r) = 2 \cdot r; \bar{u}_1(r) = (\beta + 3) - (\beta + 1)r | r \in [0; 1]);$$

$$(f_2, Y_H) = \sum_{i=1}^3 t_i y_{iH} = \underbrace{(-y_{1H} + y_{3H})}_{u_{2H}} = u_{2H} =$$

$$= (\underline{u}_2(r) = 0; \bar{u}_2(r) = (\beta - 1) - (\beta - 1)r | r \in [0; 1]).$$

В результате получим:

$$\underbrace{\begin{pmatrix} (f_1, f_1) & (f_1, f_2) \\ (f_2, f_1) & (f_2, f_2) \end{pmatrix}}_A = \underbrace{\begin{pmatrix} x_{1H} \\ x_{2H} \end{pmatrix}}_{\hat{X}_H} = \underbrace{\begin{pmatrix} u_{1H} = (f_1, Y_H) \\ u_{2H} = (f_2, Y_H) \end{pmatrix}}_{U_H} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} 3 \cdot x_{1H} + 0 \cdot x_{2H} = u_{1H} \\ 0 \cdot x_{1H} + 2 \cdot x_{2H} = u_{2H} \end{cases} \Rightarrow \hat{x}_{1H} = u_{1H}/3 =$$

$$= \left(\underline{x}_1(r) = \frac{2}{3}r; \bar{x}_1(r) = \frac{\beta+3}{3} - \frac{(\beta+1)}{3} \cdot r | r \in [0; 1] \right);$$

$$\hat{x}_{2H} = u_{2H}/2 =$$

$$= \left(\underline{x}_2(r) = 0; \bar{x}_2(r) = \frac{(\beta-1)}{2} - \frac{(\beta-1)}{2}r | r \in [0; 1] \right).$$

Оценка \hat{x}_{iH} , $i = 1, 2$, является "сильной", если компоненты $\underline{x}_i(r), \bar{x}_i(r)$ удовлетворяют условиям:

- (i) $\bar{x}_i(r)$ — монотонно убывающая (\downarrow);
- (ii) $\underline{x}_i(r)$ — монотонно возрастающая (\uparrow);
- (iii) $\bar{x}_i(r) \geq \underline{x}_i(r), \forall r \in [0; 1]$.

Для $\hat{x}_{1H}(r)$ имеем: $\underline{x}_1(r) \uparrow; \bar{x}_1(r) \downarrow, x_{1H}(r = 1) = x_1 = 2/3; \bar{x}_1(r = 0) = 1 + \beta/3$, поэтому при $\forall \beta > 0, r = 0 \Rightarrow 1 + \beta/3 > 2/3 \Leftrightarrow \bar{x}_1(r = 0) > x_1$, т. е. справедливо (iii). Это означает, что \hat{x}_{1H} является "сильной" оценкой при $\forall \beta$.

Аналогичные вычисления для \hat{x}_{2H} дают, что при $\beta \geq 1$ оценка \hat{x}_{2H} является "сильной", а при $0 < \beta < 1$ она является "слабой". Это означает, что в зависимости от значения параметра β нечеткая оценка модели $y_H(t)$ может быть либо "сильной", либо "слабой".

Выводы

- ♦ Рассмотрены нечеткие случайные данные в виде "гибридных" переменных.
- ♦ Сформулирована задача по оцениванию параметров "гибридных" данных.
- ♦ Рассмотрены модификации традиционных методов математической статистики: метод моментов; метод максимального правдоподобия; метод наименьших квадратов при обработке "гибридных" данных.
- ♦ Приведены простейшие примеры появления "сильных—слабых" оценок.

Список литературы

1. Roy Goetschel Jr and William Voxman. Elementary fuzzy calculus // Fuzzy sets and systems. 1986. N. 13. P. 31—43.
2. Menahem Friedman and etc. Fuzzy linear systems // Fuzzy sets and systems. 1998. N. 96. P. 201—209.
3. Bernhard F. A. Testing fuzzy hypotheses with craps data // Fuzzy sets and systems. 1998. N. 94. P. 323—333.
4. Горяинов В. Б., Павлов И. В., Цветкова Г. М. Математическая статистика. М.: МГТУ им. Н. Э. Баумана, 2001.
5. Ma J., Feng G. An approach to H_∞ control of fuzzy dynamic systems // Fuzzy sets and systems. 2003. N. 137. P. 367—386.

СЕТИ И СИСТЕМЫ СВЯЗИ NETWORKS AND COMMUNICATION SYSTEMS

УДК 004.9:004.7

Н. А. Балонин, д-р техн. наук, проф. Санкт-Петербургский государственный университет аэрокосмического приборостроения, e-mail: korbendfs@mail.ru,

М. Б. Сергеев, д-р техн. наук, проф., директор, НИИ информационно-управляющих систем Санкт-Петербургского национального исследовательского университета информационных технологий механики и оптики, e-mail: mbse@mail.ru

Современные сетевые технологии в научных исследованиях

Рассматривается вопрос эффективного использования сетевых технологий в процессах исследований, выполняемых распределенными в пространстве коллективами исследователей, изменяющих само представление о коллективном научном труде. Предлагается новый образ современного способа создания и редактирования научного отчета в сети.

Ключевые слова: научные исследования, глобально-распределенный научный коллектив, сетевые технологии, сетевое редактирование

N. A. Balonin, M. B. Sergeev

Modern Network Technologies in Scientific Researches

The paper discusses the effective usage of network technologies in the processes of researches carried out by distributed teams of researchers in the space that altering the concept of collective scientific work. We propose a new modern approach to the way of creating and editing a scientific reports on the network.

Keywords: scientific researches, global-distributed research team, network technologies, network editing

Введение

Бурное развитие современных сетевых технологий внесло в нашу жизнь цифровое телевидение, цифровую связь, социальные сети, Интернет-магазины и т. п., что постепенно меняет сам образ трудовой жизни и отдыха. Наблюдавшаяся еще недавно усиленная миграция научных кадров из России на Запад сменилась миграцией реальных и, добавим, виртуальных рабочих мест в Россию. Сегодня не только программирование, но и социальные, математические, и даже физические исследования можно выполнять, не выходя из дома, — в виртуальной лаборатории. Такие лаборатории оснащаются реальным, порою дорогостоящим оборудованием, но доступ к нему сотрудника можно обеспечить весьма и весьма удаленный. Ныне можно констатировать, что уровень технологий связи вырос настолько, что эффективные научные коллективы можно формировать по распределенному принципу [1]. Можно, но механизмы создания коллективного результата в научной сфере во многом пока еще остались прежними, не соответствующими близкому будущему и даже современности.

В данной работе рассматриваются принципы организация научных исследований в рамках высокопродуктивной математической сети.

Научные исследования в математической сети

Примеров стартующих и резко набирающих вес сетевых технологий немало, но надо полагать, основные события все еще впереди. Довольно очевидно, например, что глобальная математическая сеть — эффективный инструмент научных исследований, ее становление имеет большое значение для развития науки, и это то направление, в котором выгодно иметь приоритет [2].

Некоторые особенности организации глобальной математической сети таковы:

- дистанционное объединение научных сотрудников, участвующих в работе, из самых разных регионов страны, других стран;
- обеспечение дистанционного доступа к сложному дорогостоящему и дефицитному научному оборудованию и стендам, удаленное управление ими, включая математические серверы;
- обеспечение удаленного многопользовательского режима доступа к формируемым отчетным документам научными работниками;
- возможность оперативно обмениваться расчетами и результатами вычислений и моделирования по сети.

Перечислить можно и многое другое, следующее из разрушения барьеров частных учреждений, барьеров, созданных не столько их необходимостью, сколько расстоянием, условиями организации жизни

ни, подвергшимися теперь решительным изменениям. Стоит упомянуть эффективность специфических научных коллективов, возникших во Франции вокруг фигуры Мерсенна, которым пришлось бороться с теми течениями, которые их создали. Многие участники научной жизни вынуждены были жить, как Декарт в Голландии, за рубежом. Богословские школы противились размыванию своего влияния, и, тем не менее, не смогли воспрепятствовать шестью печатного станка, этого Интернета прошлого. Нынешний Интернет значительно более серьезный и готовит большие изменения, чем это возможно было столетия тому назад.

Конечно, объединить исследователей на одной сетевой площадке гораздо проще, чем каждого обеспечить научным оборудованием для этих исследований. Тем не менее, это возможно. Наиболее перспективны для расширяемой связи с "научными роботами" — стендами и прочим оборудованием — беспроводные технологии и их сенсорные завершения, которые испытывают настоящий бум своего развития. Сегодня область сенсорных завершений делят между собой множество реализаций с использованием стандартов передачи данных Wi-Fi, Bluetooth, Wireless USB, ZigBee, Home RF и т. д.

Почему технически оснащенная алгоритмическая научная сеть должна быть, все-таки, именно математическая? Да потому что математика — это основа, без нее, без квалифицированного быстрого и эффективного расчета невозможно никакое движение вперед. Следовательно, совокупно со стендами, к научной сети необходимо подключать и математическое обеспечение. Да, математические пакеты прошлого — неплохое достижение своего времени, но ныне, в связи с развитием сети Интернет, это тормоз. Сеть не может зависеть от версий и релизов разношерстного математического обеспечения на компьютерах конечных его пользователей, обмен таким научным продуктом весьма затруднен. Вычисление должно проводиться в сети.

Реализация математической сети, в отношении построения которой у авторов настоящей статьи накоплен богатый опыт, легко интегрируется в Интернет с математическими блогами и книгами, средствами удаленного проведения лабораторных исследований и представляет собой, вместе с новыми технологиями подключения стендов, основу беспроводных сетей будущего с низким энергопотреблением, предназначенных для систем управления с большим количеством робототехнических узлов. Факт, что новшество это уже внедряется, и оно неизбежно будет востребовано в большом количестве повсеместно — для того чтобы выполнить исследование, не надо лично присутствовать на эксперименте, если он управляем через сеть.

Примером проведения дистанционных исследований, начатых с более простого — с современной учебной лаборатории для выполнения работ по дисциплине "Моделирование", может служить научная математическая сеть, построенная по технологии "Живая книга" [3].

Современная система редактирования

Научное коллективное исследование в сети, естественно, должно перетекать в коллективное написание отчета, создание информативных иллюстраций и, наконец, печать в бумажном виде без дополнительных работ — пришло время таких эволюций.

Для этого необходимы соответствующие технологии поддержки. Сейчас многие крупные фирмы приступили к формированию полноценного сетевого текстового редактора. Работа идет с переменным успехом: была заявлена, но так и не осуществлена как широко приемлемая, например, весьма продвинутая система от Google. Пользователи отечественной почты относительно недавно ознакомились, но уже привыкли к новому сервису Mail.ru при написании писем. Однако для построения математических систем нужен свой текстовый редактор.

Существуют десятки версий его выполнения на имеющемся программном обеспечении. Однако такой редактор, помимо прочего, должен допускать написание и встраивание математических формул и примеров, выполняемых в реальном времени, и связь с роботами. Ничем подобным, разумеется, почтовая система и даже сложный редактор от Google не оснащены. Но это отнюдь не невозможная задача, это вопрос целесообразности. Если так именно нужно сделать, то это сделать вполне реально. Далее, система редактирования должна автоматизировать написание труда несколькими соавторами, что тоже чисто искусственная дополнительная сложность, никаким серьезным препятствием не являющаяся. Такой сетевой редактор может быть сделан.

Научная публикация — естественный выход научного исследования. Как раз сеть в состоянии ее поддержать. Необходимо создать ресурс, поддерживающий рабочие места автора и редактора текста, корректора, рецензента. Кто сталкивался с этим процессом сегодня, признает, что существующая технология на фоне современных возможностей крайне убога. Во многом это эксплуатация все той же почты, с ее ограниченным интерфейсом. Подготовка к выходу в печать, в итоге, идет неоправданно долго и трудозатратно.

Это можно и надо исправлять, причем решительно.

Решения, обеспечивающие реализацию исследований в сети

Эпоха семантического WEB приносит нам книги и документы, читаемые на букридерах и смартфонах. Немалую роль в этом играют появившиеся и быстро распространившиеся правила оформления документов с содержательными частями, выделяемыми XML-тэгами, как это делается, например, в формате FB2, а также новые расширения языка Javascript.

Отдельного рассмотрения требует техника записи в научной математической сети формул и индексов. В Интернете можно встретить итоги компиляции математических формул в формат gif, после чего математические записи теряют многое в своем качестве. Учитывая сказанное, с помощью клавиатуры (в том числе и виртуальной) определе-

ния норм векторов в пространстве R^n предлагается записывать, например, так:

$$\|x\|_1 = \sum_{i=1:n} |x_i|, \|x\|_2 = (\sum_{i=1:n} x_i^2)^{1/2}, \|x\|_\infty = \max_{i=1:n} |x_i|,$$

указывая границы индексов у сумм (или у интегралов) внизу, в строчку.

Формат, в котором реализованы указанные выше и многие другие возможности, развивающие и дополняющие формат FB2, авторы назвали FB2+ [4].

В свое время матричные операции не были включены в стандарт языка javascript. Эта ситуация исправима за счет использования, например, двойных фигурных скобок `{{...}}` для выделения матричных выражений. Исполняемый в тексте сообщения алгоритм размещается между тэгами `${{матричные операции}}$`.

С помощью предкомпилятора содержимое фигурных скобок транслируется с языка векторно-матричного исчисления в javascript, этому исполнению авторы дали название Java-MatLab. В его реализации определены все основные матричные операции: транспонирование `{{X = A'}}`, алгебраические сложение, умножение, левое `{{A = A\B}}` и правое `{{X = V/A}}` умножение на обратную матрицу, точечные операции — произведение и деление Адамара `{{X = A.*B; X = A./B}}`.

Для передачи по сети матриц предлагается использовать имеющиеся в сети тэги оформления таблиц. Для генерации, например, портретов матриц могут быть предложены тэги `<m>A = [[1, 2], [3, 4]]:опция</m>`. То же самое касается выводов графиков функций, в том числе и в 3D-виде.

Всего на Livelab.spb.ru [5] реализовано более полусотни функций для работы с графикой и анимацией, для моделирования динамических систем, для символьных и числовых вычислений, генерации матриц и векторов, решения задач линейной алгебры и многого другого. С их помощью можно решать как стандартные задачи линейной алгебры, включая решение систем линейных алгебраических уравнений и алгебраическую проблему собственных чисел, так и развернуть процедуры анализа и синтеза линейных динамических систем, проводить частотный анализ систем и сигналов, безошибочное решение целочисленных систем уравнений [6]. Реализована также возможность организации и подключения пользовательских тулбоксов.

Работа в сети со стендами

В новой технологии коммуникационные средства сети Интернет связывают не компьютер с компьютером, а компьютер с датчиками и эффекторами.

При написании сценариев работы со стендами и лабораторным оборудованием следует учитывать, что современные микропроцессоры пришли к той стадии, на которой находились некогда первые компьютеры. Их управление сводится к подаче текстовых команд, подобно тому, как это делалось в рамках дорогостоящей дисковой операционной системы (DOS). Настоящая революция началась с невинного предложения Питера Нортона (Peter Norton), заме-

нить текстовые команды, оставив командную строку, графическим интерфейсом, и, по сути, все к этой идее до сих пор сводится. Поэтому при работе со стендами необходимы режим текстовой строки и средства создания кнопочного интерфейса, к чему сеть давно готова, поскольку и графические команды, и средства инициализации кнопочного интерфейса, все это есть и достаточно развито. Остается только совместить все в едином редакторе, который в равной мере пишет как текст, так и конструирует интерфейс. Даже небольшие наработки в течение одного—двух лет дают ощутимую отдачу, не говоря о "фирменном подходе".

Наработка интерфейсов связи находится сейчас в майнстриме текущих событий. В отличие от специфичной мультимедийной технологии Flash, технология создания графических зарисовок и систем управления стендами в них в опытной математической сети [7] без проблем пережила революцию внедрения конечных систем пользователя на iPad и с Android. Вычисления и действия с лабораторными стендами исполняются на странице математической сети несколько не хуже, чем в локальной среде, а последствия этого усовершенствования самые впечатляющие.

У сети должна быть профессия

Сколько бы ни пытались математики придать своим математическим инструментам универсальность, это, в общем, не получается. У пакетов MatLab, Maple, MathCad и т. п. есть свои, довольно ясно выраженные области, где их употребление наиболее целесообразно. Создатели математических сетей тоже должны определиться со своей областью. Система MatLab продвигала в свое время тулбоксы Control, это отвечало запросам времени на конструирование систем управления. Ситуация не может быть вечной, в силу имеющейся традиции, средства Control несложно воспроизвести в сети. Но мир перешел к другой стадии. Надо отчетливо осознать, что сейчас чрезвычайно важными становятся информационные технологии переработки информации.

Матричная алгебра отвечает таким потребностям. С помощью предлагаемого синтаксиса математические вычисления, долгое время совершенно неоправданно не используемые в сети, уже сегодня шагнули на научные и обучающие форумы и блоги. В качестве примера можно привести результаты исследований ортогональных M-матриц [8, 9], проводимых авторами и давших возможность работы с плоскими и объемными портретами матриц, сформированными программными ресурсами сети. Такие иллюстрации позволили выявить особенности впервые вычисленных матриц Адамара—Мерсенна [10], Адамара—Ферма [11] и Адамара—Эйлера [12], а также их взаимосвязь.

Исследование сложных в техническом исполнении рисунков, графического 3D-отражения динамики процессов помогает выявлять новые артефакты, неизвестные науке закономерности и т. п. Опыт внедрения математических пакетов прошлого показал, сколь эффективными могут быть реализации вычислительных методов линейной алгебры

не с помощью библиотек подпрограмм, а всего лишь с помощью удачного синтаксиса, передающего особенности языка матричных вычислений типа MatLab. Язык Java-MatLab и система его реализации, разработанные и развитые в рамках выполненной НИР [4], внедрены в учебный процесс и используются на Интернет-ресурсах, предназначенных для совместных исследований в сети [5, 7, 13].

Генотип и фенотип математических сетей

За десятилетие существования глобальной информационной сети Интернет, как фактора повседневной жизни, накоплен немалый опыт создания сетевых ресурсов. Что отличает жизнеспособный развивающийся ресурс от скоро погибающего ресурса? Всего их создано было немало, и немалое же количество их исчезло во времени. Отметим родство информационного сетевого ресурса всему живому на Земле. Как и любая живая система, информационная система может быть условно разделена на части: генотип и фенотип. Генотип подобен корневой системе растения, а фенотип, это то, что собирается выше и нарастает в результате деятельности системы.

Иными словами, генотип — это набор кодов, или информация, которую сеть получила на момент создания. Фенотип, соответственно, — информация, собираемая системой, восприятие которой (трактовка) существенно зависит от генотипа. Сеть без развития и того, и другого мертва. Стартовое развитие начинается с набора фенотипа, но рано или поздно, без прогресса самого основания сети, она перестает отвечать вызовам дням. Необходима "коренная переделка", сопровождаемая нередко гибелью предыдущей информационной системы. Так или иначе, части генотипа переключиваются в новую систему, т. е. репродуцируется основное ядро кодов.

Развитие с крушениями крупных информационных ресурсов, разумеется, совершенно нежелательно в сети. Но рано или поздно наступает момент, когда основа перестает отвечать требованиям современности. Так, например, одну социальную сеть покидают пользователи, уходя в другие сети. Чтобы это предотвратить, необходимо основу время от времени менять. Такова предыстория Google+. Рациональна схема эволюции — это когда старая и новая системы сосуществуют так же, как сосуществуют родители и их дети. Такое обустройство системы или, вернее, распределенных систем, гибче.

Заметим, что способность к развитию в системе должна быть заложена заранее, как элемент необходимого конструктива. Консервирование кодов сдерживает развитие.

Например, математическая сеть MathSciNet.ru [7] начиналась с переделки и устранения недостатков, с учетом накопленного опыта, общеобразовательного портала по математике, кибернетике и программированию Artspb.com [13], внесенного в каталог системы федеральных образовательных порталов Министерства образования и науки РФ [14]. За два—три года сеть испытала несколько фаз эволюции, стартовое неизменяемое консервативное кодовое ядро с на-

чальной наработкой Интернет-MatLab сохранилось, но его уже вытесняют фитопобеги созданной принципиально многоядерной структуры LiveLab.spb.ru с Java-MatLab, Интернет-роботами и стендами [5].

Заключение

Вычислительная и иллюстративная мощь математической сети может быть чрезвычайно высока. Это удобный и простой для реализации глобально-распределенных исследований принцип развития науки отечественными и международными исследовательскими коллективами.

Наиболее востребованной общей дисциплиной для исследований в технической сфере является математика. В качестве наглядного примера быстро растущей математической системы можно сослаться также на стартовавший в сети в 2009 году ресурс, оперирующий командной строкой справки [15]. Представляется разумной идея реализации в сети уже сегодня Интернет-центра, решающего математические задачи квалифицированно, под присмотром и с участием специалистов. Для остальных сетевых ресурсов, развивающих технологию коллективных исследований, пользование математическими сервисами может осуществляться на основе их потребностей.

Такова концепция математической сети, которая запущена авторами статьи и уже растет, имея несколько разновидностей и адресов в Интернет [5, 7, 13].

Список литературы

1. **Наумова В. В.** Виртуальные научные среды для обеспечения совместной работы территориально распределенных научных работников // Информационные технологии. 2013. № 4. С. 46—51.
2. **Балонин Н. А., Сергеев М. Б.** Концепция электронного журнала с исполняемыми алгоритмами // Фундаментальные исследования. 2013. № 4—4. С. 791—795.
3. **Балонин Н. А., Сергеев М. Б., Балонин Ю. Н.** "Живая книга" / Свидетельство о государственной регистрации программы для ЭВМ № 2012661277 от 11 декабря 2012 г.
4. **Создание** основ реализации дистантных систем обучения на основе технологии "Живая книга": Отчет о НИР Гос. рег. № 01201278144 / Балонин Н. А., Сергеев М. Б., Соловьев Н. В., Востриков А. А., Балонин Ю. Н., Сергеев А. М. СПб: ГУАП, 2013. 27 с.
5. **Математическая** сеть "Живая Книга" с интернет-роботами и стендами. URL: <http://livelab.spb.ru> (основана в 2013 г.).
6. **Сергеев М. Б.** Гибридный разрядный метод решения систем уравнений в целочисленной арифметике // Информационно-управляющие системы. 2003. № 2—3. С. 16—18.
7. **Математическая** сеть "Скайнет": технологии верстки физико-математической литературы с исполняемыми алгоритмами. URL: <http://mathscinet.ru> (основана в 2012 г.).
8. **Балонин Н. А., Сергеев М. Б.** М-матрицы // Информационно-управляющие системы. 2011. № 1. С. 14—21.
9. **Балонин Ю. Н., Сергеев М. Б.** Алгоритм и программа поиска и исследования М-матриц // Научно-технический вестник информационных технологий, механики и оптики. 2013. № 3 (85). С. 82—86.
10. **Балонин Н. А., Мироновский Л. А., Сергеев М. Б.** Вычисление матриц Адамара—Мерсенна // Информационно-управляющие системы. 2012. № 5. С. 92—94.
11. **Балонин Н. А., Сергеев М. Б., Мироновский Л. А.** Вычисление матриц Адамара—Ферма // Информационно-управляющие системы. 2012. № 6 (61). С. 90—93.
12. **Балонин Н. А., Сергеев М. Б.** О двух способах построения матриц Адамара—Эйлера // Информационно-управляющие системы. 2013. № 1. С. 7—10.
13. **Образовательный** портал: математика, кибернетика и программирование. URL: <http://artspb.com> (основан в 2005 г.).
14. **Российский** общеобразовательный портал Министерства образования и науки (система федеральных образовательных порталов). URL: <http://www.school.edu.ru>.
15. **Система** математической справки. URL: <http://wolframalpha.com> (основана в 2009 г.).

МОДЕЛИРОВАНИЕ И ОПТИМИЗАЦИЯ MODELING AND OPTIMIZATION

УДК 62-50:519.7/8

В. И. Левин, д-р техн. наук, проф., e-mail: vilevin@mail.ru,
Пензенский государственный технологический университет

Устойчивость решения задач оптимизации в условиях неопределенности

Рассмотрена проблема оптимизации неполностью определенных функций — функций с параметрами, заданными с точностью до интервала. Показано, что решение этой проблемы требует также рассмотрения задачи определения устойчивости оптимума к варьированию значений параметров функции. Предложен метод нахождения оптимума функций и определения его устойчивости методами интервальной математики.

Ключевые слова: оптимизация систем, устойчивость оптимума, неопределенность, варьирование параметров, интервальная математика, математическое программирование

V. I. Levin

Solution Stability of Optimization Problems in Condition of Uncertainty

A problem of optimization of incompletely defined (non-deterministic) functions is considered. The parameters of such functions are intervals. It is shown that problem solution also requires touching of problem: finding optimum which has stability to variations of values of the function parameters. Method for determining optimum of function and its stability by methods of interval mathematics is presented.

Keywords: system optimization, uncertainty, stability of the optimum, variation of parameters, interval mathematics

Введение

На сегодняшний день в мире имеется обширная литература по оптимизации (оптимальному проектированию) различных систем с детерминированными параметрами — технических, экономических и т. д. Соответствующие задачи формулируются как задачи математического программирования с целевыми функциями и функциями ограничений, параметры которых являются детерминированными величинами. Однако на практике по объективным причинам чаще встречаются системы с недетерминированными параметрами. Оптимизация такого рода систем формализуется в виде задач математического программирования с целевыми функциями и функциями ограничений, параметры которых — различные недетерминированные величины: случайные, нечеткие, интервальные и т. д. Эти задачи, вообще говоря, сложнее детерминированных. Они требуют обобщения понятия экстремума функции, выяснения условий его существования, связанных с недетерминированностью параметров функции, и разработки специальных методов поиска экстремума таких функций.

Существует три различных подхода к решению недетерминированных задач математического про-

граммирования: детерминированный, вероятностный [1] и интервальный [2].

Детерминированный подход заключается в решении задачи для определенных значений ее параметров, выбранных внутри соответствующих заданных областей неопределенности. Например, могут быть выбраны центры (середины) областей неопределенности параметров (центральная стратегия), наихудшее сочетание значений параметров задачи (пессимистическая стратегия), их наилучшее сочетание (оптимистическая стратегия) и т. д.

Вероятностный подход заключается в решении задачи для усредненных (ожидаемых, в смысле математического ожидания) значений ее параметров, что предполагает задание вероятностной меры внутри соответствующих областей неопределенности.

Оба указанных подхода объединяет предварительная детерминизация параметров задачи, выполняемая перед ее оптимальным решением.

В отличие от них, *интервальный подход* не предполагает детерминизации параметров задачи, которые задаются в интервальной форме — в нем оптимальное решение задачи проводится в ее "естественной форме", т. е. на основе прямого сравнения недетерминированных значений целевой функции,

соответствующих различным значениям вектора аргументов, и выборе оптимального (максимального или минимального) значения этой функции. Достоинства и недостатки указанных трех подходов рассмотрены в работах [1–8].

Изложенные подходы к решению недетерминированных задач математического программирования, при всем их очевидном различии, объединяет одна общая существенная черта. А именно, все они предназначены для решения задач оптимизации, в которых параметры целевых функций и функций ограничений точно не известны. Поэтому мы не можем ограничиться просто отысканием оптимального решения нашей задачи, используя для этого один из упомянутых выше методов. В самом деле, ввиду отсутствия при решении задачи точных значений ее параметров может оказаться, что действительные значения параметров задачи несколько отличаются от тех, которые были приняты в процессе отыскания ее решения. В этих условиях для того чтобы найденное оптимальное решение задачи имело содержательный прикладной смысл, нам нужно, чтобы оно еще обладало следующим свойством: при небольшом варьировании значений параметров решаемой задачи ее оптимальное решение должно по-прежнему существовать. При этом точка, в которой достигается оптимум целевой функции, может переместиться из исходного положения в новое положение, которое, однако, должно быть близко к исходному. Другими словами, требуется, чтобы найденное оптимальное решение не полностью определенной (недетерминированной) задачи математического программирования было устойчивым относительно небольших количественных изменений ее параметров.

1. Постановка задачи

Рассмотрим сначала детерминированный случай. Пусть задана произвольная непрерывная функция n переменных

$$y = F(x_1, \dots, x_n), \quad (1)$$

где все параметры (коэффициенты) ее явного представления p_k , $k = \overline{1, l}$, известны точно. Будем рассматривать функцию (1) в ограниченной области, определяемой системой ограничений

$$\Phi_i(x_1, \dots, x_n) \leq b_i, \quad i = \overline{1, m}, \quad (2)$$

в которой параметры q_s , $s = \overline{1, t}$, явного представления функций ограничений Φ_i и правые части b_i также известны точно.

Тогда относительно указанной выше функции (1) можно сформулировать полностью определенную задачу условной оптимизации (задачу математического программирования)

$$F(x_1, \dots, x_n) = \max \quad (3)$$

при условии

$$\Phi_i(x_1, \dots, x_n) \leq b_i, \quad i = \overline{1, m}. \quad (4)$$

Решением задачи (3), (4) является некоторая точка $x^* = (x_1^*, \dots, x_n^*)$ (множество точек $M = \{x^*\}$) области (4), в которой целевая функция F достигает максимального значения F_{\max} . В современном математическом программировании разработано много различных методов эффективного решения задач вида (3), (4), ориентирующихся на тип целевой функции F и функций ограничений Φ_i , $i = \overline{1, m}$.

Предположим теперь, что в задаче оптимизации (3), (4) параметры явного представления целевой функции F , а также параметры явного представления функций ограничений Φ_i и правые части ограничений b_i известны не точно, а приближенно. Тогда, в соответствии со сказанным во введении, мы должны совместно с задачей условной оптимизации (3), (4) рассматривать еще одну: проверка устойчивости (неустойчивости) решения задачи (3), (4) относительно небольших количественных изменений ее параметров.

В отличие от существующих методов изучения устойчивости решения задач оптимизации [5] будем рассматривать все возможные количественные изменения каждого параметра задачи как единое целое. Такое рассмотрение позволяет задавать все возможные количественные изменения параметров задач оптимизации в теоретико-множественных терминах. Простейший способ такого задания заключается в том, чтобы задать совокупность указанных изменений параметров задачи в виде соответствующих числовых интервалов. Преимущество такого подхода к изучению устойчивости решения задач оптимизации состоит в том, что в его рамках изучать устойчивость задач оптимизации можно с помощью хорошо разработанных методов интервальной математики [9].

Итак, совместно с полностью определенной задачей условной оптимизации (3), (4) мы должны рассмотреть производную от нее интервальную задачу условной оптимизации

$$\tilde{F}(x_1, \dots, x_n) = \max \quad (5)$$

при условии

$$\tilde{\Phi}_i(x_1, \dots, x_n) \leq \tilde{b}_i, \quad i = \overline{1, m}. \quad (6)$$

Целевая функция \tilde{F} интервальной задачи оптимизации (5), (6) получается из целевой функции F искомой, полностью определенной задачи оптимизации (3), (4) путем замены ее точно известных параметров p_k , $k = \overline{1, l}$, соответствующими интервальными параметрами $\tilde{p}_k = [p_{k1}, p_{k2}]$, $k = \overline{1, l}$, которые и определяют интервальную целевую функцию \tilde{F} . Аналогично, любая функция ограничений $\tilde{\Phi}_i$, $i = \overline{1, m}$, интервальной задачи условной оптимизации (5), (6) получается из соответствующей функции Φ_i , $i = \overline{1, m}$, исходной полностью определен-

ной задачи (3), (4) заменой ее точно известных параметров q_{si} , $s = 1, t$, $i = 1, m$, соответствующими интервальными параметрами $\tilde{q}_{si} = [q_{si1}, q_{si2}]$, $s = 1, t$, $i = 1, m$.

Точно так же интервальные параметры \tilde{b}_i , $i = 1, m$, в ограничениях интервальной задачи условной оптимизации (5), (6) заменяют собой соответствующие точно известные параметры b_i , $i = 1, m$, в ограничениях исходной, полностью определенной задачи оптимизации (3), (4).

Будем называть полностью определенную задачу условной оптимизации (математического программирования) (3), (4) *макроустойчивой*, если она имеет решение и, кроме того, имеет решение производная от нее интервальная задача оптимизации (математического программирования) (5), (6).

Далее, будем называть полностью определенную задачу условной оптимизации (математического программирования) (3), (4) *микроустойчивой*, если она макроустойчива и, сверх того, существует пара решений (x', x'') , где $x' = (x'_1, \dots, x'_n)$ — некоторая точка решения задачи (3), (4), а $x'' = (x''_1, \dots, x''_n)$ — некоторая точка решения задачи (5), (6), расстояние между которыми $D(x', x'')$ не превосходит заданной достаточно малой величины d .

Задача настоящего исследования — разработать алгоритмы определения макро- и микроустойчивости полностью определенных задач условной оптимизации типа (3), (4).

2. Математический аппарат

В основе решения поставленной задачи лежит аппарат интервальной математики [9]. В этой математике алгебраические операции над интервальными числами $\tilde{a} = [a_1, a_2]$, $\tilde{b} = [b_1, b_2]$, ... вводятся в виде следующих теоретико-множественных конструкций:

$$\tilde{a} + \tilde{b} = \{a + b | a \in \tilde{a}, b \in \tilde{b}\},$$

$$\tilde{a} - \tilde{b} = \{a - b | a \in \tilde{a}, b \in \tilde{b}\}, k\tilde{a} = \{ka | a \in \tilde{a}\}, \dots (7)$$

и т. д.

Другими словами, любая операция над интервалами определяется на основе соответствующей операции над точечными величинами при условии, что конкретные значения величин пробегают все возможные значения из соответствующих интервалов. Из введенных алгебраических операций над интервалами вытекают простые правила выполнения этих операций:

$$\begin{aligned} [a_1, a_2] + [b_1, b_2] &= [a_1 + b_1, a_2 + b_2], \\ [a_1, a_2] - [b_1, b_2] &= [a_1 - b_2, a_2 - b_1]; \\ k[a_1, a_2] &= \begin{cases} [ka_1, ka_2], & k > 0, \\ [ka_2, ka_1], & k < 0, \end{cases} \\ [a_1, a_2] \cdot [b_1, b_2] &= [\min_{i,j} (a_i \cdot b_j), \max_{i,j} (a_i \cdot b_j)]; \\ [a_1, a_2] / [b_1, b_2] &= [a_1 \cdot a_2] \cdot [1/b_2, 1/b_1]. \end{aligned} \quad (8)$$

Введем теперь операции сравнения интервальных чисел [2, 8]. Попытаемся сравнить два интервала $\tilde{a} = [a_1, a_2]$ и $\tilde{b} = [b_1, b_2]$, рассматривая их как интервальные числа. Естественно начать со сравнения интервалов \tilde{a} и \tilde{b} на базе сравнений в отдельных парах вещественных чисел (a_i, b_j) , где $a_i \in \tilde{a}$, $b_j \in \tilde{b}$. Но это ведет к провалу, так как в общем случае одни пары чисел (a_i, b_j) будут находиться в отношении $a_i > b_j$, а другие — в противоположном отношении: $a_i < b_j$. Поэтому единственное, что остается — реализовать операцию сравнения интервалов на теоретико-множественном уровне, подобно алгебраическим операциям над интервалами (7). В соответствии со сказанным, введем операции взятия максимума \vee и минимума \wedge двух интервальных чисел $\tilde{a} = [a_1, a_2]$ и $\tilde{b} = [b_1, b_2]$ в виде теоретико-множественных конструкций

$$\tilde{a} \vee \tilde{b} = \{a \vee b | a \in \tilde{a}, b \in \tilde{b}\},$$

$$\tilde{a} \wedge \tilde{b} = \{a \wedge b | a \in \tilde{a}, b \in \tilde{b}\}. \quad (9)$$

Операция взятия максимума (минимума) из двух интервалов \tilde{a} и \tilde{b} , согласно (9), определяется как нахождение максимума (минимума) из двух точечных величин a и b при условии, что конкретные значения этих величин пробегают все возможные значения соответственно из интервалов \tilde{a} и \tilde{b} . Чтобы интервалы \tilde{a} и \tilde{b} можно было сравнить по величине, установив их отношение ($\tilde{a} \geq \tilde{b}$ или $\tilde{a} \leq \tilde{b}$), нужно, чтобы:

- введенные операции \vee , \wedge над этими интервалами существовали;
- результатом этих операций был один из операндов \tilde{a} или \tilde{b} ;
- операции \vee , \wedge были согласованы, т. е. было выполнено условие: если большим (меньшим) является один из интервалов \tilde{a} , \tilde{b} , то меньшим (большим) является другой из них.

Условие сравнимости величин двух интервалов является, очевидно, необходимым и достаточным условием. Однако легко доказать, что условие согласованности операций \vee и \wedge над интервалами выполняется всегда (для любой пары интервалов (\tilde{a}, \tilde{b})). Также всегда (для любой пары интервалов) выполняется условие существования введенных нами выше операций взятия максимума \vee и минимума \wedge двух интервалов, причем результатом операции оказывается некоторый, вообще говоря, новый интервал. В итоге необходимое и достаточное условие сравнимости двух интервалов \tilde{a} и \tilde{b} превращается в условие, по которому операции $\tilde{a} \vee \tilde{b}$ и $\tilde{a} \wedge \tilde{b}$ должны давать в результате обязательно один из интервалов-операндов: \tilde{a} или \tilde{b} . Такая формулировка условия сравнимости интервалов дает возможность получения его в конструктивной форме, пригодной, к тому же, для практического применения. Базовая форма условия здесь такова.

Теорема 1. Для сравнимости двух интервалов $\tilde{a} = [a_1, a_2]$ и $\tilde{b} = [b_1, b_2]$ и их нахождения между

собой в отношении $\tilde{a} \geq \tilde{b}$ необходимо и достаточно, чтобы одноименные границы этих интервалов удовлетворяли условиям

$$a_1 \geq b_1, a_2 \geq b_2, \quad (10)$$

а для их сравнимости и нахождения между собой в отношении $\tilde{a} \leq \tilde{b}$ — чтобы удовлетворялись условия

$$a_1 \leq b_1, a_2 \leq b_2. \quad (11)$$

По теореме 1 интервалы \tilde{a} и \tilde{b} сравнимы и находятся в определенном отношении $\tilde{a} \geq \tilde{b}$ или $\tilde{a} \leq \tilde{b}$, только когда в таком же отношении находятся их одноименные границы a_1, b_1 и a_2, b_2 . Другими словами, для сравнимости интервалов меньший интервал должен быть сдвинут обеими границами влево относительно большего интервала. Так что с помощью теоремы 1 сравнение двух интервалов и выбор большего (меньшего) из них сводится к сравнению одноименных границ этих интервалов, являющихся точно известными вещественными числами.

Теорема 2. Для несравнимости двух интервалов $\tilde{a} = [a_1, a_2]$ и $\tilde{b} = [b_1, b_2]$, иными словами, для того чтобы они не находились ни в отношении $\tilde{a} \geq \tilde{b}$, ни в отношении $\tilde{a} \leq \tilde{b}$, необходимо и достаточно, чтобы одноименные границы интервалов удовлетворяли условиям

$$a_1 < b_1, a_2 > b_2 \text{ или } b_1 < a_1, b_2 > a_2. \quad (12)$$

Условия (12) обозначают ту ситуацию, когда один интервал на числовой оси полностью "накрывает" другой.

Теорема 2 показывает существование случаев несравнимости интервалов. Несравнимость величин некоторых интервалов — естественное следствие того, что, в отличие от точных вещественных чисел, интервальные числа задаются с некоторой неопределенностью (точно известно, что вещественное число принимает некоторое значение в заданном интервале, но не известно, какое именно это значение).

Далее, теоремы 1 и 2, посвященные сравнению пар интервалов, можно обобщить на системы с произвольным числом интервалов.

Теорема 3. Для существования максимального интервала в системе из интервалов $\tilde{a}(1) = [a_1(1), a_2(1)]$, $\tilde{a}(2) = [a_1(2), a_2(2)]$, ... необходимо и достаточно, чтобы его границы располагались относительно одноименных границ всех остальных интервалов согласно следующим условиям:

$$\begin{aligned} a_1(1) &\geq a_1(2), a_1(1) \geq a_1(3), \dots; \\ a_2(1) &\geq a_2(2), a_2(1) \geq a_2(3), \dots \end{aligned} \quad (13)$$

Условия (13) записаны для случая, когда максимальным является интервал $\tilde{a}(1)$, что не ограничивает общности.

Теорема 4. Для существования минимального интервала в системе из интервалов $\tilde{a}(1) = [a_1(1), a_2(1)]$, $\tilde{a}(2) = [a_1(2), a_2(2)]$, ... необходимо и достаточно, чтобы его границы располагались относительно

одноименных границ всех остальных интервалов согласно условиям

$$\begin{aligned} a_1(1) &\leq a_1(2), a_1(1) \leq a_1(3), \dots, \\ a_2(1) &\leq a_2(2), a_2(1) \leq a_2(3), \dots \end{aligned} \quad (14)$$

Условия (14), аналогично условиям (13), записаны для случая, когда минимальным является интервал $\tilde{a}(1)$, что не ограничивает общности.

Теоремы 3, 4 означают, что интервал является максимальным (минимальным) из интервалов системы, только если максимальны (минимальны) его нижняя граница — среди нижних границ всех интервалов — и верхняя граница — среди верхних границ всех интервалов.

3. Макроустойчивость задачи условной оптимизации

Обратимся теперь к полностью определенной задаче условной оптимизации (3), (4). Опишем метод установления макроустойчивости этой задачи.

Полностью определенная задача условной оптимизации (3), (4) по определению (см. п. 1) является макроустойчивой, если она сама и производная от нее интервальная задача условной оптимизации (5), (6) имеют решения. Существование решения полностью определенной задачи условной оптимизации (3), (4) можно установить с помощью общеизвестных методов математического программирования [10—12], так что здесь нет никаких проблем. Сложнее, однако, обстоит дело с проверкой существования решения интервальной задачи условной оптимизации (5), (6). В этом случае эффективным оказывается применение детерминизационного метода решения задач интервальной оптимизации [2, 8, 13].

Интервальная задача условной оптимизации (5), (6) имеет интервальную целевую функцию $\tilde{F}(x_1, \dots, x_n)$, интервальные функции ограничений $\tilde{\Phi}_i, i = \overline{1, m}$, в левых частях ограничений и интервальные параметры $\tilde{b}_i, i = \overline{1, m}$, в правых частях. Используя формулы элементарных преобразований интервалов (8), функции $\tilde{F}, \tilde{\Phi}_i$ можно представить явно в интервальной форме. Так же можно представить и параметры \tilde{b}_i . Все эти представления записываются в виде

$$\begin{aligned} \tilde{F}(x_1, \dots, x_n) &= [F_1(x_1, \dots, x_n), F_2(x_1, \dots, x_n)]; \\ \tilde{\Phi}_i(x_1, \dots, x_n) &= [\Phi_{i1}(x_1, \dots, x_n), \Phi_{i2}(x_1, \dots, x_n)], \\ i &= \overline{1, m}; \\ \tilde{b}_i &= [b_{i1}, b_{i2}], i = \overline{1, m}. \end{aligned} \quad (15)$$

После этого всю задачу оптимизации (5), (6) также можно переписать в явном интервальном виде:

$$[F_1(x_1, \dots, x_n), F_2(x_1, \dots, x_n)] = \max, \quad (16)$$

$$[\Phi_{i1}(x_1, \dots, x_n), \Phi_{i2}(x_1, \dots, x_n)] \leq [b_{i1}, b_{i2}], i = \overline{1, m}. \quad (17)$$

От интервального представления задачи (16), (17) можно перейти к ее эквивалентному представлению в виде пары полностью определенных (детерминированных) задач условной оптимизации, которое уже поддается решению. Для этого сначала по теореме 3 представим интервальное уравнение (16) в виде эквивалентной пары детерминированных уравнений

$$F_1(x_1, \dots, x_n) = \max, F_2(x_1, \dots, x_n) = \max. \quad (18)$$

Далее, по теореме 1 представим систему интервальных неравенств (17) в виде эквивалентной системы обычных детерминированных неравенств

$$\Phi_{i1}(x_1, \dots, x_n) \leq b_{i1}, \Phi_{i2}(x_1, \dots, x_n) \leq b_{i2}, i = \overline{1, m}. \quad (19)$$

Соединив пару уравнений (18) с системой неравенств-ограничений (19), мы получим совокупность двух полностью определенных (детерминированных) задач условной оптимизации вида (3), (4):

$$\begin{cases} F_1(x_1, \dots, x_n) = \max, \\ \Phi_{i1}(x_1, \dots, x_n) \leq b_{i1}, i = \overline{1, m}, \\ \Phi_{i2}(x_1, \dots, x_n) \leq b_{i2}, i = \overline{1, m}, \end{cases} \quad (20)$$

$$\begin{cases} F_2(x_1, \dots, x_n) = \max, \\ \Phi_{i1}(x_1, \dots, x_n) \leq b_{i1}, i = \overline{1, m}, \\ \Phi_{i2}(x_1, \dots, x_n) \leq b_{i2}, i = \overline{1, m}, \end{cases} \quad (21)$$

эквивалентную исходной интервальной задаче условной оптимизации (5), (6). Задачу (20) назовем нижней граничной задачей исходной интервальной задачи (5), (6), а задачу (21) — ее верхней граничной задачей.

Итак, для получения решения интервальной задачи (5), (6) следует решить ее нижнюю (20) и верхнюю (21) граничные задачи. В общем случае полное решение нижней граничной задачи имеет вид $\{M_H(x), F_{1, \max}\}$, а верхней граничной задачи — вид $\{M_B(x), F_{2, \max}\}$.

Здесь $M_H(x)$, $M_B(x)$ — множества точек решения $x = (x_1, \dots, x_n)$ нижней и верхней граничных задач, а $F_{1, \max}$, $F_{2, \max}$ — полученные максимальные значения целевых функций этих задач. Решение интервальной задачи (5), (6) формируется из решений ее нижней и верхней граничных задач и имеет вид

$$\{x^* \in M_H(x) \cap M_B(x); \tilde{F}_{\max} = [F_{1, \max}, F_{2, \max}]\}. \quad (22)$$

Согласно (22) в качестве точки решения x^* интервальной задачи оптимизации (5), (6) выбирается любая точка из пересечения множеств точек решения ее нижней и верхней граничных задач, а в качестве максимального значения интервальной целевой функции \tilde{F}_{\max} — интервал от максимального значения целевой функции нижней задачи $F_{1, \max}$ до максимального значения целевой функции верхней задачи $F_{2, \max}$.

Из выполненного процесса построения решения интервальной задачи условной оптимизации вида

(5), (6) и определения макроустойчивости полностью определенной задачи условной оптимизации (3), (4) вытекает следующая основная теорема.

Теорема 5. Для того чтобы полностью определенная задача условной оптимизации (3), (4) была макроустойчива, необходимо и достаточно, чтобы:

- эта задача имела решение;
- интервальная задача условной оптимизации (5), (6), производная от детерминированной задачи (3), (4), имела нижнюю и верхнюю граничные задачи, обладающие решениями;
- множества решений нижней и верхней граничных задач интервальной задачи оптимизации (5), (6) пересекались.

Теорема 5 дает следующий алгоритм для проверки произвольной полностью определенной (детерминированной) задачи условной оптимизации (3), (4) на макроустойчивость.

Шаг 1. Используя подходящие известные методы решения полностью определенных (детерминированных) задач условной оптимизации [10—12], ищем решение $x' = (x'_1, \dots, x'_n)$ задачи (3), (4). Одновременно проверяем существование (несуществование) решения этой задачи.

Шаг 2. Задаваясь некоторыми подходящими значениями интервальных параметров целевой функции F , функций ограничений Φ_i , $i = \overline{1, m}$, и правых частей ограничений b_i , $i = \overline{1, m}$, в полностью определенной задаче условной оптимизации (3), (4), строим производную от нее интервальную задачу условной оптимизации (5), (6).

Шаг 3. Используя формулы интервальной математики (8), выражающие результаты элементарных преобразований интервалов, представляем целевую функцию \tilde{F} , функции ограничений $\tilde{\Phi}_i$, $i = \overline{1, m}$, а также правые части ограничений \tilde{b}_i , $i = \overline{1, m}$, интервальной задачи условной оптимизации (5), (6) в интервальной форме (15).

Шаг 4. Используя полученные на шаге 3 интервальные представления функций \tilde{F} , $\tilde{\Phi}_i$, $i = \overline{1, m}$, и параметров \tilde{b}_i , $i = \overline{1, m}$, формируем нижнюю (20) и верхнюю (21) граничные задачи интервальной задачи оптимизации (5), (6).

Шаг 5. Используя те же самые методы, что и на шаге 1, ищем решения задач (20) и (21). Одновременно проверяем существование или не существование решений этих задач. Полные решения задач имеют соответственно форму $\{M_H(x), F_{1, \max}\}$, $\{M_B(x), F_{2, \max}\}$, где $M_H(x)$ — множество точек x решения нижней, $M_B(x)$ — множество точек x решения верхней граничной задачи.

Шаг 6. Проверяется наличие либо отсутствие пересечения найденных в результате решения задач (20) и (21) множеств $M_H(x)$, $M_B(x)$.

Итог. Если в результате работы приведенного алгоритма выяснилось, что полностью определенная задача условной оптимизации (3), (4) имеет решение, а производная от нее интервальная задача

(5), (6) имеет нижнюю и верхнюю граничные задачи, обладающие решениями, причем множества этих решений пересекаются, то задача оптимизации (3), (4) является макроустойчивой. В противном случае задача (3), (4) не макроустойчива.

4. Микроустойчивость задачи условной оптимизации

Снова обратимся к полностью определенной задаче условной оптимизации (3), (4). Опишем метод установления микроустойчивости этой задачи.

Полностью определенная задача условной оптимизации (3), (4) по определению (см. п. 1) является микроустойчивой, если она макроустойчива и, кроме того, существует пара решений (x', x'') , где $x' = (x'_1, \dots, x'_n)$ — некоторое решение задачи (3), (4), а $x'' = (x''_1, \dots, x''_n)$ — некоторое решение производной от нее интервальной задачи условной оптимизации (5), (6), расстояние между которыми $D(x', x'')$ не превосходит заданной достаточно малой величины d . Из этого определения напрямую вытекает следующий алгоритм проверки произвольной полностью определенной задачи условной оптимизации (3), (4) на микроустойчивость.

Шаг 1. С помощью 6-шагового алгоритма, изложенного в п. 3, проверяем задачу (3), (4) на макроустойчивость. В случае отрицательного результата проверки (задача (3), (4) не макроустойчива) — конец алгоритма с выводом: задача (3), (4) не является микроустойчивой. В случае положительного результата проверки (задача (3), (4) макроустойчива) — переход к шагу 2.

Шаг 2. Выбираем точку решения $x' = (x'_1, \dots, x'_n)$ задачи (3), (4), найденную на шаге 1. Добавляем к ней какую-нибудь точку решения $x'' = (x''_1, \dots, x''_n)$ интервальной задачи (5), (6), также найденную на шаге 1. В результате получаем пару решений (x', x'') указанных двух задач.

Шаг 3. Вычисляем расстояние $D(x', x'')$ между точками решения x', x'' указанных двух задач, используя для этого формулу

$$D(x', x'') = \sqrt{(x'_1 - x''_1)^2 + \dots + (x'_n - x''_n)^2}. \quad (23)$$

Шаг 4. Проверяем выполнение неравенства, сравнивающего расстояние $D(x', x'')$ с заданной достаточно малой величиной d :

$$D(x', x'') \leq d. \quad (24)$$

Если условие (24) выполнено, задача оптимизации (3), (4) объявляется микроустойчивой и конец алгоритма. В противном случае совершается переход к шагу 2, в котором теперь к точке решения $x' = (x'_1, \dots, x'_n)$ задачи (3), (4), найденной на шаге 1, добавляется какая-то другая точка решения $x'' =$

(x''_1, \dots, x''_n) задачи (5), (6) из числа найденных на шаге 1. В результате получаем новую пару решений (x', x'') и т. д.

Итог. Если в результате работы алгоритма после некоторого достаточного числа шагов получена пара решений (x', x'') , удовлетворяющая неравенству (24), процедура останавливается и задача (3), (4) объявляется микроустойчивой. В противном случае процедура также останавливается, но задача (3), (4) признается не обладающей свойством микроустойчивости.

Заключение

В настоящей статье показано, что проблема оптимизации неполностью определенных функций не может ограничиться лишь отысканием точки оптимума и значения в ней нашей функции, но и должна включать в себя задачу определения устойчивости найденного оптимума. Последнее означает, что при небольшом варьировании параметров оптимизируемой функции ее оптимум должен по-прежнему существовать и находиться в точке, близкой к точке исходного оптимума. Для установления устойчивости оптимума неполностью определенных функций предложена специальная эффективная методика, основанная на аппарате интервальной математики.

Список литературы

1. **Первозванский А. А.** Математические модели в управлении производством. М.: Наука, 1975. 616 с.
2. **Левин В. И.** Интервальное дискретное программирование // Кибернетика и системный анализ. 1994. № 6. С. 91—103.
3. **Libura M.** Integer Programming Problems with Inexact Objective Function // Control and Cybernetic. 1980. Vol. 9, N 4. P. 189—202.
4. **Тимохин С. Г., Шапкин А. В.** О задачах линейного программирования в условиях неточных данных // Экономика и математические методы. 1981. Т. 17, № 5. С. 955—963.
5. **Рощин В. А., Семенова Н. В., Сергиенко И. В.** Вопросы решения и исследования одного класса задач неточного целочисленного программирования // Кибернетика. 1989. № 2. С. 42—46.
6. **Семенова Н. В.** Решение одной задачи обобщенного целочисленного программирования // Кибернетика. 1984. № 5. С. 25—31.
7. **Вощинин А. П., Сотиров Г. Р.** Оптимизация в условиях неопределенности. М.: Изд-во МЭИ, 1989. 224 с.
8. **Левин В. И.** Интервальные методы оптимизации систем в условиях неопределенности. Пенза: Изд-во Пензенского технологического института, 1999. 95 с.
9. **Алефельд Г, Херцбергер Ю.** Введение в интервальные вычисления. М.: Мир, 1987. 360 с.
10. **Юдин Д. Б., Гольдштейн Е. Г.** Задачи и методы линейного программирования. М.: Сов. радио, 1964. 350 с.
11. **Корбут А. А., Финкельштейн Ю. Ю.** Дискретное программирование. М.: Наука, 1969. 280 с.
12. **Левин В. И.** Структурно-логические методы исследования сложных систем. М.: Наука, 1987. 304 с.
13. **Левин В. И.** Дискретная оптимизация в условиях интервальной неопределенности // Автоматика и телемеханика. 1992. № 7. С. 97—106.

Э. Ю. Орехов, канд. физ.-мат. наук, доц., Уфимский государственный авиационный технический университет, e-mail: emil.orekhov@bk.ru

Равновероятная генерация целочисленных задач прямоугольного раскроя-упаковки

Предложен способ случайной равновероятной генерации целочисленных задач прямоугольного раскроя-упаковки, который может быть использован для оценки эффективности эвристических алгоритмов решения таких задач.

Ключевые слова: задача прямоугольного раскроя-упаковки, случайная равновероятная генерация, эвристический алгоритм, рекуррентные соотношения, условные вероятности

E. Yu. Orekhov

Generating Instances to Test Heuristic Algorithms for Rectangular Cutting-Packing

In this paper we suggest a method of equiprobable generation of instances for the integer rectangular cutting-packing problem. The method is intended for estimating the efficiency of heuristic cutting-packing algorithms.

Keywords: rectangular cutting-packing problem, random equiprobable generation, heuristic algorithm, recurrence relation, conditional probability

Наличие большого числа эвристических алгоритмов решения задач раскроя-упаковки делает актуальным сравнение этих алгоритмов с точки зрения выбранных критериев их эффективности. Сложившаяся практика тестирования таких алгоритмов описана в обзоре [1]. Выделяются два типа критериев эффективности: временной и качественный. Последний тип, характеризующий качество решения собственно решаемой задачи, представляет наибольший интерес. Оценивание же выбранного показателя качества основано на тестировании алгоритма на некоторой конечной выборке задач данного класса с дальнейшей интерпретацией полученных результатов. В работе [1] выделены три чаще всего используемых способа формирования такой выборки:

- использование задач, возникающих в ходе реального процесса;
- отбор задач экспертами и создание соответствующей библиотеки, примером которой является библиотека задач прямоугольного раскроя-упаковки [2];
- генерация задач данного класса с помощью так или иначе организованного случайного генератора, пример которого описан в работе [3].

Однако использование любого из описанных способов наталкивается, по крайней мере, на две трудности:

1) понятия класса решаемых задач и репрезентативности выборки оказываются не связанными друг с другом;

2) отсутствует единый общепринятый подход к тестированию эвристических алгоритмов — каж-

дый исследователь использует собственные показатели качества алгоритма и процедуры их оценки, что приводит к проблемам сопоставимости результатов, полученных разными авторами.

Указанные трудности могут быть преодолены с помощью предложенного в работе [4] подхода к определению качества эвристического алгоритма. Подход основан на введении характеристики качества алгоритма, интерпретируемой как функция распределения некоторой случайной величины. В такой ситуации задача статистического оценивания введенной характеристики качества решается на основе равновероятной генерации точек области параметрического пространства данной массовой задачи, взаимно однозначно соответствующих индивидуальным задачам этой массовой задачи. Поэтому конструирование процедуры равновероятной генерации точек упомянутой области является необходимым этапом оценки качества исследуемого эвристического алгоритма.

В настоящей работе предложена процедура равновероятной генерации индивидуальных задач целочисленной массовой задачи прямоугольного раскроя-упаковки, которая рассматривается в следующей постановке [5].

Имеются одинаковые прямоугольники (заготовки) R размерами $A \times B$, $A \leq B$, где A, B — натуральные числа, в которые нужно упаковать без перекрытия набор прямоугольников $\{r_1, \dots, r_n\}$ (прямоугольник r_i имеет размеры $c_i \times d_i$, $1 \leq c_i \leq d_i \leq K \leq A$, где c_i, d_i, K — натуральные числа, $i = 1, \dots, n$) так, чтобы число использованных заготовок оказалось наименьшим.

При заданных A, B, n, K индивидуальная задача определяется заданием размеров n упаковываемых прямоугольников, т. е. заданием набора пар натуральных чисел $\{(c_1, d_1), \dots, (c_n, d_n)\}$; отметим, что любая перестановка этих пар соответствует одной и той же индивидуальной задаче.

Построение процедуры равновероятной генерации

Задачу построения процедуры равновероятной генерации индивидуальных задач данной массовой задачи будем решать в два этапа.

На этапе I определим процедуру равновероятной генерации [5] точек некоторой области D параметрического пространства. Точки этой области представляют все индивидуальные задачи рассматриваемой массовой задачи, но условия взаимно однозначного соответствия этих индивидуальных задач и точек области D не выполнено.

На этапе II на основе подхода, предложенного в работе [6], определим процедуру равновероятной генерации индивидуальных задач с использованием их неравновероятных реализаций, полученных на этапе I.

Этап I. В параметрическом пространстве R^{2n} точек $(x_{11}, x_{12}, \dots, x_{n1}, x_{n2})$ определим область

$$D = \{(x_{11}, x_{12}, \dots, x_{n1}, x_{n2}) : \begin{cases} 1 \leq x_{11} \leq x_{12} \\ 1 \leq x_{12} \leq x_{22} \\ 1 \leq x_{21} \leq x_{22} \\ 1 \leq x_{22} \leq x_{32} \\ \dots \\ 1 \leq x_{i1} \leq x_{i2} \\ 1 \leq x_{i2} \leq x_{i+12} \\ \dots \\ 1 \leq x_{n1} \leq x_{n2} \\ 1 \leq x_{n2} \leq K \end{cases}, \quad (1)$$

где x_{ij} — натуральные числа, $i = 1, \dots, n, j = 1, 2, x_{i1}$ — меньший размер (ширина), а x_{i2} — больший размер (длина) упаковываемого прямоугольника, имеющего ранг i при ранжировании упаковываемых прямоугольников в порядке возрастания их длин.

Задача случайной равновероятной генерации точек области D сводится к получению независимых реализаций системы случайных величин $X = (X_{11}, X_{12}, \dots, X_{n1}, X_{n2})$ с законом распределения

$$P(X = x) = \begin{cases} \frac{1}{N}, & \text{если } x \in D, \\ 0, & \text{если } x \notin D, \end{cases} \quad (2)$$

где N — число точек в области D , определяемое как

$$N = \sum_{x_{n2}=1}^K \sum_{x_{i2}=1}^{x_{i+12}} \dots \sum_{x_{i2}=1}^{x_{i2}} \dots \sum_{x_{22}=1}^{x_{32}} \sum_{x_{21}=1}^{x_{22}} \sum_{x_{12}=1}^{x_{22}} \sum_{x_{11}=1}^{x_{12}} 1, \quad (3)$$

и представимое также в виде

$$N = \sum_{x_{n2}=1}^K x_{n2} \dots \sum_{x_{i2}=1}^{x_{i+12}} x_{i2} \dots \sum_{x_{22}=1}^{x_{32}} x_{22} \sum_{x_{12}=1}^{x_{22}} x_{12}. \quad (4)$$

Для получения процедуры определения N вычислим

$$S_k(y_{k+1}) = \sum_{y_k=1}^{y_{k+1}} y_k \dots \sum_{y_{k-1}=1}^{y_k} y_{i2} \dots \sum_{y_2=1}^{y_3} y_2 \sum_{y_1=1}^{y_2} y_1 \quad (5)$$

для натуральных k, y_{k+1} , а также

$$I_n(j-1) = \sum_{i=1}^n \prod_{l=0}^{j-1} (i+l) \quad (6)$$

для натуральных n, j .

Отметим, что в соответствии с [7]

$$I_n(j-1) = \frac{1}{j+1} \prod_{l=0}^j (n+l). \quad (7)$$

Используя (6), легко показать, что

$$\sum_{i=1}^n i I_i(j-1) = \frac{1}{j+1} I_n(j+1) - I_n(j), \quad (8)$$

а так как

$$S_1(y_2) = \sum_{y_1=1}^{y_2} y_1 = I_{y_2}(0),$$

то индукцией по k с использованием (8) нетрудно показать, что

$$S_k(y_{k+1}) = \sum_{i=1}^k (-1)^{i+1} a_k(2k-2-(i-1)) \times I_{y_{k+1}}(2k-2-(i-1)), \quad (9)$$

где коэффициенты $a_k(j), j = k-1, \dots, 2k-2$, определяются следующими рекуррентными соотношениями:

$$\begin{aligned} a_1(0) &= 1; \\ a_k(2k-2-(i-1)) &= \begin{cases} \frac{1}{2k-2} a_{k-1}(2k-4), & \text{если } i=1, k>1; \\ a_{k-1}(2k-2-(i-1)-1) + \frac{1}{2k-2-(i-1)} \times \\ \times a_{k-1}(2k-2-(i-1)-2), & \text{если } 1 < i \leq k-1, k>2; \\ a_k(k-1) = 1, & \text{если } i=k. \end{cases} \end{aligned} \quad (10)$$

Таким образом, в силу (4), (5) имеем

$$N = S_n(K), \quad (11)$$

где $S_n(K)$ вычисляем в соответствии с выражением (9), используя соотношения (7) и (10), и, следовательно, закон распределения (2) рассматриваемой системы дискретных случайных величин X можно считать полностью определенным, а генерация значений входящих в эту систему случайных величин в соответствии с (2) обеспечивает для любой точки области D одну и ту же вероятность попадания в выборку.

Для задания способа равновероятной генерации точек области D в соответствии с (2) необходимо определить набор условных законов распределения случайных величин, входящих в систему X , что можно сделать, зная закон распределения X .

Определим вначале безусловные законы распределения входящих в X подсистем случайных величин вида $(X_{i1}, X_{i2}, \dots, X_{n1}, X_{n2})$ и $(X_{i2}, \dots, X_{n1}, X_{n2})$. Так как, в силу (5)

$$\sum_{x_{i1}=1}^{x_{i2}} 1 = x_{i2},$$

$$\sum_{x_{i-12}=1}^{x_{i2}} \sum_{x_{i-11}=1}^{x_{i-12}} \dots \sum_{x_{i2}=1}^{x_{22}} \sum_{x_{i1}=1}^{x_{i2}} 1 = S_{i-1}(x_{i2}), \quad i = 2, \dots, n,$$

$$\sum_{x_{i1}=1}^{x_{i2}} \sum_{x_{i-12}=1}^{x_{i2}} \sum_{x_{i-11}=1}^{x_{i-12}} \dots \sum_{x_{i2}=1}^{x_{22}} \sum_{x_{i1}=1}^{x_{i2}} 1 = x_{i2} S_{i-1}(x_{i2}),$$

$$i = 2, \dots, n - 1,$$

то, учитывая (3),

$$P(X_{i2} = x_{i2}, \dots, X_{n2} = x_{n2}) = \begin{cases} \frac{x_{i2}}{N}, & \text{если } (x_{i2}, \dots, x_{n2}) \in D_{i2}, \\ 0, & \text{если } (x_{i2}, \dots, x_{n2}) \notin D_{i2}; \end{cases}$$

$$P(X_{i1} = x_{i1}, \dots, X_{n2} = x_{n2}) = \begin{cases} \frac{S_{i-1}(x_{i2})}{N}, & \text{если } (x_{i1}, \dots, x_{n2}) \in D_{i1}, \\ 0, & \text{если } (x_{i1}, \dots, x_{n2}) \notin D_{i1}; \quad i = 2, \dots, n; \end{cases}$$

$$P(X_{i2} = x_{i2}, \dots, X_{n2} = x_{n2}) = \begin{cases} \frac{x_{i2} S_{i-1}(x_{i2})}{N}, & \text{если } (x_{i2}, \dots, x_{n2}) \in D_{i2}, \\ 0, & \text{если } (x_{i2}, \dots, x_{n2}) \notin D_{i2}, \quad i = 2, \dots, n - 1; \end{cases}$$

$$P(X_{n2} = x_{n2}) = \begin{cases} \frac{x_{n2} S_{n-1}(x_{n2})}{N}, & 1 \leq x_{n2} \leq K, \\ 0, & \text{в противном случае,} \end{cases}$$

где

$$D_{i1} = \{(x_{i1}, x_{i2}, \dots, x_{n1}, x_{n2}) : \begin{cases} 1 \leq x_{i1} \leq x_{i2} \\ \dots \\ 1 \leq x_{n1} \leq x_{n2} \\ 1 \leq x_{n2} \leq K \end{cases}\}$$

для $i = 1, \dots, n$ (отметим, что $D_{11} = D$);

$$D_{i2} = \{(x_{i2}, \dots, x_{n1}, x_{n2}) : \begin{cases} 1 \leq x_{i2} \leq x_{i+12} \\ \dots \\ 1 \leq x_{n1} \leq x_{n2} \\ 1 \leq x_{n2} \leq K \end{cases}\}$$

для $i = 1, \dots, n - 1$.

Поэтому условные законы распределения случайных величин, входящих в X , с учетом (11) получим в виде

$$P(X_{11} = x_{11}/X_{12} = x_{12}, \dots, X_{n2} = x_{n2}) = \begin{cases} \frac{1}{x_{12}}, & 1 \leq x_{11} \leq x_{12}, \\ 0, & \text{в противном случае;} \end{cases}$$

$$P(X_{12} = x_{12}/X_{21} = x_{21}, \dots, X_{n2} = x_{n2}) = \begin{cases} \frac{x_{12}}{x_{22} S_1(x_{22})}, & 1 \leq x_{12} \leq x_{22}, \\ 0, & \text{в противном случае;} \end{cases}$$

$$P(X_{i1} = x_{i1}/X_{i2} = x_{i2}, \dots, X_{n2} = x_{n2}) = \begin{cases} \frac{1}{x_{i1}}, & 1 \leq x_{i1} \leq x_{i2}, \\ 0, & \text{в противном случае;} \end{cases}$$

$$P(X_{i2} = x_{i2}/X_{i+11} = x_{i+11}, \dots, X_{n2} = x_{n2}) = \begin{cases} \frac{x_{i2} S_{i-1}(x_{i2})}{S_i(x_{i+12})}, & 1 \leq x_{i2} \leq x_{i+12}, \\ 0, & \text{в противном случае;} \end{cases}$$

$$P(X_{n1} = x_{n1}/X_{n2} = x_{n2}) = \begin{cases} \frac{1}{x_{n2}}, & 1 \leq x_{n1} \leq x_{n2}, \\ 0, & \text{в противном случае;} \end{cases}$$

$$P(X_{n2} = x_{n2}) = \begin{cases} \frac{x_{n2} S_{n-1}(x_{n2})}{S_n(K)}, & 1 \leq x_{n2} \leq K, \\ 0, & \text{в противном случае.} \end{cases}$$

Вычислив значения полученных условных вероятностей с использованием (7), (9), (10), равновероятную генерацию точек области D можно осуществить, последовательно генерируя $X_{n2}, X_{n1}, \dots, X_{12}, X_{11}$ в соответствии с их условными законами распределения и используя для генерации каждой из этих случайных величин метод обратных функций [8].

Этап II. Отметим, что некоторым индивидуальным задачам взаимно однозначно соответствуют каждой "своя" точка области D , а некоторым индивидуальным задачам соответствуют более одной точки из D . Пусть, например, $A = B = K = 3, n = 2$, тогда:

- индивидуальной задаче, определяемой набором $\{(1, 2), (2, 3)\}$, соответствует единственная точка $(1, 2, 2, 3)$ из D ;
- индивидуальной задаче, определяемой набором $\{(1, 3), (2, 3)\}$, соответствуют точки $(1, 3, 2, 3)$ и $(2, 3, 1, 3)$ из D .

Вообще, индивидуальным задачам, задаваемым наборами прямоугольников, у которых все длины различны, будут соответствовать каждой единственная точка области D . Если же индивидуальная задача задается набором прямоугольников, некоторые из которых имеют одинаковые длины, то ей будет соответствовать, вообще говоря, более одной точки из D . Подсчитаем число целочисленных то-

чек области D , представляющих такую индивидуальную задачу. Именно, пусть $x = (x_{11}, x_{12}, \dots, x_{k_1 1}, x_{k_1 2}, x_{k_1+1, 1}, x_{k_1+1, 2}, \dots, x_{k_1+k_2, 1}, x_{k_1+k_2, 2}, \dots, x_{k_1+\dots+k_p, 1}, x_{k_1+\dots+k_p, 2})$ — точка области D такая, что $x_{12} = \dots = x_{k_1 2}, x_{k_1+1, 2} = \dots = x_{k_1+k_2, 2}, \dots, x_{k_1+\dots+k_{l-1}, 2} = \dots = x_{k_1+\dots+k_p, 2}$, и $k_j \geq 1, j = 1, \dots, l; \sum_{j=1}^l k_j = n$.

Пусть теперь

- во множестве чисел $x_{11}, \dots, x_{k_1 1}$ имеется m_1 подмножеств, содержащих q_{11}, \dots, q_{1m_1} одинаковых

чисел, $\sum_{r=1}^{m_1} q_{1r} = k_1, q_{1r} \geq 1, r = 1, \dots, m_1$; всего

$$\text{имеется } C_{k_1}^{q_{11}} C_{k_1 - q_{11}}^{q_{12}} \dots C_{k_1 - \sum_{r=1}^{m_1-1} q_{1r}}^{q_{1m_1}} = \frac{k_1!}{q_{11}! \dots q_{1m_1}!} =$$

$= A_1(x)$ способов образовать такое число подмножеств с данным наполнением;

- во множестве чисел $x_{k_1+\dots+k_{j-1}, 1}, \dots, x_{k_1+\dots+k_p, 1}$ имеется m_j подмножеств, содержащих q_{j1}, \dots, q_{jm_j}

одинаковых чисел, $\sum_{r=1}^{m_j} q_{jr} = k_j, q_{jr} \geq 1, r = 1, \dots, m_j$;

$$\text{всего имеется } C_{k_j}^{q_{j1}} C_{k_j - q_{j1}}^{q_{j2}} \dots C_{k_j - \sum_{r=1}^{m_j-1} q_{jr}}^{q_{jm_j}} =$$

$$= \frac{k_j!}{q_{j1}! \dots q_{jm_j}!} = A_j(x) \text{ способов образовать такое}$$

число подмножеств с данным наполнением;

- во множестве чисел $x_{k_1+\dots+k_{l-1}, 1}, \dots, x_{k_1+\dots+k_p, 1}$ имеется m_l подмножеств, содержащих q_{l1}, \dots, q_{lm_l}

одинаковых чисел, $\sum_{r=1}^{m_l} q_{lr} = k_l, q_{lr} \geq 1, r = 1, \dots, m_l$;

$$\text{всего имеется } C_{k_l}^{q_{l1}} C_{k_l - q_{l1}}^{q_{l2}} \dots C_{k_l - \sum_{r=1}^{m_l-1} q_{lr}}^{q_{lm_l}} =$$

$$= \frac{k_l!}{q_{l1}! \dots q_{lm_l}!} = A_l(x) \text{ способов образовать такое}$$

число подмножеств с данным наполнением.

Следовательно, число точек области D , представляющих данную индивидуальную задачу, есть

$$A(x) = \prod_{j=1}^l A_j(x). \quad (12)$$

Таким образом, если известна точка x , представляющая некоторую индивидуальную задачу, то можно определить число $A(x)$ точек области D , представляющих эту индивидуальную задачу.

Занумеруем теперь все индивидуальные задачи данной массовой задачи числами от 1 до N' , где N' — общее число индивидуальных задач, и пусть $u = u(x)$ — номер индивидуальной задачи, представляемой в D точкой x , а $x(u)$ — любая точка x , соответствующая индивидуальной задаче с номером u . Тогда равновероятная генерация точек в D на этапе I индуцирует реализацию случайной величины U с законом распределения

$$P(U = u) = p_u = \frac{A(x(u))}{N}, u = 1, \dots, N'.$$

Отметим, что $p = \min_{u=1, \dots, N'} p_u = \frac{1}{N}$, поэтому можно

записать $p_u = p + \Delta_u, \Delta_u = \frac{A(x(u)) - 1}{N}$.

Пусть теперь Z — дискретная случайная величина с законом распределения

$$P(Z = z) = \frac{1}{N'}, z = 1, \dots, N'.$$

В соответствии с [6], реализации равновероятно распределенной дискретной случайной величины Z можно получить с помощью алгоритма EQPR(Z).

Алгоритм EQPR(Z)

1. Генерируется реализация случайной величины U в соответствии с ее законом распределения; пусть в результате исполнения данного шага получено значение $U = u, u = 1, \dots, N'$.

2. Генерируется реализация значения вспомогательной дискретной случайной величины Y_u в соответствии с ее законом распределения

$$P(Y_u = 0) = \frac{\Delta_u}{p_u} = \frac{A(x(u)) - 1}{A(x(u))}, P(Y_u = 1) = \frac{p}{p_u} = \frac{1}{A(x(u))}.$$

3. Если $Y_u = 0$, то возврат к п. 1 данного алгоритма; иначе полагается $Z = u$, и алгоритм заканчивает работу.

Резюмируя изложенное, можно сформулировать следующую процедуру равновероятной генерации индивидуальных задач целочисленной массовой задачи прямоугольного раскрытия-упаковки в приведенной постановке.

1. Осуществить равновероятную генерацию точки $x = (x_{11}, x_{12}, \dots, x_{n1}, x_{n2}) \in D$, используя значения условных вероятностей, полученных на этапе I, любым подходящим способом.

2. Если $x_{12} \neq x_{22} \neq \dots \neq x_{n2}$, то полученная точка x есть представление равномерно генерированной индивидуальной задачи данной массовой задачи; конец работы процедуры.

3. В соответствии с (12) вычислить величину $A(x)$ так, как это указано в описании этапа II; генерировать случайную величину Y в соответствии с законом распределения

$$P(Y = 0) = \frac{A(x) - 1}{A(x)}, P(Y = 1) = \frac{1}{A(x)}.$$

4. Если $Y = 0$, то перейти к п. 1 данной процедуры; иначе полученная точка x есть представление равномерно генерированной индивидуальной задачи данной массовой задачи — конец работы процедуры.

В заключение отметим, что в качестве критерия эффективности полученной процедуры может быть выбрано математическое ожидание числа итераций данной процедуры в расчете на одну сгенерированную индивидуальную задачу. Очевидно, значение этого критерия будет тем лучше (ближе к единице), чем меньше доля индивидуальных задач, имеющих совпадающие длины размещаемых прямоугольников. В рассмотренной постановке задачи это означает, что эффективность построенного генератора будет тем выше, чем меньше параметр n и больше параметр k .

Список литературы

1. Мухачева Э. А. Задачи оптимального раскрой и родственная проблематика после 1971 г. / В кн.: Канторович Л. В., Залгаллер В. А. Рациональный раскрой промышленных материалов. — 3-е изд., испр. и доп. СПб.: Невский диалект, 2012. С. 250—292.
2. Bortfeld A. A genetic algorithm for the two-dimensional strip packing problem with rectangular pieces // Eur. J. Oper. Res. 2006. V. 172 (3). P. 814—837.
3. Schwerin P., Waechter G. The bin-packing problem: A problem generator and some numerical experiments with FFD packing and MTP // International transactions in operational research. 1997. V. 4, N 5—6. P. 337—389.
4. Орехов Э. Ю., Орехов Ю. В. Об оценке качества эвристического алгоритма на конечной массовой задаче // Информационные технологии. 2011. № 7. С. 28—33.
5. Орехов Э. Ю., Орехов Ю. В. Способ равномерно генерации целочисленных задач прямоугольного раскроя-упаковки // Принятие решений в условиях неопределенности: Межвуз. науч. сб. Вып. 4. Уфа: УГАТУ, 2008. С. 132—137.
6. Orekhov E. Yu., Orekhov Yu. V. The equiprobable generation of instances for the integer problem of scheduling jobs between unrelated parallel machines // Proc. of the 14th International workshop on computer science and informational technologies CSIT'12. Vol. 2. Ufa—Hamburg—Norwegian Fjords, 2012. P. 108—110.
7. Orekhov E. Yu., Orekhov Yu. V. Equiprobable Generation of the Integer One-Dimensional Cutting-Packing Problem Instances // Proc. of the 5th International Workshop on Computer Science and Informational Technologies. Vol. 2. Ufa, 2003. P. 41—42.
8. Соболев И. М. Численные методы Монте-Карло. М.: Наука, 1973. 311 с.

УДК 004.624

С. С. Гаврюшин, д-р техн. наук, проф., зав. каф., Г. Б. Евгеньев, д-р техн. наук, проф., МГТУ им. Н. Э. Баумана, e-mail: g.evgenyev@mail.ru

Многокритериальная оптимизация в жизненном цикле изделий

В настоящее время стоит важная задача повышения конкурентоспособности и эффективности машиностроительных производств. Для решения этой задачи необходимо создание интегрированных компьютерных систем автоматизации, охватывающих все этапы жизненного цикла изделий. Сейчас достигнуты значительные успехи в разработке программного и информационного обеспечения таких систем, включая интеграцию конструкторского и технологического проектирования. Вместе с тем этого нельзя сказать о математическом обеспечении. Наибольшие перспективы в этом отношении имеют методы многокритериальной оптимизации, которые пока используются только при проектировании изделий.

Описан пример применения метода, что может послужить активизации внедрения его в промышленность.

Ключевые слова: технология управления жизненным циклом изделия, многокритериальная оптимизация

S. S. Gavrushin, G. B. Evgenyev

Multicriteria Optimization in the Life Cycle of Products

There is the important task of improving the competitiveness and efficiency of engineering industries in this country. To solve this problem it is necessary the creation of integrated computer systems automation, covering all stages of the product life cycle. There are made significant progress in the development of software and information support of such systems, including the integration of design and process planning. It is need to make such progress at the mathematical software. The best prospects in this regard are multi-criteria optimization techniques that are still used only in the design of products.

The article describe an example of the method, which may serve to enhance the implementation in industry.

Keywords: PLM, multi-criteria optimization

Введение

Важным направлением повышения конкурентоспособности машиностроения является оптимизация параметров конструкции изделий и технологических процессов их производства. Оптимизация параметров должна проводиться на всех этапах жизненного цикла изделий, в числе которых наиболее важными являются этапы проектирования. В процессе проектирования участвуют специалисты разных направлений: расчетчики, конструкторы, технологи. Эти специалисты имеют свои критерии оптимизации, которые часто имеют противоречивый характер. В этой связи большую актуальность приобретают многокритериальные методы оптимизации, которым посвящена настоящая статья.

Структура проектирования изделий и технологических процессов

В процессе создания изделие проходит ряд этапов своего жизненного цикла. На рис. 1 представлено этапы инженерного проектирования.

Рассмотрим задачу изготовления композиционного баллона давления, эскиз которого показан на рис. 2, методом планарной намотки [1]. Изделия, изготавливаемые намоткой, представляют особый класс, у которого этапы проектирования интегрируются в один. В соответствии с требованиями заказ-

чика, баллон должен отвечать следующим требованиям и ограничениям: допустимая масса не более M_0 (обозначим α_1 , см. ниже рис. 4); объем не более V_0 (обозначим α_2); рабочее давление не менее p . При этом габаритные размеры баллона должны соответствовать следующим параметрам: R_0 (обозначим α_3) — наружный радиус; l_0 (обозначим α_4) — габаритная длина.

Помимо этого в процессе проектирования выбираются все данные (α_i), необходимые для расчетов. В их число входят радиусы полярных отверстий r_{01} , r_{02} , и параметр l , определяющие расположение плоскости намотки. Для расчетов необходимы также размеры B_1 , B_2 и R (см. рис. 2).

Параметрами расчета являются $[\sigma_{\pm 1}]$, $[\sigma_{\pm 2}]$ — допускаемые напряжения при растяжении/сжатии композиционного материала в направлении вдоль/поперек армирующего волокна; E_1 , E_2 — модули упругости в направлении вдоль/поперек армирующих элементов слоя, соответственно; ν_{12} — коэффициент Пуассона; δ — ширина лента; $\rho_{к.м}$ — плотность композитного материала.

На основе этих параметров проводится расчет свойств проектируемого объекта. Эти свойства связаны с функциональными ограничениями

$$c_l^* \leq f_l(A) \leq \tilde{c}_l^{**}, l \in [1:t]. \quad (1)$$

Здесь $f_l(A)$ — некоторые функции от параметров $A = (\alpha_1, \alpha_2, \dots, \alpha_N)$, которые могут быть заданы явно, например, в базах знаний. Для генерации чертежей деталей и сборочной единицы рассчитывают все необходимые для этой цели параметры.

Критерии качества

Как следует из рис. 1, управление каждым этапом жизненного цикла продукции (ЖЦП) осуществляется с учетом соответствующих критериев. Критерием качества $\Phi_v(A)$ называется характеристика системы, которая связана с ее качеством монотонной зависимостью [2]. Иными словами, при прочих равных условиях система тем лучше, чем больше (меньше) критерий.

При многокритериальной оптимизации с использованием критериев различных фаз жизненного цикла необходимо построение интегрированной системы конструкторско-технологического проектирования, представленного на рис. 1.

Для простоты записи в дальнейшем будем считать, что все заданные критерии $\Phi_1(A), \dots, \Phi_M(A)$ желательно уменьшить:

$$\Phi_v(A) \rightarrow \min_A, v \in [1:M].$$

Следовательно, чем меньше значение критерия $\Phi_v(A)$, тем при прочих равных условиях лучше система. Формально любой критерий можно привести к такому виду, заменяя, если это нужно, Φ_v на $1/\Phi_v$ или на $-\Phi_v$.

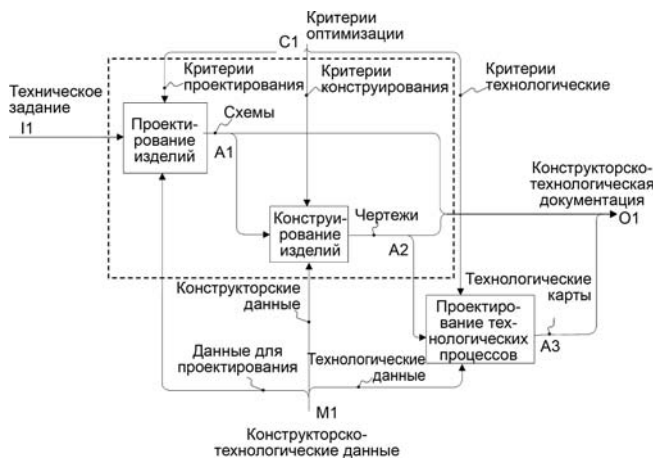


Рис. 1. Этапы проектирования изделий

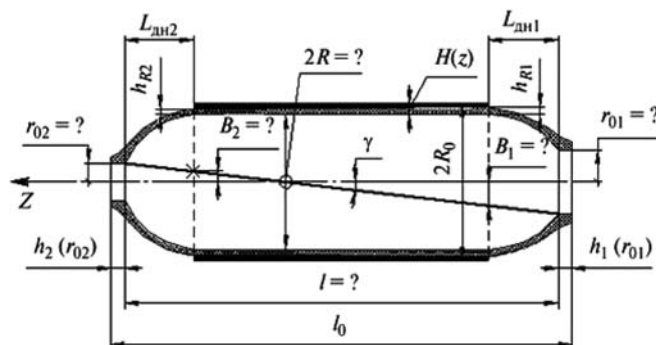


Рис. 2. Эскиз баллона давления

При конструкторско-технологическом проектировании композиционных баллонов давления можно принять критерии минимального отклонения от заданных параметров:

- $\Phi_1(A)$ — отклонение M_0 (ΔM_0);
- $\Phi_2(A)$ — отклонение V_0 (ΔV_0);
- $\Phi_3(A)$ — отклонение R_0 (ΔR_0);
- $\Phi_4(A)$ — отклонение l_0 (Δl_0).

Функциональное представление решаемой оптимизационной задачи в стандарте IDEF0 изображено на рис. 3.

Входными переменными являются технические требования заказчика, выходными — отклонения от технических требований в рассчитанном проекте. Управление проектированием осуществляется с помощью вектора *Con*, в состав компонентов которого входят r_{01} , r_{02} , l , B_1 , B_2 и R . В состав вектора *Mec* входят $[\sigma_{\pm 1}]$, $[\sigma_{\pm 2}]$, E_1 , E_2 , ν_{12} , δ и ρ .

Сформулировать математическую оптимизационную задачу при наличии нескольких критериев качества непросто, ибо критерии эти часто противоречат друг другу. Например, уменьшая массу машины (что часто очень желательно), мы в то же время уменьшаем ее прочность (что нежелательно). Или чрезмерное снижение стоимости изделия может привести к ухудшению других ее качеств.

Иногда полагают, что все дело в удачном выборе решающего критерия качества $\Phi(A)$, который должен соединить в себе (т. е. учесть) значения и важность каждого из индивидуальных критериев $\Phi_1(A)$, ..., $\Phi_M(A)$. В большинстве реальных задач такой подход себя не оправдывает.

Каждый из скалярных критериев оптимальности $\Phi_v(A)$, $v \in [1:M]$, принято называть *частным критерием оптимальности* [2]. Совокупность частных критериев оптимальности $\Phi(A) = (\Phi_1(A), \dots, \Phi_M(A))$ называют *векторным критерием оптимальности* [2]. Положим, что ставится задача минимизации каждого из частных критериев оптимальности $\Phi_1(A)$, ..., $\Phi_M(A)$ в одной и той же области допустимых значений D_α :

$$\Phi_v(A) \rightarrow \min_{A \in D_\alpha}, v \in [1:M]. \quad (2)$$

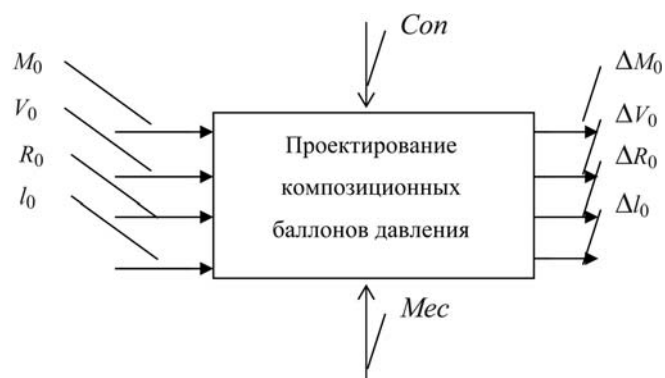


Рис. 3. Функциональная схема проектирования

Решение задачи многокритериальной оптимизации (2) в общем случае не является оптимальным ни для одного из частных критериев, а оказывается некоторым компромиссом для вектора $\Phi(A)$ в целом.

Обозначим P фронт Парето задачи (2). Множество P является подмножеством *пространства критериев* $\{\Phi\}$. Это пространство имеет размерность M (по числу частных критериев) и образуется M ортогональными осями координат, вдоль которых откладываются значения частных критериев оптимальности $\Phi_v(A)$, $v \in [1:M]$.

Векторный критерий оптимальности $\Phi(A)$ выполняет отображение множества допустимых значений D_α пространства варьируемых параметров в пространство критериев $\{\Phi\}$.

В постановке задачи многокритериальной оптимизации (2) фиксируется лишь множество допустимых значений D_α вектора варьируемых параметров и вектор частных критериев $\Phi(A) = (\Phi_1(A), \dots, \Phi_M(A))$. Этой информации недостаточно для однозначного решения задачи (2). Указанная информация позволяет лишь выделить множество Парето P (можно сказать, что решением задачи многокритериальной оптимизации в постановке (2) является множество Парето P). Для однозначного решения задачи (2) нужна дополнительная информация [2].

Метод исследования пространства параметров

Существует мнение [3, 7], что использование однокритериальных методов нередко заставляет специалистов искажать исходную постановку проблемы: подгонять реальные многокритериальные задачи под один самый важный критерий, строить необоснованные свертки критериев и т. д. При этом сведение многокритериальных задач в прокрустово ложе однокритериальных может привести к тому, что исходная проблема подменяется новой и в результате решается не та задача.

Чтобы избавить пользователя от необходимости сведения проблемы к однокритериальной разработаны различные методы. Одним из них является метод исследования пространства параметров (ИПП) [3]. В мировой литературе методам оптимизации уделяется много внимания. В то же время практически не исследована такая фундаментальная проблема, как определение допустимого множества решений. Для определения допустимого множества решений специалисту нужна помощь. Такая помощь — это одно из главных назначений метода исследования пространства параметров (метода ИПП). Метод ИПП основан на построении и анализе допустимого множества решений с помощью таблиц испытаний. Метод ИПП реализуется программным комплексом MOVI [8].

Процесс построения допустимого множества решений и его анализа проходит в интерактивном режиме, диалоге специалиста с компьютером. В ходе такого диалога многократно пересматриваются раз-

личные ограничения и в зависимости от этого определяются допустимые и Парето-оптимальные решения. Анализ допустимых решений дает уникальную информацию о ресурсах исследуемого объекта, чувствительности критериев к параметрам, взаимосвязи критериев работы параметрических, функциональных и критериальных ограничений и целесообразности их изменений. Чаще всего после такого анализа корректируется исходная постановка задачи. Нередки случаи, когда по результатам подобных исследований приходят к выводу о целесообразности коррекции математической модели объекта.

В методе ИПП постановка задачи оптимизации и ее решение — единый процесс. Интерактивное участие специалиста в процессе не только не представляет каких-либо трудностей, но и вызывает интерес. В этом плане метод является дружественным по отношению к человеку.

Формирование исходных данных применительно к конструкторско-технологическому проектированию композиционных изделий описано выше.

Для выполнения первого этапа процесса необходимо построить пространство параметров. Пространством параметров называется N -мерное пространство, состоящее из точек A с декартовыми координатами $(\alpha_1, \alpha_2, \dots, \alpha_N)$. Таким образом, каждой точке A пространства параметров соответствует конкретный набор параметров $(\alpha_1, \alpha_2, \dots, \alpha_N)$ и наоборот.

Как правило, проектировщики могут указать разумные пределы изменения каждого из параметров (рис. 4, а), которые называют *параметрическими ограничениями*:

$$\alpha_j^* \leq \alpha_j \leq \alpha_j^{**}, j \in [1:N].$$

При этом в пространстве параметров выделяется параллелепипед Π , N -мерный объем которого

$$V_{\Pi} = (\alpha_1^{**} - \alpha_1^*) \times \dots \times (\alpha_N^{**} - \alpha_N^*).$$

В дальнейшем нас будут интересовать только точки A , принадлежащие Π . Метод ИПП основан на зондировании параллелепипеда Π конечным числом пробных точек.

В качестве пробных точек рекомендуется использовать ЛП_т-последовательности [3], которые являются наиболее равномерно распределенными среди всех известных в настоящее время последовательностей.

Кроме параметрических ограничений, обычно в условия задачи включают *функциональные ограничения* (1). Если, например, функционирование системы описывается дифференциальными уравнениями, то $f_j(A)$ часто представляют собой функционалы, зависящие от интегральных кривых этих уравнений. Предполагается, что все функции $f_j(A)$ непрерывны в параллелепипеде Π .

Обозначим через G подмножество параллелепипеда Π , состоящее из точек A , удовлетворяющих

ограничениям (1). Множество G может быть любым замкнутым множеством, единственное ограничение: объем G должен быть положительным ($V_G > 0$).

Во избежание ситуации, когда некоторые значения критерия качества, по мнению проектировщиков, выходят за допустимые пределы вводят *критериальные ограничения*:

$$\Phi_v(A) \leq \Phi_v^{**}, v \in [1:M]. \quad (3)$$

Критериальное ограничение Φ_v^{**} — это худшее значение критерия $\Phi_v(A)$, которое проектировщик считает приемлемым. При этом формируется область цели (рис. 4, б).

Пусть D — множество точек A , которые удовлетворяют всем ограничениям, так что $D \subseteq G \subseteq \Pi$. Если множество D не пусто, то оно замкнуто.

Главная трудность при переходе к математической задаче многокритериальной оптимизации состоит в выборе критериальных ограничений Φ_v^{**} и в обеспечении непустоты множества допустимых точек D .

Обоснованный выбор критериальных ограничений невозможен без предварительного исследования системы: необходимо определить возможности системы по каждому из критериев в области G и, в частности, представить диапазоны изменения каждого из критериев.

Следующий этап заключается в составлении таблиц испытаний. Этот этап выполняется компьютером без вмешательства человека. Последовательно выбираются N_0 пробных точек A_1, \dots, A_{N_0} , равномерно расположенных в области G . В каждой из точек A_i вычисляются значения всех критериев $\Phi_v(A)$, $v \in [1:M]$ с проверкой на функциональные

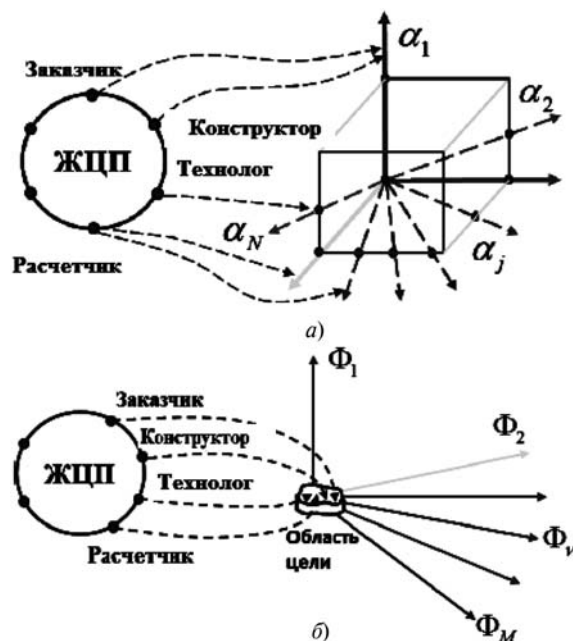


Рис. 4. Формирование пространства параметров (а) и области цели (б)

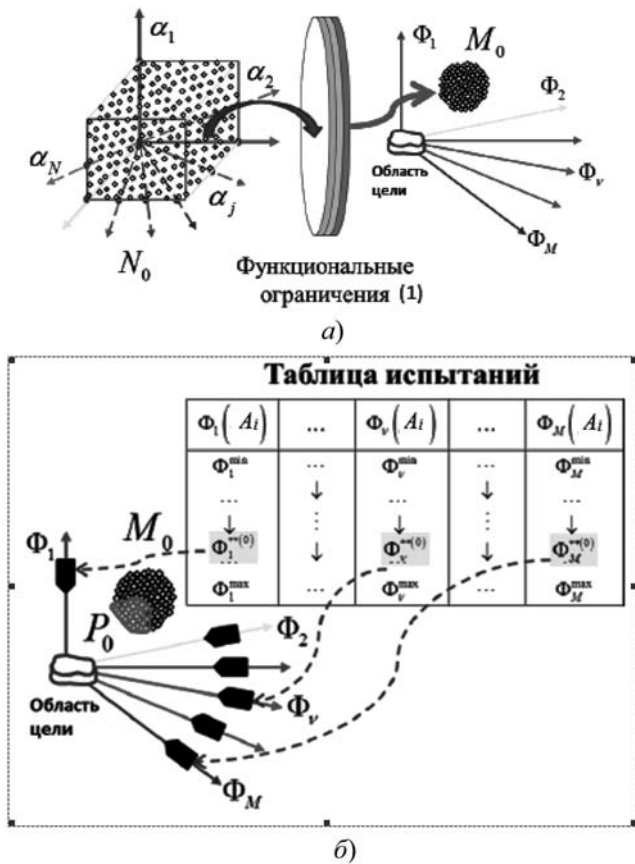


Рис. 5. Этапы формирования области G

ограничения. При этом в пространстве критериев формируется соответствующая область (рис. 5, а).

По каждому критерию составляется таблица испытаний (рис. 5, б), в которой значения $\Phi_v(A_1), \dots, \Phi_v(A_N)$ расположены в порядке возрастания

$$\Phi_v(A_{i_1}) \leq \Phi_v(A_{i_2}) \leq \dots \leq \Phi_v(A_{i_N}), v \in [1:M]$$

и указаны номера i_1, i_2, \dots, i_N соответствующих пробных точек (свой для каждого v). Таблица испытаний показывает не только приближенные значения максимума и минимума функции $\Phi_v(A)$ в области G : по таблице можно судить о частоте тех или иных значений $\Phi_v(A)$. В таблице испытаний значения критериев ранжируются в порядке ухудшения (лучшие по каждому критерию векторы расположены в верхней части таблицы).

Следующий этап процесса — проверка непустоты D . Этот этап выполняется автоматически, без вмешательства человека. Фиксируется какой-нибудь из критериев, например $\Phi_1(A)$, и рассматривается соответствующая ему таблица испытаний. Пусть s — число значений в этой таблице, удовлетворяющих выбранному критериальному ограничению Φ_1^{**} , так что

$$\Phi_1(A_{i_1}) \leq \dots \leq \Phi_1(A_{i_s}) \leq \Phi_1^{**}.$$

Путем перебора значений всех критериев в точках A_{i_1}, \dots, A_{i_s} нетрудно проверить, есть ли среди

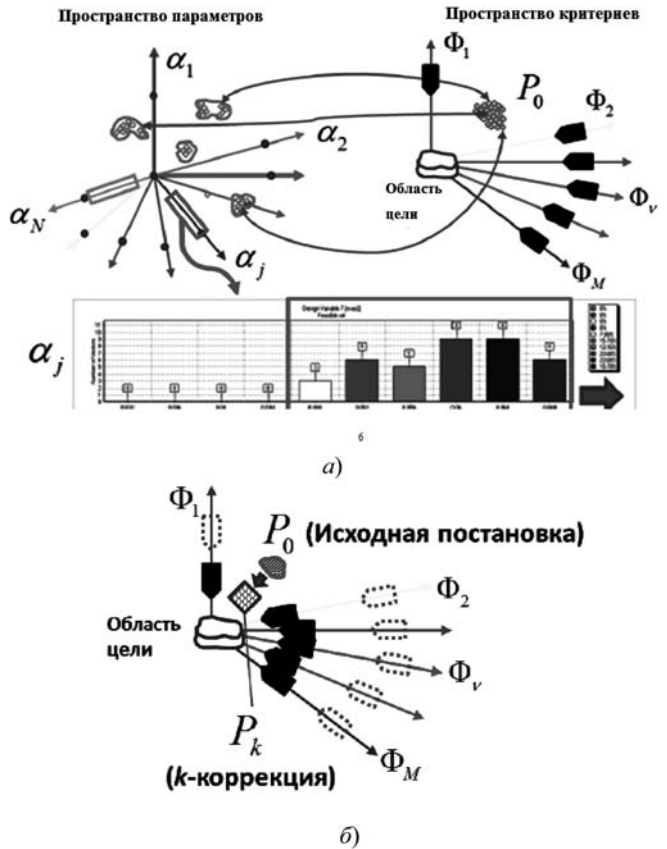


Рис. 6. Процесс оптимизации

этих точек хотя бы одна такая, в которой справедливы одновременно все неравенства

$$\Phi_v(A_{i_j}) \leq \Phi_v^{**}, v \in [1:M].$$

Если такая точка A_{i_j} существует, то множество D не пусто, и задача разрешима. В критериальном пространстве множеству D соответствует область P_0 на рис. 5, б. Она представляет собой подмножество множества M_0 .

P_0 допустимым решениям в критериальном пространстве соответствует P_0 точек в пространстве параметров (и наоборот). Диапазон изменения параметров разделяется на интервалы, каждому из которых ставится в соответствие число допустимых решений, принадлежащих данному интервалу, как это условно представлено на рис. 6, а. Рамкой отмечены расположения допустимых решений.

Оптимизация осуществляется путем коррекции границ вариации параметров, числа пробных точек и ужесточения критериальных ограничений до тех пор, пока мы либо попадем в область цели, либо максимально приблизимся к этой области. В случае достижения области цели процесс заканчивается (рис. 6, б). В противном случае необходимо выяснить причину отсутствия решения и найти приемлемый вариант, изменяя постановку задачи с привлечением всех участников жизненного цикла продукции.

Можно потребовать от участников ЖЦП уступок при назначении величин Φ_v^{**} . Если такие уступки невозможны, то необходимо вернуться к первому этапу и увеличить число N_0 пробных точек, чтобы повторить второй и третий этапы с таблицами испытаний бóльшего объема.

Наконец, если при неоднократном увеличении N_0 точки $A_{i,j}$, принадлежащие D , не обнаруживаются, то есть все основания считать, что выбранные критериальные ограничения Φ_v^{**} несовместны.

Заключение

Предложена методология использования многокритериальной оптимизации в жизненном цикле продукции. Рассмотрен метод исследования пространства параметров.

На основе методов многокритериальной оптимизации проведено конструкторско-технологическое проектирование композиционного баллона, удовлетворяющего следующим исходным данным и требованиям:

$$\begin{aligned} R_0 &= 0,15 \text{ м}; L_0 = 0,9 \text{ м}; V_0 = 0,04 \text{ м}^3; M_0 = 20 \text{ кг}; \\ p &= 30 \text{ МПа}; \delta = 0,005; [\sigma_{+1}] = 1600 \text{ МПа}; \\ [\sigma_{-1}] &= 700 \text{ МПа}; [\sigma_{+2}] = 60 \text{ МПа}; [\sigma_{-2}] = 170 \text{ МПа}; \\ E_1 &= 48 \text{ ГПа}; E_2 = 10 \text{ ГПа}; \\ \nu_{12} &= 0,25 \text{ и } \rho_{к.м} = 2200 \text{ кг/м}^3. \end{aligned}$$

В результате при значениях варьируемых параметров $r_{01} = 0,03769 \text{ м}$, $r_{02} = 0,02969 \text{ м}$, $B_1 = 0,0289 \text{ м}$, $B_2 = 0,0209 \text{ м}$, $l = 0,7309 \text{ м}$, $R = 0,133 \text{ м}$ получены следующие результирующие данные: $M_0 = 19,891 \text{ кг}$; $V_0 = 0,039 \text{ м}^3$; $R_0 = 0,1499 \text{ м}$; $l_0 = 0,9 \text{ м}$.

Список литературы

1. Данг Хоанг Минь, Гаврюшин С. С., Семисалов В. И. Анализ и синтез процесса намотки композиционного баллона в рамках концепции управления жизненным циклом продукции // Известия высших учебных заведений. Машиностроение. 2012. № 7. С. 12–16.
2. Карпенко А. П. Методы оптимизации. [Информационный ресурс]. URL: <http://bigor.bmstu.ru>
3. Соболев И. М., Статников Р. Б. Выбор оптимальных параметров в задачах со многими критериями: учеб. пособие для вузов. 2-е изд. М.: Дрофа, 2006. 175 с.
4. Figueira J., Greco S., and Ehrgott M. Multiple Criteria Decision Analysis: State of the Art Surveys. Springer's International series in Operations Research and Management Science. New York: Springer Science + Business Media, 2005.
5. Jahn J. Vector Optimization — Theory, Applications, and Extensions. Berlin: Springer Verlag, 2004.
6. T'Kindt V., Billaut J. Multicriteria Scheduling — Theory, Models and Algorithms. Berlin: Springer Verlag, 2002.
7. Statnikov R., Matusov J. Multicriteria Analysis in Engineering. Dordrecht: Kluwer Academic Publishers, 2002.
8. Statnikov R., Statnikov A. The Parameter Space Investigation Method Toolkit. Boston: Artech House, 2011.
9. Yamamoto Y. Optimization over the efficient set — Overview // Journal of Global Optimization. 2002. N 22. P. 285–317.

УДК 004.8

Е. С. Амосова, мл. науч. сотр., Институт проблем химической физики РАН, г. Черноголовка,
e-mail: alene2000@mail.ru

Разработка онтологической модели предметной области "Инновационные разработки в научных исследованиях"

Представлена онтологическая модель предметной области "Инновационные разработки в научных исследованиях". Модель разработана как расширение CERIF-2010 с учетом действующего российского законодательства. Модель будет использована для создания семантического слоя в информационной системе "Научно-технические разработки РАН".

Ключевые слова: онтология, онтологическая модель, инновационные разработки в научных исследованиях

E. S. Amosova

Development of the Ontological Domain Model for Innovative Developments in Scientific Research

The ontological domain model for innovative developments in scientific research is presented in the article. The model is developed as an extension of CERIF-2010 in accordance with the existing Russian legislation. The model will be used for creating the semantic layer of the information system "Scientific-technical developments of RAS".

Keywords: ontology, ontological model, innovative developments in scientific research

В настоящее время использование онтологий для моделирования предметных областей современных информационных систем получает все более широкое распространение [1, 2], в частности, в системах сети Интернет. Онтологическая модель позволяет разработать модель метаданных, что значительно улучшает использование системы широким кругом пользователей с точки зрения организации взаимодействия и применения семантического поиска. Поэтому задача построения онтологической модели предметной области "Инновационные разработки в научных исследованиях" является актуальной и сложной научно-практической задачей. Сложность поставленной задачи определяется, в частности, наличием множества межпредметных и междисциплинарных связей и различными целями конечных пользователей системы: ученых, экспертов, бизнесменов, политиков, сотрудников общественных и коммерческих организаций.

Цель настоящей статьи — разработка и описание онтологической модели предметной области "Инновационные разработки в научных исследованиях".

Инновационный процесс со структурной точки зрения представляет собой комплекс последовательно взаимосвязанных действий по созданию, освоению и распространению инновации, в том числе, включает в себя деятельность, обеспечивающую создание инноваций, — научно-технические услуги, маркетинговые исследования, разработка бизнес-плана проекта, оценка экономической эффективности инновации, подготовка и переподготовка кадров, организационная и финансовая деятельность.

Онтология области научных исследований представляет собой концептуальную схему процесса научно-исследовательской деятельности. Научные исследования возможны только при наличии полной и достоверной информации и наборов данных, начиная с этапа подачи заявки и заканчивая этапом публикации рецензии на разработку. Информационные системы по текущим исследованиям (CRIS) должны собирать воедино всю информацию, лежащую в основе научных исследований.

Попытки создания единой модели данных для информационных систем по научным разработкам предпринимаются не первый год [3—5]. Работы на эту тему проводят также и в Институте проблем химической физики РАН (ИПХФ РАН) [6, 7].

Составление словаря предметной области

При создании онтологии определенной предметной области, в первую очередь, следует составить глоссарий, т. е. список всех терминов, которые имеют отношение к данной предметной области. На данном этапе к работе привлекают экспертов в предметной области, которые определяют необходимость введения в онтологию того или иного термина.

Основным источником терминов (концептов) для словаря служит состав документов, сохраняе-

мых в базе данных (БД) информационной системы "Научно-технические разработки РАН", в первую очередь, "Резюме", "Технологическое предложение" и "Опросный лист", и рекомендации CERIF 1.3 [8], где определены объекты и связи между ними. Данные документы формируются на основе структуры БД, разработанной как расширение CERIF 1.3 в соответствии с принципами, введенными в работе [9].

Ввиду большого объема информации, в рамках данной статьи представляется целесообразным не приводить полный список терминов словаря, а ограничиться фрагментом, который дает общее представление о предметной области научных исследований с учетом особенностей российских научно-исследовательских учреждений (определения даны в соответствии с открытыми словарями):

ПЕРСОНА — любой человек, тем или иным образом участвующий в процессе создания и продвижения инновационных разработок;

ДИРЕКТОР — первый, главный руководитель компании, фирмы, направления деятельности. Директор управляет компанией в соответствии с ее уставом, имеет право первой подписи на денежных документах и обладает правом приема и увольнения с работы сотрудников;

ЗАМЕСТИТЕЛЬ ДИРЕКТОРА — официальное название должности помощника начальника какого-нибудь учреждения;

РУКОВОДИТЕЛЬ ПОДРАЗДЕЛЕНИЯ — лицо, заведующее структурным подразделением организации;

НАУЧНЫЙ СОТРУДНИК — проводит научные исследования, совершенствует и разрабатывает новые теории, концепции и методы и осуществляет их практическое применение;

ИНЖЕНЕР — специалист с высшим техническим образованием;

РАЗРАБОТЧИК — специалист, занимающийся разработкой проекта (схемы, механизма, аппаратуры);

АВТОР ПУБЛИКАЦИИ — человек или учреждение, ответственные за создание письменной работы, чье имя или название указано на титульной странице или в другой части публикации. У одной работы может быть несколько совместных авторов;

РЕДАКТОР ПУБЛИКАЦИИ — лицо, обрабатывающее и подготовляющее, редактирующее текст для печати или оформляющее издание;

ЭКСПЕРТ — сведущее лицо, специалист, привлекаемый для того, чтобы высказать свое мнение, дать заключение по поводу какого-нибудь дела, вопроса;

РУКОВОДИТЕЛЬ ПРОЕКТА — лицо, возглавляющее проект;

ИССЛЕДОВАТЕЛЬ — лицо, занимающееся научными исследованиями;

ИНВЕСТОР — лицо, организация или государство, производящие инвестицию;

ЗАКАЗЧИК РАБОТ — лицо или организация, размещающие заказ на выполнение определенных работ;

ВЛАДЕЛЕЦ ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ — человек или организация, которая обладает правами собственности;

ПРОЕКТ — уникальный набор процессов, состоящих из скоординированных и управляемых задач с начальной и конечной датами, предпринятых для достижения цели;

ИННОВАЦИОННАЯ РАЗРАБОТКА — разработка новых или усовершенствованных технологий, видов продукции или услуг;

ИДЕЯ — мысленный образ чего-либо, понятие о чем-либо;

НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЕ РАБОТЫ — творческая деятельность, направленная на получение новых знаний и способов их применения;

ОПЫТНО-КОНСТРУКТОРСКИЕ РАБОТЫ — комплекс работ, выполняемых при создании или модернизации продукции: разработка конструкторской и технологической документации на опытные образцы (опытную партию), изготовление и испытания опытных образцов (опытной партии);

ОПЫТНО-ТЕХНОЛОГИЧЕСКИЕ РАБОТЫ — комплекс работ по созданию новых веществ, материалов и/или технологических процессов и по изготовлению технической документации на них;

РЕЗУЛЬТАТЫ НИОКР — новые знания или новые изделия и технологии, полученные в ходе научно-исследовательских и опытно-конструкторских работ;

ПРОДУКТ — материальные или нематериальные результаты проектов (физический продукт, программное обеспечение, результаты исследований);

ОРГАНИЗАЦИЯ — общественное объединение или государственное учреждение;

НИИ — научно-исследовательский институт: учреждение, занимающееся исследованиями в области науки и техники, разработкой НИР и НИОКР, разновидность института;

КОММЕРЧЕСКАЯ СТРУКТУРА — это юридическое лицо, имеющее основной целью получение прибыли;

ПРАВООБЛАДАТЕЛЬ — физическое или юридическое лицо, обладающее какими-либо определенными правами;

ЧИСЛЕННОСТЬ СОТРУДНИКОВ — число сотрудников данной организации;

НАПРАВЛЕНИЕ ДЕЯТЕЛЬНОСТИ — вид деятельности, которым занимается организация;

ПРОГРАММА ФИНАНСИРОВАНИЯ — источник финансовых ресурсов для проекта, программы или любой другой научной деятельности;

ПУБЛИКАЦИЯ — статья или иной материал, преданный гласности в средствах массовой информации;

ПАТЕНТ — свидетельство, выдаваемое изобретателю компетентным государственным органом и удостоверяющее его авторство и исключительное право на использование изобретения в течение определенного срока.

Создание модели предметной области

Составив список терминов предметной области, необходимо распределить, какие из этих понятий являются основными классами, какие подклассами, а какие — свойствами классов (рис. 1). В классы выделены ПЕРСОНА (Person), ОРГАНИЗАЦИЯ (Organization), ПРОЕКТ (Project), ПУБЛИКАЦИЯ (Publication), ПРОДУКТ (Product), ПАТЕНТ (Patent), ПРОГРАММА ФИНАНСИРОВАНИЯ (Funding), а разновидности организаций и публикаций являются подклассами. Для создания модели используется редактор Protege.

Должности и виды деятельности определяют свойства связей между основными классами (рис. 2). Так, например, между ПЕРСОНОЙ (Person) и НИИ (Research-Institute) — подклассом организаций, можно выделить следующие связи:

— имеет_директора (has_director), является_директором (is_director_of),

— имеет_заместителя_директора (has_deputy_director), является_заместителем_директора (is_deputy_director_of),

— имеет_руководителя_подразделения (has_head_of_department), является_руководителем_подразделения (is_head_of_department_of),

— имеет_научного_сотрудника (has_researcher), работает_научным_сотрудником_в (is_researcher_at),

— имеет_инженера (has_engineer), работает_инженером_в (is_engineer_at).

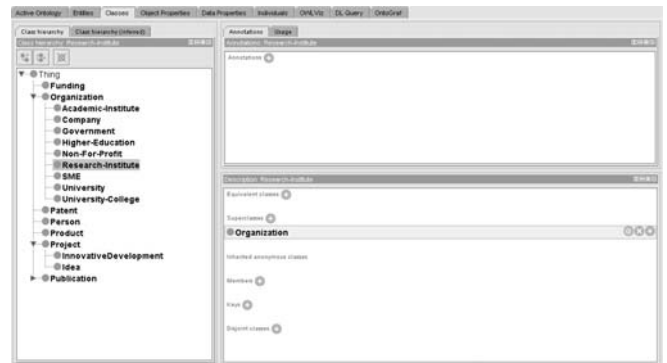


Рис. 1. Основные классы модели предметной области

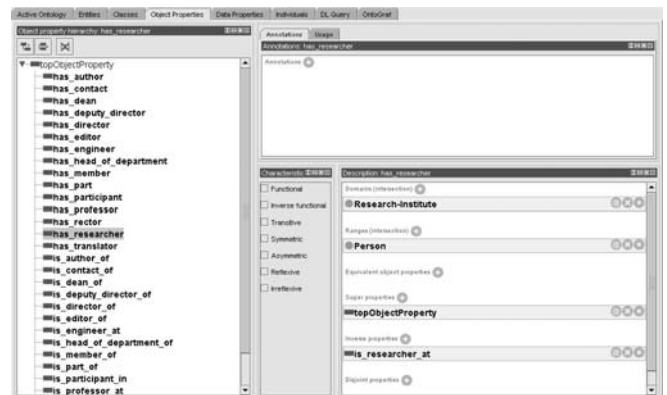


Рис. 2. Связи между классами модели предметной области

После того как заданы прямые и обратные связи между всеми классами, определяются свойства классов. Названия организаций и проектов, имена людей, выходные данные публикаций, идентификаторы проектов, даты начала и конца проектов, и другие свойства задаются как свойства данных.

Схема фрагмента модели

На схеме представлены классы модели: ПЕРСОНА (Person), ОРГАНИЗАЦИЯ (Organization), ПУБЛИКАЦИЯ (Publication), ПРОЕКТ (Project), ПРОДУКТ (Product), ПАТЕНТ (Patent), ПРОГРАММА ФИНАНСИРОВАНИЯ (Funding) (рис. 3).

Класс ОРГАНИЗАЦИЯ (Organization) разбит на подклассы по типам организаций, при этом на схеме акцент сделан на научно-исследовательских и образовательных учреждениях: НИИ (Research-Institute), ВУЗ (Higher-Education), УНИВЕРСИТЕТ (University), КОЛЛЕДЖ (University-College).

ОРГАНИЗАЦИИ (Organization) могут быть связаны друг с другом отношениями типа: включает/состоит в (has_member/is_member_of).

ПЕРСОНЫ (Person) связаны с ПУБЛИКАЦИЯМИ (Publication) отношениями типа:

является автором/имеет автора (is_author_of/has_author), является переводчиком/имеет переводчика (is_translator_of/has_translator), является редактором/имеет редактора (is_editor_of/has_editor).

ПЕРСОНЫ (Person) связаны с ПРОЕКТАМИ (Project) отношениями типа:

является участником/имеет участника (is_participant_in/has_participant).

ПЕРСОНЫ (Person) связаны с подклассом организаций НИИ (Research-Institute) отношениями типа:

имеет директора/является директором (has_director/is_director_of),

имеет заместителя директора/является заместителем директора

(has_deputy_director/is_deputy_director_of),

имеет руководителя подразделения/является руководителем подразделения

(has_head_of_department/is_head_of_department_of),

имеет научного сотрудника/работает научным сотрудником

в (has_researcher/is_researcher_at),

имеет инженера/работает инженером в (has_engineer/is_engineer_at).

ПЕРСОНЫ (Person) связаны с подклассами организаций ВУЗ (Higher-Education), УНИВЕРСИТЕТ (University), КОЛЛЕДЖ (University-College) отношениями типа:

является ректором/имеет ректора (is_rector_of/has_rector),

является деканом/имеет декана (is_dean_of/has_dean),

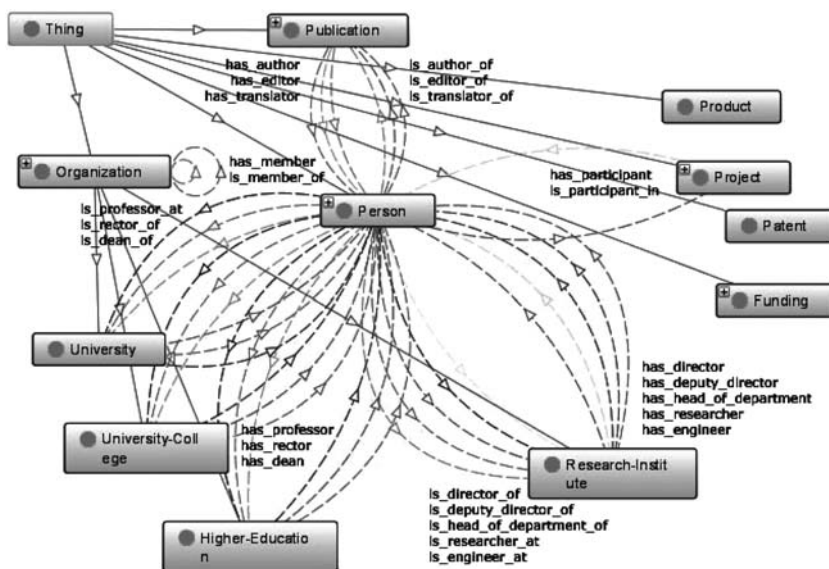


Рис. 3. Фрагмент модели предметной области

работает профессором/имеет профессора (is_professor_at/has_professor).

В работе поэтапно разработана онтологическая модель предметной области "Инновационные разработки в научных исследованиях". Разработанная модель будет использоваться для создания RDF хранилища и организации семантического поиска в информационной системе "Учет результатов интеллектуальной деятельности ИПХФ РАН".

Список литературы

1. Лапшин В. А. Онтологии в компьютерных системах. М.: Научный мир, 2010.
2. Gruber T. R. The role of common ontology in achieving sharable, reusable knowledge bases // Principles of Knowledge Representation and Reasoning. Proc. of the Second International Conference. Morgan Kaufmann, 1991. P. 601—602.
3. Бездушный А. Н., Жижченко А. Б., Кулагин М. В., Себряков В. А. Интегрированная система информационных ресурсов РАН и технология разработки цифровых библиотек // Программирование. 2000. № 4. С. 3—14.
4. Паринов С. И. Онлайн-будущее науки: наукометрическая сигнальная система. Серия препринтов "Количественный анализ в экономике". № 01. М.: ГУ ВШЭ, 2007. 52 с.
5. Зацман И. М. Информационные ресурсы для систем мониторинга в сфере науки // Системы и средства информатики. 2005. Вып. 15. С. 288—318.
6. Алдошин С. М., Амосова Е. С., Волохов В. М., Прохоров А. И., Туманов В. Е. Онтологическая модель предметной области информационной системы поддержки жизненного цикла инновационных разработок институтов РАН // Приложение к журналу "Открытое образование". Материалы XXXVII Международной конференции и дискуссионного научного клуба "Информационные технологии в науке, социологии, экономике и бизнесе". IT + SE'10. Осенняя сессия. 1—10 октября 2010 г., Украина, Крым, Ялта—Гурзуф. С. 145—146.
7. Амосова Е. С. Разработка онтологической модели предметной области "Инновационные разработки в научных исследованиях" // Материалы IX международной научно-практической конференции "Современные научные достижения — 2013". Чехия, Прага, 27 января—5 февраля 2013 г. С. 6—11.
8. CERIF 1.3 FDM: Model Introduction and Specification [электронный ресурс]. URL: http://www.eurocris.org/Uploads/Web%20pages/CERIF-1.3/Specifications/CERIF1.3_FDM.pdf
9. Бунчук М. Н. Последствия коммерциализации науки. // Коммерциализация технологий: теория и практика / Под ред. С. Я. Бабаскина, В. Г. Зинова. М.: Монолит, 2002. С. 117.

СТРУКТУРНЫЙ СИНТЕЗ И ПРИНЯТИЕ РЕШЕНИЙ STRUCTURAL SYNTHESIS AND DECISION MAKING

УДК 004.89

М. А. Николаева, канд. техн. наук, доц., **О. Ф. Зотова**, канд. техн. наук, доц.,
Н. В. Шолохова, канд. техн. наук,
Уфимский государственный авиационный технический университет,
e-mail: nadezhda.sholokhova@gmail.com

Система поддержки принятия решений при управлении депозитным портфелем физических лиц коммерческого банка

Предлагается описание общей структуры системы поддержки принятия решений (СППР) при управлении депозитным портфелем физических лиц коммерческого банка (ДПФЛ). Вводится совокупный риск управления как характеристика эффективности управления ДПФЛ. Подробно рассматривается "Сигнальная подсистема" в составе СППР, алгоритмы, математические методы.

Ключевые слова: сигнальная подсистема, совокупный риск управления, частотная логика, индикаторы, логистическая регрессия

M. A. Nikolaeva, O. F. Zotova, N. V. Sholokhova

DSS for the Commercial Bank Deposit Portfolio Management

A description of the general structure of decision support system for the commercial bank deposit portfolio management (DPFL) is proposed. The aggregate risk of management is being introduced as a characteristic of effective management of DPFL. Detail description of the "signal subsystem" within the DSS is provided as well as algorithms and mathematical methods.

Keywords: signal subsystem, aggregate risk of management, the frequency logic, indicators, logistic regression

Введение

Одним из результатов интеграции российских банков в мировую банковскую систему является ориентация на международные требования, предъявляемые к управлению рисками в банке. Так, введение нового соглашения "Базель-3" в первую очередь связано с необходимостью усиления контроля банковского риск-менеджмента.

В настоящее время наиболее разработанными являются подходы к идентификации и оценке кредитного риска, в частности, в Базельских соглашениях особое внимание к вопросу контроля за риском ликвидности было привлечено после мирового финансового кризиса конца 2000-х годов.

Таким образом, создание внутренних систем по контролю за рисками в банке, а также разработка новых методик для идентификации и оценки рисков является актуальной задачей.

Рынок программного обеспечения в настоящее время представлен большим количеством продуктов, разработанных как российскими, так и иностранными компаниями [8—13].

Практика риск-менеджмента в зарубежных банках начала развиваться значительно раньше, чем в российских банках. Об этом свидетельствует, в ча-

стности, эволюция международных Базельских соглашений. В связи с этим и программные продукты, предоставляемые зарубежными разработчиками, носят более комплексный характер охвата регулируемых рисков и реализуемых в них моделях.

В работе [15] представлен список программного обеспечения для управления рисками коммерческих организаций, разработанного зарубежными и отечественными фирмами.

Функциональность существующих автоматизированных банковских систем (АБС) позволяет как проводить комплексный анализ всех банковских рисков, так и решать отдельные задачи оценки выделенных рисков.

Совокупность розничных депозитных продуктов банка (депозитов), характеризующаяся условиями привлечения и объемом средств, образует *депозитный портфель физических лиц (ДПФЛ) коммерческого банка*.

Авторы предлагают ввести новые частные виды рисков, отражающие специфику управления ДПФЛ, а также учесть в них влияние процентного риска, риска ликвидности и риска потери доходности.

Для идентификации, анализа и оценки вводимых рисков разработан прототип системы, реали-

зующий поддержку принятия решений при управлении ДПФЛ коммерческого банка.

Целью разработки системы является повышение оперативности, качества и обоснованности принимаемых решений, связанных с определением момента внесения изменений в депозитный портфель, сравнением депозитных продуктов, размещением средств, привлеченных во вклады населения.

Бизнес-процесс "Управление депозитным портфелем физических лиц коммерческого банка"

Управление ДПФЛ может быть представлено как бизнес-процесс, в составе которого могут быть выделены несколько подпроцессов, связанных с отдельными этапами управления:

- с анализом состояния ДПФЛ и принятием решения о его пересмотре;
- с формированием структуры ДПФЛ;
- с планированием доходов от размещения депозитных средств в активы банка.

Рассматриваемая задача связана с повышением эффективности бизнес-процесса управления ДПФЛ. В качестве характеристики эффективности предлагается ввести показатель совокупного риска, который показывает, насколько, в среднем, неэффективны принятые решения.

Совокупный риск управления ДПФЛ

Введенный авторами совокупный риск управления представляет собой функцию от частных рисков, связанных с выделенными этапами управления привлеченными средствами:

$$R = \alpha_1 R_1 + \alpha_2 R_2 + \alpha_3 R_3, \quad (1)$$

где $\alpha_1, \alpha_2, \alpha_3$ — весовые коэффициенты, определяемые экспертом, $\sum_i \alpha_i = 1$;

R_1 — риск несвоевременного пересмотра депозитного портфеля, характеризует правильность и своевременность принятых менеджментом банка решений о внесении изменений в структуру депозитного портфеля;

R_2 — риск дефицита депозитных средств в составе пассивов, характеризует эффективность выбора альтернативы при изменении структуры депозитного портфеля;

R_3 — риск неэффективного размещения депозитных средств, учитывает отклонение процентного дохода от планового, получаемого в результате решения задачи размещения.

Предлагаемый подход позволяет оценить уровень риска операций, связанных со средствами, поступающими от вкладчиков, снизить уровень неопределенно-

сти параметров управления, а также охарактеризовать эффективность управления на каждом из этапов.

Мерами введенных рисков являются вероятности принятия неверных решений при управления ДПФЛ.

Структура системы поддержки принятия решений (СППР)

СППР состоит из трех основных подсистем, реализующих этапы управления ДПФЛ, связанные с идентификацией, оценкой и снижением совокупного риска.

"Сигнальная подсистема" предназначена для идентификации, оценки и снижения риска несвоевременного пересмотра ДПФЛ. В состав подсистемы входят следующие основные модули:

- "конкурентная группа";
- "депозитный портфель";
- формирования и просмотра сигналов,
- оценки качества сигналов,
- оценки вероятности пересмотра ДПФЛ,
- принятия решений, реализующий оценку риска R_1 и выдающий рекомендации лицу, принимающему решение (ЛПР) о пересмотре ДПФЛ.

"Подсистема формирования оптимальной структуры ДПФЛ" реализует модели оценки риска недостаточности депозитных средств в структуре пассивов и представлена модулями:

- построения жизненного цикла депозита,
- сравнения альтернатив для внесения в портфель,
- оценки риска R_2 .

"Подсистема размещения депозитных средств" позволяет оценить риск неэффективного распределения депозитных ресурсов в активы банка:

- эффективного размещения,
- принятия решений.

На рис. 1 приведена структура СППР.

В данной статье подробно рассматривается "Сигнальная подсистема".



Рис. 1. Структура системы поддержки принятия решений при управлении депозитным портфелем физических лиц коммерческого банка

Описание "Сигнальной подсистемы"

В "Сигнальной подсистеме" реализуется решение задачи оценки риска несвоевременного пересмотра ДПФЛ, эффективности бизнес-процесса, связанного с анализом состояния ДПФЛ и принятием решения об изменении структуры ДПФЛ.

Анализ состояния депозитного портфеля заключается в оценке меры влияния факторов внешней и внутренней среды на ДПФЛ, выявлении необходимости в изменении структуры портфеля, а также определении наилучшего момента для таких изменений. Под изменением структуры портфеля понимается введение нового депозитного продукта, останова приема средств по текущим вкладам и др.

Этап идентификации риска несвоевременного пересмотра ДПФЛ связан с выявлением внутренних и внешних факторов, влияющих на структуру ДПФЛ, а также построением системы индикаторов и формированием сигналов к пересмотру портфеля. Переход от факторов к системе индикаторов позволяет оценить меру влияния *изменений факторов* на структуру депозитного портфеля.

Для оценки качества сигналов использованы элементы частотной логики, основным инструментом которой выступают характеристики частоты элементов.

Для анализа и оценки риска несвоевременного пересмотра ДПФЛ строятся модели формирования совокупности сигналов на основе логистической регрессии.

Входными данными "Сигнальной подсистемы" являются следующие исходные данные:

- сведения о банках, работающих на рынке, включая название, контактную информацию, текущий объем активов и долю средств физических лиц в объеме пассивов;
- информация о депозитном портфеле рассматриваемого банка, включая название продукта, ставку процента, накопленную сумму средств и другие условия вкладов;
- данные о действующих депозитных продуктах других банков с описанием условий;
- таблицы данных для внешних и внутренних факторов за рассматриваемый период.

Постановка задачи

Для идентификации и оценки риска несвоевременного пересмотра ДПФЛ на предварительном этапе выделим исходную группу банков, близких по определенным показателям к исследуемому банку.

Оценим банки по двум критериям:

- объем активов;
- доля депозитов физических лиц в обязательствах банка.

Для выделенной группы банков рассмотрим основные этапы оценки риска, реализованные в со-

ставе "Сигнальной подсистемы" СППР, в том числе формирование сигналов к пересмотру ДПФЛ, построение моделей формирования совокупностей сигналов, оценку риска этапа.

Дано: X — множество факторов, оказывающих влияние на объект управления — ДПФЛ; $X_j \in X$ — j -й фактор, $j = \overline{1, m}$. При $j = \overline{1, m_1}$ — подмножество факторов внешней среды — $X_{\text{внешн}}$, при $j = \overline{m_1 + 1, m}$ — подмножество внутренних факторов — $X_{\text{внутр}}$;

x_j^t — значение j -го фактора в момент времени t ;

$y^t = \{0, 1\}$ — выходная переменная, характеризующая наличие изменений в ДПФЛ в момент времени t .

Найти: оценку уровня риска несвоевременного пересмотра депозитного портфеля, который рассчитывается по формуле:

$$R_1 = 1 - p_1 p_2, \quad (2)$$

где p_1 — вероятность принятия решения об изменении структуры ДПФЛ; p_2 — условная вероятность своевременного пересмотра ДПФЛ.

Основные этапы оценки риска

Шаг 1. Разбиение депозитного портфеля на три подпортфеля: краткосрочный, среднесрочный, долгосрочный по длительности срока вклада:

- краткосрочные — от одного до трех месяцев;
- среднесрочные — от трех месяцев до года;
- долгосрочные — свыше года.

Вероятности p_1 и p_2 рассчитываются отдельно для каждого из подпортфелей.

Шаг 2. Построение системы индикаторов и формирование сигналов.

Для каждого из факторов формируется индикатор по одному из перечисленных ниже правил, а также набор условий для генерации бинарных сигналов к пересмотру ДПФЛ.

Введем следующие обозначения:

I — множество индикаторов, $I_j \in I$ — j -й индикатор, $j = \overline{1, m}$, где m — число индикаторов, совпадающее с числом исходных факторов;

i_j^t — значение j -го индикатора в момент времени t ;

P_j — j -е пороговое значение, соответствующее индикатору I_j .

Индикаторы формируются по одному из следующих правил:

- как разность значений функции от значений фактора в момент времени t и базовый момент времени t_0 — абсолютный прирост:

$$i_j^t = F(x_j^t) - F(x_j^{t_0}), \quad (3)$$

где $F(x_{t,j})$ — значение функции j -го фактора в момент времени t ;

Формирование индикаторов и сигналов

Обозначение	Фактор	Индикатор	Сигнал
<i>Внешние факторы</i>			
X_1	Ставка рефинансирования r_{ref}	$I_1^t = r_{ref}^t - r_{ref}^{t-1}$	$I_1 > P_1$
X_2	Средняя ставка по 10 банкам r_{10}	$I_2^t = r_{10}^t - r_{10}^{t-1}$	$I_2 > P_2$
X_3	Уровень реальных доходов населения r_{inc}	$X_3 = \frac{r_{inc}^t - r_{inc}^{t-1}}{4}$ $I_3^t = X_3^t - X_3^{t-1}$	$I_3 > P_3$
X_4	Сдвиг коридора процентных ставок $[R_{min}^t, R_{max}^t]$	$I_{41}^t = R_{min}^t - R_{min}^{t-1} $ $I_{42}^t = R_{max}^t - R_{max}^{t-1} $	$I_{41} > P_{41}$ $I_{42} > P_{42}$
X_5	Ставка по межбанковским кредитам r_{MP}	$I_5^t = r_{MP}^t - r_{MP}^{t-1}$	$I_5 > P_5$
X_6	Индекс РТС r_{RTS}	$I_6^t = r_{RTS}^t - r_{RTS}^{t-1}$	$I_6 > P_6$
<i>Внутренние факторы</i>			
X_7	Объем привлеченных средств V_{liab}	$I_7^t = \frac{V_{liab}^t - V_{liab}^{t-1}}{V_{liab}^{t-1}}$	$I_7 < P_7$
X_8	Доля депозитов в привлеченных средствах V_{DP}	$X_8 = V_{DP} / V_{liab}$ $I_8^t = X_8^t - X_8^{t-1}$	$I_8 < P_8$
X_9	Средняя ставка по депозитам $r_{l.avg}$	$X_9 = r_{l.avg} - r_{MP} $ $I_9^t = X_9^t - X_9^{t-1}$	$I_9 > P_9$
X_{10}	Средняя ставка размещения $r_{a.avg}$	$X_{10} = r_{a.avg} - r_{l.avg} $ $I_{10}^t = X_{10}^t - X_{10}^{t-1}$	$I_{10} > P_{10}$
X_{11}	Чистый процентный доход NIC	$I_{11}^t = NIC^t - NIC^{t-1}$	$I_{11} < P_{11}$
X_{12}	"В рынке" σ	$\sigma^t = N_1^t / N^t$, N^t — число депозитных продуктов; N_1^t — число продуктов, со ставкой % выходящих за границы коридора процентных ставок $[R_{min}^t, R_{max}^t]$ $I_{12}^t = \sigma^t$	$I_{12} > P_{12}$
X_{13}	Нормативы ликвидности ЦБ H_2, H_3, H_4	$I_{13} = \begin{cases} H_2^t - H_2^{t-1} \\ H_3^t - H_3^{t-1} \\ H_4^t - H_4^{t-1} \end{cases}$	$I_{13} > P_{13}$
X_{14}	Дюрация D	$I_{14} = D^t - D^{t-1}$	$I_{14} > P_{14}$

- как *относительный прирост* — отношение абсолютного прироста функции от значения фактора к значению в базовый момент времени:

$$i_j^t = \frac{F(x_j^t) - F(x_j^{t_0})}{F(x_j^{t_0})}; \quad (4)$$

- если значения фактора лежат в интервале $[F(x_j^t)_{min}; F(x_j^t)_{max}]$, где $F(x_j^t)_{min}$ и $F(x_j^t)_{max}$ — нижняя и верхняя границы соответственно, то формируются два индикатора по аналогии с формулой (3) для каждой из границ в отдельности:

$$i_j^t = F(x_j^t)_{min} - F(x_j^{t_0})_{min};$$

$$i_j^t = F(x_j^t)_{max} - F(x_j^{t_0})_{max}. \quad (5)$$

Каждая из приведенных формул (3)—(5) может быть преобразована введением модуля значений.

В качестве функции может быть использовано само значение аргумента, различные виды скользящих средних и др.

В таблице представлен исходный набор внешних и внутренних факторов, а также соответствующие им индикаторы и условия для генерации сигналов, сформированные по приведенным правилам.

На рис. 2, а представлен справочник индикаторов, на рис. 2, б — графическая визуализация индикатора "Абсолютный прирост индекса РТС".

Для каждого из индикаторов устанавливается пороговое значение P_j , отклонение от которого является сигналом к изменению ДПФЛ.

Шаг 3. Анализ качества сигналов.

Для оценки качества сигналов, формируемых на предыдущем шаге, а также подбора наилучших пороговых значений используются элементы *частотной логики*.

Введем следующие обозначения:

$x_j^t = \{0; 1\}$ — наличие сигнала в момент времени t для j -го индикатора;

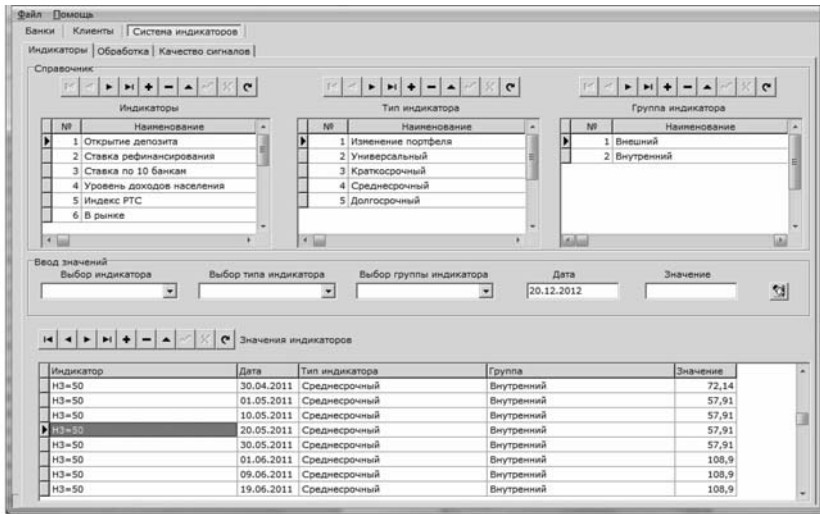
$Z(x_j^t) = \{0; 1\}$ — переменная принимает значение 1, если после возникновения сигнала было принято решение о пересмотре ДПФЛ, т. е. характеризует подтверждение сигнала внесением изменений в ДПФЛ;

$q_j = N_{Z(x_j^t)} / N$ — частота свойства $Z(x_j^t)$, где

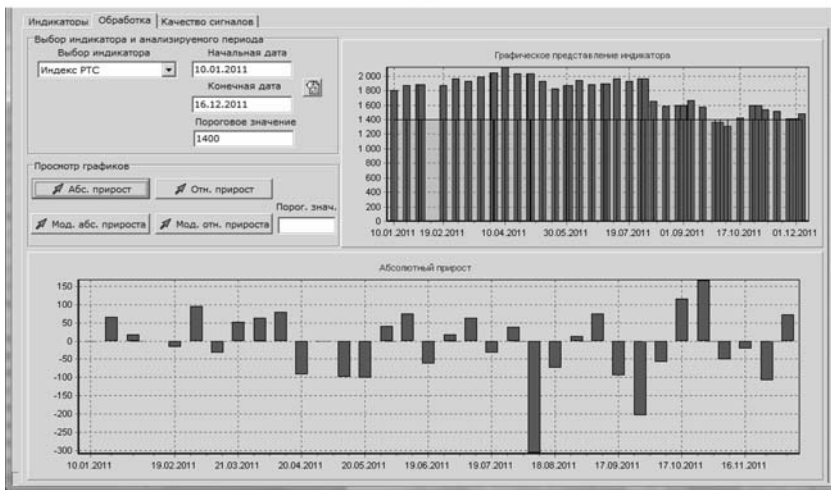
$N_{Z(x_j^t)}$ — число наблюдений, равных 1, N — общее число наблюдений [4].

Тогда наилучшему пороговому значению P_j будет соответствовать максимальное значение q_j .

На рис. 3 представлен пример изменения "порога" в подсистеме, что позволяет максимизировать значение частоты сигнала. Некоторые из сигналов характеризуются параметром "лаг", что объясняется



а)



б)

Рис. 2. Реализация сигнальной подсистемы в СПДР:

а — справочник индикаторов; б — графическое представление абсолютного прироста фактора индекса РТС

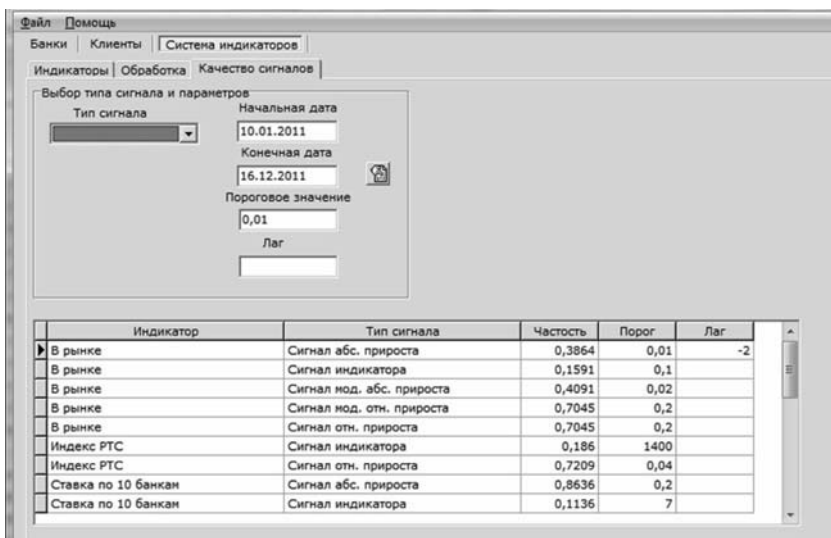


Рис. 3. Реализация модуля оценки качества сигналов в СПДР

задержкой реакции банка на произошедшие изменения фактора.

Шаг 4. Построение моделей формирования совокупности сигналов.

Для формирования системы индикаторов и выявления связей между сигналами предложены два метода: первый использует модели частотной логики, второй — инструмент логистической регрессии.

В первом методе предлагается объединить сигналы с помощью функции истинности, аргументами которой выступают частоты сигналов и их всевозможные комбинации.

В качестве функции от частот может быть выбрана как функция логической суммы переменных

$$f(q_1, q_2, \dots, q_j, \dots, q_m) = \sum_{j=1}^m q_j \quad (6)$$

так и функция, включающая попарное или множественное пересечение сигналов, в случае, если необходимо учесть возможность взаимообусловленного возникновения сигналов, например:

$$f(q_1, q_2, \dots, q_j, \dots, q_m) = \sum_{i=1}^m \sum_{j=1}^m q_i q_j, \quad i \neq j \quad (7)$$

где $f(q_1, q_2, \dots, q_j, \dots, q_m)$ — функция истинности.

Второй метод позволяет оценить вероятность принятия решения об изменении структуры ДПФЛ по имеющимся статистическим данным. Для этого в системе предлагается использование логистического регрессионного анализа [3]. При этом предполагается линейная зависимость индикаторов:

$$y^t = b_0 + b_1 i_1^t + \dots + b_m i_m^t + \varepsilon^t \quad (8)$$

где b_0, b_1, \dots, b_m — коэффициенты регрессии; ε^t — значение случайной ошибки в момент времени t .

Используя логит-преобразование, рассчитывается вероятность внесения изменений в ДПФЛ при заданных значениях индикаторов:

$$p_1 = \frac{1}{1 + e^{-y}} \quad (9)$$

Задавая пороговое значение для вероятности p_1 , получаем сигнал к переосмотру депозитного портфеля банка.

Шаг 5. Расчет вероятности своевременного пересмотра ДПФЛ.

Для оценки вероятности p_2 формируется выборка депозитов d_l , где $l = \overline{1, l'}$, l' — число депозитов в выборке, которая включает в себя удачные, *своевременно* открытые депозиты, т. е. такие депозиты, по которым реально были привлечены планируемые аналитиками банка суммы средств.

Каждому депозиту в данной выборке соответствует период времени Δt_l , прошедший от получения сигнала к пересмотру ДПФЛ до открытия депозита. На основе данной статистики и мнений экспертов строится эмпирическая функция распределения периодов, по которой определяется значение вероятности своевременного открытия депозита p_2 .

Шаг 6. Принятие решения о внесении изменений в депозитный портфель.

Учитывая рассчитанное значение p_1 и эмпирический закон распределения для оценки своевременности внесения изменений в ДПФЛ, ЛПР принимает решение об изменении ДПФЛ, обеспечивающее минимальное значение уровня риска R_1 .

Заключение

- ♦ В представленной статье предлагается авторский *риск-ориентированный* подход к управлению активами-пассивами банка, в частности, депозитами физических лиц. Среди рисков, возникающих в процессе управления, выделены как общие для банковской деятельности риски — процентный риск и риск ликвидности [1], так и риски, специфичные именно для данного вида привлекаемых средств. Так, например, риск ликвидности в данном случае проявляется в случайном характере изъятия вкладов клиентами банка.
- ♦ В "Сигнальной подсистеме" при формировании набора внешних и внутренних факторов анализируются, в том числе, нормативы ликвидности ЦБ и показатель дюрации, что позволяет учесть влияние рисков ликвидности и процентного риска на состояние ДПФЛ.
- ♦ В рамках предложенного авторами подхода сделана попытка формализовать частные виды рисков, характеризующие эффективность бизнес-процессов, связанных с управлением ДПФЛ. При анализе и исследовании динамики депозитных средств основное внимание уделялось использованию стохастических моделей.
- ♦ Результатом работы "Сигнальной подсистемы" является рекомендация пользователю к принятию решения об изменении структуры ДПФЛ, в частности, определения наилучшего момента внесения изменений.

Разработанный прототип СППР позволяет оценить введенный показатель — *совокупный риск*

управления ДПФЛ коммерческого банка, который учитывается в процессе принятия решений [6].

- ♦ Помимо задачи пересмотра ДПФЛ в системе реализуется решение задачи формирования оптимального портфеля депозитов физических лиц, связанной с выбором наилучшего депозитного продукта, который принесет банку в планируемый момент времени заданную доходность.

На заключительном этапе работы системы привлеченные в депозиты средства размещаются в активы банка для оценки потенциального дохода банка. В качестве математического аппарата используются модель стохастических ветвящихся процессов Гальтона—Ватсона [2], модель управляемых дискретных марковских процессов [7], стохастическая модель транспортной задачи [5].

- ♦ Потенциальными заказчиками системы являются банки регионального уровня с высокой долей средств физических лиц в структуре пассивов. Система предназначена для финансовых аналитиков банка, специалистов розничного управления банка, специалистов управления риском.

Список литературы

1. **Анализ** математических моделей Базель II. М.: ФИЗМАТЛИТ, 2010. 288 с.
2. **Ватугин В. А.** Лекционные курсы НОЦ. Математический институт им. В. А. Стеклова РАН (МИАН). М.: МИАН, 2008. Вып. 8: Ветвящиеся процессы и их применения. 108 с.
3. **Елисева И. И., Юзбашев М. М.** Общая теория статистики: учебник / Под ред. И. И. Елисеевой. М.: Финансы и статистика, 1995. 386 с.
4. **Зверев Г. Н.** Теоретическая информатика и ее основания. Т. 1. М.: Физматлит, 2007. 295 с.
5. **Лысак А. В., Голян В. В.** Методы линейного проектирования в транспортных задачах // Информационно-управляющая система на железнодорожном транспорте, 2009. № 1. С. 7—11.
6. **Николаева М. А., Зотова О. Ф., Шолохова Н. В.** Модели управления депозитным портфелем коммерческого банка // Управление риском. 2012. № 2.
7. **Ховард Р. А.** Динамическое программирование и Марковские процессы: Пер. с англ. В. В. Рыкова. М.: Советское радио, 1964.
8. **Информация** по АБС Банк XXI век // Материалы по ПО Инверсия. URL: <http://www.bankxxi.ru>.
9. **Официальный сайт** компании R-Style Softlab. URL: <http://www.softlab.ru>.
10. **Официальный сайт** компании DIASOFT. URL: <http://www.diasoft.ru>.
11. **Официальный сайт** компании Новая Афина. URL: <http://www.athena.ru>.
12. **Официальный сайт** компании BlackBay. URL: <http://www.blackbay.ru>.
13. **Применение** аналитических информационных систем BUSINESS OBJECTS в банковских приложениях. URL: http://www.olap.ru/desc/bobjects/bo_for_asp.
14. **Ершова Э.** Банки перестали выбирать между российскими и зарубежными АБС. URL: <http://www.cnews.ru/reviews/free/banks2009/articles/abs.shtml>.
15. **Шестопалова Н.** Российский рынок систем управления рисками в банках. PC Week/RE № 47 (605). URL: <http://www.pcweek.ru/idea/article/detail.php?ID = 105112>

ГЕОИНФОРМАЦИОННЫЕ СИСТЕМЫ GEOINFORMATION SYSTEMS

УДК 004.4

В. П. Потапов, д-р техн. наук, зам. директора филиала,
С. Е. Попов, канд. техн. наук, ст. науч. сотр., e-mail: ogidog@mail.ru,
Р. Ю. Замараев, канд. техн. наук, ст. науч. сотр.,
Институт вычислительных технологий СО РАН, г. Кемерово

ГИС-сервис для расчета интенсивности землетрясений на базе технологии Google App Engine¹

Разработан веб-сервис "Сотрясаемость" (<http://seismoshake.appspot.com>) для построения моделей макросейсмического поля при прогнозировании последствий землетрясений. Веб-сервис позволяет строить области проявлений последствий землетрясений и создавать отчеты по выбранному району. В веб-сервисе использован интегрированный подход для доступа к информации о населенных пунктах и их географической привязки на основе облачных геовычислений. Реализована программная модель интеграции служб GEONAMES.ORG с облачной архитектурой Google App Engine для расчета интенсивности землетрясений по шкале MSK-64.

Ключевые слова: сотрясаемость, веб-сервис, геовычисления, облачные технологии

V. P. Potapov, S. E. Popov, R. Yu. Zamaraev

Web Service for Calculating the Macroseismic Field Based on Google App Engine

Web-based service "Seismoshake" (<http://seismoshake.appspot.com>) are developed for modeling the macroseismic field of an earthquake. Web service allows building field displays and reporting on selected areas. Integrated approaches to access to information about locations based on cloud geocomputing are used. Implemented programming model through integrating services with cloud architecture GEONAMES.ORG Google App Engine to calculate earthquakes' intensities in MSK-64 scale.

Keywords: seismic shaking, web-service, geocomputing, cloud technologies

Введение

Проблема обеспечения сейсмической безопасности населения была и остается весьма актуальной. В ней можно выделить две основные задачи: прогнозирование высокоэнергетических сейсмических событий и оценка сейсмических рисков. Вторая задача включает в себя сейсмическое районирование территорий и прогнозирование объемов потерь и спасательных работ.

Подготовка и реализация спасательных работ находятся в ведении Министерства по чрезвычайным ситуациям и региональных служб ГО и ЧС. Практическая сторона их деятельности подкреплена нормативными документами, предполагающими в ряде задач значительный объем вычислений, в том числе на картах. К таким нормативам отно-

сится и методика прогнозирования последствий землетрясений [1], суть которой заключается в построении зон воздействия по шкале MSK-64. Открытость методики позволяет всем заинтересованным организациям реализовать доступный и оперативный расчет интенсивности землетрясений, построение зон воздействия и итоговых отчетов для заданного региона.

Развитие облачных технологий [2] позволяет создавать веб-сервисы, представляющие собой клиент-серверную технологию, в которой клиент использует ресурсы группы серверов в сети Интернет. При этом наличие нескольких источников используемых ресурсов, с одной стороны, позволяет повышать доступность системы клиент—сервер за счет возможности масштабирования, а с другой стороны, снижает риск неработоспособности виртуальных серверов, что является неоспоримым преимуществом с позиций оперативности расчетов.

¹ Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 13-05-98091 p_сибирь_a.

Поэтому создание специализированного веб-сервиса для расчета интенсивности землетрясений на базе облачных технологий с учетом доступности и оперативности в вопросах прогнозирования последствий землетрясений является актуальной проблемой в рамках интеграции геоинформационных систем и облачных технологий.

Постановка задачи

Была поставлена задача разработать ГИС-сервис для расчета интенсивности землетрясений с возможностью построения зон воздействия в виде круговых областей согласно бальной шкале MSK-64 на электронной карте предполагаемого географического района возникновения сейсмического события. При этом функционал сервиса должен удовлетворять следующим требованиям:

- работа с базой данных населенных пунктов (редактирование, добавление/удаление);
- поиск в географических базах данных имен объектов (топонимики);
- формирование отчетов в pdf-формате;
- эргономика работы с GUI, в частности, ввод расчетных параметров (магнитуда, глубина и географические координаты сейсмического события, выбор района расчетов), заполнение информационных сообщений для формирования отчета (энергетический класс, ближайший географический объект, тип события, время события).

Методика расчета интенсивности землетрясения

В работе была использована методика прогнозирования последствий землетрясений (МППЗ) [1]. Методика прошла экспертизу Межведомственного координационного научного совета (МВКНС) по проблемам гражданской обороны и чрезвычайных ситуаций (протокол от 19 сентября 2000 г. № 3) и может быть использована для разработки программных средств сторонними организациями.

Частью этой методики является расчет интенсивности землетрясения I , которая определяется как мера воздействия колебания грунта на внешнюю среду, оцениваемая по двенадцатибальной шкале. Соотношение между магнитудой M и интенсивностью I определяется эмпирической зависимостью, которая в литературе известна как формула макросейсмического поля

$$I = bM - \sqrt{\Delta^2 + h^2} + c, \quad (1)$$

где Δ — эпицентральное расстояние до географического объекта, км; h — глубина очага, км; M — магнитуда землетрясения, b , v , c — коэффициенты уравнения макросейсмического поля, которые приведены в таблице.

Преобразуя выражение (1) относительно Δ , можно получить выражение эпицентрального расстояния

Коэффициенты уравнения макросейсмического поля

Регион	b	v	c
Карпаты	1,5	3,5	3,0
Крым и Южная Кубань	1,5	3,5	3,0
Северный Кавказ	1,6	3,1	2,1
Дагестан	1,5	3,6	3,1
Закавказье	1,4	3,5	4,2
Азербайджан	1,4	3,5	3,5
Западная Туркмения	1,5	3,8	3,5
Средняя Азия и Казахстан	1,5	3,5	3,0
Алтай и Саяны	1,5	3,5	3,0
Прибайкалье	1,5	4,0	4,0
Якутия и Северо-Восток	1,5	3,5	3,0
Приморье и Приамурье	1,5	3,5	3,0
Сахалин	1,6	4,3	3,3
Курилы	1,5	4,5	4,5
Чукотка	1,5	3,5	3,0
Балтийский щит	1,5	3,5	3,0
Европейская часть РФ, Урал и Западная Сибирь	1,5	3,5	3,0

через параметры магнитуды M , интенсивности I и глубины сейсмического события h :

$$\Delta = \sqrt{10^{\frac{2(bM+c-I)}{v}} - h^2}, \quad (2)$$

где $10^{\frac{2(bM+c-I)}{v}} - h^2 > 0$; $0 \leq h \leq 10$ — требование МППЗ. Для промышленных взрывов предполагается $h = 0$ км; $1 \leq M \leq 8$ — требование МППЗ.

Выражение (2) позволяет рассчитывать максимальное значение параметра Δ для показателя интенсивности землетрясения I , изменяющегося в целочисленном интервале от 1 до 12 включительно, при заданных значениях магнитуды M и глубине h .

Для расчета параметра Δ в формуле (1) до заданного географическими координатами населенного пункта используется сферическая теорема косинусов в модификации для точек-антиподов:

$$\Delta = \arctan \sqrt{\frac{[\cos \varphi_2 \sin \Delta \lambda]^2 + [\cos \varphi_1 \sin \varphi_2 - \sin \varphi_1 \cos \varphi_2 \cos \Delta \lambda]^2}{\sin \varphi_1 \sin \varphi_2 + \cos \varphi_1 \cos \varphi_2 \cos \Delta \lambda}}, \quad (3)$$

где φ_1 , φ_2 , λ_1 , λ_2 — широта и долгота двух точек в радианах — центр сейсмического события и координаты населенного пункта, соответственно; $\Delta \lambda$ — разница координат по долготе.

Программная реализация ГИС-сервиса

В качестве облачной платформы для разработки и размещения ГИС-сервиса была выбрана технология Google App Engine. Существует большое количество ресурсов [3, 4], приводящих сравнение достоинств и недостатков облачных технологий. Можно утверждать, что по степени реализации и проработанности сервисов, API, языков програм-

"Сейсматика". GWT позволяет быстро писать и отлаживать AJAX-приложения на языке Java, используя инструментарий отладки Java. Компилятор GWT переводит код Java-приложения в соответствующий браузеру JavaScript и HTML, что избавляет разработчика от рутинной работы верстки и дизайна веб-элементов. Это позволило эргономично и оперативно дорабатывать и отлаживать ГИС-сервис "Сейсматика".

Пакет GMap позволил визуализировать результаты расчетов непосредственно на карте географической местности без использования сторонних настраиваемых Map-серверов (например MapInfo). Последняя версия продукта Google Map API предоставляет богатые возможности для разработки визуально-картографического контента для свободно распространяемого программного обеспечения, работающего как облачный сервис, что в контексте поставленной задачи вполне достаточно, учитывая выдвигаемые требования доступности расчетов.

Технология GWT-RPC позволила организовать взаимодействие клиентской части с облачной архитектурой Google App Engine посредством асинхронных вызовов удаленных методов сервлетов, содержащих программную логику расчета математической модели (1) с передачей полей веб-формы, а также выполнения http-запросов к сервисам GeoNames без перезагрузки графического интерфейса пользователя. Данный подход снял проблему сохранения состояния переменных как на стороне сервера, так и на стороне клиента, что позволило инкапсулировать функционал и сделать модель поведения сервиса приближенной к стационарному приложению.

Классы пакета java.net совместно с технологией GWT-RPC применяли для взаимодействия серверных компонентов с сервисами.

Алгоритм работы сервиса

Алгоритм работы ГИС-сервиса можно разделить на два этапа.

1. Расчет сейсмической сотрясаемости (1) и построение проявлений последствий землетрясений в виде круговых областей согласно выражению (2) по заданным значениям параметров расчета: долгота/широта, магнитуда и глубина сейсмического события (рис. 3, см. вторую сторону обложки).

2. Формирование отчета согласно расчетам и дополнительной информации, введенной пользователем: дата/время сейсмического события, энергетический класс, тип события, ближайшие географические объекты (рис. 4, см. вторую сторону обложки).

Согласно выражению (2) рассчитывают максимальные значения параметра Δ для показателя интенсивности землетрясения I , изменяющегося в целочисленном интервале от 1 до 12 включительно, при заданных значениях магнитуды M и глубине h . Круговые области образуются между двумя соседними окружностями, соответствующими значениям $\Delta(I_j)$ и $\Delta(I_{j+1})$, где $j = 1...11$ (см. рис. 3).

Формирование отчета происходит в виде генерации HTML-кода, представляющего веб-контент в GWT-элементе HTMLPanel. К интерфейсу сервиса добавляется выпадающая панель, содержащая этот элемент с html-страницей отчета (см. рис. 4 на второй стороне обложки). Модель размещения контента HTMLPanel позволяет выводить его на программный PDF-принтер, например BullZip PDF Printer, тем самым формируя аналогичный документ.

Описание получаемых результатов работы алгоритма и сформированного отчета представлено в справочной системе ГИС-сервиса (см. рис. 4).

Заключение

ГИС-сервис "Сотрясаемость" является новым средством для расчета интенсивности землетрясений на базе облачных технологий Google App Engine, который интегрирует технологии Google Maps, инструментарий Google Web Toolkit и базу данных географических объектов GeoNames для обеспечения оперативности и доступности расчетов при прогнозировании последствий природных и техногенных сейсмических событий.

Сервис может быть использован в практической деятельности работниками РСЧС и гражданской обороны, научными сотрудниками организаций, в том числе входящих в Федеральную систему сейсмологических наблюдений, а также специалистами центров мониторинга и прогнозирования чрезвычайных ситуаций.

Список литературы

1. **Методика** прогнозирования последствий землетрясений // М.: Изд. ВНИИ ГО ЧС, 2000. 27 с.
2. **Облачные** сервисы. Взгляд из России / Под ред. Е. Гребнева. М.: CNews, 2011. 282 с.
3. **Google** App Engine (GAE) против Amazon Web Services (AWS) [Электронный ресурс] // HABRAHABR.RU: Новостной сайт, коллективные блоги. URL: http://habrahabr.ru/blogs/cloud_computing/117146/ (дата обращения: 03.02.2012).
4. **Google** App Engine (GAE) vs Amazon Elastic Computing (EC2) vs Microsoft Azure [Электронный ресурс] // Danny Tuppeny/DanTup. URL: <http://blog.dantup.com/2010/10/google-app-engine-gae-vs-amazon-elastic-computing-ec2-vs-microsoft-azure> (дата обращения: 03.02.2012).
5. **The GeoNames** geographical database [Электронный ресурс] // GeoNames. URL: <http://ws.geonames.org/search> (дата обращения 21.01. 2013).
6. **The GeoNames** geographical database [Электронный ресурс] // GeoNames. URL: <http://ws.geonames.org/findNearby> (дата обращения 21.01. 2013).

Ф. Г. Агаев¹, д-р техн. наук, проф., директор,
Н. Г. Джавадов², д-р техн. наук, проф., ген. директор,
Х. Г. Асадов³, д-р техн. наук, нач. отд., **Р. В. Халафов**⁴, аспирант,
e-mail: asadzade@rambler.ru

¹ Институт космических исследований природных ресурсов

² ПО Промавтоматика, г. Баку

³ НИИ Аэрокосмической информатики

⁴ Национальное аэрокосмическое агентство, г. Баку, Азербайджан

Новый способ оптимизации дистанционного зондирования неоднородных участков природных и техногенных сред

Предложено обобщение и формализация задачи оптимальной оценки и вычета неоднородных участков различных составных объектов природных и техногенных сред методом дистанционного зондирования на базе нормализованных дифференциальных индексов.

Предложен способ оптимизации особого класса спектральных систем дистанционного зондирования, предназначенных для решения множества прикладных задач, формализуемых по типу задачи оптимальной оценки и вычета неоднородных участков составных объектов.

Ключевые слова: дистанционное зондирование, оптимизация, спектральные индексы, неоднородные участки, информационные системы

F. G. Agayev, N. H. Javadov, H. H. Asadov, R. V. Khalafov

New Technique of Optimization for Remote Sensing of Non-Homogenous Zones of Natural and Technogenic Medium

Generalization and formalization of the task of optimal assessment and subtraction of non-homogenous zones of various composed objects of natural and technogenic media using remote sensing at the basis of normalized differential indices is suggested.

The technique of optimization of special class of spectral system of remote sensing designated for solution of many applied tasks, formalized as task of optimal assessment and subtraction of many-homogenous zones of composed objects is suggested.

Keywords: remote sensing, optimization, spectral indices, non-homogenous areas, information systems

Хорошо известно, что способы оптимизации информационных систем могут быть как общими для всех классов систем, так и ориентированными для отдельных классов всего множества информационных систем. Если системы дистанционного зондирования рассматривать в качестве отдельного класса информационных систем, то можно ставить вопрос о разработке специфических способов оптимизации систем данного класса и отдельных его подклассов, оперирующих важнейшими атрибутами и характеристиками процедур дистанционного зондирования. В настоящей статье мы изложим предлагаемый новый способ оптимизации отдельного подкласса систем дистанционного зондирования, предназначенных для локализации неоднородных участков экологических объектов. Следует отметить, что проблема выделения и обработки сигналов, получаемых с неоднородных участков, может быть рассмотрена и решена с разных позиций. Например, с точки зрения формальной циф-

ровой обработки информационных сигналов здесь вполне применимы различные методы линейной и нелинейной фильтрации, сглаживания, подчеркивания контуров и т. д. [1]. При геоинформационном подходе, например, к неоднородностям электрических свойств сейсмоактивного района могут быть применены методы активного и пассивного электромагнитного мониторинга [2]. Если речь идет о биоразнообразии, то здесь широко применяют мультифрактальный анализ, исследование микроландшафтов с помощью космических снимков [3]. При геоинформационном подходе к проблеме анализа неоднородных сцен часто используют объектно-ориентированные методы моделирования пространственных данных [4]. В технике дистанционного зондирования основным спектральным индексом был и остается нормализованный дифференциальный вегетационный индекс, широко применяемый для анализа неоднородных сцен [5, 6]. При этом следует подчеркнуть, что методический прием, ис-

пользованный при формировании этого индекса, был широко применен к различным исследуемым объектам реального мира и фактически было сформировано целое множество нормализованных дифференциальных индексов. Предлагаемый нами способ также оперирует нормализованными дифференциальными индексами. Чтобы подтвердить важность и роль нормализованных дифференциальных индексов для подкласса спектральных систем дистанционного зондирования приведем краткий обзор широко применяемых на практике спектральных нормализованных дифференциальных индексов.

1. *Нормализованный дифференциальный вегетационный индекс (NDVI)* впервые предложен Rouse (1973) [1] и концептуально представленный Krieger (1969) [2], он определяется как

$$NDVI = \frac{R_{NIR} - R_{RED}}{R_{NIR} + R_{RED}},$$

где R_{NIR} и R_{RED} — соответственно, интенсивности отраженных с растительности излучений на длинах волн 857 и 645 нм.

2. *Нормализованный дифференциальный водный индекс (NDWI)* [3], определяется как

$$NDWI = \frac{R_{red} - R_{SWIR}}{R_{red} + R_{SWIR}},$$

где R_{SWIR} — радиация, излучаемая в коротковолновой части (2,08...2,35 мкм) ИК диапазона; R_{red} — излучение в диапазоне 0,63...0,69 мкм.

3. *Нормализованный дифференциальный влажностный индекс (NDMI)* [5] определяется как

$$NDMI = \frac{R_{NIR} - R_{mid IR}}{R_{NIR} + R_{mid IR}},$$

где $R_{mid IR}$ — интенсивность отраженного излучения в среднем ИК диапазоне (5-й канал ТМ; 1,5...1,89 мкм).

4. *Вегетационный индекс, устойчивый к влиянию атмосферы (ARVI)*,

$$ARVI = \frac{R_{NIR} - R_b}{R_{NIR} + R_b},$$

где $R_b = R_{RED} - a(R_{RED} - R_{BLUE})$, обычно $a = 1$, при малом покрытии растительностью и неизвестном типе атмосферы $a = 0,5$ [4].

5. *Нормализованный дифференциальный песочный индекс (NDSI)* [3]

$$NDSI = \frac{R_{red} - R_{SWIR2}}{R_{red} + R_{SWIR2}}.$$

Согласно работе [3],

$$-1 \leq NDSI \leq 1.$$

6. *Нормализованный дифференциальный почвенный индекс (NDSI₂)*, определяемый как [5]

$$NDSI_2 = \frac{R_{SWIR} - R_{NIR}}{R_{SWIR} + R_{NIR}}.$$

7. *Нормализованный дифференциальный облачный индекс (NDCI)*

$$NDCI = \frac{R(\lambda_2) - R(\lambda_1)}{R(\lambda_2) + R(\lambda_1)},$$

где $R(\lambda_2)$ — радиация Солнца, последовательно переотраженная от растительности и облачности, $\lambda_2 = 0,86$ мкм; $R(\lambda_1)$ — то же, при $\lambda_1 = 0,65$ мкм [6].

8. *Нормализованный дифференциальный индекс снега (NDSI₁)*

$$NDSI_1 = \frac{R_{red} - R_{SWIR}}{R_{red} + R_{SWIR}}.$$

Для снега $NDSI_1 > 0,4$ [7].

9. *GNDVI* — Зеленый (Green) *NDVI* [8]:

$$NDVI_g = \frac{R_{NIR} - R_{green}}{R_{NIR} + R_{green}}.$$

Очевидно, что используемые на практике нормализованные дифференциальные индексы не ограничиваются приведенным списком, он может быть значительно дополнен.

Рассмотрим типичную задачу дистанционного зондирования (рис. 1) различных объектов, суть которой заключается в локализации и оптимальном вычете неоднородных участков из сформированной общей оценки. Предлагаемый способ оптимизации сформулируем следующим образом. Исходно задается объект дистанционного зондирования с площадью S_1 , характеризуемый спектральным нормализованным дифференциальным индексом NDI_1 .

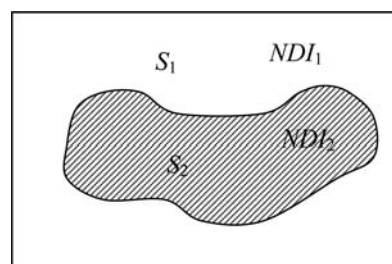


Рис. 1. Иллюстрация составного объекта дистанционного зондирования с неоднородным участком в задаче исследования

В пределах объекта S_1 содержится объект другого рода, имеющий площадь S_2 и характеризующийся индексом NDI_2 .

При условиях $S_1 = \text{const}$; $S_2 = \text{const}$

$$(S_1 \cdot NDI_1)^2 + (S_2 \cdot NDI_2)^2 = \text{const}. \quad (1)$$

Следует найти такое взаимное соответствие между элементами NDI_1 и NDI_2 , при котором квадрат разности

$$(S_1 \cdot NDI_1 - S_2 \cdot NDI_2)^2 \rightarrow \text{max}. \quad (2)$$

Поясним, что далее понимается под элементами NDI . NDI_1 определим как

$$NDI_1 = \frac{I_{\lambda_1}(S_1) - I_{\lambda_2}(S_1)}{I_{\lambda_1}(S_1) + I_{\lambda_2}(S_1)}, \quad (3)$$

где $I_{\lambda_1}(S_1)$ — сигнал дистанционного зондирования, полученный с объекта площадью S_1 на длине волны λ_1 ; $I_{\lambda_2}(S_1)$ — сигнал дистанционного зондирования, полученный с объекта площадью S_1 на длине волны λ_2 .

В результате зондирования этого объекта получаются, соответственно, сигналы

$$I_{\lambda_1}(S_2) = k_1 I_{\lambda_1}(S_1), \quad (4)$$

$$I_{\lambda_2}(S_2) = k_2 I_{\lambda_2}(S_1), \quad (5)$$

где k_1 и k_2 — коэффициенты спектрального различия на λ_1 и λ_2 соответственно.

Следовательно, объект с площадью S_2 характеризуется индексом NDI_2 :

$$NDI_2 = \frac{k_1 I_{\lambda_1}(S_1) - k_2 I_{\lambda_2}(S_1)}{k_1 I_{\lambda_1}(S_1) + k_2 I_{\lambda_2}(S_1)}. \quad (6)$$

С учетом соотношений (1)–(6) предлагаемый способ оптимизации применительно к изложенной выше типичной задаче базируется на следующем утверждении.

Утверждение. Сформулированная выше типичная задача локализации и оптимального вычета неоднородного участка имеет решение, т. е. условие (2) выполняется при выполнении следующего равенства:

$$I_{\lambda_2} = I_{\lambda_1} \cdot \sqrt{k_1/k_2}. \quad (7)$$

Доказательство. Из условия (2) легко можно получить следующее эквивалентное условие:

$$2S_1 \cdot S_2 \cdot NDI_1 \cdot NDI_2 \rightarrow \text{min}. \quad (8)$$

С учетом $S_1 = \text{const}$; $S_2 = \text{const}$ условие (8) принимает следующий вид:

$$NDI_1 \cdot NDI_2 \rightarrow \text{min}. \quad (9)$$

С учетом выражений (3) и (6) имеем

$$F = NDI_1 \cdot NDI_2 = \frac{I_{\lambda_1}^2 k_1 - I_{\lambda_1} I_{\lambda_2} (k_1 + k_2) + k_2 I_{\lambda_2}^2}{I_{\lambda_1}^2 k_1 + I_{\lambda_1} I_{\lambda_2} (k_1 + k_2) + k_2 I_{\lambda_2}^2}. \quad (10)$$

Заметим, что в выражении (10) и далее для упрощения записи аргументы S_1 и S_2 не указываются.

Введем в рассмотрение искомую функцию

$$I_{\lambda_2} = f(I_{\lambda_1}). \quad (11)$$

С учетом (11) выражение (10) может быть рассмотрено в качестве функционала, подлежащего минимизации путем соответствующего подбора функции (11).

Исследуя функционал (10) на минимум от $I_{\lambda_2} = f(I_{\lambda_1})$, нетрудно получить следующее выражение:

$$f^2(I_{\lambda_1}) = \frac{I_{\lambda_1}^2 k_1}{k_2}. \quad (12)$$

Из выражения (12) получаем

$$f(I_{\lambda_1}) = I_{\lambda_1} \sqrt{k_1/k_2}. \quad (13)$$

Проверка найденного решения выполнена вычислительным путем. С учетом выражений (3) и (6) произведение $NDI_1 \cdot NDI_2$ при $k_1 = 4$; $k_2 = 1$; $I_{\lambda_1} = 1$; $f(I_{\lambda_1}) = 2I_{\lambda_1} = x$ может быть выражено в виде

$$NDI_1 \cdot NDI_2 = \frac{(1-x)(4-x)}{(1+x)(4+x)}. \quad (14)$$

Вычисленные значения $NDI_1 \cdot NDI_2$ при разных значениях x приведены в таблице.

x	0,5	1	2	2,5	3	5
$NDI_1 \cdot NDI_2$	0,27	0	-0,11	-0,088	-0,077	0,077

График зависимости $NDI_1 \cdot NDI_2 = f(x)$ приведен на рис. 2.

Как видно из графика, приведенного на рис. 2, значение произведения $NDI_1 \cdot NDI_2$ имеет минимум при $x = 2$, что и следовало ожидать из условия (14) при заданных значениях остальных параметров.

Таким образом, согласно предложенному способу оптимизации решение различных задач дис-

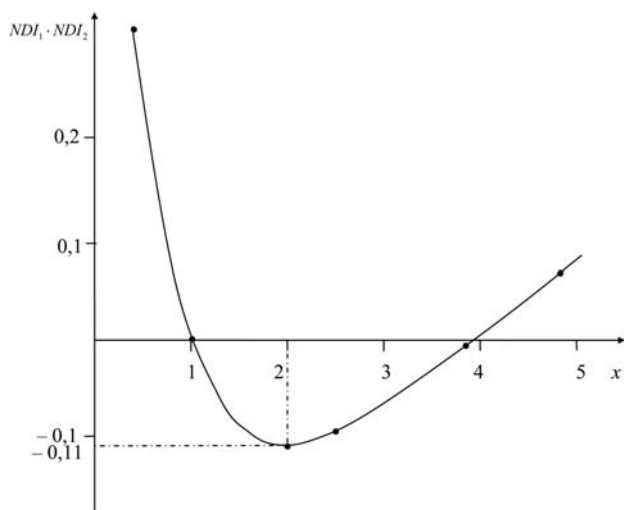


Рис. 2. График зависимости $NDI_1 \cdot NDI_2 = f(x)$

танционного зондирования неоднородных природных и техногенных объектов с применением нормализованных дифференциальных индексов может быть осуществлено с учетом соотношений (1)–(7), доказывающих существование оптимального режима локализации и вычета неоднородных участков рассматриваемых объектов.

Изложенный способ может быть применен при решении следующих практических задач:

- в лесоведении: максимально достоверная оценка компонентов двухсоставных вегетационных участков (т. е. лесов, полей);
- в почвоведении: максимально достоверная оценка зон с различными типами земляного покрова;
- в сельском хозяйстве: максимально достоверная оценка полей, где выращены различные типы урожая;
- в экологии: максимально достоверная оценка территорий, где расположены полигоны отходов, мусорные скопления и т. д.

Очевидно, что область применения предлагаемого способа оптимизации спектрального дистанционного зондирования не ограничивается указанными выше сферами и может быть значительно расширена.

В заключение сформулируем основные выводы и положения проведенного исследования:

1. Предложено обобщение и формализация задачи оптимальной оценки и вычета неоднородных участков различных составных объектов природных и техногенных сред методом дистанционного

зондирования на базе нормализованных дифференциальных индексов.

2. Предложен способ оптимизации особого класса спектральных систем дистанционного зондирования, предназначенных для решения множества прикладных задач, формализуемых по типу задачи оптимальной оценки и вычета неоднородных участков составных объектов.

Список литературы

1. **Волосюк В. К., Волошук Р. П., Нежальская К. Н., Павликов В. В., Усс М. Л.** Цифровые методы обработки информационных процессов. Харьков: ХАИ, 2011. URL: <http://rsd.khai.edu/posobiya.php>
2. **Дашевский Ю. А., Мартынов А. А.** Обратные задачи электрических зондирований в сейсмоактивных районах: учеб.-метод. пособие. Новосибирск, 2002.
3. **Артюшенко М. В., Зуб Л. Н., Подгорецкая Л. В., Федоровский А. Д.** Мультифрактальный анализ биоразнообразия и ценотической структуры сообщества растений по данным дистанционного зондирования // *Доповіді Національної Академії Наук України*. 2011. № 9. С. 132–136.
4. **Основы геоинформатики** / Под ред. В. С. Тикунова. М.: Академия, 2004. 352 с.
5. **Барталев С. С.** Разработка методики региональной экологической оценки состояния лесов по данным спутниковых наблюдений: дис. ... канд. техн. наук. Москва, 2006.
6. **Корниенко С. Г.** Изучение неоднородностей деятельного слоя криолитозоны с использованием данных теплового дистанционного зондирования. URL: http://d33.infospace.ru/d33_conf/2008_pdf/2/40.pdf
7. **Rouse J. W., Hass R. H., Schell J. A., Deering D. W.** Monitoring vegetation systems in the great plains with ERTS // *Third ERTS Symposium, NASA SP-351*. 1973. Vol. 1. P. 309–317.
8. **Kriegler F. J., Malila W. A., Nalepka R. F., Richardson W.** Preprocessing transformations and their effects on multispectral // *Proc. of the Sixth International Symposium on Remote Sensing of Environment*. An Arbor, MI: University of Michigan, 1969. P. 97–131.
9. **Майорова В. И., Банников А. М., Гришко Д. А., Жаренов И. С., Леонов В. В., Топорков А. Г., Харлан А. А.** Контроль состояния сельскохозяйственных полей на основе прогнозирования динамики индекса // *Наука и образование*. Эл. N ФС77-48211. URL: <http://technomag.edu.ru/doc/577991.html>
10. **Kauffman Y. J., Tanre D., Нагорный В. Д.** Определение и мониторинг продуктивного потенциала сельскохозяйственных земель // *Науки о земле*. 2011. № 01. С. 68–75. URL: <http://geo-science.ru>
11. **Potither S., Ishii R., Suzuki R.** The potential of Normalized Difference Soil Index (NDSI) for soil water content estimation in Mongolia // *Proc. of 30th Asian Conference on Remote Sensing 2009 (ACRS 2009)* Beijing, China, 18–23 October. 2009. Vol. 1 of 3. P. 356–359.
12. **Marshak A., Knyazikhin Y., Davis A., Wiscombe W., Pilewskie P.** Cloud – vegetation interaction: use of Normalized Difference Cloud Index for estimation of cloud optical thickness // *Geophysical Research Letters*. 2000. June. Vol. 27. P. 1695–1698.
13. **Индексные изображения**. URL: <http://nrcgit.ru/aster/methods/arithmet.htm>
14. **Gitelson A. A., Kauffman Y. J., Merzlyak M. N.** Use of green channel in remote sensing of global vegetation from EOS-MODIS // *Remote Sensing of Environment*. 1996. No. 58. P. 289–298.

А. В. Черняев, д-р техн. наук, проф., **А. А. Павлов**, канд. техн. наук, инж.,
МАТИ — Российский государственный технологический университет имени К. Э. Циолковского,
e-mail: alex_v_ch@bk.ru

Географическая информационная система малого речного бассейна и особенности ее формирования в задачах моделирования нефтяных загрязнений

Рассмотрены вопросы построения геоинформационной математической модели для информационной системы малого речного бассейна для применения в моделировании и прогнозировании аварийных нефтяных загрязнений.

Ключевые слова: разлив нефти, математическое моделирование, береговая поверхность, геоинформационные системы, техногенная безопасность

A. V. Chernyaev, A. A. Pavlov

Geographic Information System of the Small River Basin and Features of its Generation in Oil Pollution Modeling Problems

The article deals with geoinformation mathematical model generation problems. The geoinformation mathematical model under consideration is for small river basin information system intended for use in accidental oil pollution modeling and forecasting.

Keywords: oil flood, mathematical modeling, a coastal surface, geoinformation systems, technogenic safety

Введение

Для компьютерной обработки и анализа информации по аварийным загрязнениям рек применяются географические информационные системы (ГИС) [1]. Эти технологии объединяют традиционные операции при работе с базами данных (запрос, статистический анализ и др.) с преимуществами полноценной визуализации и географического (пространственного) анализа, которые предоставляет карта. Особенно перспективным представляется интеграция ГИС с моделями аварийных загрязнений бассейнов малых рек [1, 2].

На сегодняшний день для моделирования геоэкологических процессов ГИС используют недостаточно, особенно для территорий бассейнов малых рек. Поэтому в данной работе исследованы вопросы создания ГИС малого речного бассейна.

Разработка картографической основы ГИС малого водотока

Для создания ГИС предлагается использовать систему настольной картографии MapInfo, позволяющую решать сложные задачи географического анализа, придавать графический вид статистическим данным, выдавать необходимую информацию по специальным запросам. Пространственно распределенная информация хранится в виде географических слоев (объекты обустройства, гидросеть, пункты оп-

робования снежного покрова и др.), которые можно совмещать в любой последовательности и подвергать обработке с использованием ряда аналитических процедур. Хранящаяся в базе данных информация может быть использована как для просмотра и отображения, так и для организации сложных тематических запросов, расчетов и моделирования различных ситуаций в целях выбора оптимального решения.

В связи с тем, что исследуемые загрязнения носят разномасштабный характер и могут распространяться вниз по течению на большие расстояния, необходимо многоуровневое представление данных на картах разных масштабов. Информационное наполнение ГИС состоит из нескольких уровней:

- 1-й уровень — обзорная карта всей территории бассейна исследуемой реки (масштабы 1:500 000, 1:1 000 000);
- 2-й уровень — карты притоков первого порядка (масштабы 1:100 000, 1:200 000);
- 3-й уровень — карты притоков второго и третьего порядков (масштабы 1:25 000, 1:10 000);
- 4-й уровень — крупномасштабные карты малых водотоков (масштабы 1:5 000, 1:2 000), масштаб которых зависит от размера ключевого участка.

Информационное содержание цифровой географической основы в ГИС организовано в виде множества слоев, функциональным назначением которых является объединение пространственных объектов (точнее набора данных, характеризующих их

в визуальной базе данных), имеющих какие-либо общие свойства. Такими свойствами являются:

— принадлежность к одному типу пространственных объектов (слой "растительность", слой "гидро-сеть", слой "железные дороги" и т. д.);

— представление на карте одинаковыми графическими примитивами (линиями, точками, полигонами).

Слой ГИС по значению для дальнейшего моделирования можно подразделить на два блока: основные и вспомогательные слои. Основные слои используют непосредственно при моделировании, вспомогательные дают общую картину исследуемой территории.

К основным слоям относятся:

— слои, описывающие объект исследования (малые водотоки);

— слои, на которых показаны потенциальные источники загрязнения нефтепродуктами;

— слои, на которых показаны реципиенты негативных воздействий.

Объектом моделирования являются малые реки. На мелко- и среднемасштабных картах они будут показаны линиями, на крупномасштабных — линиями и/или полигонами. В атрибутивной таблице указываются: название, тип (реки, притоки с указанием порядка, каналы, ручьи), характеристика (постоянные, пересыхающие), глубина, ширина (в случае линейного объекта), узловые точки русла водотока, оси русла водотока. На отдельных слоях показывают информацию о береговой поверхности, включающую: тип грунта береговой поверхности и размер уклона береговой поверхности. Также для моделирования необходимы данные о растительном покрове берегов, поэтому на отдельном слое показана береговая линия и дана ее характеристика (наличие растительности, тип растительности, видовой состав, густота, высота).

К объектам исследования мы относим также параметры загрязнителя. Эту информацию наносят на несколько отдельных слоев, а именно слой пространственного положения нефтяного загрязнения и слои, отображающие градации цвета, объемы испарившейся, диспергированной и эмульсифицированной нефти, а также участки донной поверхности, подверженные загрязнению при затоплении тяжелых фракций нефти.

К потенциальным источникам загрязнения относят, в первую очередь, промышленные объекты, которые в зависимости от масштаба показывают в виде точечных или полигональных объектов. В атрибутивной таблице указаны: название, подчинение, максимально возможный объем загрязнения, физические и химические характеристики потенциального загрязнителя.

Автомобильные и железные дороги показываются линейными объектами. В атрибутивной таблице приведены: категория, грузопоток нефтепродуктов и т. п.

Нефтепроводы и продуктопроводы показываются линейными объектами. В атрибутивной таблице должны быть указаны: название, подчинение, год сооружения или реконструкции, объем перекачиваемого продукта, физические и химические характеристики потенциального загрязнителя.

Реципиентами нефтяного загрязнения могут быть населенные пункты, водозаборы, ООПТ, водоохранные зоны и прибрежно-защитные полосы. Населенные пункты на мелко- и среднемасштабных картах будут показаны точечными объектами и полигонами, на крупномасштабных — полигонами. В атрибутивной таблице указывают: название, тип (города, поселки городского типа, сельские поселения и т. п.), административный статус, число жителей.

Водозаборы показываются точечными объектами с указанием объема водозабора, цели забора воды.

ООПТ показываются точечными объектами или полигонами в зависимости от масштаба карт и размера ООПТ. В атрибутивной таблице указывают: название, подчинение, категория, обитающие около воды виды, занесенные в Красную книгу и др.

Водоохранные зоны и прибрежно-защитные полосы на мелко- и среднемасштабных картах показаны линиями, на крупномасштабных — полигонами. В атрибутивной таблице указаны название и статус.

К вспомогательным слоям относят: рельеф, растительность, прочие коммуникации, прочие промышленные объекты, сельскохозяйственные объекты и некоторые другие.

Рельеф представлен в виде следующего набора слоев: основные и дополнительные горизонталы, бергштрихи; отметки высот; промоины, обрывы и т. д.

Проектируемая система является открытой к дополнениям и расширениям, что позволяет создавать новые слои и расширять базу данных за счет включения новых атрибутов.

Следует отметить важность своевременного обновления информации.

Представление исходных данных

Описание параметров модели водотока начинается с выбора файла карты или фотоплана и импортирования топографической основы в ГИС. После чего по середине исследуемого русла водотока наносят узловые точки. Ввиду того, что расстояние между узловыми точками различное, при их выборе необходимо выполнение требования кратности расстояния между узловыми точками протяженности линейного участка. Таким образом, на карте задается ось водотока с помощью указания узловых точек, относительно которых в дальнейшем проводят раз-

Атрибутивная таблица оси водотока

Поле таблицы	Значение
Номер узловой точки	
Широта узловой точки	
Долгота узловой точки	
Абсолютная высота узловой точки	

Таблица 2

Атрибутивная таблица линейных участков водотока

Поле таблицы	Значение
Номер линейного участка	
Номер узловой точки, к которой идет привязка	
Ширина линейного участка в межень	
Ширина линейного участка при максимальном уровне	
Ширина линейного участка при среднем уровне	
Тип профиля русла	
Глубина в межень	
Глубина при максимальном уровне	
Глубина при среднем уровне	
Ширина основания дна	
Коэффициент шероховатости дна	
Плотность воды	
Температура воздуха	
Температура воды	
Скорость ветра	
Кинематическая вязкость воды	

Таблица 3

Атрибутивная таблица элементарных площадок водотока

Поле таблицы	Значение
Номер элементарной площадки	
Номер линейного участка, к которому идет привязка	
Номер узловой точки, к которой идет привязка	
Ширина элементарной площадки в межень	
Широта центра элементарной площадки в межень	
Долгота центра элементарной площадки в межень	
Ширина элементарной площадки при максимальном уровне	
Широта центра элементарной площадки при максимальном уровне	
Долгота центра элементарной площадки при максимальном уровне	
Ширина элементарной площадки при среднем уровне	
Широта центра элементарной площадки при среднем уровне	
Долгота центра элементарной площадки при среднем уровне	
Тип профиля русла	
Глубина в межень	Перечисляемое
Глубина при максимальном уровне	
Глубина при среднем уровне	
Ширина основания дна	
Коэффициент шероховатости донной поверхности	Перечисляемое
Плотность воды	
Температура воздуха	
Температура воды	
Скорость ветра	
Кинематическая вязкость воды	

биение водотока на линейные участки и элементарные площадки (рис. 1, табл. 1).

Водоток, заданный с помощью указания узловых точек, аппроксимируется с помощью линейных отрезков. Представление водотока в виде набора линейных участков имеет вид, показанный на рис. 2.

Затем создается атрибутивная таблица, описывающая параметры водотока в виде линейных участков для отображения на крупномасштабной карте (табл. 2).

Для того чтобы иметь возможность рассматривать ламинарный режим течения жидкости, линейные участки разбивают на элементарные площадки. Представление водотока в виде набора элементарных площадок показано на рис. 3.

После разбиения линейных участков на элементарные площадки заполняется атрибутивная таблица (табл. 3).

Для внесения информации, характеризующей береговую поверхность, создается отдельный слой (рис. 4) и соответствующая таблица с атрибутивной информацией (табл. 4).

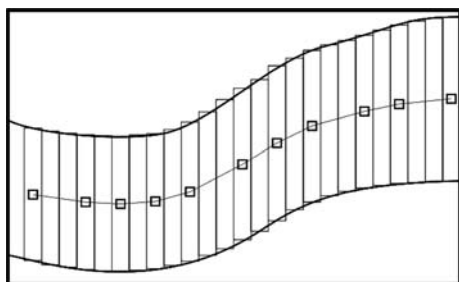


Рис. 2. Представление водотока в виде линейных участков

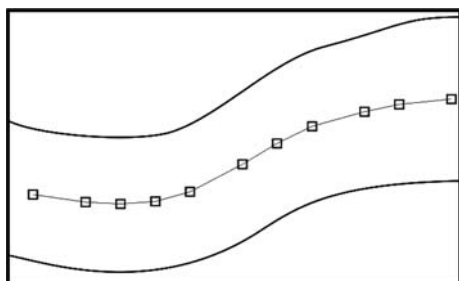


Рис. 1. Нанесение узловых точек и оси водотока

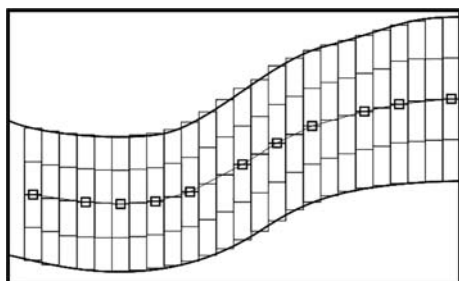


Рис. 3. Представление водотока в виде набора элементарных площадок

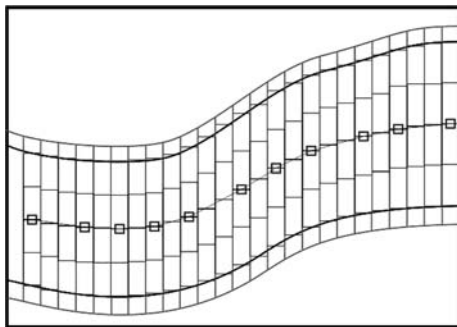


Рис. 4. Представление береговой поверхности

Таблица 4

Атрибутивная таблица береговой поверхности водотока

Поле таблицы	Значение
Номер участка берега	Перечисляемое Перечисляемое
Крутизна берега	
Номер линейного участка	
Номер узловой точки, к которой идет привязка	
Тип берегового грунта	
Тип растительности	
Площадь поверхности побега растения	
Число побегов на 1 м ²	

"Перечисляемое" значение означает то, что в случае описания типа профиля русла выбирается сечение прямоугольной, треугольной, эллиптической или трапецидальной формы (см. табл. 3). Коэффициент шероховатости донной поверхности: выбирается как одно из значений коэффициента шероховатости (см. табл. 3). В табл. 4 "перечисляемое": выбирается из списков тип грунта береговой поверхности, из другого списка выбирается тип растений, произрастающих на береговой поверхности.

Представление результатов моделирования

В зависимости от цели моделирования используют несколько разновидностей представления результатов. В случае, когда необходимо отображение результатов расчетов распространения и трансформации нефтяного загрязнения через определенные интервалы времени, создается множество атрибутивных таблиц, каждая из которых является слоем. Каждая таблица представляет результаты распространения и трансформации нефтяного разлива в зависимости от времени с начала аварийного разлива. Таким образом, появляется возможность, последовательно загружая слои, показывать динамику распространения нефтяного загрязнения. Данный механизм реализуют путем добавления из расчетной программы в таблицы атрибутивной информации линейных участков и элементарных площадок полей, приведенных в табл. 5.

Отображение результатов расчета объема нефтепродуктов, осевших на береговой поверхности, осуществляется путем добавления из расчетной программы в таблицу атрибутивных данных о береговой поверхности полей, приведенных в табл. 6.

Для отображения результатов моделирования на мелкомасштабных картах ГИС используется представление в виде закрашенных определенным образом линейных отрезков, на крупномасштабных картах — элементарных площадок. В случае, когда нефтяное загрязнение находится на указанном линейном участке или элементарной площадке, данный участок или площадка закрашивают цветом, в противном случае — оставляют незакрашенным. Значение загрязнения в случае работы с мелкомасштабными картами отображается с помощью толщины линий, в случае работы с крупномасштабными картами — насыщенностью цвета заливки полигона. Толщина линии и насыщенность цвета определяют как процентное отношение отображаемого значения к максимальному значению.

Таблица 5

Дополнительные поля к атрибутивным таблицам линейных участков и элементарных площадок для отображения результаты расчетов распространения и трансформации нефтяного загрязнения через определенные интервалы времени

Поле таблицы	Значение
Распространение загрязнения по акватории за время T_1	
...	
Распространение загрязнения по акватории за время T_N	
Объем испарившейся нефти за время T_1	
...	
Объем испарившейся нефти за время T_N	
Объем осевшей нефти на дно за время T_1	
...	
Объем осевшей нефти на дно за время T_N	
Объем диспергировавшей нефти и эмульсифицированной воды за время T_1	
...	
Объем диспергировавшей нефти и эмульсифицированной воды за время T_N	

Таблица 6

Дополнительные поля к атрибутивной таблице береговой поверхности водотока для отображения результаты расчетов распространения и трансформации нефтяного загрязнения через определенные интервалы времени

Поле таблицы	Значение
Объем осадения нефти на растениях	
Объем инфильтрата в грунт берега за время T_1	
...	
Объем инфильтрата в грунт берега за время T_N	

Графическое представление отображения результатов моделирования в ГИС показано на рис. 5, где градациями серого цвета показаны степень загрязнения береговой поверхности и объем испарившейся нефти в течение времени T_i после начала разлива.

В случае, когда целью моделирования является определение максимально возможного ареала рас-

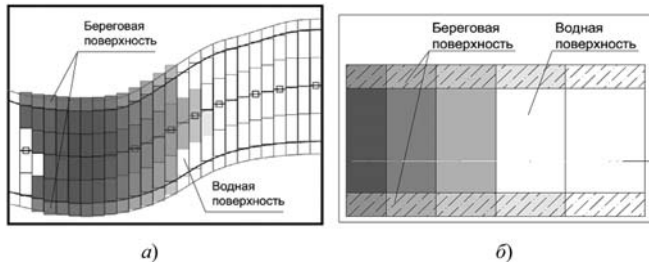


Рис. 5. Представление результатов моделирования для мелко-масштабной (а) и крупномасштабной (б) карт

Таблица 7

Дополнительные поля к атрибутивным таблицам линейных участков и элементарных площадок для определения максимально возможного ареала распространения нефтяного загрязнения

Поле таблицы	Значение
Распространение загрязнения по акватории	
Объем испарившейся нефти	
Объем осевшей нефти на дно	
Объем диспергировавшей нефти и эмульсифицированной воды	

Таблица 8

Дополнительные поля к атрибутивной таблице береговой поверхности водотока для определения максимально возможного ареала распространения нефтяного загрязнения

Поле таблицы	Значение
Объем осадения нефти на растениях	
Объем инфильтрата в грунт берега	

Таблица 9

Дополнительные поля к атрибутивным таблицам линейных участков и элементарных площадок для определения интервалов времени, через которые аварийный разлив достигнет заданных створов

Поле таблицы	Значение
Распространение загрязнения по акватории до створа C_1	
Распространение загрязнения по акватории до створа C_M	
Объем испарившейся нефти до створа C_1	
Объем испарившейся нефти до створа C_M	
Объем осевшей нефти на дно до створа C_1	
Объем осевшей нефти на дно до створа C_M	
Объем диспергировавшей нефти и эмульсифицированной воды до створа C_1	
Объем диспергировавшей нефти и эмульсифицированной воды до створа C_M	

пространения нефтяного загрязнения, из расчетной программы в таблицы атрибутивной информации линейного участка и элементарной площадки добавляются поля, приведенные в табл. 7.

Для отображения результатов расчета в таблицу атрибутивной информации о береговой поверхности добавляются поля, показанные в табл. 8.

Когда целью моделирования является определение интервалов времени, через которые аварийный разлив достигнет заданных створов, из расчетной программы в таблицы атрибутивной информации линейного участка и элементарной площадки добавляются поля согласно табл. 9.

Так как отображение в ГИС интервалов времени, необходимых нефтяному загрязнению для попадания в интересующие створы водотока, представляется затруднительным, то возникает необходимость в создании файла-отчета, содержащего информацию об этих интервалах времени. Кроме этого в файле-отчете содержатся результаты расчетов общей площади поверхности аварийного разлива, объемы испарившейся и осевшей на дно и на береговую поверхность нефти.

Заключение

Разработана ГИС малого речного бассейна, позволяющая осуществлять ввод исходной информации, характеризующей русло водотока, береговую поверхность, погодные-климатические условия, параметры нефтепродуктов и объекты инфраструктуры в единой формализованной форме, необходимой для расчета распространения и трансформации аварийного нефтяного разлива в условиях малого водотока. Использование предложенной ГИС позволило автоматизировать процесс ввода исходной информации, а также повысило информативность результатов моделирования; сделало отображение результатов моделирования более наглядным.

Разработанная ГИС служит средством представления результатов моделирования, распространения и трансформации нефтяных загрязнений по акваториям малых водотоков, что позволяет оптимальным образом планировать операции по локализации и ликвидации аварийных нефтяных разливов, а также более адекватно оценивать размер экологического ущерба.

Список литературы

1. Ларионов В. А. Моделирование аварийных разливов нефти на суше с применением ГИС-технологий: методика. М.: МНТЦ БЭСТС, 2004. С. 11–13.
2. Павлов А. А., Черняев А. В. Моделирование процессов осаждения нефтяных загрязнений на береговую поверхность малых рек // Информационные технологии. 2009. № 11. С. 37–40.
3. Черняев А. В., Павлов А. А. Моделирование процессов распространения нефтяных загрязнений по акватории малых водотоков // Записки горного института. 2013. Т. 203. Проблемы рационального природопользования. С. 237–240.

АВТОМАТИЗИРОВАННЫЕ СИСТЕМЫ УПРАВЛЕНИЯ AUTOMATED CONTROL SYSTEMS

УДК 621.391

А. В. Волынская, канд. техн. наук, доц.,
Уральский государственный университет путей сообщения, г. Екатеринбург,
e-mail: anna-volinskaya@mail.ru

Разработка устройств формирования и оптимальной обработки сложных двоичных сигналов для информационных систем железнодорожного транспорта

Приведены результаты поиска двоичных сигналов с оптимальной автокорреляционной функцией, а также компьютерного моделирования в целях оценки их помехоустойчивости; представлены варианты технической реализации генератора и оптимального фильтра-обнаружителя этих новых сигналов.

Ключевые слова: сложные двоичные сигналы, корреляционный прием, коды Баркера, компьютерное моделирование помехоустойчивости, генератор сигналов, оптимальный фильтр-обнаружитель сигналов

A. V. Volynskaya

Development of Devices of Formation and Optimum Processing of Difficult Signals for the Information Systems of Railway Transport

Search results of binary signals with optimum autocorrelation function, and also computer modeling for the purpose of an assessment of their noise stability are given; options of technical realization of the generator and the optimum filter detector of these new signals are presented.

Keywords: binary signals, the correlation reception, Barker's codes, computer modeling of a noise stability, the generator of signals, the optimum filter detector of signals

Постановка задачи

В цифровых каналах связи решается или задача различения сигналов, или задача обнаружения сигнала, которую можно рассматривать как частный случай задачи обнаружения. Для решения этих задач оптимальным образом следует применять корреляционный прием. Корреляционный прием тем эффективней, чем сложнее полезный сигнал [1, 2]. Но не все сложные сигналы одинаково эффективны для решения задачи обнаружения. Наилучшими являются те, у которых отношение N главного пика функции автокорреляции (АКФ) к боковым пикам — наибольшее. Такие сигналы (двоичные последовательности) известны и широко применяются, их называют сигналами (коды) Баркера [3]. Сигналов Баркера всего восемь, самый сложный из них состоит из 13 символов и имеет отношение высоты главного пика АКФ к боковым $N = 13$. Это свойство позволяет надежно обнаруживать такой сигнал при отношениях сигнал/помеха $C/P < 1$. Однако в наиболее "тяжелых" (в смысле помехо-

устойчивости) каналах, например в каналах телемеханики электрифицированного железнодорожного транспорта, даже сигналы Баркера не обеспечивают требуемой надежности их обнаружения.

Результаты теоретических исследований

Нами найдены сигналы, имеющие большее, чем у сигналов Баркера отношение N , способы поиска таких последовательностей описаны в работе [4]. Один из способов основан на комбинировании сигналов Баркера. В качестве "материнской" последовательности берется последовательность Баркера, а затем каждый элемент материнской последовательности заменяется прямой или инверсной "дочерней" последовательностью Баркера же, в зависимости от того, ноль или единица в материнской последовательности. Назовем эти новые сигналы *комбинированными сигналами Баркера*.

Из 38 возможных парных комбинаций материнских и дочерних последовательностей нашему тре-

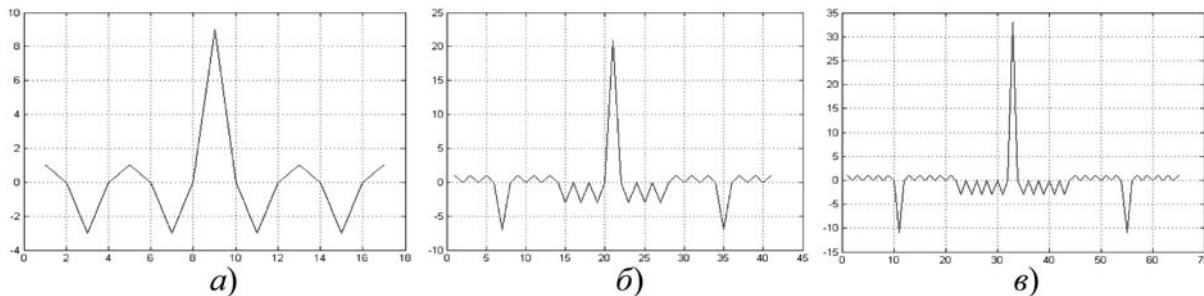


Рис. 1. АКФ сигналов 3—3 (а), 3—7 (б) и 3—11 (в)

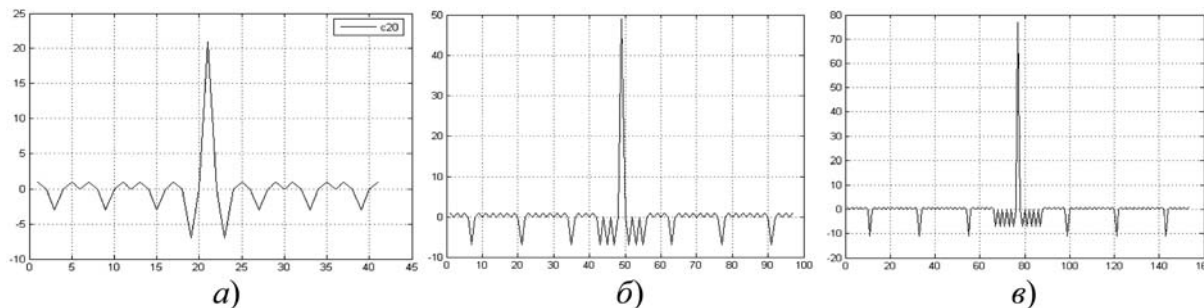


Рис. 2. АКФ сигналов 7—3 (а), 7—7 (б) и 7—11 (в)

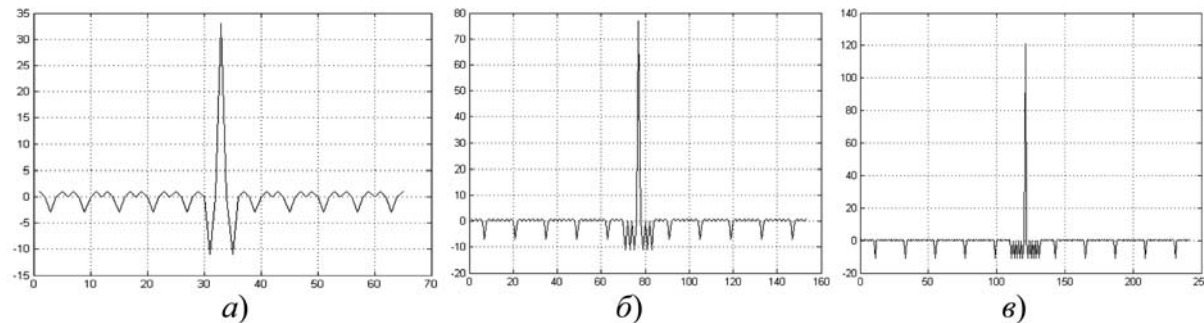


Рис. 3. АКФ сигналов 11—3 (а), 11—7 (б) и 11—11 (в)

бованию удовлетворяют только 10 последовательностей:

3—4,1; 3—3; 3—7; 3—11; 7—3; 7—7;
7—11; 11—3; 11—7; 11—11,

где первое число — материнская последовательность, второе — дочерняя. Например, для последовательности 3—7 материнская последовательность — 110, а дочерняя — 1110010, тогда новая последовательность имеет вид:

1 1 1 0 0 1 0 1 1 1 0 0 1 0 0 0 0 1 1 0 1
"1" "1" "0"

На рис. 1—3 приведены АКФ новых сигналов.

Помехи снижают главный пик корреляционной функции и поднимают боковые пики, поэтому, чем больше отношение высоты главного пика АКФ к боковым, тем выше вероятность правильного приема сигналов. Если сравнить это отношение у сигнала Баркера 13 с комбинированным сигналом Баркера 11—11, то видим, что оно в $121/13 = 9,308...$ раз больше.

Результаты моделирования

Автором проведено исследование помехоустойчивости новых сигналов путем моделирования в программной среде LabVIEW [5]. Результаты приведены на рис. 4—9.

На рис. 7 видим, что при отношении С/П = 1/19 сигнал Баркера 11 не обнаруживается, так как

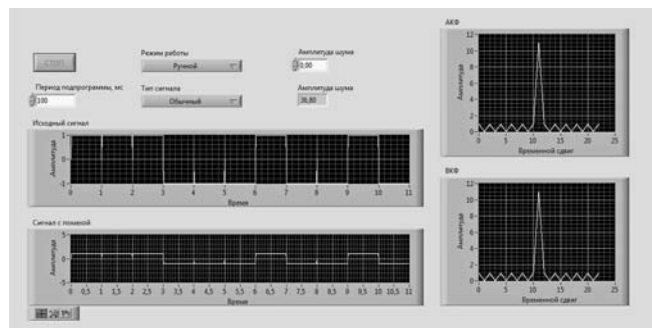


Рис. 4. 11-элементный сигнал Баркера и его АКФ

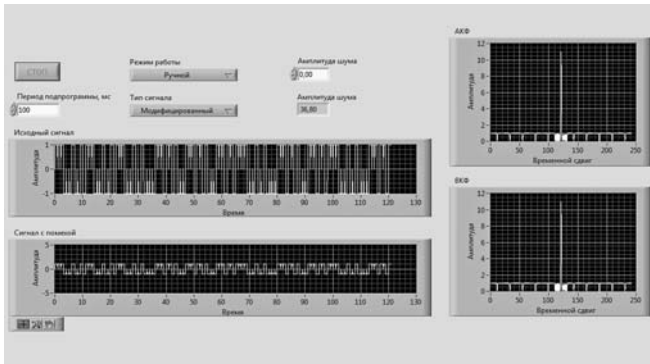


Рис. 5. Комбинированный сигнал Баркера 11–11 и его АКФ

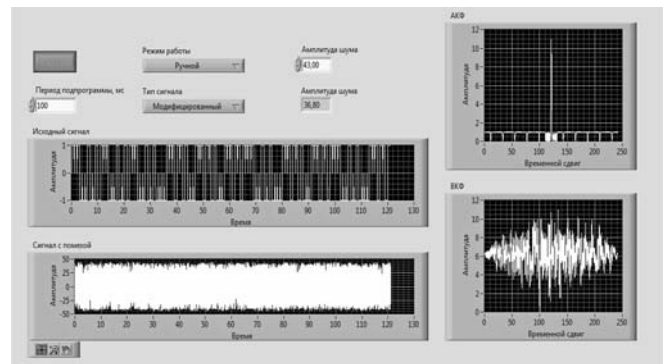


Рис. 9. Комбинированный сигнал Баркера 11–11 + помеха (С/П = 1/43) и ФВК

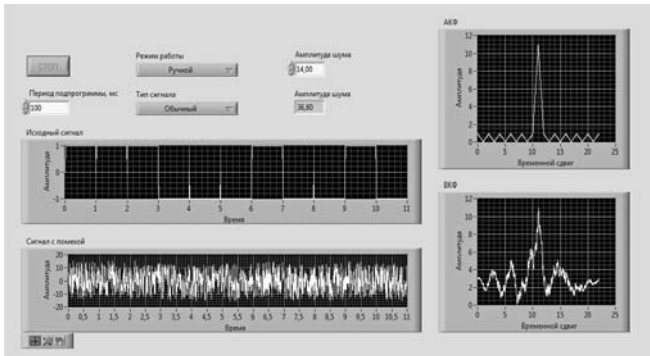


Рис. 6. 11-элементный сигнал Баркера + помеха (С/П = 1/14) и функция взаимной корреляции (ФВК)

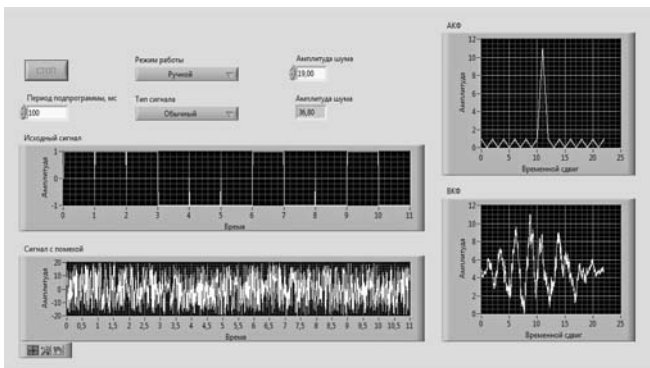


Рис. 7. 11-элементный сигнал Баркера + помеха (С/П = 1/19) и функция взаимной корреляции (ФВК)

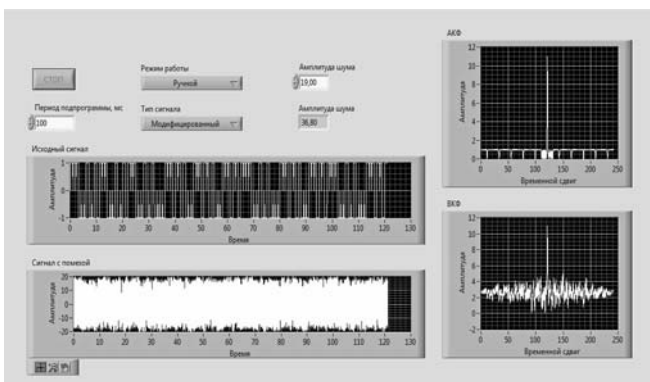


Рис. 8. Комбинированный сигнал Баркера 11–11 + помеха (С/П = 1/19) и ФВК

главный пик ФВК соизмерим с боковыми. Кроме того, это может привести к "ложной тревоге".

Комбинированный сигнал Баркера при этом же отношении С/П надежно обнаруживается, так как главный пик ФВК существенно превышает боковые (рис. 8).

Обнаружение нового сигнала становится затруднительным только при отношении С/П = 1/43.

Техническая реализация

На рис. 10 приведена структурная схема генератора комбинированных сигналов Баркера, который содержит:

- Д3, Д7, Д11 — формирователи дочерних (быстрых) канонических последовательностей Баркера;
- М3, М7, М11 — формирователи материнских (медленных) канонических последовательностей Баркера;
- F — генератор тактовых импульсов;
- F/3, F/7, F/11 — делители частоты на 3, на 7 и на 11;
- ⊕ — сумматор по модулю два;
- П1, П2, П3 — переключатели.

Схемы дочерних и материнских формирователей канонических последовательностей Баркера одинаковые, но работают с разной тактовой частотой.

Генератор работает следующим образом. На верхний вход сумматора по модулю два через переключатель П1 подается сигнал с одного из дочерних формирователей. Например, на рис. 11 показана схема формирователя 7-элементной канонической последовательности Баркера, выполненная автором в программной среде NI Multisim 12.0.1.

На вход IO8 подается одиночный импульс, который продвигается по цепочке триггеров D_FF с частотой тактовых импульсов, подаваемых на вход IO13. Импульсы с выходов триггеров поступают на первые входы схем "И" AND2, вторые входы которых (IO1—IO7) соединены либо с плюсом питания, либо с землей. Далее сигналы с выходов AND2 поступают через схему "ИСКЛЮЧАЮЩЕЕ ИЛИ" OR7 на выход формирователя и далее через пере-

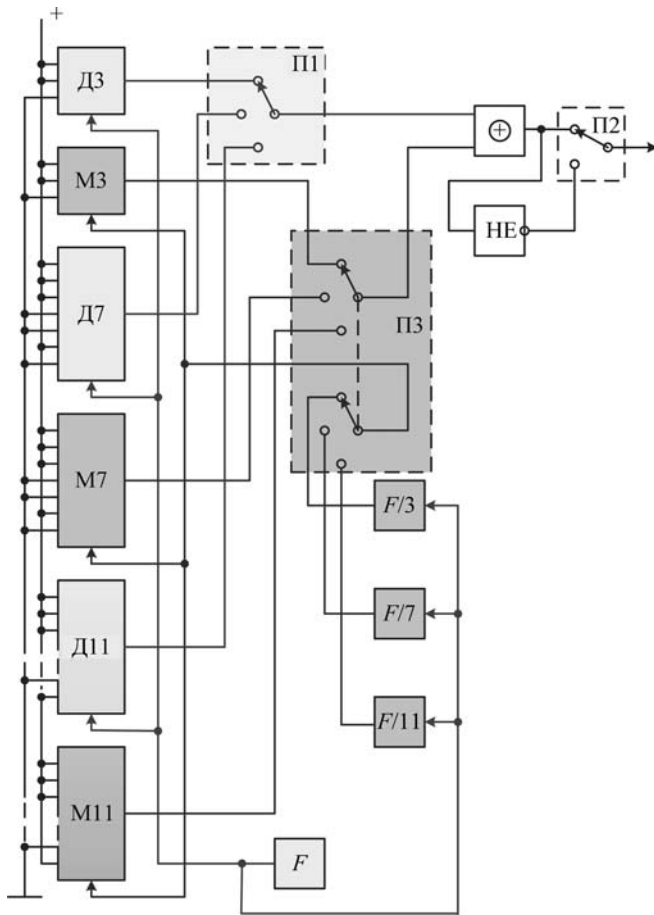


Рис. 10. Структурная схема генератора комбинированных сигналов Баркера

ключатель П1 на верхний вход сумматора по модулю два \oplus , на нижний вход которого поступает какая-либо из материнских последовательностей через верхнюю секцию двухсекционного переключателя П3.

Логика работы сумматора по модулю два такова, что когда на его нижнем входе логический ноль, то для второго входа он работает как повторитель, но, если на нижнем входе — логическая единица, то для второго входа он работает как инвертор. Поэтому на выходе сумматора по модулю два дочерняя быстрая последовательность будет либо неинвертированная, либо инвертированная в соответствии с медленной материнской последовательностью.

На рис. 12 приведен комбинированный сигнал Баркера 3—7 на выходе сумматора по модулю два. Он, в свою очередь, может быть проинвертирован элементом "НЕ" в нижнем положении переключателя П2.

На рис. 13 приведена структурная схема оптимального фильтра-обнаружителя комбинированных сигналов Баркера. Он содержит следующие элементы:

- РС — цифровая линия задержки, выполненная в виде регистра сдвига;
- СМ — аналоговые сумматоры с числом входов 121, 77, 77, 49, 33, 33, 21, 21, 9.
- П — двухсекционный переключатель на девять направлений;

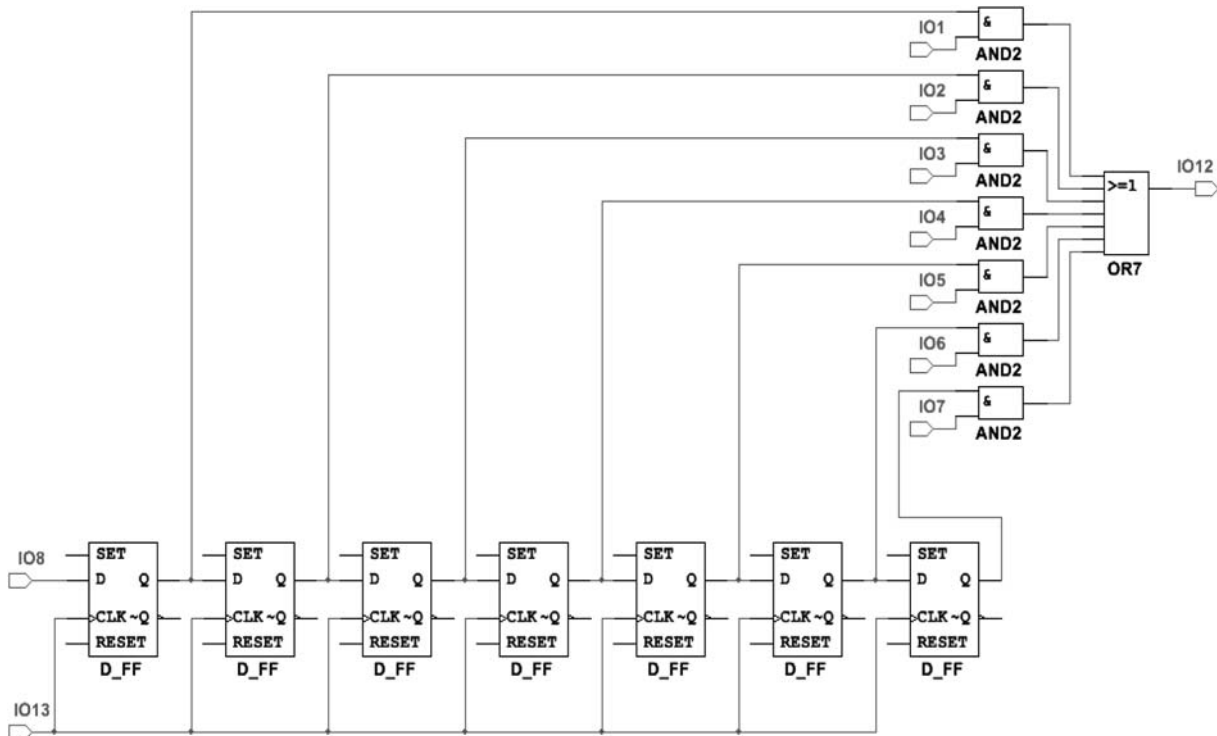


Рис. 11. Формирователь 7-элементной канонической последовательности Баркера (NI Multisim 12.0.1)

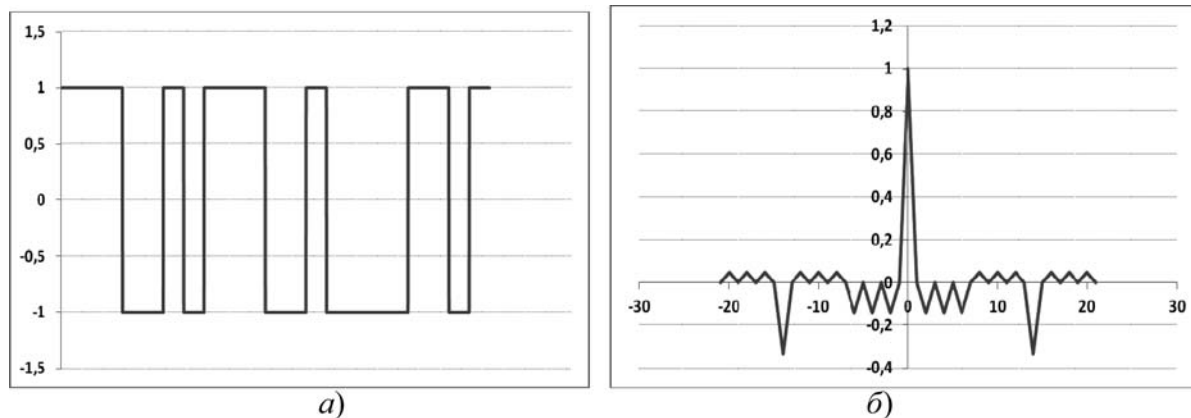


Рис. 12. Комбинированный сигнал Баркера 3–7 (а) и его нормированная АКФ (б)

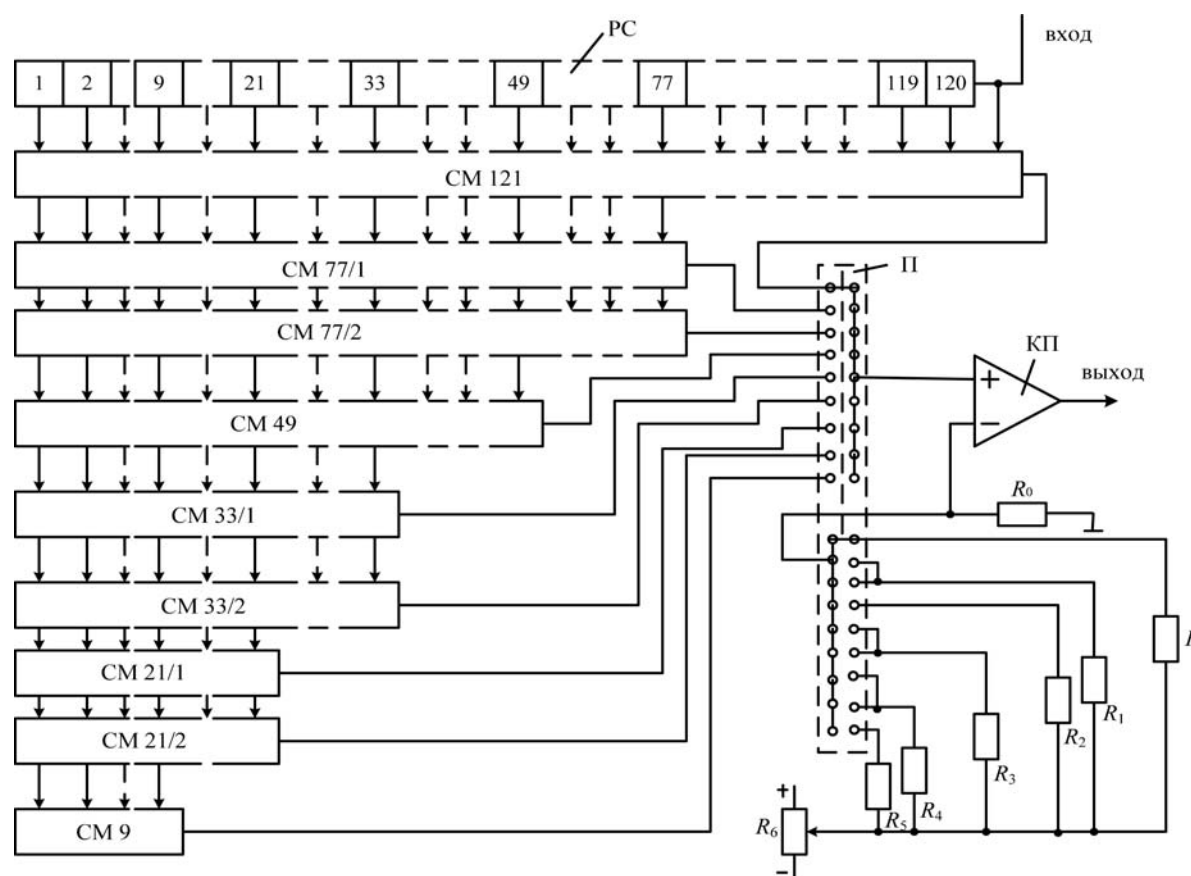


Рис. 13. Оптимальный фильтр-обнаружитель комбинированных сигналов Баркера

- КП — компаратор, выполняющий совместно с резистивным делителем напряжения $R_0—R_6$ роль порогового устройства.

Схема аналогового сумматора CM 21/2, для примера, приведена на рис. 14. Входы IO1—IO21 соединены с первыми соответствующими выходами регистра сдвига. Они, в свою очередь, соединены с инвертирующим входом операционного усилителя либо непосредственно через резистор 21кΩ, либо дополнительно через аналоговый инвертор. Таким образом, сумматор CM 21/2 выполняет роль опти-

мального фильтра для комбинированного сигнала Баркера 7–3, а на его выходе IO22 формируется функция взаимной (при наличии помех в канале) либо авто- (помех нет) корреляции, что показано на рис. 15.

Эта функция имеет главный пик, высота которого зависит от отношения сигнал/помеха, который может быть зафиксирован при превышении им порога, задаваемого потенциометром R_6 , появлением на выходе компаратора КП уровня логической единицы.

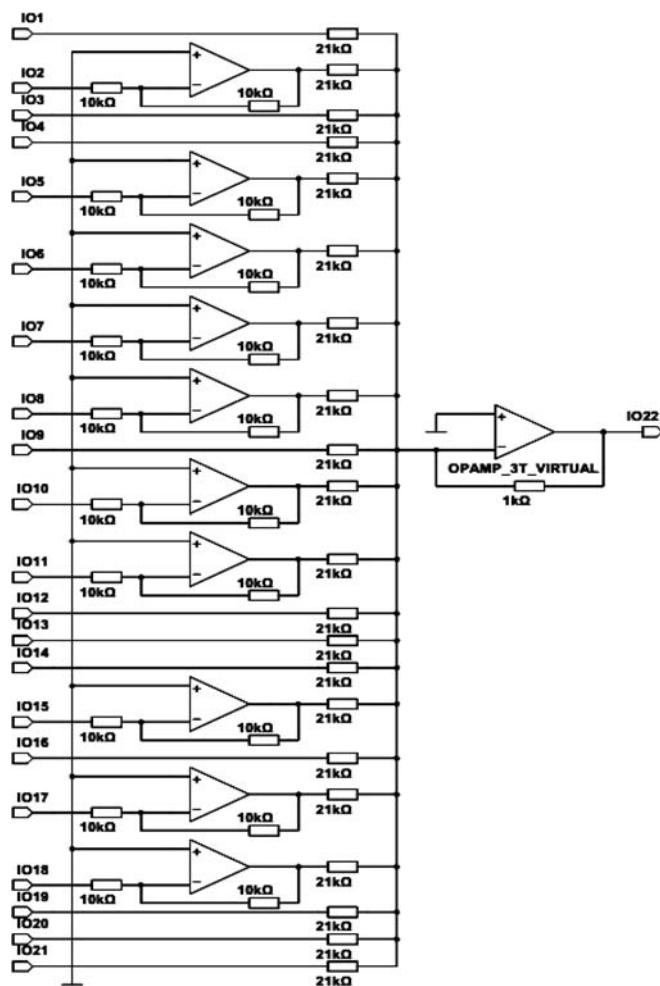


Рис. 14. Аналоговый сумматор CM 21/2 для комбинированного сигнала Баркера 7–3 (NI Multisim 12.0.1)

Такие сигналы предполагается применять в новых альтернативных каналах телемеханики железнодорожного транспорта, например, в PLC-каналах (по контактному проводу электрифицирован-

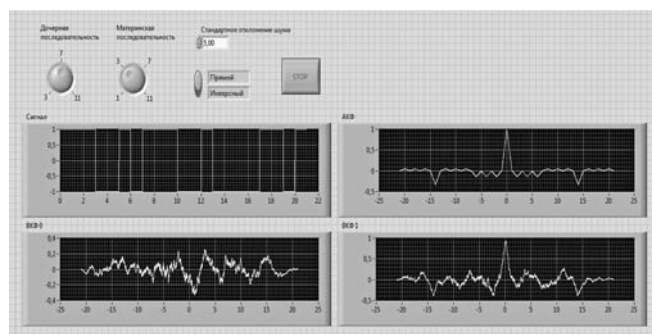


Рис. 15. Комбинированный сигнал Баркера 7–3, его АКФ, помеха и ФВК

ных железных дорог), где уровень электромагнитных помех беспрецедентно высок [6].

Список литературы

1. Ли И., Читем Т., Виснер Дж. Применение корреляционного анализа для обнаружения периодических сигналов в шуме // Теория информации и ее приложения / Под ред. А. А. Харкевича. М.: Государственное издательство физико-математической литературы, 1959. С 138–158.
2. Питерсон В., Бердсал Т., Фокс В. Теория обнаружения сигналов // Теория информации и ее приложения / Под ред. А. А. Харкевича. — М.: Государственное издательство физико-математической литературы, 1959. С. 210–275.
3. Barker R. H. Group synchronizing of binary digital system // Communication theory. London, 1953. 273 p.
4. Волынская А. В. Результаты математического моделирования процесса поиска кодовых последовательностей с заданными корреляционными свойствами // Вестник Уральского государственного университета путей сообщения: Науч.-техн. журнал. Екатеринбург: УрГУПС, 2009. № 3–4. С. 64–71.
5. Волынская А. В., Калинин П. М. Новые помехоустойчивые сигналы для интеллектуального канала телемеханики // Фундаментальные исследования: Науч. журнал РАЕ. 2012. № 11. Ч. 4. С. 922–926.
6. Волынская А. В., Сергеев Б. С. Предпосылки применения псевдослучайных сигналов-переносчиков в каналах телемеханики железнодорожного транспорта // Транспорт: наука, техника, управление: Научный информационный сборник РАН ВИНТИ, 2011. Вып. 6. С. 39–41.

XVI Всероссийская научная конференция RCDL-2014

"Электронные библиотеки: перспективные методы и технологии, электронные коллекции"

13—16 октября 2014 года в Объединенном институте ядерных исследований (г. Дубна) при поддержке Российского фонда фундаментальных исследований, Института проблем информатики РАН, Московской секции ACM SIGMOD будет проводиться очередная XVI Всероссийская научная конференция с международным участием RCDL-2014 — "Электронные библиотеки: перспективные методы и технологии, электронные коллекции".

Электронные библиотеки являются объектом исследований и разработок, направленных на развитие теории и практики обработки, распространения, хранения, поиска и анализа цифровых объектов различной природы. Эти современные репозитории знаний представляют собой сложные онлайновые информационные системы.

Серия Всероссийских научных конференций RCDL, труды которых представлены на постоянном сайте <http://rcdl.ru>, нацелена на формирование российского корпуса международного сообщества ученых, развивающих это научное направление.

Научные электронные библиотеки играют важную роль в реализации активно поддерживаемых научным сообществом инициатив открытого доступа к результатам исследований, в формировании открытой науки.

В рамках конференции RCDL-2014 планируется проведение традиционного семинара молодых ученых "Диссертационные исследования по тематике информационных технологий, связанных с электронными библиотеками", на котором авторам работ, отобранных на основе предварительного рецензирования, будет дана возможность представить текущие результаты своих исследований, а также обсудить их сильные и слабые стороны с более опытными коллегами.

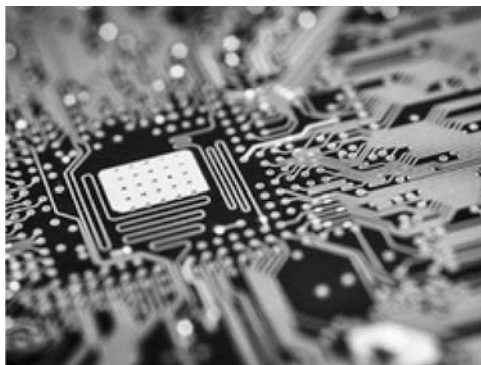
Рабочие языки конференции — русский и английский, представление расширенных тезисов и доклада возможно на любом из них. Программный комитет ожидает представления не публиковавшихся ранее работ, имеющих преимущественно *исследовательский характер* и содержащих *оригинальные* результаты.

Труды конференции будут опубликованы в виде сборника текстов принятых докладов, а также в электронном виде на постоянном сайте конференций RCDL (<http://rcdl.ru>) и в европейском репозитории трудов научных конференций CEUR Workshop Proceedings (<http://ceur-ws.org/>). Тексты лучших докладов, представленных на конференцию, будут рекомендованы к публикации (в доработанном виде) в периодических изданиях, включенных в список ВАК таких, как "Информатика и ее применения", "Системы и средства информатики", "Программная инженерия".

Расширенные тезисы докладов объемом не менее 2000 слов с библиографией принимаются Программным комитетом в электронном виде до 15 мая 2014 года, заявки на проведение тьюториалов и сопутствующих семинаров — до 20 мая 2014 года. Заявленные доклады будут рецензироваться членами Программного комитета. По результатам рецензирования Программный комитет примет решение о принятии или отклонении заявленных докладов и известит о нем авторов до 15 июня 2014 года. Готовый для публикации в трудах конференции текст принятого доклада должен быть представлен в Программный комитет не позднее 1 августа 2014 года.

Более подробная информация о конференции содержится на ее сайте (<http://rcdl2014.jinr.ru>).

Математические школы и IT-прорыв



*Уральский федеральный университет создаст мировой центр информационных технологий на базе матмеха и радиофака
Математические школы университета имеют необходимый для создания мирового IT-центра опыт и потенциал для развития*

Уральский федеральный университет прошел жесткий отбор двух федеральных министерств и теперь создает центр прорывных исследований в области информационных технологий, объединив три научные группы Института математики и компьютерных наук (ранее — матмех УрГУ) и Института радиоэлектроники и информационных технологий — РтФ (ранее — радиофак УПИ).

Центр получит название "Теория автоматов и конечных языков, алгоритмы анализа данных и обработки информации, суперкомпьютеры, компьютерное зрение, робототехника". Его финансирование — по прогнозам университета, 20—30 млн рублей в год в течение пяти лет — будет осуществляться за счет средств федеральных целевых программ "Информационное общество (2011—2020 годы)" Минкомсвязи РФ и "Научные и научно-педагогические кадры инновационной России" на 2014—2020 годы Минобрнауки РФ.

Проект УрФУ вошел в число 19 отобранных федеральными министерствами по поручению председателя правительства из проектов, предложенных 130 вузами и научными организациями. Отбор был направлен на формирование на базе россий-

ских научных организаций и высших учебных заведений исследовательских центров, в которых помимо исследований и разработок мирового уровня в области IT будет осуществляться подготовка кадров, реализующих эффективные принципы и формы интеграции науки, образования и бизнеса.

"Научные школы Уральского федерального университета благодаря многолетнему опыту работы и желанию развиваться способны вывести российскую науку на мировой уровень,— уверен проректор УрФУ по науке Владимир Кружаев. — Центр в области информационных технологий позволит обеспечить новые результаты научных исследований в этой сфере и их прорывы в практику".

Напомним, что 2014-й год объявлен в УрФУ Годом IT. Ключевым событием станет финал чемпионата мира по спортивному программированию, участие в котором примут более тысячи лучших программистов из 120 университетов планеты, ведущих подготовку в сфере информационных технологий.

*Эдуард Никульников,
Тел.: (343) 3507464
www.urfu.ru*

Адрес редакции:

107076, Москва, Стромьинский пер., 4

Телефон редакции журнала (499) 269-5510

E-mail: it@novtex.ru

Дизайнер *Т.Н. Погорелова*. Технический редактор *Е.В. Конова*.

Корректор *Т.В. Пчелкина*.

Сдано в набор 05.12.2013. Подписано в печать 22.01.2014. Формат 60×88 1/8. Бумага офсетная.

Усл. печ. л. 8,86. Заказ IT214. 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.