

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

3(211)
2014

ТЕОРЕТИЧЕСКИЙ И ПРИКЛАДНОЙ НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

Издается с ноября 1995 г.

УЧРЕДИТЕЛЬ
Издательство "Новые технологии"

СОДЕРЖАНИЕ

ОБЩИЕ ВОПРОСЫ

Зацман И. М. Информационно-компьютерная наука: технологические предпосылки становления. 3

МОДЕЛИРОВАНИЕ И ОПТИМИЗАЦИЯ

Божко А. Н., Сюсюкалов Б. С. Математические модели базирования и избыточности в механических системах. 13
Васюков А. В., Ермаков А. С., Петров И. Б., Потапов А. П., Фаворская А. В., Шевцов А. В. Комбинирование сеточно-характеристического метода и метода сглаженных частиц в задачах компьютерного моделирования упругопластических тел. 19

ИНФОРМАЦИОННО-ИЗМЕРИТЕЛЬНЫЕ СИСТЕМЫ

Ермачихин А. В., Литвинов В. Г. LabVIEW в современной индустрии измерений (обзор). 25

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ЭКОНОМИЧЕСКИХ СИСТЕМАХ

Токмакова А. А. Алгоритм стохастического отбора объектов и признаков в задаче банковского кредитного скоринга. 30
Исмагилова Л. А., Орлова Е. В. Методологические основы и инструментарий бюджетно-налогового моделирования и прогнозирования. 35

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ОБРАЗОВАНИИ

Иванников А. Д. Тематические интернет-порталы как средство агрегации электронного контента в заданной предметной области. 43

ПРИКЛАДНЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Ивашенко А. В., Пейсахович Д. Г. Метод проактивной диспетчеризации в информационной среде для транспортных операторов 5PL. 49

Журнал в журнале НЕЙРОСЕТЕВЫЕ ТЕХНОЛОГИИ

Скорородов А. В. Обнаружение воздушного судна по спектру его звукового сигнала с использованием нейросетевого классификатора. 56
Игнатьев Д. И., Храбров А. Н. Нейросетевое моделирование нестационарных продольных аэродинамических характеристик самолета. 60
Кинарский А. В., Жернаков С. В. Стендовая отладка параметров авиационного газотурбинного двигателя на основе нейросетевых технологий. 69
Приложение. Инсаров В. В., Тихонова С. В., Михайлов И. И. Проблемы построения систем технического зрения, использующих комплексирование информационных каналов различных спектральных диапазонов

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

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

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

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

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

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

INFORMATION TECHNOLOGIES

INFORMACIONNYYE TEHNOLOGII

3(211)
2014

THEORETICAL AND APPLIED SCIENTIFIC AND TECHNICAL JOURNAL

Published since November 1995

УЧРЕДИТЕЛЬ
Издательство "Новые технологии"

CONTENTS

GENERAL QUESTIONS

Zatsman I. M. Computer and Information Science: Technological Prerequisites for Formation 3

MODELING AND OPTIMIZATION

Bozhko A. N., Syusyukalov B. S. Mathematical Models for Basing and Redundancy in Mechanical Systems 13
Vasyukov A. V., Ermakov A. S., Petrov I. B., Potapov A. P., Favorskaya A. V., Shevtsov A. V. Combining the Grid-Characteristic Method and the Method of Smoothed Particle for Computer Modeling of Elastoplastic Bodies. 19

INFORMATION-MEASURING SYSTEMS

Ermachikhin A. V., Litvinov V. G. LabVIEW in Modern Measurement Industry (Summary) 25

INFORMATION TECHNOLOGIES IN ECONOMIC SYSTEMS

Tokmakova A. A. Stochastic Optimization Algorithm for Feature and Object Selection with Application to Credit Scoring 30
Ismagilova L. A., Orlova E. V. Methodological Basis and Tools for Fiscal Modeling and Forecasting 35

INFORMATION TECHNOLOGIES IN EDUCATION

Ivannikov A. D. Subject Devoted Internet Portals as an Instrument for Digital Content Aggregation 43

APPLIED INFORMATION TECHNOLOGIES

Ivaschenko A. V., Peysakhovich D. G. Proactive Dispatching in the Integrated Information Space for SPL Providers in Transportation Logistics 49

Journal-in-journal NEUROTECHNOLOGIES

Skorokhodov A. V. Aircraft Detection by its Audio Signal Spectrum using Neural Network Classifier 56
Ignatyev D. I., Khrabrov A. N. Modeling of Unsteady Longitudinal Aerodynamic Characteristics of Aircraft using Neural Networks 60
Kinarskiy A. V., Zhernakov S. V. Bench Debugging of Aviation Gas Turbine Engine Parameters Based on the Neuronet Technology 69

APPLICATION. Insarov V. V., Tikhonova S. V., Mikhaylov I. I. The Problems of Constructing a Machine Vision Systems Using Fusion of Information Incoming from Channels of Different Spectral Ranges.

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"

ОБЩИЕ ВОПРОСЫ GENERAL QUESTIONS

УДК 002:004

И. М. Зацман, д-р техн. наук, зав. отделом, e-mail: izatsman@yandex.ru,
ИПИ РАН

Информационно-компьютерная наука: технологические предпосылки становления

Рассматриваются концептуальные и программные документы 7-й Рамочной программы Евросоюза, в которых обобщается необходимость разработки новой парадигмы создания информационных технологий будущих поколений. Анализируются идеи Горна и Шрейдера об интеграции информационной и компьютерной наук в интересах формирования единой области знаний "информационно-компьютерная наука". Приводится описание характерных черт этой области знаний, которая позиционируется как возможный вариант новой парадигмы. В интересах дальнейшего развития теоретических оснований информационно-компьютерной науки формулируется аксиома герметичности ее сред: ментальной, социально-коммуникационной и цифровой электронной. Анализируются примеры из медицинской информатики и компьютерной лингвистики, которые иллюстрируют роль этой аксиомы в процессе создания информационных технологий.

Ключевые слова: информационная наука, компьютерная наука, информатика как компьютерная и информационная наука, информационные технологии будущих поколений

I. M. Zatsman

Computer and Information Science: Technological Prerequisites for Formation

Conceptual and program documents of the 7th Framework Programme of the European Union in which need of new paradigm development for information technology creation of future generations are considered. Gorn's and Shreider's approaches to the coalescence of computer and information sciences in order to form a unified knowledge domain "computer and information science" are analyzed. The description of characteristic features of this knowledge domain as possible paradigm is provided. In interests of further development of the theoretical foundations for computer and information science, the hermeticity axiom of its media (mental, social and communication, digital electronic) is formulated. Examples from medical informatics and computer linguistics, which illustrate a role of this axiom in the course of information technology creation, are considered.

Keywords: information science, computer science, informatics as computer and information science, future generation information technology

Введение

Необходимость фундаментального переосмысления теоретических оснований создания информационных технологий (ИТ) будущих поколений проходит "красной нитью" в концептуальных и программных документах 7-й Рамочной программы Евросоюза [1–6]. Необходимость их переосмысления является, в частности, следствием экспериментальных исследований степени взаимосвязанности научных дисциплин, включая компьютерную науку, и патентуемых ИТ. В результате этих исследований были определены те научные дисциплины (на-

правления научных исследований), результаты которых используются в процессе создания ИТ и других технологий. Были вычислены индикаторы взаимосвязанности, каждый из которых является интенсивностью цитирования результатов исследований одной научной дисциплины в описаниях патентуемых технологий [7].

Для европейских патентов по ИТ были получены следующие значения этих индикаторов:

- компьютерная наука — 4,55 % (т. е. менее 5 % публикаций, которые цитируются в ИТ-патентах, относятся к этой науке);
- ядерная физика — 2,89 %;

- прикладная физика — 2,69 %;
 - акустика — 2,69 %;
 - биохимия и молекулярная биология — 2,27 %.
- Для патентов США по ИТ были получены следующие значения:
- прикладная физика — 3,23 %;
 - компьютерная наука — 2,35 %;
 - акустика — 2,21 %;
 - оптика — 1,99 %;
 - радиология — 1,84 %;
 - биохимия и молекулярная биология — 1,71 % [8].

Для патентуемых ИТ эти данные иллюстрируют несущественные отличия в доле научных публикаций для разных научных дисциплин. При этом индикаторы взаимосвязанности для компьютерной науки, являющейся профильной для ИТ, не сильно отличаются от индикаторов для физических дисциплин и биохимии. Существуют небольшие отличия в региональном разрезе, например, доля научных публикаций по компьютерной науке, цитируемых в патентах США по ИТ, равна 2,35 %, что на 2,2 % меньше, чем доля научных публикаций по компьютерной науке, цитируемых в европейских патентах.

Кроме ИТ аналогичные индикаторы взаимосвязанности были вычислены для технологий полупроводниковой промышленности (58,7 % всех научных публикаций относятся к физике), биотехнологий (52,8 % всех научных публикаций относятся к наукам о жизни), технологий химии полимеров (42,6 % всех научных публикаций относятся к химии) и для других групп технологий [9]. Сопоставление значений индикаторов взаимосвязанности профильных наук с соответствующими группами технологий свидетельствует о том, что для компьютерной науки их значения, как минимум, в десять раз меньше, чем для других профильных наук. Приведенные экспериментальные данные подтверждают необходимость фундаментального переосмысления теоретических оснований создания ИТ будущих поколений, в частности, в интересах увеличения степени взаимосвязанности информационно-компьютерных научных результатов с патентуемыми техническими решениями в сфере ИТ.

В данной статье в качестве концептуальной основы переосмысления предлагается использовать идею интеграции информационной науки [10, 11] с компьютерной наукой [12—14] и формирования на этой основе единой области знаний [15]. Эта идея была предложена в 1963 г. Солом Горном [16, 17]. Похожая формулировка этой идеи была предложена Юлием Шрейдером независимо от Горна [18].

Интеграция информационной и компьютерной наук как основа переосмысления теоретических оснований создания ИТ рассматривалась в работах [19, 20]. При этом были определены общие объекты предметных областей этих наук следующим образом:

"Имеется непустое пересечение множества символьных выражений "языка" компьютерной науки и множества слов естественных языков информаци-

онной науки, являющихся знаковой информацией, то есть знаковыми формами представления знаний и главными объектами социокультурных коммуникаций. Это объектное пересечение (то есть пересечение объектов исследования в компьютерной и информационной науках) относится одновременно к предметным областям обеих наук. Однако эти объекты в компьютерной науке трактуются и обрабатываются как абстрактное множество символьных выражений, а в информационной науке эти же объекты трактуются и обрабатываются как множества конкретных слов естественных языков ... Для интеграции двух подходов к трактовке и обработке этого объектного пересечения необходима новая научная парадигма информационно-компьютерной науки".

Анализ первых работ по интеграции информационной и компьютерной наук [16—18] показал, что попытки реализации идеи интеграции неизбежно требуют дать ответы на следующие вопросы:

- ✓ На каких теоретических основаниях должна строиться информационно-компьютерная наука как единая область знаний?
- ✓ В какой системе аксиом и с использованием каких терминов следует строить информационно-компьютерную науку как единую область знаний?

Таким образом, научная парадигма информационно-компьютерной науки должна включать описание ее системы терминов, аксиом и теоретических оснований. Кроме того, из них должны логически следовать существующие основания обеих наук: информационной и компьютерной.

Основная цель статьи заключается в описании одного из возможных подходов к построению теоретических оснований информационно-компьютерной науки, предложенного в работе [21], и его развитии. Этот подход представляет собой дальнейшее развитие идей Горна и Шрейдера. Их краткому обзору посвящен следующий первый раздел статьи¹. Похожие идеи продолжают предлагаться и обсуждаться в начале XXI века [15]. Иногда они заново формулируются в редуцированной форме и предлагаются как теоретическая основа для разработки новых информационных и конвергентных технологий. При этом работы Горна и Шрейдера, являющихся авторами фундаментальных и более общих подходов к описанию информационно-компьютерной науки, часто не упоминаются из-за их забвения.

Второй раздел статьи содержит описание современных технологических предпосылок становления информационно-компьютерной науки, а также пример редуцированного подхода к интеграции компьютерной и информационной наук. В разделе 3 формулируется аксиома герметичности ее фундаментальной, социально-коммуникационной и циф-

¹ Идеи Горна и Шрейдера были практически полностью преданы забвению. Более подробное описание их подходов можно найти в работах [19, 20], материалы которых использованы в этой статье.

ровой электронной. Показано, что эта аксиома в неявном виде использовалась и ранее в процессе построения теоретических оснований информационно-компьютерной науки.

1. Горн и Шрейдер об информатике как информационно-компьютерной науке

В апреле 2013 г. исполнилось 50 лет, как была опубликована статья Сола Горна "Информационно-компьютерные науки как новая фундаментальная область знаний" [16]. В этой работе Горн использовал множественное число для ее названий. Единственное число появилось в 1983 г. в названии его статьи "Информатика как информационно-компьютерная наука: ее идеология, методология и социальные аспекты" [17]. Далее в статье термины "информатика" и "информационно-компьютерная наука" рассматриваются как синонимы.

Пятьдесят лет назад Горн рассматривал результат интеграции информационной науки (information science) и компьютерной науки (computer science) как новую фундаментальную дисциплину, что нашло отражение в названии его работы, включающей следующие слова "a new basic discipline" [16]. Аналогичный подход к определению информатики как единой области знаний был предложен Ю. А. Шрейдером [18]. Его результаты были получены независимо от Горна, но идейно их подходы во многом весьма близки. В статье Шрейдера говорится, что не существует двух информатик (информационная наука и компьютерная наука), а есть два облика информатики. Первый из них (информационная наука) дополнительно нагружен представлениями о традиционном информационном обслуживании специалистов-ученых и инженеров в области их профессиональных интересов. Второй облик (компьютерная наука) неправомерно искажен чисто программистскими проблемами, не специфичными для информатики. Специфические же проблемы информатики оказываются там, где возникают задачи информационного представления знаний в форме, удобной для обработки, передачи и творческого реконструирования знаний в результате усилий пользователя [18, стр. 51].

В 1963 г. Горн перечислил ряд вопросов, которые относятся к информационно-компьютерной науке: "Примерами основных вопросов исследования в этой области могут быть системы программирования, проектирование компьютерных систем, искусственный интеллект, информационный поиск и т. д. Вероятностная теория информации Шеннона определенно принадлежит к этой области знания, но помимо нее существует теория компьютерной информации и процессов ее обработки, которую также необходимо включить в предметную область этой науки. Одним из центральных вопросов этой новой науки, скорее всего, станет анализ и синтез машинных языков и процессоров их обработки" [16, стр. 150].

После перечисления направлений исследований он рассматривает информационно-компьютерную науку как учебную дисциплину. Горн сопоставляет новую область знаний с другими дисциплинами с точки зрения образовательного процесса и формулирует следующие вопросы. Каким образом студент может узнать, относится ли сфера его интересов именно к этой новой области знаний, а не к одной из уже устоявшихся дисциплин? Какое ему необходимо образование для того, чтобы углубиться в эту новую область знаний? Горн рассматривает эти вопросы, отталкиваясь в своих рассуждениях от прагматических аспектов исследований в этой новой области знаний [16, стр. 154].

Горн приводит перечень тех дисциплин, которые должны преподаваться студентам, изучающим информационно-компьютерную науку: математика, физика, философия, лингвистика, психология, вычислительная техника и компьютерное программирование. Предложенный им подход к изучению информационно-компьютерной науки уже тогда начал опробоваться в Пенсильванском университете. Через двадцать лет, в 1983 г., когда уже накопился опыт ее преподавания, Горн пишет, что его понимание парадигмы информационно-компьютерной науки заключается в том, что эта область знания не является ветвью математики, так как она должна соотносить себя с прагматическими вопросами, от которых математика не должна зависеть [17, стр. 137].

Главный вывод Горна, которым он завершает статью 1983 г., состоит в следующем: "...нам не следует отделять компьютерную науку от информационной науки, а следует пытаться отстаивать единую область знания — информатику. Любая попытка поощрить разделение ... повлечет за собой отделение практики от знаний ... Такое разделение будет причиной прекращения деятельного кипения, которое поддерживается сплавом знаний и практической деятельности ..." [17, стр. 139—140].

В 1988 г., т. е. через пять лет после публикации 1983 г., независимо от Горна аналогичный подход был предложен Ю. А. Шрейдером. В своей статье он формулирует ряд ключевых положений научной парадигмы информатики: "...информация есть общественное достояние, она в принципе социальна, в то время как знание, вообще говоря, соотносено с конкретной личностью, с тем, кто им владеет и непосредственно пользуется. ... Информация должна пройти через "когнитивный экран" тех, для кого она представляет ценность. Так возникает необходимость считаться не только с существованием мира объективированного социализированного знания, т. е. информации как превращенной формы знания, но и с феноменом личностного знания. ... Тождественность информации и знания при этом исключается, но информация как превращенная форма знания сохраняет следы своего происхождения. ... Наиболее принципиальные вопросы информатики

всегда возникали на стыке информации и знания, там, где речь шла о превращении одного в другое". Далее Шрейдер пишет о пропасти, разделяющей *информацию и знания как сущности разной природы* (выделено мной — ИЗ) [18, стр. 50—51].

Именно это положение Шрейдера послужило отправной точкой для формулировки аксиомы герметичности сред предметной области информационно-компьютерной науки. Эта аксиома, которая будет описана далее в разделе 3, является краеугольным камнем ее теоретических оснований.

Через пятьдесят лет после публикации первой работы Горна [16] наблюдается следующий парадокс. С одной стороны, фундаментальные работы Горна, который предсказал становление новой области знаний, редко упоминаются (по данным на март 2012 г. индекс цитируемости работы [16] по Web of Science равен 16, а работы [17] — 11).

С другой стороны, в наши дни предлагаются редуцированные подходы к интеграции информационной и компьютерной наук, являющиеся частными случаями подходов Горна и Шрейдера. Ссылки на их работы, как правило, отсутствуют, что говорит о забвении их идей. Один пример предлагаемого редуцированного подхода рассматривается в следующем разделе.

2. Новые направления исследований

В наши дни появляются новые направления исследований и разработок, которые оказывают существенное влияние на процессы интеграции информационной и компьютерной наук. Например, в начале XXI века стало формироваться комплексное направление исследований, получившее название "когнитивная информатика" [22—25]. Ее предметная область включает актуальные проблемы целенаправленной генерации и эволюции новых знаний, находящиеся на стыке когнитивной, компьютерной и информационной наук [26, 27].

В рамках рассматриваемого подхода к интеграции эти проблемы отнесены к предметной области информационно-компьютерной науки, что существенно ее расширяет. Такое расширение необходимо учитывать при построении теоретических оснований информационно-компьютерной науки. Поэтому сначала предлагается уточнить границы предметной области новой области знаний, а только потом вносить изменения в ее теоретические основания. С этой целью далее будут рассмотрены следующие концептуальные и программные документы, которые были использованы в процессе уточнения границ предметной области:

- аналитический отчет по вопросам обеспечения конкурентоспособности США в XXI веке, подготовленный Консультативным комитетом по информационным технологиям при Президенте США [28] (информацию о содержании этого отчета на русском языке можно найти в работе [29]);

- описание приоритетных направлений исследований и разработок по созданию информационно-коммуникационных технологий (ИКТ) 7-й Рамочной программы ЕС, принятой на период 2007—2013 гг.;
- описание тематики ориентированных фундаментальных исследований по конкурсу Российского фонда фундаментальных исследований (РФФИ) 2011 г. в части разработки основ создания конвергентных технологий.

2.1. Аналитический отчет

Основная идея аналитического отчета заключается в том, что конкурентоспособность США в XXI веке будет во многом определяться широкомасштабным применением новых информационных технологий во всех областях знаний и сферах практической деятельности. Согласно отчету для их создания потребуются дальнейшее развитие вычислительной отрасли в целом и информационно-компьютерной науки в частности. В сфере научного познания информационные технологии представлены в отчете как составляющая триады "научная теория — научный эксперимент — информационные технологии, обеспечивающие моделирование и проведение экспериментов".

Предлагаемое в отчете определение вычислительной отрасли имеет следующий вид [28, стр. 10]: "это быстро растущая мультидисциплинарная предметная область, в которой используются возможности передового компьютеринга (advanced computing) для понимания и решения сложных проблем, и которая включает три компонента:

- алгоритмы (численные и нечисленные), программное обеспечение, разработанное для решения проблем естественных, гуманитарных и инженерных наук;
- *информационно-компьютерная наука* (выделено мной — ИЗ), которая разрабатывает и оптимизирует современные аппаратные, программные и сетевые средства, необходимые для решения вычислительно сложных проблем;
- вычислительная инфраструктура, которая поддерживает решение научных и инженерных проблем, а также развитие информационно-компьютерной науки".

В этом определении используется словосочетание "информационно-компьютерная наука", с помощью которого идея интеграции была отражена С. Горном. В аналитическом отчете информационно-компьютерная наука трактуется как один из трех компонентов вычислительной отрасли. В этом отчете ее предметная область существенно редуцирована по сравнению с работами Горна и Шрейдера. Ее предназначение определяется необходимостью решения вычислительно сложных проблем. Перефразируя Шрейдера, можно сказать, что в этом подходе облик информационно-компьютерной науки сильно редуцирован из-за ориентации на вычислительно сложные проблемы.

2.2. Приоритетные направления исследований и разработок по ИКТ

В документах 7-й Рамочной программы ЕС сформулировано восемь приоритетных направлений исследований и разработок, включая направление "Перспективные ИКТ" [1—6]. В этих документах констатируется необходимость фундаментального переосмысления теоретических оснований создания ИКТ будущих поколений. При этом не рассматриваются концептуальные подходы к такому переосмыслению и термин "информационно-компьютерная наука", выражающий базовый концепт интеграции, не используется.

Цели проектов, финансируемых в рамках направления "Перспективные ИКТ", сформулированы следующим образом: "Своевременная идентификация и обоснование новых тематических направлений исследований и разработок, которые имеют большой научно-технический потенциал и могут стать основой для разработки ИКТ будущих поколений. Эти проекты должны включать междисциплинарные исследования новых и альтернативных подходов к разработке ИКТ будущего и быть нацеленными на *фундаментальное переосмысление системы теоретических, прикладных, методологических и технологических принципов, подходов и понятий, используемых сегодня в сфере ИКТ* (выделено мной — ИЗ)" [3, стр. 54].

Цитируемый документ содержит описание нескольких конкурсных тем направления "Перспективные ИКТ", включая тему "ИКТ долговременного применения". В описании этой темы говорится о необходимости разработки новых подходов к генерации, представлению и сохранению личностных знаний человека, их интеграции в цифровой электронной среде и глобальному использованию, а также формулируется ряд актуальных исследовательских проблем [3, стр. 57—63].

Приведем формулировки двух проблем, иллюстрирующих актуальность рассматриваемых в статье вопросов расширения предметной области новой науки за счет проблем, находящихся на стыке когнитивной, компьютерной и информационной наук [3, стр. 62—63]:

1) разработать теоретические и прикладные основы создания долговечных систем, обеспечивающих их эволюцию при минимизации затрат на их развитие в условиях многократной смены поколений программно-аппаратных и сетевых средств и/или форматов данных. Долговечные системы должны быть способны к обеспечению своей первоначальной социально-значимой функциональности в течение долгого периода времени и изменению ее в случае необходимости. Методы сохранения и изменения функциональных возможностей должны быть компьютерно-независимыми и должны обеспечивать устойчивую эволюцию долговечных систем;

2) разработать новые подходы к представлению и сохранению знаний, ориентированные на долго-

временный и безотказный к ним доступ в условиях локальной генерации отдельных "квантов" знаний, их интеграции, глобального использования систем представления и сохранения знаний с учетом контекста и временной эволюции систем. Должна быть обеспечена долговременная устойчивость систем представления и сохранения знаний в условиях многообразия их использования и *эволюции семантики во времени* (выделено мной — ИЗ).

Таким образом, в документах 7-й Рамочной программы ЕС сформулирована потребность в разработке теоретических оснований создания ИКТ будущих поколений, учитывающих следующие обстоятельства и ограничения:

- устойчивая эволюция долговечных [информационно-компьютерных] систем;
- локальная генерация новых "квантов" знаний, их интеграция, компьютерное представление новых знаний в динамике их формирования;
- глобальное использование систем представления и сохранения новых знаний с учетом контекста и временной эволюции [информационно-компьютерных] систем;
- эволюция семантики [а также форм и структур представления знаний человека] во времени.

Сформулируем ряд вопросов для того, чтобы детализировать перечисленные обстоятельства в части генерации и представления новых знаний. На эти вопросы нужно ответить тем или иным образом в рамках любого предлагаемого подхода к построению теоретических оснований информационно-компьютерной науки. Явное описание ответов на эти вопросы нужно для сравнения разных вариантов их построения.

Во-первых, допускает ли предлагаемый подход категоризацию ментальных знаний по степени их конвенциональности? Есть ли категория личностных ментальных знаний человека, которые их автором ни с кем не согласованы? Если ответ "да", то каковы механизмы и средства распространения личностных знаний и в результате каких процессов происходит их преобразование в коллективные (согласованные) и конвенциональные знания?

Во-вторых, определяются ли в предлагаемом подходе в явном виде источники и цели генерации новых ментальных знаний человека, а также пополнения существующих систем знаний?

В-третьих, можно ли влиять на процессы генерации, пополнения и эволюции систем знаний (сделать эти процессы целенаправленными)? Каковы механизмы и средства этого влияния?

В-четвертых, допускает ли предлагаемый подход разделение ментальных знаний человека на имплицитные, которые не выражены в явном виде, и эксплицитные знания?

Все четыре перечисленных вопроса являются ключевыми. Они не исчерпывают перечень тех вопросов, которые необходимо решить в процессе описания сопоставляемых вариантов теоретиче-

ских оснований создания ИКТ будущих поколений, в частности, и научной парадигмы информационно-компьютерной науки в целом.

2.3. Ориентированные фундаментальные исследования

В марте 2011 г. РФФИ провел конкурс ориентированных фундаментальных исследований по 23 междисциплинарным темам [30]. По всем 23 темам поступило 1137 заявок. Авторы 468 заявок получили гранты на выполнение проектов. По теме "Фундаментальные основы конвергентных технологий" поступило чуть более 40 заявок, из которых почти половина получила гранты на выполнение проектов, включая следующие [31, 32]:

- Исследование и разработка нейроморфных средств и сложных антропоморфных технических систем (рассматриваются модели восприятия информации, накопления знаний о среде и поведении путем обучения в реальном времени).
- Исследование ритмических кодов мыслительной деятельности и создание на этой основе модели когнитивного пространства человека и интерфейса мозг-компьютер высокого уровня (разрабатываются технологии определения типа совершаемых в уме мысленных действий по рисунку электроэнцефалограммы, исследуется распределение ритмов мозга для отдельных видов мышления: пространственного, образного, вербально-логического и смешанных форм, определяется количественная оценка расстояния между разными видами мышления с вычислением их координат на модели когнитивного пространства).
- Когнитивные основы конвергентной технологии распознавания речи (строится обучающее множество паттернов для технологии автоматического распознавания речи).

Названия и краткие аннотации позволяют получить первое представление о проблематике конкурса по теме "Фундаментальные основы конвергентных технологий". В этих проектах исследуются био-, нейро- и когнитивные технологии и рассматриваются вопросы их интеграции с ИКТ. В рамках этого конкурса РФФИ конвергенция технологий разных видов позиционируется как фактор формирования новой технологической базы цивилизации, основанной на воспроизведении систем и процессов живой природы в виде технических систем и технологических процессов [30—32], в том числе в виде информационно-компьютерных систем и технологий.

В XXI веке исследования по конвергенции технологий разных видов позиционируются как актуальная проблематика. В частности, в ежегодных программных документах под названием "NSF Investments and Strategic Goals" Национального научного фонда (ННФ) США вопросы конвергенции технологий появились в 2003 г. Отметим, что одно из приоритетных направлений исследований, финан-

сируемых ННФ, имеет название "Computer and Information Science and Engineering" [33].

Исследования по конвергенции технологий направлены на решение крупных социальных задач, что отражено в самих названиях исследовательских программ, проектов, научных мероприятий и отчетных материалов. Так, отчет 2003 г. по проекту ННФ № CTS-0128860, подготовленный ведущими специалистами США по конвергенции технологий, называется "Converging Technologies for Improving Human Performance" [34], а аналитический отчет 2004 г., подготовленный ведущими специалистами Европейского союза, — "Converging technologies — shaping the future of European societies" [35].

РФФИ финансирует проекты в области интеграции ИКТ с био-, нейро- и когнитивными технологиями на конкурсной основе с 2011 г. Далее в статье результат интеграции всех перечисленных видов технологий кратко будем обозначать как когнитивные ИКТ. Создание фундаментальных основ их разработки является сегодня актуальной тематикой. В рамках рассматриваемого подхода эта тематика также относится к предметной области информационно-компьютерной науки (ИКН).

3. Аксиома герметичности сред предметной области ИКН

Понятие среды было введено К. К. Колиным в интересах описания структуры научных исследований, относящихся к комплексу наук об информации [36—38]. В этом разделе в рамках рассматриваемого подхода обосновывается необходимость принятия аксиомы герметичности сред как базового элемента теоретических оснований ИКН.

В работах [39—43], содержащих описание оснований ИКН, эта аксиома в явном виде не формулировалась. Однако эта аксиома использовалась в неявном виде в описании следующих оснований ИКН:

1) деление всей предметной области ИКН на перечисленные ниже среды;

2) определение двух классов понятий ИКН: первый класс — это по определению однородные понятия (например, знания), второй — двуединые понятия (например, семиотические знаки);

3) распределение однородных понятий — по средам, а двуединых понятий — по границам между средами;

4) использование двух взаимосвязанных стадий представления ментальных знаний человека и составляющих их концептов в цифровой электронной среде в виде кодов.

Приведем краткое описание перечисленных оснований и отношений между ними, используя систему терминов, построенную в работах [39—43], и отметим случаи неявного использования аксиомы герметичности сред, чтобы затем сформулировать ее в явном виде.

Используемая система терминов включает два основных класса понятий: однородные (знания, зна-

ковая информация, данные, цифровая информация, цифровые данные и компьютерные коды) и двуединые (семиотические знаки, формокоды и семокоды). В начале процесса построения теоретических оснований ИКН рассматривались только три среды (затем число сред было увеличено) и следующие соответствующие им однородные понятия:

- *ментальная среда знаний человека*, которая включает выраженные и невыраженные (имплицитные) знания;
- *социально-коммуникационная среда*, к которой относятся отчужденные от человека сенсорно-воспринимаемые формы представления его знаний (знаковая информация) и данные;
- *цифровая электронная среда*, к которой относятся цифровая информация, цифровые данные и компьютерные коды, включая коды концептов и форм представления выраженных знаний.

Таким образом, в процессе построения теоретических оснований ИКН каждое однородное по определению понятие было соотнесено только с одной из перечисленных трех сред [39—43]. Граница между ментальной и социально-коммуникационной средами является, по словам Ю. Шрейдера, той "пропастью", которая разделяет знаковую информацию и знания как сущности разной природы [18, стр. 50—51].

В процессе распределения однородных понятий по средам *неявно предполагалось*, что каждое из этих понятий может быть определено *только в пределах одной среды*. Затем были описаны две последовательные стадии представления знаний в цифровой электронной среде в виде компьютерных кодов. Первая стадия представления знаний относится к границе между ментальной и социально-коммуникационной средами. *Неявно предполагается*, что для концептов и знаковых форм их представления (или форм знаков) эта граница является *непроницаемой для сущностей граничащих сред*. Знаки как двуединые по своей природе понятия принадлежат именно этой границе². Вследствие непроницаемости границы возникает необходимость в ассоциативном соотнесении знаний и концептов с формами их представления с помощью семиотических знаковых систем, включая языковые системы (рис. 1).

Этот рисунок иллюстрирует ситуацию, когда в системе знаний человека могут быть выделены несколько планов содержания в зависимости от числа тех языков или других знаковых систем, которыми он владеет и которые использует для представления своих знаний в отчужденной форме. В используемой системе терминов такие формы по определению являются знаковой информацией. На рис. 1 изображено два плана содержания в ментальной среде (обозначены как А и Б), соответствующие двум разным языковым знаковым системам.

² Здесь использована диадическая модель семиотического знака.

Например, для описания одних и тех же знаний могут использоваться русский язык (план содержания А) и английский язык (план содержания Б). Тогда каждому из этих двух планов содержания, полученных в результате членения одной и той же системы знаний, будет соответствовать свой план выражения в социально-коммуникационной среде. *Неявно предполагается*, что в общем случае это *соответствие является асимметричным* [44].

На рис. 1 каждый план выражения обозначен двумя книгами (две книги на русском языке и две книги их переводов на английский). Граница между ментальной и социально-коммуникационной средами обозначена штриховой линией. Вторая стадия представления знаний относится к границе между социально-коммуникационной и цифровой электронной средами, которая обозначена штрихпунктирной линией (рис. 2).

Неявно предполагалось, что эта граница также является *непроницаемой для сущностей граничащих сред*.

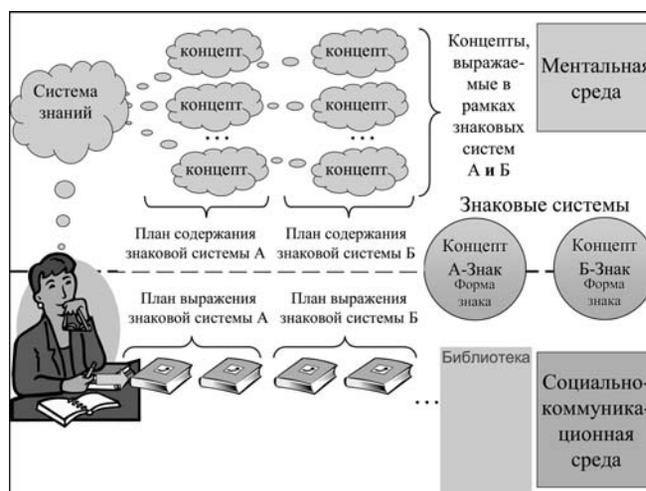


Рис. 1. Две среды, планы содержания и выражения, знаковые системы

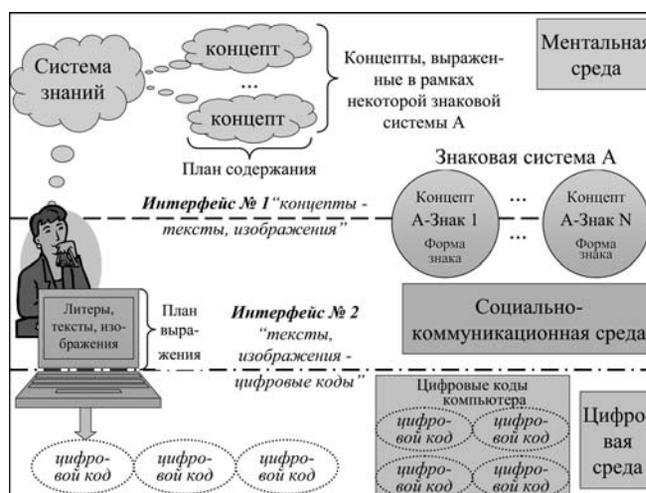


Рис. 2. Три среды, два интерфейса и две стадии представления в цифровой электронной среде знаний и составляющих их концептов в виде кодов

На этой границе используются таблицы компьютерного кодирования литер, необходимые для представления текстов в цифровой электронной среде. Эти таблицы принадлежат границе между социально-коммуникационной и цифровой электронной средами (таблицы кодирования условно обозначены компьютером на рис. 2). Для вербализуемых знаний последовательное использование сначала языковых знаковых систем в процессе генерации текстов, а затем таблиц компьютерного кодирования является традиционным способом представления ментальных знаний человека и составляющих их концептов в цифровой электронной среде.

Рис. 2 иллюстрирует ситуацию, когда изображениям и текстам (как вербальным формам представления знаний) ставятся в соответствие цифровые коды. Важно отметить, что в процессе интеграции информационной науки с компьютерной наукой и формировании единой предметной области получились не две, а три границы между средами (рис. 3):

- между ментальной средой знаний человека и социально-коммуникационной средой (на рис. 2 и 3 обозначена как интерфейс № 1);
- между социально-коммуникационной средой и цифровой электронной средой (на рис. 2 и 3 обозначена как интерфейс № 2);
- между ментальной средой знаний человека и цифровой электронной средой (на рис. 3 обозначена как интерфейс № 3).

Определение в работе [39] двуединых понятий на второй и третьей границах, которые в используемой системе терминов названы, соответственно, формокодами и семокодами, дало возможность предложить методы компьютерного кодирования тех концептов, которые не имеют вербальных форм представления или имеют несколько разных вербальных форм их представления из-за асимметрии языковых знаковых систем [44]. Для каждого из трех интерфейсов *неявно предполагалось*, что соответствующая ему граница является *непроходимой для однородных сущностей граничащих сред*. Семiotические модели, разработанные на основе этих трех интер-

фейсов, дали возможность описать количественно процессы генерации и эволюции новых экспертных знаний в динамике их формирования [26, 27].

Аксиома герметичности представляет собой обобщение на все среды ИКН положения Ю. Шрейдера о пропасти, которая разделяет знаковую информацию и знания как сущности разной природы, принадлежащие социально-коммуникационной и ментальной средам соответственно. Используя выделенные выше неявные предположения, предлагается следующая формулировка *аксиомы герметичности трех сред*: любая однородная сущность, которая определяется в системе терминов ИКН для ментальной, социально-коммуникационной или цифровой электронной среды, существует только в пределах одной из этих сред, но может быть ассоциативно связана с помощью двуединых сущностей с одной или более однородными сущностями других сред.

Из принятия этой аксиомы следует два свойства однородных сущностей:

- любая граница является непроницаемой для однородных сущностей;
- ассоциативные связи однородных сущностей граничащих сред в общем случае являются асимметричными.

Сформулированная аксиома герметичности позволяет уточнить теоретические основания ИКН, эксплицируя те исходные положения, которые ранее использовались неявно. Кроме того, эта аксиома может быть использована непосредственно при разработке тех видов ИКТ, которые охватывают однородные сущности разных сред. При анализе и выборе средств поддержки интерфейсов на границах между средами важно уже на этапе концептуального проектирования выяснить, обладают ли ассоциативные связи, установленные в рамках выбранных интерфейсов, свойством асимметрии, и определить, как это свойство должно использоваться на последующих этапах разработки ИКТ. Приведем два примера технологий, которые охватывают однородные сущности разных сред и учитывают свойство асимметрии.

В работе [39, с. 222–225] была рассмотрена технология обработки файла цифровых данных одного плоского изображения³, который формируется компьютерным томографом. В этой технологии для одного файла, принадлежащего цифровой электронной среде, формально могут быть сгенерированы тысячи разных вариантов его представления, принадлежащих социально-коммуникационной среде. Разные варианты соответствуют разным участкам спектра плотностей тканей или органов, исследуемых с помощью компьютерного томографа.

³ В этом примере файл цифровых данных содержит значения плотности тканей или органов одного слоя (среза), исследуемых с помощью компьютерного томографа. Значения плотностей изменяются по шкале Хаунсфилда от –1024 до 3071, т. е. всего 4096 значений плотности. Нулевое значение по этой шкале соответствует плотности воды.

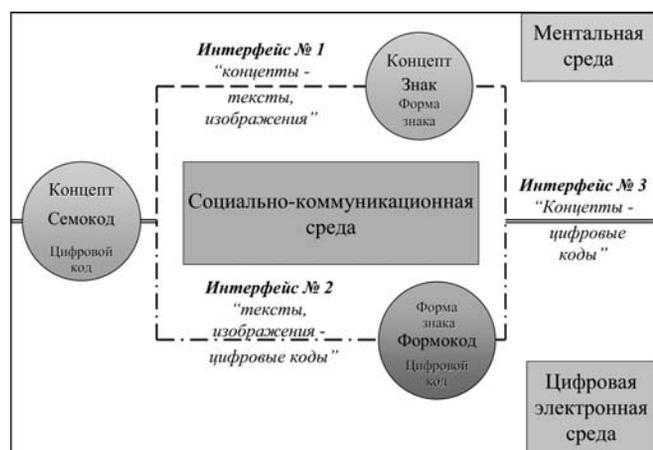


Рис. 3. Три среды и три интерфейса между ними

В этом примере из медицинской информатики свойство асимметрии ассоциативных связей используется для организации исследования разных видов тканей (например, костной ткани, мозговой ткани и других видов мягких тканей) на основе обработки одного файла цифровых данных. При этом увеличение числа вариантов его представления, например, за счет использования специальных мониторов, отображающих 1024 оттенка серого цвета, повышает точность диагнозов.

Второй пример возьмем из компьютерной лингвистики. При разработке технологий поиска актуальными являются задачи редуцирования асимметрии в языковых знаковых системах, например, снятие омонимии при обработке текстов на естественном языке.

Таким образом, в первом примере при проектировании технологий обработки изображений разработчики стремятся в максимально возможной степени использовать асимметрию ассоциативных связей для повышения точности диагнозов, а во втором примере — максимально ее редуцировать при обработке текстов на естественном языке в процессе снятия омонимии.

Заключение

В статье рассматривались только три среды предметной области ИКН. Однако отсюда не следует, что для ее описания всегда будет достаточно именно этих сред. Например, в задачах нейроинформатики иногда необходимо включать в рассмотрение совокупность сущностей четырех сред: ментальной, социально-коммуникационной, цифровой электронной и нейрофизиологической [45]. Соответственно, объекты исследования предметной области нейроинформатики могут в общем случае принадлежать этим четырем средам. В процессе исследования могут использоваться двуединые понятия, принадлежащие следующим шести границам между этими средами:

- между ментальной средой знаний человека и социально-коммуникационной средой;
- между социально-коммуникационной и цифровой электронной средами;
- между ментальной и цифровой электронной средами;
- между нейрофизиологической и социально-коммуникационной средами;
- между нейрофизиологической и цифровой электронной средами;
- между ментальной и нейрофизиологической средами.

Для первой границы традиционным примером таких понятий являются семиотические знаки (см. рис. 2). Для второй и третьей границ ранее были определены аналогичные двуединые понятия "формкод" и "семокод" (см. рис. 3). Сегодня актуальным является вопрос, как определить аналогичные двуединые понятия для оставшихся трех границ.

В настоящее время отсутствует общее описание средового деления предметной области ИКН, учитывающее объекты исследований в нейроинформатике, когнитивной информатике и биоинформатике. Это не позволяет сегодня определить общее число сред, которые охватывает предметная область ИКН.

Необходимость в переосмыслении системы теоретических, прикладных, методологических и технологических принципов, подходов и понятий, используемых сегодня в сфере разработки ИКТ стало ключевой предпосылкой становления ИКН как общенаучной области знаний. В настоящее время осознана и эксплицирована с научной, технологической и образовательной позиций потребность в построении теоретических оснований ИКН как общенаучной области знаний. На основе идей Горна и Шрейдера, в статье рассмотрен только один из возможных подходов к построению ее теоретических оснований, и в рамках этого подхода сформулирована аксиома герметичности сред предметной области ИКН.

Предметные области других областей знаний редко включают подобное разнообразие сред и, соответственно, разнообразие природы объектов исследования. Исключение составляет семиотика. В семиотическом треугольнике Фреге денотат является одной из его вершин и может быть любой природы (материальной, ментальной, социально-коммуникационной, цифровой электронной и т. д.). Если задачей семиотики является изучение знаков и знаковых систем, используемых во всех областях знаний, но *только на одной границе* между ментальной средой знаний человека и социально-коммуникационной средой, то задачей ИКН — создание теоретических основ разработки ИКТ для всех областей знаний и широкого спектра сфер практической деятельности. При этом ИКТ будущих поколений могут охватывать однородные сущности пяти и более сред, а также двуединые сущности на *нескольких границах* между ними.

Предлагаемое развитие идей Горна и Шрейдера является одним из возможных вариантов построения теоретических оснований ИКН, не зависящих от числа сред. Естественно, что могут быть предложены и другие варианты описания ее оснований. Разработка и сопоставительный анализ разных вариантов построения теоретических оснований ИКН является сегодня актуальной общенаучной проблемой, решение которой станет фундаментом для разработки образовательных программ по ИКН и подготовки специалистов для создания ИКТ будущих поколений.

Список литературы

1. **Decision** N 1982/2006/EC of the European Parliament and of the Council of 18 December 2006 concerning the Seventh Framework Programme of the European Community for research, technological development and demonstration activities (2007—2013) // Official Journal of the European Union L412 30.12.2006. P. 1—41.

2. **CORDIS ICT Programme Home.** URL: http://cordis.europa.eu/fp7/ict/programme/home_en.html (состояние страницы на 23.05.2011).
3. **ICT FP7 Work Programme 2007-08.** URL: ftp.cordis.europa.eu/pub/fp7/ict/docs/ict-wp-2007-08_en.pdf (состояние файла на 23.05.2011).
4. **ICT FP7 Work Programme 2009-10.** URL: ftp.cordis.europa.eu/pub/fp7/ict/docs/ict-wp-2009-10_en.pdf (состояние файла на 23.05.2011).
5. **ICT FP7 Work Programme 2011-12.** URL: ftp.cordis.europa.eu/pub/fp7/ict/docs/ict-wp-2011-12_en.pdf (состояние файла на 23.05.2011).
6. **FP7 Exploratory Workshop 4 "Knowledge Anywhere Anytime".** URL: http://cordis.europa.eu/ist/directorate_f/f_ws4.htm (состояние страницы на 23.05.2011).
7. **Verbeek A., Debackere K., Luwel M., Andries P., Zimmermann E., Deleus D.** Linking science to technology: Using bibliographic references in patents to build linkage schemes // *Scientometrics*. 2002. Vol. 54, N 3. P. 399—420.
8. **Научно-технический отчет** по проекту "Linking Science to Technology — Bibliographic References in Patents", выполненному в рамках контракта № ERBHPV2-СТ-1993-03, финансировавшего Еврокомиссии. В 9 т. Т. 7 "Information Technology". Leuven: Katholieke Universiteit Leuven, 2002. 114 p.
9. **Third European Report on Science & Technology Indicators.** Luxembourg: Office for Official Publications of the European Communities, 2003. 451 p.
10. **Гиляревский Р. С.** Основы информатики. М.: Экзамен, 2003.
11. **Информатика как наука об информации:** Информационный, документальный, технологический, экономический, социальный и организационный аспекты / Под ред. Р. С. Гиляревского. М.: ФАИР-ПРЕСС, 2006.
12. **Королев Л. Н., Миков А. И.** Информатика. Введение в компьютерные науки. М.: Высшая школа, 2003.
13. **Final Report of the Computing Curricula 2001 project:** Computer Science (National Science Foundation. Grant N 0003263). URL: http://www.computer.org/portal/c/document_library/get_file?p_l_id=2814020&folderId=3111026&name=DLFE-57603.pdf (состояние страницы на 11.03.2013).
14. **Рекомендации по преподаванию программной инженерии и computer science в университетах (Software Engineering 2004: Curriculum Guidelines for Undergraduate Degree Programs in Software Engineering; Computing Curricula 2001: Computer Science):** Пер. с англ. М.: Интернет-Университет Информационных Технологий, 2007. 462 с.
15. **Snir M.** Computer and information science and engineering: one discipline, many specialties // *Communication of the ACM*, March 2011. P. 38—43.
16. **Gorn S.** The computer and information sciences: a new basic discipline // *SIAM Review*. 1963. Vol. 5, N 2. April. P. 150—155.
17. **Gorn S.** Informatics (computer and information science): its ideology, methodology, and sociology. In: *The studies of information: Interdisciplinary messages* / Ed. by F. Machlup and U. Mansfield. New York: Wiley, 1983. P. 121—140.
18. **Шрейдер Ю. А.** Информация и знание // *Системная концепция информационных процессов*. М.: ВНИИСИ, 1988. С. 47—52.
19. **Зацман И. М., Кожунова О. С.** Предпосылки и факторы конвергенции информационной и компьютерной наук // *Информатика и ее применения*. 2008. Т. 2. Вып. 1. С. 77—98.
20. **Зацман И. М., Кожунова О. С.** Предпосылки конвергенции компьютерной и информационной наук // *Системы и средства информатики: Спец. вып. Научно-методологические проблемы информатики*. М.: ИПИ РАН, 2006. С. 112—139.
21. **Зацман И. М.** Система семиотических аксиом информатики как фундаментальной науки // *Философия и будущее цивилизации: Тез. докл. и выступлений IV Российского философского конгресса (Москва, 24—28 мая 2005 г.)*. В 5 т. Т. 1. М.: Современные тетради, 2005. С. 715—716. URL: <http://www.aintell.info/elib/w2/4.pdf> (состояние страницы на 12.03.2013).
22. **Bryant A.** Cognitive Informatics, Distributed Representation and Embodiment // *Brain and Mind*. 2003. Vol. 4, N 2. P. 215—228.
23. **Wang Y., Kinsner W.** Recent Advances in Cognitive Informatics // *IEEE Transactions on Systems, Man, and Cybernetics (Part C)*. 2006. Vol. 36, N 2. P. 121—123.
24. **Wang Y.** The theoretical framework of cognitive informatics // *International Journal of Cognitive Informatics and Natural Intelligence*. 2007. Vol. 1, N 1. P. 1—27.
25. **Wang Y., Widrow B., Zhang B., Kinsner W., Sugawara K., Sun F., Lu J., Weise T., Zhang D.** Perspectives on the Field of Cognitive Informatics and Its Future Development // *International Journal of Cognitive Informatics and Natural Intelligence*. 2011. Vol. 5, N 1. P. 1—17.
26. **Zatsman I.** Denotatum-Based Models of Knowledge Creation for Monitoring and Evaluating R & D Program Implementation // *Proc. of the 11th IEEE International Conference on Cognitive Informatics & Cognitive Computing*. Los Alamitos, CA: IEEE Computer Society Press. 2012. P. 27—34.
27. **Zatsman I.** Tracing Emerging Meanings by Computer: Semiotic Framework // *Proc. of the 13th European Conference on Knowledge Management*. 2012. Vol. 2. P. 1298—1307.
28. **Computational Science: Ensuring America's Competitiveness.** Report to the President. Arlington, VA: National Coordination Office for Information Technology Research and Development, 2005.
29. **Коллин К. К.** Новая стратегическая компьютерная инициатива США и задачи России в области развития фундаментальной информатики // *Информационные технологии*. 2006. № 7. С. 2—5.
30. **Аннотация и рубрикатор** темы "Фундаментальные основы конвергентных технологий" в конкурсе РФФИ ориентированных фундаментальных исследований по актуальным междисциплинарным темам 2011 года. URL: http://www.rfbr.ru/rffi/getimage/20.Фундаментальные_основы_конвергентных_технологий.pdf?objectId=31173.
31. **Мысляков Д.** Практика широкого формата (интервью с П. Пашиным) // *Газета "Поиск"*. 2012. № 12 (1190). С. 6—7.
32. **Список проектов** конкурса "ОФИ-М-2011", получивших финансовую поддержку РФФИ. URL: <http://www.rfbr.ru/rffi/getimage?objectId=38202>.
33. **The National Science Foundation's FY 2003 funding request to Congress.** URL: <http://www.nsf.gov/about/budget/fy2003/goals.htm>.
34. **Roco M., Bainbridge W. S.** (Eds). *Converging technologies for improving human performance: nanotechnology, biotechnology, information technology and cognitive science*. U.S. National Science Foundation. Dordrecht, The Netherlands: Kluwer Academic Publishers, 2003. 482 p.
35. **Nordmann A.** (Ed). *Converging technologies — shaping the future of European societies. A report from the high level expert group on "foresighting the new technology wave"*. Brussels: European Communities, 2004. 68 p.
36. **Коллин К. К.** О структуре научных исследований по комплексной проблеме "Информатика" // *Сб. науч. тр. "Социальная информатика"*. М.: ВКШ при ЦК ВЛКСМ, 1990. С. 19—33.
37. **Коллин К. К.** Эволюция информатики и проблемы формирования нового комплекса наук об информации // *Научно-техническая информация. Сер. 1*. 1995. № 5. С. 1—7.
38. **Коллин К. К.** Становление информатики как фундаментальной науки и комплексной научной проблемы // *Системы и средства информатики: Спец. вып. "Научно-методологические проблемы информатики"*. М.: ИПИ РАН, 2006. С. 7—58.
39. **Зацман И. М.** Концептуальный поиск и качество информации. М.: Наука, 2003. 271 с.
40. **Зацман И. М.** Семиотические основания и элементарные технологии информатики // *Информационные технологии*. 2005. № 7. С. 18—31.
41. **Зацман И. М., Косарик В. В., Курчавова О. А.** Задачи представления личностных и коллективных концептов в цифровой среде // *Информатика и ее применение*. 2008. Т. 2. Вып. 3. С. 54—69.
42. **Зацман И. М.** Семиотическая модель взаимосвязей концептов, информационных объектов и компьютерных кодов // *Информатика и ее применение*. 2009. Т. 3. Вып. 2. С. 65—81.
43. **Зацман И. М.** Нестационарная семиотическая модель компьютерного кодирования концептов, информационных объектов и денотатов // *Информатика и ее применение*. 2009. Т. 3. Вып. 4. С. 87—101.
44. **Гак В. Г.** Асимметрия // *Большой энциклопедический словарь "Языкознание"*. М.: Большая российская энциклопедия, 1998. С. 47.
45. **Shepard R. N.** Perceptual-Cognitive Universals as Reflections of the World // *Behavioral and Brain Sciences*. 2001. Vol. 24, N 3. P. 581—601.

МОДЕЛИРОВАНИЕ И ОПТИМИЗАЦИЯ MODELING AND OPTIMIZATION

УДК 004.942

А. Н. Божко, канд. техн. наук, доц., МГТУ им. Н. Э. Баумана,
e-mail: abozhko@inbox.ru,

Б. С. Сюсюкалов, гл. специалист, ОТП банк

Математические модели базирования и избыточности в механических системах

Работа посвящена моделированию позиционных механических связей в технических системах. Предлагается гиперграфовая модель, которая адекватно описывает условия базирования деталей в составе конструкции. В терминах данной модели можно выразить множество проектных ситуаций, возникающих в процессе технологической подготовки сборочного производства. Рассматривается эффект перебазирования, который на языке гиперграфов представляется в виде избыточных ребер гиперграфа. Предлагается эвристический алгоритм удаления лишних ребер, продуцирующий неизбыточную структуру с максимальной свободой собираемости.

Ключевые слова: сборка, механическая связь, сопряжение, базирование, граф, гиперграф, стягивание, узел, последовательность сборки, сборочная единица, стягивание вершин, перебазирование

A. N. Bozhko, B. S. Syusyukalov

Mathematical Models for Basing and Redundancy in Mechanical Systems

The paper deals with modeling of positional mechanical linkages in technical systems. We consider a hypergraph model that adequately describes the conditions basing components in the construction. In terms of this model can be expressed many design situations arising in the process of technological preparation of assembly production. We consider the effect of overbasing, which modeling is represented as a redundant relationships hypergraph. The article offers a heuristic algorithm of removing excess edges of hypergraph, which produces not excessive structure with the maximum freedom assembly.

Keywords: assembly, mechanical connection, coupling, basing, graph, hypergraph, contraction of the vertices, assembly sequence, assembly unit, the contraction of vertices, overbasing

Введение

Одной из важнейших задач технологической подготовки машиностроительного производства является проектирование рациональной последовательности сборки изделия. Данное проектное решение фиксируется в технологическом документе, который называется схемой сборки (СС). Схема сборки определяет способ декомпозиции изделия на независимо собираемые сборочные единицы (СЕ), вертикальную соподчиненность сборочных единиц и деталей, а также последовательность сборки всех составляющих изделия. В классической монографии, которая специалистами по сборочному делу считается самым авторитетным источником по технологии сборки среди публикаций на русском языке, утверждается: "...Технологическая схема сборки является основным оперативным докумен-

том, по которому персонал сборочного цеха знакомится с последовательностью сборки новой машины, организует выполнение сборочного процесса, проводит комплектование машины, подачу сборочных единиц и деталей в надлежащей последовательности, ведет учет, планирует производство, вносит усовершенствования в конструкцию, технологический процесс и организацию производства машины" [11].

Последовательность сборки изделия является конструктивно реализуемой, если для каждого устанавливаемого элемента (детали или СЕ) выполняются условия базирования и ограничения, накладываемые геометрической разрешимостью. Под *базированием* понимается взаимная координация деталей, которая достигается фиксацией положения устанавливаемой детали относительно собранного фрагмента изделия. Координатные системы, которые используются для определения положения и

установки деталей, называются *конструкторскими базами*. Они обычно реализуются в виде участков поверхностей, контактных точек, осей симметрии, искусственной разметки и других явных или скрытых конструктивных элементов [12].

Геометрической разрешимостью (геометрическим доступом) называется ситуация, когда собранная конфигурация не содержит геометрических препятствий, делающих невозможным установку детали в выбранное положение. Базирование и геометрический доступ — это необходимые условия существования последовательности сборки любого узла или изделия. Их нарушение автоматически исключает альтернативу из дальнейшего рассмотрения. Множество последовательностей сборки, для которых данные условия выполняются, образует "операционное поле" для выбора лучшей альтернативы по различным дополнительным критериям: технологическим, производственным или экономическим.

С точки зрения формализации и компьютерного моделирования условия базирования можно считать первичными. Именно они резко ограничивают число вариантов, которые надлежит проверить на геометрическую разрешимость и технологическую реализуемость. Без учета базирования множество допустимых вариантов растет экспоненциально в зависимости от числа деталей изделия ($n!$, если быть точным), поскольку постановка задачи геометрической разрешимости допускает любую тестовую конфигурацию элементов трехмерной сцены (даже несвязную).

В данной работе ограничимся обсуждением условий базирования и способов формализации этой важнейшей структурной и конструктивной характеристики изделий.

Гиперграфовая модель базирования

В технических системах базирование достигается реализацией механических связей (соединений и сопряжений), которые закладываются на стадии конструирования, а реализуются в сборочном производстве. Данные связи доставляют изделию механическую связность, а также функциональную и геометрическую определенность. В работах по автоматизации проектирования были предложены многочисленные структурные модели, предназначенные для описания механических связей изделия [1—5]. Если отвлечься от различий в наименованиях и технике генерации, то большинство моделей представляют собой парафраз графа механических связей.

Этот носитель проектной информации задается следующим образом: детали представляются в виде вершин графа, а соединения и сопряжения описываются ребрами. Две вершины соединены ребром тогда и только тогда, когда на соответствующие детали наложена механическая связь. На рис. 1 показана конструкция крепления промежуточного вала цилиндрического редуктора и граф механических связей этой конструкции. На этом рисунке верши-

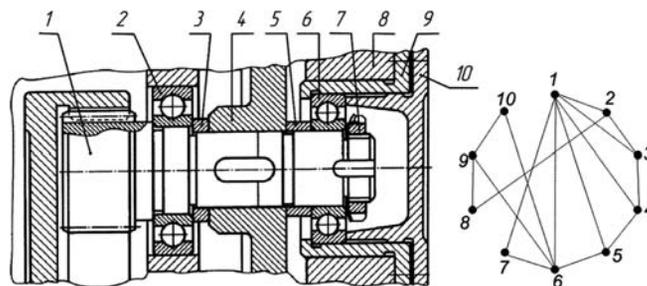


Рис. 1. Промежуточный вал цилиндрического редуктора внутреннего зацепления

ны графа описывают детали, а ребра представляют соединения и сопряжения изделия.

Практика показала, что данная структурная модель не позволяет точно описать условия базирования во всех проектных ситуациях. Главная причина неполной адекватности заключается в следующем. Граф механических связей задает бинарное отношение, тогда как отношение базирования в общем случае является многоместным. Действительно, существуют конструктивные решения, когда для точной координации детали требуется реализовать несколько связей одновременно, т. е. полный комплект баз включает в себя фиксированную группировку, состоящую из нескольких деталей. В качестве простейшего примера можно привести способ установки подшипника на вал, показанный на рис. 1. Внутреннее кольцо 6 этой сборочной единицы позиционируется относительно торца втулки 5 и наружной поверхности вала 1. Две эти связи, реализованные совместно, определяют положение подшипника, но ни одна из них в отдельности таким свойством не обладает. Для описания многоместных отношений и отношений переменной местности удобно использовать аппарат гиперграфов и гиперсетей [13].

Определение 1. Подмножество деталей будем называть геометрически определенным в составе изделия, если положение всех входящих в него элементов задано относительно друг друга согласно комплексу сборочных чертежей и техническим требованиям.

Геометрическая определенность не предполагает относительной или абсолютной неподвижности деталей. Элементы кинематических пар и детали механизмов могут обладать различными степенями подвижности и, вместе с тем, составлять геометрически определенные конфигурации в составе изделия. В частности, геометрически определенными в смысле определения 1 будут все изделие, любая его сборочная единица, а также все комплектующие и детали, входящие в состав конструкции.

Определение 2. Геометрически определенное подмножество деталей является минимальным, если исключение любой детали нарушает свойство взаимной координации составных частей.

Определение 3. Машине или механическому прибору поставим в соответствие гиперграф $H = (X, R, W)$, где множество вершин $X = \{x_i\}_{i=1}^n$ описывает элементы технической системы (детали и комплектующие), множество гиперребер $R = \{r_j\}_{j=1}^m$ представляет минимальные геометрически определенные группировки элементов, а $W: R \rightarrow 2^X$ — инцидентор, который связывает гиперребра с входящими в него вершинами. Будем называть эту модель *гиперграфом механических связей* [7].

На рис. 2, а приведен гиперграф механических связей конструкции крепления промежуточного вала.

Существуют проектные ситуации, требующие совместного анализа механических связей конструкции и отношения базирования между ее элементами. Такая необходимость, в частности, возникает в процессе синтеза рациональной размерной схемы изделия. Адекватной математической моделью изделия, доставляющей всю необходимую для такого типа задач проектную информацию, будет гиперсеть. Она представляет собой композицию графа и гиперграфа механических связей на общем множестве вершин [10]. На рис. 2, б показан пример гиперсети для изделия, представленного на рис. 1.

Гиперсеть — это содержательная, но чрезвычайно громоздкая модель. Ее генерация и анализ становятся затруднительными даже для изделий со сравнительно небольшим числом деталей (примерно от 20 и более). Вся проектную информацию, необходимую для формализации отношения базирования при сборке, дает гиперграф механических связей, поэтому далее будем обсуждать только эту модель.

С позиций системного анализа последовательность сборки изделия можно рассматривать как последовательность реализации механических связей, заложенных на стадии конструирования и доставляющих деталям определенность положения в составе конструкции. Если отвлечься от технологического содержания сборочных операций, то оказывается, что установка детали заключается в ре-

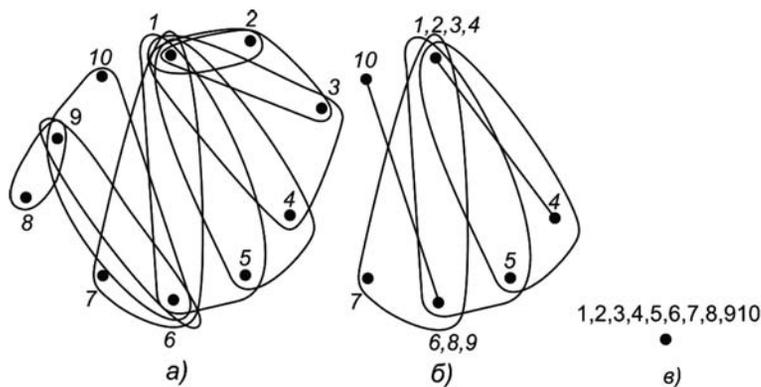


Рис. 3. Гиперграфы, описывающие стартовое (а), промежуточное (б) и финальное (в) состояния сборочного процесса

ализации базовых инцидентор, которые связывают ее с собранным фрагментом изделия. Это действие можно описать как стягивание гиперребра, которое заключается в отождествлении всех инцидентных этому ребру вершин и удалении реализованной связи. Каждый отдельный "акт сборки" заключается в установке детали (СЕ) на собранный фрагмент изделия, который в терминах гиперграфовой модели представляется одной вершиной. Это значит, что гиперребра высоких степеней в процессе стягиваний "схлопываются" и превращаются в гиперребра степени 2, как показано на рис. 3, б. Стартовое состояние процесса сборки задается исходным гиперграфом (рис. 3, а), финальное состояние — одновершинным гиперграфом (рис. 3, в), в котором реализованы все связи.

Сказанное дает основание для следующего определения.

Определение 4. Гиперграф $H = (X, R, W)$ назовем стягиваемым (s -гиперграфом), если существует последовательность $P(H) = (H_0, \dots, H_{n-1})$, для элементов которой выполняются следующие требования [9, 10]:

- 1) $H_0 = H$;
- 2) H_{n-1} представляет собой одновершинный гиперграф;
- 3) для всех H_j и $H_{j+1} \in P(H)$, $j = \overline{0, n-2}$, справедливо соотношение $|R_j| - 1 = |R_{j+1}|$;
- 4) каждый элемент последовательности H_{j+1} получается из предыдущего H_j стягиванием ребра кратности 2, $j = \overline{0, n-2}$. Будем называть такое стягивание нормальным.

Значительный практический и научный интерес представляют формулировки необходимых и достаточных условий стягиваемости гиперграфов, задающих механические структуры изделий различных классов. В работах [7–10] приводятся несколько утверждений, относящихся к стягиваемости гиперграфов механических связей изделий, элементы которых ведут себя как абсолютно твердые тела в процессе сборки. Приведем основной результат.

Утверждение. Пусть существует последовательность нормальных стягиваний гиперграфа

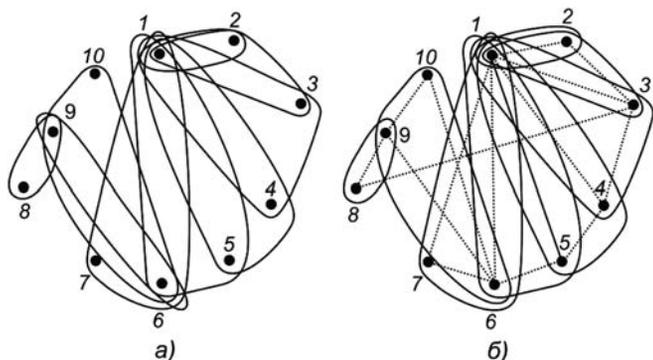


Рис. 2. Гиперграф (а) и гиперсеть (б), описывающие структуру механических связей

$H = (X, R, W)$, переводящая H в одновершинный гиперграф без петель. Тогда выполняются условия:

- 1) среди ребер H существует по крайней мере одно ребро степени 2;
- 2) гиперграф является связным;
- 3) число вершин $|X|$ и ребер $|R|$ гиперграфа H удовлетворяют линейному ограничению $|X| = |R| + 1$.

Пропуская достаточно громоздкое формальное доказательство этого важного утверждения, приведем содержательные аргументы в пользу его справедливости.

Необходимость первого условия доказывает первая операция нормального стягивания, которая применяется к вершинам исходного гиперграфа. Связность следует из существования последовательности нормальных стягиваний, переводящей гиперграф в точку. Рассмотрим более подробно третье условие утверждения.

Если $X > R + 1$, то гиперграф превращается в плохо скоординированную структуру, поскольку его наличных связей оказывается недостаточно для фиксации положения каждой вершины в смысле определений 1–4. Если $X \gg R + 1$, то гиперграф может стать настолько разреженным, что потеряет формальную связность.

Обсудим очень важный с прикладной и теоретической точек зрения случай, когда $X < R + 1$. В процессе сборки любого изделия координация и закрепление первой детали (СЕ) выполняется посредством внешнего приспособления, поэтому она не нуждается во внутреннем базировании. Для установки любого элемента (детали или СЕ) требуется только один комплект конструкторских баз, который определяет его положение относительно собранного фрагмента. Это значит, что число полных комплектов баз должно быть на единицу меньше общего числа деталей изделия. Нарушение этого требования влечет за собой эффект, который в конструкторской и технологической практике называется *перебазированием*.

В процессе сборки перебазирование приводит к необходимости установки некоторой детали или сборочной единицы с одновременной координацией по двум или большему числу полных комплектов конструкторских баз. Такая ситуация избыточной скоординированности является крайне нежелательной. Ее можно устранить посредством дополнительных технологических операций или локальным изменением конструкции, с очевидными негативными последствиями для жизненного цикла изделия.

Избыточность и перебазирование

Эффект перебазирования легко описать в гиперграфовых терминах. Если $X < R + 1$, то в последовательности нормальных стягиваний гиперграфа H неизбежно возникнет ситуация, подобная изображенной на рис. 4, а. На этом рисунке серый овал означает стянутый в одну вершину фрагмент гиперграфа, а вершина t символизирует устанавливаемый элемент.

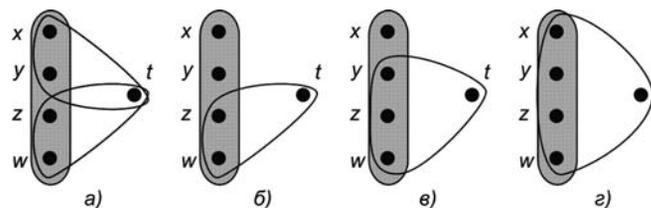


Рис. 4. Гиперграфовая модель эффекта перебазирования и варианты его устранения

Отказ от перебазирования заключается в нормализации операции стягивания. Это значит, что связь, которая нарушает это условие, должна быть отброшена (рис. 4, б), реконструирована (рис. 4, в) либо включена в состав валидной связи (рис. 4, г).

Практика технологической подготовки производства показывает, что в конструкциях машин средней и высокой сложности, с большим числом деталей и механических связей достаточно часто встречаются избыточные механические связи, вызывающие эффект перебазирования. Тому имеется множество причин, например, погрешности проектирования, толкование кинематических связей как позиционных, условности изображения конструктивных элементов на сборочных чертежах и противоречия в ЕСКД и пр. Не свободно от ошибок данного типа и компьютерное геометрическое моделирование и автоматизированное проектирование, поскольку современные САД-системы не имеют средств анализа механических структур на корректность и избыточность.

Рассмотрим следующую постановку задачи. Пусть гиперграф $H = (X, R, W)$, описывающий механические связи некоторого изделия, содержит избыточные ребра, т. е. для H выполняется неравенство $X < R + 1$. Требуется устранить избыточность и привести гиперграф к линейному ограничению 3 определения 4 удалением $R - X + 1$ ребер, не нарушив при этом условий 1 и 2.

Определение 5. Операцию удаления избыточных ребер в гиперграфе H , в результате которой он становится s -гиперграфом, будем называть *линеаризацией*.

Сочетание различных критериев оптимальности и ограничений на удаление ребер позволяет сформулировать несколько важных с практической и теоретической точек зрения разновидностей задачи линеаризации. Перечислим основные варианты ограничений:

- 1) разрешается удалять любые ребра гиперграфа;
- 2) некоторое подмножество ребер объявляется защищенным от удаления;
- 3) на ребрах гиперграфа определена некоторая нечеткая функция или задано вероятностное распределение, описывающие предпочтительные варианты удаления;
- 4) условия корректной редукции описываются логическими зависимостями. Например, удаление ребра x требует обязательного исключения ребра y , или ребра x и y не могут быть удалены одновременно.

Реальные проектные ситуации, возникающие в технологической практике, дают основания для формулировки различных критериев оптимальности в задаче линейаризации. Например, можно искать такой вариант удаления ребер, который гарантирует существование заданной последовательности сборки или схемы сборочного состава. Интересны с практической и теоретической точек зрения варианты линейаризации, которые продуцируют гиперграфовые структуры максимальной собираемости или расчленяемости и др.

Рассмотрим важный частный случай поставленной задачи. Задан гиперграф H , содержащий избыточные ребра. Требуется получить линейаризованный гиперграф H^* , допускающий максимальное число нормальных стягиваний. В операции разрешается удалять любые ребра, исключение которых не нарушает условия 1–2 определения 5.

Действительно, число различных нормальных стягиваний может служить косвенным измерителем "свободы" при сборке технической системы. Чем это число больше, тем менее жесткими являются условия базирования. Это значит, что последующий выбор по технологическим или экономическим критериям будет располагать более богатым набором альтернатив.

На рис. 5, *а* изображен избыточный гиперграф с четырьмя вершинами и пятью ребрами. Два варианта его линейаризации представлены на рис. 5, *б* и рис. 5, *в*. Оба гиперграфа являются s -гиперграфами, но допускают различное число нормальных стягиваний: у варианта *б* — их пять, вариант *в* имеет всего лишь одну такую последовательность.

В работе предлагается эвристический алгоритм, позволяющий получить субоптимальное решение задачи для вариантов не очень высокой размерности. Алгоритм основан на следующем предположении. Наибольшее комбинаторное разнообразие (по числу различных вариантов нормальных стягиваний) имеют гиперграфы, структура которых обладает максимальной центральностью (минимальным радиусом) [13].

Справедливость этого предположения очевидна для деревьев — важного частного случая s -гиперграфов. Рассмотрим два полярных варианта: звезду (рис. 6, *а*) и цепь (рис. 6, *б*), каждая из которых состоит из n вершин. Радиус звезды равен 1, радиус цепи — $r \geq \lfloor \frac{n}{2} \rfloor$, где n — число вершин. Каждое из этих деревьев содержит по $n - 1$ ребер. Подсчитаем число нормальных стягиваний, которые будем рассматривать как последовательность реализации ребер. Легко видеть, что звезда, обладающая меньшим радиусом, имеет $(n - 1)!$ таких последовательностей, а цепь — всего лишь $2(n - 1)$.

Данная тенденция (увеличения числа нормальных стягиваний с ростом центральности структуры) остается справедливой и для гиперграфов. На рис. 7 приведены примеры гиперграфов с максимальной

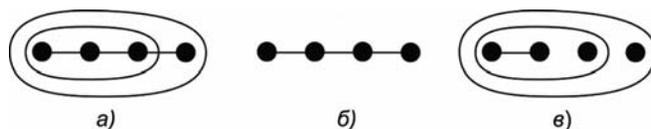


Рис. 5. Избыточный гиперграф (*а*) и варианты его линейаризации (*б*, *в*)

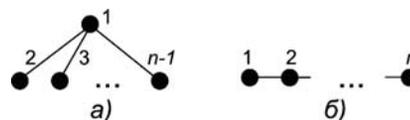


Рис. 6. Простейшие стягиваемые структуры: звезда (*а*) и цепь (*б*)

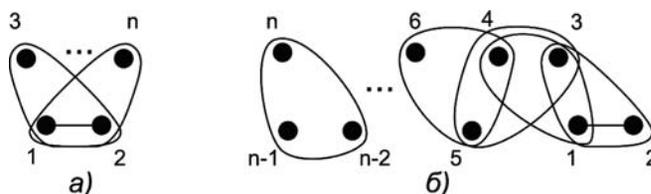


Рис. 7. Примеры гиперграфов с максимальной (*а*) и минимальной (*б*) центральностью

(рис. 7, *а*) и минимальной центральностью (рис. 7, *б*). Для первой структуры число нормальных стягиваний равно $(n - 2)!$, тогда как для второй оно равно всего лишь 1.

Алгоритм линейаризации

Основная идея алгоритма заключается в следующем. Дан гиперграф, содержащий избыточные ребра. Найдем в нем максимальный по числу вершин подграф, обладающий тремя свойствами:

- 1) все ребра подграфа имеют степень два (т. е. являются простыми ребрами);
- 2) подграф является ациклическим (в общем случае лесом);
- 3) подграф обладает минимальным радиусом (максимальной центральностью) среди всех возможных подграфов данного вида.

Преобразуем исходный гиперграф стягиванием каждой компоненты леса в одну мультивершину. Будем повторять описанную операцию до тех пор, пока гиперграф не превратится в вершину без петель. Решение задачи образуют множество вершин исходного гиперграфа и все ребра, которые использовались в данной процедуре. Неиспользуемые ребра считаются избыточными и удаляются.

Формальное описание алгоритма.

Дан гиперграф $H = (X, R, W)$. Попытаемся выделить подгиперграф G гиперграфа H , обладающий наибольшим числом нормальных стягиваний.

Шаг 1. Положим $G = \emptyset$.

Шаг 2. Выделим в H всевозможные ребра степени два и инцидентные им вершины. В выделенном подграфе найдем остовный лес, компоненты которого имеют минимальный радиус. Занесем эти ребра вместе с инцидентными вершинами в G .

Шаг 3. Преобразуем H стягиванием каждой компоненты (дерева) леса в одну мультивершину.

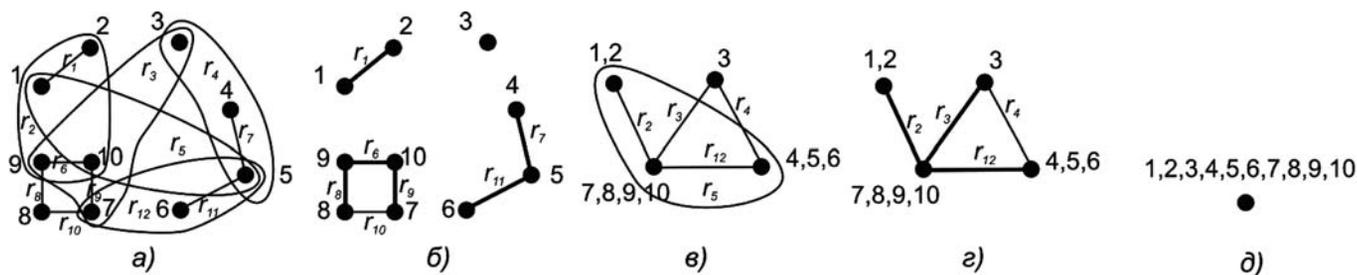


Рис. 8. Операции алгоритма по удалению избыточных ребер

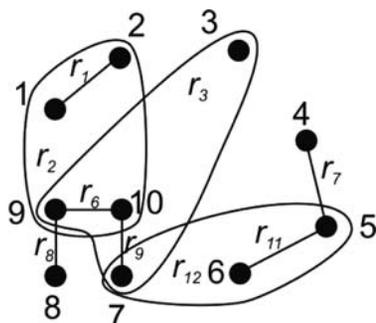


Рис. 9. Стыгиваемый гиперграф, полученный в результате работы алгоритма удаления избыточности

Шаг 4. Проверка условий останова. Содержит ли H ребра степени два?

Да: выполнить операцию 2.

Нет: выполнить операцию 5.

Шаг 5. Гиперграф H состоит из одной вершины?

Да: подгиперграф G является s -гиперграфом и представляет собой решение поставленной задачи.

Нет: гиперграф $H = (X, R, W)$ не содержит подгиперграфа G , допускающего нормальное стягивание. Задача не имеет решения.

На рис. 8 приведен пример, иллюстрирующий операции алгоритма удаления избыточности. На рис. 8, а показан гиперграф H в исходном состоянии. Операции поиска максимального подграфа с ребрами степени два и выделения леса минимального радиуса иллюстрирует рис 8, б. Ребра леса изображены жирными линиями. На рис. 8, в показан результат стягивания ребер леса, а на рис. 8, г — начало второй итерации алгоритма. Результат преобразования представлен на рис. 9.

Заключение

Отношение базирования, определенное на множестве деталей изделия, — это важнейшая конструктивная характеристика, решающим образом влияющая на способность машины или прибора разделяться на сборочные единицы и собираться в заданной последовательности. В общем случае это отношение является многоместным. Его свойства достаточно точно описывает гиперграф механических связей. На языке данной математической структуры можно представить множество проектных решений, возникающих на стадии технологической подготовки сборочного производства, в ча-

стности перебазирования. Пребазирование — это ситуация избыточной скоординированности деталей или сборочных единиц, которая приводит к необходимости установки конструктивных элементов с ориентацией по нескольким полным комплектам конструкторских баз одновременно. Это явление приводит к появлению неразрешимых размерных цепей, нарушает технологические инструкции и вносит возмущения в организацию производственного процесса. Математическим описанием перебазирования служит гиперграф механических связей с избыточными ребрами. В работе предлагается алгоритм удаления избыточных ребер, приводящий к структуре с максимальной свободой собираемости. Множество последовательностей сборки, реализуемых по условию базирования, представляет собой исходное множество альтернатив в задачах выбора рациональных последовательностей по дополнительным критериям и решающим правилам: технологическим, точностным, экономическим и пр.

Список литературы

1. Baldwin D. F., Abell T. E., Lui M. M., De Fazio T. L., Whitney D. E. An integrated computer aid for generating and evaluating assembly sequences for mechanical products // IEEE Transactions on Robotics and Automation. 1991. N 7. P. 78—94.
2. Whitney D. E. Mechanical Assemblies: Their Design, Manufacture, and Role in Product Development. Oxford University Press, 2004. 518 p.
3. De Fazio T. L., Whitney D. E. Simplified generation of all mechanical assembly sequences // IEEE Journal of Robotics and Automation. 1987. N 3 (6). P. 640—658.
4. Antonsson E. K., Cagan J. Formal Engineering Design Synthesis. Cambridge University Press, 2005. 500 p.
5. Homem De Mello L. S., Sanderson A. C. A correct and complete algorithm for the generation of mechanical assembly sequences // IEEE Transactions on Robotics and Automation. 1991. N 7. P. 228—240.
6. Айгнер М. Комбинаторная теория. М.: Мир, 1982. 558 с.
7. Божко А. Н., Бетин Е. А. Анализ стягиваемости гиперграфов // Информационные технологии. 2005. № 5. С. 6—12.
8. Божко А. Н. Моделирование механических связей изделия. Условия стягиваемости // Электронное научно-техническое издание "Наука и образование". 2011. № 5.
9. Божко А. Н. Моделирование механических связей изделия // Электронное научно-техническое издание "Наука и образование". 2011. № 3.
10. Божко А. Н. Моделирование позиционных связей в механических системах // Информационные технологии. 2012. № 10. С. 27—33.
11. Новиков М. П. Основы технологии сборки машин и механизмов. М.: Машиностроение, 1980. 298 с.
12. Сборка и монтаж изделий машиностроения: справочник в 2-х т. / Под ред. В. С. Корсакова, В. К. Замятина. М.: Машиностроение, 1983. 480 с., 360 с.
13. Харари Ф. Теория графов / Под ред. Г. П. Гаврилова. Изд. 2-е. М.: Едиториал УРСС, 2003. 296 с.

А. В. Васюков, ассистент, **А. С. Ермаков**, аспирант, **И. Б. Петров**, чл.-корр. РАН, зав. каф.,
А. П. Потапов, доц., **А. В. Фаворская**, аспирант, e-mail: aleanera@yandex.ru, **А. В. Шевцов**, аспирант,
 Московский физико-технический институт, e-mail: petrov@mipt.ru

Комбинирование сеточно-характеристического метода и метода сглаженных частиц в задачах компьютерного моделирования упругопластических тел

Рассматривается комбинированный метод моделирования упругопластических тел, призванный объединить преимущества двух методов: сглаженных частиц и сеточно-характеристического. Имеются два семейства методов, являющихся оптимальными для двух разных групп задач. Однако реальная задача часто может оказаться смешанной, что потребует идти на существенный компромисс при выборе численного метода. Для решения таких задач разрабатывается комбинированный численный метод GCM-SPH, объединяющий преимущества и частично устраняющий недостатки двух базовых методов.

Ключевые слова: сеточно-характеристический метод, метод сглаженных частиц, численное моделирование, неструктурированные сетки, комбинированный численный метод GCM-SPH, высокопроизводительные вычислительные системы, пространственные динамические задачи

A. V. Vasyukov, A. S. Ermakov, I. B. Petrov, A. P. Potapov, A. V. Favorskaya, A. V. Shevtsov

Combining the Grid-Characteristic Method and the Method of Smoothed Particle for Computer Modeling of Elastoplastic Bodies

In this paper the combined method for the modeling of elastoplastic bodies is considered. This method is designed to combine the advantages of two methods: the SPH method and grid-characteristic method. There are two families of methods, which are optimal for the two different groups of tasks. However, the real problem may often be mixed, which would require to make substantial compromises during the choosing a numerical method. In order to solve such problems the combined numerical method GCM-SPH is developed, connecting the advantages and partially eliminating the disadvantages of two basic techniques.

Keywords: grid-characteristic method, method of smoothed particle hydrodynamics, numerical simulation, unstructured grid, the combined method GCM-SPH, high-performance computing systems, spatial dynamic problems

Введение

При моделировании упругих и упругопластических взаимодействий на протяжении значительного времени широкое распространение имеют сеточные методы, в частности, метод конечных объемов [1] и сеточно-характеристический [1–5]. Задачи ударного взаимодействия деформируемых твердых тел рассматриваются в работах [6–14].

В последнем десятилетии XX века с ростом вычислительных мощностей стали приобретать распространение адаптации метода сглаженных частиц (SPH) к моделированию упругопластических тел [15–18], предложенного первоначально для моделирования звезд и претерпевшего затем ряд усовершенствований [19, 20]. Также для решения задач в системе "ударник—мишень" используются адаптивные сетки [21].

Постановка задачи

Суть комбинированного метода GCM-SPH [22, 23] состоит в моделировании частей тел, подверженных существенным деформациям, методом сглаженных частиц [24], в то время как остальные части тел моделируются сеточно-характеристическим методом [2–5]. Более того, путем соответствующего разделения удается добиться отсутствия пластических деформаций в области, моделируемой сеткой, и вблизи границы раздела методов. Поэтому можно ограничиться решением системы уравнений, характеризующей состояние бесконечно малого объема сплошной линейно-упругой среды [25] с использованием сеточно-характеристического метода, в то время как метод сглаженных частиц используется для решения аналогичной системы линейной упругости [25] в сочетании с критерием пластического течения Мизеса [26].

Состояние бесконечно малого объема сплошной линейно-упругой среды подчиняется следующим уравнениям [25]:

$$\rho \partial_t \mathbf{v} = (\nabla \cdot \mathbf{T})^T, \quad (1)$$

$$\partial_t \mathbf{T} = \lambda (\nabla \cdot \mathbf{v}) \mathbf{I} + \mu (\nabla \otimes \mathbf{v} + (\nabla \otimes \mathbf{v})^T). \quad (2)$$

Уравнение (1) является локальным уравнением движения. В нем ρ — плотность материала, \mathbf{v} — скорость движения, \mathbf{T} — тензор напряжений Коши, являющийся симметричным в силу закона парности касательных напряжений. Уравнение (2) выводится из закона Гука путем дифференцирования по времени. В нем λ , μ — параметры Ляме, определяющие свойства упругого материала.

В формулах (1), (2) использованы следующие обозначения:

\mathbf{I} — единичный тензор второго ранга;

$\partial_t \mathbf{a} \equiv \frac{\partial \mathbf{a}}{\partial t}$ — частная производная поля \mathbf{a} по t ;

$\mathbf{a} \otimes \mathbf{b}$ — тензорное произведение векторов \mathbf{a} и \mathbf{b} , $(\mathbf{a} \otimes \mathbf{b})^{ij} = a^i b^j$.

Поскольку для обоих базовых методов решение в каждой точке пространства на каждом шаге по времени зависит только от решения в близлежащей области на предыдущем шаге, вдали от границы раздела двух методов каждый из них работает независимо. Поэтому наибольший интерес представляет организация взаимодействия двух методов вдоль границы раздела.

Как известно, основой сеточно-характеристического метода является перенос решения между временными слоями вдоль характеристик [2—5], а конкретный способ вычисления значения в точке пересечения характеристики с предыдущим временным слоем не играет роли. Это позволяет реализовать учет решения в области с частицами при поиске решения в области сетки, граничащей с частицами: при попадании характеристики в область с частицами интерполяция компонент скорости и тензора напряжений проводится с использованием того же ядра сглаживания, которое используется в реализованной вариации метода сглаженных частиц.

Несколько сложнее оказывается организация учета решения в области с сеткой, при поиске решения в области с частицами, граничащей с сеткой. Суть метода сглаженных частиц [24] состоит в вычислении временных производных компонент решения, аналитически выраженных через численное решение в соседних частицах, и последующем явном интегрировании по времени. Для решения этой проблемы используются "сеточные" частицы — тонкий слой особых частиц, расположенных вдоль границы раздела методов внутри области с сеткой. В отличие от обычных частиц решение в этих частицах, а также их смещение за шаг, не вычисляется обычным методом [24], а интерполируется по сетке [2].

Максимальный шаг по времени для метода в целом определяется как минимум из максимально возможных шагов для каждого из методов.

Используемые обозначения

Пусть рассматриваемая область интегрирования разделена на две части. В правой части применяется сеточно-характеристический метод. Далее будем называть ее GCM-областью. В левой части применяется гибридный метод сглаженных частиц. Далее будем называть ее SPH-областью.

Обозначим $\mathbf{q}_{\text{GCM}}(t)$ вектор-функцию, описывающую состояние вещества в GCM-области. При этом $\mathbf{q}_{\text{GCM}}(t)$ характеризует совокупность значений величин, описывающих состояние вещества в GCM-области во всех узлах используемой тетраэдральной сетки на данном временном слое t . Обозначим тетраэдральную сетку, используемую в GCM-области на данном временном слое t , как $\text{Mesh}_{\text{GCM}}(t)$.

Обозначим $\mathbf{q}_{\text{SPH}}(t)$ вектор-функцию, описывающую состояние вещества в SPH-области. При этом $\mathbf{q}_{\text{SPH}}(t)$ характеризует совокупность значений величин, описывающих состояние вещества в SPH-области во всех используемых частицах на данном временном слое t .

Вдоль границы между GCM-областью и SPH-областью в GCM-области вводятся два слоя вспомогательных мнимых частиц. При этом обозначим $\mathbf{q}_{\text{SPH}}(t, \text{im})$ ту часть совокупности значений $\mathbf{q}_{\text{SPH}}(t)$, которая характеризует состояние вещества в данных вспомогательных мнимых частицах. Двух слоев мнимых частиц достаточно для поиска недостающих соседей для частиц из SPH-области.

Обозначим $\tau_{\text{GCM}}(t)$ максимально возможный на данном временном слое t шаг интегрирования в GCM-области, а максимально возможный на данном временном слое t шаг интегрирования в SPH-области — как $\tau_{\text{SPH}}(t)$.

Обозначим \mathbf{N}_{GCM} оператор перехода на новый временной слой в GCM-области. Тогда справедливо

$$\mathbf{q}_{\text{GCM}}(t + \tau(t)) = \mathbf{N}_{\text{GCM}} \cdot \mathbf{q}_{\text{GCM}}(t).$$

Данный оператор \mathbf{N}_{GCM} определяет всю совокупность действий над значениями предыдущего временного слоя $\mathbf{q}_{\text{GCM}}(t)$ в соответствии с использованием сеточно-характеристического метода [3] на используемой тетраэдральной сетке $\text{Mesh}_{\text{GCM}}(t)$.

Обозначим \mathbf{N}_{SPH} оператор перехода на новый временной слой в SPH-области. Тогда справедливо

$$\mathbf{q}_{\text{SPH}}(t + \tau(t)) = \mathbf{N}_{\text{SPH}} \cdot \mathbf{q}_{\text{SPH}}(t).$$

Данный оператор \mathbf{N}_{SPH} определяет всю совокупность действий над значениями предыдущего временного слоя $\mathbf{q}_{\text{SPH}}(t)$ в соответствии с использованием гибридного метода сглаженных частиц [24]. При этом его действия разбиты на два этапа:

1. Предварительные вычисления в данной части области интегрирования с целью определить максимально возможный временной шаг:

$$\tilde{\mathbf{q}}_{\text{SPH}}(t) = \tilde{\mathbf{N}}_{\text{SPH}} \cdot \mathbf{q}_{\text{SPH}}(t),$$

где $\tilde{\mathbf{q}}_{\text{SPH}}$ — промежуточное значение вектор-функции, описывающей состояние вещества GCM-области;

$$\tau_{\text{SPH}}(t) = f(\tilde{\mathbf{q}}_{\text{SPH}}(t)).$$

2. Расчет функций, описывающих состояние вещества в SPH-области:

$$\mathbf{q}_{\text{SPH}}(t + \tau(t)) = \hat{\mathbf{N}}_{\text{SPH}} \cdot \tilde{\mathbf{q}}_{\text{SPH}}(t),$$

т. е. оператор \mathbf{N}_{SPH} представим в виде произведения двух вспомогательных операторов $\hat{\mathbf{N}}_{\text{SPH}}$ и $\tilde{\mathbf{N}}_{\text{SPH}}$:

$$\mathbf{N}_{\text{SPH}} = \hat{\mathbf{N}}_{\text{SPH}} \tilde{\mathbf{N}}_{\text{SPH}}.$$

Обозначим $\mathbf{I}_{\text{SPH} \leftarrow \text{GCM}}$ оператор интерполяции в мнимых частицах по значениям в тетраэдральной сетке:

$$\mathbf{q}_{\text{SPH}}(t, \text{im}) = \mathbf{I}_{\text{SPH} \leftarrow \text{GCM}} \cdot \mathbf{q}_{\text{GCM}}(t).$$

Можно представить данный оператор $\mathbf{I}_{\text{SPH} \leftarrow \text{GCM}}$ и так, что он действует на всю совокупность $\mathbf{q}_{\text{SPH}}(t)$, при этом изменения будут проведены им только в мнимых частицах $\mathbf{q}_{\text{SPH}}(t, \text{im})$:

$$\mathbf{q}_{\text{SPH}}(t) = \mathbf{I}_{\text{SPH} \leftarrow \text{GCM}} \cdot \mathbf{q}_{\text{GCM}}(t).$$

Обозначим $\mathbf{I}_{\text{GCM} \leftarrow \text{SPH}}$ оператор интерполяции в точках пересечения характеристик, выходящих из GCM-области в SPH-область:

$$\mathbf{q}_{\text{GCM}}(t) = \mathbf{I}_{\text{GCM} \leftarrow \text{SPH}} \cdot \mathbf{q}_{\text{SPH}}(t).$$

При этом оператор $\mathbf{I}_{\text{GCM} \leftarrow \text{SPH}}$ затрагивает $\mathbf{q}_{\text{GCM}}(t)$ только путем восполнения значений вдоль данных выходящих характеристик.

Здесь были введены следующие обозначения:

$\tau(t)$ — шаг интегрирования, вычисляемый на каждом временном слое как минимум между $\tau_{\text{GCM}}(t)$ и $\tau_{\text{SPH}}(t)$;

$\mathbf{N} \cdot \mathbf{q}(t)$ — действие оператора \mathbf{N} на временной слой t , описываемый вектор-функцией $\mathbf{q}(t)$.

Введем также оператор $\mathbf{O}(\eta^p)$ — "О большое":

$$\mathbf{O}(\eta^p) \cdot \mathbf{q}(\eta) = \mathbf{O}(\eta^p).$$

Система уравнений (1), (2) представляется в виде

$$\partial_t \mathbf{q} + \mathbf{A}_1 \partial_{\xi_1} \mathbf{q} + \mathbf{A}_2 \partial_{\xi_2} \mathbf{q} + \mathbf{A}_3 \partial_{\xi_3} \mathbf{q} = 0. \quad (3)$$

Здесь (ξ_1, ξ_2, ξ_3) — некая декартова система координат, в которой на данном временном слое решается данная система уравнений; $\mathbf{q} = \{v_1, v_2, v_3, t_{11}, t_{22}, t_{33}, t_{12}, t_{13}, t_{23}\}^T$ — вектор, составленный из компонент скорости \mathbf{v} и симметричного тензора напряжений \mathbf{T} ; матрицы $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3$ получены путем представления (1), (2) в матричной форме.

Описание комбинированного метода GCM-SPH

Приведем описание алгоритма действия метода GCM-SPH на одном временном шаге t с акцентом на описание расчета границы между GCM-областью и SPH-областью.

Шаг 1. Интерполяция функций, характеризующих состояние вещества в мнимых частицах, по значениям функций, характеризующих состояние вещества в тетраэдральной сетке для сеточно-характеристического метода.

$$\mathbf{q}_{\text{SPH}}(t, \text{im}) = \mathbf{I}_{\text{SPH} \leftarrow \text{GCM}} \cdot \mathbf{q}_{\text{GCM}}(t).$$

Шаг 2. Предварительные вычисления в SPH-области с целью определить максимально возможный временной шаг на данном временном слое t :

$$\tilde{\mathbf{q}}_{\text{SPH}}(t) = \tilde{\mathbf{N}}_{\text{SPH}} \cdot \mathbf{q}_{\text{SPH}}(t),$$

$$\tau_{\text{SPH}}(t) = f(\tilde{\mathbf{q}}_{\text{SPH}}(t)).$$

Шаг 3. Определение максимально возможного временного шага в тетраэдральной сетке для сеточно-характеристического метода и выбор минимального из двух:

$$\tau_{\text{GCM}}(t) = f(\text{Mesh}_{\text{GCM}}(t));$$

$$\tau(t) = \min(\tau_{\text{GCM}}(t), \tau_{\text{SPH}}(t)).$$

Шаг 4. Расчет функций, описывающих состояние вещества в GCM-области:

$$\mathbf{q}_{\text{GCM}}(t) = \mathbf{I}_{\text{GCM} \leftarrow \text{SPH}} \cdot \mathbf{q}_{\text{SPH}}(t);$$

$$\mathbf{q}_{\text{GCM}}(t + \tau(t)) = \mathbf{N}_{\text{GCM}} \cdot \mathbf{q}_{\text{GCM}}(t).$$

Шаг 5. Расчет функций, описывающих состояние вещества в SPH-области:

$$\mathbf{q}_{\text{SPH}}(t + \tau(t)) = \hat{\mathbf{N}}_{\text{SPH}} \cdot \tilde{\mathbf{q}}_{\text{SPH}}(t).$$

Шаг 6. Сохранение результатов интегрирования на данном временном слое.

Доказательство аппроксимации первого порядка

Строго говоря, порядок аппроксимации определяется для гладких решений [27], однако многочисленные численные эксперименты, проводимые методом сквозного счета, показали удовлетворительное соответствие расчетных и экспериментальных данных. Дело в том, что в рассматриваемых задачах численные решения всегда "размазываются" на несколько расчетных ячеек, поэтому более правильно говорить не о разрывах, а о больших градиентах численного решения. Это также соответствует и физике рассматриваемых процессов, поскольку в них отсутствуют такие явления, как, например, ударные волны в задачах сверхзвукового обтекания в атмосфере.

Пусть некий оператор \mathbf{N}_1 при переходе на следующий временной слой задает аппроксимацию решения первого порядка. Тогда для решения должно быть справедливо следующее разложение в ряд Тейлора:

$$\begin{aligned} \mathbf{q}(t + \tau, \xi_1, \xi_2, \xi_3) &= \mathbf{N}_1 \cdot \mathbf{q}(t, \xi_1, \xi_2, \xi_3) = \\ &= \mathbf{q}(t, \xi_1, \xi_2, \xi_3) + \tau \frac{\partial \mathbf{q}(t, \xi_1, \xi_2, \xi_3)}{\partial t} + \mathbf{O}(\tau^2). \end{aligned}$$

Запишем это разложение в форме операторов:

$$\mathbf{N}_1 = \mathbf{I} + \tau \left(\frac{\partial}{\partial t} \right)_0 + \mathbf{O}(\tau^2).$$

Так как система уравнений (3) выполняется, то должно быть справедливым следующее соотношение:

$$\left(\frac{\partial}{\partial t} \right)_0 = -\mathbf{A}_1 \left(\frac{\partial}{\partial \xi_1} \right)_0 - \mathbf{A}_2 \left(\frac{\partial}{\partial \xi_2} \right)_0 - \mathbf{A}_3 \left(\frac{\partial}{\partial \xi_3} \right)_0.$$

Пусть оператор \mathbf{N}_{GCM} задает аппроксимацию решения в GCM-области первого порядка точности по времени. Тогда для решения в GCM-области должно выполняться соответствующее разложение в ряд Тейлора. Распишем его в форме операторов:

$$\mathbf{N}_{\text{GCM}} = \mathbf{I} + \tau \left(\frac{\partial}{\partial t} \right)_{\text{GCM}} + \mathbf{O}(\tau^2).$$

Так как система уравнений (3) выполняется, то должно быть справедливым следующее соотношение:

$$\left(\frac{\partial}{\partial t} \right)_{\text{GCM}} = -\mathbf{A}_1 \left(\frac{\partial}{\partial \xi_1} \right)_{\text{GCM}} - \mathbf{A}_2 \left(\frac{\partial}{\partial \xi_2} \right)_{\text{GCM}} - \mathbf{A}_3 \left(\frac{\partial}{\partial \xi_3} \right)_{\text{GCM}}.$$

Здесь $\left(\frac{\partial}{\partial \eta} \right)_{\text{GCM}}$ — оператор дифференцирования по η функции $\mathbf{q}_{\text{GCM}}(t, \xi_1, \xi_2, \xi_3)$, $\eta = t, \xi_1, \xi_2, \xi_3$, т. е.

$$\mathbf{N}_{\text{GCM}} = \mathbf{I} - \tau \mathbf{A}_1 \left(\frac{\partial}{\partial \xi_1} \right)_{\text{GCM}} - \tau \mathbf{A}_2 \left(\frac{\partial}{\partial \xi_2} \right)_{\text{GCM}} - \tau \mathbf{A}_3 \left(\frac{\partial}{\partial \xi_3} \right)_{\text{GCM}} + \mathbf{O}(\tau^2).$$

При этом оператор $\mathbf{I}_{\text{GCM} \leftarrow \text{SPH}}$ представляет собой интерполяцию по частицам, и, следовательно, функция $\mathbf{I}_{\text{GCM} \leftarrow \text{SPH}} \cdot \mathbf{q}_{\text{SPH}}(t, \xi_1, \xi_2, \xi_3)$ кусочно-дифференцируема по координатам. Таким образом, справедливо:

$$\begin{aligned} \mathbf{N}_{\text{GCM}} \mathbf{I}_{\text{GCM} \leftarrow \text{SPH}} &= \mathbf{I}_{\text{GCM} \leftarrow \text{SPH}} - \\ &- \tau \mathbf{A}_1 \left(\frac{\partial}{\partial \xi_1} \right)_{\text{GCM}} \mathbf{I}_{\text{GCM} \leftarrow \text{SPH}} - \tau \mathbf{A}_2 \left(\frac{\partial}{\partial \xi_2} \right)_{\text{GCM}} \mathbf{I}_{\text{GCM} \leftarrow \text{SPH}} - \\ &- \tau \mathbf{A}_3 \left(\frac{\partial}{\partial \xi_3} \right)_{\text{GCM}} \mathbf{I}_{\text{GCM} \leftarrow \text{SPH}} + \mathbf{O}(\tau^2) \mathbf{I}_{\text{GCM} \leftarrow \text{SPH}}, \end{aligned}$$

т. е.

$$\begin{aligned} \mathbf{N}_{\text{GCM}} \mathbf{I}_{\text{GCM} \leftarrow \text{SPH}} &= \mathbf{I}_{\text{GCM} \leftarrow \text{SPH}} + \\ &+ \tau \left(\frac{\partial}{\partial t} \right)_{\text{GCM}} \mathbf{I}_{\text{GCM} \leftarrow \text{SPH}} + \mathbf{O}(\tau^2) \mathbf{I}_{\text{GCM} \leftarrow \text{SPH}}. \end{aligned}$$

Таким образом, оператор \mathbf{N}_{GCM} дает для решения $\mathbf{q}_{\text{GCM}}(t) = \mathbf{I}_{\text{GCM} \leftarrow \text{SPH}} \cdot \mathbf{q}_{\text{SPH}}(t)$ первый порядок аппроксимации.

Пусть оператор \mathbf{N}_{SPH} задает аппроксимацию решения в SPH-области первого порядка точности по времени. Тогда для решения в SPH-области должно

выполняться соответствующее разложение в ряд Тейлора. Распишем его в форме операторов:

$$\mathbf{N}_{\text{SPH}} = \mathbf{I} + \tau \left(\frac{\partial}{\partial t} \right)_{\text{SPH}} + \mathbf{O}(\tau^2).$$

Так как система уравнений (3) выполняется, то должно быть справедливым следующее соотношение:

$$\left(\frac{\partial}{\partial t} \right)_{\text{SPH}} = -\mathbf{A}_1 \left(\frac{\partial}{\partial \xi_1} \right)_{\text{SPH}} - \mathbf{A}_2 \left(\frac{\partial}{\partial \xi_2} \right)_{\text{SPH}} - \mathbf{A}_3 \left(\frac{\partial}{\partial \xi_3} \right)_{\text{SPH}}.$$

Здесь $\left(\frac{\partial}{\partial \eta} \right)_{\text{SPH}}$ — оператор дифференцирования по η функции $\mathbf{q}_{\text{SPH}}(t, \xi_1, \xi_2, \xi_3)$, $\eta = t, \xi_1, \xi_2, \xi_3$, т. е.

$$\begin{aligned} \mathbf{N}_{\text{SPH}} &= \mathbf{I} - \tau \mathbf{A}_1 \left(\frac{\partial}{\partial \xi_1} \right)_{\text{SPH}} - \tau \mathbf{A}_2 \left(\frac{\partial}{\partial \xi_2} \right)_{\text{SPH}} - \\ &- \tau \mathbf{A}_3 \left(\frac{\partial}{\partial \xi_3} \right)_{\text{SPH}} + \mathbf{O}(\tau^2). \end{aligned}$$

При этом если оператор $\mathbf{I}_{\text{SPH} \leftarrow \text{GCM}}$ представляет собой интерполяцию на тетраэдральной сетке полиномом первого порядка, то, следовательно, функция $\mathbf{I}_{\text{SPH} \leftarrow \text{GCM}} \cdot \mathbf{q}_{\text{GCM}}(t, \xi_1, \xi_2, \xi_3)$ кусочно-дифференцируема по координатам. Таким образом, справедливо:

$$\begin{aligned} \mathbf{N}_{\text{SPH}} \mathbf{I}_{\text{SPH} \leftarrow \text{GCM}} &= \mathbf{I}_{\text{SPH} \leftarrow \text{GCM}} - \\ &- \tau \mathbf{A}_1 \left(\frac{\partial}{\partial \xi_1} \right)_{\text{SPH}} \mathbf{I}_{\text{SPH} \leftarrow \text{GCM}} - \tau \mathbf{A}_2 \left(\frac{\partial}{\partial \xi_2} \right)_{\text{SPH}} \mathbf{I}_{\text{SPH} \leftarrow \text{GCM}} - \\ &- \tau \mathbf{A}_3 \left(\frac{\partial}{\partial \xi_3} \right)_{\text{SPH}} \mathbf{I}_{\text{SPH} \leftarrow \text{GCM}} + \mathbf{O}(\tau^2) \mathbf{I}_{\text{SPH} \leftarrow \text{GCM}}, \end{aligned}$$

т. е.

$$\begin{aligned} \mathbf{N}_{\text{SPH}} \mathbf{I}_{\text{SPH} \leftarrow \text{GCM}} &= \mathbf{I}_{\text{SPH} \leftarrow \text{GCM}} + \\ &+ \tau \left(\frac{\partial}{\partial t} \right)_{\text{SPH}} \mathbf{I}_{\text{SPH} \leftarrow \text{GCM}} + \mathbf{O}(\tau^2) \mathbf{I}_{\text{SPH} \leftarrow \text{GCM}}. \end{aligned}$$

Таким образом, оператор \mathbf{N}_{SPH} дает для решения $\mathbf{q}_{\text{SPH}}(t) = \mathbf{I}_{\text{SPH} \leftarrow \text{GCM}} \cdot \mathbf{q}_{\text{GCM}}(t)$ первый порядок аппроксимации.

Аппроксимация второго порядка для гладкого решения доказывается аналогично.

Экспериментальные результаты

В работе [22] были приведены результаты численных экспериментов на прототипе метода для одномерной постановки задачи распада разрыва, показавшие принципиальную перспективность метода. В работе [23] были представлены результаты численного решения трехмерной модельной задачи распада разрыва на усовершенствованной однопроцессорной версии метода GCM-SPH. К моменту написания данной работы был реализован многопоточный метод GCM-SPH для решения трехмерных задач и проведены серии численных экспериментов. Во всех проведенных расчетах все внешние границы являлись свободными.

Распад разрыва

В первую очередь было проведено сравнение численного решения задачи распада разрыва с известным аналитическим решением. Моделирование осуществлялось в двух основных постановках: для продольных и упругих волн внутри железного куба размером $4 \times 4 \times 4$ м, при этом исходный фронт волны проходил через центр куба под наклоном ко всем координатным осям. Разрыв рассматривался в простейшем варианте: начальная скорость условно слева от фронта разрыва бралась равной некоторой ненулевой константе, а справа — нулевой. Проверке подвергалось решение, снятое через равные промежутки вдоль отрезка, проходящего через центр куба перпендикулярно исходному фронту волны и незначительно выходящему из куба. Обе постановки решались каждым из базовых методов в отдельности и двумя вариациями метода GCM-SPH: частицы справа от сетки, и наоборот (границей методов при этом являлась плоскость).

На рис. 1 приведены распределения Y -компоненты скорости v вдоль контрольного отрезка для случая продольных волн в различные моменты времени и для различных численных методов.

Приведенные графики наглядно иллюстрируют корректность численных решений, полученных с помощью метода GCM-SPH. Аналогичные результаты получают и для других компонент скорости, а также для случая поперечных волн. Кроме того, численные эксперименты на постановках задач, отличающихся от описанных выше формой исходного импульса (треугольник и синус), также демонстрируют приемлемое поведение метода GCM-SPH.

Пробой тонкой пластины ударником

Были выполнены расчеты следующих двух задач: численное моделирование пробоя пластины

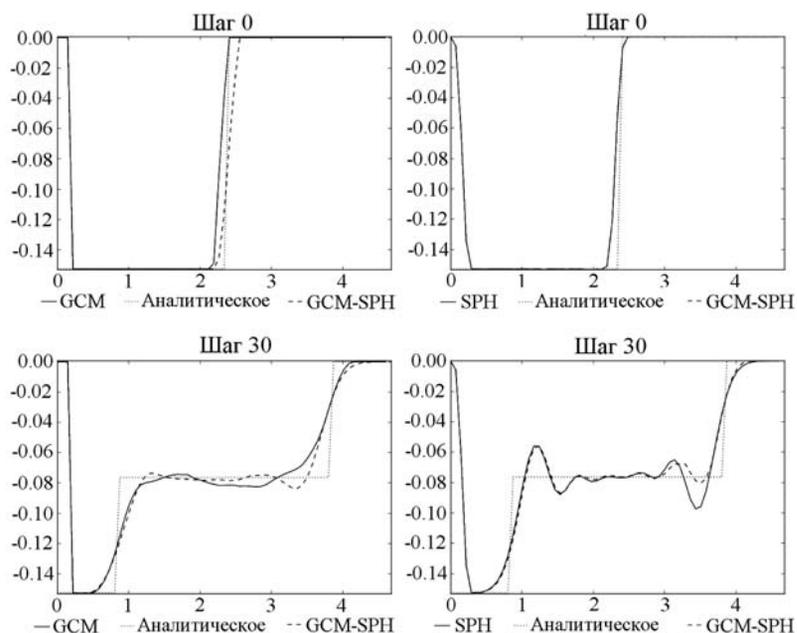


Рис. 1. Сравнение численного решения задачи распада разрыва с аналитическим

кубом, падающим под прямым углом со скоростью 1200 м/с, и шаром, падающим под углом 60° к поверхности. Пластина имела размеры $30 \times 30 \times 6,2$ мм и выполнена из стали 09Г2С, в то время как ударник выполнен из стали 20. Подвергаемая удару поверхность пластины находится в плоскости XU . Длина ребра куба в первой постановке составляла 4,6 мм, а радиус шара во второй постановке — 3,6 мм. Обе постановки решались каждым из двух методов: методом сглаженных частиц и методом GCM-SPH (ударник целиком и часть пластины в области удара моделируются методом сглаженных частиц, а остальная часть пластины — сеточно-характеристическим методом). На контакте между ударником и пластиной задается условие слипания, что является унаследованной особенностью оригинального метода сглаженных частиц.

На рис. 2, 3 изображено движение ударника и пластины в плоскости XZ , полученное при числен-

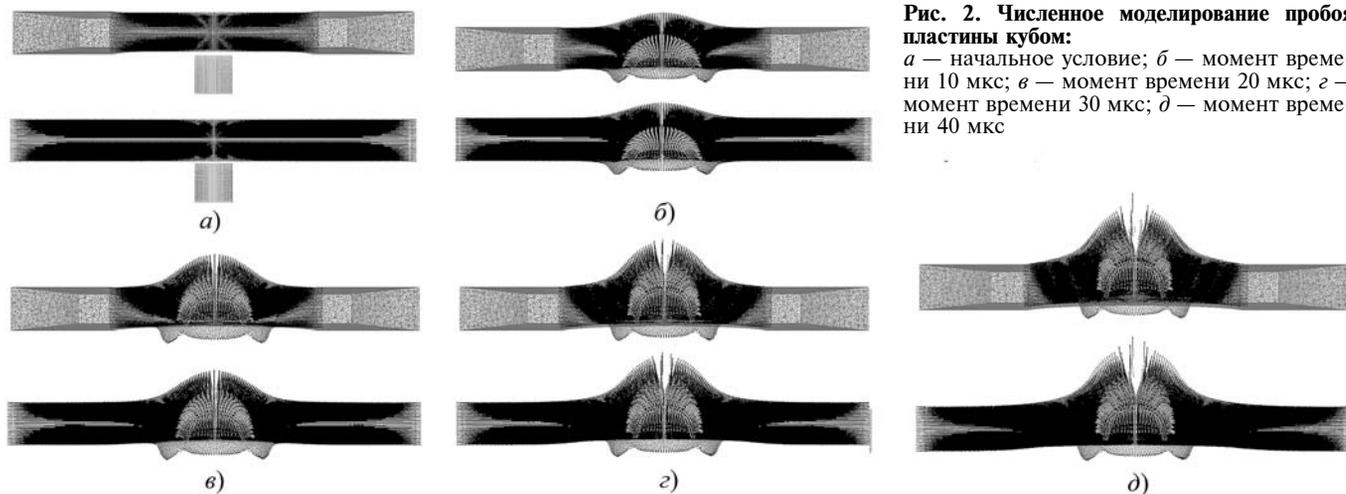


Рис. 2. Численное моделирование пробоя пластины кубом:

a — начальное условие; $б$ — момент времени 10 мкс; $в$ — момент времени 20 мкс; $г$ — момент времени 30 мкс; $д$ — момент времени 40 мкс

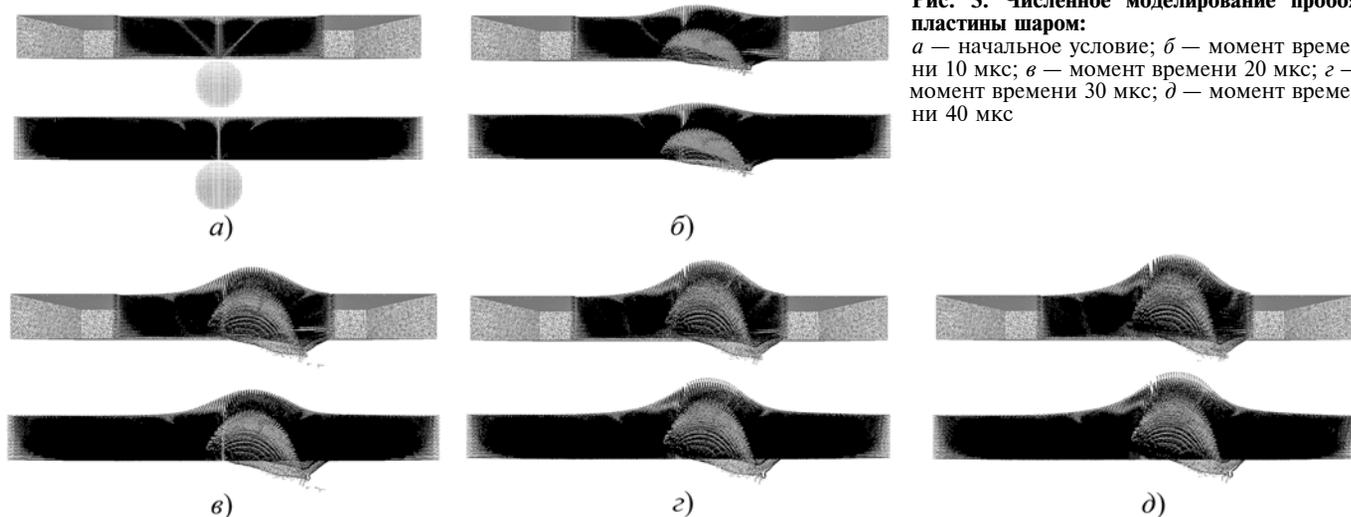


Рис. 3. Численное моделирование пробоя пластины шаром:

a — начальное условие; *б* — момент времени 10 мкс; *в* — момент времени 20 мкс; *г* — момент времени 30 мкс; *д* — момент времени 40 мкс

ном моделировании указанными выше методами, а именно: на рис. 2 представлены результаты численного моделирования пробоя пластины кубом, а на рис. 3 — пробоя шаром. На каждом из рис. 2, 3 внизу изображен расчет методом сглаженных частиц, а сверху — методом GCM-SPH.

Заключение

Ввиду доказанности сохранения комбинированным методом GCM-SPH порядка аппроксимации исходных методов и положительных результатов численных экспериментов ожидается успешное применение метода в задачах моделирования ударных нагрузок, вызывающих частичное разрушение сложных конструкций, для чего необходимы дальнейшие исследования применимости метода к решению задач различного рода.

Список литературы

1. Куликовский А. Г., Погорелов Н. В., Семенов А. Ю. Математические вопросы численного решения гиперболических систем уравнений. М.: ФИЗМАТЛИТ. 2001. 608 с.
2. Петров И. Б., Фаворская А. В. Библиотека по интерполяции высоких порядков на неструктурированных треугольных и тетраэдральных сетках // Информационные технологии. 2011. № 9. С. 30–32.
3. Квасов И. Е., Петров И. Б., Санников А. В., Фаворская А. В. Компьютерное моделирование пространственных динамических процессов сеточно-характеристическим методом на неструктурированных тетраэдральных сетках // Информационные технологии. 2011. № 9. С. 28–30.
4. Петров И. Б., Холодов А. С. Численное исследование некоторых динамических задач механики деформируемого твердого тела сеточно-характеристическим методом // Журн. выч. мат. и мат. физ. 1984. Т. 24, № 5. С. 722–739.
5. Петров И. Б., Челноков Ф. Б. Численное исследование волновых процессов и процессов разрушения в многослойных преградах // Журн. выч. мат. и мат. физ. 2003. Т. 43, № 10. С. 1562–1579.
6. Фомин В. М., Гулидов А. И., Сапожников Г. А. и др. Высокоскоростное взаимодействие тел. Новосибирск: Изд. СО РАН, 1999. 600 с.
7. Уилкинс М. Расчет упруго-пластических течений // Вычислительные методы в гидродинамике. М.: Мир, 1967. С. 212–263.
8. Кукуджанов В. Н. Численное моделирование динамических процессов деформирования и разрушения упругоэластических сред // Успехи механики. 1985. Т. 8. Вып. 4. С. 21–64.

9. Канель Г. И., Фортгов В. Е. Механические свойства конденсированных сред при интенсивных импульсных воздействиях // Успехи механики. 1987. Т. 10, № 3. С. 3–82.
10. Одинцов В. А. Механизм разрушения цилиндров // Вопросы физики взрыва и удара. М.: МВТУ им. Н. Э. Баумана, 1980. Вып. 1. С. 22–70.
11. Сагомонян А. Я. Аналитическое решение задач проникания // М.: Изд-во МГУ, 1974. 299 с.
12. Белов Н. Н., Корнеев А. И., Николаев А. П. Численный анализ разрушения в плитах под действием импульсных нагрузок // ПМТФ. 1985. № 3. С. 132–136.
13. Баженов В. Г., Ломунов В. К., Чекмарев Д. Т. Численное моделирование нелинейных задач динамики упругопластических конструкций // Матем. моделирование. 2006. Т. 18, № 1. С. 10–16.
14. Бабкин А. В., Селиванов В. В. и др. Физика взрыва / Под ред. Л. П. Орленко. В 2 т. Т. 1. М.: ФИЗМАТЛИТ, 2002. 832 с.
15. Libersky L. D., Petschek A. G. Smoothed particles hydrodynamics with strength of materials // Proceedings of The Next Free Language Conference. NY, 1991. P. 248–257.
16. Randles P. W., Libersky L. D., Carney T. C., Sandstrom F. W. SPH simulation of fragmentation in the MK82 bomb // AIP Conference Proceedings, 1996. V. 370. P. 331–334.
17. Liu G. R., Liu M. B. Smoothed Particles Hydrodynamics. Singapore. World Scientific Publishing Co. Pte. Ltd. 2003.
18. Медин С. А., Паршиков А. Н. Развитие метода SPH и его применение в задачах гидродинамики конденсированных сред // Теплофизика высоких температур. 2010. Т. 48, № 6. С. 973–980.
19. Monaghan J. J. An introduction to SPH // Computer Physics Communications. Amsterdam. 1988. V. 48. P. 89–96.
20. Monaghan J. J. SPH without a tensile instability // Journal of Computational Physics. 2000. P. 290–311.
21. Гильманов А. Н., Кулачкова Н. А. Численное исследование двумерных течений газа со скачками методом TVD на физически адаптивных сетках // Матем. моделирование. 1995. Т. 7, № 3. С. 97–106.
22. Шевцов А. В., Черников Д. В. Комбинирование метода сглаженных частиц и сеточно-характеристического метода // Труды 55-й научной конференции МФТИ. 2012. С. 111–112.
23. Шевцов А. В., Фаворская А. В., Потапов А. П., Васюков А. В. Трехмерный комбинированный метод численного моделирования упругопластических тел // Сб. науч. тр. МФТИ: Математические и информационные модели управления. 2013. С. 39–44.
24. Потапов А. П., Ройз С. И., Петров И. Б. Моделирование волновых процессов методом сглаженных частиц (SPH) // Математическое моделирование. 2009. Т. 21, № 7. С. 20–28.
25. Новацкий В. К. Теория упругости. М.: Мир, 1975. 872 с.
26. Новацкий В. К. Волновые задачи теории пластичности. М.: Мир, 1978.
27. Самарский А. А. Введение в теорию разностных схем. М.: Наука, 1971. 552 с.

ИНФОРМАЦИОННО-ИЗМЕРИТЕЛЬНЫЕ СИСТЕМЫ INFORMATION-MEASURING SYSTEMS

УДК 004.05, 004.3-6

А. В. Ермачихин, аспирант, мл. науч. сотр., ассистент, e-mail: al.erm@mail.ru,
В. Г. Литвинов, канд. физ.-мат. наук, доц.,
Рязанский радиотехнический университет

LabVIEW в современной индустрии измерений (обзор)

Дано краткое описание графической среды программирования LabVIEW, ее истории и вклада в развитие современных лабораторных стендов. Описаны основные области применения и иерархия сертификации специалистов. Приведены примеры программ по управлению и автоматизации научных экспериментов.

Ключевые слова: LabVIEW, автоматизация измерений, сертификация, labview portal

A. V. Ermachikhin, V. G. Litvinov

LabVIEW in Modern Measurement Industry (Summery)

This paper provides a brief description of the graphical programming environment of LabVIEW, the history and contribution to the development of modern laboratory stands. The basic application and a hierarchy of certification experts. The examples of programs for the management and automation of scientific experiments.

Keywords: LabVIEW, measurement automation, certification, labview portal

Введение

LabVIEW — это аббревиатура **L**aboratory **V**irtual **I**nstrumentation **E**ngineering **W**orkbench (среда инженерно-графического программирования). Это объектно-ориентированная среда программирования, созданная компанией National Instruments (NI) в 1986 г. [1]. Разработанные в LabVIEW программы называются виртуальными приборами (ВП), имеют расширение **.vi** и состоят из двух частей: лицевой панели и блок-диаграммы. На лицевой панели находится интерфейс программы, в то время как блок-диаграмма отвечает за логику работы программы [1–3].

В основе данной среды программирования лежит визуальный язык G. Главное преимущество визуальных языков программирования — то, что они позволяют наглядно представить такие программные структуры, как, например, алгоритмы и данные. Это противоречит традиционным текстовым языкам программирования, где такие многомерные структуры закодированы в одномерные строки с помощью достаточно сложного синтаксиса. Визуальные языки убирают этот слой абстракции, позволяя программисту непосредственно наблюдать и манипулировать сложными программными структурами [4].

Области применения LabVIEW [3]

LabVIEW находит применение в следующих областях:

- ❖ *в промышленных задачах:*
 - измерение сигналов с датчиков [5];
 - встраиваемые системы измерения и управления [6, 7];
 - автоматические системы тестирования [8, 9];
 - мехатроника и робототехника [10];
- ❖ *в образовании* [11, 12]:
 - управление приборами [13];
 - системы имитации [14, 15];
 - модернизация лабораторных практикумов и установок [16, 17];
 - дистанционное обучение [18];
 - практикумы по техническим дисциплинам [19, 20];
- ❖ *в научных исследованиях:*
 - моделирование [15];
 - автоматизация эксперимента [13, 21, 22];
 - обработка результатов эксперимента [14].

В последнее время LabVIEW используют для построения некоторых из самых крупных и сложных приложений в мире [1], например, приложений для

истребителей пятого поколения F-35, европейской организации ядерных исследований (ЦЕРН) и др.

Благодаря использованию передовой архитектуры программного обеспечения и аппаратных средств компания NI смогла предоставить авиакосмической компании Lockheed Martin Aeronautics расширяемую систему с очень гибкими возможностями настройки для удовлетворения текущих и будущих потребностей F-35 [1].

Чтобы перенаправить любые потенциально опасно заблудившиеся частицы в большом адронном коллайдере, инженерами ЦЕРНа были выбраны LabVIEW и высокопроизводительные модульные платформы NI PXI. Конечная система использует более 100 PXI-систем, которые синхронизированы с точностью 1 мс [1].

Сертификаты от NI

Certified LabVIEW Associate Developer (CLAD) является первым сертификатом. Наличие данного сертификата указывает на широкий круг знаний о среде LabVIEW, базовое понимание кодирования, а также возможность читать и интерпретировать существующий код.

Certified LabVIEW Developer (CLD) указывает на возможность проектирования и разработки функциональных программ при минимизации времени разработки и обеспечения ремонтпригодности через надлежащую документацию.

Certified LabVIEW Architect (CLA) является ключевым этапом в процессе сертификации, где указывается возможность построения разумной иерархии виртуального прибора и план проекта, отвечающего определенным требованиям.

27 февраля 2013 г. [1] компания NI анонсировала новую возможность для сертификации экспертов по встраиваемым системам контроля и мониторинга. Certified LabVIEW Embedded Systems Developer (CLED) демонстрирует профессионализм и опыт в анализе требований, проектировании, разработке, отладке и развертывании приложений для встраиваемых систем на основе NI CompactRIO, NI Single-Board RIO и модулей NI R-серии.

LabVIEW как средство автоматизации экспериментов

Компания NI является одной из передовых в разработке и продвижении технологий компьютерных приборов для измерений, управления и автоматизации [16]. Наличие различных библиотек, тулкетов и экспресс-приборов существенно сокращает время для разработки автоматизированной программы управления экспериментом или обработки данных.

Существенным преимуществом виртуального прибора является его модульность. При проектировании автоматизированной системы можно разработать отдельные виртуальные приборы для каж-

дого модуля, которые затем объединить в один прибор, реализующий функции системы в целом.

Далее приведены уже существующие решения для управления приборами и экспериментами в лабораториях регионального центра зондовой микроскопии коллективного пользования (РЦЗМкп) при ФГБОУ ВПО "РГРТУ".

Автоматизированная установка низкочастотной спектроскопии. Для управления установкой по спектроскопии низкочастотного (НЧ) шума была разработана программа [12], блок-диаграмма которой показана на рис. 1 (см. третью сторону обложки). Обработка результатов НЧ спектроскопии обычно занимает достаточно много времени, но с применением вычислительной техники результаты могут быть получены практически в режиме реального времени. Усовершенствование методов обработки результатов НЧ шумовой спектроскопии позволяют получать более надежные результаты и за меньшее время.

Всю блок-диаграмму условно можно разделить на несколько частей:

- задание начальных установок, выбор файла записи результатов (показано на рис. 1, см. третью сторону обложки);
- изменение напряжения смещения на образце (рис. 2, см. третью сторону обложки);
- временная задержка, вывод спектральной плотности мощности (СПМ) шума в реальном времени (показано на рис. 1);
- запись данных в файл, построение усредненной СПМ и вольт-шумовой характеристики (ВШХ) (рис. 3, см. четвертую сторону обложки).

Как правило, шумовые параметры структур измеряются в стационарных условиях, т. е. при установленном значении тока через структуру и при постоянной температуре. Поэтому алгоритмы управления установкой должны содержать временные паузы, необходимые для полной релаксации переходных процессов, вызванных, например, изменением напряжения постоянного смещения на структуре.

Анализ сигнала заключается в вычислении зависимости СПМ от частоты. Сигнал подвергается умножению на весовую функцию "окна". В качестве такой функции используется окно Блекмана—Харрисона. Далее вычисляется автокорреляционная функция от сигнала. СПМ вычисляется как Фурье-преобразование от автокорреляционной функции.

Вычисление частотной зависимости СПМ с наложением весовой функции "окна" в среде LabVIEW представлено единым функциональным блоком, настройка которого заключается в выборе параметров "окна". Далее вычисляются несколько значений СПМ, проводится их усреднение путем вычисления среднего арифметического по каждой частотной составляющей и сохранение результатов в файл на жесткий диск ЭВМ (рис. 3).

В процессе работы программы на монитор выводится графическая зависимость СПМ на вы-

бранной частоте от напряжения электрического смещения на исследуемой структуре. Как показали эксперименты, эта зависимость позволяет качественно контролировать источник шумового сигнала (образец или установку) в течение всего эксперимента. После записи результатов в файл устанавливается новое значение напряжения смещения на образце и процесс измерений повторяется до того момента, пока не будет остановлен оператором.

В результате работы измерительной установки формируется файл данных "СПМ — частота — напряжение смещения", обработка которого позволяет выявить особенности спектра НЧ шума в изучаемом образце.

Автоматизированный измерительно-аналитический комплекс для исследования иммитанса. Для исследования полупроводниковых структур в современных электрофизических лабораториях часто применяют измерительные комплексы, состоящие из различных приборов. Чем сложнее комплексы и чем больше информации об образцах необходимо сохранить, тем больше риск возникновения ошибки из-за человеческого фактора. Поэтому существует необходимость в создании автоматизированных измерительно-аналитических комплексов.

Программное обеспечение комплекса реализовано в среде инженерно-графического программирования LabVIEW с использованием библиотек Agilent E4980.lvlib и Lake Shore 335.lvlib [1], что позволило снизить временные затраты на разработку управляющей программы для измерительного комплекса. Кроме того, имеются широкие возможности для представления результатов эксперимента в той или иной форме и их математической обработки.

На рис. 4 (см. четвертую сторону обложки) представлена блок-диаграмма программы, которая управляет комплексом, изменяет температуру, снимает, обрабатывает и сохраняет результаты измерений в файл.

На рис. 5 представлена блок-диаграмма виртуального подприбора, который управляет RLC измерителем Agilent E4980A.

На рис. 6 приведена блок-диаграмма виртуального подприбора, который проверяет наличие заданного файла данных. Если файл отсутствует, то создает его, а также в зависимости от выбранного измерения и числа частот создает необходимые подписи для графиков.

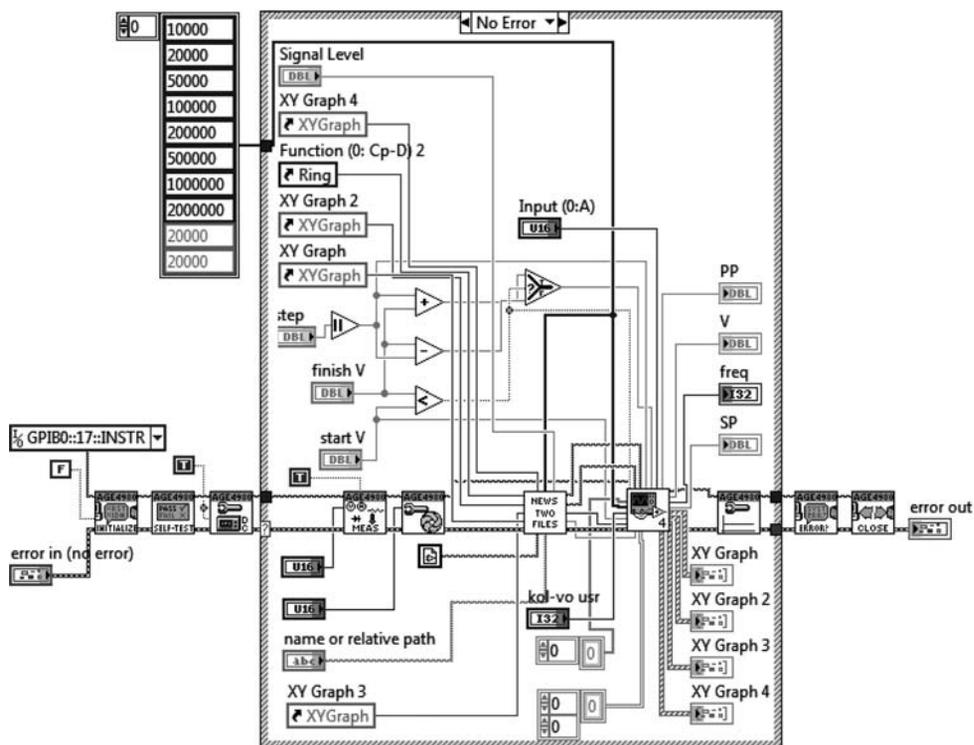


Рис. 5. Блок-диаграмма виртуального подприбора, управляющего Agilent E4980A

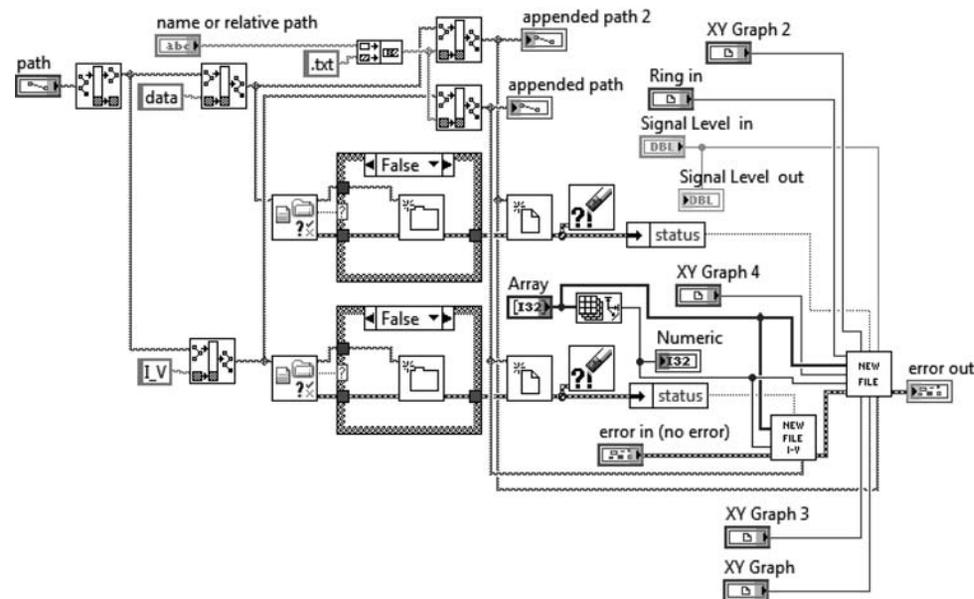


Рис. 6. Блок-диаграмма виртуального подприбора, создающего новые файлы для записи

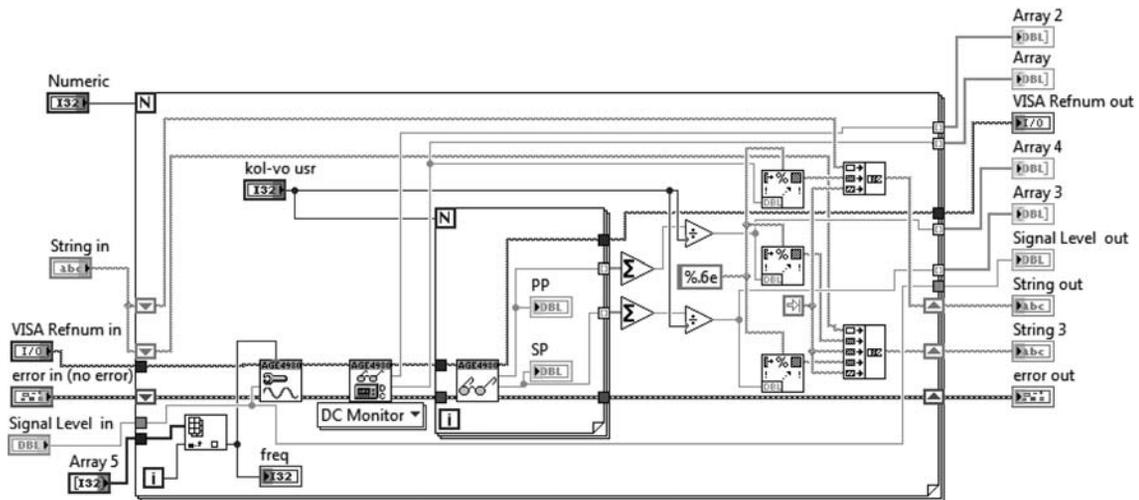


Рис. 7. Блок-диаграмма виртуального подприбора, осуществляющего установку необходимых амплитуды и частоты измерительного сигнала

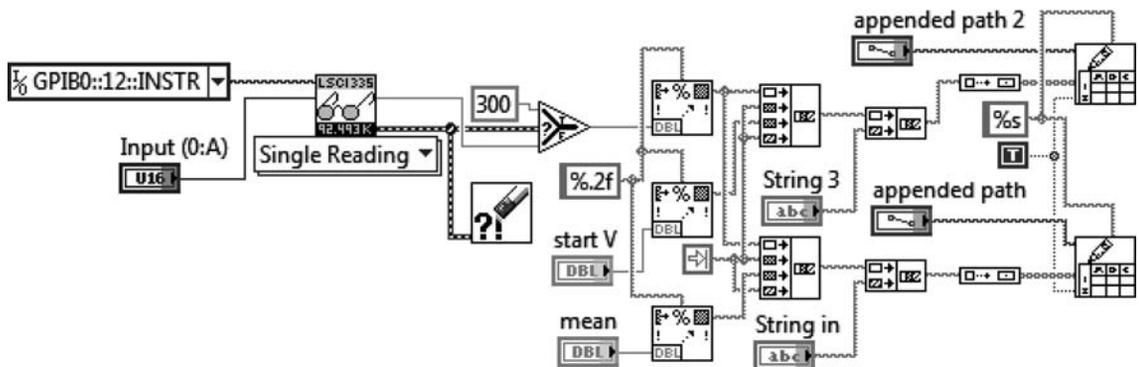


Рис. 8. Блок-диаграмма виртуального подприбора, осуществляющего запись данных в файл

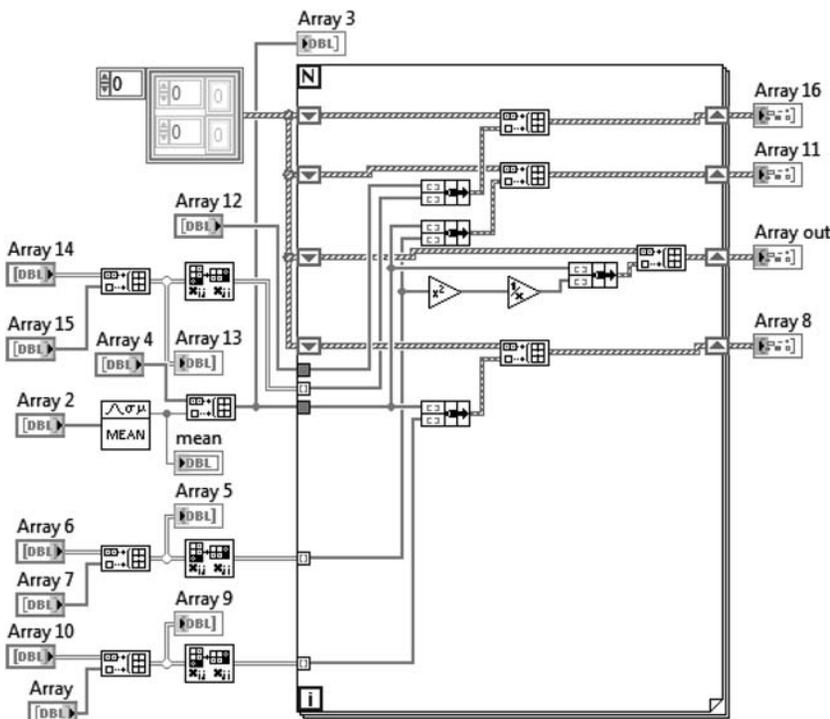


Рис. 9. Блок-диаграмма виртуального подприбора, осуществляющего вывод данных в виде графиков на лицевую панель виртуального прибора

На рис. 7 представлена блок-диаграмма виртуального подприбора, который отвечает за установку необходимых амплитуды и частоты измерительного сигнала и изменение постоянного напряжения смещения на образце. Запись данных в файл показана на рис. 8, вывод данных в виде графиков на лицевую панель виртуального прибора — на рис. 9.

Заключение

Почти все современные измерительные приборы имеют совместимость с LabVIEW через интерфейсы GPIB, RS-232 или шину USB. Многие из этих приборов уже в комплекте имеют программы, написанные на LabVIEW для управления и визуализации измеренных данных.

LabVIEW имеет достаточно низкий порог вхождения, заметно сокращает время на написание управляющих или аналитических программ.

К минусам LabVIEW можно отнести:

- узкую направленность решаемых задач (написать утилиту как игру или текстовый редактор практически невозможно, хотя есть исключения [22]);
- меньшую производительность по сравнению с другими языками программирования;
- отсутствие драйверов и тулкетов для не-Windows платформ (Mac, Linux);
- для запуска программы на "чистом" ПК необходима LabVIEW Run-Time библиотека [1].

В рейтинге языков программирования **TIОBE** [23] LabVIEW вышла из 50 лидирующих позиций в 2013 г., причиной этого можно назвать узкий круг решаемых задач, но для научных экспериментов, автоматизации измерительных и производственных комплексов данный продукт подходит как нельзя лучше.

Существует много учебной и технической литературы для разного уровня подготовки. Также существует много Интернет-ресурсов по данной тематике, стоит отметить один из лучших — **LabVIEW Portal** [22]. Данный портал включает в себя специалистов разного уровня и направлений деятельности, которые живут в разных уголках нашей планеты. Это дает возможность оценить любую проблему и задачу под разными углами. Пользователи данного ресурса являются людьми, которые занимаются учебной и научной деятельностью, например [21, 24, 25].

Работа выполнена при финансовой поддержке Министерства образования и науки РФ с использованием оборудования РЦЗМкп при ФГБОУ ВПО "РГРТУ" в НОЦ неупорядоченных и наноструктурированных материалов и устройств на их основе.

Список литературы

1. Сайт компании National Instruments / www.ni.com.
2. Русскоязычный сайт LabVIEW. URL: <http://www.labview.ru>.
3. Smedley Trevor J., Cox Philip T. Visual Languages for the Design and Development of Structured Objects // Journal of Visual Languages and Computing — VLC. 1997. Vol. 8, N 1. P. 57—84.
4. Смутко С. В., Неймак В. С., Параска Г. Б. Экспериментальные динамические исследования одностепенного зубчато-важильного планетарного механизма с использованием программной среды LabVIEW // Вестник Хмельницкого национального университета. Технические науки. 2013. № 3. С. 241—244.
5. Григорьев И. Single-board RIO — платформа компании National Instruments для встраиваемых систем // Электроника: Наука, технология, бизнес. 2009. № 2. С. 70—73.
6. LabVIEW для разработки встраиваемых систем // Электроника: Наука, технология, бизнес. 2006. № 3. С. 50—53.
7. Синютин Е. С. Система автоматической проверки работоспособности мобильных полиграфов с применением инструментария LabVIEW // Изв. Южного федерального университета. Технические науки. 2011. Т. 118, № 5. С. 51—55.
8. Манонина И. В. Применение программы LabVIEW для изучения вопросов поверки измерительных приборов // Т-Comm: Телекоммуникации и транспорт. 2012. № 8. С. 50—52.

9. Бутусов Д. Н., Красильников А. В., Райцин С. Б., Миронченко Е. А., Клушников Ю. В. Модельное проектирование системы управления шасси роботизированной платформы // Современные проблемы науки и образования. 2013. № 1.

10. Константиян Т. К. Фактор активности в процессе обучения графическому программированию с использованием LabVIEW // Вестник Российского университета дружбы народов. Сер. "Информатизация образования". 2009. № 4. С. 102—106.

11. Сепоян П., Соков О. А. Образовательная программа National Instruments // Информационные системы и технологии. 2006. № 1—5. С. 34—37.

12. Ермачихин А. В., Литвинов В. Г., Кострюков С. А., Мальченко С. И. Программа для автоматизированного измерения НЧ шумов и спектральной плотности мощности шума в элементах электронной техники. Свидетельство о государственной регистрации программы для ЭВМ № 2012617503. Зарегистрировано в Реестре программ для ЭВМ 20 августа 2012 г.

13. Ермачихин А. В., Литвинов В. Г., Мальченко С. И., Рыбин Н. Б. Программа для моделирования спектров релаксационной спектроскопии глубоких уровней полупроводниковых диодных структур. Свидетельство о государственной регистрации программы для ЭВМ № 2013611095. Зарегистрировано в Реестре программ для ЭВМ 9 января 2013 г.

14. Ермачихин А. В., Литвинов В. Г., Мальченко С. И., Рыбин Н. Б. Программа для построения и обработки спектров релаксационной спектроскопии глубоких уровней полупроводниковых диодных структур. Свидетельство о государственной регистрации программы для ЭВМ № 2013611098. Зарегистрировано в Реестре программ для ЭВМ 9 января 2013 г.

15. Петров М. Н., Белехов Я. С. Построение интерфейса виртуального измерительного прибора // Вестник Новгородского государственного университета им. Ярослава Мудрого. 2003. № 23. С. 96—99.

16. Сорокин С. С. Цифровые лаборатории в профориентированном обучении инженеров в комплексе "ЛИЦЕЙ-ВУЗ" // Педагогическое образование и наука. 2008. № 9. С. 57—60.

17. Прибылов Н. Н., Прибылова Е. И., Прицепова С. А. Лабораторный практикум по физике для дистанционного обучения // Физическое образование в вузах. 2003. Т. 9, № 2. С. 108—112.

18. Соков О. А., Корячкин В. П., Горбачев Н. Б., Галаган П. В. Принципы построения учебных лабораторных комплексов на базе программно-управляемых информационно-измерительных систем // Информационные системы и технологии. 2004. № 2 (3). С. 80—84.

19. Горбачев Н. Б., Галаган П. В., Мосин Ю. В., Тарасова М. А. Лабораторный практикум по физике в среде LabVIEW в системе инженерного образования // Информационные системы и технологии. 2006. № 1—3. С. 45—48.

20. Холомина Т. А., Кострюков С. А., Литвинов В. Г., Ермачихин А. В. Спектроскопия низкочастотных шумов полупроводниковых приборов // Датчики и системы. 2013. № 5. С. 15—21.

21. Ермачихин А. В., Литвинов В. Г., Паршутин В. О. Установка для исследования параметров диодных структур методом двух гармоник // Инженерные и научные приложения на базе National Instruments—2012. Сб. тр. XI междунар. науч.-практ. конф. М.: ДМК Пресс. 2012. С. 398—400.

22. Сайт LabVIEW Portal. URL: <http://www.labviewportal.eu>.

23. Сайт TIОBE (индекс популярности различных сред программирования). URL: <http://www.tiobe.com/index.php/content/paperinfo/tpci/index.html>.

24. Петренко П. Б., Кривоzubов П. А., Морозова Е. С. Повышение точности импульсной рефлектометрии на основе согласованной фильтрации с учетом экспериментальных оценок статистических характеристик шума // Вопросы защиты информации. 2011. № 3. С. 35—40.

25. Кострюков С. А., Ермачихин А. В., Литвинов В. Г., Холомина Т. А., Рыбин Н. Б. Измерительный комплекс спектроскопии низкочастотных шумов полупроводниковых диодных структур // Измерительная техника. 2013. № 9. С. 61—64.

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ЭКОНОМИЧЕСКИХ СИСТЕМАХ INFORMATION TECHNOLOGIES IN ECONOMIC SYSTEMS

УДК 519.584

А. А. Токмакова, студент,
МФТИ, e-mail: aleksandra-tok@yandex.ru

Алгоритм стохастического отбора объектов и признаков в задаче банковского кредитного скоринга

Рассматривается задача совместного отбора информативных признаков и объектов в задаче кредитного скоринга. Для оценки вероятности риска невозврата кредита используется логистическая регрессия. Отбор выполняется с помощью стохастической процедуры оптимизации. Для снижения размерности признакового пространства решена задача группировки признаков, возникающая в результате бинарного представления порядковых и номинальных признаков. Для повышения качества классификации проводится разбиение объектов на группы. Для оценки качества кластеризации предложена модификация ROC-кривой для разбиения множества объектов более чем на два класса. Вычислительный эксперимент выполнен на исторических данных о потребительских кредитах за несколько последних лет.

Ключевые слова: банковский скоринг, отбор признаков, отбор объектов, стохастическая оптимизация, номинальная шкала

A. A. Tokmakova

Stochastic Optimization Algorithm for Feature and Object Selection with Application to Credit Scoring

We consider feature and object selection problem for logistic regression with application to credit scoring, and propose the selection procedure that involves stochastic optimization. To improve prediction quality, we suggest to cluster the objects and to train the model separately on different clusters. To estimate clusterization quality, we propose the modification of the ROC curve for multiclass case. Another improvement comes from the binary representation of ordered and nominal features, which allows additional dimension reduction by dummy variables grouping. The performance of the procedure is evaluated on consumer loans data.

Keywords: feature selection, object selection, credit scoring, stochastic optimization, nominal scales

Введение

При построении скоринговых моделей альтернативой экспертным методам [1] являются алгоритмы интеллектуального анализа данных [2]. Требуется спрогнозировать вероятность невозврата кредита. Заемщик (далее — объект выборки) предоставляет сведения в виде ответов на вопросы кредитной анкеты, именуемые признаками. Требуется решить проблему исключения избыточного числа признаков [2, 3], поскольку их наличие может привести к ухудшению качества классификации и снижению устойчивости прогностической модели.

На качество прогнозирования вероятности риска невозврата кредита влияет набор объектов [1, 4], т. е. результат оптимизации может зависеть не только от вопросов анкеты, но и от состава выборки заем-

щиков. Проблема заключается в том, что для разных классов людей набор признаков, существенных при принятии решения о выдаче кредита, может быть разным [5].

На текущий момент основным методом отбора признаков в задаче банковского скоринга является пошаговая логистическая регрессия [6, 7]. Для группировки значений признаков в номинальных шкалах используется метод, основанный на WOE (Weight of Evidence) — характеристике, отображающей статистическую значимость признака [1]. Модификации и комбинации этих алгоритмов применяются не только для отбора признаков, но и для кластеризации объектов. Однако применение пошаговой регрессии приводит к завышенным оценкам значимости признаков регрессии и переобучению [8].

В данной работе для отбора признаков и объектов используется генетический алгоритм [9, 10]. Качество классификации объектов оценивается с помощью модификации ROC-кривой — графика, позволяющего оценить качество бинарной классификации, отображающий соотношение между долей верных положительных классификаций от общего числа положительных классификаций и долей ошибочных положительных классификаций от общего числа отрицательных классификаций при варьировании порога решающего правила [11]. Предлагается новый подход группировки значений признаков в номинальных шкалах, позволяющий значительно снизить размерность признакового пространства. В вычислительном эксперименте выполнено сравнение результатов классификации классических методов и подхода, описанного в работе.

1. Постановка задачи об отборе признаков

Измеряемая выборка представляет собой множество значений зависимой переменной y_i и независимой переменной x_i . Выборка $D = (\mathbf{X}, \mathbf{y}) = \{(\mathbf{x}_i, y_i)\}_{i=1}^m$ — множество m пар, состоящих из вектора $\mathbf{x}_i = [x_{ij}]_{j=1}^n$ значений n свободных переменных и соответствующего этому вектору значения зависимой переменной $y_i \in \{1, 0\}$, где y_i — индикатор возврата заемщиком кредита.

Принимается гипотеза о том, что \mathbf{y} — многомерная случайная величина, имеющая распределение Бернулли: $\mathbf{y} \propto \text{Be}(\boldsymbol{\sigma})$, причем все y_i независимы в совокупности. Здесь σ_i — вероятность возврата кредита заемщиком с индексом i . Вектор $\boldsymbol{\sigma}$ определяется с помощью модели логистической регрессии:

$$\mathbf{y} = \boldsymbol{\sigma}(\mathbf{w}) + \boldsymbol{\varepsilon};$$

$$\boldsymbol{\sigma}(\mathbf{X}, \mathbf{w}) = f(\mathbf{X}, \mathbf{w}) = \frac{1}{1 + \exp(-\mathbf{X}\mathbf{w})}, \quad (1)$$

где $\mathbf{w} \in R^n$ — фиксированный вектор весов признаков.

Зададим индексацию на множестве объектов $I = \{1, \dots, m\}$. Разобьем исходную выборку $D = \{(\mathbf{x}_i, y_i)\}_{i=1}^m$ на обучающую, $i \in L$, и контрольную, $i \in T$, причем множества L и T не пересекаются и образуют полное покрытие множества рассматриваемых объектов.

Учитывая гипотезу о распределении зависимой переменной \mathbf{y} (1), запишем выражение для функции правдоподобия:

$$P(D|\mathbf{w}) = \prod_{i \in L} \sigma(\mathbf{x}_i, \mathbf{w})^{y_i} (1 - \sigma(\mathbf{x}_i, \mathbf{w}))^{1 - y_i}.$$

Будем использовать логарифм функции правдоподобия как критерий оценки качества модели:

$$-\ln P(D|\mathbf{w}) = - \sum_{i \in L} (y_i \ln \sigma(\mathbf{x}_i, \mathbf{w}) + (1 - y_i) \ln(1 - \sigma(\mathbf{x}_i, \mathbf{w}))) = \ln S(\mathbf{w}), \quad (2)$$

где $S(\mathbf{w})$ — функция ошибки.

Воспользуемся методом максимального правдоподобия и запишем оценку вектора параметров модели в следующем виде:

$$\hat{\mathbf{w}} = \underset{\mathbf{w} \in R^n, i \in L}{\operatorname{argmin}} \ln S(\mathbf{w}). \quad (3)$$

Таким образом, задача отбора признаков является задачей оптимизации. Информативными будем называть те признаки, которые повышают качество классификации. Введем индексацию на множестве признаков $J = \{1, \dots, n\}$. Пусть $A \subseteq J$ — множество индексов информативных признаков. Тогда задача выбора признаков модели (1) имеет вид

$$A = \underset{A \subseteq J}{\operatorname{argmin}} \ln S(\hat{\mathbf{w}}|\mathbf{X}(T, A), \mathbf{y}(T)), \quad (4)$$

где оценка вектора параметров модели получена минимизацией выражения

$$\hat{\mathbf{w}} = \underset{\mathbf{w} \in R^{|A|}}{\operatorname{argmin}} \ln S(\mathbf{w}|\mathbf{X}(L, A), \mathbf{y}(L)).$$

Наблюдаемые признаки с индексами из множества A позволят повысить качество классификации, а также устойчивость и интерпретируемость прогностической модели.

Также для повышения качества классификации предлагается разбить множество объектов на классы, каждый из которых будет описываться индивидуальным набором оптимальных параметров.

В качестве альтернативного критерия согласно банковским требованиям для оценки качества разбиения клиентов на классы предлагается использовать площадь под ROC-кривой. ROC-кривая — график, позволяющий оценить качество бинарной классификации, отображающий соотношение между долей верных положительных классификаций TPR (англ. True Positive Rate, называемой чувствительностью алгоритма классификации) от общего числа положительных классификаций и долей ошибочных положительных классификаций FPR (англ. False Positive Rate) от общего числа отрицательных классификаций при варьировании порога решающего правила.

Введем обозначение $a_i = \mathbf{x}_i^T \mathbf{w}$. Тогда доля верных положительных классификаций

$$\text{TPR}(a_i, X) = \frac{\sum_{i=1}^n [a_i = +1][y_i = +1]}{\sum_{i=1}^m [y_i = +1]}.$$

Доля ложных положительных классификаций:

$$\text{FPR}(a_i, X) = \frac{\sum_{i=1}^n [a_i = +1][y_i = -1]}{\sum_{i=1}^m [y_i = -1]}.$$

2. Стохастический отбор признаков

Для решения задачи (4) предлагается использовать следующую процедуру.

Множество индексов признаков A , используемых в модели, задается бинарным вектором \mathbf{a} :

$$\begin{cases} \mathbf{a}_j = 1, & \text{если } j \in A; \\ \mathbf{a}_j = 0, & \text{иначе.} \end{cases}$$

Рассмотрим некоторое множество бинарных векторов $F = \{\mathbf{a}_i\}$. Для каждого вектора \mathbf{a}_i из множества F оценивается вектор параметров (3) модели (1) и вычисляется значение функции ошибки (2):

$$c_i = S(\mathbf{a}_i | \mathbf{X}(L, A), \mathbf{y}(L)). \quad (5)$$

1. Отбирается P векторов с наименьшим значением c_i , где P — параметр процедуры. Без ограничения общности будем считать, что P — четное число. Полученное множество векторов обозначим $F' = \{\mathbf{a}_1^T, \dots, \mathbf{a}_P^T\}$.

2. Множество F' случайным образом разбивается на пары $\{\mathbf{a}_i^T, \dots, \mathbf{a}_j^T\}$.

3. С каждой парой $\{\mathbf{a}_i^T, \dots, \mathbf{a}_j^T\}$ проводится операция скрещивания:

- генерируется случайное число $\xi \in \{1, \dots, n-1\}$;
- векторы разделяются на две части и смешиваются следующим образом:

$$[a_i^1, \dots, a_i^\xi, a_j^{\xi+1}, \dots, a_j^n] \rightarrow \mathbf{a}'_i,$$

$$[a_j^1, \dots, a_j^\xi, a_i^{\xi+1}, \dots, a_i^n] \rightarrow \mathbf{a}'_j.$$

4. Операция модификации:

- генерируется $t < n$ попарно различных случайных чисел $\mu_1, \dots, \mu_t \in \{1, \dots, n\}$;
- инвертируется значение позиций μ_1, \dots, μ_t векторов \mathbf{a}'_i и \mathbf{a}'_j . Полученное множество векторов обозначим как $F'' = \{\mathbf{a}''_i\}_{i=1}^P$.

5. Для каждого вектора из множества F'' оценим значение вектора параметров модели (3) и вычислим значение функции ошибки $S(\mathbf{a}''_i | \mathbf{X}(L, A), \mathbf{y}(L))$.

6. При стабилизации множества векторов F'' процедура останавливается. В противном случае процедура переходит на новую итерацию, начиная с шага 1.

Для того чтобы определить, насколько стабилизировался набор признаков, задаваемых векторами \mathbf{a}' из множества F' введем понятие энтропии W . Найдем нормированные расстояния Хэмминга между всеми парами векторов \mathbf{a}'_i и \mathbf{a}''_k , где $i, k \in \{1, \dots, P\}$, из множеств F' и F'' :

$$\rho(\mathbf{a}'_i, \mathbf{a}''_k) = \frac{\sum_{j=1}^n \text{XOR}(\mathbf{a}'_i^j, \mathbf{a}''_k^j)}{n},$$

где \mathbf{a}^j — обозначение для j -й компоненты вектора.

Далее необходимо сформировать $\frac{P}{2}$ пар $(\mathbf{a}'_i, \mathbf{a}''_k)$ из наборов с минимальным расстоянием Хэмминга ρ .

Тогда величина энтропии W , показывающая насколько стабилизировалось множество векторов F' , записывается в виде:

$$W = - \sum_{l=1}^{P/2} \rho_l \ln \rho_l,$$

где суммирование ведется по сформированным парам.

Итерационная процедура выбора признаков выполняется до тех пор, пока значение энтропии W не стабилизируется. Вектор \mathbf{a}'' , полученный на последней итерации, задает множество A информативных признаков.

3. Постановка задачи о группировке значений признаков в номинальных шкалах

Так как значения номинального признака не сравнимы между собой, все значения номинальной шкалы K необходимо представить в виде бинарных признаков, называемых фиктивными переменными. В этом разделе описан подход, при котором индексируются бинарные векторы, описывающие значения номинального признака. Размерность признакового пространства снижается за счет группировки фиктивных переменных.

Рассмотрим признак ξ_γ , заданный в номинальной шкале K . При использовании фиктивных переменных для учета признака ξ_γ в модели получим:

$$\xi_\gamma = [\chi_1, \dots, \chi_{|K|-1}],$$

где $\chi_i \in \{0, 1\}^m$.

Решим задачу снижения размерности признакового пространства с помощью группировки векторов фиктивных переменных на некоторое число групп $|\Gamma|$. Для этого построим сюръективное отображение $h: K \rightarrow \Gamma$, которое каждому элементу из номинальной шкалы ставит в соответствие номер группы:

$$\begin{array}{ccccccc} K = & 1 & 2 & 3 & \dots & |K| \\ & \downarrow & \downarrow & \downarrow & & \downarrow \\ \Gamma = & 1 & 1 & 2 & \dots & |\Gamma|. \end{array}$$

Обозначим все множество индексов признаков, кроме γ , как $W = J \setminus \{\gamma\}$, а множество индексов признаков, задаваемое с помощью группировки бинарных векторов как V . Требуется найти функцию $h: K \rightarrow \Gamma$, обеспечивающую минимум функции ошибки:

$$(h, \Gamma) = \operatorname{argmin}_{h \in H} S(\mathbf{w} | \mathbf{X}(T, W \cup V), \mathbf{y}(T)), \quad (6)$$

здесь H — множество всех возможных отображений $K \rightarrow \Gamma$.

Задача (6) также решается с помощью стохастической процедуры оптимизации, описанной в разделе 2. Для этого фиксируется набор информативных признаков A и рассматриваются только признаки в номинальных шкалах. Итерационную процедуру необходимо выполнять для каждого номинального признака в отдельности. Осуществляется перебор

числа всех возможных групп $|\Gamma|$, причем из множества F' вектору ξ_γ ставятся в соответствие векторы $\mathbf{a}'_i \in \{0, \dots, |\Gamma|\}^{|\mathcal{K}|}$. При расчете значения функции ошибки (5) признаки χ_i с одинаковыми индексами групп, получаемыми из вектора \mathbf{a}'_i , суммируются.

4. Описание стохастической процедуры кластеризации объектов

В работе [12] показано, что удаление из выборки нехарактерных, шумовых объектов значительно повышает качество классификации. Таким образом, при оптимизации целевой функции (2) необходимо сформировать множество информативных объектов B .

Для разбиения объектов более чем на две группы необходимо модифицировать предложенный критерий. Зафиксируем множество активных признаков модели A .

- Зададим индексацию на множестве объектов с помощью вектора $\mathbf{a}_i \in \{0, \dots, k\}^m$, где k — число классов. Рассмотрим некоторое множество векторов $Z = \{\mathbf{a}_i\}$.
- Будем считать, что каждая группа объектов описывается отдельной моделью. Рассмотрим вектор \mathbf{a}_j . Для всех пар варьирующего порога каждой из моделей построим ROC-кривую, суммируя TPR и FPR моделей. Подсчитаем площадь AUC (англ. Area Under the Curve) под получившейся ROC-кривой. Поставим в соответствие каждому вектору из множества Z значение площади под соответствующей ROC-кривой.
- Далее процедура работает аналогично процедуре, описанной в разделе 2, за исключением того, что на третьем шаге отбирается P векторов с наибольшим значением критерия AUC.

5. Описание базового подхода

Классическим решением задачи кредитного скоринга считается подход, состоящий из трех ступеней:

1. Фильтрация признаков [1]. В качестве алгоритма, формирующего множество информативных признаков, используется пошаговая логистическая регрессия.

2. Группировка значений признаков в номинальных шкалах основана на подсчете критерия WOE (Weight of Evidence), отображающего статистическую значимость каждого значения номинального признака [1]:

$$WOE_i = \ln \frac{\text{DistrGood}_i}{\text{DistrBad}_i},$$

где DistrGood_i — отношение числа людей, которым выдали кредит, имевших данное значение номинального признака, к общему числу людей, которым выдали кредит; DistrBad_i —

отношение числа людей, которым не выдали кредит, имевших данное значение номинального признака, к общему числу людей, которым не выдали кредит. В одну группу объединяются значения признака в номинальной шкале с близкими значениями WOE.

В скоринговых моделях в качестве меры информативности номинальных признаков выступает критерий IV (Information Value):

$$IV = \sum_{i=1}^k (\text{DistrGood}_i - \text{DistrBad}_i) \ln \frac{\text{DistrGood}_i}{\text{DistrBad}_i}.$$

3. Кластеризация объектов [13]. На этом этапе происходит отбор информативных объектов и их разделение по группам. В качестве базового в работе выбран метод k -ближайших соседей.

6. Вычислительный эксперимент

В этом разделе показано преимущество использования процедуры стохастической оптимизации перед базовыми алгоритмами. Исходная выборка [5] содержит информацию о 1680 клиентах, подававших заявления на потребительские кредиты. Матрица плана \mathbf{X} состоит из 14 признаков: 5 числовых, 8 номинальных и 1 бинарный. Проводился отбор информативных признаков, группировка векторов номинальных признаков и кластеризация объектов.

Отбор информативных признаков. Каждый номинальный признак необходимо представить в виде множества бинарных векторов, так как его значения не сравнимы между собой. После представления всех номинальных признаков полная матрица плана \mathbf{X} стала содержать 65 признаков. Стохастическая процедура фильтрации признаков позволила снизить размерность признакового пространства до 48 признаков. Итерационный процесс для функции энтропии W и функции ошибки $S(\mathbf{w})$ изображен на рис. 1 и рис. 2.

Группировка векторов номинальных признаков. В табл. 1 приведен размер номинальной шкалы K для каждого номинального признака. В третьей колонке показано, какие признаки были включены

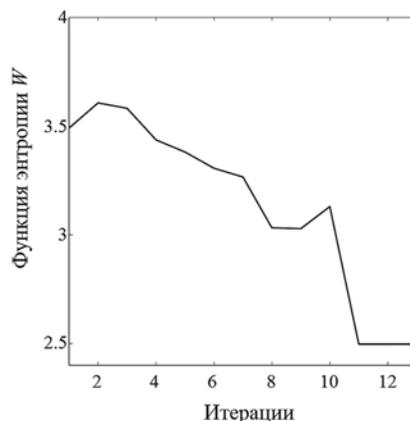


Рис. 1. Сходимость функции энтропии W

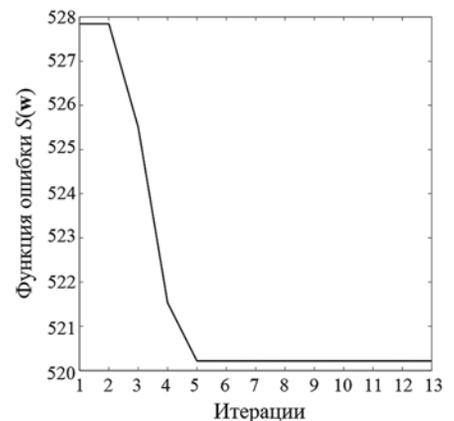


Рис. 2. Сходимость функции ошибки $S(\mathbf{w})$

Таблица 1

Группировка признаков в номинальных шкалах

Наименование	Размер номинальной шкалы	Результаты фильтрации и группировки
Регион проживания	7	3
День подачи заявления	7	3
Семейное положение	4	Отсутствует
Уровень образования	5	1
Форма занятости	4	1
Профессия	26	2
Автомобиль	3	Отсутствует
Кредитные карты	3	Отсутствует

Таблица 2

Сравнение результатов работы базовых алгоритмов и стохастической процедуры оптимизации

Название этапа	Базовые алгоритмы, AUC	Предложенная процедура, AUC
Фильтрация признаков	0,57 (50)	0,64 (48)
Группировка признаков	0,60 (24)	0,61 (12)
Кластеризация объектов	0,62 (2)	0,84 (4)

Примечание. Для этапа фильтрации и группировки признаков в скобках указано итоговое число признаков в матрице плана X. Для этапа кластеризации объектов в скобках указано число кластеров.

в итоговую модель и каким числом групп $|\Gamma|$ они представлены в матрице плана X.

После повторного запуска стохастической процедуры фильтрации со сгруппированными номинальными признаками число признаков в модели снизилось до 12.

Кластеризация объектов. Для подсчета AUC был использован метод кросс-валидации, проведены исследования по устойчивости оценок значения критерия в зависимости от числа измерений. Будем считать, что значение AUC при каждом случайном разбиении выборки на обучение и контроль — случайная величина, имеющая нормальное распределение. Тогда для того чтобы в 95 % случаях отклонение 2σ (для левого и правого хвостов распределения) не превышало 0,01, необходимо усреднять показания по 100 случайным разбиениям (рис. 3).

Значения критерия AUC, которые достигаются при разбиении выборки заемщиков на $k = 2, \dots, 6$ групп, представлены на рис. 4.

Таким образом, наибольшее значение критерия AUC достигается при $k = 4$. При дальнейшем разбиении объектов на группы модель начинает переобучаться. Сходимость значения AUC для $k = 4$ показана на рис. 5.

Сравнение результатов базовых алгоритмов из раздела 5 для фильтрации признаков и группировки признаков в номинальных шкалах с результатом работы стохастической процедуры оптимизации приведено в табл. 2.

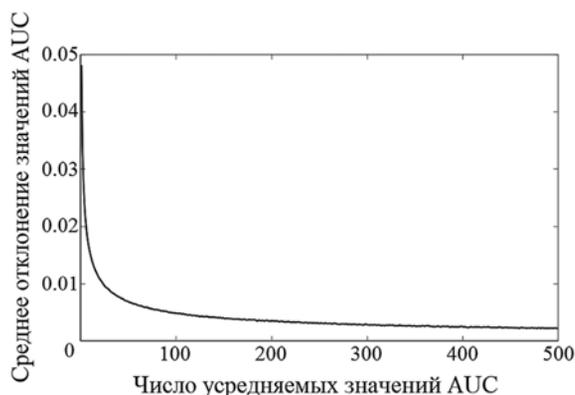


Рис. 3. Число значений, необходимых к усреднению в методе кросс-валидации

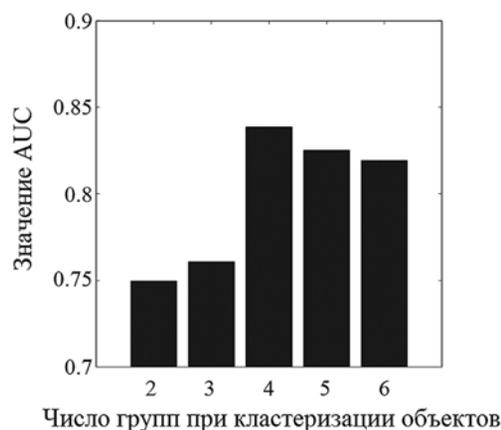
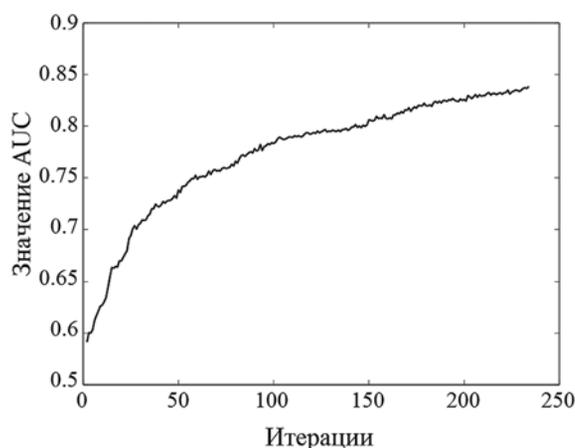


Рис. 4. Зависимость значения AUC при разбиении выборки заемщиков на классы

Рис. 5. Сходимость значения AUC при $k = 4$

Снижение значения критерия стохастической процедуры оптимизации при группировке признаков наблюдается в связи со значительным снижением размера признакового пространства.

Заключение

Решена задача совместного отбора признаков и объектов, а также разбиение множества объектов на классы. В вычислительном эксперименте пока-

зано, что процедура стохастической оптимизации превосходит базовые алгоритмы по качеству классификации. Несмотря на то что сама процедура работает достаточно медленно, в задачах банковского кредитного скоринга нет необходимости быстрого нахождения результатов оптимизации: необходимо как можно точнее предсказать вероятность невозврата кредита. Полученные модели используются для прогнозирования риска невозврата кредита в реальном времени.

Работа поддержана РФФИ № 14-07-31045 и № 12-0731095.

Список литературы

1. **Siddiqi N.** Credit risk scorecards: developing and implementing intelligent credit scoring. New Jersey: Wiley and Sons, 2006.
2. **Goldberg D.** Genetic algorithms in search, optimization and machine learning. New York: Adison-Wesley, 1989.
3. **Thomas L.** A survey of credit and behavioural scoring: forecasting financial risk of lending to consumers // International Journal of Forecasting. 2000. Vol. 16. P. 149—172.

4. **Никаненкова В. В.** Кредитный скоринг как инструмент оценки кредитоспособности заемщика // Вестник Адыгейского государственного университета. Сер. 5: Экономика. 2012. № 2. С. 32—38.

5. **German** Credit data. — Institute of Statistics and Econometrics, Hamburg, 1994. URL: <http://archive.ics.uci.edu/ml/machine-learning-databases/statlog/german/>.

6. **Hosmer D., Lemeshow S.** Applied logistic regression. New York: Wiley and Sons, 2000.

7. **Bolton C.** Logistic regression and its application in credit scoring. Pretoria: University of Pretoria, 2009.

8. **Roecker E.** Prediction error and its estimation for subset-selected model // Technometrics. 1991. Vol. 33. P. 459—468.

9. **Leardi R., Boggia R., Terrile M.** Genetic algorithms as a strategy for feature selection // Journal of chemometrics. 1992. Vol. 5. N 5. P. 267—281.

10. **Oh I. S., Lee J. S., Moon B. R.** Hybrid genetic algorithms for feature selection // IEEE transactions on pattern analysis and machine intelligence. 2004. Vol. 26, N 11. P. 1424—1437.

11. **Hand D., Till R.** A simple generalization of the area under the ROC curve for multiple class classification problems // Machine learning. 2001. Vol. 45. P. 171—186.

12. **Адуенко А. А.** Выбор признаков и шаговая логистическая регрессия для задачи кредитного скоринга // Машинное обучение и анализ данных. 2012. № 3. С. 279—291.

13. **МакКей D.J.C.** Information Theory, Inference, and Learning Algorithms. Cambridge, UK: Cambridge University Press, 2003.

УДК 330.24; 004.89

Л. А. Исмагилова, д-р техн. наук, проф., директор Института экономики и управления, зав. каф.,
e-mail: ismagilova_ugatu@mail.ru,

Е. В. Орлова, докторант, канд. техн. наук, доц., e-mail: ekorl@mail.ru,
Уфимский государственный авиационный технический университет

Методологические основы и инструментарий бюджетно-налогового моделирования и прогнозирования¹

Рассматриваются проблемы моделирования и прогнозирования бюджетно-налоговых потоков в условиях неопределенности и риска. Сформированы принципы разработки интеллектуальной технологии моделирования и прогнозирования в системе налогообложения. Приведен пример оценки налогового потенциала и прогнозирования налоговых поступлений в бюджет административно-территориального образования.

Ключевые слова: бюджетно-налоговое планирование, моделирование, инструментальные средства, неопределенность, методы искусственного интеллекта

L. A. Ismagilova, E. V. Orlova

Methodological Basis and Tools for Fiscal Modeling and Forecasting

The problem of modeling and forecasting fiscal flows under uncertainty and risk is considered in the article. A set of principles was used for development an intelligent technology in taxation management system. The example of the tax potential assessment and tax revenues prediction of the administrative-territorial unit is carried out.

Keywords: fiscal and tax planning, modeling, tools, uncertainty, methods of artificial intelligence

¹ Работа выполнена при финансовой поддержке РГНФ, проект № 12-12-02000.

Введение

Моделирование систем бюджетно-налогового планирования и прогнозирования составляет основу решения практических задач оценивания налогового потенциала, планирования и прогнозирования налоговых поступлений в бюджет. Целями системы бюджетно-налогового планирования и прогнозирования является оценка соотношения налогового потенциала с фактическими поступлениями налогов и определение на этой основе экономически обоснованных объемов и структуры поступлений налогов в бюджетную систему в планируемом периоде. При этом налоговое планирование осуществляется с ориентацией на обоснованно высокий уровень мобилизации налогов и сборов в бюджетную систему.

Процесс налогообложения должен комплексно рассматриваться в рамках системы налогового менеджмента [1, с. 51], [2, с. 19], т. е. системы государственного и корпоративного управления налоговыми потоками на основе применения обоснованных рыночных форм и методов принятия решений в области управления налоговыми доходами и налоговыми расходами. Основными свойствами системы налогового планирования и прогнозирования как сложной, динамической, многоуровневой, стохастической и нестационарной системы являются:

- взаимозависимость уровней дохода бюджета и налогообложения;
- взаимосвязь с системой управления экономикой и финансами государства;
- комплексный и стратегический характер принимаемых управленческих налоговых решений;
- динамический характер системы планов и прогнозов;
- многовариантность управленческих решений в области налогообложения;
- влияние на налоговые потоки множества факторов риска и неопределенности.

Указанные свойства предопределяют формирование рациональной системы налогового менеджмента и должны учитываться при разработке моделей и алгоритмов управления.

Существующие подходы и модели формирования систем бюджетно-налогового планирования и прогнозирования неполностью отражают перечисленные свойства. Прогнозы налоговых доходов строят на основе статистических методов [10, с. 103], [12, с. 270], не учитывающих трансформаций в системах цен, характера финансово-хозяйственных сделок, разнообразия договоров и др. Подходы, направленные на решение частных задач налогового контроля [7, с. 89], ограничены их применением на этапах камеральных и выездных проверок. Методы регулирования налогового механизма [3, с. 77] не обладают прогностическими способностями и не могут быть использованы в системах бюджетного прогнозирования. При этом техноло-

гии налогового планирования и прогнозирования должны, во-первых, отражать сложный динамический и многофакторный характер систем налогового менеджмента, во-вторых, воспроизводить существенные свойства и особенности функционирования системы в целях обоснованного планирования доходной и расходной частей бюджета.

1. Источники данных и особенности функционирования бюджетно-налоговой системы в условиях неопределенности

Учет при моделировании качественных оценок, экспертных знаний, а также законов и закономерностей функционирования социально-экономических систем во внешней среде позволяет отразить при моделировании их гетерогенные свойства. Адекватность моделей, способность их к адаптации и ориентированность на имеющиеся данные определяют точность решения задач. Быстроразвивающиеся в последнее время методы анализа неопределенности и рисков в экономических системах базируются на различных понятиях и трактовках категорий "неопределенность" и "риск". С позиций управления различие между неопределенностью и риском в том, что риск, как проявление вероятности наступления какого-либо события из множества возможных, может быть измерен или рассчитан. Риск объективно существует и связан с вероятностной природой многих процессов, многовариантностью материальных, социальных, информационных отношений, в которые вступают субъекты социально-экономических систем.

Статистические методы, используемые при моделировании для решения задач оценки, анализа, планирования и прогнозирования, наиболее распространены, хотя они справедливы в предположении о возможности описания экономического показателя нормальным законом распределения. В реальной системе состояние описывается множеством взаимосвязанных показателей, природу формирования которых нельзя рассматривать как влияние многих равнозначных факторов, воспроизводимых во времени и ситуации. Существенное значение имеют также синергетические системные эффекты, когда случайный характер каждого из элементов системы при их взаимодействии меняется непредсказуемым образом.

Неопределенность как экономическая категория отражает невозможность точного учета и расчета изменений, возникающих в системе под воздействием сложных причинно-следственных связей, социальных, политических, экономических, демографических и иных причин. Неопределенность нельзя измерить и ею нельзя управлять в силу ограниченности научных знаний о процессах и закономерностях развития экономики. В то же время проявление неопределенности в экономических ситуациях приводит к отклонениям показателей,

потере точности расчетов и прогнозов, а потому должно быть учтено.

Учет неопределенности при решении задач статистической обработки данных традиционными методами регрессионного анализа невозможен в силу указанных обстоятельств. Известные методы искусственного интеллекта, такие как нечеткие логики, нейронные сети, экспертные системы позволяют решать отдельные задачи обработки данных.

Комплексный подход к эффективному моделированию систем бюджетно-налогового планирования предполагает разработку интегрированной многоэтапной информационной технологии, анализ разнородных сведений о системе, ее элементах и внешней среде [1, с. 46]. Можно выделить четыре категории сведений, характерных для социально-экономической системы, которые необходимо учитывать при моделировании:

1) данные официальной статистики о функционировании системы внутрисистемного и внесистемного уровней за несколько временных периодов (чем больше, тем точнее оценки);

2) статистические данные об изменении основных экономических показателей в регионе, в стране и в мире (рост цен, уровень инфляции, курсы валют, демографическая ситуация и т. п.);

3) знания экспертов об основных закономерностях функционирования моделируемой системы, а также аналогичной экономической ситуации;

4) знания экспертов о законах и закономерностях экономического развития общества на региональном, государственном и мировом уровне.

Смысл этой классификации в том, что такой подход интегрирует сведения о функционировании системы, отражая все основные процессы, формирующие свойства системы, в том числе неопределенность и риск поведения. Каждая из категорий предполагает свои методы обработки данных, соответствующие характеру данных, а поэтому наиболее эффективные.

Моделирование системы бюджетно-налогового планирования сводится к разработке структуры информационной технологии, отражающей особенности данных, использующей математические и эвристические методы обработки данных и знаний, определяющей последовательность анализа и способы получения результатов.

Технология анализа данных в общем виде может быть представлена как $\inf\{E_I\} \xrightarrow{M_d} \{E_{i \in I}\}$; $M_d: \{O, R, C, V\}$, где $\{E_I\}$ — множество данных об элементах E_I системы; $i \in I$ — подмножество элементов, сгруппированных по принципу общности свойств; M_d — технологии, методы анализа данных, алгоритмы, схемы отбора, классификации данных; O — критерии или цели анализа; R — правила отбора экономических объектов; C — действия по

управлению, возможные в сложившейся ситуации; V — оценки показателей функционирования и развития системы. В зависимости от особенностей решаемых задач инструментарий моделирования и прогнозирования включает методы анализа данных M_{du} ($u = \overline{1, 4}$), которые соответствуют разным категориям сведений, используемых при решении задач анализа, планирования и прогнозирования социально-экономических систем.

Технология анализа данных $M_{d1}: \inf\{E_I\} \xrightarrow{M_{d1}} \{E_{i \in I}\}$; $M_{d1}: \{O, R\}$ предполагает выявление экономически схожих объектов, имеющих общие тенденции функционирования и, соответственно, управления. Обобщение как один из итогов анализа статистических данных об объектах экономической системы имеет важное значение для выявления тенденций при прогнозировании в условиях неопределенности данных. Парировать неблагоприятное (чаще) влияние некоторых видов рисков на экономическую ситуацию можно, выявляя заблаговременно рискованные ситуации. Наиболее эффективно задачи обобщения решаются на нейронных сетях. По сравнению с традиционными методами математической статистики, классификации и аппроксимации нейронные сети обеспечивают достаточно высокое качество решений при меньших затратах. Они позволяют выявлять нелинейные закономерности в неоднородных избыточных данных, дают хорошие результаты при большом числе входных параметров и обеспечивают хорошие решения при относительно небольших объемах данных.

Технология анализа статистических данных

$M_{d2}: \inf\{P\} \xrightarrow{M_{d2}} \{P_i \in P_k\}$; $M_{d2}: \{R, V\}$; $P_k \in P$ об изменении основных показателей объекта реализуется методами статистического анализа, поскольку всегда имеется достаточное количество данных об общих закономерностях развития. Отличие этого этапа от предыдущего в том, что выявляются тенденции изменения значений P_i показателей $\{P_k\}$, правила R и действия C соответствуют выбранному методу статистического анализа, а критерии O и оценки V определяют точность аппроксимации. Применение традиционных методов статистического анализа обосновано в том случае, если статистики действительно достаточно и если данные однородны. В условиях нестабильной экономики задачи аппроксимации также можно успешно решать на нейронных сетях.

Технология анализа плохо формализуемых знаний экспертов $M_{d3}: \inf\{S\} \xrightarrow{M_{d3}} \{\text{str}_i(S)\}$; $M_{d3}: \{R, C\}$ реализуется методами искусственного интеллекта, когда разрабатываются экспертные системы, в том числе нечеткие экспертные системы. Знания подвергаются морфологическому, синтаксическому и семантическому анализу и структурируются $\text{str}(S)$ в соответствии с множеством правил R и действий C .

Технология анализа знаний экспертов о законах и закономерностях развития общества M_{d4} :

$\text{inf}\{S\} \xrightarrow{M_{d4}} \{\text{str}_i(S)\}$; $M_{d4} : \{O, R, C, V\}$ соответствует рассмотренной технологии M_{d3} и может отличаться количеством и видом выделенных структур, а также методами построения экспертных систем.

2. Технология бюджетно-налогового моделирования и прогнозирования

Принципы построения. Особенность решения задачи оценки потенциальных налоговых сборов (налогового потенциала региона, территории) и прогнозирования налоговых поступлений в бюджет состоит в том, что необходимо рассматривать разнородные по финансово-экономическим, технико-экономическим и организационным показателям объекты налогообложения, разные условия функционирования схожих (например, предприятий одной отрасли) объектов в разных административно-территориальных образованиях, а также разные условия дифференциации налоговых изъятий.

Анализ исходных данных о сумме собранных за предыдущий период налогов, структуре налоговых доходов, об отраслевой принадлежности и общественной значимости ведущих предприятий регионального образования позволяет выявить проблемы, усложняющие решение задачи бюджетно-налогового планирования и прогнозирования. Таких проблем две: статистическая неоднородность данных и отсутствие сопоставимых данных за некоторый временной интервал. Обе эти проблемы имеют объективную природу и не могут игнорироваться при применении известных методов статистического анализа. Объективные сложности измерения социально-экономических показателей во времени порождаются динамическими процессами развития систем, которые особенно заметны в переходный период экономики. В условиях быстрых экономических и институциональных преобразований теряется статистическая однородность данных о деятельности экономических субъектов, становятся несопоставимыми условия получения данных и, как следствие, неточными прогнозы.

В связи с этим необходимо провести анализ данных о налоговых поступлениях в бюджет с позиции их достаточности, однородности, зависимости или независимости друг от друга, достоверности, и наконец, с позиции их воспроизводимости, т. е. возможности получить и подтвердить одни и те же данные из разных источников. Используются следующие системные принципы построения системы бюджетно-налогового моделирования и прогнозирования:

- принцип классификации предприятий по структуре налоговых обязательств, отражающих структуру налогооблагаемой базы; такой принцип по-

зволяет сгруппировать предприятия, находящиеся в схожих условиях хозяйствования и, следовательно, имеющие похожие тенденции развития;

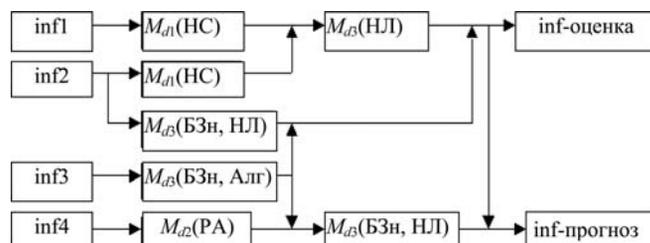
- принцип обобщения описания каждого из классов некоторой обобщенной характеристикой, который позволяет заменить решение общей задачи оценки и прогнозирования решением нескольких подзадач, более точно отражающих специфику налогообложения разнородных предприятий;
- принцип использования объективных знаний о закономерностях изменения внешнеэкономических факторов и об их влиянии на налогооблагаемые базы и на сумму налоговых поступлений.

Информационная технология моделирования налогового потенциала административно-территориального образования и прогнозирования собираемости налогов построена с учетом выделенных категорий следующих данных: inf1 — налогооблагаемая база основных налогов (на прибыль, на добавленную стоимость, на имущество, подоходного налога, единого социального налога и др.); inf2 — показатели финансово-хозяйственной деятельности экономических объектов налогообложения; inf3 — данные об изменении курсов валют, инфляции, цены на нефть; inf4 — институциональные правила формирования бюджетов.

Технология моделирования оценки налогового потенциала и прогнозирования налоговых поступлений в бюджет региона (или административно-территориального образования) может быть представлена без учета сервисных информационных блоков в виде схемы информационного взаимодействия. На рисунке показаны этапы и последовательность обработки информации в соответствии с технологиями анализа M_d и методами обработки.

Технология основана на использовании следующих методов анализа данных: нейронные сети (НС), нечеткие логики (НЛ), базы знаний (БЗн), расчеты (Алг) и статистический регрессионный анализ (РА).

Для решения поставленной задачи необходимо провести анализ данных о налоговых поступлениях в бюджет: во-первых, с позиции их достаточности; во-вторых, с позиции их однородности; в-третьих, с позиции их зависимости или независимости друг от друга; в-четвертых, с позиции их достоверности;



Технология бюджетно-налогового моделирования и прогнозирования

в-пятых, с позиции их воспроизводимости, т. е. возможности получить и подтвердить одни и те же данные из разных источников.

Достаточные, в смысле репрезентативности данные не являются достаточными для прогнозирования, поскольку прогнозирование с использованием временных рядов возможно и обоснованно только при наличии статистически однородных данных за ряд лет. Анализ однородности данных предполагает выявление исходных данных, которые получены в одинаковых или хотя бы сопоставимых условиях. При оценке налогового потенциала в имеющиеся данные о налоговых поступлениях включены фактически уплаченные налоги с учетом переходящих задолженностей перед бюджетом за прошлые периоды, что не соответствует схеме расчета налогооблагаемой базы и планирования налоговых поступлений.

Данные о налогооблагаемых базах и данные о собранных налогах являются недетерминированно-зависимыми. Так, увеличение прибыли приводит не только к увеличению налога на прибыль, но и опосредованно влияет на изменение других видов налогов. Экономические показатели, описывающие налогооблагаемую базу объектов налогообложения, также находятся в явной или неявной зависимости друг от друга. Для выявления характера зависимостей факторов и использования их для решения задач бюджетно-налогового моделирования и прогнозирования ограниченно применимы методы регрессионного анализа, с учетом того, что регрессионный анализ накладывает ограничения на качество представляемых данных, такие как однородность, репрезентативность и воспроизводимость.

Достоверность данных определяется способом получения данных, правильностью постановки эксперимента (технологии работы с документами) и добросовестностью представления данных. Поскольку в экономических системах, в отличие от технических, не существует объективных физических экспериментов и экономические данные представляются в виде статистической отчетности самих предприятий, а также контролирующих служб, приходится исходить из предположения, что достоверность при сборе данных обеспечена.

Воспроизводимость данных как необходимое условие эффективности методов анализа и прогнозирования при оценке налоговых поступлений не обеспечивается полностью, так как сведения о налоговых поступлениях могут существенно изменяться применительно к конкретным условиям и времени, например, за счет разнообразных налоговых льгот и "местных" видов налогов.

Таким образом, для выявления основных тенденций изменения налогового потенциала региона необходимо ставить и решать задачи обобщения, способные сгладить нежелаемые при статистической обработке свойства исходных данных (неоп-

ределенность, неоднородность, достаточность и взаимозависимость).

На основе сформулированных выше системных принципов решаются следующие задачи.

1. Классификация предприятий как объектов налогообложения. Выявление однородных по структуре налогооблагаемой базы предприятий. Определение обобщенного описания каждого класса однородных по структуре налогооблагаемой базы экономических объектов.

2. Анализ факторов внешней среды, влияющих на объекты налогообложения и на размер бюджетных поступлений. Это внешнеэкономические факторы; факторы, определяющие финансово-экономическое состояние предприятий; институциональные факторы в виде изменений налогового кодекса и правовых межбюджетных отношений.

3. Оценка изменений налоговых поступлений в текущем и прогнозирование собираемости налогов в следующем году.

Этапы реализации технологии. Информационные технологии, основанные на нейронных сетях и теории нечетких множеств, позволяют получать качественно новые результаты при решении таких задач, как классификация, прогнозирование, принятие решений и т. п. и поэтому могут быть применены для решения задачи оценки и прогнозирования налоговых поступлений в бюджет административно-территориального образования [8, с. 50]. Реализация информационной технологии осуществляется в несколько этапов, каждый из которых предполагает решение одной из самостоятельных подзадач. При этом методы решения изменяются при переходе от одного этапа к другому и соответствуют особенностям подзадач [4, с. 312].

Этап 1. Анализ исходной информации и обоснование принципов классификации. Анализ данных об объектах налогообложения (предприятиях) показывает, что, во-первых, они относятся к разным отраслям производства, во-вторых, имеют разный уровень налогов (в том числе вследствие налоговых льгот), в-третьих, имеют разную структуру налогооблагаемой базы и, в-четвертых, в разной степени зависят от влияния внешней среды. Это означает, что данные об уплате налогов за предыдущий год не являются статистически однородной совокупностью, и предварительное разделение предприятий на классы по принципу однородности структуры налогооблагаемой базы позволит выявить схожие по экономической сути объекты, объединить их в классы и исследовать отдельно.

Этап 2. Разделение объектов на классы. В результате разделения объектов налогообложения (предприятий) на классы определяются группы предприятий со схожими налоговыми возможностями. При этом объем уплаченных налогов зависит от производственных мощностей и размеров предприятий. В группе могут быть крупные и мелкие

предприятия разных отраслей. Выполнение данной подзадачи базируется на использовании аппарата нейросетей. При решении задачи классификации на смоделированной и обученной нейронной сети для каждого вектора уплаченных в городской (республиканский или иной) бюджет налогов можно определить класс, к которому объект может быть отнесен с наибольшей уверенностью.

Этап 3. Формирование обобщенного описания классов. Обобщение показателей и построение осредненной, наиболее вероятной характеристики каждого класса позволит в дальнейшем решить задачу прогнозирования более обоснованно, поскольку изучаются однородные объекты налогообложения с наиболее близкими законами развития.

Этап 4. Уточнение описания классов. Уточнение классификации. После проведенной классификации можно получить описание нескольких классов в виде интервалов значений, описывающих в нормированных величинах границы изменения каждого вида налогов по всем классам. Задача решается с помощью методов нечетких множеств, т. е. методов, устанавливающих соотношение между качественными и количественными значениями показателей экономической деятельности.

Введение понятия "эталонный", "типичный" представитель и построение центра класса позволяет в дальнейшем работать не с каждым объектом налогообложения в отдельности, а с обобщенным представителем, что будет в значительной степени соответствовать тенденции изменений показателей у обобщенного объекта налогообложения.

Этап 5. Разработка базы знаний для оценки и прогнозирования. Принципиально важно, что недостаточное для прогнозирования количество информации о собираемых за длительный период налогах восполняется за счет использования объективных знаний о тех или иных факторах, поддающихся точному измерению по временному ряду, что позволяет рассматривать экономические и другие процессы в динамике. Это три группы факторов, прямо или косвенно влияющих на налоговую базу и сумму собранных налогов: внутренние (дебиторская задолженность, прибыль, себестоимость продукции, численность работающих и др.), внешние (цены на нефть, курс доллара, инфляция и др.), институциональные (изменение налогового законодательства и межбюджетных отношений на региональном и федеральном уровнях). Для выявления качественной зависимости между факторами и структурой налоговых обязательств (основными налогами) разрабатывается база знаний (правил, устанавливающих взаимосвязь значений факторов и показателей и качественные значения факторов).

Этап 6. Анализ и расчет влияния факторов на налоги. Прогнозирование уровня поступления налогов. Для вычисления прогнозных значений изменений налогов в зависимости от изменений

внешней среды используется база знаний, в которой описываются экономические закономерности влияния выделенных внешних факторов и внутренних факторов, характеризующих налогооблагаемую базу, на значение собираемых налогов.

4. Численный пример реализации технологии

Для оценки налогового потенциала и прогнозирования налоговых поступлений в бюджет города Уфы выбраны предприятия, составляющие представительную выборку, результаты анализа которой могут быть распространены на всю совокупность экономических субъектов налогообложения. Обоснованность такого предположения определяется общим объемом налогов, собранных на этих градообразующих предприятиях и составляющих около 50 % от всех налоговых поступлений.

Анализируются статистические данные об уплаченных крупными предприятиями города Уфы налогах: налог на прибыль, налог на добавленную стоимость, подоходный налог, налог на имущество, прочие налоги в городской бюджет. Кроме того, имеются данные по каждому предприятию города о следующих факторах: 1 — выручка; 2 — себестоимость; 3 — среднегодовая стоимость имущества; 4 — среднесписочная численность работающих; 5 — денежные средства на оплату труда; 6 — прибыль-убыток отчетного периода; 7 — дебиторская задолженность; 8 — кредиторская задолженность.

Классификация субъектов налогообложения. Для решения задачи классификации субъектов налогообложения по принципу однородности структуры уплаченных налогов построена нейросетевая модель M_T , которая представляется в следующем виде [5, с. 116]:

$$M_T = \{In\{5, НДС, НП, НИ, ПН, ПР\}, Layer\{(2, 10), (3, 10)\}, Out\{5, Z^1, Z^2, Z^3, Z^4, Z^5\}\},$$

где НДС, НП, НИ, ПН, ПР — исходные значения основных налогов, пронормированных на интервале (0,1); Z^1, Z^2, Z^3, Z^4, Z^5 — классы однородных по структуре уплаченных налогов предприятий; число классов определяется видами налогов, принятыми в качестве основных; $Layer\{(2, 10), (3, 10)\}$ — два скрытых слоя нейронной сети, содержащих по 10 нейронов. Тип активационной функции нейросетевого отображения — гиперболический тангенс. Входными переменными для модели нейросети являются пронормированные значения собранных налогов, а выходными переменными — принадлежность к классам однородных по структуре уплаченных налогов субъектов налогообложения. Обучение нейросети осуществляется на основе обучающей выборки, представленной в табл. 1 и состоящей из 10 обучающих пар. Метод обучения — обратное распространение ошибки. Оценку качест-

Обучающая выборка по нормированным данным о собранных налогах

Направление	Вид налога	Номер эксперимента									
		1	2	3	4	5	6	7	8	9	10
Входы	НДС	0,401	0,808	0	0,449	0,021	0,273	0,014	0,299	0	0,330
	НП	0	0,359	0,220	0,812	0	0,158	0	0,269	0,016	0,326
	НИ	0,006	0,347	0,001	0,165	0,250	0,686	0,005	0,223	0,001	0,009
	ПН	0	0,249	0	0,105	0,063	0,243	0,224	0,819	0	0,089
	Прочие	0,012	0,481	0,050	1	0,036	0,375	0,028	0,303	0,451	1
Выходы	Z1	1	1	0	0	0	0	0	0	0	0
	Z2	0	0	1	1	0	0	0	0	0	0
	Z3	0	0	0	0	1	1	0	0	0	0
	Z4	0	0	0	0	0	0	1	1	0	0
	Z5	0	0	0	0	0	0	0	0	1	1

ва обучения проводят с помощью служебной информации, записываемой при обучении нейросети.

В результате разделения субъектов налогообложения на классы определены группы предприятий со схожими налоговыми возможностями, т. е. с похожим финансово-экономическим портретом. При этом объем уплаченных налогов зависит от производственных мощностей и размеров предприятий. В группе есть как крупные, так и мелкие предприятия.

После проведенной классификации получено описание пяти классов в виде интервалов, описывающих в нормированных значениях границы изменения каждого вида налогов по всем классам. Для первого класса вектор интервальных значений Z_1 имеет вид: $Z_1 = \{\text{НДС}(0,296; 0,590), \text{ПН}(0,114; 0,238), \text{НИ}(0,002; 0,229), \text{НП}(0,028; 0,249), \text{ПР}(0,052; 0,268)\}$. По представленным в векторе Z_1 границам изменений налогов в каждом классе строится вектор центра классов Z_1^0 , каждое значение которого оп-

ределяется как значение функции принадлежности. В этом случае можно составить описание классов в виде логической формулы, определяющей уровень налогов через функцию принадлежности, устанавливающую соотношение измеренных значений и качественных оценок. База правил включает правила следующего типа: $Z_1^{FS} = \{(\text{НДС} = \text{"очень большое"}) \& (\text{Налог на прибыль} = \text{"среднее"}) \& (\text{Подходный налог} = \text{"очень малое"}) \& (\text{Налог на имущество} = \text{"малое"}) \& (\text{Прочие налоги} = \text{"малое"})\}$.

Прогнозирование налоговых поступлений. На этапе прогнозирования налоговых поступлений в бюджет вычисляются значения приращения налогов за определенный период времени для субъектов каждого класса [6, с. 52]. Для вычисления прогнозных значений изменения налогов в зависимости от изменений внешней среды строится база знаний, в которой описываются экономические закономерности влияния внешних и внутренних

Таблица 2

Исходные данные и результаты моделирования

Показатель	Темп прироста, %				
	Группа 1	Группа 2	Группа 3	Группа 4	Группа 5
Исходные данные для моделирования					
Добавленная стоимость	252	137	170	191	130
Дебиторская задолженность	-20	-10	30	-26	160
Прибыль	120	60	160	-180	124
Средства на оплату труда	72	43	53	-0,41	67
Выручка	48	77	176	71	68
Себестоимость	39	0,38	77	37	17
Инфляция	6	6	6	6	6
Курс доллара	11	11	11	11	11
Цены на нефть	55	55	55	55	55
Среднесписочная численность работающих	-2	-14	-5	-46	-58
Среднегодовая стоимость имущества	1	-13	3	2	-19
Результаты моделирования					
Налог на добавленную стоимость	36	63,2	73,6	21,3	28,1
Налог на прибыль	25,2	31,5	36,1	2,3	9,2
Подходный налог	26,3	33,1	45	41,3	51,2
Налог на имущество	21,2	20,1	13,6	12,2	3,2
Прочие	36,2	31,2	24,1	12,3	41,2

факторов, характеризующих налогооблагаемую базу, на значение налогов.

В основу прогнозирования положена модель, использующая методы нечетких множеств, которая адекватна поставленной задаче, так как оперирует количественными и качественными оценками переменных. Прогнозирование приращения налогов основано на экспертных оценках, законах развития экономических систем и эвристических рассуждениях. Фрагмент базы правил, построенных по принципу "если-то", представлен ниже.

Если дебиторская задолженность = "растет" & инфляция = "растет" **То** НДС = "падает". **Если** курс доллара = "растет" & цены на нефть = "растут" & ДС = "растет" **То** НДС = "растет". **Если** денежные средства на оплату труда = "растут" & "среднесписочная численность работающих" = "растет" **То** ПН = "растет".

Исходные данные и результат моделирования показаны в табл. 2. Исходные данные для моделирования представлены в виде изменений факторов-показателей по группам субъектов. С учетом найденных значений приростов налогов по каждой группе субъектов рассчитываются прогнозные значения налоговых поступлений по группам и видам налогов.

Затем необходимо провести коррекцию налоговых сумм путем учета влияния изменений межбюджетных отношений в виде изменений ставок зачисления налоговых доходов по разным уровням бюджетов. Были получены оценки значений налоговых доходов (табл. 3) за предыдущий период времени ($n - 1$) для оценки точности прогнозирования, а также на два следующих периода ($n + 1$, $n + 2$).

Анализ результатов моделирования показал, что ошибка прогноза по описанной модели не превышает 8 %, что является доказательством высокой точности полученных оценок. Была проведена серия экспериментов по прогнозированию налоговых поступлений на основе эконометрических методов. Результаты прогнозирования по многофакторной регрессионной модели показали точность аппроксимации 19 % [9, с. 70]. Сопоставление результатов использования этих двух методов получения прогнозных оценок налоговых доходов доказывает преимущество предлагаемого подхода, основанного на применении нейросетевых и нечетких моделей.

Таблица 3

Фактические и оцененные по модели значения налоговых доходов

Налоговые доходы	Период времени (n)			
	$n - 1$	n	$n + 1$	$n + 2$
Фактическое значение, млн руб.	1,024	2,177	3,516	2,745
Оцененное значение, млн руб.	0,996	2,296	3,715	2,530
Относительное отклонение, %	2,78	5,48	5,65	7,82

Заключение

Рассмотрены базовые принципы разработки информационной технологии бюджетно-налогового моделирования и прогнозирования, представляющие собой интегрированную совокупность интеллектуальных методов и моделей анализа разнородных сведений о системе, учитывающие неопределенность, риск и обеспечивающие высокую степень адекватности модельного описания системы. Предлагается технология бюджетно-налогового моделирования и прогнозирования, которая формирует более точные прогнозы налоговых поступлений в бюджет, во-первых, за счет использования, кроме фактической информации о собранных налогах объектов налогообложения, дополнительной содержательной информации о закономерностях формирования и изменения налоговой структуры предприятий каждого класса (число классов может меняться); об экономических и институциональных изменениях внешней среды; о взаимном влиянии факторов на налогооблагаемую базу; во-вторых, за счет применения методов обработки адекватных природе данных в системе налогообложения.

Список литературы

1. Барулин С. В., Ермакова Е. А., Степаненко В. В. Налоговый менеджмент. М.: Дашков и К, 2012. 331 с.
2. Барулин С. В., Барулина Е. В. Парадоксы российской системы налогообложения // Финансы и кредит. 2012. № 12. С. 15–23.
3. Варакса Н. Г. Моделирование налоговой политики и налогового механизма в рамках стратегии экономического развития государства // Управленческий учет. 2013. № 8. С. 75–84.
4. Исмагилова Л. А., Орлова Е. В. Эффективное управление разноуровневыми экономическими системами. М.: Машиностроение, 2012. 382 с.
5. Исмагилова Л. А., Орлова Е. В. Моделирование эффективной налоговой нагрузки по группам субъектов налогообложения // Известия РАН. Теория и системы управления. 2007. № 1. С. 113–117.
6. Исмагилова Л. А., Орлова Е. В. Информационная технология оценки и прогнозирования налогового потенциала // Экономика и управление. 2001. № 1. С. 50–53.
7. Нейросетевое математическое моделирование в задачах ранжирования и кластеризации в бюджетно-налоговой системе регионального и муниципального уровней. Уфа: РИЦ БГУ, 2011. 221 с.
8. Орлова Е. В. Экономико-математический инструментальный управления экономической системой в условиях неопределенности. Уфа: Изд. УГАТУ, 2012.
9. Орлова Е. В., Исмагилова Л. А. Управление налогообложением региональной системы на основе интеллектуальных методов и моделей // Нейрокомпьютеры: разработка и применение. 2013. № 3. С. 68–72.
10. Скляр А. В., Цацулин А. Н. Моделирование налоговых поступлений в региональные бюджеты // Управленческое консультирование. 2012. № 1. С. 100–111.
11. Татаркин А. И., Пыхова И. А. Методологические принципы и основные направления бюджетной и налоговой политики государства на территории в условиях административно-территориальной реформы // Известия Уральского государственного экономического университета. 2004. № 9. С. 3–13.
12. Троянская М. А. Моделирование временных рядов налоговых поступлений адаптивными методами // Вестник Оренбургского государственного университета. 2006. № 8. С. 268–274.

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ОБРАЗОВАНИИ INFORMATION TECHNOLOGIES IN EDUCATION

УДК 004.738.5

А. Д. Иванников, д-р техн. наук, зам. директора по научной работе, e-mail: ADI@iprm.ru,
Институт проблем проектирования в микроэлектронике РАН

Тематические интернет-порталы как средство агрегации электронного контента в заданной предметной области

Анализируются методы создания тематических интернет-порталов для эффективного доступа к электронным ресурсам заданной тематики. Осуществляется сравнение тематических порталов с универсальными поисковыми системами с точки зрения релевантности получаемой выборки ресурсов. Рассматриваются принципы эффективной рубрикации агрегируемых ресурсов, дающей возможность организовать атрибутно-контекстный поиск. В качестве примера агрегации электронных ресурсов с помощью тематических порталов описывается система федеральных образовательных порталов.

Ключевые слова: агрегация электронного контента, тематические порталы, атрибутный поиск, рубрикация ресурсов, образовательные порталы

A. D. Ivannikov

Subject Devoted Internet Portals as an Instrument for Digital Content Aggregation

The methods to create subject devoted web portals for efficient access to electronic resources of specified domain are analyzed. Subject portals are compared with universal search engines in terms of obtained resources list relevance. The principles of aggregated resources effective categorization that make the attribute — contextual search possible are discussed. As an example, the system of Russian Federal educational portals is described.

Keywords: aggregation of digital content, subject devoted portals, attribute search, categorization of resources, educational portals

Введение

В настоящее время наблюдается ярко выраженная тенденция отхода от представления информации на бумаге к ее электронному представлению. В метро посмотришь, все в основном электронные книги читают, бумажные реже. Электронные образовательные ресурсы не заменили обычные учебники, но существенно потеснили их. Все большее число научных журналов имеют электронные версии, а целый их ряд выходит только в электронном виде. И если вам нужно узнать что-то, вы уже не звоните в справочную, а входите в интернет и ищите информацию там. Именно там, в интернете, и располагается большинство электронных ресурсов. Но поскольку интернет — это паутина сайтов, наброшенная на весь земной шар, то весьма актуальным вопросом становится осуществление эффективно поиска нужных электронных ресурсов.

Тематические порталы и их основные свойства

В интернет найти что-то нужное, причем не несколько полезных ресурсов, а весь их набор, кластер, относящийся к делу, где было бы представлено все или большинство необходимых материалов или идей, довольно сложно.

Универсальные поисковики — Google, Yandex, Yahoo — используют семантический или контекстный поиск, т. е. поиск по ключевым словам. Они выдают список ссылок на имеющиеся в интернете электронные ресурсы, использующие заданные ключевые слова, причем на первых местах стоят те, которые были наиболее востребованы. Понятие *востребованность* в разных поисковиках интерпретируется по-разному, но суть одна и та же: первыми стоят те, к которым больше обращаются или на которые больше ссылаются.

Есть и другой подход. Назовем его — *агрегация электронных ресурсов* (электронного контента) для какого-то применения, анализа или использования. Для этого необходимо, чтобы в одном месте, более точно — в одной точке входа, можно было бы найти всю информацию, находящуюся в интернете. А лучше, чтобы в эту точку входа сложили все, что к этой теме относится, даже если в интернете этого не было. Это тематические порталы. В тематический портал заносятся только электронные ресурсы (или информация о них), в явном виде относящиеся к тематике портала. Именно в этом его существенное отличие от нетематических порталов.

Если это образовательные порталы, то их широко используют в современном образовательном процессе [1, 2], причем только в этих порталах находятся образовательные ресурсы; если порталы по научной деятельности, то в них находятся только научные статьи и результаты научных исследований.

А в тематическом портале имеются ресурсы только по данной тематике. Кроме того, там обычно возможен поиск по атрибутам. В отобранных ресурсах с указанными атрибутами можно было бы использовать поиск по ключевому слову. В этом случае степень релевантности выданных ссылок на электронные ресурсы явно будет выше и перечень предлагаемых ресурсов будет явно короче.

Таким образом, отличительными особенностями тематических порталов, обеспечивающих агрегацию электронного контента заданного назначения, является следующее:

- поиск осуществляется исключительно по электронным ресурсам заданной тематики;
- особенностью таких систем является атрибутивный и контекстный поиски в любом сочетании;
- существенно более высокая релевантность найденных электронных ресурсов;
- отсутствие необходимости просмотра большого числа ссылок на электронные ресурсы в целях выбора действительно релевантных ресурсов;
- наличие более совершенных сервисов, отвечающих целям работы с агрегированным контентом заданного назначения.

Поясним последнюю упомянутую особенность. Обычно на тематических порталах кроме электронных ресурсов основного назначения — непосредственно образовательных ресурсов или научных публикаций — имеются еще смежные вспомогательные материалы: образовательные стандарты и рабочие программы в случае образовательного портала или перечень рейтинговых научных журналов в случае результатов научно-исследовательских работ. И между электронными ресурсами основного назначения и вспомогательными базами данных формируется система взаимных ссылок, с использованием которых реализуются дополнительные сервисы, востребованные при работе с агрегированным контентом.

Формирование множества электронных ресурсов тематических порталов

Для поисковых систем ответ на вопрос о том, как в их поле зрения попадают электронные ресурсы, очевиден: их в интернете отыскивают с помощью поисковика, причем он ищет все, что есть в интернете по данному ключевому слову.

Для тематических порталов дело обстоит иначе. Электронные ресурсы должны из интернета попасть в поле зрения тематического портала, да еще и с прикрепленными атрибутами в соответствии с принятой на тематическом портале классификацией. Более того, желательны специальные меры по выставлению в ареале этих тематических порталов (а следовательно, и в интернете) дополнительных электронных ресурсов в целях охвата не только того, что уже есть в интернете по данной тематике, а всей целевой области тематического портала в целом. Причем надо подчеркнуть, что тематические порталы работают не просто с электронными ресурсами, а с ресурсами, привязанными к разделам принятого на тематическом портале классификатора или классификаторов. Только в этом случае можно организовать дающий большие преимущества атрибутивный поиск.

Итак, при создании какого-либо тематического портала необходимо разработать систему классификаторов (рубрикаторов) с перечнем рубрик, собрать, сгруппировать целевые электронные ресурсы и отнести каждый электронный ресурс к какой-либо рубрике всех используемых классификаторов, причем электронный ресурс может быть отнесен одновременно к нескольким рубрикам классификатора. Так, какой-либо учебный материал может быть полезен при изучении как географии, так и физики.

Сбор электронных образовательных ресурсов в большой степени различается на начальной стадии создания тематического портала — начальной агрегации электронного контента, а также в процессе функционирования и развития. Начальная агрегация электронного контента предполагает большую долю ручного труда, что несмотря на существенную автоматизацию этого процесса требует временных и финансовых затрат. В зависимости от тематики портала процесс агрегации естественно различается. Но в любом случае на начальном этапе создания агрегация должна осуществляться как можно большими порциями для скорейшего создания некоторой "критической" массы, уже интересной для потенциальных пользователей. Так, в случае создания образовательных порталов www.edu.ru и window.edu.ru [3] первоначальный набор образовательных ресурсов формировался путем просмотра сайтов вузов и других порталов и сайтов, на которых существовали целые разделы ресурсов, так или иначе привязанных к образованию. При этом осуществлялся анализ и отбор необходимых электронных ресурсов.

После создания необходимой критической массы агрегированных электронных ресурсов в процессе функционирования и развития задача пополнения набора электронных ресурсов в целях охвата всей предметной области портала, а также включения в агрегированное множество новейших ресурсов остается актуальной. В случае высокой авторитетности портала, что обуславливается большим объемом качественных релевантных ресурсов на портале и рядом дополнительных факторов, авторы или организации, генерирующие соответствующие электронные ресурсы, будут заинтересованы предложить свои ресурсы для публикации на портале.

После аккумуляции на тематическом портале достаточно большой критической массы качественных агрегированных ресурсов для дальнейшего расширения перечня ресурсов, кроме, естественно, продолжения поиска новых качественных ресурсов в интернете, могут быть использованы следующие методы:

1. Создание авторитетной и эффективно функционирующей редакции портала.

2. Принятие мер к повышению рейтинга тематического портала в соответствующих разделах наиболее используемых поисковых систем.

3. Регистрация тематического портала как средства массовой информации или как научного или методического электронного журнала с соответствующим признанием размещаемых на портале публикаций.

4. В случае высокого рейтинга тематического портала признание соответствующими государственными органами необходимости обязательной публикации электронных ресурсов открытого доступа, разработанных по государственным контрактам, на соответствующих тематических порталах.

5. Заключение договоров и проведение совместных работ редакции тематического портала и редакций родственных или смежных порталов и сайтов по включению в каталог тематического портала материалов или информации о материалах родственных или смежных порталов и сайтов.

6. Сотрудничество с оргкомитетами наиболее авторитетных конференций по тематике портала в части публикации на портале наиболее интересных материалов.

7. Создание на портале форумов, блогов и гостиных для обсуждения публикуемых на портале материалов и формирование тематических интернет-сообществ в соответствии с разделами тематического портала.

8. Проведение круглых столов и секций на соответствующих конференциях по поиску и использованию материалов тематического портала, его популяризации.

9. Подготовка, печать, рассылка и раздача бумажных методических пособий по поиску и использованию тематических порталов.

В обязательном порядке необходимо предусмотреть простой и понятный способ предложения автором или организацией своего электронного ресурса для публикации. В частности, каким бы сложным не был какой-либо из классификаторов, используемых на тематическом портале, автору или его представителю должно быть достаточно из спускающегося меню выбрать один или несколько рубрик, которые становятся атрибутами предлагаемого к публикации электронного ресурса.

Хранить на портале сами ресурсы или только метаданные?

Итак, тематический портал — это та сущность в интернете, выйдя на которую и задав каким-то образом условия отбора ресурсов из всего множества имеющихся на портале, пользователь получает доступ к отобранным электронным ресурсам. При этом пользователю все равно, сам ресурс находился тут же на сервере тематического портала или был считан с какого-то другого сайта, находящегося географически совсем в другом месте.

Таким образом, существуют следующие подходы к формированию тематического портала:

- электронные ресурсы можно хранить на самом тематическом портале;
- на самом тематическом портале можно хранить только каталог, т. е. систему карточек или других видов метаописаний электронных ресурсов со ссылками на те адреса в интернете, где хранятся сами электронные ресурсы;
- можно сочетать перечисленные выше два подхода.

Выбор одного из вышеперечисленных вариантов определяется ментальными предпочтениями сообщества, юридическими и техническими факторами. Так, не всегда автор соглашается размещать разработанный им ресурс в централизованной библиотеке портала, а предпочитает разместить его на сайте своего, например учебного, заведения. В этом случае на тематическом портале размещается метаописание данного ресурса, включая ссылку на его расположение. Кроме того, если локальных пользователей разработанного ресурса достаточно количество, то размещение ресурса в какой-то центральной удаленной точке повышает дальний трафик, например до Москвы, который дороже локального трафика.

Если электронный ресурс размещается в интернете, то за соблюдение авторских и имущественных прав отвечает владелец домена, на котором ресурс размещен. То есть при размещении электронного ресурса на самом тематическом портале именно редакция портала должна получить разре-

шение автора и/или владельца. Конечно, в каждом конкретном случае можно определить принадлежность прав и заключить соответствующий договор. Однако при достаточно большом количестве электронных ресурсов, скажем, порядка нескольких десятков тысяч ресурсов, объем задачи с учетом возможных вариантов определения прав и сроков становится очень большим.

Если на самом тематическом портале хранятся только метаданные, включающие ссылку на сами электронные ресурсы, то вопрос о правомерности открытого размещения в интернет электронных ресурсов перед редакцией тематического портала не стоит.

Рубрикаторы электронных ресурсов

При разработке тематического портала (или системы тематических порталов, см. ниже) важным вопросом является выбор набора классификаторов и их рубрик, в соответствии с которыми осуществляется атрибутивный поиск. Например, чтобы была возможность при задании соответствующего запроса к некоторому тематическому образовательному portalу получить набор электронных версий учебников для 10 класса средней школы по географии, необходимо, чтобы электронные ресурсы на этом портале имели классификаторы по уровню образования, по типу материала (учебник, учебное пособие, задачник, атлас, энциклопедия и т. д.) — и по предметам. Необходимо также наличие в соответствующих классификаторах рубрик: 10 класс, учебник и география. Именно удобство и эффективность работы пользователей при поиске ресурсов через тематический портал являются определяющими факторами при выборе набора классификаторов и их рубрик. В частности, предметные классификаторы, наиболее эффективные при реализации в тематических порталах, могут отличаться от классификаторов, общепринятых в библиотечном деле или в научно-технической информации.

Так, общепринятый для печатной литературы классификатор УДК скорее является классификатором системы знаний вообще и не учитывает особенности того или иного тематического портала. При работе в интернете пользователю удобно, чтобы в каждой рубрике было одинаковое по порядку число единиц хранения, чего при использовании УДК не получается. Кроме того, качественное представление рубрик УДК при создании электронного образовательного ресурса, особенно междисциплинарного характера, весьма сложно и в идеале должно осуществляться с привлечением квалифицированного библиотечного работника. А это вступает в противоречие с требованием простоты представления к публикации на тематическом портале созданного электронного ресурса. Другой широко

используемый классификатор ГРНТИ, который явился основой для разработки номенклатуры научных специальностей ВАКа, в общем-то подходит в качестве предметного рубрикатора образовательных ресурсов для высшего образования. А вот для образовательных ресурсов средней школы — не совсем.

Так, в школе имеется предмет физика, одним из разделов которой является механика, которая, в свою очередь, состоит из статики, кинематики и динамики. Вполне естественно было бы иметь в тематическом образовательном портале на первом уровне рубрику "Физика", которая на втором уровне включала бы рубрику "Механика", которая на третьем уровне включала бы три рубрики: "Статика", "Кинематика" и "Динамика". Но в ГРНТИ как "Физика", так и "Механика" стоят на первом уровне рубрикации. В рубрике "Механика" отнести куда-то "Статику", "Кинематику" и "Динамику" можно разве что в "Общую механику", а там разве что в рубрику уже третьего уровня "Механика точки, системы и твердого тела". Это явно не для классификации школьных образовательных ресурсов.

Так, при создании порталов "Российское образование" www.edu.ru и "Единое окно доступа к образовательным ресурсам window.edu.ru" предметный рубрикатор ресурсов [4] был разработан на основе действующих образовательных стандартов. В то же время на тех же порталах для образовательных ресурсов высшего образования был принят рубрикатор ГРНТИ. Таким образом, рубрикаторы для тематического портала, а обычно в этом случае используются несколько рубрикаторов, могут не являться независимыми. Так, в рассмотренном случае используемый предметный рубрикатор зависит от рубрики другого рубрикатора — рубрикатора "Уровень образования".

Другим примером взаимозависимости используемых в тематическом портале рубрикаторов является Федеральный портал по научной и инновационной деятельности www.sci-innov.ru [5]. На этом портале кроме научно-технической информации по приоритетным направлениям науки и техники имеются материалы по инфраструктуре поддержки инновационной деятельности, нормативно-правовой базе научно-технической и инновационной деятельности, информации о выставках и конференциях и др. Перечисленные выше разделы являются рубриками одного из используемых классификаторов, при этом в зависимости от рубрики этого классификатора используются те или иные рубрики других классификаторов. В этом случае взаимозависимость рубрикаторов проявляется гораздо более сильно, чем в случае упомянутых выше образовательных порталов.

Один портал или система порталов?

Более серьезно проблема выбора классификаторов встает, когда создается не тематический портал, а система тематических порталов.

Чтобы портал был действительно тематическим, имеющаяся информация должна достаточно полно охватывать все вопросы тематики. Наполнение, анализ и развитие информационного наполнения портала должна осуществлять редакция и сообщество, куда входят профессиональные специалисты по тематике портала, не только систематизаторы, но и обязательно предметники по тематике портала.

Однако требование профессионализма членов сообщества, осуществляющего смысловое редактирование и ведение портала, вступает в противоречие с достаточно полным охватом тематики имеющимся материалом. Например, если рассматривать образовательный портал, который должен включать электронные образовательные ресурсы по всем основным предметам, то можно сразу обнаружить, что даже только для среднего образования в России имеется сообщество преподавателей по физике, несколько сообществ по гуманитарным предметам, в частности, сообщество преподавателей русского языка и т. д. Поэтому логично, чтобы существовали не только разделы единого образовательного портала по физике, русскому языку и т. д., а отдельные порталы, скажем, по естественно-научным дисциплинам, гуманитарным дисциплинам, экономике и менеджменту и т. д. Особенно это требуется для высшего образования, поскольку курсы этого уровня образования близко смыкаются с научными исследованиями, где специализация еще более выражена.

Для каждого предметного сообщества существуют общепризнанные лидирующие организации, в которых работают известные авторитеты в своей области. Именно в таких организациях и целесообразно создавать тематические порталы с привлечением общепризнанных в сообществе авторитетов, естественно, если эти авторитеты поддерживают идеологию создания тематических порталов.

Такая система будет эффективной при наличии некоторой обобщающей структуры — портала, где будет храниться база данных метаописаний всех электронных ресурсов, создаваемых на всех порталах, в нашем примере — по физике, по русскому языку, по экономике и т. д. Назовем такие порталы вертикальными или предметными, а обобщающий портал горизонтальным. Горизонтальный портал кроме сбора всех метаописаний в единую базу данных и обеспечения атрибутно-контекстного поиска выполняет еще ряд функций для всей системы порталов в целом. При такой организации на всех порталах системы должны использоваться одинаковые форматы метаописаний и рубрикаторы.

С точки зрения автора статьи такая организация, в частности образовательных порталов, является наиболее эффективной.

Так, в 2001—2005 гг. в соответствии с Федеральной целевой программой "Развитие единой образовательной информационной среды" Министерства образования и науки Российской Федерации была создана и введена в эксплуатацию система федеральных образовательных порталов (см. рисунок на второй стороне обложки) [6], включающая:

- горизонтальный портал "Российское образование" www.edu.ru (Государственный НИИ информационных технологий и телекоммуникаций);
- вертикальные профильные порталы: "Экономика, социология, менеджмент" www.ecsocman.edu.ru (теперь www.ecsocman.hse.ru) (Высшая школа экономики), правовой портал "Юридическая Россия" www.law.edu.ru (Юридический факультет Санкт-Петербургского государственного университета), "Социально-гуманитарное и политологическое образование" (Российский университет дружбы народов), "Естественно-научный образовательный портал (физика, химия, биология, математика)" (Санкт-Петербургский институт точной механики и оптики); "Инженерное образование" (МВТУ им. Н. Э. Баумана); "Российский общеобразовательный портал" www.school.edu.ru (Издательство "Просвещение"); "Информационно-телекоммуникационные технологии в образовании" www.ict.edu.ru (Государственный НИИ информационных технологий и телекоммуникаций);
- вертикальные специализированные порталы: портал "Информационная поддержка ЕГЭ" www.ege.edu.ru; "Российский портал открытого образования" www.openet.edu.ru (теперь www.openet.ru); "Образовательный портал по поддержке процессов обучения в странах СНГ"; портал "Здоровье и образование"; портал "Международное образование".

Основным содержанием горизонтального портала "Российское образование" [7] является библиотека метаописаний, т. е. сформированных определенным образом библиографических карточек электронных образовательных ресурсов, содержащих кроме данных об авторах, аннотаций и ключевых слов также значения атрибутов из рубрикаторов по предметной области, типам ресурсов, уровням образования и аудитории, для которой электронный ресурс предназначен прежде всего, а также указатель — адрес в интернете, по которому размещается сам электронный образовательный ресурс.

На вертикальных профильных порталах размещаются библиотеки метаописаний электронных образовательных ресурсов по профилю порталов, причем метаописания имеют ту же структуру, что

и метаописания на горизонтальном портале "Российское образование". На вертикальных специализированных порталах размещаются библиотеки аналогичных по структуре метаописаний электронных ресурсов, связанных с решением той конкретной задачи, в соответствии с которой создавались специализированные порталы.

Кроме того, на всех порталах размещается и другая информация, связанная с задачами и профилем созданных порталов. На каждом портале существует также и библиотека полнотекстовых образовательных ресурсов, в этом случае указателем местоположения ресурса в метаописании является адрес в библиотеке самого портала.

Однако для настоящего рассмотрения важно, что форматы и рубрикаторы метаописаний электронных ресурсов на всех порталах одинаковы. Это позволило осуществить автоматическую репликацию сформированных на вертикальных порталах ресурсов соответствующего профиля в общую базу данных метаописаний горизонтального портала "Российское образование" и, соответственно, атрибутивный поиск электронных образовательных ресурсов по всем метаописаниям системы образовательных порталов. Более того, если на порталах образовательных учреждений были бы использованы аналогичные метаописания и рубрикаторы, то при автоматическом добавлении этих метаописаний в базу данных портала "Российское образование" атрибутивный поиск осуществлялся бы и по порталам образовательных учреждений.

Портал "Российское образование" при рассматриваемой схеме работы обеспечивает атрибутивно-контекстный поиск по всем метаописаниям системы образовательных порталов, а также обеспечивает постоянную проверку доступности электронных образовательных ресурсов в соответствии с указателями — интернет-адресами, имеющимися в метаописаниях ресурсов.

Такая организация сбора данных об имеющихся в интернете образовательных ресурсах, так же как и электронных ресурсах тематических порталов другого назначения, является наиболее эффективной. Сложностью такого подхода является необходимость использования всем образовательным сообществом (или каким-либо сообществом другого профиля) одного и того же стандарта на метаописания ресурсов.

Также нелегкой задачей является необходимость убедить редакцию какого-либо профильного портала передавать метаописания своих электронных ресурсов в общую базу данных метаописаний,

хотя при этом увеличивается посещаемость самого профильного портала. В настоящее время описанная система федеральных образовательных порталов распалась. Ряд порталов из перечисленных прекратили свое существование. Порталы, для которых в списке приведены интернет-адреса, в настоящее время существуют и являются востребованными. К сожалению, между ними не осуществляется обмена метаописаниями, т. е. каждый портал существует сам по себе и решает собственные задачи. Однако одним из важных преимуществ всех из них является возможность атрибутивно-контекстного поиска электронных образовательных ресурсов.

Заключение

Тематические порталы являются эффективным средством агрегации электронных ресурсов по определенной тематике, которая может быть достаточно широкой (образовательная тематика является хорошим примером). Техники использования метаданных электронных интернет-ресурсов получают все большее развитие. Универсальные поисковики также начинают их использовать. Представляется, что в недалеком будущем мы увидим использование атрибутивно-контекстного поиска и универсальными поисковыми порталами.

Список литературы

1. Тихонов А. Н., Иванников А. Д. Технологии дистанционного обучения в России // Высшее образование в России. 1994. № 3. С. 3—10.
2. Усков В. Л., Иванников А. Д., Усков А. В. Перспективные технологии для электронного образования // Информационные технологии. 2007. № 2. С. 32—38.
3. Булгаков М. В., Гридина Е. Г., Иванников А. Д., Старых В. А. Федеральная система информационно-образовательных ресурсов // Информационные ресурсы России. 2009. № 2. С. 25—27.
4. Булгаков М. В., Гридина Е. Г., Иванников А. Д., Тихонов А. Н. Реализация каталога Федерального портала "Российское образование" // Дистанционное и виртуальное обучение. 2004. № 6. С. 25—40.
5. Булгаков М. В., Иванников А. Д., Ковальский М. С., Носов В. П., Рогачев А. А., Таран А. Н. Федеральный портал по научной и инновационной деятельности // Сб. "Интернет-порталы: содержание и технологии". М.: Просвещение, 2007. Вып. 4. С. 323—344.
6. Гридина Е. Г., Иванников А. Д., Булгаков М. В., Чиннова И. И., Сигалов А. В. Система Федеральных образовательных порталов: 3 года в Интернет // Открытое образование. 2005. № 1. С. 35—54.
7. Абрамов А. Г., Булакина М. Б., Булгаков М. В., Внотченко С. С., Зыбарев Е. Ю., Иванников А. Д., Сигалов А. В. Информационные разделы и сервисы Федеральных образовательных порталов: опыт разработки и сопровождения // Информатизация образования и науки. 2009. № 3. С. 51—64.

УДК 004.413.2

А. В. Ивашенко, д-р техн. наук, доц., e-mail: anton.ivashenko@gmail.com, **Д. Г. Пейсахович**, аспирант, Самарский государственный аэрокосмический университет

Метод проактивной диспетчеризации в информационной среде для транспортных операторов 5PL

С учетом современных требований к деятельности оператора 5PL предлагается и обосновывается новый подход к автоматизации планирования грузоперевозок в условиях единого информационного пространства. Описывается метод проактивной диспетчеризации и многофакторная информационная среда, позволяющая реализовать данный метод управления. Практическая реализация состоит в построении алгоритмов информационного управления обстоятельствами взаимодействия пользователей логистических услуг.

Ключевые слова: транспортная логистика, 5PL, интеллектуальные системы планирования, проактивная диспетчеризация, мультиагентные системы, кондиционное управление

A. V. Ivaschenko, D. G. Peysakhovich

Proactive Dispatching in the Integrated Information Space for 5PL Providers in Transportation Logistics

Considering the modern trends in transportation logistics in this paper it is proposed and validated a new approach for cargo flow scheduling automation in the integrated information space of 5 Party Logistics operator. There is described the technology of proactive dispatching and the multi-actor integrated information space structure that allows providing such an approach for transportation logistics. Its implementation is based on the information management technology for interaction of the users of logistics services.

Keywords: transportation logistics, 5PL, intelligent scheduling systems, proactive dispatching, multi-agent systems, scheduling by conditions

Введение

Развитие информационно-коммуникационных технологий позволяет существенно повысить точность планирования и контроля исполнения заказов на перевозку грузов наземным транспортом. Оснащение водителей грузовых транспортных средств устройствами спутниковой навигации и терминалами с выходом в Интернет обеспечивает техническую возможность информационного взаимодействия с диспетчерами в режиме реального времени, что определяет новые требования к интеллектуальным системам планирования транспортных ресурсов. В связи с этим в последнее время растет популярность открытых порталов, предоставляющих услуги как транспортным компаниям, так и экспедиторам, по получению, планированию и диспетчеризации перевозок. Эти порталы реализуются

с помощью программных платформ, которые функционируют по принципу SAAS (Software as a service), что делает их открытыми и доступными для небольших транспортных компаний и отдельных водителей, которые за сравнительно малую плату получают доступ в единое информационное пространство международного уровня [1].

Такая тенденция привела к появлению нового типа логистических операторов 5PL, оказывающих комплекс логистических услуг [2] за счет организации единого информационного пространства среди заказчиков (грузоотправителей и грузополучателей) и исполнителей — транспортно-экспедиционных компаний. Наименование 5PL (PL — Party Logistics) дается по аналогии с услугами 3PL и 4PL [3]. Основное отличие оператора 5PL состоит в том, что у такой компании может не быть собственных материальных или финансовых ресурсов, а услуги

по организации перевозок оказываются путем накопления и обработки информации об участниках рынка и построения для них оптимальных логистических цепочек.

Так как оператор 5PL может не иметь собственных ресурсов и работает в основном с внешними перевозчиками, возникает новая проблема распределения транспортных ресурсов в условиях неопределенности их количества и времени доступности, которая не решается в полной мере в существующих автоматизированных системах управления транспортом. Таким образом, актуальной является задача разработки новых методов и средств интеллектуальной диспетчеризации, одно из решений которой предлагается в данной статье.

Современные подходы по автоматизации управления грузовым транспортом

Задача построения Интернет-портала для транспортно-экспедиционных компаний обычно рассматривается в контексте применения WEB-технологий и компьютерных сетей. В качестве примера одного из наиболее известных порталов можно привести систему "АвтоТрансИнфо" [4]. Также весьма распространены автоматизированные системы планирования транспорта, в которых ограничивается фиксированный пул (перечень) ресурсов и интервал времени (горизонт) планирования. Основное внимание разработчиков систем планирования в этой области сводится к решению задачи оптимизации загрузки транспортных ресурсов, для чего используют традиционные подходы [5], основанные на точном планировании или (в случае большого объема данных или необходимости функционирования в реальном времени) интеллектуальных эвристических алгоритмов [6, 7].

При решении данной задачи строится централизованный план, оптимизированный по группе критериев. При отклонении от данного плана диспетчеры пытаются либо наверстать сроки, либо перестроить расписание. Поскольку делать это в режиме реального времени сложно, актуальной считается задача сопоставления и совместного использования режимов планирования: *batch* (пакетный режим) и *incremental* (с приращениями). Эти режимы характеризуют разные подходы к обработке событий, требующих пересмотра планов, например, появление нового заказа, который должен быть выполнен в течение времени, на которое ресурсы уже запланированы, или существенное отклонение от выполнения запланированного заказа, связанное с поломкой либо опозданием транспорта. В первом случае происходит накопление этих событий и пересмотр всего плана целиком, а во втором случае будут локальные изменения части плана, обеспечивающие сохранение его в целом.

Поскольку деятельность оператора 5PL требует постоянной обработки множества событий, поступающих в режиме реального времени, достаточно интересен опыт разработки систем планирования в режиме *incremental*. Здесь наиболее популярны и перспективны мультиагентные технологии [8—10], которые позволяют реализовать локальные переговоры агентов ресурсов и заказов и таким образом обеспечить пересмотр части расписания без полного перепланирования. Это в большинстве случаев позволяет сохранить значения интегральных показателей качества расписания, используемых в качестве оптимизационных критериев.

Совместное функционирование группы транспортных компаний, часть из которых является экспедиторами и передает исполнение заказов подрядчикам, а другая часть пользуется услугами оператора 5PL эпизодически в случае недозагрузки ресурсов собственными заказами, не позволяет на этапе планирования однозначно определить перечень доступных ресурсов и возможности по их загрузке. Вместе с тем решения по назначению заказов должны быть согласованы с конечными исполнителями.

В этих условиях задача распределения транспортных ресурсов сводится к задаче обеспечения эффективного обмена информацией между многими участниками процесса планирования — лицами, принимающими решения, в ходе которого за определенное время обеспечивается выработка общего согласованного расписания. В этих условиях могут показать хорошие результаты принципы управления, основанные на организации P2P взаимодействия и проведения серий аукционов [11]. Для реализации информационного управления в мультиагентной среде хорошие результаты показывает метод кондиционального управления [12], реализующий стратегии управления содержимым и временем получения информации. Данный метод использует подходы, широко применяющиеся в социальных сообществах сети Интернет и ориентированные на сохранение определенного уровня внимания (интереса) пользователей к общей тематике или процессу общения, что соответствует специфике поставленной задачи.

Определение проактивной диспетчеризации

Рассмотрим жизненный цикл заказа на транспортировку грузов w_j , $i = 1 \dots N_w$, — период времени с момента его поступления до исполнения. Обозначим участников жизненного цикла заказов (грузоотправителей, грузополучателей, экспедиционные и транспортные компании) u_j , $j = 1 \dots N_u$. Каждую такую компанию будем называть актором — субъектом, представляющим интересы соответствующей компании и выступающим в качестве лица, принимающего решения по обработке заказов.

Обозначим $c_{i,j}$ стоимость исполнения заказа w_i актором u_j , а $\Delta t_{i,j}$ продолжительность его выполнения при полной загрузке ресурсов (один актор не может выполнять одновременно несколько заказов, один заказ не может выполняться несколькими акторами).

Опишем следующие события жизненного цикла заказа w_i на интервале времени $(T_{\text{нач}}, T_{\text{зав}})$ с помощью булевых переменных:

$\varepsilon_i = \varepsilon(w_i, t_i^*) \in \{0, 1\}$ — появление заказа w_i в системе в момент времени t_i^* ;

$e_{i,j} = e(w_i, u_j, t_{i,j}) \in \{0, 1\}$ — предложение заказа w_i актору u_j в момент $t_{i,j}$;

$e'_{i,j} = e'(w_i, u_j, t'_{i,j}) \in \{0, 1\}$ — выбор (назначение) w_i актору u_j в момент $t'_{i,j}$.

Значение 0 соответствует отсутствию события, время наступления такого события будет равно $T_{\text{зав}}$.

Рассмотрим задачу распределения ресурсов, которую решает оператор 5PL. Деятельность оператора связана с поиском транспортных ресурсов сторонних компаний для выполнения поступающего потока заказов, при этом оплата услуги по планированию каждого заказа может быть фиксированной или зависеть от его стоимости. Наиболее высокий доход оператора будет при максимальном числе успешно обработанных (т. е. переданных исполнителям) заказов и их суммарной стоимости. Поскольку ресурсы (водители и транспортные средства) предоставляются в основном сторонними компаниями на определенный период времени, не все заказы могут быть обработаны в срок: возникают отказы или существенные задержки. Для того чтобы их уменьшить необходимо обеспечить привлекательность сервиса для транспортных компаний, которая состоит в построении "выгодных" маршрутов за счет консолидации заказов, снижения "холостого пробега", простоя и планирования "кругорейсов" (обратной загрузки).

Специфика деятельности оператора 5PL не позволяет строить согласованное расписание на всем горизонте планирования. Поэтому процесс распределения заказов более соответствует задаче диспетчеризации, которая рассматривает не процесс планирования, но процесс управления распределением ресурсов в реальном времени. При определении цели диспетчеризации обычно указывают обеспечение согласованной работы за счет равномерности, непрерывности, ритмичности и экономичности выполнения всех процессов. Диспетчеризация призвана обеспечить высокий уровень адаптивности посредством оперативного регулирования бизнес-процесса, нарушенного какими-либо внешними событиями.

Наиболее близкой к решаемой задаче является задача о назначениях в открытой форме. Модифи-

цировав эту задачу при условии распределения событий по времени, получим

$$\begin{aligned} & \sum_{i=1}^{N_w} \sum_{j=1}^{N_u} e'(w_i, u_j, t'_{i,j}) c_{i,j} \rightarrow \min; \\ & \sum_{j=1}^{N_u} e'(w_i, u_j, t'_{i,j}) = 1, i = 1 \dots N_w; \end{aligned} \quad (1)$$

$$\sum_{i=1}^{N_w} \sum_{j=1}^{N_u} \varepsilon(w_i, t_i^*) e'(w_i, u_j, t'_{i,j}) \cdot (t'_{i,j} - t_i^*) \rightarrow \min. \quad (2)$$

Задача (1), (2) описывает поиск наиболее эффективного распределения ресурсов. Будем считать стоимость выполнения работы, равносильной доходу актора-исполнителя. При этом для транспортных компаний-исполнителей справедлива следующая цель:

$$\sum_{i=1}^{N_w} \sum_{j=1}^{N_u} e'(w_i, u_j, t'_{i,j}) c_{i,j} \rightarrow \max. \quad (3)$$

Каждому актору относительно потока заказов $e(w_i, u_j, t_{i,j})$ нужно выбрать такую стратегию $e'(w_i, u_j, t_{i,j})$, при которой выполняется (3). Соответственно, центру относительно потока заказов $\varepsilon(w_i, t_i^*)$ нужно выбрать такую стратегию (план предложения) $e(w_i, u_j, t_{i,j})$ для нескольких u_j с учетом (2), при которой выполнится (1).

Частная цель каждого актора для заданного потока $e(w_i, u_j, t_{i,j})$ может быть представлена в следующем виде:

$$\begin{aligned} \forall u_j : & \sum_{i=1}^{N_w} \sum_{k=1}^{N_u} e(w_i, u_j, t_{i,j}) e'(w_i, u_j, t'_{i,j}) \times \\ & \times (1 - e'(w_i, u_k, t'_{i,k})) c_{i,j} \rightarrow \max. \end{aligned} \quad (4)$$

Допустим, что в процессе диспетчеризации актор начинает исполнять заказ сразу после получения, т. е. справедливо ограничение

$$\begin{aligned} \forall u_j : & \sum_{i=1}^{N_w} \sum_{j=1}^{N_w} e'(w_i, u_j, t'_{i,j}) \cdot e'(w_i, u_j, t'_{i,j}) \times \\ & \times (1 - \theta(t'_{i,j} - t'_{i,j}) \cdot \theta(t'_{i,j} + \Delta t_{i,j} - t'_{i,j})) = 0, \end{aligned}$$

где $\theta(x)$ — ступенчатая функция Хэвисайда:

$$\theta(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0. \end{cases}$$

Назовем задачу (4) задачей *проактивной диспетчеризации*. Ее основной особенностью является выделение целей центра (оператора 5PL) и акторов (исполнителей), что позволяет реализовать управ-

ление распределением ресурсов, учитывая фактор времени, с помощью мультиагентной технологии. Решения о выборе заказов $e'(w_p, u_j, t'_{i,j})$ распределены на горизонте планирования и зависят от событий появления заказов $e(w_p, u_j, t_{i,j})$.

Такой подход позволяет декомпозировать задачу проактивной диспетчеризации: решить задачу оптимизации относительно каждого актора, которая с учетом неопределенности будет выражаться в стратегии взаимодействия (в частности, может быть применена игровая стратегия); построить систему управления взаимодействием акторов с учетом фактора времени, которая бы позволила обеспечивать выполнение целевых функций (1), (2) и (3) в условиях конкуренции.

Стратегии проактивной диспетчеризации

Пусть поступающие заказы сразу предлагаются всем акторам.

Рассмотрим задачу (3) в условиях убывающей стоимости заказов (каждый следующий заказ дешевле предыдущего):

$$w_p, \forall u_j; t_{i+1,j} > t_{i,j} : c_{i+1,j} < c_{i,j}$$

Учитывая составляющие суммы (3) наилучшей стратегией каждого актора будет при соблюдении (4) следующая:

$$\forall u_j, e(w_p, u_j, t_{i,j}), e'(w_p, u_j, t'_{i,j}) : (t'_{i,j} - t_{i,j}) \rightarrow \min. (5)$$

При этом соблюдается минимум суммарного времени жизни заказов

$$\sum_{i=1}^{N_w} \sum_{j=1}^{N_u} e(w_p, u_j, t_{i,j}) \cdot e'(w_p, u_j, t'_{i,j}) \cdot (t'_{i,j} - t_{i,j}) \rightarrow \min. (6)$$

Следовательно, соблюдается (2), а для соблюдения (1) достаточно выбирать наиболее дешевый ресурс среди свободных на момент появления заказов и формировать для этого актора $e(w_p, u_j, t_{i,j})$.

Эта стратегия соответствует "жадному" потреблению заказов и часто оправдывает себя, поскольку относительный рост стоимости обычно всегда сопровождается появлением новых возможных исполнителей и ростом конкуренции. Актор с такой стратегией не будет выжидать более выгодного заказа, если он свободен и у него есть доступный заказ. Учитывая (4), для этого случая критерий (показатель) эффективной работы актора будет выглядеть как минимум времени простоя акторов. Для упо-

рядоченных последовательно событий (для каждой пары $e_{i,j}, e_{i+1,j} : t'_{i+1,j} > t'_{i,j}$):

$$\sum_{i=1}^{N_w-1} e'(w_p, u_j, t'_{i,j}) \cdot e'(w_{i+1}, u_j, t'_{i+1,j}) \times (t'_{i+1,j} - \Delta t_{i,j} - t'_{i,j}) \rightarrow \min. (7)$$

Рассмотрим задачу (4) в условиях нарастающей стоимости заказов:

$$w_p, \forall u_j; t_{i+1,j} > t_{i,j} : c_{i+1,j} > c_{i,j}$$

В этом случае увеличение времени ожидания каждого заказа приведет к повышению суммарной стоимости, однако общий доход может быть снижен вследствие роста суммарного времени простоя.

На практике события поступления заказов обычно подчиняются определенному закону распределения: существуют интервалы времени высокой активности заказчиков и интервалы времени жесткой конкуренции исполнителей, которые обычно можно определить в результате, например, интервально-корреляционного анализа временных рядов [13] поступления заказов. В результате акторы вырабатывают эвристики, которые позволяют им решать задачу оптимизации собственной загрузки. При этом они принимают решения в условиях неполной определенности: с одной стороны, во время выполнения заказа могут поступить $e(w_p, u_j, t_{i,j})$ с более высокой стоимостью, с другой стороны, долгое ожидание снижает суммарную экономическую эффективность работы.

В реальной среде 5PL каждый актор может иметь собственную модель поведения в зависимости от выбранной стратегии выстраивания последовательности действий во времени. Модель поведения определяется также спецификой предметной области. Так, для небольших внутригородских перевозчиков наблюдается предпочтение дальних поездок, несмотря на то что в случае удачи может быть выгоднее выполнение нескольких коротких поездок.

Например, водители такси, несмотря на весьма жесткую конкуренцию, придерживаются стратегии выжидания выгодных заказов по высокой цене (в зоне вокзалов или крупного скопления ожидаемых клиентов), что существенно снижает общий уровень сервиса. При этом конкуренция приводит к возникновению очереди исполнителей, что вовсе не означает недостатка актуальных заказов. Логика централизованной диспетчерской в этом случае должна быть направлена на ситуационное разрушение таких тенденций в процессе передачи заказов водителям.

Это означает, что в ответ на поток заказов $\varepsilon(w_p, t_i^*)$ путем введения задержек будет формироваться $e(w_p, u_j, t_{i,j}), t_{i,j} > t_i^*$ такой, чтобы с учетом стратегии акторов выполнялось условие (1).

При организации проактивной диспетчеризации прямое управление распределением ресурсов невозможно, поскольку перечень ресурсов и заказов меняется со временем. Поэтому перед центром — оператором 5PL необходимо ставить следующие задачи.

1. Привлечение заказчиков и исполнителей в целях максимизации их числа для обеспечения возможности выбора наилучшего исполнения каждого заказа.

2. Формирование обстоятельств взаимодействия в целях поддержки кооперации и конкуренции, полезных для его участников и/или центра.

3. Оценка и анализ ключевых показателей эффективности бизнеса в целях поиска путей повышения уровня сервиса.

С учетом этой специфики, обусловленной влиянием фактора времени, сформулируем следующие показатели эффективности деятельности оператора 5PL:

- суммарная стоимость заказов, обработанных за период

$$\sum_{i=1}^{N_w} \sum_{j=1}^{N_u} e'(w_p, u_j, t'_{i,j}) c_{i,j} \quad (8)$$

- средняя загрузка ресурсов

$$\frac{1}{N_u} \sum_{i=1}^{N_w} \sum_{j=1}^{N_u} e'(w_p, u_j, t'_{i,j}) \Delta t_{i,j} \quad (9)$$

- связанный объем заказов — стоимость заказов и среднее время ожидания их распределения

$$\Theta = \frac{1}{N_w} \sum_{i=1}^{N_w} \sum_{j=1}^{N_u} \varepsilon(w_p, t_i^*) \cdot e'(w_p, u_j, t'_{i,j}) \cdot (t'_{i,j} - t_i^*) \quad (10)$$

Основная идея балансировки этих показателей состоит в том, что в случае увеличения стоимости запланированных заказов нельзя допускать высокого роста привлекаемых ресурсов, так как для поддержания их интереса нужно будет еще больше увеличивать число новых заказов. В соответствии с (10) для повышения числа вариантов распределения ресурсов, новые заказы необходимо передавать исполнителю не сразу. Ожидание в течение определенного периода времени позволит лучше подобрать ресурсы среди освобождающихся акторов. Однако акторы в это время могут забирать другие заказы согласно жадной стратегии, описанной выше. В связи с этим целесообразно поддерживать данный показатель на определенном уровне, небольшие колебания в зависимости от пропорции числа заказов к числу ресурсов в данный период времени обеспечат адаптивность 5PL оператора к изменениям внешней среды.

Реализация проактивной диспетчеризации

Транспортную сеть оператора 5PL обычно представляют в виде графа, узлами/вершинами которого являются пункты погрузки и разгрузки, а ребрами/дугами — пути, соединяющие эти пункты. Современные геоинформационные сервисы позволяют определить маршруты перемещения между двумя узлами с достаточно высокой точностью, причем учитывают разную проходимость дорог, скоростные режимы, наличие пробок и возможных ограничений передвижения.

Такая сеть представляет собой основу для любой системы управления ресурсами в транспортной логистике, объективно описывает исходную ситуацию и позволяет оптимизировать выполнение заказов с учетом информации о необходимых географических пунктах загрузки и выгрузки.

Для решения задачи проактивной диспетчеризации предлагается дополнить этот граф, построив на его основе так называемую оверлейную сеть [9]. Этот метод широко применяется в мультиагентных системах маршрутизации и позволяет строить имитационные модели по аналогии с природными механизмами самоорганизации, что полезно в условиях решаемой задачи. Оверлейная сеть представляет собой совокупность узлов, соответствующих как реальным географическим точкам, так и виртуальным площадкам, в которых могут находиться акторы, ожидающие новых заказов. Узлы оверлейной сети связаны между собой логическими ребрами, их "сложность" прохождения может быть отлична от реального времени или стоимости, которые требуют перемещение по пути между реальными географическими пунктами.

Информация об оверлейной сети становится доступна акторам, участвующим во взаимодействии. Отметим, что для разных групп акторов могут быть построены отличающиеся оверлейные сети. Таким образом обеспечивается информационное управление их взаимодействием, цель которого — сохранение требуемой ритмичности потребления заказов.

Учитывая особенности решаемой задачи, можно сформулировать принцип изменения весов ребер в условиях проактивной диспетчеризации. В отличие от традиционного подхода, основанного на отсечении заведомо невыигрышных решений, здесь необходимо искать и отсекал решения, выигрышные с позиций акторов, но не выгодные с позиций баланса (8)—(10). Такой подход основан на том, что акторы сами заинтересованы в поиске наилучшего для себя решения, поэтому из предложенных вариантов выберут наиболее выгодный. Центр же должен проводить корректировки, направленные на сохранение справедливого распределения заказов, что соответствует принципам управления обстоятельствами.

В рамках предлагаемого подхода исследовали два алгоритма. Первый алгоритм основан на выделении в оверлейной сети участков ответственности акторов и увеличении сложности перехода из этих участков далее по сети. Фактически, можно реализовать централизованное решение задачи о назначениях с последующим ограничением движения акторов в сети только теми переходами, которые ведут их к назначенным товарам. Для реализации такого решения система назначает попарно заказы и акторы и перекрывает все ребра сети, за исключением тех, что приведут эти пары друг к другу. Данный метод требует большого числа переключений активности ребер графа, вследствие чего его эффективность тем ниже, чем выше условная стоимость таких переключений.

Второй алгоритм — формирование ограниченный — основан на том, что перекрываются все вершины за исключением минимального числа вершин, содержащих товары с самым длительным сроком ожидания обработки (согласно (10)). Такое управление направлено на минимизацию вероятности ухода товара из системы, связанного со срывом срока обработки. При этом учитывается только актуальный срок ожидания, так как по условиям задачи центр не владеет информацией о реальной максимальной длительности ожидания обработки товаров.

В качестве примера реализации описанных выше принципов приведем систему интеллектуального планирования и поддержки принятия решений в транспортной логистике Haulier Management Solution, разработанную в компании *Magenta Technology* [14]. Проект разрабатывался для оператора 4PL, однако с учетом приведенной во введении классификации, а также особенностей реализации программного обеспечения, поддерживающего принцип SAAS, данное решение следует отнести к классу 5PL. Внедрение данного продукта позволило обеспечить повышение показателей эффективности распределения ресурсов: сокращение времени на взаимодействие до 10 %, времени обработки заказов до 25 %, вызванное, в первую очередь, автоматизацией основных бизнес-процессов компании по планированию и диспетчеризации транспортных ресурсов. Кроме этого, реализация описанного метода при предоставлении услуг на базе SAAS платформы с увеличением числа транспортно-экспедиционных компаний до пяти позволяет сократить число потерянных заказов более чем на 10 %, что подтверждается как результатами имитационного моделирования, так и в ходе анализа реальных потоков заказов, обрабатываемых оператором 5PL.

В результате обобщения современных тенденций и практического опыта по автоматизации грузовых перевозок удалось сформулировать и формализовать метод проактивной диспетчеризации, который позволяет по-новому подойти к решению актуальной задачи управления транспортными ресурсами в условиях неопределенности, свойственной деятельности операторов 5PL. Для реализации этого метода предложено построить оверлейную сеть для интеллектуальной системы планирования в составе портала оператора 5PL, предоставляющего услуги в области транспортной логистики.

Список литературы

1. **Dmitriev A.** Maxoptra dynamic real time scheduling solution / New Magenta Papers: Scientific papers. Vol. 2. Yelm, WA, USA: Science Book Publishing House, 2013. P. 7—12.
2. **Чернова Д. В., Токманев С. В.** Комплексная оценка экономической эффективности управления запасами оптово-средних организаций // Вестник Самарского государственного экономического университета. 2009. № 10 (60). С. 107—110.
3. **Hickson A., Wirth B., Morales G.** Supply chain intermediaries study. University of Manitoba Transport Institute, 2008. 56 p.
4. **АвтоТрансИнфо** [Электронный ресурс]. URL: <http://www.ati.su>
5. **Leung J. Y.-T., Anderson J. H.** Handbook of scheduling: algorithms, models and performance analysis / Ed. by J. Y.-T. Leung. CRC computer and information science series. Boca Raton, FL, USA: Chapman & Hall, Published by CRC, 2004. 1120 p.
6. **Ghiani G., Guerriero F., Laporte G., Musmanno R.** Real-time vehicle routing: solution concepts, algorithms and parallel computing strategies // European Journal of Operational Research, Elsevier. 2003. Vol. 151. P. 1—11.
7. **Giaglis G. M., Minis I., Tatarakis A., Zimpekis V.** Minimizing logistics risk through real-time vehicle routing and mobile technologies: research to-date and future trends // International Journal of Physical Distribution and Logistics Management. 2004. Vol. 34, N 9. P. 749—764.
8. **Шереметов Л. Б.** Децентрализованное управление адаптивными сетями поставок на основе теории коллективного интеллекта и агентной технологии. Часть I: Модель сети поставок // Информационно-управляющие системы. 2009. № 4. С. 13—20.
9. **Городецкий В. И.** Самоорганизация и многоагентные системы. II. Приложения и технология разработки // Известия РАН. Теория и системы управления. 2012. № 3. С. 55—75.
10. **Andreev V., Glaschenko A., Ivaschenko A., Inozemtsev S., Rzevski G., Skobelev P., Shveykin P.** Magenta multi-agent systems for dynamic scheduling // Proc. of the first International Conference on Agents and Artificial Intelligence (ICAART 2009). Porto, Portugal, 19—21 January, 2009. P. 489—496.
11. **Ivaschenko A., Lednev A.** Auction model of P2P interaction in multi-agent software // Proc. of the 5th International Conference on Agents and Artificial Intelligence (ICAART 2013), Barcelona, Spain. 2013. Vol. 1. P. 431—434.
12. **Ивашенко А. В.** Метод кондиционального управления взаимодействием в мультиагентной среде // Системы управления и информационные технологии. 2013. № 1. С. 39—43.
13. **Прикладной анализ случайных процессов** / Под ред. С. А. Прохорова. Самара: Изд-во СНЦ РАН, 2007. 582 с.
14. http://www.magenta-technology.com/news_and_pr/hms_brings_csl_closer.html (доступ 31.08.2013).

ЖУРНАЛ В ЖУРНАЛЕ



**НЕЙРОСЕТЕВЫЕ
ТЕХНОЛОГИИ**

№ 3
МАРТ
2014

Главный редактор:

ГАЛУШКИН А. И.

Редакционная коллегия:

АВЕДЬЯН Э. Д.
БАЗИЯН Б. Х.
БЕНЕВОЛЕНСКИЙ С. Б.
БОРИСОВ В. В.
ГОРБАЧЕНКО В. И.
ЖДАНОВ А. А.
ЗЕФИРОВ Н. С.
ЗОЗУЛЯ Ю. И.
КРИЖИЖАНОВСКИЙ Б. В.
КУДРЯВЦЕВ В. Б.
КУЛИК С. Д.
КУРАВСКИЙ Л. С.
РЕДЬКО В. Г.
РУДИНСКИЙ А. В.
СИМОРОВ С. Н.
ФЕДУЛОВ А. С.
ЧЕРВЯКОВ Н. И.

**Иностранные
члены редколлегии:**

БОЯНОВ К.
ВЕЛИЧКОВСКИЙ Б. М.
ГРАБАРЧУК В.
РУТКОВСКИЙ Л.

Редакция:

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

Скороходов А. В.

Обнаружение воздушного судна по спектру его звукового сигнала с использованием нейросетевого классификатора 56

Игнатьев Д. И., Храбров А. Н.

Нейросетевое моделирование нестационарных продольных аэродинамических характеристик самолета 60

Кинарский А. В., Жернаков С. В.

Стендовая отладка параметров авиационного газотурбинного двигателя на основе нейросетевых технологий 69

Вниманию читателей!

Статья Е. А. Сирота "Прогнозирование нестационарных временных последовательностей на основе иерархических многослойных нейронных сетей", опубликованная в журнале "Информационные технологии" № 1, 2014 г. (стр. 67—71), подготовлена по результатам работы, выполненной при поддержке гранта Минобрнауки РФ по программе "Развитие кооперации российских вузов и производственных предприятий" (Постановление Правительства " 218 от 09.04.2010 г. — 3 очередь, № 02.G25.31.0002), а также программы стратегического развития Воронежского государственного университета.

А. В. Скороходов, мл. науч. сотр.,
Федеральное государственное бюджетное учреждение науки Институт оптики атмосферы
им. В. Е. Зуева Сибирского отделения РАН, г. Томск, e-mail: vazime@yandex.ru

Обнаружение воздушного судна по спектру его звукового сигнала с использованием нейросетевого классификатора

Разработан алгоритм автоматической идентификации реактивных гражданских самолетов по спектру их звуковых сигналов с применением самоорганизующейся нейронной сети Кохонена. Показано, что при их обнаружении информативной частью спектра является диапазон 50...3000 Гц. Предложенный алгоритм позволяет идентифицировать воздушные суда при скорости ветра более 5 м/с и созданном ими уровне шума в точке наблюдения, превышающем менее чем на 3...5 дБ уровень фонового шума. Представлено описание архитектуры нейронной сети и алгоритма ее обучения. Обсуждаются результаты тестирования программы обнаружения воздушных судов, вылетающих из международного аэропорта Шереметьево (г. Москва).

Ключевые слова: воздушное судно, идентификация, нейронная сеть, спектр

A. V. Skorokhodov

Aircraft Detection by its Audio Signal Spectrum using Neural Network Classifier

An algorithm for automatic identification of civil jet aircraft by its audio signal spectrum using a self-organizing neural network is developed. Shown that the range of 50—3000 Hz is an informative spectrum part for aircraft detection. The proposed algorithm can identify the aircraft at a wind speed of 5 m/s, and creating noise in the observation point greater than less than 3—5 dB background noise. The neural network architecture description and a learning algorithm are presented. The program test results of aircraft detection departing from Sheremetyevo International Airport (Moscow) are discussed.

Keywords: aircraft, identification, neural network, spectrum

Введение

В настоящее время одной из значительных экологических проблем в крупных населенных пунктах является повышенный уровень шума. Особенно это актуально для больших городов, вблизи которых расположены крупные аэропорты (Казань, Москва, Новосибирск, Санкт-Петербург и др.), из которых выполняется несколько десятков вылетов в час. Несмотря на то что существуют заранее определенные маршруты пролетов авиатранспорта, реактивные пассажирские самолеты в силу различных причин могут их нарушать, курсируя над территорией населенных пунктов. Поэтому возникает необходимость в автоматическом контроле шума воздушных судов (ВС) в целях определения нарушений ими установленных маршрутов пролета и передачи этой информации диспетчерским пунктам аэропортов. Особенно актуальным контроль шума авиатранспорта стал после принятия постановления Правительства Российской Федерации № 138 от 11 марта 2010 г. "Об утверждении Федеральных правил использования воздушного пространства Российской Федерации".

Задача обнаружения ВС в пределах контролируемой зоны может быть решена на основе следующих подходов.

1. Использование визуальной информации, получаемой с помощью панорамных видеорегистраторов или видеокамер с полусферическим углом обзора типа "all sky" [1]. Основным недостатком такого способа обнаружения ВС является их малоприспособленность при наличии облачного покрова, плотного тумана и других атмосферных явлений, значительно ухудшающих видимость.

2. Применение радиолокационных станций обнаружения, управления и слежения. Главные недостатки данного подхода: высокая стоимость радарных установок, невозможность их размещения в произвольной точке наблюдения и недоступность информации от уже имеющихся станций.

3. Использование акустических средств контроля шума авиатранспорта. К недостаткам этого способа относятся небольшой радиус контролируемой территории (1—2 км) и снижение качества результатов обнаружения ВС при некоторых метеорологических условиях, в первую очередь сильном ветре и ливневых дождях. Однако преимуществом является то,

что с помощью этих средств можно определять как уровень шума, создаваемый ВС при полете, так и их тип. При этом обеспечивается успешное функционирование акустических средств контроля при наличии облачности, туманов и снегопадов. Еще одной отличительной особенностью является низкая (по сравнению с остальными подходами) стоимость регистрирующих устройств.

В данной работе для решения задачи обнаружения ВС используется акустический метод. Следует отметить, что в соответствии с существующим ГОСТ [2] измерение шума ВС необходимо выполнять при скорости ветра до 5 м/с. Для того чтобы минимизировать влияние, оказываемое ветром, на результаты измерения шума ВС на местности, следует использовать ветрозащиту микрофонов.

Цель данной работы — разработка алгоритма идентификации ВС по спектру звукового сигнала, основанного на применении технологии искусственных нейронных сетей, а также оценка перспективности его использования для решения задачи обнаружения реактивных пассажирских самолетов, создающих уровень шума при полете, превышающий менее чем на 3...5 дБ уровень фонового шума.

Исходные данные

Используемая в работе аппаратная часть, показанная на рис. 1, состоит из микрофона, имеющего ветрозащиту, аналогового усилителя с полосовым фильтром (АУ) (полоса пропускания 12...12 500 Гц), блока аналого-цифрового преобразователя (АЦП) (частота оцифровки — 25 000 Гц) и мобильного вычислительного устройства — персонального компьютера.

Основная идея обнаружения ВС заключается в анализе спектров акустических сигналов, регистрируемых приведенным на рис. 1 аппаратным комплексом. В результате проведенных исследований звукозаписей, полученных в районе аэропортов Богашево (г. Томск) и Шереметьево (г. Москва), установлено, что спектры звука двигателей различного типа летательных аппаратов существенно отличаются друг от друга [3]. Так, например, на рис. 2 показаны спектры звуковых сигналов в диапазоне

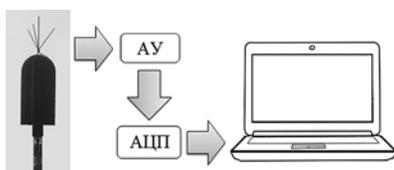


Рис. 1. Аппаратный комплекс обнаружения ВС по спектру их акустических сигналов

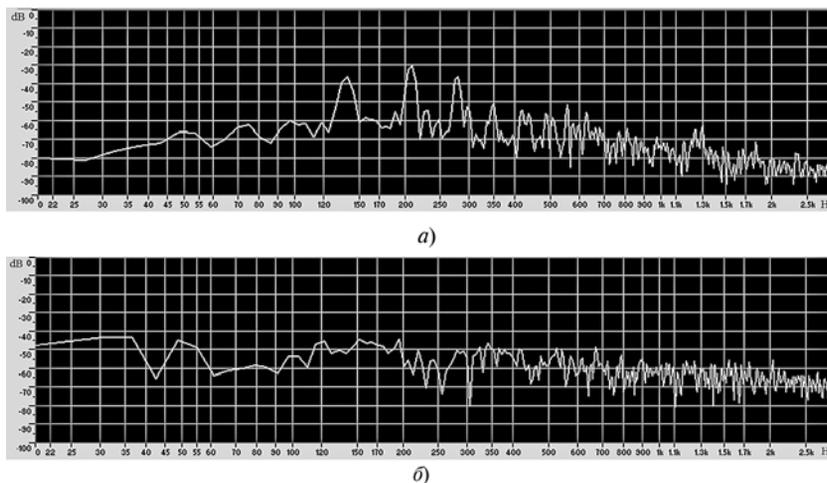


Рис. 2. Графики спектров вертолета Ми-8 на расстоянии 800...1000 м от точки приема в районе аэропорта Богашево (а) и реактивного пассажирского самолета А-320 на расстоянии 1000...1500 м от точки приема в районе аэропорта Шереметьево (б)

до 3000 Гц вертолета (Ми-8) и реактивного самолета (А-320). Как видно из рис. 2, а, график спектра звука вертолета имеет оригинальную форму в диапазоне 50...400 Гц (три высоких пика). Для реактивных самолетов характерной особенностью является насыщенная область спектра в диапазоне 700...3000 Гц (рис. 2, б).

Согласно схеме, показанной на рис. 3, после выполнения процедуры аналого-цифрового преобразования акустического сигнала информация сохраняется в бинарные файлы. Для вычисления спектров используется быстрое преобразование Фурье (БПФ).

Результатом выполнения процедуры БПФ является массив спектральных отсчетов, который и будет использован в качестве идентификационных характеристик ВС.

Первым этапом при решении задачи обнаружения ВС по спектру его звукового сигнала является составление обучающей выборки, на основе которой выполняется настройка алгоритма идентификации. В качестве тестовых образцов были использованы записи, сделанные в непосредственной близости от аэропорта Шереметьево (г. Москва). В тестовую выборку попали записи как с пролетами различных типов реактивных пассажирских самолетов, когда уровень шума ВС превышал на 10...40 дБ уровень фонового шума, так и записи только фонового шума, при наличии и отсутствии ветра.



Рис. 3. Схема работы программной части системы обнаружения ВС

Алгоритм обнаружения

Для идентификации ВС был выбран алгоритм, основой которого является самоорганизующаяся нейронная сеть Кохонена. Основными преимуществами сетей данного типа являются достаточно высокая эффективность при работе с зашумленными данными (фоновый уровень шума), способность "извлекать" ключевые характеристики сигналов (особенности спектра реактивного самолета) и высокая производительность. Поскольку основным назначением сетей Кохонена является предварительный анализ данных (кластеризация), в алгоритме реализован интерпретатор откликов нейронов, позволяющий принимать решение относительно принадлежности спектра акустического сигнала к одному из следующих классов:

- "Пусто": ВС нет и отсутствует ветер;
- "Ветер": ВС нет и присутствует ветер;
- "Ветер/ВС": по спектру звукового сигнала невозможно однозначно принять решение о его принадлежности к одному из классов: "ветер" или "ВС";
- "ВС": в воздушном пространстве обнаружено ВС.

Архитектура используемой нейронной сети Кохонена показана на рис. 4. На вход сети предъявляется N спектральных отсчетов $x_1, x_2, x_3, x_4, \dots, x_N$ акустического сигнала из заданного спектрального диапазона, полученных после БПФ. Сама сеть состоит из четырех нейронов, определяющих некий кластер, соответствие которого определенному классу назначается интерпретатором. Каждый нейрон $i = 1, \dots, 4$ соединен со всеми входами сети $j = 1, 2, \dots, N$ с помощью синаптической связи, имеющей вес w_{ij} . Выходом сети является информация о классе, к которому принадлежит анализируемый спектр звукового сигнала.

Для настройки нейронной сети использовался приведенный ниже алгоритм Conscience Winner Takes All (CWTA), признанный одним из наиболее эффективных алгоритмов самоорганизации [4].

1. Значения весовых коэффициентов w_{ij} инициализируются случайными значениями. Поскольку $N \gg 200$, то нормализация входного вектора $x = (x_1, x_2, x_3, x_4, \dots, x_N)$ и векторов весовых коэффициентов $w_i = (w_{i,1}, w_{i,2}, \dots, w_{i,N})$ не оказывает

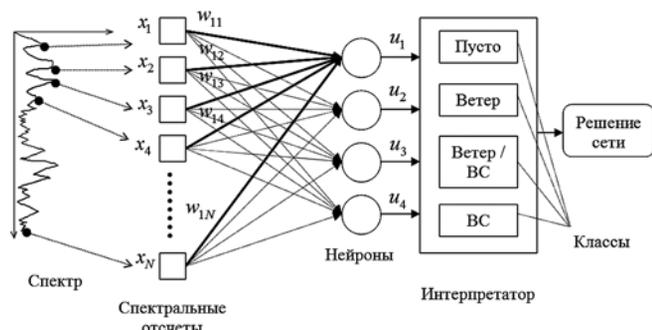


Рис. 4. Архитектура используемой нейронной сети Кохонена

влияния на процесс самоорганизации [4], и, следовательно, этой процедурой можно пренебречь.

2. Процесс обучения сети разбивается на эпохи. При этом эпохой считается предъявление на вход нейронной сети всей обучающей выборки. После поступления обучающего образца на вход сети Кохонена вычисляется суммарный входной сигнал u_i для каждого нейрона:

$$u_i = \sum_{j=1}^N x_j w_{ij}$$

3. Определяется нейрон-победитель, имеющий по результатам сравнения наибольшее значение $u_i/g_i(l)$, где $g_i(l)$ — число побед i -го нейрона на l -й итерации. Такой способ определения победителя реализует "механизм утомления", позволяющий задействовать все нейроны за счет введения их в область с большим объемом данных [4]. При этом "механизм" утомления используется только на начальных итерациях обучения, после которых принятие решения о нейроне-победителе принимается путем сравнения значений u_i . В данной работе "механизм утомления" отключался после того, как число побед $g_i(l)$ у каждого нейрона достигало 100.

4. Нейрон-победитель активируется, вырабатывая на своем выходе состояние 1, а остальные нейроны принимают значение 0. Для победившего нейрона происходит коррекция его весовых коэффициентов по правилу Гроссберга

$$w_w(l+1) = w_w(l) + \eta(l)[x - w_w(l)],$$

где $\eta(l)$ — параметр настройки нейронной сети, определяющий скорость и точность обучения. Согласно рекомендациям, приведенным в работе [5], коэффициент $\eta(l) = 0,01$ на протяжении всего процесса обучения.

5. В конце каждой эпохи (кроме первой) проводится проверка на завершение обучения сети, согласно критерию останова

$$\sqrt{\sum_{j=1}^N (w_{ij}(e+1) - w_{ij}(e))^2} < 0,05 \quad \forall i,$$

где e — номер эпохи обучения.

Процесс настройки нейронной сети заключается в коррекции весовых коэффициентов таким образом, что при предъявлении сходных наборов спектральных отсчетов акустических сигналов победителем оказывался один и тот же нейрон. При этом настройка сети Кохонена протекает в полностью автоматическом режиме (самоорганизация процесса обучения).

Для создания интерпретатора откликов нейронной сети была заново проанализирована обучающая выборка. На ее вход предъявлялись образцы с заранее известной принадлежностью к одному из четырех классов. После того как был определен

номер нейрона-победителя, ему ставилось в соответствие название класса, к которому принадлежит обучающий образец. Таким образом, после обработки всей обучающей выборки для каждого нейрона было подсчитано число соответствий каждому классу. Затем каждому нейрону было назначено имя строго одного класса, число соответствий с которым было наибольшим.

Процедура идентификации ВС аналогична процессу обучения сети Кохонена с той лишь разницей, что после выявления нейрона-победителя спектральному отсчету звукового сигнала ставилось в соответствие имя класса, к которому он принадлежит. Таким же образом обрабатывается вся аудиозапись, полученная с микрофона.

Результаты идентификации воздушных судов

Результатом выполнения процедуры обнаружения ВС являются файлы, в которых записаны временные отсчеты с соответствующим каждому из них одним из четырех классов. Графический пример записи результатов идентификации показан на рис. 5, из которого видно, что было зафиксировано два события, соответствующих пролетам ВС: $t \in (34, 36)$ и $t \in (40, 55)$. В результате исследований звуковых сигналов реактивных самолетов и шума ветра было установлено, что их спектры могут иметь достаточно высокую степень сходства. При этом в результате проведенных натурных наблюдений было установлено, что длительность пролета реактивных самолетов над контролируемой зоной превышает 7 с. Поэтому первое событие из рис. 5 не является событием пролета ВС, а второе — ему соответствует. Поэтому в данной работе событие, соответствующее обнаружению ВС, считалось достоверным, если его длительность превышала 7 с. Аналогичные временные ограничения для остальных классов не вводили.

Для оценки эффективности предложенного алгоритма обнаружения ВС по спектру их звукового сигнала была составлена тестовая выборка, состоящая из 3669 архивных записей, собранных с использованием аппаратуры, установленной в г. Зеленограде (Московская обл.) на крыше высотного здания. Длительность каждой такой записи (тестового образца) составляла 60 с. При этом в тестовую выборку были включены образцы, относящиеся к различным классам, определенным ранее. Таким образом алгоритм проверяли не только на эффективность обнаружения ВС по 199 записям их пролетов, но и на ложные срабатывания (ЛС) по 3470 записям без присутствия ВС в условиях с различными уровнями фонового шума и скоростью ветра. При этом ЛС считалось событием, когда алгоритм 7 раз подряд оши-

Результаты оценки достоверности обнаружения ВС

Спектральный диапазон, Гц	ВС ($N_F = 199$, # $N_F = 90$)	ЛС ($N_F = 3470$)
0...12 500	$N_R = 189$, # $N_R = 81$ $E = 95\%$, # $E = 90\%$	$N_R = 3387$ $100\% - E = 2,4\%$
50...3000	$N_R = 184$, # $N_R = 79$ $E = 92\%$, # $E = 88\%$	$N_R = 3444$ $100\% - E = 0,7\%$

бочно относил спектр звукового сигнала, преимущественно шума ветра, к ВС. Следует отметить, что в записях с пролетами ВС присутствовали данные, когда уровень шума ВС превышал более чем на 8...10 дБ уровень фонового шума, а скорость ветра была менее 5 м/с, и также образцы, где звук летящего ВС был едва различим или скорость ветра была больше 5 м/с.

В работе было выполнено сравнение результатов обнаружения ВС при использовании всего диапазона спектра акустического сигнала (0...12 500 Гц) и части спектра (50...3000 Гц). В таблице приведены процентные соотношения правильного обнаружения, ЛС и общей достоверности результатов идентификации согласно выражению

$$E = \frac{N_R}{N_F} 100\%,$$

где N_R — число правильно принятых алгоритмом решений; N_F — общее число рассматриваемых ситуаций. Результаты оценки эффективности алгоритма представлены в таблице. Знаком "#" в таблице отмечены события, когда превышение уровней шума ВС над уровнем фонового шума составляло менее 3...5 дБ или скорость ветра была более 5 м/с.

Как видно из таблицы результаты обнаружения ВС сходны при использовании различных спектральных диапазонов, однако число ложных обнаружений меньше при диапазоне 50...3000 Гц. Исходя из этого, можно сделать вывод о том, что диапазон спектра 50...3000 Гц является информативным для идентификации реактивного пассажирского авиатранспорта. Следовательно, для уменьшения объема

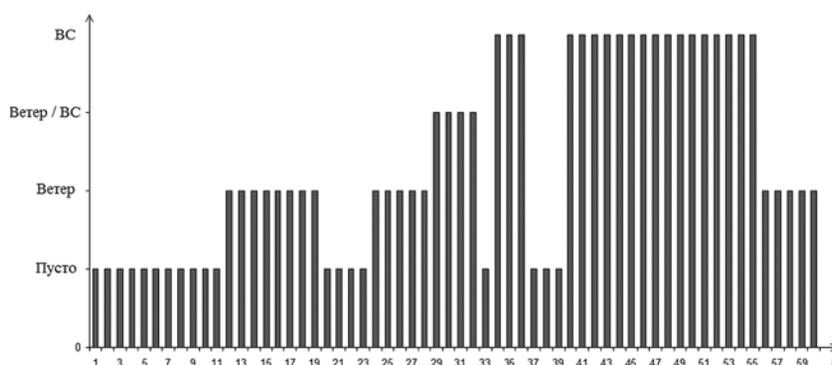


Рис. 5. Результат обнаружения ВС по спектру его акустического сигнала

вычислений и числа ложных обнаружений, а также для повышения производительности программы оставшейся частью спектра можно пренебречь.

В работе была проведена оценка быстродействия разработанного программного обеспечения, в ходе которой было установлено, что на обработку записи длительностью 60 с было затрачено чуть более 7 с. Следовательно, можно сделать вывод о том, что разработанный алгоритм обнаружения ВС является эффективным для использования в режиме реального времени и натуральных условиях.

Заключение

Результаты исследований подтверждают эффективность предложенного подхода к автоматическому обнаружению гражданских реактивных самолетов по спектрам их звуковых сигналов не только при благоприятных условиях (слабый ветер, уровень шума ВС превышает уровень фонового шума более чем на 8...10 дБ), но и в условиях с низкими уровнями сигналов от ВС и при скорости ветра более 5 м/с. При этом показано, что информативной областью спектра для идентификации гражданского реактивного авиатранспорта является диапазон 50...3000 Гц. В результате проведения тестовых испытаний установлено, что время получения отклика сети на один спектр не превышает одной секунды, что позволяет использовать алгоритм в режиме реального времени. Также высказано предположение,

что для уверенного принятия решения о наличии ВС достаточно 7 с.

Автор выражает признательность сотрудникам Института оптики атмосферы им. В. Е. Зуева СО РАН Белану Д. Б. и Мамышеву В. П. за предоставленную возможность использовать в своих исследованиях архив аудиозаписей с пролетами ВС, а также д-ру физ.-мат. наук Астафурову В. Г. и канд. физ.-мат. наук Одинцову С. Л. за полезное обсуждение полученных результатов.

Работа выполнена при частичной финансовой поддержке Программы ОФН РАН "Фундаментальные основы акустической диагностики искусственных и природных сред" (проект 3.10.1).

Список литературы

1. **Форсайт Д. А., Понс Ж.** Компьютерное зрение. Современный подход / Пер. с англ. М.: Вильямс, 2004. 928 с.
2. **ГОСТ 22283—88.** Шум авиационный. Допустимые уровни шума на территории жилой застройки и методы его измерения. Введ. 1988—12—22. М.: Госстандарт России: Изд-во стандартов, 1988. 16 с.
3. **Скороходов А. В.** Распознавание типа воздушного судна по спектру его звукового сигнала с применением нейросетевых технологий // Материалы X Всероссийской научной конференции молодых ученых "Наука. Технологии. Инновации", Новосибирск, 3—5 декабря 2010. Ч. 1. Новосибирск: Изд-во Новосибирского государственного технического университета, 2010. С. 37—38.
4. **Осовский С.** Нейронные сети для обработки информации / Пер. с польск. И. Д. Рудинского. М.: Финансы и статистика, 2002. 344 с.
5. **Хайкин С.** Нейронные сети: полный курс / Пер. с англ. Н. Н. Кузусуль, А. Ю. Шелестовой. М.: Вильямс, 2008. 1103 с.

УДК 519.7:004.67

Д. И. Игнатьев, мл. науч. сотр., e-mail: d.ignatyev@mail.ru, **А. Н. Храбров**, нач. отдела, ФГУП Центральный аэрогидродинамический институт им. Н. Е. Жуковского

Нейросетевое моделирование нестационарных продольных аэродинамических характеристик самолета

Для анализа динамики полета самолета и синтеза алгоритмов управления требуются математические модели его нестационарных аэродинамических характеристик. Для разработки этих моделей используются результаты различных динамических экспериментов в аэродинамических трубах при различных значениях кинематических параметров. Представлен подход для моделирования нестационарных аэродинамических характеристик самолета, в котором используются нейронные сети. Проведено сравнение многослойного перцептрона и рекуррентной нейронной сети NARX. Показано, что второй тип нейронных сетей является более подходящим для моделирования нестационарных аэродинамических характеристик в задачах динамики полета. Для повышения точности моделей, разрабатываемых по результатам различных экспериментов, предложен алгоритм обучения нейронных сетей, учитывающий гетероскедастичность данных, в основе которого лежит байесовская регуляризация. Показано, что его применение позволяет повысить точность получаемых моделей.

Ключевые слова: нестационарные аэродинамические характеристики, рекуррентная нейронная сеть NARX, обучение нейронных сетей, байесовская регуляризация, гетероскедастичность

Modeling of Unsteady Longitudinal Aerodynamic Characteristics of Aircraft Using Neural Networks

Mathematical models of unsteady aerodynamics are required for flight dynamic problems. Various dynamic experiments in wind tunnels with different values of kinematic parameters are used for the model development. A technique for modeling of aircraft unsteady aerodynamic characteristics using neural networks is presented in the paper. Comparison of multilayer perceptron and NARX is fulfilled. NARX is shown to be more appropriate for unsteady aerodynamics modeling in the flight dynamics problems. To improve accuracy of the models that are developed using various experiments the training algorithm for heteroscedastic data based on Bayesian regularization is proposed. The algorithm is shown to improve the accuracy of neural network models.

Keywords: unsteady aerodynamic characteristics, recurrent neural network NARX, neural network training, Bayesian regularization, heteroscedastic data

Введение

По данным EASA (European Aviation Safety Agency) за последнее десятилетие основной причиной авиакатастроф стала потеря самолетом управляемости вследствие ухудшения его аэродинамических характеристик при выходе на большие углы атаки и неподготовленности летчика к пилотированию в таких условиях [1]. В связи с этим возникает необходимость в детальном исследовании динамики самолета в критических режимах, таких как сваливание и штопор, а также в разработке математических моделей аэродинамики в расширенной области режимов полета для обучения летчиков технике пилотирования в критических условиях [2]. В частности, для решения этих задач крайне необходимы математические модели, описывающие нестационарную аэродинамику самолетов на больших углах атаки.

С развитием вычислительной техники и численных методов решения уравнений в частных производных наметился существенный прогресс в области прямого численного моделирования нагрузок, действующих на самолет с использованием уравнений Навье—Стокса [3]. Однако современные возможности не позволяют решать совместно уравнения механики жидкости и движения самолета в задачах динамики полета. Так, например, для моделирования динамики и синтеза алгоритмов управления самолетом необходимо проведение большого числа параметрических исследований, что возможно только при использовании простых и не требующих значительных вычислительных и временных ресурсов моделей аэродинамики. Помимо вышеизложенного не представляется возможным использование методов вычислительной аэродинамики при проведении полунатурного моделирования полета самолета на пилотажных стендах с участием летчика в реальном времени.

Для решения этих и других важных задач динамики полета необходимы упрощенные математические модели нестационарной аэродинамики, учиты-

вающие сложные эффекты отрывного и вихревого обтекания, которые работают в реальном времени. Такие математические модели должны быть способны описывать существенные для динамики полета нелинейные явления в широком диапазоне изменения кинематических параметров. На практике исследования особенностей нестационарной аэродинамики самолета проводятся для ограниченного набора кинематических параметров в динамических экспериментах в аэродинамических трубах (АДТ) с использованием различных экспериментальных установок, после чего упрощенные математические модели нестационарной аэродинамики, способные работать в реальном времени, разрабатываются с использованием полученных экспериментальных данных.

Как известно, универсальными аппроксимирующими свойствами обладают нейронные сети. Использование данного математического аппарата имеет ряд преимуществ. Колмогоровым А. Н. показано, что любая непрерывная функция многих переменных может быть представлена с помощью суперпозиции непрерывных функций одной переменной [4]. Этот результат позволяет использовать нейронные сети для аппроксимации функций с любой наперед заданной точностью. Для разработки математических моделей с помощью искусственных нейронных сетей не требуются значительные упрощающие предположения. Это существенным образом облегчает работу исследователя по разработке математических моделей [5]. В связи с этим последнее время появилось большое число работ, в которых авторы для идентификации и моделирования нелинейной аэродинамики используют нейронные сети [6—13].

В настоящей работе предложен нейросетевой подход для моделирования нестационарных аэродинамических характеристик. Нейронная сеть используется для описания данных, полученных при различных динамических экспериментах в АДТ с моделью самолета. Моделирование проводится на примере момента тангажа, так как именно для этой

аэродинамической характеристики в случае неустановившегося движения наблюдаются наиболее сложные нелинейные зависимости от кинематических параметров.

Помимо точности аппроксимации важной задачей при разработке нейросетевых моделей является повышение точности описания данных вне той области, на которых нейронная сеть была настроена, т. е. повышение обобщающей способности нейронных сетей. В работе [14] был предложен метод регуляризации, основанный на байесовском решающем правиле, позволяющий повысить обобщающую способность математической модели за счет снижения эффективного числа ее параметров. В дальнейшем в работе [15] был разработан эффективный алгоритм обучения нейронных сетей GNBR (Gauss-Newton approximation to Bayesian Regularization). Данный подход хорошо себя зарекомендовал для повышения обобщающей способности нейронной сети.

В описанных работах рассматривались случаи, когда данные представлены однотипным набором данных. Вместе с тем, такой подход не всегда может быть применен корректно для гетероскедастичных данных, т. е. данных, характеризующихся неоднородностью наблюдений. Как уже было сказано выше, при создании математических моделей нестационарной аэродинамики используются данные различных экспериментов, которые проводятся на различных установках, при различных значениях кинематических параметров, в каждом из экспериментов получается разное число точек, с различной точностью и т. д. В настоящей работе предлагается метод обучения нейронных сетей, позволяющий повысить точность разрабатываемых моделей по результатам разнотипных экспериментов. В основе подхода лежит метод, основанный на байесовском решающем правиле [14, 15], позволяющий автоматически определять параметры регуляризации. В отличие от вышеизложенных методов [14, 15] в представленной работе учитывается, что данные, получаемые в различных экспериментальных исследованиях одного явления, могут различаться, например, областью исследуемых признаков, погрешностью измерений, числом экспериментальных точек и т. д.

2. Краткое описание эксперимента

Для решения практических задач в области динамики полета в тех случаях, когда рассматривается движение самолета на малых углах атаки, используется упрощенное представление аэродинамических коэффициентов в виде линейного разложения в ряд Тейлора по кинематическим параметрам движения [16]. Например, при продольных колебаниях с закрепленным центром тяжести (в таком

случае $\bar{\alpha}(t) \equiv \bar{\omega}_z$) математическая модель момента тангажа m_z имеет вид

$$m_z = m_{z_{ст}}(\alpha_0) + m_z^\alpha(\alpha_0)(\alpha - \alpha_0) + (m_z^{\bar{\omega}_z} + m_z^{\bar{\alpha}})(\alpha_0)\bar{\omega}_z, \quad (1)$$

где $m_{z_{ст}}(\alpha_0)$ — стационарное значение момента тангажа для среднего при колебании угла атаки α_0 ; $m_z^\alpha(\alpha_0)$ — стационарная аэродинамическая производная, определяющая продольную статическую устойчивость; $(m_z^{\bar{\omega}_z} + m_z^{\bar{\alpha}})(\alpha_0)$ — комплекс вращательных и нестационарных аэродинамических производных, отвечающий за затухание продольных колебаний; $\bar{\omega}_z = \frac{\omega_z b_A}{V_\infty}$ — безразмерная угловая скорость; $\bar{\alpha}(t) = \frac{\dot{\alpha} b_A}{V_\infty}$ — безразмерная производная угла атаки по времени; b_A — средняя аэродинамическая хорда; V_∞ — скорость потока.

Для определения аэродинамических производных $m_z^\alpha(\alpha_0)$ и $(m_z^{\bar{\omega}_z} + m_z^{\bar{\alpha}})(\alpha_0)$ можно применять экспериментальные методы, в частности, метод вынужденных гармонических колебаний модели с малой амплитудой в потоке АДТ.

Однако данный подход (1) справедлив только в области линейного изменения аэродинамических коэффициентов, в частности, при малых углах атаки и при малых амплитудах колебаний. При больших углах атаки и больших амплитудах колебаний наблюдаются нелинейные изменения аэродинамических коэффициентов, в частности, аэродинамические производные зависят от частоты и амплитуды колебаний, наблюдаются гистерезисы полных аэродинамических характеристик, связанные с запаздыванием развития отрывного обтекания, которые не могут быть описаны с использованием линейной модели (1). Для того чтобы уточнить модель аэродинамики в области нелинейного изменения аэродинамических коэффициентов, проводятся дополнительные экспериментальные исследования, в частности, вынужденные гармонические колебания с большой амплитудой.

В ходе реализации проекта SimSAC 6-й Европейской рамочной программы был проведен ряд экспериментов по определению нестационарных аэродинамических характеристик модели транзвукового летательного аппарата Transonic Cruiser — TCR [17]. Общий вид аэродинамической модели представлен на рис. 1. Исследования проводились в АДТ Т-103 ЦАГИ при скорости потока $V_\infty = 40$ м/с в два этапа. При проведении первого этапа на ди-



Рис. 1. Общий вид аэродинамической модели TCR

наимической установке осуществлялись колебания модели с малыми амплитудами, при этом измеряли аэродинамическую производную m_z^α в фазе с изменением угла атаки α и комплекс аэродинамических производных $m_z^{\bar{\omega}_z} + m_z^{\bar{\alpha}}$ в фазе с изменением угловой скорости $\bar{\omega}_z$.

На втором этапе экспериментальных исследований в АДТ использовалась другая динамическая установка, предназначенная для проведения испытаний при колебаниях с большими амплитудами. При данном типе эксперимента исследовали нелинейные нестационарные зависимости полного момента тангажа. На рис. 2 представлены зависимости $m_z(t)$, которые имеют вид сложных гистерезисных петель.

Гармонические колебания задавались для различных частот в широком диапазоне амплитуд по закону:

$$\alpha = \alpha_0 + A_\alpha \sin \omega t, \quad (2)$$

где α_0 — средний угол атаки; A_α — амплитуда колебаний; $\omega = 2\pi f_k$ — круговая частота; f_k — частота колебаний. Для каждой определенной частоты и амплитуды колебаний в эксперименте были получены показания аэродинамических нагрузок за период колебания. Период колебания был разбит на

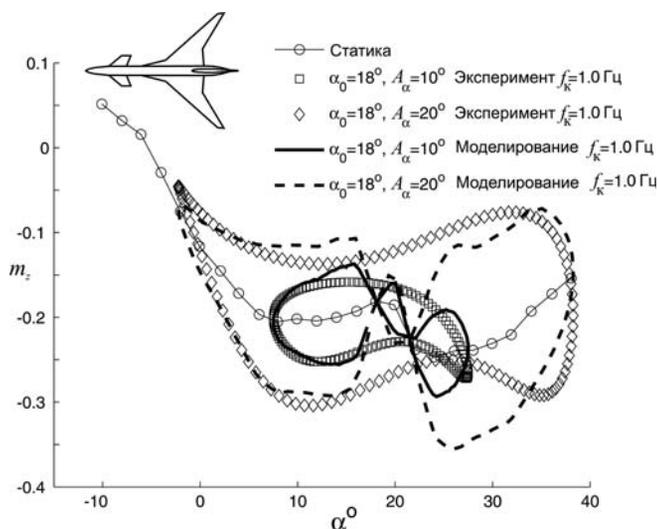


Рис. 2. Момент тангажа при вынужденных колебаниях с большой амплитудой

128 последовательных положений аэродинамической модели самолета в потоке АДТ, различающихся углом атаки, по которым в дальнейшем настраивались нейронные сети.

Как было продемонстрировано в работе [18], применение традиционной модели (1) в широком диапазоне изменения кинематических параметров приводит к существенным ошибкам. Так, на рис. 2 показаны результаты моделирования эксперимента при вынужденных колебаниях с большой амплитудой с использованием традиционной модели, идентифицированной по данным эксперимента с малой амплитудой ($A_\alpha = 3^\circ$), в сравнении с данными эксперимента ($A_\alpha = 10^\circ$ и $A_\alpha = 20^\circ$). В области углов атаки $\alpha = 16...24^\circ$ линейная модель (1) показывает качественно и количественно неверный результат, заключающийся в том, что модель дает дополнительные петли гистерезиса, которые не наблюдаются в эксперименте.

Таким образом, традиционные для задач динамики модели вида (1) не могут быть использованы для описания нестационарных аэродинамических характеристик в области больших углов атаки. В общем случае нестационарные значения аэродинамических характеристик зависят от большого числа кинематических параметров и от предыстории движения. В связи с этим модель нестационарных аэродинамических характеристик, которая строится только по результатам только одного вида эксперимента, не может описать поведение исследуемой характеристики во всем диапазоне исследованных параметров. Для создания математической модели с большей предсказательной способностью необходимо использовать как можно больше различных экспериментов, моделирующих различные движения летательного аппарата.

В работе ставится задача разработать подход, который бы позволил описать все результаты экспериментов с использованием нейросетевого аппарата.

Выбор конфигурации нейронной сети

Зависимости нестационарных аэродинамических производных от амплитуды и частоты колебаний, а также наличие гистерезисов полных аэродинамических характеристик демонстрируют зависимость нестационарной аэродинамики от предыстории движения. Распространенным подходом для моделирования нелинейных процессов с наличием внутренней динамики является использование нейронных сетей с обратными связями [12, 19, 20]. С учетом этих особенностей в работе [12] для моделирования нестационарных аэродинамических характеристик предложено использовать рекуррентную нейронную сеть.

В настоящей работе была использована рекуррентная нейронная сеть NARX (nonlinear autoregressive network with exogenous inputs) с одним скрытым слоем. Схема данной нейронной сети представ-

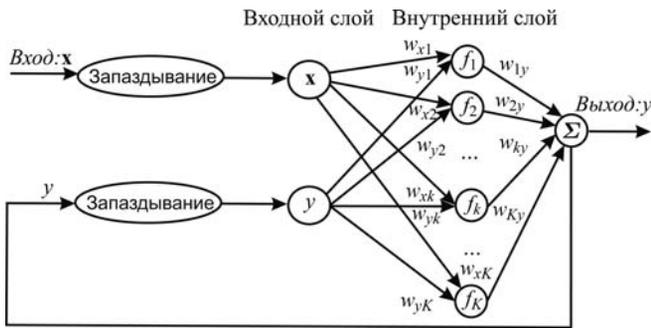


Рис. 3. Схема нейронной сети NARX

лена на рис. 3. Для того чтобы данная нейронная сеть смоделировала искомую величину y в момент времени t на вход в нейронную сеть подается вектор состояния $\mathbf{x}(t)$, а также ряд его предшествующих дискретных значений $\mathbf{x}(t-1), \mathbf{x}(t-2), \dots, \mathbf{x}(t-T_{\text{вх}})$. Для случая продольного движения модели самолета на установке вынужденных колебаний в АДТ компонентами вектора состояния являются кинематические параметры $\mathbf{x}(t) = (\alpha(t), \dot{\alpha}(t))^T$. Помимо этого при моделировании на каждом шаге нейронная сеть использует значения выходного сигнала, полученные нейронной сетью на предыдущих этапах моделирования $y(t-1), y(t-2), \dots, y(t-T_{\text{вых}})$, где $T_{\text{вх}}, T_{\text{вых}}$ — задержка по входному и выходному сигналам. На рисунке w_{xi}, w_{yi}, w_{pj} , где $i, j = 1, \dots, K$ — веса связей между нейронами, f_i — функция активации нейрона.

Метод обучения нейронной сети

При обучении нейронной сети использовался метод обратного распространения ошибки. Обычно в рамках данного метода минимизируется целевая функция F — сумма квадратов ошибок E_D :

$$F = E_D = \frac{1}{2} \sum_{i=1}^N (a_i - y_i)^2,$$

где a_i — исходные данные обучающего множества; y_i — результат работы нейронной сети; N — число паттернов обучающего множества. Часто для повышения обобщающей способности нейронных сетей в выражение для целевой функции F может добавляться дополнительный член, называемый регуляризатором, которым может выступать сумма квадратов весов w_j связей и порогов (далее вместо "связи и пороги" нейронной сети будем использовать "связи"):

$$E_W = \frac{1}{2} \sum_{j=1}^K w_j^2,$$

где K — число связей нейронной сети.

Таким образом, целевая функция представляется в виде линейной комбинации суммы квадратов ошибок и суммы квадратов связей:

$$F = \beta E_D + \alpha E_W,$$

где α и β — параметры целевой функции.

Основной задачей при регуляризации является определение параметров целевой функции α и β . Метод, который хорошо себя зарекомендовал для определения параметров целевой функции, предложен в работе [14]. В ней рассматривалась задача поиска интерполирующей функции для однородного набора данных $D = (a_m, x_m)$, который можно представить в виде

$$a_m = g(x_m) + v_m, \quad (3)$$

где g — функция, аппроксимирующая исходный набор данных; v_m — шум, имеющий гауссово распределение с нулевым математическим ожиданием и стандартным отклонением σ_m . Как было показано в работе [14], $\beta = 1/\sigma_m^2$.

Попробуем расширить данный подход на случай, когда исходные данные могут быть представлены в виде ряда однородных наборов данных. Таким образом, требуется найти функцию y , которая аппроксимирует экспериментальные данные, полученные в n различных экспериментах

$$(\mathbf{x}_1, \mathbf{a}_1), (\mathbf{x}_2, \mathbf{a}_2), \dots, (\mathbf{x}_n, \mathbf{a}_n),$$

где $\mathbf{x}_i = (x_{i1}, \dots, x_{iN_i})$ — вектор параметров исследуемого явления, соответствующий i -му эксперименту; $\mathbf{a}_i = (a_{i1}, \dots, a_{iN_i})$ — вектор значений наблюдаемой величины, соответствующий i -му эксперименту, $\{x_{i m_i}, a_{i m_i}\} = D_i, m_i = 1, \dots, N_i$ — набор данных, полученных в однотипном эксперименте:

$$\begin{cases} a_{1 m_1} = y(x_{1 m_1}) + v_{1 m_1}, m_1 = 1, \dots, N_1; \\ a_{2 m_2} = y(x_{2 m_2}) + v_{2 m_2}, m_2 = 1, \dots, N_2; \\ \dots \\ a_{n m_n} = y(x_{n m_n}) + v_{n m_n}, m_n = 1, \dots, N_n, \end{cases}$$

где $v_{i m_i}, m_i = 1, \dots, N_i$ — ошибка, соответствующая i -му набору данных.

При этом также предполагается, что ошибки в каждом из отдельных экспериментов независимы, имеют гауссово распределение с нулевым математическим ожиданием, но с разными стандартными отклонениями $\sigma_i^2 = 1/\beta_i$, т. е. данные предполагаются гетероскедастичными. Таким образом, вместо параметра β , мы получаем некоторую матрицу \mathbf{B} , состоящую из параметров $\beta_1, \beta_2, \dots, \beta_n$.

Используя байесовское решающее правило, мы можем записать функцию плотности апостериорной вероятности для весов нейронной сети $\mathbf{w} = (w_1 w_2 \dots w_K)^T$ при заданном наборе исходных данных D_1, D_2, \dots, D_n , параметрах целевой функции α, \mathbf{B} , данной конфигурации M нейронной сети:

$$P(\mathbf{w}|D_1, D_2, \dots, D_n, \alpha, \mathbf{B}, M) = \frac{P(D_1, D_2, \dots, D_n|\mathbf{w}, \mathbf{B}, M)P(\mathbf{w}|\alpha, M)}{P(D_1, D_2, \dots, D_n|\alpha, \mathbf{B}, M)}, \quad (4)$$

где $P(\mathbf{w}|\alpha, M)$ — плотность априорной вероятности, которая определяет наши представления о том, какие значения могут принимать веса нейронной сети до того, как были использованы данные; $P(D_1, D_2, \dots, D_n|\mathbf{w}, \mathbf{B}, M)$ — функция правдоподобия; $P(D_1, D_2, \dots, D_n|\alpha, \mathbf{B}, M)$ — нормирующий множитель.

В предположении о том, что ошибка аппроксимации имеет гауссово распределение, мы можем записать выражение для функции правдоподобия в виде

$$P(D_1, D_2, \dots, D_n|\mathbf{w}, \mathbf{B}, M) = \frac{\sqrt{\det \mathbf{B}}}{(2\pi)^{N/2}} \exp\left(-\frac{1}{2} \mathbf{e}^T \mathbf{B} \mathbf{e}\right) = \frac{1}{Z_D} \exp\left(-\frac{1}{2} \mathbf{e}^T \mathbf{B} \mathbf{e}\right), \quad (5)$$

где $Z_D = \frac{(2\pi)^{N/2}}{\beta_1^{N_1/2} \beta_2^{N_2/2} \dots \beta_n^{N_n/2}}$, $\beta_i = \frac{1}{\sigma_i^2}$, $\mathbf{e} = (e_1, \dots, e_N)^T$ — вектор ошибок, $e_i = (y(x_i) - a_i)$ — ошибка на i -й паре данных, \mathbf{B} — матрица размера $N \times N$, на диагонали которой расположены β_i :

$$\mathbf{B} = \begin{pmatrix} \beta_1 & 0 & \dots & & 0 \\ 0 & \beta_1 & 0 & \dots & 0 \\ & & \dots & & \\ 0 & \dots & 0 & \beta_i & 0 & \dots & 0 \\ 0 & \dots & 0 & \beta_i & 0 & \dots & 0 \\ & & \dots & & & & \\ 0 & & & & 0 & \beta_n & 0 \\ 0 & & & & 0 & & \beta_n \end{pmatrix}.$$

Предполагая, что до момента предъявления данных D_1, D_2, \dots, D_n веса нейронной сети \mathbf{w} — случайная величина, получаем, что априорная вероятность имеет гауссово распределение:

$$P(\mathbf{w}|\alpha, M) = \frac{1}{\left(\frac{2\pi}{\alpha}\right)^{K/2}} \exp\left(-\frac{1}{2} \alpha \mathbf{w}^T \mathbf{w}\right) = \frac{1}{Z_W} \exp\left(-\frac{1}{2} \alpha \mathbf{w}^T \mathbf{w}\right), \quad (6)$$

где K — общее число весовых элементов нейронной сети.

Таким образом, можно получить выражение для функции плотности апостериорной вероятности (4):

$$P(\mathbf{w}|D_1, D_2, \dots, D_n, \alpha, \mathbf{B}, M) = \frac{\frac{1}{Z_D} \frac{1}{Z_W} \exp\left(-\frac{1}{2} \mathbf{e}^T \mathbf{B} \mathbf{e} - \frac{1}{2} \alpha \mathbf{w}^T \mathbf{w}\right)}{P(D_1, D_2, \dots, D_n|\alpha, \mathbf{B}, M)}, \quad (7)$$

или

$$P(\mathbf{w}|D_1, D_2, \dots, D_n, \alpha, \mathbf{B}, M) = \frac{1}{Z_F} \exp(-F).$$

Оптимальные значения весовых коэффициентов будут получены максимизацией апостериорной вероятности $P(\mathbf{w}|D_1, D_2, \dots, D_n, \alpha, \mathbf{B}, M)$, что эквивалентно минимизации регуляризованной целевой функции F . Таким образом, мы можем получить выражение для нее:

$$F = \frac{1}{2} \mathbf{e}^T \mathbf{B} \mathbf{e} + \frac{1}{2} \alpha \mathbf{w}^T \mathbf{w}.$$

В выражении для апостериорной вероятности (7) имеются неизвестные значения α, \mathbf{B} . Для поиска оптимальных значений параметров целевой функции также можно снова использовать байесовское решающее правило. Для этого можно записать:

$$P(\alpha, \mathbf{B}|D_1, D_2, \dots, D_n, M) = \frac{P(D_1, D_2, \dots, D_n|\alpha, \mathbf{B}, M)P(\alpha, \mathbf{B}|M)}{P(D_1, D_2, \dots, D_n|M)}. \quad (8)$$

Поскольку заранее мы ничего не знаем о значениях параметров целевой функции, то предположим, что они имеют равномерное распределение $P(\alpha, \mathbf{B}|M)$. Тогда для максимизации апостериорной вероятности (8) необходимо максимизировать функцию правдоподобия $P(D_1, D_2, \dots, D_n|\alpha, \mathbf{B}, M)$. Вместе с тем, эта функция является нормирующим множителем в формулах (4), (7). Выражая ее оттуда, получим:

$$P(D_1, D_2, \dots, D_n|\alpha, \mathbf{B}, M) = \frac{P(D_1, D_2, \dots, D_n|\mathbf{w}, \mathbf{B}, M)P(\mathbf{w}|\alpha, M)}{P(\mathbf{w}|D_1, D_2, \dots, D_n, \alpha, \mathbf{B}, M)} = \frac{Z_F}{Z_W Z_D}. \quad (9)$$

Константы Z_W, Z_D определены выше в формулах (5), (6).

Для определения Z_F разложим целевую функцию в окрестности точки минимума (MP) в ряд Тейлора с точностью до членов второго порядка малости:

$$F \approx F(\mathbf{w}_{MP}) + \frac{1}{2} (\mathbf{w} - \mathbf{w}_{MP})^T \mathbf{H} (\mathbf{w} - \mathbf{w}_{MP}),$$

где $\mathbf{H} = \nabla^2 F$ — матрица Гесса целевой функции. Поскольку

$$Z_F = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \exp(-F) dw_1 \dots dw_K,$$

то можно оценить Z_F :

$$Z_F \approx \exp(-F_{MP})(2\pi)^{K/2}(\det \mathbf{H}_{MP})^{-1/2}.$$

Таким образом, подставляя в (9) значения Z_F , Z_W , Z_D , получим:

$$P(D_1, D_2, \dots, D_n | \alpha, \mathbf{B}, M) \approx \frac{\exp(-F_{MP})(2\pi)^{K/2}(\det \mathbf{H}_{MP})^{-1/2}}{\left(\frac{2\pi}{\alpha}\right)^{K/2} \left(\frac{2\pi}{\beta_1}\right)^{N_1/2} \dots \left(\frac{2\pi}{\beta_n}\right)^{N_n/2}}. \quad (10)$$

Оптимальное значение параметров можно найти, если приравнять нулю соответствующие производные логарифма выражения (10) по этим параметрам. Тогда

$$\ln P(D_1, D_2, \dots, D_n | \alpha, \mathbf{B}, M) = -F_{MP} - \frac{1}{2} \ln \det \mathbf{H}_{MP} + \frac{K}{2} \ln \alpha + \frac{N_1}{2} \ln \beta_1 + \dots + \frac{N_n}{2} \ln \beta_n - \frac{N}{2} \ln 2\pi.$$

Тогда для α можно получить выражение

$$\alpha \approx \frac{K - \alpha \text{Sp}(\mathbf{H}^{-1})}{\mathbf{w}^T \mathbf{w}},$$

где Sp — след матрицы.

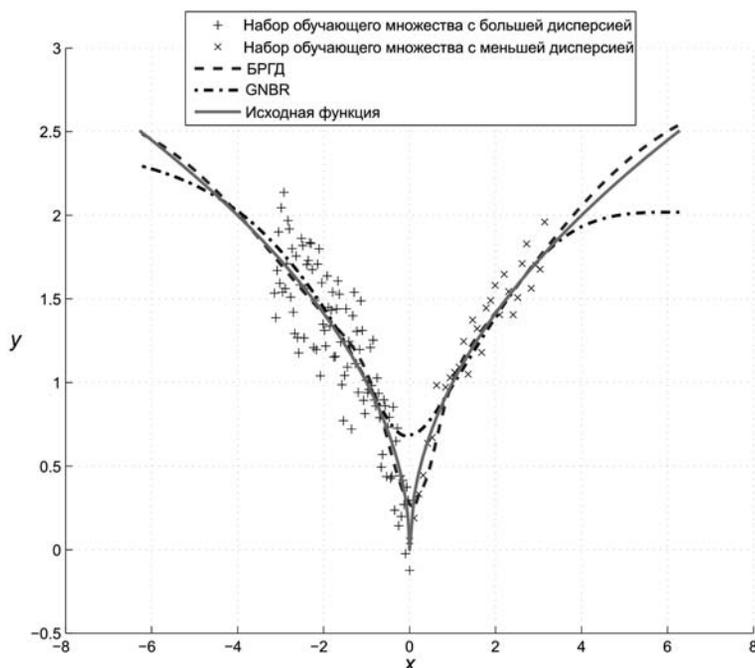


Рис. 4. Сравнение методов обучения нейронных сетей на тестовом примере

Аналогичным образом, для β_i можно получить:

$$\beta_i = \frac{N_i}{\mathbf{e}^T \frac{d\mathbf{B}}{d\beta_i} \mathbf{e} + \text{Sp}\left(\frac{d\mathbf{H}}{d\beta_i} \mathbf{H}^{-1}\right)}.$$

Таким образом, в рамках предложенного подхода коррекция коэффициентов регуляризации для каждого из подмножеств множества обучения происходит с учетом ошибки на данном подмножестве.

Для корректировки весов может быть использован метод Левенберга—Марквардта:

$$\mathbf{w}_i = \mathbf{w}_{i-1} - (\mathbf{H} + \mu \mathbf{E})^{-1} \mathbf{g},$$

где $\mathbf{g} = \nabla F$ градиент целевой функции; $\mu > 0$ — коэффициент для улучшения сходимости в методе Левенберга—Марквардта. Для аппроксимации гессиана можно воспользоваться методом Ньютона—Гаусса

$$\mathbf{H} \approx \mathbf{J}^T \mathbf{B} \mathbf{J} + \alpha \mathbf{E},$$

где \mathbf{J} — матрица Якоби. Градиент можно вычислить как $\mathbf{g} \approx \mathbf{J}^T \mathbf{B} \mathbf{e} + \alpha \mathbf{w}_{i-1}$. Тогда выражение для корректировки весов на i -й итерации

$$\mathbf{w}_i = \mathbf{w}_{i-1} - (\mathbf{J}^T \mathbf{B} \mathbf{J} + (\alpha + \mu) \mathbf{E})^{-1} (\mathbf{J}^T \mathbf{B} \mathbf{e} + \alpha \mathbf{w}_{i-1}).$$

Таким образом, построен алгоритм реализации метода обучения нейронных сетей, основанный на байесовской регуляризации с учетом гетероскедастичности данных (далее — БРГД).

Использование метода для решения тестовой задачи

Проведем оценку того, насколько учет гетероскедастичности данных влияет на точность нейросетевых моделей. В качестве элементарного примера рассмотрим функцию $y = \sqrt{|x|}$. К функции был добавлен шум, имеющий гауссово распределение. Используемый шум имел нулевое математическое ожидание, но различные стандартные отклонения на разных интервалах: при $x < 0$ — большее значение, при $x > 0$ — меньшее. Кроме того, точек с меньшим стандартным отклонением "получено" меньше, чем с большим. Таким образом, были смоделированы два вида эксперимента, с помощью которых определяют искомую физическую зависимость в разных диапазонах независимой величины с разной точностью, что довольно часто встречается на практике.

Описанный набор данных был использован для обучения рекуррентной нейронной сети NARX. Обучение было проведено двумя методами. Вначале использовался алгоритм, который строится на предположении о гомоскедастичности данных — GNBR [15], затем

использовался вышеизложенный подход БРГД. После этого с помощью полученных нейронных сетей была смоделирована исходная функция не только на данных, которые использовались при обучении, но также на данных, выходящих за этот диапазон. Результаты моделирования представлены на рис. 4. Там же даны исходные наборы данных, использованные при обучении, а также для сравнения изображена искомая зависимость.

Представленный график демонстрирует, что новый алгоритм обучения позволяет повысить точность описания искомой функции не только на множестве, которое использовалось для настройки коэффициентов нейронной сети ($-3 < x < 3$), но также и на множестве, которое не использовалось для этих целей ($-6 < x < -3, 3 < x < 6$).

Модель нестационарных аэродинамических характеристик

В рассмотренном тестовом примере метод обучения, учитывающий структуру экспериментальных данных, позволил повысить как аппроксимирующие, так и обобщающие свойства нейронной сети. Рассмотрим теперь задачу разработки математической модели нестационарного момента тангажа m_z по результатам эксперимента в АДТ, проведенного для аэродинамической модели TCR.

Для обучения и тестирования обобщающей способности сети были использованы экспериментальные результаты по колебанию модели с малыми и большими амплитудами. При этом число тестов с малыми амплитудами составляло 78, с большими — 12. В каждом из тестов описан один период колебания модели с разными параметрами колебательного процесса (2): среднее значение α_0 , амплитуда колебаний $\Delta\alpha$, частота ω . Каждый тест состоит из 128 последовательных положений модели самолета в потоке в течение одного периода колебания.

Из всего экспериментального объема данных две трети были использованы на обучение сети и

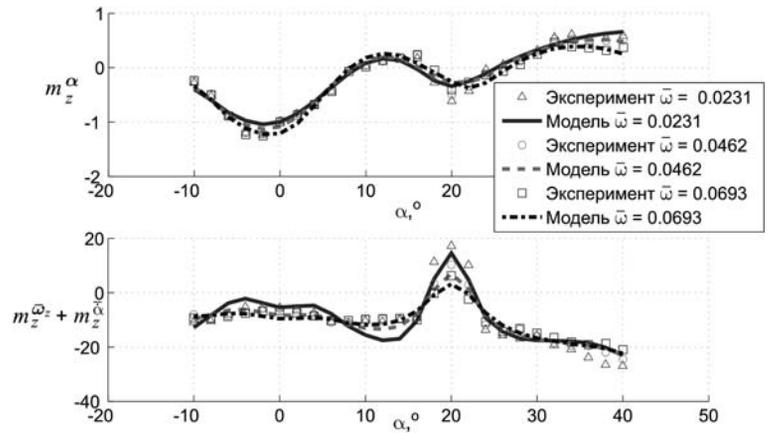


рис. 5. Аэродинамические производные момента тангажа m_z

треть на проверку обобщающей способности — способности моделировать m_z на данных, которые не использовались при обучении. Таким образом, для обучения нейронной сети было использовано 5632 точки с экспериментальными данными, для тестирования — 5888. Функцией активации нейронов скрытого слоя был сигмоид.

На рис. 5 представлены результаты нейросетевого моделирования нестационарных производных в сравнении с экспериментом. Результаты представлены для различных безразмерных частот колебаний

$$\bar{\omega} = \frac{2\pi f_k b a}{V_\infty}$$

развития отрывного обтекания с переднего горизонтального оперения $16^\circ < \alpha < 24^\circ$ наблюдается "расщепление" производных в зависимости от частоты колебаний, которое хорошо моделируется нейронной сетью.

На рис. 6 (см. вторую сторону обложки) представлены результаты моделирования момента тангажа при колебаниях с большой амплитудой в сравнении с экспериментом. Нейросетевая модель качественно и количественно описывает наблюдаемые в эксперименте гистерезисы.

Для того чтобы оценить, насколько разработанный метод обучения оказывает влияние на точность модели, на том же рисунке даны результаты модели, полученной с использованием метода GNBR. Сравнение демонстрирует, что с использованием предложенного подхода удается получить более точную модель момента тангажа. Проведем дополнительный сравнительный анализ нейросетевых моделей, полученных с использованием методов GNBR и БРГД.

Для количественного сравнения в таблице приведены ошибки для моделей момента тангажа m_z и

производной $m_z^{\bar{\omega}} + m_z^{\bar{\alpha}}$ отдельно на множествах обучения и тестирования. В качестве меры ошибки

Ошибки моделей

Модель	Величина	Обучение, %	Тестирование, %
NARX (GNBR)	$m_z^{\bar{\omega}} + m_z^{\bar{\alpha}}$	7,09	8,58
	m_z (большие амплитуды)	5,59	8,3
NARX (БРГД)	$m_z^{\bar{\omega}} + m_z^{\bar{\alpha}}$	5,65	5,77
	m_z (большие амплитуды)	4,53	6,34
Много-слойный перцептрон	$m_z^{\bar{\omega}} + m_z^{\bar{\alpha}}$	3,88	3,96
	m_z (большие амплитуды)	15,25	27,42

использовалась средняя квадратическая погрешность, отнесенная ко всему диапазону измеряемой величины Δy :

$$\text{err}_i = \frac{\sqrt{\frac{1}{N_i-1} \sum_{j=1}^{N_i} (y_j^{\text{эксп}} - y_j^{\text{мод}})^2}}{\Delta y}$$

Сравнивая модели NARX (GNBR) и NARX (БРГД), можно отметить существенное улучшение точности моделирования. Так, ошибка аппроксимации m_z уменьшилась на 23 %, ошибка обобщения — на 31 %. Ошибки описания $m_z^{\bar{\omega}_z} + m_z^{\bar{\alpha}}$ уменьшились на 25 и 49 % соответственно.

На рис. 7, 8 (см. вторую сторону обложки) представлены диаграммы рассеяния для $m_z^{\bar{\omega}_z} + m_z^{\bar{\alpha}}$ и m_z на множествах тестирования. Из них видно, что разработанная методика позволяет добиться меньшего разброса значений искомым величин.

В работе [13] проведено моделирование нестационарных аэродинамических характеристик TCR с использованием многослойного перцептрона (МНП). Нейронная сеть имела два скрытых слоя, для обучения и проверки обобщающей способности был использован тот же набор данных, что и для NARX. На вход в нейронную сеть подавался вектор состояния:

$$\mathbf{I} = (\alpha(t), \dot{\alpha}(t), t, \alpha_0, \Delta\alpha, \omega),$$

где $\alpha(t)$, $\dot{\alpha}(t)$ — "текущие" значения угла атаки и его производной по времени; t — "текущее" время каждого теста, т. е. время, которое прошло от начала колебания модели до момента, когда модель оказалась в текущем положении; α_0 , $\Delta\alpha$, ω — параметры колебательного процесса (2). Таким образом, задавался полный цикл движения от начального положения до конечного.

Для того чтобы сравнить МНП и NARX в таблице также приведены значения ошибки многослойного перцептрона на множестве тестирования.

Ошибка МНП по динамической производной $m_z^{\bar{\omega}_z} + m_z^{\bar{\alpha}}$ несколько меньше по сравнению с NARX для обоих методов обучения, однако ошибка МНП на втором множестве (большие амплитуды) существенно больше как по сравнению с ошибкой МНП на малых амплитудах, так и по сравнению с ошибками рекуррентных сетей на обоих множествах. Таким образом, рекуррентная сеть позволяет добиться большей обобщающей способности в целом на обоих множествах за счет незначительного снижения обобщающей способности на одном из множеств. Важно отметить, что для моделирования МНП задавался весь период колебания модели, кроме того, во входной вектор входили амплитуда и частота колебаний A_α , ω , что крайне осложняет задачу моде-

лирования произвольного движения летательного аппарата на пилотажных стендах. С учетом вышеизложенного можно сделать вывод, что для моделирования нестационарных аэродинамических характеристик в задачах динамики полета предпочтительнее использовать рекуррентные нейронные сети.

Заключение

Актуальность задачи разработки математических моделей нестационарной аэродинамики на больших углах атаки определяется требованием обеспечения безопасности полетов самолетов. Для решения задач динамики полета необходимы упрощенные и работающие в реальном времени математические модели нестационарной аэродинамики. В работе предложен и обоснован подход для разработки нейросетевых моделей нестационарной аэродинамики для задач динамики полета.

Построена модель момента тангажа для разрабатываемого летательного аппарата TCR с использованием рекуррентной нейронной сети NARX. Проведено сравнение данной нейронной сети и многослойного перцептрона. Показано, что нейронная сеть NARX является более подходящей для моделирования нестационарных аэродинамических характеристик в задачах динамики полета.

В работе предложен эффективный метод обучения нейронных сетей, учитывающий гетероскедастичность данных (БРГД). Метод, в основе которого лежит байесовская регуляризация, реализован в виде алгоритма обучения нейронных сетей, использующий для вычисления гессиана метод Ньютона—Гаусса в рамках метода поиска минимума целевой функции Левенберга—Марквардта. На простом синтезированном примере, а также на примере разработки нейросетевой модели коэффициента момента тангажа для TCR показано, что данный метод позволяет существенно уменьшить ошибку моделирования нейронной сетью по сравнению с методом обучения, который не учитывает гетероскедастичность данных.

Работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований (проект № 12-08-00679-а).

Список литературы

1. Annual Safety Review 2011 [Электронный ресурс] / EASA — Europa. URL: <https://www.easa.europa.eu/communications/docs/annual-safety-review>
2. Abramov N. B., Goman M. G., Khrabrov A. N., Kolesnikov E. N., Fucke L., Soemarwoto B., Smaili H. Pushing Ahead — SUPRA Airplane Model for Upset Recovery // AIAA Modeling and Simulation Technologies Conference. Minneapolis, Minnesota, USA. 2012. AIAA 2012—4631.
3. Da Ronch A., Vallespin D., Ghoreyshi M., Badcock K. J. Computation of Dynamic Derivatives Using CFD // 28th AIAA Applied Aerodynamics Conference. Chicago, IL, USA. 2010. AIAA—2010—4817.
4. Колмогоров А. Н. О представлении непрерывных функций нескольких переменных в виде суперпозиции непрерывных функций одного переменного и сложения // Докл. АН СССР. 1957. Т. 114, № 5. С. 953—956.

5. **Галушкин А. И.** Нейронные сети: основы теории. М.: Горячая линия Телеком, 2010. 496 с.

6. **Linse D. J., Stengel R. F.** Identification of aerodynamic coefficients using computational neural networks // AIAA Paper. 1992. N 92-0172. P. 1—11.

7. **Raol J. R., Jategaonkar R. V.** Aircraft parameter estimation using recurrent neural networks. A critical appraisal // AIAA Paper. 1995. N 95-3504. P. 119—128.

8. **Reisenhel P. H.** Development of nonlinear indicial model using response functions generated by a neural network // 35th Aerospace Sciences Meeting and Exhibit. 1997. AIAA 97-0337.

9. **Peyada N. K., Ghosh A. K.** Aircraft Parameter Estimation using Neural Network based Algorithm // AIAA Atmospheric Flight Mechanics Conference, Chicago, Illinois, USA. 2009. AIAA 2009-5941.

10. **Нейрокомпьютеры** в авиации (самолеты) / Под ред. В. И. Васильева, Б. Г. Ильасова, С. Т. Кусимова. Кн. 14: учеб. пособие для вузов. М.: Радиотехника, 2003. 496 с.

11. **Faller W. E., Schrek S. J.** Real-time prediction of unsteady aerodynamics: Application for aircraft control and maneuverability enhancement // IEEE Transactions on neural networks. 1995. Vol. 6, N 6. P. 1461—1468.

12. **Planckaert L.** Model of Unsteady Aerodynamic Coefficients of a Delta Wing Aircraft at High Angles of Attack // RTO-MP-069(I), RTO AVT Symposium on "Advanced Flow Management: Part A — Vortex Flows and High Angle of Attack for Military Vehicles". Loen, Norway. 2001.

13. **Игнатъев Д. И., Храбров А. Н.** Использование искусственных нейронных сетей для моделирования динамических эф-

фектов аэродинамических коэффициентов трансзвукового самолета // Ученые записки ЦАГИ, 2011. Т. XLII, № 6. С. 84—91.

14. **MacKay D. J. C.** Bayesian Interpolation // Neural Computation. 1992. № 4. P. 415—447.

15. **Foresee F. D. and Hagan M. T.** Gauss-Newton approximation to Bayesian regularization // Proceedings of the International Joint Conference on Neural Networks. 1997. P. 1930—1935.

16. **Бюшгенс Г. С., Студнев Р. В.** Аэродинамика самолета. Динамика продольного и бокового движения. М.: Машиностроение, 1979. 349 с.

17. **Mialon B., Khrabrov A., Da Ronch A., Cavagna L., Zhang M., Ricci S.** Benchmarking the prediction of dynamic derivatives: wind tunnel tests, validation, acceleration methods // AIAA Atmospheric Flight Mechanics Conference. Toronto, Ontario, Canada. AIAA Paper 2010-8244. 2010.

18. **Виноградов Ю. А., Жук А. Н., Колянко К. А., Храбров А. Н.** Математическое моделирование динамических эффектов нестационарных аэродинамических характеристик, вызванных запаздыванием развития отрыва потока на переднем горизонтальном оперении // Ученые записки ЦАГИ. 2011. Т. XLII, № 5. С. 72—81.

19. **Dzielinski A.** Neural network-based NARX models in non-linear adaptive control // Int. J. Appl. Math. Comput. Sci. 2002. Vol. 12, N 2. P. 235—240.

20. **Ahn K. K., Anh H. P. H., Phuc N. H.** Neural NNARX model identification of PAM-based robot arm using recurrent neural networks and genetic algorithm // International Symposium on Electrical & Electronics Engineering. HCM City, Vietnam. 2007. P. 104—109.

УДК 004.3.069:004.3.015

А. В. Кинарский, аспирант, e-mail: kinarskiy@yandex.ru, **С. В. Жернаков**, д-р техн. наук, проф.,
ФГБОУ ВПО "Уфимский государственный авиационный технический университет"

Стендовая отладка параметров авиационного газотурбинного двигателя на основе нейросетевых технологий

Рассматривается один из подходов к отладке параметров авиационного газотурбинного двигателя на этапах стендовых и доводочных испытаний с использованием нейросетевых технологий. Разработана инженерная методика, которая может быть использована для стендовой отладки параметров газотурбинного двигателя.

Ключевые слова: отладка, авиационный газотурбинный двигатель, нейронные сети

A. V. Kinarskiy, S. V. Zhernakov

Bench Debugging of Aviation Gas Turbine Engine Parameters Based on the Neuronet Technology

One of the approaches, based on the neuronet technology, to debugging of aviation gas turbine engine parameters during bench testing was described. The engineering technique for bench debugging of gas turbine engine parameters was developed.

Keywords: debugging, aviation gas turbine engine, neural networks

Введение

Отладка параметров авиационного двигателя является важным и необходимым этапом в процессе его эксплуатации, особенно в период планово-профилактических и ремонтных работ. Известно, что процесс проведения таких работ занимает зна-

чительное время и требует высокой квалификации обслуживающего персонала. Кроме этого, процесс отладки сопряжен с большими трудностями "согласования" в работе различных агрегатов, узлов и подсистем двигателя. Наличие погрешностей определения параметров при стендовых испытаниях,

в ряде случаев соизмеримых с допусками на отладку, создает ситуации, когда после отладки часть двигателей, имеющих параметры в пределах норм, признается негодной и проходит переотладку, и наоборот, параметры другой части двигателей, признанных годными, оказываются не в норме. В первом случае растягивается во времени и удорожается процесс отладки, во втором снижается его качество. Поэтому задача удешевления процесса стендовой отладки, повышения его качества и, как следствие, сокращения рассеивания высотно-скоростных характеристик двигателей является актуальной.

Актуальность

Анализ работ в данной области [1–3] показывает, что существующие в настоящее время алгоритмы и программы, реализующие процесс отладки, не лишены недостатков, среди которых основными являются:

- отсутствие универсальной методики, реализующей данную задачу (большинство предприятий отрасли ориентируется на собственные разработки);
- требование наличия больших объемов априорной и апостериорной информации по парку газотурбинных двигателей (ГТД);
- назначение жестких допусков для каждого отлаживаемого параметра;
- значительные временные затраты на процесс отладки параметров ГТД, связанные с необходимостью решения оптимизационной задачи: минимизация функционала качество/время и т. д.

В целях устранения указанных недостатков ниже предлагается метод решения задачи отладки параметров ГТД с использованием нейросетевых технологий. Особенность постановки и решения этой задачи заключается в том, что при построении нейронной сети, как и ранее, используется лишь экспериментально полученная информация, в то время как классические методы решения данной задачи [1–3] требуют применения среднестатистических математических моделей ГТД, описания физики протекающих процессов и т. д.

Постановка задачи

Допустим, что параметры отрегулированного, нормально функционирующего ГТД в пространстве контролируемых параметров, например, на плоскости параметров P_2^* и T_4^* (рис. 1) соответствуют заданному номинальному режиму работы двигателя: $P_2^* = P_{2ном}^*$; $T_4^* = T_{4ном}^*$.

Будем полагать, что характеристики парка исправных двигателей для тех же измеряемых параметров дают некоторый разброс относительно указанной номинальной точки, образуя эллипс (в многомерном пространстве — эллипсоид) рассеивания (рис. 1). Выход рабочей точки за пределы этого эллипса соответствует аномальным изменениям па-



Рис. 1. Эллипс рассеивания параметров двигателя

раметров индивидуального двигателя. Тогда целью отладки параметров ГТД является возвращение "выпавшей" точки в эллипс (эллипсоид) путем плавной регулировки элементов конструкции ГТД, например, путем корректировки диаметра D реактивного сопла [2].

Методика исследования

Решение данной задачи на основе метода статистического моделирования, подробно изложенного в работах [1, 3], предполагает выполнение следующих шагов:

- в процессе экспериментальных исследований (в ходе стендовых испытаний) выделяются параметры, максимально влияющие на процесс отладки: тяга R , температура газа перед турбиной T_3^* , удельный расход топлива $C_{уд}$;
- определяется выполнение условия нахождения перечисленных параметров в поле допуска:

$$\begin{aligned} R_n &\leq R \leq R_v; \\ T_{3н}^* &\leq T_3 \leq T_{3в}^*; \\ C_{уд.н} &\leq C_{уд} \leq C_{уд.в}, \end{aligned} \quad (1)$$

где индексы "н" и "в" — нижняя и верхняя допустимые границы указанных параметров соответственно;

- определяются параметры среднестатистического по парку ГТД: R_S , T_{3S}^* и $C_{удS}$;
- формируется целевая функция

$$I = (\bar{R} - \bar{R}_S)^2 + (\bar{T}_3^* - \bar{T}_{3S}^*)^2 + (\bar{C}_{уд} - \bar{C}_{удS})^2, \quad (2)$$

где значения параметров приведены к относительным (безразмерным):

$$\begin{aligned} \bar{R} &= R/R_{max}; \\ \bar{T}_3^* &= T_3^*/T_{3max}^*; \\ \bar{C}_{уд} &= C_{уд}/C_{удmax}; \end{aligned}$$

- находится такое значение варьируемого параметра (в данном случае диаметра D реактивного сопла), при котором достигается минимум целевой функции (2) при выполнении ограничений (1).

Решение данной задачи в *нейросетевом базисе* может быть представлено в виде следующей последовательности шагов:

- формирование обучающей выборки по результатам испытаний парка ГТД;
- определение границ изменения варьируемого параметра:

$$D_{\text{н}} \leq D \leq D_{\text{в}};$$

- построение нейросетевой модели среднестатистического ГТД, входными параметрами которой является диаметр \bar{D} реактивного сопла, а выходными — параметры двигателя \bar{R} , \bar{T}_3^* и $\bar{C}_{\text{уд}}$;
- построение регулировочной кривой

$$I = f(\bar{D}),$$

где I — целевая функция (невязка) (2); \bar{D} — относительное значение диаметра реактивного сопла:

$$\bar{D} = D/D_{\text{max}};$$

- вычисление требуемой поправки диаметра реактивного сопла индивидуального ГТД

$$\Delta D = D - D_{\text{опт}},$$

где $D_{\text{опт}}$ — оптимальное значение диаметра D реактивного сопла, соответствующее минимуму целевой функции (2);

- уточнение значений параметров отрегулированного двигателя \bar{R} , \bar{T}_3^* и $\bar{C}_{\text{уд}}$ для скорректированного значения диаметра реактивного сопла:

$$D = D_{\text{опт}}.$$

На рис. 2 приведена схема нейронной сети (НС), реализующей вышеописанный алгоритм.

Данная сеть хранит информационный "портрет" среднестатистического ГТД на одном из режимов его работы (например, максимальном) и при подаче на ее вход значения \bar{D} вычисляет значения приведенных параметров \bar{R} , \bar{T}_3^* и $\bar{C}_{\text{уд}}$.

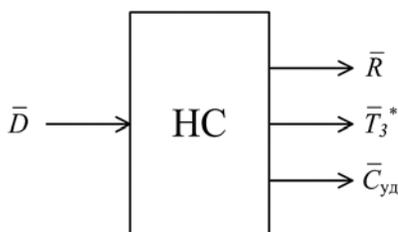


Рис. 2. Структура нейронной сети

Результаты

Рассмотрим более подробно пример решения задачи отладки.

В качестве обучающей выборки используются результаты стендовых испытаний парка авиационных двигателей на максимальном режиме работы, представленные в таблице, где значения диаметра \bar{D} реактивного сопла, тяги \bar{R} , температуры газов перед турбиной \bar{T}_3^* и удельного расхода топлива $\bar{C}_{\text{уд}}$ приведены в относительных единицах.

Параметры среднестатистического по парку двигателя:

$$\bar{R}_S = 0,9889;$$

$$\bar{C}_{\text{уд}S} = 0,9762;$$

$$\bar{T}_{3S}^* = 0,9714;$$

$$\bar{D}_S = 0,9937.$$

В качестве архитектуры НС принималась трехслойная сеть прямого распространения. График изменения ошибки обучения НС в зависимости от числа нейронов скрытого слоя приведен на рис. 3, откуда следует, что ошибка обучения НС минимальна при числе нейронов скрытого слоя, равном 3.

Наилучшая сходимость процесса обучения НС обеспечивается при использовании алгоритма сопряженных градиентов [4, 5]. Зависимость целевой функции (2) от параметра \bar{D} приведена на рис. 4.

В данном случае $D_{\text{опт}} = 0,9918$; $I_{\text{min}} = 2,14 \cdot 10^{-6}$. Допустим, что индивидуальный ГТД имеет диаметр реактивного сопла $\bar{D} = 0,9945$, тогда из графика на

Фрагмент обучающей выборки

\bar{D}	\bar{R}	$\bar{C}_{\text{уд}}$	\bar{T}_3^*
0,9982	0,9876	0,9718	0,9663
0,9982	0,9941	0,9783	0,9747
0,9927	0,9937	0,9642	0,9658
0,9927	0,9807	0,9837	0,976
0,9908	0,992	0,9978	0,9867
0,9927	0,9856	0,9794	0,9697
0,9945	0,9895	0,9696	0,9661
0,9908	0,9861	0,9902	0,9796
0,9945	0,9846	0,9653	0,961
0,9908	0,9985	0,9902	0,9881
0,9963	0,9902	0,9674	0,9708
0,9982	0,9937	0,9642	0,9674
0,9954	0,9934	0,9674	1
0,9954	0,9873	0,9685	0,9672
0,9945	0,9946	0,9783	0,9767
0,9982	0,989	0,9577	0,9569
0,989	0,9949	0,987	0,9805
0,9973	0,9988	0,9707	0,9674
0,989	0,9827	0,9805	0,9701

Заключение

Разработанный нейросетевой метод отладки параметров ГТД отличается простотой реализации и универсальностью применения [6, 7].

Применение аппарата НС оказывается эффективным при решении широкого круга плохо формализуемых задач, связанных с классификацией режимов, идентификацией, контролем, диагностикой, прогнозированием технического состояния ГТД, отладкой его параметров.

Результаты решения задачи отладки параметров ГТД показывают, что:

- процесс отладки параметров ГТД легко формализуется в нейросетевом базисе;
- для вычисления требуемого значения диаметра реактивного сопла можно воспользоваться регулировочной кривой, построенной на основе результатов обучения нейронной сети, с использованием данных стендовых испытаний парка авиационных двигателей.

Список литературы

1. **Арьков Ю. Г., Алаторцев В. П.** Об отладке основного контура ТРДФ на форсажных режимах в стендовых условиях // Испытания авиационных двигателей. Уфа: УАИ, 1978. № 6. С. 51–59.
2. **Гумеров Х. С., Алаторцев В. П., Горюнов И. М.** Оценка и отладка параметров турбореактивного двигателя на самолетах // Авиационная техника. 1998. № 3. С. 72–77.
3. **Дегтярев Ю. Д., Алаторцев В. П., Гумеров Х. С., Афанасьев И. П.** Статистическое моделирование технологического процесса отладки одновальных ГТД при стендовых испытаниях // Испытания авиационных двигателей. Уфа: УАИ, 1982. № 10. С. 97–102.
4. **Васильев В. И., Ильясов Б. Г., Кусимов С. Т., Жернаков С. В.** и др. Нейрокомпьютеры в авиации. М.: Радиотехника, 2004. 496 с.
5. **Галушкин А. И.** Теория нейронных сетей. М.: Радиотехника, 2000. 416 с.
6. **Жернаков С. В.** Применение нейросетей для отладки параметров ГТД в процессе эксплуатации // Нейрокомпьютеры и их применение: VII Всеросс. конф. с междунар. участием. М.: ИПУ, 2001. С. 161–165.
7. **Жернаков С. В.** Отладка параметров авиационного газотурбинного двигателя на основе нейронных сетей // Приборы и системы. Управление. Контроль. Диагностика. М.: Научтехлитиздат, 2003. № 12. С. 33–40.



Рис. 3. Зависимость ошибки обучения НС от числа нейронов скрытого слоя

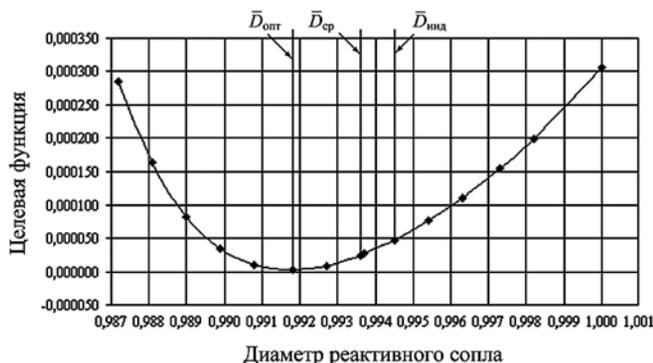


Рис. 4. Зависимость целевой функции от диаметра реактивного сопла

рис. 4 следует: $\Delta \bar{D} = 0,0027$, что эквивалентно, в свою очередь, изменению диаметра реактивного сопла:

$$\Delta D = \Delta \bar{D} D_{\max} = 1,5 \text{ мм.}$$

Отлаженному двигателю (после корректировки параметра D) будут соответствовать следующие параметры:

$$\bar{R} = 0,9876;$$

$$\bar{C}_{\text{уд}} = 0,9755;$$

$$\bar{T}_3^* = 0,9707.$$

Адрес редакции:

107076, Москва, Стромынский пер., 4

Телефон редакции журнала (499) 269-5510

E-mail: it@novtex.ru

Дизайнер *Т.Н. Погорелова*. Технический редактор *Е.В. Конова*.

Корректор *Е.В. Комиссарова*.

Сдано в набор 09.01.2014. Подписано в печать 20.02.2014. Формат 60×88 1/8. Бумага офсетная.

Усл. печ. л. 8,86. Заказ IT314. 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.