

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Том 22
2016
№ 6

ТЕОРЕТИЧЕСКИЙ И ПРИКЛАДНОЙ НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

Издается с ноября 1995 г.

УЧРЕДИТЕЛЬ
Издательство "Новые технологии"

СОДЕРЖАНИЕ

ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ И СЕТИ

- Бритвин Н. В., Карпухин Е. О.** Разработка имитационной модели атаки "человек посередине" для исследования эффективности протоколов информационного взаимодействия 403
- Богатырев В. А., Богатырев А. В.** Надежность функционирования кластерных систем реального времени с фрагментацией и резервированным обслуживанием запросов 409
- Неустроев М. Ю.** Анализ показателей эффективности и скорости обслуживания в центрах обработки вызовов 416

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ И ИЗОБРАЖЕНИЙ

- Филимонов А. Б., Фам Фьонг Кыонг.** Динамическое распознавание воздушных целей по радиолокационным дальностным портретам 423
- Поперечный П. С.** Сравнительная оценка корректирующей способности кодов Рида—Соломона и Боуза—Чоудхури—Хоквингема 431
- Дворников С. С., Дворников С. В.** Формирование сигналов с непрерывной фазой для передачи дискретной информации 435

СИСТЕМЫ АВТОМАТИЗИРОВАННОГО ПРОЕКТИРОВАНИЯ

- Петров Ю. И.** Система поддержки принятия решений по созданию конкурентоспособной технически сложной продукции 441

БАЗЫ ДАННЫХ

- Родионов А. Н.** Системная интерпретация ключевых категории концептуальных моделей баз данных 447
- Кузнецова М. В., Стрижов В. В.** Локальное прогнозирование временных рядов с использованием инвариантных преобразований 457
- Чегырбоцкий А. Н.** Автоматическая классификация объектов многомерной выборки рекурсивными методами построения кривых Пеано (на примере выборки Р. Фишера) 463

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ОРГАНИЗАЦИОННЫХ И СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ СИСТЕМАХ

- Мамедова М. Г., Джабраилова З. Г.** Методологический подход к многокритериальному принятию решений в задачах управления человеческими ресурсами . . 467

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.
РЯБОВ Г. 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.
ВИШНЕКОВ А. В., д. т. н., проф.
ГАЛУШКИН А. И., д. т. н., проф.
ДИМИТРИЕНКО Ю. И., д. ф.-м. н., проф.
ДОМРАЧЕВ В. Г., д. т. н., проф.
ЗАБОРОВСКИЙ В. С., д. т. н., проф.
ЗАГИДУЛЛИН Р. Ш., к. т. н., доц.
ЗАРУБИН В. С., д. т. н., проф.
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.
ЛЬВОВИЧ Я. Е., д. т. н., проф.
МИХАЙЛОВ Б. М., д. т. н., проф.
НЕЧАЕВ В. В., к. т. н., проф.
ПОЛЕЩУК О. М., д. т. н., проф.
СОКОЛОВ Б. В., д. т. н., проф.
ТИМОНИНА Е. Е., д. т. н., проф.
УСКОВ В. Л., к. т. н. (США)
ФОМИЧЕВ В. А., д. т. н., проф.
ШИЛОВ В. В., к. т. н., доц.

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

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

INFORMATION TECHNOLOGIES

INFORMACIONNYYE TEHNOLOGII

Vol. 22
2016
No. 6

THEORETICAL AND APPLIED SCIENTIFIC AND TECHNICAL JOURNAL

Published since November 1995

ISSN 1684-6400

CONTENTS

COMPUTING SYSTEMS AND NETWORKS

- Britvin N. V., Karpukhin E. O.** Development a Simulation Model of the Attack "Man in the Middle" for Studies the Effectiveness Data-Driven Interaction Protocols . . . 403
- Bogatyrev V. A., Bogatyrev A. V.** The Reliability of the Cluster Real-Time Systems with Fragmentation and Redundant Service Requests 409
- Neustroev M. Yu.** Analysis of Performance of Efficiency and Speed of Service in Call Centers 416

DIGITAL PROCESSING OF SIGNALS AND IMAGES

- Filimonov A. B., Pham Phuong Cuong.** Dynamic Aerial Target Recognition by Using Radar Range Profiles 423
- Poperechny P. S.** Error-Correcting Capability Estimating for RS and BCH Codes . . 431
- Dvornikov S. S., Dvornikov S. V.** Forming Signals with Continuous Phase for Broadcast Discrete Information 435

CAD-SYSTEMS

- Petrov Yu. I.** Decision Support System for Development of Technically Complex Manufactured Goods 441

DATABASE

- Rodionov A. N.** System Interpretation of Conceptual Data Model Key Categories . . 447
- Kuznetsova M. V., Strijov V. V.** Local Forecasting of Time Series with Invariant Transformations 457
- Chetyrbotsky A. N.** Automatic Classification of Objects of Multidimensional Sampling Recursive Method of Constructing Curves of Peano (for Example, Fisher's Iris Data) 463

INFORMATION TECHNOLOGIES IN THE ORGANIZATIONAL AND SOCIO-ECONOMIC SYSTEMS

- Mammadova M. H., Jabrayilova Z. G.** The Methodological Approach to Multi-Criteria Decision-Making in Human Resource Management 467

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.
Ryabov G. 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.
Vishnekov A. V., Dr. Sci. (Tech.), Prof.
Galushkin A. I., Dr. Sci. (Tech.), Prof.
Dimitrienko Yu. I., Dr. Sci. (Phys.-Math.), Prof.
Domrachev V. G., Dr. Sci. (Tech.), Prof.
Zaborovsky V. S., Dr. Sci. (Tech.), Prof.
Zagidullin R. Sh., Cand. Sci. (Tech.), Ass. Prof.
Zarubin V. S., Dr. Sci. (Tech.), Prof.
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.
Ljvovich Ya. E., Dr. Sci. (Tech.), Prof.
Mikhailov B. M., Dr. Sci. (Tech.), Prof.
Nechaev V. V., Cand. Sci. (Tech.), Ass. Prof.
Poleschuk O. M., Dr. Sci. (Tech.), Prof.
Sokolov B. V., Dr. Sci. (Tech.)
Timonina E. E., Dr. Sci. (Tech.), Prof.
Uskov V. L. (USA), Dr. Sci. (Tech.)
Fomichev V. A., Dr. Sci. (Tech.), Prof.
Shilov V. V., Cand. Sci. (Tech.), Ass. Prof.

Editors:

Bezmenova M. Yu.
Grigorin-Ryabova E. V.
Lysenko A. V.
Chugunova A. V.

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

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

УДК 004.056

Н. В. Бритвин¹, специалист, e-mail: britvin.nickita@yandex.ru,

Е. О. Карпухин^{1, 2}, канд. техн. наук, ст. науч. сотр., e-mail: ret1987@yandex.ru

¹ Федеральное государственное бюджетное учреждение науки

Центр информационных технологий в проектировании Российской академии наук, Одинцово

² Федеральное государственное бюджетное образовательное учреждение

высшего профессионального образования "Московский авиационный институт
(национальный исследовательский университет)", Москва

Разработка имитационной модели атаки "человек посередине" для исследования эффективности протоколов информационного взаимодействия

Задачи, связанные с оценкой и повышением эффективности применения протоколов и средств защиты от различных атак, актуальны для большинства предприятий, в том числе авиационной и ракетно-космической промышленности. Совершенствование протоколов информационного взаимодействия, а также средств защиты, основанных на детектировании аномалий в сети путем наблюдения за меняющимися характеристиками информационных потоков при воздействии на них злоумышленника, требует разработки адекватной модели для описания реализации такой актуальной атаки, как "человек посередине". Проведен анализ основных характеристик процесса передачи данных по сети, на которые оказывает влияние атака "человек посередине". Определены принцип воздействия атакующего на моделируемую систему передачи данных и последствия реализации атаки. Результатом работы является программное средство для имитации присутствия злоумышленника в сети, позволяющее исследовать влияние атакующего на процесс информационного взаимодействия между абонентами в виде задержек, потерь, искажений пакетов и снижения скорости передачи данных. Также приводятся методы и особенности применения разработанной модели.

Ключевые слова: атака "человек посередине", информационное взаимодействие, оптимизация протоколов, модель злоумышленника, TCP, UDP

Введение

Популярность и распространенность мобильных и сетевых технологий с каждым годом растет. Однако вместе с проникновением технологий в повседневную деятельность возрастают угрозы, связанные с перехватом данных при информационном взаимодействии двух сторон, что может повлечь за собой серьезные последствия. Одним из результатов реализации такого рода угроз является атака "человек посередине". Атака "человек посередине" (англ. Man in the middle, MitM-атака) представляет собой ситуацию, когда злоумышленник может читать, видоизменять, замещать, удалять информацию, которой обмениваются два адресата, при этом взаимодействующие стороны не ощущают присутствия третьего лица [1]. Этот вид атаки остается популярным среди хакеров, так как предоставляет возможность модификации информации в любых сферах (банковские транзакции, личная переписка, вторжение в частную жизнь).

Чтобы предупредить последствия данной атаки, необходимо комплексное решение, состоящее не только из средств защиты информации от несанкционированного доступа, но и систем обнаружения и предотвращения вторжений. Последние могут обнаруживать злоумышленника по аномалиям в сети, которые возникают из-за изменения характеристик информационных потоков в результате деятельности атакующего. К этим характеристикам следует отнести потери и искажения пакетов, увеличение задержек при передаче данных по сети, а также снижение скорости передаваемой информации. Имитация данных характеристик сети позволит воспроизвести атакующие воздействия со стороны злоумышленника на процесс информационного взаимодействия и повысить эффективность систем обнаружения вторжений.

Для имитации воздействия злоумышленника, реализующего атаку "человек посередине", на потоки данных, передаваемых по сети, можно вос-

пользоваться различными моделями, включая математические и имитационные.

Математические модели базируются на графах атак [2], а также на системах массового обслуживания [3]. Однако для применения моделей на основе графов атак требуются экспертные оценки вероятностей реализации того или иного проявления атаки (потерь, перемешивания пакетов и т. д.), которые отсутствуют для рассматриваемого вида атаки. Модель злоумышленника на основе систем массового обслуживания позволяет воспроизвести только один параметр — задержку, вносимую атакующей стороной.

Имитационная модель атаки "человек посередине" в силу возможности усложнения модели злоумышленника при сохранении принципов построения модели позволяет воспроизвести основные характеристики проявления этого вида атаки на систему передачи информации и поэтому будет рассмотрена в данной работе.

1. Особенности разработки имитационной модели атаки "человек посередине"

Наша задача состоит в разработке модели, имитирующей характеристики сети, в которой находится злоумышленник. Данная модель не предполагает операций по дешифровке сетевого трафика, если таковой зашифрован, перехват паролей и иных деструктивных действий, которые могут нарушить работоспособность сети. Разрабатываемая модель может быть использована для имитации не только действий третьего лица — злоумышленника, но и процессов, происходящих в сетях. Примерами таких процессов являются потери пакетов в беспроводных сетях из-за искажения данных при их передаче по физическому каналу, увеличение времени передачи пакетов до получателя и обратно при смене маршрута между передающей и принимающей сторонами.

Для создания приложения, которое будет имитировать атаку "человек посередине", необходимо провести анализ особенностей реализации программного обеспечения (ПО) данного типа. Были выбраны и исследованы следующие программы:

- *Dsniff* — мониторинг сети для сбора паролей, файлов и т. д.;
- *Cain&Abel* — выявление паролей путем перехвата информационных пакетов;
- *Ettercap* — ПО для перехвата и прослушивания сетевых пакетов.

Dsniff [4] специализируется на перехвате паролей, сеансов, сообщений электронной почты и адресов веб-страниц. Данная программа работает в локальных сетях и позволяет извлекать информацию из пакетов. На извлечение информации затрачивается некоторое время, что приводит к появлению задержек при передаче пакетов по сети и ограничению скорости передаваемых данных [5], а в ряде случаев и к потерям пакетов.



Рис. 1. Общая схема модели, имитирующей поведение злоумышленника в сети

Cain&Abel [6] — утилита для восстановления паролей, использующихся при входе в систему, и перехвата данных. Основным интересом представляет функция подмены связи IP и MAC адресов (ARP Spoofing), осуществление которой приводит к перенаправлению информационных потоков между абонентами через злоумышленника, что влечет модификацию сетевых пакетов с возможностью отправки дубликатов некоторых из них.

Ettercap [7] является более мощным решением в проведении атак "человек посередине". Данное ПО обладает возможностью анализа сетевого трафика и фильтрации контента, а также модификации пакетов в реальном времени. Применение фильтра приводит к перемешиванию пакетов в исходящем от злоумышленника потоке данных по сравнению с входным.

По результатам проведенного анализа представленных выше программ было решено обеспечить имитационную модель злоумышленника (рис. 1) следующим функционалом: искажение пакетов, потеря пакетов, перемешивание пакетов, повтор пакетов, ограничение пропускной способности, предоставление возможности модифицирования полезной нагрузки пакета в "реальном" времени.

Имитатор атаки "человек посередине" реализован в виде программы на языке высокого уровня C# [8] и использует четыре сокета: два на прием данных от отправителя и получателя и два на передачу данных — к получателю и отправителю. Далее будут рассмотрены основные функции имитационной модели злоумышленника на примерах работы отдельных алгоритмов.

2. Модификация пакетов в "реальном" времени

Так как наша имитационная модель работает на прикладном уровне стека TCP/IP, то это позволяет



Рис. 2. Модель модификации сетевых пакетов с использованием сторонних средств

обрабатывать трафик любого типа — от HTML-страниц до сообщений, передаваемых по сети. Основной идеей данного алгоритма является запись передаваемых данных между абонентами в текстовый файл — логирование пакетов. С помощью разработанной программы можно модифицировать пакеты с их последующей отправкой получателю (РС 2 на рис. 2) с привлечением сторонних средств. Это позволит более гибко использовать данный имитатор атаки, моделируя процесс передачи данных для абонентов в режиме "реального" времени.

3. Ограничение пропускной способности канала между абонентами

Во время атаки "человек посередине" злоумышленник вносит временную задержку, которая связана с анализом передаваемого трафика. Увеличение времени обработки каждого пакета снижает скорость передачи данных и приводит к эффекту ограничения пропускной способности канала между абонентами.

Алгоритм ограничения пропускной способности канала приближает нашу модель к функционированию реальной сети с находящимся в ней злоумышленником. Особенностью данного алгоритма является предотвращение проблем, которые харак-

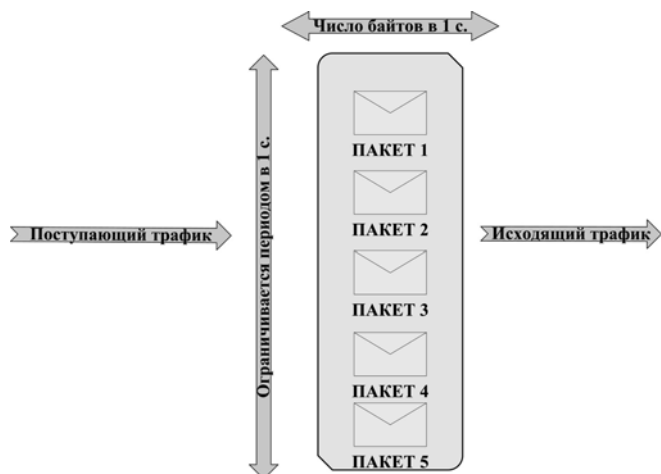


Рис. 3. Схема работы "виртуального" буфера

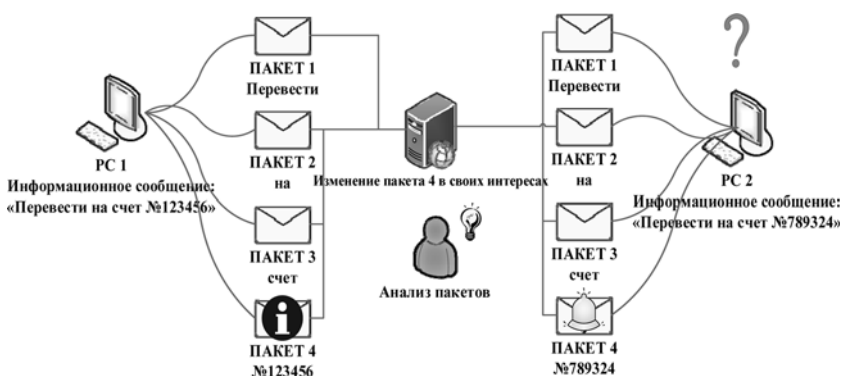


Рис. 4. Модель искажения пакетов злоумышленником

терны для стандартных решений по ограничению трафика. В стандартных алгоритмах существует проблема переполнения буфера. Это происходит в том случае, если буфер, отведенный под пакеты, заполняется быстрее, чем опустошается. Мы решили избежать данных проблем путем создания "виртуального" буфера — в нем может храниться только один пакет, который сразу же передается получателю. В случае ожидания отправки нескольких пакетов под каждый из разрешенных к передаче пакетов создается свой "виртуальный" буфер. По достижении максимального числа разрешенных к отправке пакетов за 1 с наша программа отбрасывает вновь приходящие. Число сетевых пакетов, которое наше приложение обрабатывает за 1 с определяет пропускную способность канала (рис. 3). Такой подход позволяет более практично подойти к задаче ограничения трафика.

4. Искажение пакетов злоумышленником

Довольно интересно проявляет себя алгоритм искажения пакетов. Его практическая ценность для исследования воздействия атакующего на информационные потоки в сети состоит в том, что злоумышленник может сфальсифицировать содержимое пакета в личных целях. Отличие данного алгоритма от модификации пакетов в "реальном" времени состоит в том, что при включении данного параметра от пользователя не требуется никакого вмешательства и предварительной записи набора пакетов. Программа сама изменит пакет, имитируя действие третьего лица (рис. 4).

5. Имитация воздействия атакующего на информационный поток в виде потерь, повторов и перемешивания пакетов

Потери, повторы и перемешивание сетевых пакетов происходят в результате деструктивной деятельности злоумышленника и являются следствием внедрения атакующего между абонентами. Эти функции нежелательны для злоумышленника, так как позволяют обнаружить его, но в то же время интенсивная обработка сетевого трафика требует от него огромных ресурсов, из-за чего и возникают указанные выше потери, повторы и перемешивания пакетов.

В некоторых случаях злоумышленник, анализируя трафик, может посчитать ненужным ретрансляцию пакета адресату. Это приведет к снижению скорости передачи данных между абонентами, что повысит эффективность обработки информационного потока атакующим (рис. 5). Данную функцию также можно расценивать как один из параметров, характеризующих сеть, — ведь в обычном информационном взаимодействии возможна потеря пакетов.

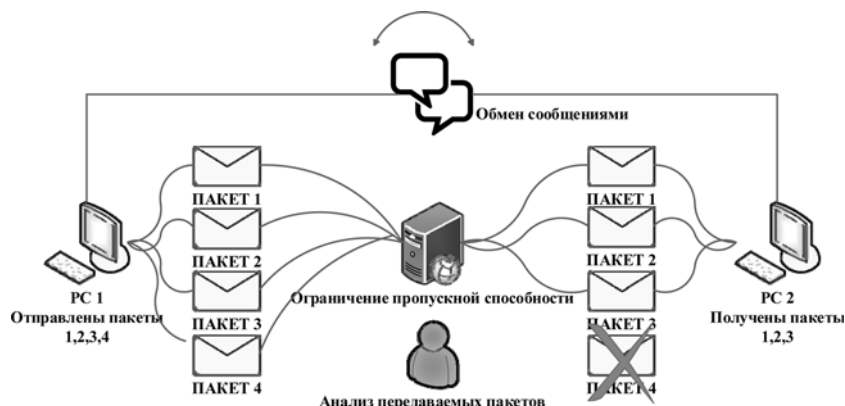


Рис. 5. Модель системы передачи информации с потерями пакетов

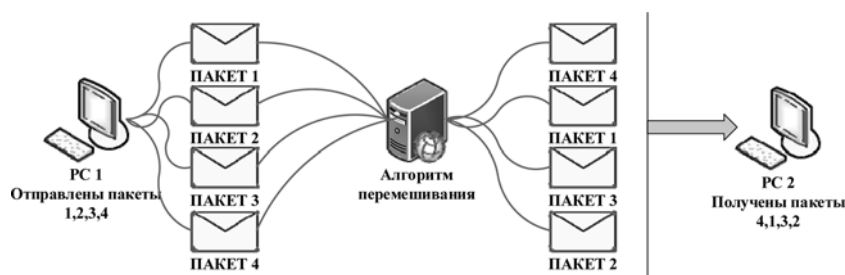


Рис. 6. Схема реализации алгоритма перемешивания пакетов

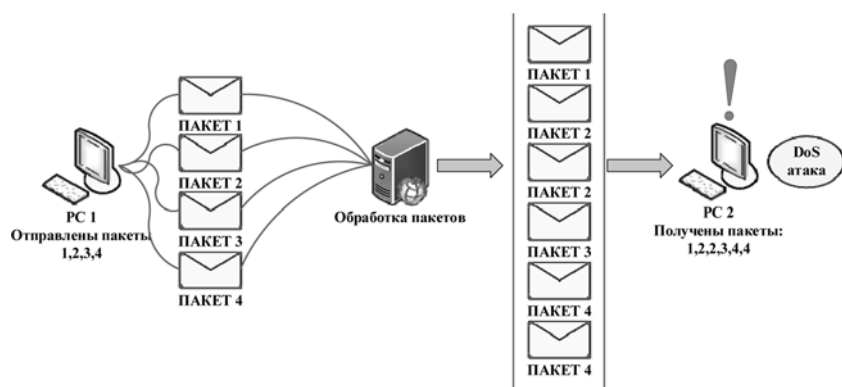


Рис. 7. Реализация алгоритма повтора пакетов

Для того чтобы как можно дольше не привлекать к себе внимание, злоумышленник должен обеспечить непрерывное взаимодействие между адресатами. В связи с этим время, потраченное на модификацию пакетов, влечет за собой перемешивание последовательности пакетов (рис. 6).

Принимающая сторона будет осуществлять прием пакетов, адресованных только ей и, возможно, по определенной маске пакета. Злоумышленник, прослушивающий трафик, может провести атаку "отказ в обслуживании" путем многократной отправки повторяющихся пакетов, причем принимающая сторона ничего не заподозрит (рис. 7).

6. Основные характеристики разработанной программы

Все изложенные выше алгоритмы имитационной модели атаки "человек посередине" реализованы в одном приложении MITM.exe (рис. 8).

В меню "Настройка" можно указать следующие характеристики имитатора атаки:

- потеря пакетов;
- повтор пакетов;
- перемешивание пакетов;
- искажение пакетов;
- ограничение скорости передаваемых данных (рис. 9).

В меню "Конфигурация" пользователь может выбрать вид используемого протокола (TCP или UDP), поменять порты и IP-адреса отправителя и получателя. После выбора и установки параметров работы имитатора атаки "человек посередине" необходим его запуск путем нажатия кнопки "Start".

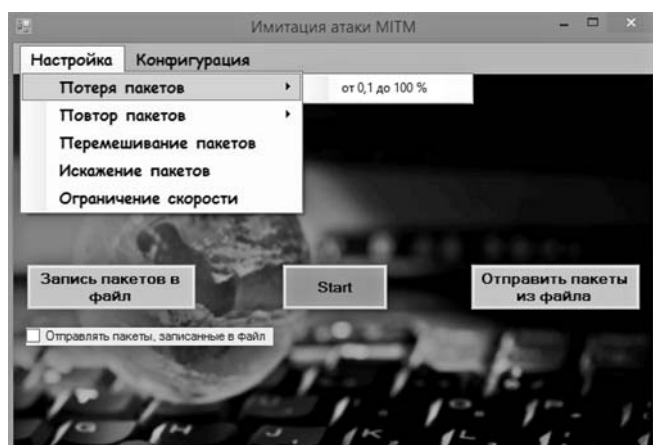


Рис. 8. Основные настройки приложения

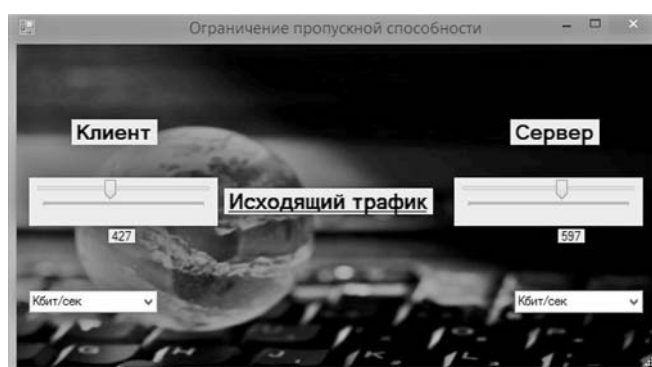


Рис. 9. Настройка ограничения скорости передаваемых данных

7. Методы и особенности применения разработанной модели

Один из вариантов применения разработанной модели, который представлен на рис. 10, заключается в передаче данных между отправителем и получателем через имитатор атаки, который расположен между ними и воспроизводит воздействие злоумышленника на сетевые протоколы, такие как TCP и UDP, в виде потерь, повторов, перемешивания и искажения пакетов, а также ограничения скорости передаваемых данных. Чтобы симулировать атаку, необходимо подключить отправителя и получателя к имитатору атаки "человек посередине", указав IP-адрес и порт последнего в одной из программ, протоколы которой предполагается протестировать, например iperf. Для имитации атаки в программе MITM.exe указывается процент пакетов (от 0,1 до 100 %), которые будут отброшены или продублированы, число искаженных и перемешанных пакетов, а также проводится ограничение скорости передаваемых данных (см. рис. 9). Меняя данные параметры на имитаторе атаки "человек посередине", можно изучить влияние атакующего на такие характеристики, как скорость и надежность информационного взаимодействия между абонентами, которые, в свою очередь, определяются механизмами управления перегрузкой и контроля целостности заголовков сетевых пакетов и данных, расположенных в этих пакетах. Рассмотренный метод применения имитационной модели атаки "человек посередине" пригодится для исследования новых протокольных элементов и компонентов на предмет их эффективности противодействия атакующему.

Еще одним вариантом применения разработанной имитационной модели является перехват данных с их последующей записью и ретрансляцией записанных пакетов в сеть. Для этого осуществляется перехват сеанса между взаимодействующими сторонами с помощью программы-сниффера, а затем эти пакеты через некоторое время воспроизводятся с помощью опции "Отправлять пакеты, записанные в файл". Для записи пакетов также можно использовать разработанную программу MITM.exe, предварительно выполнив настройку отправителя и получателя, как было описано в первом варианте применения имитационной модели. При использовании в сети коммутаторов фирмы Cisco серий 4500 и 6500, повторной отправкой пакетов можно протестировать работу системы противодействия ключевым узлам атаки "человек посередине", таким как DHCP Snooping [9] и IP Source Guard [10].

Заключение

Поставленная во введении задача по разработке имитационной модели атаки "человек посередине"

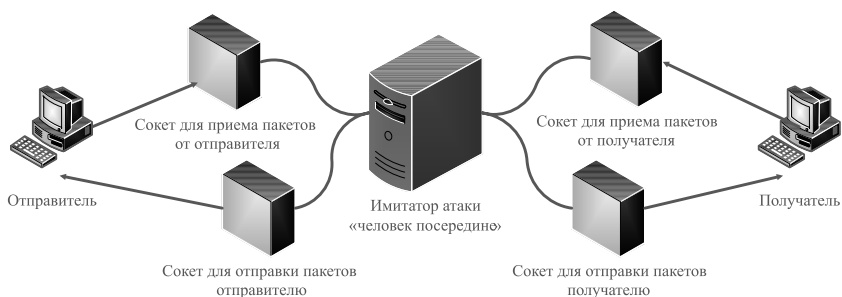


Рис. 10. Структурная схема имитатора атаки для исследования влияния воздействия злоумышленника на протоколы информационного взаимодействия

была достигнута полностью. Ее результатом является программа MITM.exe. Она написана под ОС Windows со средой .NET Framework 4.5. Данная модель может использоваться для исследования методов и средств защиты, основанных на детектировании аномалий в сети, путем наблюдения за меняющимися характеристиками информационных потоков при воздействии на них злоумышленника. Также она будет полезна разработчикам сетевых протоколов, так как позволяет оценить воздействие злоумышленника на работу протокольной машины и вносить необходимые дополнения в ее функционирование.

Стоит отметить, что компонент программы, описанный в п. 2, делает ее актуальной в течение длительного времени. Не важно, какие протоколы прикладного уровня будут использовать абоненты, так как злоумышленник всегда сможет обрабатывать пакеты по своему усмотрению в силу функционирования имитационной модели на транспортном уровне. Также необходимо отметить, что со скорым выходом Visual Studio 2015 программа получит кроссплатформенность. Этот статус позволит использовать ее на большинстве современных вычислительных устройств.

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 14-07-31247 мол_a.

Список литературы

1. **Понятие** термина MITM-атака [Электронный ресурс] // Сайт It-Sektor: [Электронный ресурс], 2015. URL: <http://it-sektor.ru/ponyatie-termina-mitm-ataka.html> (дата обращения 06.03.2015).
2. **Sheyner O., Haines J., Jha S., Lippmann R., Wing J.** Automated Generation and Analysis of Attack Graphs // Proceedings of the IEEE Symposium on Security and Privacy. Oakland, CA, USA, 2002. P. 273–284.
3. **Мазепа Р. Б., Михайлов В. Ю., Большов О. А., Карпухин Е. О., Корнилов А. М.** Методы обеспечения безопасности информационных процессов. М.: МАИ-ПРИНТ, 2012. 168 с.
4. **Сниффинг** сети на коммутаторах [Электронный ресурс] // Сайт Nag.ru: [Электронный ресурс], 2015. URL: <http://nag.ru/articles/reviews/15770/sniffing-seti-na-kommutatorah.html> (дата обращения 06.03.2015).
5. **Шоуэнберг Р.** Атака на банки. Аналитическая статья "Лаборатории Касперского" [Электронный ресурс], 2015. URL: http://www.itsec.ru/articles2/Inf_security/ataka-na-banki (дата обращения 06.03.2015).

6. **Использование** сниффера Cain & Abel [Электронный ресурс] // Сайт Записки it Guy: [Электронный ресурс], 2015. URL <http://itguy-note.blogspot.de/2010/03/cain-abel.html> (дата обращения: 06.03.2015).

7. **Анализ** безопасности компьютерных сетей [Электронный ресурс] // Сайт Zen Way: [Электронный ресурс], 2015. <http://zenway.ru/page/ettercap> (дата обращения 06.03.2015).

8. **Нейгел К., Ив'ен Б., Глинн Д., Уотсон К., Скиннер М.** С# 5.0 и платформа .Net 4.5 для профессионалов: Пер. с англ. М.: Издат. дом "Вильямс", 2014. 1440 с.

9. **Конфигурация** DHCP Snooping [Электронный ресурс] // Сайт Cisco.com: [Электронный ресурс], 2015. URL: [.cisco.com/c/en/us/td/docs/switches/datacenter/sw/4_1/nx-os/security/configuration/guide/sec_nx-os-cfg/sec_dhcpsnoop.html](http://www.cisco.com/c/en/us/td/docs/switches/datacenter/sw/4_1/nx-os/security/configuration/guide/sec_nx-os-cfg/sec_dhcpsnoop.html) (дата обращения 05.11.2015).

10. **Конфигурация** IP Source Guard [Электронный ресурс] // Сайт Cisco.com: [Электронный ресурс], 2015. URL: http://www.cisco.com/c/en/us/td/docs/switches/lan/catalyst3650/software/release/3e/security/configuration_guide/b_sec_3e_3650_cg/b_sec_3e_3650_cg_chapter_01101.pdf (дата обращения 05.11.2015).

N. V. Britvin¹, expert, britvin.nickita@yandex.ru,

E. O. Karpukhin^{1, 2}, Senior Staff Scientist, ret1987@yandex.ru

¹ Design Information Technologies Center, Russian Academy of Sciences

² Moscow Aviation Institute (National Research University)

Development a Simulation Model of the Attack "Man in the Middle" for Studies the Effectiveness Data-Driven Interaction Protocols

The tasks, related to evaluating and improving the effectiveness of protocols and protection means from various attacks, are relevant for majority of enterprises, including aerospace industry. Improving data-driven interaction protocols and protection means, based on the detection of anomalies in the network by monitoring the changing characteristics of information flows when exposed to an attacker, requires the development of an adequate model to describe the actual implementation of such an attack, as a "man in the middle". To achieve this target, in this work was carried the analysis of the main characteristics of the process data transferring over a network, that are affected by the attack "man in the middle". It defines the principles of attacks to simulated data transmission system and the implications of this attack. The result of this work is a software tool to simulate the presence of attacker in the network, allowing the attacker to investigate the influence on the process of information exchange between subscribers in the form of delays, loss, distortion and reduce packet data rate. It also provides methods and features of application developed model.

Keywords: attack "man in the middle", data-driven interaction, protocol optimization, model of the attacker, TCP, UDP

References

1. **Ponjatie** termina MITM-ataka [Definition of the term MITM-attack]. Site It-Sektor: [electronic resource], 2015. URL: <http://it-sektor.ru/ponyatie-termina-mitm-ataka.html> (date of the application 06.03.2015).

2. **Sheyner O., Hames J., Jha S., Lippmann R., Wing J. M.** Automated Generation and Analysis of Attack Graphs // *Proceedings of the IEEE Symposium on Security and Privacy*. Oakland, CA, USA, 2002. P. 273–284.

3. **Mazepa R. B., Mihajlov V. Ju., Bol'shov O. A., Karpukhin E. O., Kornilov A. M.** *Metody obespechenija bezopasnosti informacionnyh processov* [Methods of safety information processes]. Moscow: MAI-PRINT, 2012, 168 p.

4. **Sniffing** seti na kommutatorah [Sniffing network switches] // Site Nag.ru: [electronic resource], 2015. URL: <http://nag.ru/articles/reviews/15770/sniffing-seti-na-kommutatorah.html> (date of the application; 06.03.2015).

5. **Shouenberg R.** *Ataka na banki. Analiticheskaja stat'ja "Laboratorii Kasperskogo"* [Attacks on banks. The analytical article "Kaspersky Lab"], 2015. URL: http://www.itsec.ru/articles2/Inf_security/ataka-na-banki (date of the application 06.03.2015).

6. **Ispol'zovanie** sniffera Cain & Abel [Using sniffer Cain & Abel]. Site Zapiski it Guy: [electronic resource], 2015. URL <http://itguy-note.blogspot.de/2010/03/cain-abel.html> (date of the application 06.03.2015).

7. **Analiz** bezopasnosti komp'yuternyh setej [Analysis of the security of computer networks]. Site Zen Way: [electronic resource], 2015. <http://zenway.ru/page/ettercap> (date of the application 06.03.2015).

8. **Nejgel K., Iv'ен B., Glinn D., Uotson K., Skinner M.** *С# 5.0 i platforma .Net 4.5 dlja professionalov: Per. s angl. Moscow: OOO "I.D. Vil'jams", 2014. 1440 pp.: il. [Christian Nagel, Bill Iven, Jay Glynn, Karli Watson, Morgan Skinner. С # 5.0 and .Net 4.5 platform for professionals, Trans, from English. M.: OOO "ID Williams"]*

9. **Konfiguracija** DHCP Snooping [Configuring DHCP Snooping]. Site Cisco.com: [electronic resource], [2015]. URL: [.cisco.com/c/en/us/td/docs/switches/datacenter/sw/4_1/nx-os/security/configuration/guide/sec_nx-os-cfg/sec_dhcpsnoop.html](http://www.cisco.com/c/en/us/td/docs/switches/datacenter/sw/4_1/nx-os/security/configuration/guide/sec_nx-os-cfg/sec_dhcpsnoop.html) (date of the application 05.11.2015).

10. **Konfiguracija** IP Source Guard [Configure IP Source Guard]. Site Cisco.com: [electronic resource], 2015. URL: http://www.cisco.com/c/en/us/td/docs/switches/lan/catalyst3650/software/release/3e/security/configuration_guide/b_sec_3e_3650_cg/b_sec_3e_3650_cg_chapter_01101.pdf (date of the application 05.11.2015).

В. А. Богатырев, д-р техн. наук, проф., e-mail: vladirnir.bogatyrev@gmail.com, **А. В. Богатырев**, аспирант Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики

Надежность функционирования кластерных систем реального времени с фрагментацией и резервированным обслуживанием запросов

Предложены модели функциональной надежности кластера при обслуживании запросов реального времени с разделением запроса на фрагменты (части) и возможностью их параллельного резервированного выполнения в разных узлах кластера. Проанализировано влияние фрагментации и резервного выполнения копий фрагментов запроса на вероятность выполнения кластером запроса в директивные сроки с учетом возможных сбоев, отказов и ошибок вычислений.

Ключевые слова: надежность, резервированные вычисления, реальное время, кластер, запросы, система массового обслуживания, очередь, фрагментация

Введение

Проектирование компьютерных систем ответственного назначения требует обеспечения высокой производительности, отказоустойчивости, живучести, информационной и функциональной надежности и безопасности [1–6].

Надежность компьютерных систем реального времени определяется не только безотказностью и отказоустойчивостью структуры, но и устойчивостью вычислительного процесса, требующей безошибочности и своевременности обслуживания функциональных запросов [7–10].

Устойчивость (надежность) вычислительного процесса к отказам и ошибкам может быть повышена в результате отправки копий поступающих в компьютерную систему запросов на резервированное обслуживание в несколько узлов одновременно [11–14].

Эффективность резервирования запросов во многом зависит от интенсивности потока запросов и выбора дисциплины диспетчеризации. Для систем, представимых многоканальными системами массового обслуживания с общей очередью бесконечной длины, в [11, 12] показано, что резервированное обслуживание запросов при малой загрузке позволяет сократить среднее время ожидания запросов. С учетом этого в [11] предложена дисциплина обслуживания, названная "широковещательное обслуживание с копированием запроса" (Broadcasting with a customer Copying), при которой каждый запрос в момент поступления направляется на резервированное выполнение в свободные узлы (приборы). Если число занятых приборов в момент поступления запроса меньше некоторого порогового значения, то запрос резервируется (копируется) во все свободные узлы, иначе он обслуживается в одном узле, причем запрос считается обслуженным при успешном выполнении хотя бы одной из его копий. Эффективность обслуживания с резервированием запросов в [11, 12] оценивается по критерию снижения среднего времени ожидания запросов.

Для компьютерных систем, функционирующих в реальном времени, эффективность определяется не столько средним временем ожидания запросов, сколько вероятностью безошибочного выполнения запросов в директивные сроки (своевременность безошибочного обслуживания).

Своевременность обслуживания запросов в вычислительных системах реального времени может оцениваться по вероятности времени ожидания запросов в очереди, меньшего предельно допустимого времени t_0 [13, 14] или по вероятности получения результатов к директивному сроку T_0 .

Для повышения устойчивости к отказам и ошибкам вычислительного процесса в кластерных компьютерных системах реального времени копии запросов могут распределяться на выполнение в несколько узлов кластера, чем достигается резервирование вычислительного процесса.

В кластерных системах, серверные узлы которых соответствуют локальным системами массового обслуживания с локальными очередями (в отличие от многоканальных систем обслуживания с общей очередью [11]), возможен существенный разброс времен ожидания в разных очередях, что потенциально может привести к увеличению вероятности своевременного выполнения резервированных запросов хотя бы в одном узле кластера, несмотря на увеличение нагрузки узлов из-за резервирования запросов.

В работе [13] исследованы варианты дисциплин резервированного обслуживания запросов в кластерных системах реального времени, предусматривающих распределение каждого запроса на выполнение в два или несколько (k) узлов, с выдачей результатов в директивный срок, а в [14] поставлена и решена задача оптимизации резервированного обслуживания запросов реального времени в системах кластерной архитектуры.

Вместе с тем, в работах [13, 14] не проанализированы возможности повышения вероятности свое-

временного и безошибочного резервированного обслуживания запросов реального времени в результате их фрагментации (разделения на части) при распараллеливании выполнения фрагментов разными узлами кластера.

Цель работы — исследование возможностей повышения вероятности своевременного и безошибочного обслуживания запросов реального времени в результате их фрагментации и резервирования.

Для достижения цели ставится задача исследования дисциплин обслуживания, предусматривающих:

- возможность разделения (фрагментации) запроса на части, выполняемые в разных узлах кластера (распараллеливание обслуживания запроса);
- возможность фрагментации запроса с резервированным выполнением копий фрагментов в разных узлах кластера.

Условием успешного выполнения запроса в кластерных системах реального времени для исследуемых дисциплин обслуживания является своевременное безошибочное выполнение запросов, при котором каждый фрагмент запроса должен быть безошибочно выполнен к директивному сроку хотя бы в одном из выделенных для этого узлов.

Ограничения применения исследуемой дисциплины обслуживания могут быть вызваны невозможностью распараллеливания выполнения запросов на разные узлы кластера. К сдерживающим факторам эффективного применения рассматриваемой дисциплины могут быть отнесены дополнительные издержки на диспетчеризацию запросов и невозможность деления задачи (запроса) на равные части (фрагменты), число которых может быть ограничено.

Предлагаемая дисциплина может быть потенциально эффективно применена для класса потенциально распараллеливаемых задач, например, для матричных вычислений, для обработки массивов данных, вычислений, связанных с большим числом перебираемых вариантов решений.

Сложность проводимых исследований обусловлена наличием технического противоречия. Действительно, резервированное выполнение запросов несколькими узлами приводит к повышению надежности вычислений при необходимости получения безошибочного результата хотя бы одним узлом, но вызывает возрастание загрузки узлов и, как следствие, увеличение вероятности задержки запросов в очереди сверх предельно допустимого времени t_0 . В то же время резервированное выполнение запросов разными узлами, с учетом стохастичности обслуживания, может повысить вероятность выполнения запроса в требуемый срок хотя бы одним из узлов. Разделение запросов на части с их выполнением в разных узлах приводит к ускорению вычислений, но в силу той же стохастичности обслуживания может привести к недопустимой задержке выполнения одной из частей при необходимости безошибочного выполнения в директивный срок всех фрагментов запроса.

В качестве объекта исследований рассматривается вычислительный кластер, объединяющий n идентичных компьютерных узлов (серверов), в каждом из которых организуется собственная локальная очередь запросов. Будем считать известными: среднее время выполнения запросов v , интенсивность отказов λ_0 и ошибок вычислений λ_1 в узлах. Поступающий в кластер запрос может быть распределен на обслуживание в любой компьютерный узел, представляемый системой массового обслуживания типа М/М/1 с бесконечной очередью [15]. Предположим, что запрос может быть разделен на s равных частей, причем выполнение различных частей может осуществляться независимо (параллельно) в разных узлах кластера, при этом время выполнения фрагмента (и всего запроса с учетом распараллеливания) будет равно v/s . Каждый запрос или его фрагмент может копироваться и направляться на резервное выполнение в очередь разных узлов кластера. При поступлении запроса (или его фрагмента) в очередь узла в нем запускается таймер, отсчитывающий директивное (предельно допустимое) время ожидания t_0 , при этом если фрагментация запроса не проводится, то $t_0 = T_0 - v$, а если проводится, то $t_0 = T_0 - (v/s)$. Считывание результатов осуществляется после отсчета таймерами директивного времени T_0 . Потерями времени на диспетчеризацию запросов, их фрагментацию и дефрагментацию, а также на контроль безошибочности результатов будем пренебрегать.

Постановка задачи исследования дисциплины обслуживания с фрагментацией и резервированием обслуживания запросов

Для узлов кластера, представимых системами массового обслуживания типа М/М/1 с бесконечной очередью, вероятность r того, что время ожидания запросов без фрагментации и резервированного обслуживания меньше предельно допустимого значения t_0 , вычисляется [15] по формуле

$$r_0 = 1 - \lambda v e^{-t_0(v^{-1} - \lambda)}. \quad (1)$$

В случае объединения n компьютеров в кластер при балансировке их нагрузки формула (1) оценки вероятности своевременности обслуживания запросов может быть модифицирована [13, 14]:

$$r_1 = 1 - \frac{\lambda v}{n} e^{-t_0\left(\frac{1}{v} - \frac{\lambda}{n}\right)}. \quad (2)$$

Формула (2) позволяет вычислить вероятность своевременности обслуживания запросов, когда каждый запрос направляется на выполнение в один из компьютерных узлов, а резервирование выполнения копий запроса не происходит.

С учетом увеличения загрузки узлов при резервировании вычислений в k узлах вероятность превышения времени ожидания запросов предела t_0 в некотором (конкретном) узле вычисляется как

$$r_2 = 1 - \frac{\Lambda vk}{n} e^{-t_0 \left(\frac{1}{v} - \frac{\Lambda k}{n} \right)},$$

где $\rho = \Lambda vk/n$ — загрузка узла.

При независимости процессов обслуживания в разных узлах вероятность того, что хотя бы в одном из k узлов, выполняющих запрос, его задержка в очереди меньше предельно допустимой величины t_0 , определяется выражением

$$R_2 = 1 - (1 - r_2)^k = 1 - \left[\frac{\Lambda vk}{n} e^{-t_0 \left(\frac{1}{v} - \frac{\Lambda k}{n} \right)} \right]^k,$$

отсюда можно найти среднее время ожидания

$$w = \int_0^{\infty} \left(\left[\frac{\Lambda vk}{n} e^{-t_0 \left(\frac{1}{v} - \frac{\Lambda k}{n} \right)} \right]^k \right) dt.$$

Результаты исследований [11–14] не затрагивают фрагментацию запросов с возможностью резервного выполнения фрагментов разными узлами кластера.

Цель работы — исследование возможностей повышения эффективности кластерных систем реального времени в результате фрагментации запросов и резервного выполнения копий фрагментов в различных узлах кластера, когда требуется безошибочное обслуживание запросов к директивным срокам (с учетом предельно допустимого времени ожидания).

В статье ставится задача построения моделей кластера реального времени и исследование эффективности резервированного выполнения фрагментированных запросов несколькими узлами при требовании выдачи результатов в директивные сроки.

Проводимые исследования предполагают анализ:

- своевременности резервированного обслуживания запросов с возможностью фрагментацией в предположении отсутствия отказов, сбоев и ошибок во время пребывания запроса (идеализированный случай для анализа кластера как системы массового обслуживания);
- надежности вычислительного процесса при требовании своевременности и безошибочности выдачи результатов резервированного обслуживания запросов с их возможной фрагментацией в условиях отказов, сбоев и ошибок выполнения запросов.

Показателем эффективности рассматриваемых кластерных систем является вероятность безошибочного обслуживания резервированного запроса реального времени с выполнением каждого выделяемого в нем фрагмента хотя бы в одном из узлов кластера за время, меньшее предельно допустимого времени ожидания обслуживания.

Вероятность своевременности резервированного обслуживания фрагментированных запросов

Рассмотрим оценку вероятности своевременности резервированного вычислительного процесса в предположении безотказности узлов и идеальности контроля, когда к моменту выполнения запроса без дополнительных задержек вырабатывается информация о правильности (достоверности) вычислений. Издержками на диспетчеризацию, фрагментацию и дефрагментацию запросов будем пренебрегать. Будем считать, что все запросы могут при фрагментации быть разделены на s равных частей.

При разделении каждого запроса на s равных частей (фрагментов), независимо (возможно одновременно) выполняемых разными узлами кластера, интенсивность фрагментированных запросов относительно исходного потока запросов Λ увеличивается в s раз. Вместе с тем, в результате сокращения среднего времени выполнения фрагментированных запросов в s раз в итоге средняя нагрузка узлов остается равной $\rho = \Lambda v/n$.

Таким образом, при фрагментации запросов без их резервированного выполнения разными узлами вероятность превышения времени ожидания запросов предела t_0 в некотором (одном, конкретном) узле определяется выражением

$$r_3 = 1 - \frac{\Delta v}{n} \exp\left(-t_0 \left(\frac{s}{v} - \frac{s\Lambda}{n} \right)\right),$$

а при резервированном независимом выполнении каждого фрагмента запроса в k узлах кластера (что увеличивает загрузку узла в k раз) вероятность превышения времени ожидания запросов предела t_0 в некотором (конкретном) узле вычисляется как

$$r_{23} = 1 - \frac{\Delta vk}{n} \exp\left(-t_0 \left(\frac{s}{v} - \frac{s\Lambda k}{n} \right)\right).$$

При фрагментации запроса на s частей время ожидания всех s фрагментов не должно превосходить директивный срок ожидания t_0 , поэтому вероятность своевременности обслуживания фрагментированных запросов без их резервированного обслуживания в кластере вычисляется по формуле

$$R_3 = r_3^s = \left[1 - \frac{\Delta v}{n} \exp\left(-t_0 \left(\frac{s}{v} - \frac{s\Lambda}{n} \right)\right) \right]^s.$$

При k -кратном резервировании фрагментированного запроса, т. е. с независимым вычислением каждого из s фрагментов в k узлах кластера, вероятность своевременности обслуживания фрагментированного запроса при условии, что время ожидания каждого фрагмента не превосходит предельное время ожидания t_0 хотя бы в одном из k узлов, определяется выражением

$$R_{23} = [1 - (1 - r_{23})^k]^s = \left\{ 1 - \left[\frac{\Lambda kv}{n} \exp\left(-t_0 \left(\frac{s}{v} - \frac{ks\Lambda}{n} \right)\right) \right]^k \right\}^s.$$

Среднее время ожидания запросов для рассматриваемых случаев можно вычислить следующим образом:

$$w_3 = \int_0^{\infty} \left(1 - \left[1 - \frac{\Lambda v}{n} \exp\left(-t_0 \left(\frac{s}{v} - \frac{s\Lambda}{n} \right)\right) \right]^s \right) dt;$$

$$w_{23} = \int_0^{\infty} \left(1 - \left\{ 1 - \left[\frac{\Lambda k v}{n} \exp\left(-t_0 \left(\frac{s}{v} - \frac{k s \Lambda}{n} \right)\right) \right]^k \right\}^s \right) dt.$$

С учетом того, что при обслуживании запросов с фрагментацией и без нее допустимое время ожидания определяется соответственно как $t_0 = T_0 - v$ и $t_0 = T_0 - v/s$, вероятности своевременного обслуживания запросов для исследуемых вариантов организации вычислительного процесса оценим как

$$R_1 = r_1 = 1 - (\Lambda/n)v \exp(-(T_0 - v)(v^{-1} - \Lambda/n));$$

$$R_2 = 1 - (1 - r_2)^k =$$

$$= 1 - \left[\frac{\Lambda k v}{n} \exp\left\{-\left(T_0 - v\right)\left(\frac{1}{v} - \frac{\Lambda k}{n}\right)\right\} \right]^k; \quad (3)$$

$$R_3 = r_3^s = \left[1 - \frac{\Lambda v}{n} \exp\left(-\left(T_0 - \frac{v}{s}\right)\left(\frac{s}{v} - \frac{s\Lambda}{n}\right)\right) \right]^s; \quad (4)$$

$$R_{23} = [1 - (1 - r_{23})^k]^s =$$

$$= \left\{ 1 - \left[\frac{\Lambda k v}{n} \exp\left(-\left(T_0 - \frac{v}{s}\right)\left(\frac{s}{v} - \frac{k s \Lambda}{n}\right)\right) \right]^k \right\}^s. \quad (5)$$

Последняя формула является обобщением формул для оценки вероятности допустимого времени ожидания $R_1 - R_3$ с учетом значений k и s .

Вероятность своевременности и безошибочности обслуживания запросов

Определим вероятности своевременного и безошибочного выполнения запросов в кластере для рассматриваемых дисциплин обслуживания с резервированием фрагментированных запросов в предположении, что отказ или ошибка выполнения запроса могут возникать от момента его помещения в очередь узла до момента считывания результата (этот интервал времени составляет T_0).

Вероятность своевременности и безошибочности за время T_0 нахождения запроса (без резервирования и фрагментации) в узле кластера определяется выражением [13]:

$$B = r_1 \exp(-\lambda T_0) =$$

$$= [1 - (\Lambda/n)v \exp(-(T_0 - v)(v^{-1} - \Lambda/n))] \exp(-\lambda T_0),$$

где r_1 — вероятность своевременности обслуживания запроса в узле кластера, вычисляемая по формуле (2), а $\exp(-\lambda T_0)$ — вероятность безошибочности и безотказности узла за время нахождения в нем запроса T_0 , при этом $\lambda = \lambda_0 + \lambda_1$ — суммарная интенсивность отказов узлов и ошибок их вычислений.

Вероятность своевременного и безошибочного обслуживания запросов в кластере при резервиро-

вании запросов без фрагментации вычисляется по формуле

$$P_2 = 1 - (1 - B_2)^k,$$

где B_2 — вероятность своевременного и безошибочного выполнения запроса в одном узле кластера, определяемая выражением

$$B_2 = r_2 \exp(-\lambda T_0) =$$

$$= \left[1 - \frac{\Lambda k v}{n} \exp\left(-\left(T_0 - v\right)\left(\frac{1}{v} - \frac{\Lambda k}{n}\right)\right) \right] \exp(-\lambda T_0). \quad (6)$$

Вероятность своевременного и безошибочного обслуживания запросов в кластере с фрагментацией запросов без их резервирования вычисляется как

$$B_3 = (r_3 \exp(-\lambda T_0))^s =$$

$$= \left[\left\{ 1 - \frac{\Lambda v}{n} \exp\left(-\left(T_0 - \frac{v}{s}\right)\left(\frac{s}{v} - \frac{s\Lambda}{n}\right)\right) \right\} \exp(-\lambda T_0) \right]^s. \quad (7)$$

Вероятность своевременного и безошибочного обслуживания запросов в кластере при резервировании и фрагментации запросов в предположении, что все запросы могут быть разделены на s равных частей, выполняемых в разных узлах, при независимом резервном обслуживании в k узлах каждого из s фрагментов, вычисляется по формуле

$$P_2 = [1 - (1 - B_{23})^k]^s,$$

где B_{23} — вероятность своевременного и безошибочного выполнения фрагмента запроса в одном узле кластера, определяемая выражением

$$B_{23} = \left\{ 1 - \frac{k \Lambda v}{n} \exp\left(-\left(T_0 - \frac{v}{s}\right)\left(\frac{s}{v} - \frac{k s \Lambda}{n}\right)\right) \right\} \exp(-\lambda T_0). \quad (8)$$

Вероятность своевременности и безошибочности обслуживания запросов с различной возможностью фрагментации

Рассмотренные выше модели соответствуют идеализированному случаю, когда все запросы допускают возможность фрагментации на одинаковое число частей, выполняемых разными узлами кластера.

Рассмотрим теперь случай неоднородности запросов по возможному числу частей, выделяемых при их фрагментации, и по кратности их резервирования обслуживания в кластере.

Допустим, что в общем случае имеется n типов запросов, которые допускают разбиение соответственно на $i = 1, 2, 3, \dots, n$ фрагментов, независимо выполняемых в разных узлах кластера (запрос i -го типа допускает разбиение на i независимо выполняемых частей). Будем считать заданными кратности резервирования соответствующих типов запросов ($g_1, g_2, g_3, \dots, g_n$) и известными вероятности поступления соответствующих типов запросов

($b_1, b_2, b_3, \dots, b_n$) $\left(\sum_{i=1}^n b_i \right) = 1$, причем некоторые из вероятностей b_i могут быть нулевыми.

Математическое ожидание числа выделяемых фрагментов

$$S = \sum_{i=1}^n b_i i,$$

а средняя кратность резервированного выполнения запросов

$$K = \sum_{i=1}^n b_i k_i.$$

Интенсивность запросов, поступающих в некоторый узел, и интенсивность их обслуживания при фрагментации определим как $S\Lambda/n$ и S/v , а при резервировании запросов без фрагментации эти интенсивности равны $K\Lambda/n$ и $1/v$, если при этом также реализуется фрагментация, то рассматриваемые интенсивности определяются как $K S\Lambda/n$ и S/v .

С учетом того, что при обслуживании запросов с фрагментацией допустимое время ожидания для запросов i -го типа $t_0 = T_0 - (v/s_i)$, а без фрагментации — $t_0 = T_0 - v$, на основе модификации формул (3)—(5) вычислим вероятности своевременного обслуживания запроса i -го типа для вариантов:

а) с резервированием выполнения запросов без их фрагментации

$$A_{2i} = 1 - \left[\frac{\Lambda v K}{n} \exp\left\{-\left(T_0 - v\right)\left(\frac{1}{v} - \frac{\Lambda K}{n}\right)\right\} \right]^{k_i};$$

б) с фрагментацией без резервирования

$$A_{3i} = \left[1 - \frac{\Lambda v}{n} \exp\left(-\left(T_0 - \frac{v}{s_i}\right)\left(\frac{S}{v} - \frac{S\Lambda}{n}\right)\right) \right]^{s_i};$$

в) с резервированием и фрагментацией выполнения запросов

$$A_{23i} = \left\{ 1 - \left[\frac{\Lambda K v}{n} \exp\left(-\left(T_0 - \frac{v}{s_i}\right)\left(\frac{S}{v} - \frac{K S \Lambda}{n}\right)\right) \right]^{k_i} \right\}^{s_i}.$$

На основе модификации формул (6)—(8) вычислим вероятности своевременного и безошибочного обслуживания запроса i -го типа для вариантов:

а) с резервированным выполнением запросов без их фрагментации

$$P_{2i} = 1 - (1 - D_2)^{k_i},$$

где

$$D_2 = \left[1 - \frac{\Lambda v K}{n} \exp\left(-\left(T_0 - v\right)\left(\frac{1}{v} - \frac{\Lambda K}{n}\right)\right) \right] \exp(-\lambda T_0);$$

б) с фрагментацией запросов без их резервирования

$$P_{3i} = \left[\left\{ 1 - \frac{\Lambda v}{n} \exp\left(-\left(T_0 - \frac{v}{s_i}\right)\left(\frac{S}{v} - \frac{S\Lambda}{n}\right)\right) \right\} \exp(-\lambda T_0) \right]^{s_i};$$

в) при резервировании и фрагментации запросов

$$P_{23i} = \left[1 - (1 - D_{23})^{k_i s_i} \right],$$

где вероятность своевременного и безошибочного выполнения фрагмента запроса в одном узле кластера

$$D_{23} = \left\{ 1 - \frac{K \Lambda v}{n} \exp\left(-\left(T_0 - \frac{v}{s}\right)\left(\frac{S}{v} - \frac{K S \Lambda}{n}\right)\right) \right\} \exp(-\lambda T_0).$$

Примеры расчета вероятности своевременного и безошибочного обслуживания запросов

Расчеты проведем для случая, когда все запросы могут быть разделены на s равных частей, выполняемых в разных узлах.

Зависимость вероятности своевременного обслуживания при фрагментации запроса (без резервирования) от интенсивности потока запросов Λ представлена на рис. 1. Запрос считается успешно выполненным в кластере при условии, что время ожидания всех s фрагментов не превосходит директивный срок ожидания $t_0 = T_0 - (v/s)$. Расчет проведен при $v = 1$ с и числе узлов кластера $n = 5$.

Кривые 1—4 соответствуют разбиению запроса на $s = 1, 2, 3, 4$ частей при директивном сроке пребывания запросов $T_0 = 1,5$ с, а кривые 5—8 — разбиению на $s = 1, 2, 3, 4$ при директивном сроке $T_0 = 3$ с.

Из представленных графиков видно существование границы интенсивности входного потока, ниже которой разбиение запроса на части с их параллельным обслуживанием в разных узлах эффективно (приводит к увеличению вероятности обслуживания запроса в директивные сроки), а выше — нет. Причем при увеличении числа фрагментов, на

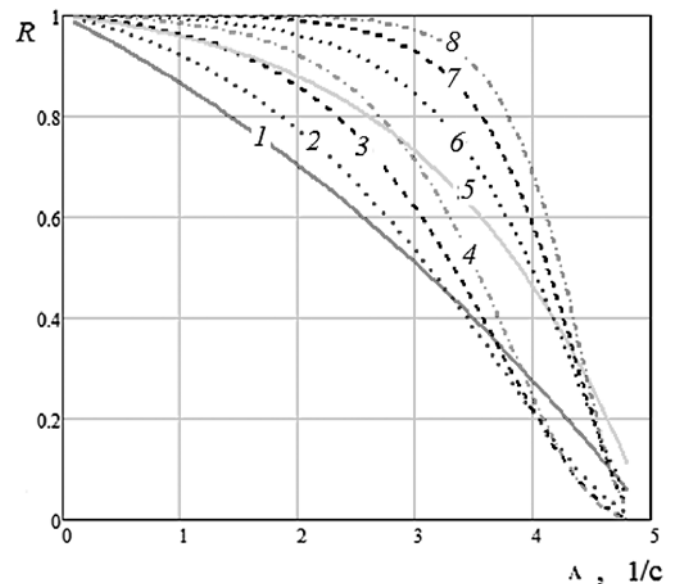


Рис. 1. Зависимость вероятности своевременного обслуживания при фрагментации запросов без их резервирования

которые разбивается запрос, вероятность своевременного обслуживания запросов при нагрузке ниже граничной увеличивается. Приведенные расчеты подтверждают, что чем менее жесткие директивные сроки выполнения запросов, тем больше вероятность их своевременного обслуживания.

Зависимость вероятности своевременного обслуживания запросов от интенсивности их входного потока Λ при резервировании копий фрагментированных запросов представлена на рис. 2. Запрос считается успешно выполненным в кластере, если время ожидания хотя бы одной из k резервных копий, выполняемых в разных узлах кластера, всех s фрагментов не превосходит директивный срок ожидания $t_0 = T_0 - (v/s)$. Расчет проведен при $v = 1$ с, директивном сроке обслуживания $T_0 = 1,5$ с и числе узлов кластера $n = 15$.

Кривые 1–3 соответствуют разбиению запроса на $s = 1, 2, 3$ частей без резервирования их выполнения, кривые 4–6 — разбиению запроса на $s = 1, 2, 3$ частей при дублировании их выполнения в разных узлах ($k = 2$), а кривые 7–9 — разбиению запроса на $s = 1, 2, 3$ частей при кратности резервирования вычислений $k = 3$.

Результаты расчетов показывают возможность увеличения вероятности своевременного выполнения запросов при разбиении запросов на части (фрагменты) с резервированным параллельным выполнением фрагментов запроса в разных узлах кластера. При этом при низкой нагрузке эффективность обслуживания запросов по вероятности своевременного обслуживания повышается при увеличении числа фрагментов и кратности резервирования их выполнения в разных узлах кластера. По мере увеличения интенсивности входного потока запросов (загрузки узлов) для увеличения вероятности своевременного безошибочного выполнения запросов число частей (фрагментация) разбиения запроса для параллельного обслуживания и кратность резервирования при их выполнении

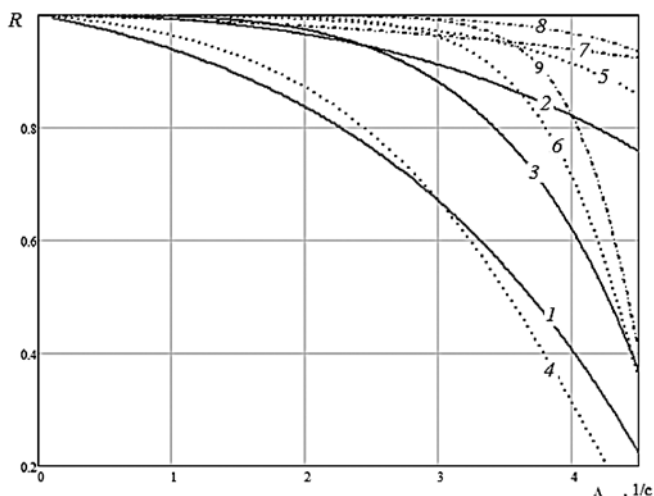


Рис. 2. Зависимость вероятности своевременного обслуживания при резервировании и фрагментации запросов

должна уменьшаться. При этом выбор оптимального числа частей разбиения запроса и кратность их резервирования может определяться на основе рассмотренных моделей.

Предложенный подход позволяет реализовать адаптивное обслуживание запросов с выбором числа выделяемых фрагментов и кратности их резервирования в зависимости от интенсивности входного потока и загруженности узлов кластера.

Для решения целесообразности применения предлагаемых дисциплин обслуживания в конкретных прикладных компьютерных системах необходим дополнительный комплекс работ, связанных с учетом конкретных особенностей реализуемых систем и прикладных процессов реального времени.

Достоверность принятых решений о целесообразности рассматриваемых дисциплин обслуживания может быть повышена при построении моделей с учетом:

- взаимозависимости запросов и возможностей их распараллеливания (в том числе на неравные части, распределение случайной величины времени, которое может отличаться от экспоненциального);
- влияния на эффективность фрагментации процессов распараллеливания, диспетчеризации и распределения фрагментированных запросов между узлами кластера;
- синхронизации вычисления при возможной взаимной зависимости результатов выполнения резервированных фрагментов.

Решение о целесообразности применения предлагаемых дисциплин обслуживания в конкретных условиях реализации кластеров, помимо использования предлагаемых аналитических моделей, может потребовать их уточнение, использование имитационного моделирования или натуральных испытаний.

Заключение

Для кластерных систем реального времени, консолидирующих ресурсы серверов, с организацией локальных очередей в каждом из них предложена оценка вероятности того, что время ожидания запросов при их фрагментации и резервированном параллельном выполнении фрагментов в разных узлах кластера меньше предельно допустимой величины t_0 .

Проанализировано влияние фрагментации и кратности резервирования вычислений на вероятность своевременного выполнения запросов в условиях возможных отказов и ошибок вычислений.

Показана возможность увеличения вероятности своевременного выполнения запросов при разбиении запросов на части (фрагменты) с резервированным параллельным выполнением фрагментов запроса в разных узлах кластера. При этом при низкой нагрузке вероятность своевременного обслуживания повышается при увеличении числа фрагментов и

кратности резервирования их выполнения в разных узлах кластера.

Показано существование границы интенсивности входного потока, ниже которой разбиение запроса на части с их параллельным выполнением разными узлами приводит к увеличению вероятности обслуживания запроса в директивные сроки, причем эффективность фрагментации возрастает при увеличении кратности резервирования при выполнении фрагментов.

Предложенные модели могут быть использованы при оценке функциональной надежности и выборе организации распараллеливания и резервирования вычислительного процесса реального времени для вычислительных и управляющих систем кластерной архитектуры.

Список литературы

1. **Kopetz H.** Real-Time Systems: Design Principles for Distributed Embedded Applications. Springer. 2011. 396 p.
2. **Черкесов Г. Н.** Живучесть и отказобезопасность ответственных технических систем // Проектирование и технология электронных средств. 2015. № 1. С. 15–24.
3. **Шубинский И. Б.** Функциональная надежность информационных систем: методы анализа // Надежность, 2012. 296 с.
4. **Gurov S. V., Utkin L. V.** An inverse problem of the load-sharing system reliability analysis: Constructing the load function // Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability. 2015. P. 1–10.
5. **Aleksanin S. A., Zharinov I. O., Korobeynikov A. G., Perezyabov O. A., Zharinov O. O.** Evaluation of chromaticity coordinate shifts for visually perceived image in terms of exposure to external illuminance // ARPN Journal of Engineering and Applied Sciences. 2015. T. 10. № 17.
6. **Aliev T. I., Rebezova M. I., Russ A. A.** Statistical Methods for Monitoring Travel Agencies // Automatic Control and Computer Sciences. 2015. Vol. 49, N. 6. P. 321–327.
7. **Богатырев В. А., Богатырев А. В., Богатырев С. В.** Оценка надежности выполнения кластерами запросов реального времени // Изв. вузов. Приборостроение. 2014. Т. 57. № 4. С. 46–48.
8. **Богатырев В. А., Богатырев А. В., Богатырев С. В.** Перераспределение запросов между вычислительными кластерами при их деградации // Изв. вузов. Приборостроение. 2014. Т. 57, № 9. С. 54–58.
9. **Богатырев В. А.** Оценка надежности и оптимальное резервирование кластерных компьютерных систем // Приборы и системы. Управление, контроль, диагностика. 2006. № 10. С. 18–21.
10. **Богатырев В. А., Богатырев С. В., Богатырев А. В.** Оптимизация кластера с ограниченной доступностью кластерных групп // Научно-технический вестник ИТМО. 2011. № 1 (71). С. 63–67.
11. **Дудин А. Н., Сунь Б.** Многолинейная ненадежная система с управляемым ширококвотельным обслуживанием // Автоматика и телемеханика. 2009. Т. 70. № 12. С. 147–160.
12. **Dudin A. N., Sun B.** A multiserver MAP/PH/N system with controlled broadcasting by unreliable servers // Automatic Control and Computer Sciences. 2009. N. 5. P. 32–44.
13. **Bogatyrev V. A., Bogatyrev A. V.** Functional Reliability of a Real-Time Redundant Computational Process in Cluster Architecture Systems // Automatic Control and Computer Sciences. 2015. Vol. 49, N. 1. P. 46–56.
14. **Богатырев В. А., Богатырев А. В.** Оптимизация резервированного распределения запросов в кластерных системах реального времени // Информационные технологии. 2015. Т. 21, № 7. С. 495–502.
15. **Вишневецкий В. М.** Теоретические основы проектирования компьютерных сетей М.: ТЕХНОСФЕРА, 2003. 512 с.

V. A. Bogatyrev, Professor, e-mail: Vladimir.bogatyrev@gmail.com,

A. V. Bogatyrev, Post Graduate

Saint Petersburg National Research University of Information, Technologies, Mechanics and Optics

The Reliability of the Cluster Real-Time Systems with Fragmentation and Redundant Service Requests

For clustered systems, real-time, consolidating server resources, with the organization of local queues in each of them proposed the assessment of the probability that the waiting time of requests in their fragmentation and redundant parallel execution fragments in different nodes of a cluster is lower than the maximum permissible value.

Purpose — to improve the timeliness of service requests in real time the result of fragmentation and the reservation requests.

To achieve the goal of the research disciplines of service at the requirement of delivery of results within the deadlines, providing:

- *splitting (fragmentation) of the query into parts that are executed in different nodes of a cluster (paralleling service the request);*
- *fragmentation of the query with redundant execution copies of fragments in different nodes of a cluster.*

Analyzed the impact of fragmentation and multiplicity of redundancy calculations on the probability of timely fulfillment of requests in terms of possible failures and errors of calculations.

The possibility of increasing the probability of timely fulfillment of requests when splitting a query into pieces (fragments) with redundant parallel execution of query fragments to different nodes of a cluster. In this case, when the low probability of timely maintenance increases with the increase of number of fragments and the multiplicity of redundancy they will run on different nodes of a cluster.

The existence of the border intensity input stream, below which the query is broken into pieces with their concurrent execution of different nodes leads to an increase in the probability of servicing the request within the deadlines, and the fragmentation efficiency increases with increasing multiplicity of redundancy when performing fragments.

The proposed model can be used to assess the functional reliability and the choice of the paralleling and redundancy of the computational process for real-time computing and control systems of cluster architecture.

Keywords: reliability, redundant computation, real-time, cluster, requests, Queuing system, Queuing, fragmentation

References

1. **Kopetz H.** *Real-Time Systems: Design Principles for Distributed Embedded Applications*. Springer, 2011, 396 p.
2. **Cherkesov G. N.** Zhivuchest' i otkazobezopasnost' otvetstvennykh tehnikeskikh sistem, *Proektirovanie i tehnologija jelektronnykh sredstv*, 2015, no. 1, pp. 15–24.
3. **Shubinskij I. B.** Funkcional'naja nadezhnost' informacionnykh sistem: metody analiza. *Nadezhnost'*, 2012. 296 p.
4. **Gurov S. V., Utkin L. V.** An inverse problem of the load-sharing system reliability analysis: Constructing the load function, *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 2015, pp. 1–10.
5. **Aleksanin S. A., Zharinov I. P., Korobeynikov A. G., Perezyabov O. A., Zharinov O. O.** Evaluation of chromaticity coordinate shifts for visually perceived image in terms of exposure to external illuminance, *ARPN Journal of Engineering and Applied Sciences*, 2015, vol. 10, no. 7.
6. **Aliev T. I., Rebezova M. I., Russ A. A.** Statistical Methods for Monitoring Travel Agencies, *Automatic Control and Computer Sciences*, 2015, vol. 49, no. 4, pp. 321–327.
7. **Bogatyrev V. A., Bogatyrev A. V., Bogatyrev S. V.** Ocenka nadezhnosti vypolnenija klasterami zaprosov real'nogo vremeni, *Izv. vuzov. Priborostroenie*, 2014, vol. 57, no. 4, pp. 46–48.
8. **Bogatyrev V. A., Bogatyrev A. V., Bogatyrev S. V.** Pereraspredelenie zaprosov mezhdru vychislitel'nymi klasterami pri ih degradacii, *Izv. vuzov. Priborostroenie*, 2014, vol. 57, no. 9, pp. 54–58.
9. **Bogatyrev V. A.** Ocenka nadezhnosti i optimal'noe rezervirovanie klasternykh komp'yuternykh sistem, *Pribory i sistemy. Upravlenie, kontrol', diagnostika*, 2006, no. 10, pp. 18–21.
10. **Bogatyrev V. A., Bogatyrev S. V., Bogatyrev A. V.** Optimizacija klastera s ogranichennoj dostupnost'ju klasternykh grupp, *Nauchno-tehnicheskij vestnik ITMO*, 2011, no. 1 (71), pp. 63–67.
11. **Dudin A. N., Sun' B.** Mnogolinejnaja nenadezhnaja sistema s upravljajemym shirokoveshhatel'nym obsluzhivaniem, *Avtomatika i telemekhanika*, 2009, vol. 70, no. 12, pp. 147–160.
12. **Dudin A. N., Sun' B.** A multiserver MAP/PH/N system with controlled broadcasting by unreliable servers, *Automatic Control and Computer-Sciences*, 2009, no. 5, pp. 32–44.
13. **Bogatyrev V. A., Bogatyrev A. V.** Functional Reliability of a Real-Time Redundant Computational Process in Cluster Architecture Systems, *Automatic Control and Computer Sciences*, 2015, vol. 49, no. 1, pp. 46–56.
14. **Bogatyrev V. A., Bogatyrev A. V.** Optimizacija rezervirovanogo raspredelenija zaprosov v klasternykh sistemah real'nogo vremeni, *Informacionnye tehnologii*, 2015, vol. 21, no. 7, pp. 495–502.
15. **Vishnevskij V. M.** *Teoreticheskie osnovy proektirovanija komp'yuternykh setej*. Moscow: TEHNOSFERA, 2003. 512 p.

УДК 004.09

М. Ю. Неустроев, аспирант, e-mail: m.neustroev@gmail.com

Государственное бюджетное образовательное учреждение высшего образования Московской области "Технологический университет", г. Королев, Московская область

Анализ показателей эффективности и скорости обслуживания в центрах обработки вызовов

Проводится анализ уже существующих моделей центров обслуживания вызовов с применением IVR-систем, показана необходимость дальнейших исследований и разработки алгоритмов интеллектуального управления вызовами IVR-системой без участия операторов для снижения загрузки операторов.

Ключевые слова: Центр обслуживания вызовов (ЦОВ), интерактивное голосовое меню (IVR), эффективность ЦОВ, ASA, время ожидания, время обслуживания, загрузка операторов, голосовые сообщения

Введение

Чтобы выжить в современной конкурентной среде, компаниям необходимо найти способы повышения качества обслуживания клиентов. Современный потребитель услуг знает, чего он хочет. Теперь если предприятие хочет повысить уровень оказываемых услуг и пойти дальше, чем простое выживание, а также добиться стабильного успеха, оно должно модернизировать и внедрять новые сервисы повышения услуг обслуживания своих клиентов. Интеграция Центров обслуживания вызовов (ЦОВ) в текущую информационную структуру помогает повысить производительность и эффективность обслуживания клиентов [1, 2]. Создание собственного ЦОВ — серьезный шаг, требующий значительных затрат. И если для крупной компании такой шаг может быть оправдан, для среднего и особенно малого бизнеса попытка самостоятель-

но обслуживать входящие и совершать исходящие звонки может стать дорогой ошибкой с потерей не только денег, но и клиентов.

Большинство современных ЦОВ используют одну или более технологий обработки вызовов, таких как автоматическое распределение звонков (ACD), интерактивное голосовое меню (IVR), распознавание речи (VRU). Для качественного и гарантированного обслуживания поступающих звонков каждый из модулей должен быть изучен и правильно запрограммирован. Новые типовые системы распределения вызовов значительно отличаются от систем предыдущего поколения своими алгоритмами и процессами [3].

Показатели эффективности обслуживания в ЦОВ

Для прогнозирования эффективности работы, выполняемой ЦОВ, на передний план выдвигаются

подсистемы, с помощью которых организована и построена архитектура сбора, хранения и первичной обработки информации, с последующим предоставлением отчетности о комплексном функционировании оборудования и составлением будущего прогноза поведения технических систем и служб в различных ситуациях в целом. Выполнение данных задач не может повлиять на качество и уровень услуг по предоставлению информации клиенту, но может натолкнуть управляющий состав call-центра на выявление механизмов и средств по оптимизации показателя эффективности или регулирования потоков данных и позволяет систематизировать работу так, чтобы обеспечить выполнение большинства запросов клиентов [4].

В международной практике основными показателями эффективности и скорости обслуживания входящих вызовов являются:

- средняя скорость ответа (Average Speed of Answer — ASA);
- процент звонков, прерванных абонентами во время ожидания (Abandon Rate — AR);
- уровень обслуживания (Service Level — SL);
- автоматическое распределение вызовов (Automatic Call Distribution — ACD);
- средняя продолжительность разговора (Average Talk Time — ATT);
- интегрирование компьютерной телефонии (Computer Telephony Integration — CTI).

В режиме реального времени по каждому из этих показателей могут быть составлены подробные отчеты в хронологическом порядке, в виде графиков или полноценных структурированных таблиц на каждый момент времени (на каждое событие). Но конкретный выбор контролируемых параметров, видов статистики, поставляющих информацию об этих параметрах, зависит от организации бизнес-процессов в конкретном ЦОВ. Высший уровень качества обслуживания большинство ЦОВ достигает, используя ключевые показатели эффективности KPI (Key Performance Indicators), которые позволяют управлять обслуживанием более эффективно, так как выполнение каждого из ключевых показателей напрямую влияет на достижение общей цели.

В мировой практике KPI является неизменным элементом не только для оценки сложных технологических процессов, но и системы управления в целом. KPI — оценка качественных и количественных показателей, итоговых результатов выполнения бизнес-процессов и сравнения их с тактическими, стратегическими и операционными целевыми ориентирами для определения разницы между фактическим и целевым показателем. KPI-система дает возмож-

ность прозрачно для каждого сотрудника или подразделения оценить эффективность работы в целом. Это позволяет оценить эффективность деятельности всех структур в режиме реального времени и спрогнозировать то, как будут решаться задачи и достигаться цели.

Система IVR

При поступлении входящего вызова на любую из очередей в ЦОВ он автоматически попадает в IVR-систему, которая представляет собой программное обеспечение, предназначенное для автоматизированной обработки вызовов, поступающих в ЦОВ. Среди всей массы вызовов, поступающих в ЦОВ, можно выделить несколько типов, обслужить которые возможно и без участия операторов. К таким типам относятся:

- получение стандартной справочной информации (время работы компании, адреса ее офисов, описание предоставляемых услуг, цены и т. д.);
- запрос абонентом своего текущего баланса;
- заказ стандартных услуг;
- получение копий документов и др.

Данные вызовы может обрабатывать IVR-система. Кроме того, внедрение IVR-системы обеспечивает наибольшую продуктивность при создании сложных составных голосовых меню, авторизации абонентов для использования платного контента или услуг, создания автоматизированного оператора, предназначенного для маршрутизации звонков на основе добавочного номера абонента и т. д. Правильная реализация интерактивного меню позволяет в значительном объеме снизить нагрузку на очереди и операторов, высвобождая время для обработки более сложных запросов. На рис. 1 приве-

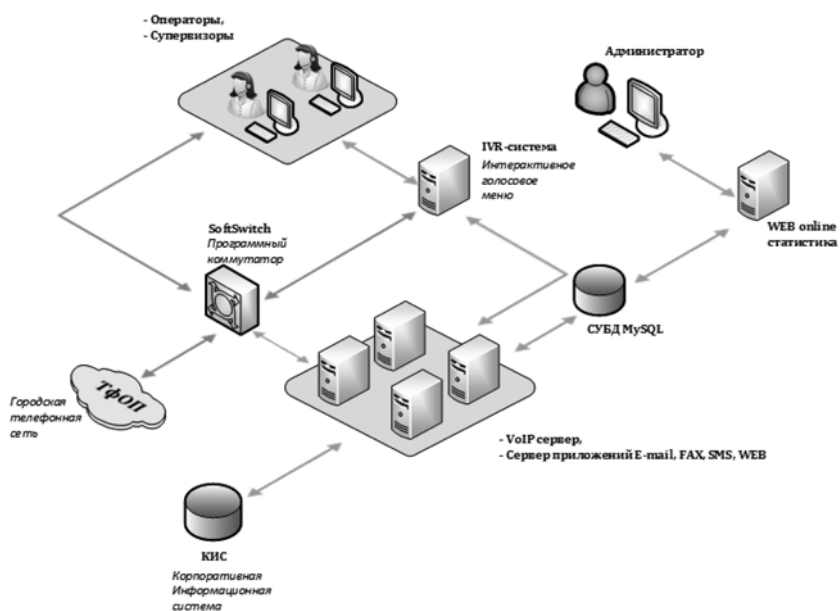


Рис. 1. Схема современного центра обслуживания вызовов

дена структура современного центра обслуживания вызовов.

Расчет эффективности обслуженных звонков IVR-системой в ЦОВ без учета соединения с оператором рассчитывается по формуле

$$P_{IVR} = \frac{N_{IVR}}{N_R} \cdot 100 \%, \quad (1)$$

где N_{IVR} — число вызовов, которые были обслужены IVR-системой; N_R — число вызовов, поступивших в ЦОВ.

ЦОВ позволяют организовать индивидуальное обслуживание для каждого клиента или заказчика. Повышение качества такого обслуживания осуществляется с использованием интегрированной компьютерной телефонии (СТІ). Необходимость многократного выяснения у клиента или заказчика некоторых общих сведений и сути вопроса полностью отпадает с использованием системы СТІ, так как эти сведения накапливаются в буфер и каждый раз посылаются вместе с вызовом. Например, если позвонить в контактный центр и получить приглашение от автоинформатора, заказчику потребуются ввести номер своего лицевого счета лишь один раз. Ему не придется вводить номер повторно, поскольку при дальнейших переадресациях его вызова (если такие будут необходимы) номер будет передаваться вместе с вызовом. Использование СТІ во всех видах коммуникаций (в сетях, на АСD, на IVR, на Web-серверах и рабочих местах операторов) позволяет получать, обрабатывать и пересылать все сведения о клиенте вместе с его идентификатором вызова [5].

Основным аспектом СТІ является предоставление оператору сведений о клиенте при входящем звонке. Обычно данный функционал организован в виде всплывающей подсказки. Для получения данных СТІ получает некий определитель номера (ANI — Automatic Number Identification) и идентификатор набранного номера (DNIS — Dialed Number Identification String) с городской телефонной сети или от IVR-системы и далее отправляет данные в приложение рабочего места оператора. Поиск информации осуществляется в текущей базе данных клиентов с последующим выводом в виде всплывающей подсказки. Обработка данных и поиск информации, как правило, осуществляются с такой скоростью, что всплывающая подсказка появляется у оператора одновременно с поступившим вызовом. Таким образом, оператор обслужит клиента быстрее, поскольку до начала разговора он уже имеет информацию по данному клиенту, а также всю хронологическую информацию по нему.

Средняя скорость ответа ASA

Показатель ASA демонстрирует уровень производительности при обращении в контактный центр,

а также его способность обработать существующую нагрузку. Как только входящий вызов попадает в систему АСD, начинается расчет средней скорости ответа. Данный показатель измеряется как отношение общего времени ожидания в очереди к общему числу обслуженных вызовов операторами:

$$ASA = \frac{T_q}{N_R}, \quad (2)$$

где T_q — общее время ожидания в той или иной очереди; N_R — общее число обслуженных вызовов.

Рекламные акции организации, запуск новых продуктов, проведение акций и различных мероприятий являются фактором повышенной нагрузки на контактный центр. Интерес компании — сделать так, чтобы каждый позвонивший получил информацию или услугу и при этом не повесил трубку. В качестве одного из КРІ контактного центра должен присутствовать механизм управления производительностью системы ASA, которая зависит от ряда факторов. Сюда можно отнести "мощность" контактного центра и уровень нагрузки. Недостатком показателя ASA является то, что при ответе на большинство звонков в очень короткие сроки по отношению к общему числу звонков показатель ASA будет искажаться. Например, при принятых 95 вызовов в течение 3 с и 5 вызовов в течение 55 с средняя скорость ответа будет 11,03 с. В данном случае показатель производительности значительно изменяется по отношению к общему числу вызовов и нарушает отчетность.

Можно ли улучшить показатель ASA? Работа над оптимизацией значения требует комплексного технологического решения, которое ориентировано на постоянное повышение компетенции персонала и продуктивность процессов обслуживания клиентов. К таким процессам можно отнести правильно построенную систему IVR в целях снижения нагрузки на очередь и получения необходимой информации из автоответчика или интегрированного в систему IVR. Чем больше сервисов, которые позволяют выполнять запросы клиентов, тем меньше клиентов, которые стремятся переключиться на оператора, а следовательно, ожидать в очереди. Оперирование данным сервисом имеет критическое значение в самообслуживании, так как неудобный IVR объективно создает спрос на "живое" общение.

Опыт в области создания систем голосового самообслуживания на базе технологий синтеза и распознавания речи, а также записи телефонных переговоров и их комплексного речевого анализа позволяют решить следующие задачи:

- автоматизировать наиболее типовые запросы клиентов с помощью системы голосового самообслуживания;

- внедрить процесс мониторинга эргономики голосовых приложений с помощью речевой аналитики и анализа эмоционального состояния;
- внедрить систему мониторинга компетенций операторов.

Среднее время обслуживания вызова АНТ

Повышение скорости обработки обращения контролируется с помощью показателя АНТ (Average Handle Time — АНТ), формирующего общие представления о скорости обработки обращения в контактном центре. Это время складывается из времени общения с клиентом, удержания вызова и времени на постобработку вызова (ввод данных в карточку CRM (Customer Relationship Management — CRM), заполнение форм и заявок, формирование отчетов и т. д.). То есть значение показателя складывается из двух компонентов — взаимодействие оператора с клиентом и с прикладными информационными системами на рабочем месте. Значение показателя АНТ позволяет оценить степень достаточности выделяемого персонала на каждый конкретный проект или очередь, т. е. АНТ — удобный инструмент прогнозирования нагрузки и распределения трудовых ресурсов. Вместе с тем, динамика АНТ отражает компетенцию контактного центра в целом, эффективность имеющихся процессов обслуживания, ИТ-инфраструктуры контактного центра, процессов развития и обучения персонала. В связи с этим сокращение АНТ (в рамках одних и тех же проектов) демонстрирует общий рост производительности контактного центра. Снижение среднего времени диалога (Average Talk Time — АТТ) оказывает существенное влияние на ключевые показатели эффективности работы контактного центра: уровень обслуживания (Service Level — SL) и долю потерянных вызовов (Lost Call Rate — LCR). Для эффективной работы по оптимизации данных показателей необходимо автоматически контролировать количественно-временные параметры разговора операторов: АНТ, время до снятия трубки, число и длительность пауз в разговоре (hold, mute).

Распознавание голосовых обращений ASR

Технология автоматического распознавания речи (Automated Speech Recognition — ASR) позволяет преобразовывать различные произнесенные слова в язык машинных кодов. Процесс распознавания включает в себя предобработку и декодирование речи.

На этапе предварительной обработки осуществляются запись и оцифровка речи с дальнейшим акустическим анализом. На стадии декодирования применяются лексическая, акустико-фонетическая и языковая модели, которые позволяют получить граф, содержащий множество вероятных слов в последовательности.

Существуют два варианта использования технологии ASR. Первый вариант "закрытая грамматика" предполагает, что клиент произносит определенные ключевые слова или фразы, которые заранее определены, например, "справка", "личный счет", "баланс". В варианте "открытая грамматика" клиент произносит фразу в произвольной форме, в которой распознаются ключевые слова в предложении, к примеру: "Мне необходима выписка лицевого счета".

ASR в голосовом приложении применяется в двух режимах:

- стандартный режим — клиент произносит ключевую фразу после окончания голосовой подсказки или информационного блока меню;
- режим с прерыванием (Barge-In) — клиент произносит ключевую фразу в любой момент прослушивания голосовой подсказки меню IVR, не дожидаясь ее окончания.

Важнейшая особенность и преимущество технологии ASR — это простота в использовании и управлении. Клиенту, обратившемуся в ЦОВ, гораздо проще словесно произнести фразу или слово, чем использовать нужные пункты IVR и переключать клавишами телефоны между ними. Тоновый набор команд для переключения между сервисами центра гораздо более затруднителен, нежели использование автоматического режима использования. В данном случае имеются в виду случаи использования необходимой информации, которую следует заложить в IVR, использовав всего 10 клавиш телефонного аппарата.

Множество российских компаний еще не знают, как можно применять данную технологию для получения выгоды от ее использования. Существующие IVR-системы, по возможности, заменяют ASR-технологией, перенося существующие пункты меню на голосовой набор. Данный способ не рассчитан на речевое взаимодействие системы и человека, в результате чего клиенты не могут в полном объеме использовать новые автоматизированные сервисы, тем самым увеличивая нагрузку на IVR-систему. Организации от данного внедрения только увеличивают расходы на обслуживание системы, не извлекая при этом выгоды с возможной экономией затрат.

На западе опыт применения ASR-технологии показывает реальные изменения принципа общения клиентов с голосовыми системами. Для достижения видимого результата необходим глубокий анализ психологического поведения клиентов в целях оптимизации IVR-системы с подстройкой под голосовые сервисы. С учетом данных факторов и анализов поведения клиентов число обработанных вызовов автоматическими системами может быть увеличен на 40 %.

Еще одним важным аспектом, который требует особого внимания при внедрении ASR, — локальное тестирование с последующим переносом на гло-

бальные группы операторов. Результаты в период тестового испытания подвергаются корректировке с учетом желаний клиентов. ASR-технология должна обеспечивать удобство использования клиентом с получением наибольшего эффекта автоматизации.

Синтез речи (TTS)

Синтез речи TTS (Text-to-Speech — воспроизведение текста голосом) — одно из самых популярных и востребованных решений. Яркий пример — база знаний в ЦОВ. TTS-ресурсы в центрах обработки вызовов используются около семи лет. Они легко интегрируются в состав многоканальных IVR-систем и системы голосового обслуживания с использованием стандартных протоколов обмена данными и языков разметки. Системы голосового самообслуживания значительно расширяют возможности по созданию интеллектуальных центров, которые дают возможность самостоятельно, без участия оператора, в интерактивном режиме получить информацию или заказать услуги. На большинство запросов к информационно-справочной системе в той или иной форме способен отвечать компьютер. Важно подчеркнуть, что с особенной полнотой достоинства ЦОВ, использующих синтез речи, проявляются в тех ситуациях, когда информация обновляется периодически (еженедельно, ежедневно, еже часно). Такие услуги могут быть интересны потребителям только в том случае, если информация актуальна. Вот некоторые преимущества внедрения TTS в ЦОВ:

- снижение затрат;
- сокращение времени обслуживания вызова;
- обеспечение непрерывного цикла обслуживания 24 ч в сутки;
- возможность интеграции с различными информационными системами;
- снижение влияния человеческого фактора;
- снижение затрат на персонал и его обучение.

Плюсы данного внедрения в том, что клиент слышит человеческую речь, а недостаток — динамично меняющаяся информация. Технология синтеза речи использует сложные лингвистические алгоритмы обработки выходного сигнала из текста, что позволяет достичь максимально приближенного к реальности звучания. Сейчас технология развивается быстрыми темпами. TTS способна воспринимать сложные фразеологизмы и воспроизводить их в единой смысловой конструкции с особой интонацией. Особое внимание уделяется паузам и интонации синтетической речи, ведь именно эти параметры во многом обеспечивают естественность голоса. Лингвистическая обработка позволяет в автоматическом режиме расшифровывать множества используемых сокращений.

Сравнительный анализ IVR-систем

Сравнительный анализ показателей эффективности показал, что использование IVR без использования технологий ASR и TTS (классический) наиболее эффективно. Это обусловлено тем, что внедрение новых технологий с совершенствованными IVR-системами в контактном центре показывает принципиально новый уровень эффективности, который едва ли можно достигнуть на основе нового набора (DTMF). На рис. 2 отражены усредненные показатели эффективности при анализе методов классического IVR и модифицированной системы обработки вызовов с помощью ASR и TTS.

В среднем классический IVR способен обрабатывать 10...20 % всех поступающих вызовов, не зависящих от числа очередей. Такой низкий уровень автоматизации обусловлен наличием только маршрутизации вызовов по различным направлениям или статическим фразам, представляющим стандартную справочную информацию. Остальная часть запросов не дает полных ответов клиентам и, как правило, распределяется между операторами и потерянными вызовами. Внедрение систем распознавания и синтеза речи по отдельности или в комплексе позволяет автоматизировать до 80 % всех поступающих вызовов клиентов. Таким образом, нагрузка на операторов уменьшается и лишь часть голосовых запросов достигает обработки 20 %. Данные механизмы предоставляют максимально раскрытую динамическую информацию и позволяют управлять услугами в автоматическом режиме, без привлечения операторов. Конечно же, основным преимуществом является наименьшее число операторов, но при этом требуется наибольшая компетентность операторов для обслуживания запросов, которое не может быть автоматизировано [6].

Использование классической модели архитектуры IVR в сравнении с моделью голосового управления наделена очевидными ограничениями:

- необходимость прослушивания всех пунктов меню IVR;
- ограниченность пунктов меню IVR наличием кнопок на телефонном аппарате;

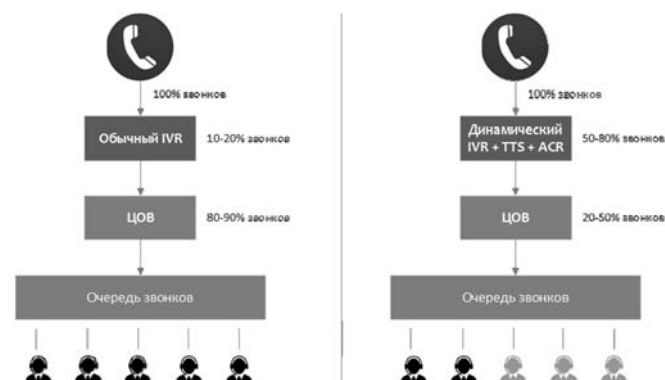


Рис. 2. Усредненные показатели эффективности

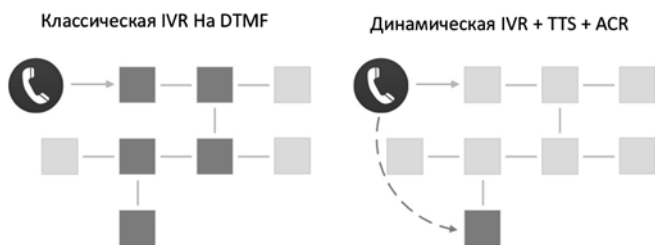


Рис. 3. Сравнение динамической и статической IVR-систем

- сложная многоуровневая архитектура голосового меню; дополнительные вызовы в ЦОВ за счет ошибочности выбора нужного пункта меню;
- ограниченность и доступность использования DTMF-набора.

Опыт внедрения и исследование различных комплексов IVR показывают, что до каждого следующего пункта меню не доходит в среднем 10 % звонящих (рис. 3). Это означает, что при использовании пяти уровней меню информация получают только половина звонящих.

Значительное число IVR-систем неверно функционируют из-за нарушений логики, которые остаются незамеченными для ЦОВ. Периодичность появления нарушений довольно ощутима, так как граница между вариантами выбора нечеткая, и звонящий не уверен в точности выбора необходимого направления пунктов меню. Динамическая IVR-система становится более гибкой и комплексной при использовании распознавания речи и также позволяет автоматизировать процессы с выполнением большего числа операций. Анализ внедрения комплекса, состоящего из ACR и TTS, показывает следующие преимущества:

- упрощенный поиск и способ получения информации;
- уменьшенное среднее время звонка;
- автоматизация большего числа вызовов;
- возможность получения абонентами динамической информации;
- лояльность клиентов перед использованием голосовых сервисов.

Динамическая система рассчитана на извлечение требуемой информации из базы данных на основе правил маршрутизации, которые разработаны в компании (состояние лицевого счета, акции, справки). Построение такой системы — достаточно сложный процесс несмотря на преимущественно программную реализацию решения. Модели сильно различаются не только функциональными возможностями, но и ценовой характеристикой.

Контактные центры, не использующие IVR-системы, крайне не эффективны с точки зрения финансирования, так как заработная плата операторов — самый дорогой ресурс в обслуживании данного ЦОВ. И никакая система не сможет обеспечить "живое" общение в целях решения персонализированной проблемы клиентов, так как иногда клиенты

не в силах ее сформулировать. Применение автоматизированных комплексов и интегрирование модели IVR с использованием TTS, ACR и других механизмов существенно снизят операционные затраты в обслуживании и значительно упростят организационные бизнес-процессы в работе ЦОВ.

Заключение

Правильный подход при внедрении технологий в ЦОВ и правильное распределение ресурсов приносят ощутимый результат. Внедрение ЦОВ — очень длительный процесс, от которого эффект не сразу виден, поэтому внедрять ЦОВ необходимо поэтапно, расширяя его функциональность. Если же разработка проекта и внедрение затруднены, а прогноз использования показывает малую эффективность, следует воспользоваться услугами аутсорсинга. Проблема при внедрении ЦОВ сводится к неумению правильно поставить задачу, определить функции и задачи центра, что является существенной помехой при внедрении и разработке технического задания. Также остаются многие нерешенные проблемы и задачи при использовании нескольких компаний-производителей в построении ЦОВ. В результате руководитель ЦОВ испытывает массу проблем в обслуживании систем и интеграции их между собой, кроме того затрудняется функционирование и заметно усложняется управление центром.

Комплексный подход при построении и внедрении ЦОВ объединяет в себе такие основные компоненты, как создание логической схемы обслуживания и обработки данных по клиентам, объединение всех компонентов разных производителей, взаимодействие с корпоративной информационной системой (КИС), интеграции с возможными платежными системами.

Современные системы КИС, CRM, ERP (Enterprise Resource Planning — планирование ресурсов предприятия) и ЦОВ используют стандартные протоколы и открытые интерфейсы, поэтому, как правило, технических проблем и вопросов не возникает.

По моему мнению, разработка и внедрение ЦОВ без интеграции с текущими информационными системами, корпоративными системами — неоправданная трата финансов. ЦОВ должен являться неотъемлемой частью информационной системы. Наибольший эффект можно достигнуть при совместном использовании разных источников данных и различных приложений (АСР, CRM, ERP, Service Desk), позволяющих автоматизировать все процессы и ускорить обслуживание клиентов, а также упростить работу операторов ЦОВ. Супервизоры, в свою очередь, при построении комплексных отчетов могут выстроить правильную политику обработки вызовов и изменить текущие схемы взаимодействия операторов и клиентов, распределяя нагрузку и снижая риски потери информации в процессе внутрикорпоративного общения.

Список литературы

1. **Аббасова Т. С.** Принципы и средства оптимизации высокоскоростных информационных каналов // Информационно-технологический вестник. 2014. № 2 (02). С. 10–16.
2. **Аббасова Т. С.** Информационное и программное обеспечение для тестирования производительности сетевых узлов территориально-распределенных центров обработки данных // Электротехнические и информационные комплексы и системы. 2010. Т. 6, № 3. С. 10–13.
3. **Андреев Р. В.** Обособленное обслуживание вызовов для повышения эффективности работы ЦОВ // T-Comm Telecommunications and Transport. 2012. № 1. 32 с.
4. **Артюшенко В. М., Аббасова Т. С., Белюченко В. М.** Системный анализ в области управления и обработки информации. Королев: Изд-во ГБОУ ВО МО "ТУ". 2015. 168 с.
5. **Гольдштейн Б. С.** Контроль показателей качества обслуживания с учетом перехода к сети связи следующего поколения // Техника связи. 2009. № 1.
6. **Привалов В. И., Бокковой Ю. В., Зимин В. М., Харламова Е. С., Четкин В. А., Шевченко Е. П.** Повышение эффективности центров обработки данных для информационных систем // Двойные технологии. 2014. № 4. С. 75–78.

M. Yu. Neustroev, Graduate Student, e-mail: m.neustroev@gmail.com

State Educational Institution of Higher Education Moscow Region "University of technology",
Korolev, Moscow region

Analysis of Performance of Efficiency and Speed of Service in Call Centers

In the article the analysis of already exist of existing models samples using IVR-systems, shows the need for further research and development of algorithms to intelligently manage calls the IVR-system without participation of operators to reduce the loading call centers.

Keywords: Service call center, interactive voice response (IVR), call center efficiency, ASA, waiting time, time for maintenance, loading operators, voice message

References

1. **Abbasova T. S.** Principy i sredstva optimizacii vysokoskorostnyh informacionnyh kanalov, *Informacionno-tehnologicheskij vestnik*, 2014, no. 2 (02), pp. 10–16. (Abbasova T. S. The principles and means of optimizing the high-speed data channels, *Information Technology Bulletin*, 2014, no. 2 (02), pp. 10–16.)
2. **Abbasova T. S.** Informacionnoe i programmnoe obespechenie dlja testirovanija proizvoditel'nosti setevyh uzlov territorijno-raspredeleennyh centrov obrabotki darmyh, *Jelektrotehnicheskie i informacionnye komplekсы i sistemy*, 2010, vol. 6, no. 3, pp. 10–13. (Abbasova T. S. Information and software for testing the performance of network nodes geographically distributed data centers, *Electrical and information systems and systems*, 2010, vol. 6, no. 3, pp. 10–13.)
3. **Andreev R. V.** Obosoblennoe obsluzhivanie vyzovov dlja povyshenija jeffektivnosti raboty COV, *T-Comm — Telecommunications and Transport*, 2012, no. 1, p. 32. (Andreev, R. V. Separate call service to improve the efficiency of call centers, *T-Comm — Telecommunications and Transport*, 2012, no. 1, 32 p.)
4. **Artjushenko V. M., Abbasova T. S., Beljuchenko V. M.** *Sistemnyj analiz v oblasti upravlenija i obrabotki informacii*, Korolev: Izd-vo GBOU VO MO "TU", 2015, 168 p. (Artushenko V. M., Abbasova T. S., Beljuchenko V. M. Systems analysis in the field of management and information processing, *Is king Univ GBOU IN MO "TU"*, 2015, 168 p.)
5. **Gol'dshtejn B. S.** Kontrol' pokazatelej kachestva obsluzhivaniya s uchetom perehoda k seti svjazi sledujushhego pokolenija, *Tehnika svjazi*, 2009, no. 1. (Goldstein, B. C. Monitoring of service quality indicators taking into account the transition to communication networks of the next generation, communication technology, 2009, no. 1.)
6. **Privalov V. I., Bokovoj Ju., Zimin V. M., Harlamova E. S., Четкин В. А., Шевченко Е. П.** Povyshenie jeffektivnosti centrov obrabotki dannyh dlja informacionnyh system, *Dvojnye tehnologii*, 2014, no. 4, pp. 75–78.

Новая книга

Игорь И. Абрамов. Основы моделирования элементов микро- и нанoeлектроники. LAP LAMBERT Academic Publishing, Germany, 2016. 444 с.

В монографии рассмотрены физические основы, проблемы и принципы моделирования приборных структур (элементов) микро- и нанoeлектроники. С единых позиций проведена систематизация моделей элементов. Основное внимание уделено наиболее адекватным численным моделям. Книга предназначена для специалистов в области микро- и нанoeлектроники, преподавателям, аспирантам, магистрантам и студентам соответствующих специальностей. Монография написана на основе статей автора, опубликованных в ведущих научно-технических журналах Российской Федерации в области микро- и нанoeлектроники, и обобщает его сорокалетний опыт работы в данной области науки.

Информацию о книге и ее приобретение можно осуществить, обратившись на сайт:
<https://www.ljubljuknigi.ru/>

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ И ИЗОБРАЖЕНИЙ DIGITAL PROCESSING OF SIGNALS AND IMAGES

УДК: 004.9:621.396.96

А. Б. Филимонов, д-р техн. наук, профессор, filimon_ab@mail.ru,
Фам Фыонг Кыонг, стажер, cuongbkedu@yahoo.com,
МГТУ МИРЭА, г. Москва

Динамическое распознавание воздушных целей по радиолокационным дальностным портретам

Рассматривается проблема динамического распознавания воздушных целей по радиолокационным дальностным портретам. Обсуждается структура информационной базы систем распознавания. Предлагаются динамические схемы распознавания, основанные на агрегировании радиолокационной информации, получаемой в процессе слежения за целью.

Ключевые слова: радиолокационное наблюдение, распознавание воздушных целей, дальностные портреты, информативные признаки, динамические образы, агрегирование данных, схемы динамического распознавания

Задачами радиолокации являются обнаружение и распознавание объектов с помощью радиоволн, а также определение их местоположения и параметров движения в пространстве. Их решение возлагается на особый тип радиотехнических средств — радиолокационные системы (РЛС) [1, 2]. Одно из актуальных направлений в теории и практике автоматического распознавания протяженных воздушных целей (ВЦ) связано с анализом характеристик их радиолокационных дальностных портретов (ДП), называемых в англоязычной литературе "range profile". Современное состояние данной области исследований представляют работы [3—9]. Данной проблематике посвящена работа авторов [10] — в ней рассматривались вопросы конструирования информативных признаков для идентификации ДП.

Радиолокационные цели обладают, как правило, достаточно сложной конфигурацией, и их ДП могут сильно изменяться при изменении угла наблюдения. В этих условиях наибольшую надежность и качество распознавания ВЦ возможно достигнуть лишь на основе методологии распознавания *динамических* образов [11].

Данная статья посвящена вопросам построения динамических схем распознавания движущихся ВЦ, основанных на агрегировании данных анализа ДП, полученных для последовательных моментов времени наблюдения.

Радиолокационные дальностные портреты ВЦ

Объекты наблюдения в радиолокации называются *радиолокационными целями* или просто *целями*.

Дальностный портрет является одномерной характеристикой протяженных радиолокационных целей. Это — отклик объекта наблюдения во временной области, полученный при его облучении широкополосными радиоимпульсами. Обычно это импульсы с внутриимпульсной частотной модуляцией, т. е. их высокочастотное заполнение имеет переменную частоту.

В проведенных исследованиях использовалась программа моделирования эталонных ДП воздушных целей "Radar Target Back Scattering Simulation" (BSS), разработанная под руководством Я. Д. Ширмана [12]. В качестве зондирующего радиосигнала, излучаемого антенной радиолокационной станции, был выбран класс ЛЧМ-сигналов (chirp signal): радиоимпульсы с линейной частотной модуляцией (ЛЧМ) и прямоугольной огибающей.

Изменение частоты f внутри ЛЧМ-радиоимпульса происходит по линейному закону:

$$f(t) = f_0 + \Delta f \cdot \frac{t}{\tau}, \quad 0 \leq t \leq \tau,$$

где f_0 — начальная частота колебаний; Δf — ширина полосы частот; τ — длительность импульсов.

При генерации ДП в симуляторе BSS принимались следующие параметры зондирующего сигнала: $\tau = 30$ мкс, $\Delta f = 160$ МГц, $f_0 = 10$ ГГц (т. е. длина радиоволны на начальной частоте равна 3 см), частота следования импульсов 2000 Гц.

Заметим, что дальностный отсчет определяется удаленностью зондирующего радиоимпульса в момент его отражения от ВЦ, а амплитуда принятого сигнала для каждого дальностного отсчета — мощностью отраженной радиоволны.

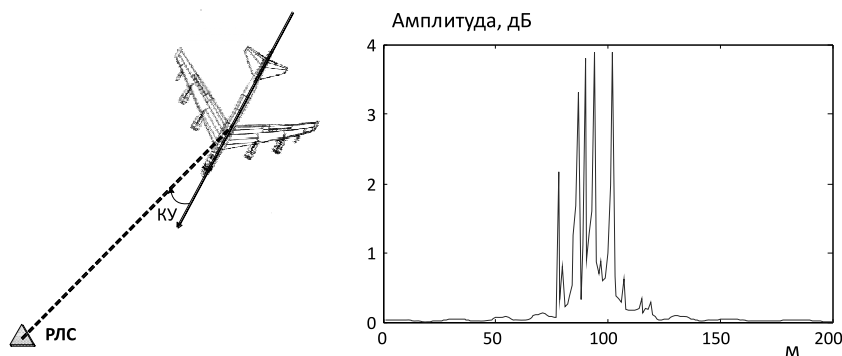


Рис. 1. Дальностный портрет бомбардировщика В-52

Далее дальностные портреты рассматриваются как одномерные функции:

$$A = \pi(\xi), \quad (1)$$

где ξ — дальность отсчета (м), A — амплитуда отклика (дБ). Предполагаем, что $0 \leq \xi \leq L$ (L — ширина строка дальности зондирующих сигналов [12]).

Радиолокационную цель можно представлять в виде множества точечных рассеивателей радиоволн [13], таких как поверхности, края, двугранные углы, трехгранники, впадины и пр. Тогда ДП имеет смысл эпюры амплитуд принятого эхосигнала, развернутой по пространственной координате, определяющей удаленность сформировавших их рассеивателей.

На рис. 1 в качестве примера приведен ДП бомбардировщика В-52.

Задача распознавания ВЦ

Процедура распознавания ВЦ на основе анализа их ДП предполагает решение двух задач:

- 1) формирование комплекса информативных признаков ВЦ по ее ДП;
- 2) классификация ВЦ на основе данных информативных признаков.

Введем обозначения:

- Ω — множество наблюдаемых ВЦ;
- $\omega \in \Omega$ — объекты наблюдения;
- $\{\Omega_1, \Omega_2, \dots, \Omega_m\}$ — алфавит классов ВЦ;
- $\Theta = \{\theta_1, \theta_2, \dots, \theta_m\}$ — алфавит меток (имен) классов ВЦ;
- $\theta_i \in \Theta$ ($i = 1:m$) — метки (имена) классов;
- Π — множество дальностных портретов ВЦ;
- $\pi \in \Pi$ — дальностные портреты.

Каждая ВЦ относится к некоторому классу и, следовательно, ей априори соответствует определенная метка класса:

$$\omega \rightarrow \theta.$$

Результатом радиолокационного наблюдения является дальностный портрет ВЦ:

$$\omega \rightarrow \pi.$$

Распознавание объектов наблюдения основано на формировании комплекса информативных признаков x_1, x_2, \dots, x_n , которые образуют систему координат в пространстве признаков X . В итоге объект наблюдения $\omega \in \Omega$ в некоторый фиксированный момент времени $t \in T$ представляется кортежем $\mathbf{x} \in X$:

$$\mathbf{x} = (x_1, x_2, \dots, x_n). \quad (2)$$

На этапе формирования информативных признаков осуществляется преобразование

$$\pi \rightarrow \mathbf{x}.$$

Далее посредством анализа информативных признаков алгоритм распознавания вычисляет соответствующую метку класса $y \in \Theta$, которая и *присваивается* наблюдаемой ВЦ. В случае правильной работы алгоритма будет получено истинное значение метки ВЦ:

$$y = \theta.$$

Будем различать *статические* и *динамические* схемы распознавания движущихся ВЦ. Первые рассчитаны на анализ статического образа ВЦ, полученного в результате одноразового наблюдения ВЦ, а во вторых анализируется динамический образ ВЦ, представляющий временную последовательность ДП, полученных в течение определенного периода времени слежения за целью.

Алгоритм статической классификации осуществляет функциональное преобразование

$$\mathbf{x} \rightarrow y,$$

т. е. он для каждого входного образа $\mathbf{x} \in X$ вычисляет соответствующую метку класса $y \in \Theta$, которая и *присваивается* наблюдаемой ВЦ.

В схемах динамического распознавания ВЦ агрегируются данные наблюдения для различных курсовых углов, получаемых в разные моменты времени отслеживания движения ВЦ.

Информативные признаки

Надежность процедуры распознавания в значительной мере зависит от выбора системы информативных признаков. Для идентификации дальностных портретов воздушных целей предлагается применять различные информативные признаки. В работе авторов [10] обсуждались морфологические, геометрические и вейвлетные характеристики дальностных портретов. Здесь мы ограничимся изложением весьма простого способа конструирования информативных характеристик ВЦ, основанного на гармоническом анализе ДП.

Рассмотрим задачу аппроксимации ДП конечным рядом Фурье:

$$\pi(\xi) \approx \hat{\pi}_m(\xi),$$

где $\hat{\pi}_m(\xi)$ — тригонометрический полином n -го порядка:

$$\hat{\pi}_m(\xi) = \frac{1}{2} a_0 + \sum_{k=1}^n \left[a_k \cos\left(\frac{2\pi k}{L} \xi\right) + b_k \sin\left(\frac{2\pi k}{L} \xi\right) \right]. \quad (3)$$

В силу ортогональности тригонометрического базиса в (3) для коэффициентов Фурье справедливы формулы:

$$a_k = \frac{2}{L} \int_0^L \pi(\xi) \cos\left(\frac{2\pi k}{L} \xi\right) d\xi, \\ b_k = \frac{2}{L} \int_0^L \pi(\xi) \sin\left(\frac{2\pi k}{L} \xi\right) d\xi. \quad (4)$$

Динамический аспект распознавания воздушных целей

На эффект отражения радиоволн от ВЦ в сторону РЛС существенно влияет *курсовой угол* (КУ) — угол между продольной осью ЛА и направлением на РЛС в плоскости истинного горизонта последней, отсчитываемый по ходу часовой стрелки. Таким образом, ДП протяженных ВЦ определяется не только их конфигурацией, но и курсовым углом ВЦ. Для учета этого фактора в выражение (1) необходимо внести параметр γ — значение курсового угла ($^\circ$):

$$A = \pi(\gamma, \xi).$$

Далее предполагаем, что $0 \leq \gamma \leq 360^\circ$.

Отметим, что ДП на рис. 1 отвечает КУ = 20° .

Пусть ВЦ движется по некоторой траектории, как показано на рис. 2. Поскольку в процессе движения меняется ее угловое положение относительно РЛС, то различным моментам времени наблюдения

$$t_1 < t_2 < \dots < t_d$$

будут отвечать различные курсовые углы

$$\gamma(t_i), \quad i = 1:d,$$

так что на выходе приемника РЛС мы сможем получить семейство соответствующих ДП

$$\pi(\gamma(t_i), \xi), \quad i = 1:d, \quad (5)$$

совместное использование которых и лежит в основе динамического распознавания ВЦ.

Далее T — множество моментов времени съема радиолокационной информации:

$$T = \{t_i, \quad i = 1:d\}. \quad (6)$$

Отметим два эффекта в процессах радиолокационного наблюдения, существенные для решения обсуждаемой задачи радиолокационного распознавания ВЦ.

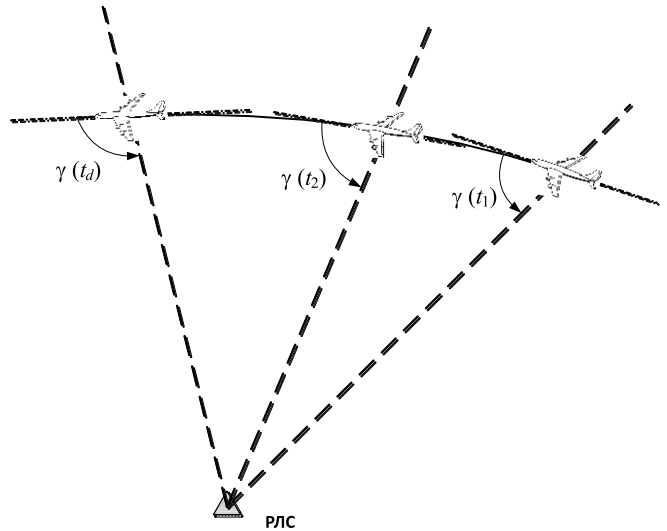


Рис. 2. Изменение КУ ЛА в процессе слежения за ВЦ

Во-первых, ДП могут существенно изменяться при малых изменениях КУ. Это обстоятельство иллюстрирует рис. 3, где представлено семейство ДП самолета В-52 в узком диапазоне КУ: $30 \dots 30,6^\circ$.

Во-вторых, разнотипные ВЦ могут иметь схожие ДП при некоторых значениях КУ и вследствие этого давать неразличимые образы в пространстве признаков X . Данный эффект иллюстрирует рис. 4. На нем представлены дальностные портреты самолетов ТУ-16 и В-52 для КУ $\approx 179^\circ$. Видно их большое сходство.

Однако необходимая различимость ДП тех же ВЦ может проявляться при новых условиях наблюдения. Это обстоятельство подтверждает сравне-

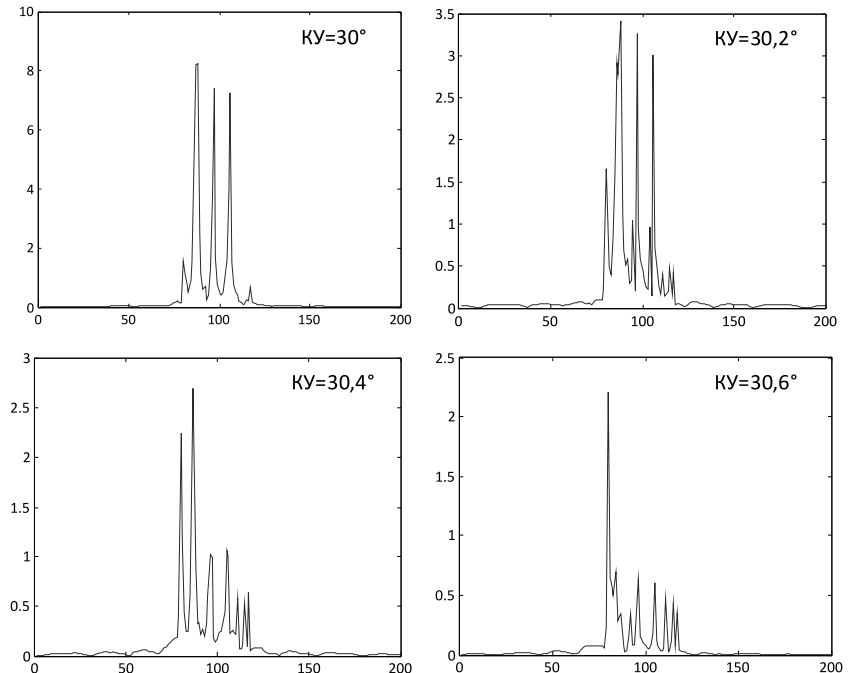


Рис. 3. Семейство ДП В-52 для ряда значений КУ

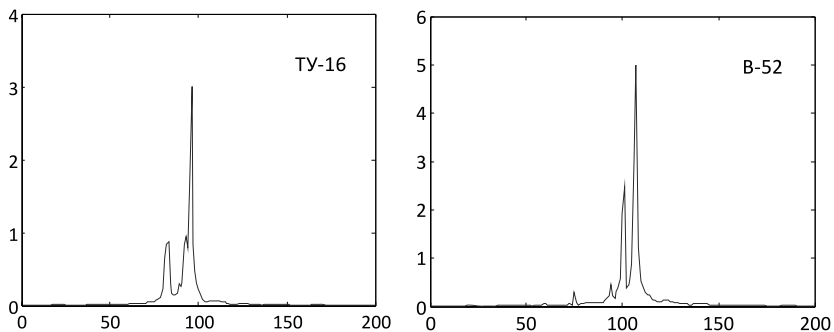


Рис. 4. ДП бомбардировщика ТУ-16 и В-52 для КУ ≈ 179°

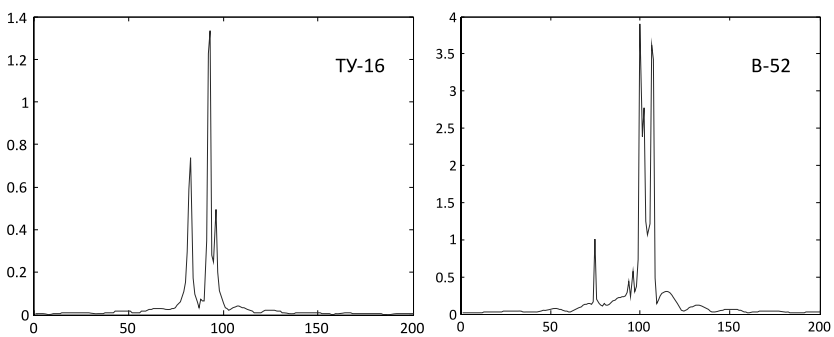


Рис. 5. ДП бомбардировщика ТУ-16 и В-52 для КУ = 177°

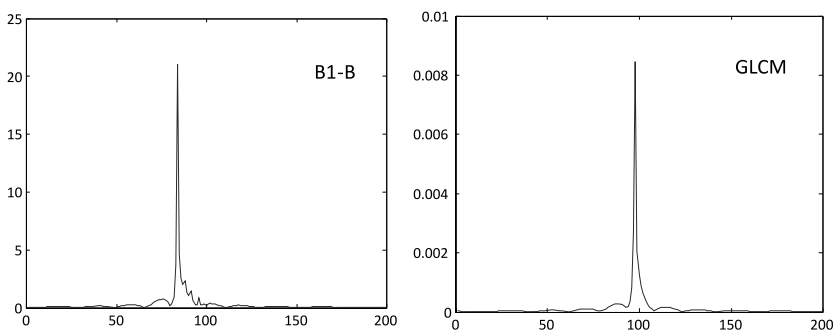


Рис. 6. ДП самолета В1-В и ракеты GLCM, КУ ≈ 179°

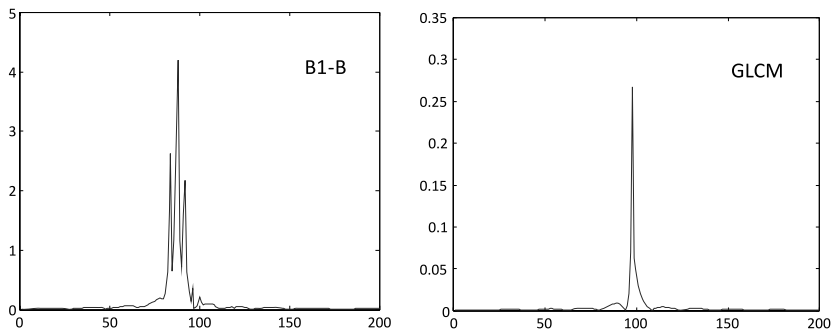


Рис. 7. ДП самолета В1-В и ракеты GLCM, КУ = 175°

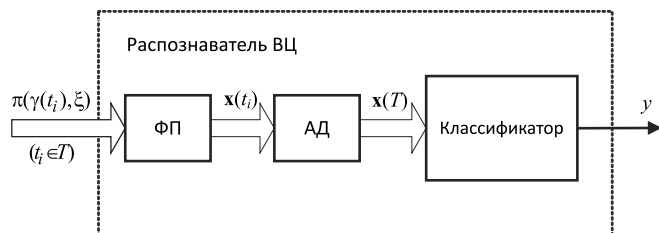


Рис. 8. Схема распознавания с агрегированием признаков

ние ДП В-52 и ТУ-16 для КУ = 177°, представленных на рис. 5.

Эффект сходства ДП, если не учитывать уровень мощности принятых сигналов, весьма наглядно демонстрирует рис. 6: здесь представлены ДП самолета В1-В и ракеты GLCM для КУ ≈ 179°. Однако в других условиях наблюдения (КУ = 175°), как показывает рис. 7, мы получаем несхожие ДП.

Таким образом, статические схемы распознавания ВЦ могут давать ошибочные ответы вне зависимости от выбора комплекса информативных признаков. Негативные последствия схожести ДП для отдельных значений КУ можно избежать посредством схемы *динамического* распознавания ВЦ, в которой используется совокупность данных наблюдения для различных КУ, получаемых в разные моменты времени отслеживания движения ВЦ.

Возможны два способа организации динамического распознавания ВЦ, различающихся порядком обработки данных, полученных для разных условий наблюдения.

В первом способе используется агрегирование входных данных, подаваемых на вход классификатора. Во втором способе используется двухступенчатая схема классификации, включающая первичный и вторичный классификаторы, и процедура агрегирования применяется по отношению к выходным данным первичного классификатора.

Метод агрегирования информативных признаков

Суть первого способа состоит в расширении пространства информативных признаков. Соответствующую схему распознавания поясняет рис. 8. Она включает три блока: блок формирования информативных признаков (ФП), блок агрегирования данных (АД) и классификатор.

Для каждого ДП в (5) формируется кортеж признаков $\mathbf{x}(t_i)$ согласно (2), которые далее объединяются в массив:

$$\mathbf{x}(T) = [\mathbf{x}(t_1), \mathbf{x}(t_2), \dots, \mathbf{x}(t_d)]. \quad (7)$$

Он представляет динамический образ объекта в *расширенном пространстве признаков* \bar{X} — декартовой степени исходного пространства признаков:

$$\mathbf{x}(T) \in \bar{X},$$

где

$$\bar{X} = X^d.$$

На вход классификатора подается вектор признаков (7).

Существенным недостатком такого подхода является d -кратное увеличение размерности признакового пространства, что будет неизбежно сказываться на росте требуемых вычислительных ресурсов на программную реализацию и обучение классификатора.

Двухступенчатая схема классификации

Второй способ построения динамического распознавания выгодно отличается от первого тем, что не вызывает кратного увеличения размерности решаемых задач распознавания ВЦ. Он предполагает применение двух классификаторов: первичного и вторичного.

Первичный классификатор выполняет предварительную классификацию ДП на основе анализа информативных признаков ВЦ — он обрабатывает каждый кортеж признаков

$$\mathbf{x}(t_i) \quad (i = 1:d),$$

присваивая им *расчетные* метки класса принадлежности ДП:

$$\mathbf{z}(t_i) \quad (i = 1:d).$$

Из них в блоке агрегирования формируется кортеж меток:

$$\mathbf{z}(T) = [z(t_1), z(t_2), \dots, z(t_d)]. \quad (8)$$

Заметим, что

$$\mathbf{z}(T) \in \bar{\Theta},$$

где

$$\bar{\Theta} = \Theta^d.$$

Это пространство логично именовать *расширенным пространством* меток.

Данные (8), накопленные за период наблюдения (6), представляют динамический образ ВЦ, но не в пространстве информативных признаков, а в пространстве меток. Далее они используются во *вторичном классификаторе*, который вырабатывает итоговую оценку метки класса принадлежности ВЦ.

Предложенную схему распознавания ВЦ представляет рис. 9. Она включает четыре блока: блок формирования информативных признаков (ФП), первичный классификатор, блок агрегирования данных (АД) и вторичный классификатор.

Идея предлагаемой двухступенчатой схемы классификации состоит в том, чтобы первичную классификацию считать в качестве промежуточной.

При этом результаты ошибочной классификации на первой ступени не отбрасываются, а учитываются как *фактор неопределенности*, который затем устраняется переходом к динамическому образу ВЦ и действием второй ступени классификации.

Классификаторы на основе деревьев решений

Возможны различные варианты выбора классификаторов в рассматриваемых схемах распознавания ВЦ. Ограничимся обсуждением опыта применения авторами деревьев решений (ДР) для реализации двухступенчатой схемы распознавания, представленной на рис. 9.

Деревья решений представляют собой последовательные иерархические структуры, состоящие из узлов, которые содержат правила вида ЕСЛИ—ТО. Конечными узлами дерева являются "листья", соответствующие найденным решениям и объединяющие некоторое число объектов классифицируемой выборки [14].

В дереве имеется один особый узел — *корневой*. От него можно перейти по дереву к любому другому узлу. В конце любой цепочки подряд идущих ребер находятся терминальные узлы — *листья*, которые соответствуют определенным классам. Нетерминальные узлы — узлы принятия решений. Они содержат критерии выбора, а выходящие из них ребра выражают взаимоисключающие результаты проверки соответствия этим критериям. По существу, при каждой проверке условия происходит сортировка выборок данных таким образом, что каждый элемент данных определяется как соответствующий только одному ребру.

Если переменная числового типа, то в узле формируется правило вида

$$x \leq c,$$

где c — некоторый порог, который чаще всего выбирается как среднее арифметическое двух соседних упорядоченных значений переменной x обучающей выборки.

В работе использовался алгоритм построения дерева решений CART. Программная реализация полученных алгоритмов и их апробация выполнялись в системе MATLAB, причем в качестве инструментария построения и применения деревьев решений использовался пакет прикладных программ Statistic Toolbox.

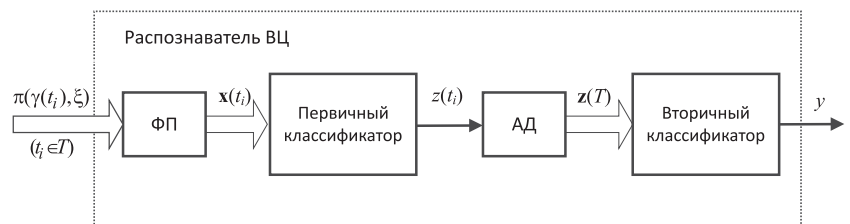


Рис. 9. Двухступенчатая схема классификации

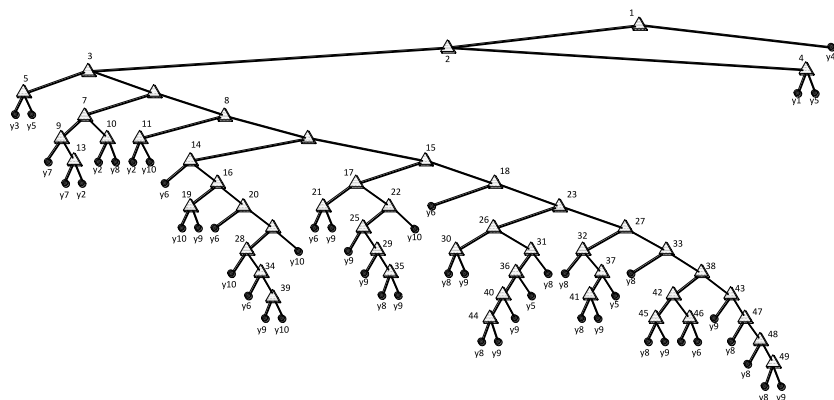


Рис. 10. Дерево решений для ступени первичной классификации

Исходные ДП воздушных целей получены посредством программы BSS. Анализировались 10 типов ВЦ: АН-26, АН-64, МИГ-21, АLCM, В-1В, В-52, F-15, GLCM, TORNADO, ТУ-16.

Рассматриваемым типам самолетов сопоставляется множество меток

$$Y = \{y_i, i = 1:m\},$$

причем $m = 10$,

$$\begin{aligned} y_1 &= \text{АН-64}; & y_6 &= \text{F-15}; \\ y_2 &= \text{ALCM}; & y_7 &= \text{GLCM}; \\ y_3 &= \text{АН-26}; & y_8 &= \text{МИГ-21}; \\ y_4 &= \text{В-1В}; & y_9 &= \text{TORNADO}; \\ y_5 &= \text{В-52}; & y_{10} &= \text{ТУ-16}. \end{aligned}$$

Критериальные характеристики узлов дерева решений

№ узла	Критерий выбора	№ узла	Критерий выбора
1	$x_6 (26,65)$	26	$x_{18} (-29,65)$
2	$x_9 (26,55)$	27	$x_{19} (15,45)$
3	$x_{11} (-6,8)$	28	$x_7 (20,85)$
4	$x_{14} (-18,95)$	29	$x_{19} (20,6)$
5	$x_5 (64,4)$	30	$x_1 (109,2)$
6	$x_3 (20,05)$	31	$x_4 (25,95)$
7	$x_{13} (2,25)$	32	$x_3 (25,8)$
8	$x_7 (11,85)$	33	$x_3 (36,25)$
9	$x_{10} (-0,9)$	34	$x_{17} (81)$
10	$x_9 (4,3)$	35	$x_3 (44,1)$
11	$x_1 (41,25)$	36	$x_1 (109,7)$
12	$x_4 (-64,25)$	37	$x_1 (70,9)$
13	$x_{11} (0,15)$	38	$x_{19} (26,25)$
14	$x_{19} (40,25)$	39	$x_1 (294,6)$
15	$x_9 (8,25)$	40	$x_3 (42,1)$
16	$x_{13} (62,4)$	41	$x_{10} (-9,5)$
17	$x_6 (-25,9)$	42	$x_5 (30,3)$
18	$x_{11} (2,15)$	43	$x_{12} (-23)$
19	$x_5 (50,2)$	44	$x_{16} (-17,55)$
20	$x_9 (-35,95)$	45	$x_{18} (-23,8)$
21	$x_{19} (22,2)$	46	$x_1 (79)$
22	$x_1 (85)$	47	$x_{18} (-38)$
23	$x_{10} (-10,55)$	48	$x_3 (51,15)$
24	$x_9 (-10,1)$	49	$x_{18} (-36,35)$
25	$x_9 (5,55)$		

Рассматриваемый диапазон изменения курсового угла:

$$1^\circ \leq \text{КУ} < 11^\circ. \quad (9)$$

Было выбрано 19 информативных признаков, состоящих из 10 косинусных и 9 синусных коэффициентов (4) в аппроксимирующем тригонометрическом полиноме (3).

Для обучения ДР первой ступени классификации используются все ДП, отвечающие угловому диапазону (9) с шагом изменения угла $0,05^\circ$. Таким образом, обучающая выборка включает 200 ДП для каждого типа ВЦ и, следовательно, общий объем обучающей выборки равен 2000.

Построенное ДР (первичный классификатор) представлено на рис. 10.

Атрибуты построенного ДР: x_1, x_2, \dots, x_{19} — спектральные характеристики ДП (число гармоник $m = 9$ в тригонометрическом многочлене (3)).

Узлы дерева пронумерованы. Их критериальные характеристики отражает таблица: здесь для каждого узла указан атрибут (признак), по которому происходит сортировка данных, и соответствующее пороговое значение (приведено в скобках).

Тестирование построенного ДР (первичный классификатор) показывает, что на обучающей выборке из 2000 образов правильно распознаются 1973 образа (99,6 %) и неправильно — 27 образов (1,4 %).

Результаты тестирования обученной ДР для $\text{КУ} = 1,3^\circ$ представлены на рис. 11. Видно, что в данном случае неправильно классифицируется TORNADO — данный самолет идентифицируется как МИГ-21.

Однако для $\text{КУ} = 3,1^\circ$ ситуация другая — осуществляемая ДР классификация является точной (рис. 12).

Данный результат поясняет смысл агрегирования данных наблюдения — это позволяет корректировать результаты ошибочной работы первичного классификатора.

В исследуемой схеме распознавания агрегирование выполнялось для трех последовательных значений КУ (9) с шагом $0,05^\circ$.

ДР для второй ступени классификации представлено на рис. 13.

Для анализа эффективности разработанной схемы динамической классификации ВЦ тестирующая выборка формировалась из ДП, курсовые углы которых не были включены в обучающую выборку:

- диапазон курсового угла: $3,02^\circ \leq \text{КУ} \leq 3,97^\circ$;
- шаг изменения угла $0,05^\circ$.

Результаты тестирования показывают, что построенная на базе ДР двухступенчатая схема распознавания ВЦ дает 100 % правильных ответов.

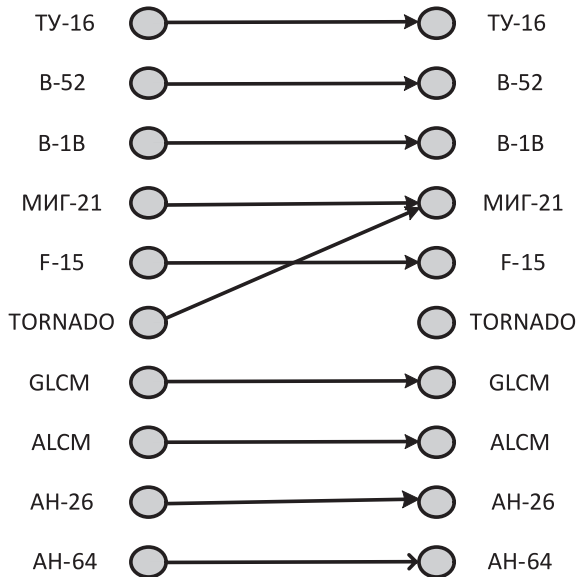


Рис. 11. Результат тестирования ДР для $KV = 1,3^\circ$

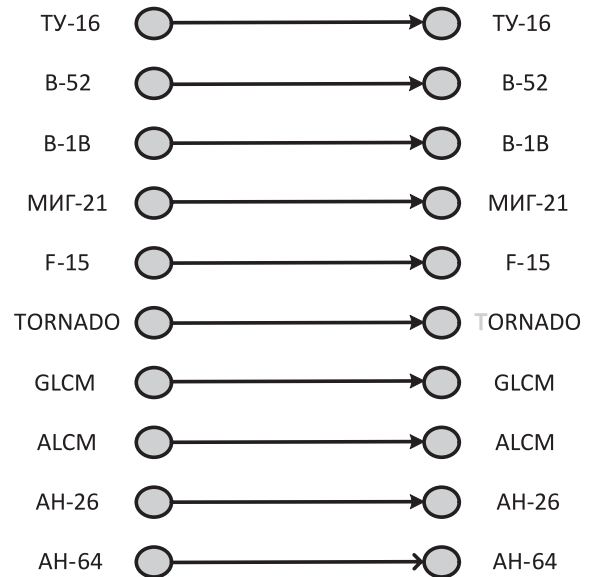


Рис. 12. Результат тестирования ДР для $KV = 3,1^\circ$

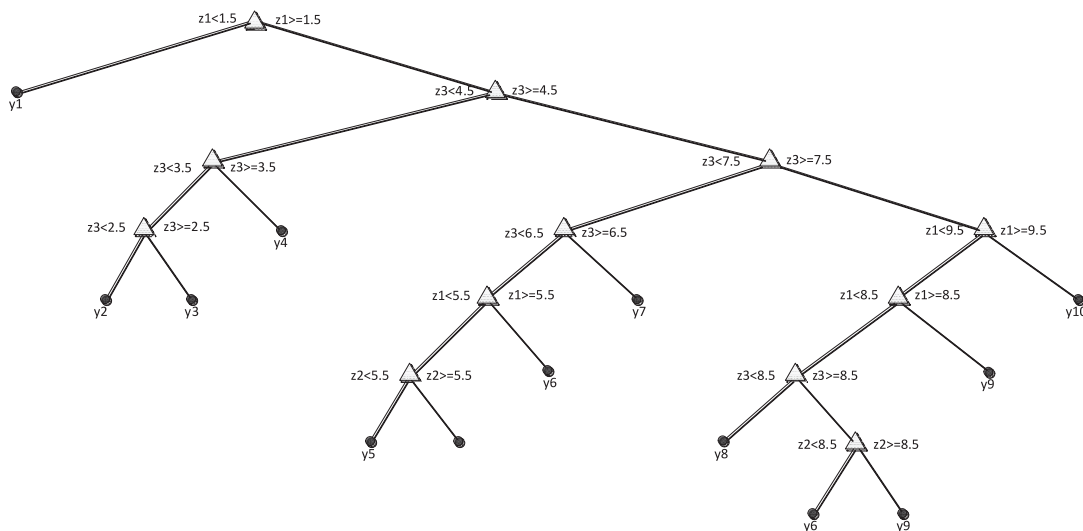


Рис. 13. Дерево решений для вторичной классификации

Отметим, что избыточная входная информация об объектах наблюдения может существенно усложнять алгоритмы их классификации. В связи с этим несомненный научный интерес представляют вопросы оптимального комплексирования информативных признаков различных типов, рассмотренных в работе [10] и в данной статье, для инженерного проектирования систем автоматического распознавания ВЦ.

Список литературы

1. Белоцерковский Г. Б. Основы радиолокации и радиолокационные устройства. М.: Советское радио, 1975. 336 с.
2. Бакулев П. А. Радиолокационные системы. М.: Радиотехника, 2007. 376 с.

3. Tait P. Introduction to radar target recognition. London: Institution of Electrical Engineers, IET radar series, 2005. N. 18. 396 p.
4. Лещенко С. П. Информативность и достоверность радиолокационного распознавания воздушных целей по дальностным портретам // Збірник наукових праць ХВУ. Вип. 7 (37). X., 2001. С. 63—72.
5. Лещенко С. П. Развитие теории и техники радиолокационного распознавания воздушных целей // Прикладная радиоэлектроника, 2009, Т. 8, № 4. С. 490—496.
6. Тоцкий А. В., Молчанов П. А., Поспелов Б. Б. Распознавание летательных аппаратов по радиолокационным дальностным профилям // Авиационно-космическая техника и технология. 2010. № 5 (72). С. 77—82.
7. Li H. J., Yang S. H. Using range profiles as feature vectors for identify aerospace objects // IEEE Trans. Antennas and Propagation. 1993. Vol. 41, N. 3. P. 261—268.
8. Jacobs S. P., O'Sullivan A. Automatic target recognition using sequences of high resolution radar range profiles // IEEE Trans. on Aerospace and Electronic Systems. 2000. Vol. 36, N. 2. P. 364—381.

9. Zhou D., Liu G., Wang J. Spatio-temporal target identification method of high-range resolution radar // *Pattern Recognition*. 2000. Vol. 33, N. 1. P. 1–7.

10. Филимонов А. Б., Фам Ф. К. Методы формирования информативных признаков радиолокационных дальностных портретов воздушных целей // *Мехатроника, автоматизация, управление*. 2016. № 4. С. 273–281.

11. Васильев В. И. Распознающие системы: Справочник. Киев: Наукова думка, 1983. 422 с.

12. Shirman Y. D., Gorshkov S. A., Leshchenko S. P., Orlenko V. M., Sedyshev S. Y., Sukharevskiy O. I. *Computer Simulation of Aerial Target Radar Scattering, Recognition, Detection, and Tracking* / Ed. Y. D. Shirman. Boston—London: Artech house, 2002. 294 p.

13. August W. R., Stephen J. H. *Theory and practice of radar target identification*, Artech House, 2000. 738 p.

14. Breiman L., Friedman J. H., Olshen R. A., Stone C. J. *Classification and regression trees*. Monterey, CA: Wadsworth & Brooks. 1984. 358 p.

A. B. Filimonov, Professor, filimon_ab@mail.ru, Fam Fuong Kuong, an intern, cuongbkedu@yahoo.com, Moscow Technological University

Dynamic Aerial Target Recognition by Using Radar Range Profiles

This article considers the problem of aerial target recognition based on analysis characteristics of radar range profiles and also information base structure of recognition system. We will compare static and dynamic recognition schemes. Firstly static pattern of target is built, this pattern is data received at a moment. Secondly dynamic patterns of target are built, they are tracking data of target in a specific period of time.

There is an important feature of radar range profiles: they have complex configuration and their range profiles can be greatly changed when change observation angle. Moreover, some aerial targets of different types in different course angles may have similar range profiles and therefore generate indistinguishable pattern in the signature space. Thus, the static recognition schemes may give wrong answers regardless of the choice of informative signatures and make it difficult or even exclude the possibility of their practical use.

Studies show that required reliability and quality of moving target recognition can be obtained by applying dynamic recognition scheme. In this scheme, dynamic pattern is built by data aggregation at different moments of radar observation.

Radar observation occurs in noise and interference conditions, which limits resolution and accuracy of radar, uncertainties and variability of reflection properties of targets. Therefore, only robust recognition algorithms can have practical value, they are designed to work in conditions of uncontrolled disturbing factors. These dynamic recognition schemes allow to minimizing the impact of uncertainty factors by expanding data base of recognition algorithms.

The first stage of radar recognition procedure is forming informative signatures, which is considered as an input pattern tuple of target observation, then it solves the problem pattern classification. There are two ways of dynamic aerial target recognition, which differ in the processing order of observation data. The first method uses aggregation of input data, which is supplied to the input of the classifier. In the second method, aggregation procedure applies with respect to the output data of the classifier.

Keywords: radar, radar observation, aerial target recognition, range profiles, informative signatures, dynamic patterns, data aggregation, dynamic recognition schemes

References

1. Belocerkovskij G. B. *Osnovy radiolokacii i radiolokacionnye ustrojstva* (Radar basics and radar devices). Moscow: Sovetskoe radio, 1975. 336 p.

2. Bakulev P. A. *Radiolokacionnye sistemy* (Radar systems). Moscow: Radiotekhnika, 2007. 376 p.

3. Tait P. *Introduction to radar target recognition*. London: Institution of Electrical Engineers, IET radar series, 2005, no. 18. 396 p.

4. Leshchenko S. P. Informativnost' i dostovernost' radiolokacionnogo raspoznavaniya vozdušnyh celej po dal'nostnym portretam (Information capability and reliability of aerial target radar recognition), *Zbirnik naukovih prac' HVU*, 2001, Iss. 7 (37). H., pp. 63–72.

5. Leshchenko S. P. Razvitie teorii i tekhniki radiolokacionnogo raspoznavaniya vozdušnyh celej (Development of the theory and technology of aerial target radar recognition), *Prikladnaya radioelektronika*, 2009, vol. 8, no. 4, pp. 490–496.

6. Tockij A. V., Molchanov P. A., Pospelov B. B. Raspoznavanie letatel'nyh apparatov po radiolokacionnym dal'nostnym profilam (Aircraft recognition using radar range profiles), *Aviacionno-kosmicheskaya tekhnika i tekhnologiya*, 2010, no. 5 (72), pp. 77–82.

7. Li H. J., Yang S. H. Using range profiles as feature vectors for identify aerospace objects, *IEEE Trans. Antennas and Propagation*, 1993, vol. 41, no. 3, pp. 261–268.

8. Jacobs S. P., O'Sullivan A. Automatic target recognition using sequences of high resolution radar range profiles, *IEEE Trans. on Aerospace and Electronic Systems*, 2000, vol. 36, no. 2, pp. 364–381.

9. Zhou D., Liu G., Wang J. Spatio-temporal target identification method of high-range resolution radar, *Pattern Recognition*, 2000, vol. 33, no. 1, pp. 1–7.

10. Filimonov A. B., Fam F. K. Metody formirovaniya informativnyh priznakov radiolokacionnyh dal'nostnyh portretov vozdušnyh celej (Methods for forming informative signatures of radar range profiles of aerial targets), *Mehatronika, avtomatizacija, upravlenie*, 2016, no. 4, pp. 273–281.

11. Vasil'ev V. I. *Raspoznajushhie sistemy* (Recognition systems). Spravochnik. Kiev: Naukova dumka, 1983. 422 p.

12. Shirman Y. D., Gorshkov S. A., Leshchenko S. P., Orlenko V. M., Sedyshev S. Y., Sukharevskiy O. I. *Computer Simulation of Aerial Target Radar Scattering, Recognition, Detection, and Traking* / Ed. Shirman Y. D. Boston—London: Artech house, 2002. 294 p.

13. August W. R., Stephen J. H. *Theory and practice of radar target identification*, Artech House, 2000. 738 p.

14. Breiman L., Friedman J. H., Olshen R. A., Stone C. J. *Classification and regression trees*. Monterey, CA: Wadsworth & Brooks. 1984. 358 p.

П. С. Поперечный, аспирант ИППМ РАН, e-mail: prpoperechny@elvees.com
Институт проблем проектирования в микроэлектронике РАН

Сравнительная оценка корректирующей способности кодов Рида—Соломона и Боуза—Чоудхури—Хоквингема

Приведены сравнительные характеристики кодов Рида—Соломона (РС) и Боуза—Чоудхури—Хоквингема (БЧХ), которые находят широкое применение, в том числе для помехоустойчивого хранения данных. Предложен способ сравнения корректирующей способности кодов исправлять пакетные ошибки. Приведено аналитическое выражение для построения передаточной функции декодеров РС и БЧХ для двоичного симметричного канала (ДСК). Показано, что при одинаковых длинах кодированных данных и скоростях кода кодек БЧХ исправляет лучше случайно распределенные ошибки, однако проигрывает кодеку РС, если ошибки сгруппированы. Также проведен синтез логической схемы двух кодеков, приведено их сравнение по занимаемой площади и задержек на критическом пути.

Ключевые слова: РС, БЧХ, корректирующая способность, ДСК, поле Галуа

Введение

Помехоустойчивые коды Рида—Соломона (РС) и Боуза—Чоудхури—Хоквингема (БЧХ) находят широкое применение в системах передачи-приема информации, хранения данных и т. д. Данные коды относятся к блочным помехоустойчивым кодам, кодирующим относительно большие блоки данных и способным исправлять множественные ошибки в этих блоках [1, 2].

Ввиду того, что данные коды с самого начала широко применялись в системах передачи-приема информации в радиосвязи, существуют различные способы оценки корректирующей способности кодов, основанные на математических моделях каналов связи [1]. Классификация каналов связи — достаточно широкая в зависимости от применения, однако все модели наиболее приближены к системам передачи информации: двоичный симметричный канал (ДСК), канал с аддитивным белым гауссовым шумом (АБГШ), канал с замиранием и т. д. Модели каналов делятся на непрерывные и дискретные, с памятью и без памяти. Например, дискретный канал без памяти — это такой канал, для которого выходной символ в данный момент статистически зависит лишь от соответствующего входного символа. Дискретные каналы с памятью — это такие каналы, в которых выход в данный момент статистически зависит как от текущего входа, так и от предыдущих входов и выходов [3].

В статье рассмотрены коды РС и БЧХ применительно к устройствам хранения информации, таким как flash-память, SSD-диск [4], для которых известные модели каналов связи не подходят ввиду разной природы ошибок.

Зачастую составить полную модель канала сложно из-за различных видов искажений [5, 6]. Во flash-памяти причинами искажений бита данных могут быть: цикл записи/стирания, парное соединение битов, ошибки сохранения данных, инверсия битов, случайно распределенные дефектные блоки, ограни-

ченный ресурс, уменьшение технологии производства и применение многоуровневой ячейки и, как следствие, уменьшение уровня напряжения на бит [7]. Ввиду множества причин искажения информации не исключено появление пакетных ошибок. При этом такие способы, как переспрос или перемеживание, могут оказаться неэффективными. В статье предложена модель канала с пакетными ошибками и показано преимущество кодов РС в этом случае.

Эквивалентные коды БЧХ и РС для применения в носителях информации

Согласно стандарту ONFI [8], flash-память делится на область хранения непосредственно информационных данных и на дополнительную область (spare-область) для хранения контрольных битов. Выберем наиболее подходящие и эквивалентные коды РС и БЧХ в части избыточности и длины кодированных данных.

Коды РС являются недвоичными, для систем хранения данных предпочтительнее, чтобы символ представлялся байтом [9]. Таким образом, в поле Галуа $GF(2^8)$ выберем для рассмотрения код РС(255, 239) с числом битов на символ $m = 8$. То есть $n = 255$ — число закодированных символов, $k = 239$ — число информационных символов. Данный код позволяет исправлять восемь символов в кодовом слове. Каждый символ является 8-битным. На рис. 1 показана структура кода и одна из

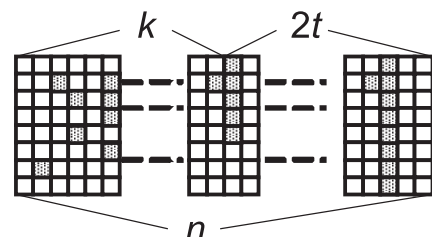


Рис. 1. Структура кода РС с возможными ошибками

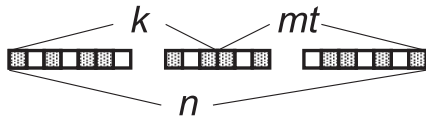


Рис. 2. Структура кода БЧХ с возможными ошибками

комбинаций ошибок (черным выделены ошибочные биты), которую может исправить данный код.

Символы могут быть испорчены частично либо полностью, любую комбинацию код РС(255, 239) способен исправить, если число испорченных символов не более $t = 8$. Если представить этот код в битах, т. е. в поле $GF(2)$, то получится код $(n \cdot m, k \cdot m) = (255 \cdot 8, 239 \cdot 8) = (2040, 1912)$.

Эквивалентный код по скорости и длине данных подходит БЧХ (2047, 1915) в поле $GF(2^{11})$, способный исправлять $t = 12$ случайных битовых ошибок. На рис. 2 представлена структура кода и возможная комбинация исправляемых ошибок.

Таким образом, из вышеприведенного следует, что код БЧХ(2047, 1915) исправляет 12 случайных битовых ошибок, в то время как код РС(255, 239) примерно с тем же числом информационных и контрольных битов исправляет лишь $t = 8$ случайных ошибок, но байтовых.

Для оценки исправляющей способности кода строят передаточную функцию — зависимость числа неисправленных ошибок на выходе из декодера от числа ошибок в канале, т. е. на входе декодера [10, 11]. Аналитическое выражение для вероятности ошибки на выходе декодера от вероятности битовой ошибки на входе приближенно выражается следующим образом [9, 10, 12, 13]:

$$p_b = \frac{1}{n} \sum_{j=t+1}^n j \binom{n}{j} p^j (1-p)^{n-j}, \quad (1)$$

где $\binom{n}{j} = \frac{n!}{j!(n-j)!}$ — число различных способов выбора

из n битов j ошибочных битов; $\binom{n}{j} p^j (1-p)^{n-j}$ —

вероятность появления j ошибок в блоке из n битов. По сути, выражение (1) показывает вероятность появления ошибок кратности $t + 1$ и выше в блоке из n битов вне зависимости от вида кодиро-

вания. Очевидно, если ошибок больше t , то заявленный кодек не справится, и ошибки появятся уже на выходе декодера. Поэтому формулу (1) для битовой ошибки можно применить и для недвоичных кодов РС, если его рассматривать как двоичный $(n \cdot m, k \cdot m) = (255 \cdot 8, 239 \cdot 8) = (2040, 1912)$. Однако для недвоичных кодов существует несколько другая формула [10], выражающая вероятность появления ошибки в символе на выходе (ошибки в символе показаны на рис. 1):

$$p_E = \frac{1}{2^{m-1}} \sum_{j=t+1}^{2^m-1} j \binom{2^m-1}{j} p_s^j (1-p_s)^{2^m-1-j}, \quad (2)$$

где $p_s = 1 - (1-p)^m$ — вероятность ошибки на символ в зависимости от битовой ошибки [9]. Так как в символе РС содержится m битов, то битовая вероятность на выходе декодера РС выражается формулой

$$p_b = \frac{p_E}{m}. \quad (3)$$

Выражение (1) можно упростить, чтобы не вычислять огромного порядка факториалы для построения зависимостей на графике. Из [10] известно:

$$\frac{1}{n} \sum_{j=2}^n j \binom{n}{j} p^j (1-p)^{n-j} = p - p(1-p)^{n-1}.$$

Потому выражение (1) можно расписать на две суммы и потом упростить:

$$\begin{aligned} p_b &= \frac{1}{n} \sum_{j=t+1}^n j \binom{n}{j} p^j (1-p)^{n-j} = \\ &= \frac{1}{n} \sum_{j=2}^n j \binom{n}{j} p^j (1-p)^{n-j} - \frac{1}{n} \sum_{j=2}^t j \binom{n}{j} p^j (1-p)^{n-j} = \\ &= p - p(1-p)^{n-1} - \frac{1}{n} \sum_{j=2}^t j \binom{n}{j} p^j (1-p)^{n-j} = \\ &= p - p(1-p)^{n-1} - \frac{1}{n} \sum_{j=2}^t j \prod_{i=1}^{j-1} \frac{(n-j+i)}{j!} p^j (1-p)^{n-j}. \end{aligned} \quad (4)$$

Для подтверждения выражений (1), (3), (4) можно собрать схему в среде моделирования MATLAB, как представлено на рис. 3, с примене-

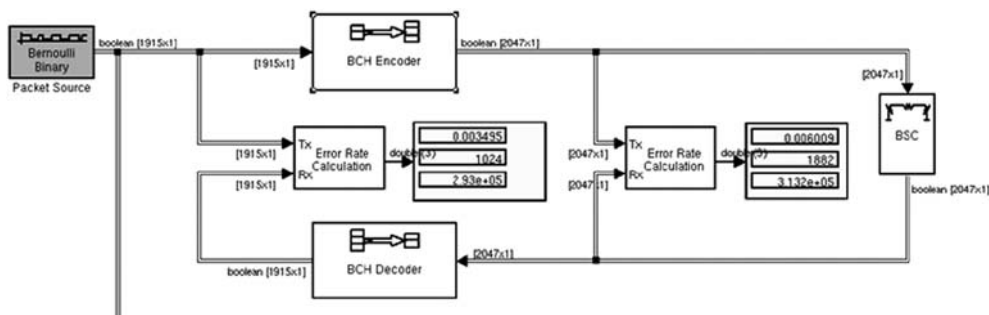


Рис. 3. Схема для оценки передаточной функции кода БЧХ

нием библиотечных кодеков БЧХ и РС с задаваемыми параметрами, генератора случайных данных и двоичного симметричного канала с задаваемой вероятностью ошибки. Аналогичную схему можно собрать для кода РС.

Изменяя в схеме вероятность появления ошибок в канале, можно снять зависимость $BER_{out}(BER_{in})$ — скорость появления ошибки на выходе декодера от скорости на входе. На рис. 4 (см. вторую сторону обложки) изображены данные эксперимента (VSP-b1 RS-b1), а также аналитические зависимости (VCH-b1-calc, RS-b1-calc), описываемые выражениями (3), (4) для кодов БЧХ и РС. Из графика видно, что БЧХ кодек для исправления случайных ошибок обладает лучшей исправляющей способностью (почти 2 порядка при вероятности ошибки на входе $2 \cdot 10^{-3}$). Однако в работах [6, 9] утверждается, что коды РС обладают наилучшей исправляющей способностью среди кодов с такой же избыточностью. Такое утверждение следует из [9], а именно: минимальное расстояние кода (d^*) удовлетворяет неравенству

$$d^* \geq d = 2t + 1 = n - k + 1.$$

Но для любого линейного кода имеет место граница Синглтона:

$$d^* \leq n - k + 1 \Rightarrow d^* = n - k + 1 \Rightarrow d^* = d.$$

Коды РС чрезвычайно эффективны для исправления пакетов ошибок [10]. Поскольку данная схема эксперимента не раскрывает возможности кода РС, необходимо составить другую модель канала.

Оценка корректирующей способности исправлять пакетные ошибки

Существуют способы избегания пакетных, т. е. зависимых, ошибок: перемеживанием данных [3, 12, 14] либо переспросом [15]. В случае работы с flash-памятью нельзя однозначно сказать, что эти способы однозначно устраняют зависимые ошибки ввиду сложной природы их возникновения (см. введение).

Согласно [3] пакетной ошибкой следует считать как минимум два подряд ошибочных бита, отделенных защитным интервалом относительно другого пакета ошибок. Для генерации пакетных ошибок предлагается использовать ДСК подобно схеме рис. 3 с отличием в дополнительном блоке. А именно, предлагается каждый ошибочный бит как бы "размазывать" по данным на задаваемую ширину — собственно ширину пакета ошибок. Ввиду того, что в нашем случае число исправляемых кодеками ошибок значительно меньше длины данных, можно считать, что ошибки в ДСК достаточно прорежены, и защитный интервал будет иметь место. Работа канала ДСК с пакетными ошибками продемонстрирована на рис. 5.

Столь небольшое усложнение канала ДСК позволяет провести ряд экспериментов с целью оценить

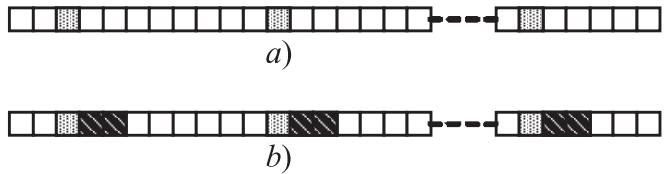


Рис. 5. Ошибки в данных после канала ДСК (а); пакетные ошибки (длина пакета — 3 бита) (б)

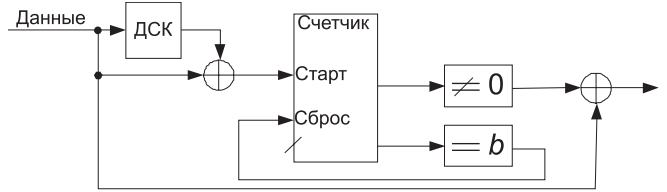


Рис. 6. Схема канала ДСК с пакетами ошибок длины b

корректирующую способность сравниваемых кодов БЧХ и РС. Предлагаемый канал является вырожденным случаем канала с конечным числом состояний [3] или канала Гильберта—Эллиота [16]. При этом вероятность возникновения пакета теперь задается вероятностью появления битовой ошибки или эту же вероятность можно интерпретировать как вероятность перехода в "плохое" состояние, при этом в "хорошем" состоянии данные не портятся вовсе, а в "плохом" портятся с вероятностью 1. В подходах, предложенных в работах [17—20], рассматривается канал Гильберта—Эллиота, однако анализ результатов оказывается весьма сложным ввиду большого числа переменных. В предлагаемом вырожденном канале добавляется лишь один переменный параметр b — длина пакета ошибок. На рис. 6 представлена схема канала ДСК с пакетными ошибками.

Вероятность битовой ошибки на выходе декодера БЧХ в этом случае будет выражаться следующим образом:

$$p_b = \frac{b}{n} \sum_{j=t/b+1}^n \binom{n}{j} p_L^j (1 - p_L)^{n-j}, \quad (5)$$

где $p_L = pb$ — вероятность появления битовых ошибок на входе декодера; p — вероятность появления пакетов ошибок (параметр ДСК); b — длина пакета ошибок.

Экспериментальные результаты

Семейство зависимостей (5) для пакетов ошибок в 1, 2, 6, 12 бит, а также экспериментальные данные приведены на рис. 7 (см. вторую сторону обложки). Очевидно, что при увеличении длины пакета ошибки кодек декодирует все хуже, устремляясь по поведению к не декодированным данным. Это объясняется тем, что для кодов БЧХ нет разницы — сгруппированы ошибки или нет, влияет лишь общее число неисправных битов.

Подобное семейство экспериментальных данных для кода РС(255, 239) представлено на рис. 8 (см. вторую сторону обложки). Видно, что в данном случае нельзя однозначно сказать, что при увеличении длины пакета кодек исправляет ошибки все хуже.

На рис. 9 (см. вторую сторону обложки) представлены сравнительные характеристики кодеков БЧХ и РС в канале ДСК с пакетными ошибками (длина пакетов 2, 8, 13). На рис. 4 было очевидно, что БЧХ для случайных ошибок выигрывает у РС почти 2 порядка, однако уже в случае если ошибки сгруппированы по два бита, кодек РС показывает преимущество. Далее преимущество увеличивается при группировании ошибок по 8 бит (два порядка при вероятности на входе 0,002). При этом характеристика БЧХ при увеличении длины пакета стремится к характеристике не декодированных данных, а РС улучшается. В предельном случае, т. е. при длине пакета ошибок ($b = 13$), превышающей исправляющую способность кода БЧХ (12 исправляемых ошибок), характеристика практически не отличается от характеристики без декодера, однако кодек РС продолжает выполнять успешное декодирование.

Результаты синтеза логической схемы на базе ПЛИС фирмы Altera (Agria II) приведены в таблице.

Заключение

Несмотря на то, что для случайных ошибок преимущества остаются за кодом БЧХ (почти два порядка при вероятности ошибки на входе $2 \cdot 10^{-3}$), предложенная модель канала ДСК с пакетными ошибками наглядно показывает преимущества кода РС (те же два порядка) в сравнении с кодом БЧХ, при прочих равных условиях, а именно скорости кода и длине данных. Даже при группировании ошибок по два бита кодек РС исправляет лучше БЧХ.

Результаты синтеза логической схемы на ПЛИС показали, что для кода РС требуется в 2,5 раза меньше логических элементов, при этом задержка на критическом пути несколько больше. Это объясняется тем, что арифметика для кода РС(255, 239) выполняется в поле Галуа $GF(2^8)$, однако для эквивалентного кода БЧХ (2047, 1912) — в поле $GF(2^{11})$, что усложняет такие элементы, как умножители. Алгоритм РС сложнее БЧХ на один дополнительный шаг (поиск значения ошибки Форни [2]),

Ресурсы ПЛИС, задержка на критическом пути для кодов РС и БЧХ

Сравниваемые параметры	РС	БЧХ
Логические ячейки, шт.	2885	6676
Регистры, шт.	866	1132
Задержка, нс	9,8	7,5

в котором осуществляется затратная по скорости выполнения операция деления в поле Галуа.

В статье представлены как аналитические выражения для битовой ошибки на выходе декодеров БЧХ и РС, так и их упрощенная интерпретация для построения на ЭВМ. Также предложена относительно простая модель канала с пакетными ошибками, наглядно демонстрирующая преимущества одного кода в сравнении с другим. Данный способ может использоваться для проверки различных блоковых, сверточных и других кодов, без ограничения на длину данных и скорость кода.

Список литературы

1. **Блейхут Р.** Теория и практика кодов, контролируемых ошибок. М.: Мир, 1986. 576 с.
2. **Морелос-Сарагоса Р.** Искусство помехоустойчивого кодирования. Методы алгоритмы, применение. М.: Техносфера, 2005. 320 с.
3. **Галлагер Р.** Теория информации и надежная связь. М.: Советское радио, 1974. 720 с.
4. **What Types of ECC Should Be Used on Flash Memory?** Application note, Spansion, 2011, URL: [//www.spansion.com/Support/Application%20Notes/Types_of_ECC_Used_on_Flash_A.N.pdf](http://www.spansion.com/Support/Application%20Notes/Types_of_ECC_Used_on_Flash_A.N.pdf)
5. **Березюк Н. Т.** Кодирование информации. Двоичные коды. Харьков: Вища школа. 1978, 252 с.
6. **Петерсон У., Уэлдон Э.** Коды, исправляющие ошибки. М.: Мир, 1976. 580 с.
7. **Баишев А.** Твердотельные диски — надежное решение для ответственных применений // Современные технологии автоматизации. 2007. № 4. С. 68—71.
8. **Open NAND Flash Interface Specification.** URL: www.onfi.org
9. **Ратайчук И. А., Шульгин В. И.** Исследование помехоустойчивости цифрового канала связи с использованием недвоичных кодов Рида—Соломона // Радиоэлектронные и компьютерные системы. 2011. № 1.
10. **Скляр Б.** Цифровая связь. Теоретические основы и практическое применение. М.: Изд. дом "Вильямс", 2003. 1106 с.
11. **Forward error correction for high bit-rate DWDM submarine systems.** ITU-T Recommendation G.975.1. URL: <https://www.itu.int/rec/T-REC-G.975.1/en>
12. **Odenwalder J. P.** Error Control Coding Handbook. Linkabit Corporation, 1976. 213 с.
13. **Jordanova L., Laskov L., Dobrev D.** Influence of BCH and LDPC Code Parameters on the BER Characteristic of Satellite DVB Channels // Engineering, Technology&Applied Science Research. 2014. N. 1.
14. **Hafmann W., Pless V.** Fundamentals of Error-correcting Codes. Cambridge University press, 2003.
15. **Галлагер Р.** Коды с малой плотностью проверок на четность. М.: Мир, 1966.
16. **Ryan W. E., Lin S.** Channel Codes Classical. Cambridge University press, 2009.
17. **Мальцев Г. Н., Джумков В. В.** Обобщенная модель дискретного канала передачи информации с группированием ошибок // Информационно-управляющие системы. 2013. № 1.
18. **Певнев В. Я., Пуранов М. В.** Экспериментальное исследование моделей групповых ошибок в канале связи. Харьков: Изд. ХНУВД, 2006.
19. **Золотарев В. В., Овечкин Г. В., Овечкин П. В.** Исследование недвоичных многопороговых декодеров в каналах связи с пакетизируемыми ошибками. М.: Изд. института космических исследований, 2009.
20. **Фрейман В. И., Пирожков А. П.** Исследование моделей дискретных каналов с памятью в рамках лабораторного практикума, построенного на базе пакета MatLab // Электротехника, информационные технологии, системы управления. 2013. № 7.

Errr-Correcting Capability Estimating for RS and BCH Codes

This paper proposes the approach of estimating of error-correcting capability by means of changed Binary Symmetric Channel (BSC). It's known that Reed-Solomon codes are more powerful than any codes with the same code-rate and data length. But to prove it there is a special BSC with packet error. Actually it's shown that RS codes have advantages both in correcting-level capability and in apparatus resources.

Keywords: RS, BCH, correcting level capability, BSC, Galois field

References

1. **Blahut R.** *Teoriya i praktika kodov, kontrolirujushih oshibki.* Moscow: Mir, 1986, 576 p.
2. **Morelos-Saragosa R.** *Iskusstvo pomehoustojchivogo kodirovaniya. Metodi algoritmi primeneniya.* Moscow: Tehnosfera, 2005, 320 p.
3. **Gallager R.** *Theoria informacii I nadejnaja svjaz.* Moscow: Sovetskoe radio, 1974, 720 p.
4. **What** Types of ECC Should Be Used on Flash Memory? Application note, Spansion, 2011, URL: [//www.spansion.com/Support/Application%20Notes/Types_of_ECC_Used_on_Flash_AN.pdf](http://www.spansion.com/Support/Application%20Notes/Types_of_ECC_Used_on_Flash_AN.pdf)
5. **Berezjuk N. T.** *Kodirovanie informacii. Dvoichnie kodi.* Kharkov: Visha shkola, 1978, 252 p.
6. **Peterson W., Weldon A.** *Kodi ispravljajushie oshibki.* Moscow: Mir, 1976, 580 p.
7. **Baishev A.** Tverdotelnye diski — nadejnoe reshenije otvetstvennih primenenij, *Sovremennije tehnologii avtomatizacii*, 2007, no. 4, pp. 68–71.
8. **Open** NAND Flash Interface Specification. URL: www.onfi.org
9. **Ratajchuk I. A., Shulgin V. I.** Issledovanie pomehoustojchivosti cifrovogo kanala svjazi s ispolzovaniem nedvoichnih kodov Reed—Solomon, *Radioelektronnye komputernije sistemi*, 2011, no. 1.
10. **Sklar B.** *Cifirovaja svjaz. Teoriticheskie osnovy i prakticheskoe primeneniya.* Moscow: Williams, 2003, 1106 p.
11. **Forward** error correction for high bit-rate DWDM submarine systems. ITU-T Recommendation G.975.1. URL <https://www.itu.int/rec/T-REC-G.975.1/en>
12. **Odenwalder J. P.** *Error Control Coding Handbook.* Linkabit Corporation, 1976, 213 c.
13. **Jordanova L., Laskov L., Dobrev D.** Influence of BCH and LDPC Code Parameters on the BER Characteristic of Satellite DVB Channels, *Engineering, Technology & Applied Science Research*. 2014, no. 1.
14. **Hafmann W., Pless V.** *Fundamentals of Error-correcting Codes.* Cambridge University press, 2003.
15. **Gallager R.** *Kodi s maloj plotnostju proverok na chetnost.* Moscow: Mir, 1966.
16. **Ryan W. E., Lin S.** *Chahnel Codes Classical.* Cambridge University press, 2009.
17. **Malcev G. N., Djumkov V. V.** Obobshhennaja model diskretnogo kanala peredachi informacii s gruppировaniem oshibok, *Informacionno-upravljajushie sistemi*, 2013, no. 1.
18. **Pevnev V. J., Curanov M. V.** *Eksperimentalnoje issledovanie modelej gruppovih oshibok v kanale svjazi.* Kharkov: Izdatelstvo KNUVD, 2006.
19. **Zolotarjev V. V., Ovechkin G. V., Ovechkin P. V.** *Issledovanie nedvoichnih mnogoporogovih dekoderov v kanalah svjazi c pajetirujushimisja oshibkami.* Moscow: Institut kosmicheskikh issledovanij. 2009.
20. **Frejman V. I., Pirojkov A. P.** Issledovanie modelej diskretnih kanalov s pamatju v ramkah laboratornogo praktikuma, postroennogo na baze paketa MatLab. Perm, *Elektrotehnika, informacionne tehnologii, sistemi upravlenija*, 2013, no. 7.

УДК 621.391

С. С. Дворников, лаборант, e-mail: dvornikov_s_s@mail.ru,

С. В. Дворников, д-р техн. наук, проф., профессор кафедры, e-mail: practicsv@yandex.ru,
Военная академия связи имени С. М. Буденного, г. Санкт-Петербург

Формирование сигналов с непрерывной фазой для передачи дискретной информации

Рассмотрены вопросы формирования сигналов каналов управления с фазовой модуляцией. Представлены результаты, характеризующие компактность спектра описываемых сигналов. Даны предложения по практическому использованию результатов. Описаны дальнейшие этапы исследований.

Ключевые слова: частотная модуляция, фазовая модуляция, канал управления, спектр, спектрально-эффективные сигналы

Введение

Ограниченность частотного ресурса стимулирует разработчиков информационно-технических систем (ИТС) к поиску спектрально-эффективных сигнальных конструкций, используемых в интересах передачи информации. При этом к основным показателям, характеризующим тот или иной мо-

дуляционный формат, следует отнести помехоустойчивость и спектральную эффективность [1]. Данные показатели тесно взаимосвязаны между собой, поэтому сложно обеспечить высокие показатели по одному из них, без снижения требований по остальным. Между тем, широко известные модуляционные форматы, используемые для передачи

информации, в ряде случаев не устраивают как производителей ИТС, так потребителей, в том числе и специалистов, занимающихся вопросами практической эксплуатации прикладных информационных систем [2].

Следовательно, вопросы разработки новых видов спектрально эффективных сигналов остаются актуальными, в том числе и в базисах, традиционно отличающихся от получившего широкое распространение базиса Фурье [3, 4]. Между тем, по мнению зарубежных специалистов [5, 6], традиционные виды модуляции еще до конца не исчерпали свои возможности и в рамках революции цифровой обработки сигналов требуют дополнительного переосмысления. В связи с этим в настоящей статье рассматриваются подходы к формированию сигналов с плавным изменением фазы для передачи дискретной информации. По мнению авторов, это открывает новые возможности по формированию спектрально-эффективных конструкций для различных задач прикладных информационных систем и радиотехники.

Обоснование выбора исходной конструкции сигнала

Любой модуляционный формат должен быть ориентирован на ИТС, в которой он будет использован. В свою очередь, качество каналов ИТС определяется пропускной способностью, которая характеризует количество достоверно передаваемой информации в единицу времени [1]. Однако верхняя граница пропускной способности C (бит/с) определяется отношением средней мощности сигнала S к средней мощности шума N ($h^2 = S/N$) в канале ИТС и полосой его пропускания ΔF (Гц). В результате для ИТС передачи дискретной информации пропускная способность будет определяться в соответствии с теоремой Шеннона следующим выражением [7]:

$$C = \Delta F \log_2(1 + h^2). \quad (1)$$

Основание логарифма как раз и указывает на дуальный характер передаваемой информации (т.е. предполагает два состояния).

Поскольку мощность шума определяется полосой пропускания канала ИТС $N = N_0 \Delta F$, где N_0 — спектральная плотность мощности шума, а на практике используют сложные сигнальные конструкции, в которых одним символом переносится несколько битов информации, то более широкое распространения получила величина

$$h_0^2 = E_b/N_0, \quad (2)$$

где E_b — энергия, приходящаяся на бит. Выражение (2) характеризует энергетические характеристики ИТС.

В общем случае величина h_0^2 связана с h^2 следующим соотношением:

$$h_0^2 = h^2 \frac{\Delta F}{R}, \quad (3)$$

где R — скорость передачи информации.

В дальнейшем величину h_0^2 определим как отношение сигнал/шум (ОСШ).

Пропотенцировав формулу (1), с учетом (2) можно получить более удобное выражение для характеристики ИТС:

$$h_0^2 = \frac{\Delta F}{C} (2^{C/\Delta F} - 1). \quad (4)$$

Фактически выражение устанавливает зависимость отношения пропускной способности к доступной полосе от энергетической характеристики ИТС, графическая интерпретация которой представлена на рис. 1.

Очевидно, что количество информации и, соответственно, скорость передачи информации зависит от качества канала ИТС, определяемого значением ОСШ, т.е. чем выше h_0^2 , тем больше информации можно передавать в одной и той же полосе. Другими словами, чем хуже качество канала, тем ниже в нем скорость передачи информации.

Кривая, представленная на рис. 1, определяет область допустимых значений передачи информации для сигнальных конструкций. Поскольку величина R всегда меньше C , то значение $h_0^2 \approx -1,6$ дБ, называемое пределом Шеннона [7], определяет предельное качество канала ИТС, при котором $R \rightarrow 0$, т.е. передача информации становится невозможной.

В практических приложениях допустимо незначительное искажение передаваемой информации, характеризуемое таким показателем, как достоверность передачи информации, численно определяемая вероятностью ошибки на бит P_b , которая зависит от вида модуляции.

Так, для сигналов двоичной фазовой манипуляции (ФМн2)

$$P_b^{\text{ФМн2}} = Q(\sqrt{2} h_0^2). \quad (5)$$

А для сигналов двоичной частотной манипуляции (ЧМн2)

$$P_b^{\text{ЧМн2}} = Q(h_0^2), \quad (6)$$

где $Q(z) = \frac{1}{\sqrt{2\pi}} \int_z^\infty e^{-t^2/2} dt$ — гауссова функция [7].

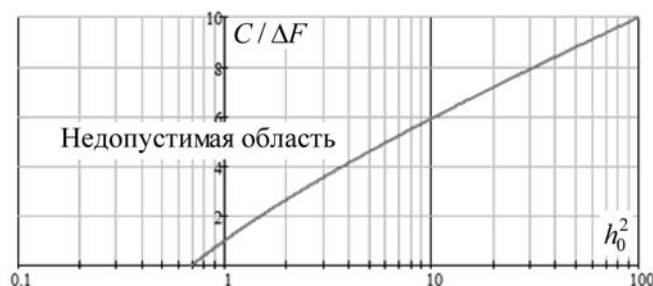


Рис. 1. Зависимость пропускной способности в полосе канала от ОСШ

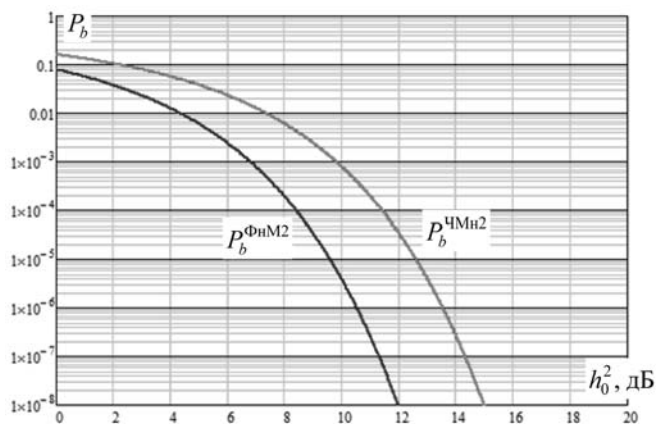


Рис. 2. Зависимость вероятности битовой ошибки в канале от ОСШ

Условие допустимости ошибки в канале ИТС позволяет на практике перейти от показателя пропускной способности C к скорости передачи R , при заданном уровне битовой ошибки P_b . На рис. 2 показаны зависимости вероятности ошибки на бит для сигналов ЧМн2 и ФМн2 от ОСШ.

Таким образом, рассмотренные показатели каналов ИТС позволяют определить критерии эффективности модуляционных форматов, к которым следует отнести спектральную и энергетическую эффективность.

Итак, энергетическая эффективность характеризует энергию, которую необходимо затратить для передачи информации с заданной достоверностью (т.е. вероятностью ошибки).

Спектральная эффективность характеризует полосу частот, необходимую для передачи информации с определенной скоростью.

Не менее важным показателем является устойчивость вида модуляции к различным видам деструктивных воздействий, обусловленных как естественными, так и преднамеренными помехами. Однако эта тематика требует отдельного рассмотрения.

Очевидно, что с увеличением позиционности модуляции возрастает скорость передачи информации, но при этом увеличивается битовая ошибка.

Это объясняется тем, что уменьшаются физические различия между параметрами сигналов, определяющими значение информационных символов, которые при низких значениях ОСШ приводят к ошибкам детектирования.

В настоящее время компромиссным решением для цифровых ИТС по критерию спектральной и энергетической эффективности являются модуляционные форматы на основе ФМн2 и ФМн4 [5]. Поэтому именно их целесообразно рассматривать в качестве исходных для формирования сигналов с плавным изменением фазы при передаче дискретной информации.

Формирование сигналов с плавным изменением фазы при передаче дискретной информации

Анализ структуры сигналов ФМн2 и ФМн4 показывает, что они по своей сути являются фрагментами радиоимпульсов с различными значениями начальной фазы. Фрагменты соответствуют информационной комбинации символов [4], т. е. их спектры будут иметь значительное число компонентов, не содержащих информационное наполнение и ухудшающих частотные характеристики каналов при их совместной организации в рамках единой ИТС.

В то же время сигналы фазовой модуляции, т.е. сигналы с плавным изменением фазы, лишены указанных недостатков [5]. В связи с этим рассмотрим более подробно синтез фазомодулированных сигналов (ФМ).

В общем случае модуляцией называется изменение по закону модулирующего сигнала (информационного сообщения) во времени одного из параметров гармонического колебания [7]:

$$s(t) = U_m \cos(\varphi(t)), \quad (7)$$

где U_m — амплитудный параметр несущего колебания; $\varphi(t)$ — фазовый (угловой) параметр несущего колебания.

Так, в отсутствие модуляции мгновенная фаза $\varphi(t)$ изменяется с постоянной скоростью $\omega_H = 2\pi f_H$ и является линейной функцией времени:

$$\varphi(t) = \omega_H t + \varphi_0. \quad (8)$$

Именно процесс модуляции полной фазы $\varphi(t)$ (учитывается начальное значение фазы φ_0 в (8)) модулирующим сигналом $u_\Omega(t)$ за счет изменения мгновенной фазы называется *фазовой модуляцией* [5, 7].

Важным моментом является то, что при модуляции линейность изменения $\varphi(t)$ нарушается, и в каждый момент времени t скорость изменения мгновенной фазы будет определяться мгновенной частотой.

В работе [6] обосновано, что мгновенная фаза и мгновенная частота связаны между собой следующими соотношениями:

$$\omega(t) = \frac{d}{dt} \varphi(t); \quad \varphi(t) = \int_0^t \omega(t) dt. \quad (9)$$

Из выражений (9) следует, что изменение фазы колебания по закону $\varphi(t)$ вызывает изменение мгновенной частоты по закону производной от фазы, а изменение мгновенной частоты по закону $\omega(t)$ приводит к изменению фазы по закону интеграла от мгновенной частоты.

Таким образом, сигнал ФМ получается путем приращения начальной фазы $\Delta\varphi(t)$ колебания (8), пропорционального модулирующему сигналу $u_\Omega(t)$. В результате полная фаза $\varphi(t)$ принимает вид:

$$\varphi(t) = \omega_H t + \varphi_0 + \Delta\varphi(t) = \omega_H t + \varphi_0 + \alpha u_\Omega(t). \quad (10)$$

Если модулирующий сигнал является гармоническим колебанием, то с увеличением его значения полная фаза фазомодулированного колебания

будет возрастать быстрее, чем по линейному закону. В то же время при уменьшении значения модулирующего сигнала происходит спад скорости роста $\varphi(t)$ во времени. В [7] максимальное отклонение фазы $\Delta\varphi(t)$ от первоначального ее значения определено как девиация частоты.

Аналитически процесс формирования сигналов ФМ можно представить, подставив значение (10) в выражение (7):

$$s_{\text{ФМ}}(t) = U_m \cos[\omega_{\text{н}}t + \varphi_0 + \alpha u_{\Omega}(t)]. \quad (11)$$

В (11) величина $\alpha u_{\Omega}(t) = \Delta\varphi(t)$ как раз и есть девиация, т.е. отклонение фазы модулированного сигнала от линейного закона, согласно которому первоначально происходило изменение несущего колебания.

В работе [5] значение девиации определено как

$$\Delta\varphi(t)_{\text{max}} = \alpha |u_{\Omega}(t)_{\text{max}}| = m_{\text{ФМ}}. \quad (12)$$

В выражении (12) значение $m_{\text{ФМ}}$, характеризующее максимальное отклонение фазы, называется индексом фазовой модуляции.

Важным моментом является то, что если модуляция осуществляется гармоническим колебанием (тональная модуляция) $u_{\Omega}(t) = U_{\Omega} \cos(\Omega t)$ (где Ω — частота модулирующего сигнала), то аналитическую запись такого сигнала можно представить как

$$s_{\text{ФМ}}(t) = U_m \cos[\omega_{\text{н}}t + \varphi_0 + \alpha U_{\Omega} \cos(\Omega t)] = U_m \cos[\omega_{\text{н}}t + \varphi_0 + m_{\text{ФМ}} \cos(\Omega t)]. \quad (13)$$

Величина $m_{\text{ФМ}} = \alpha U_{\Omega}$ в (13) пропорциональна амплитуде модулирующего колебания.

Анализ аналитической модели сигнала ФМ (11) показывает, что она является обобщающей по отношению к сигналам ФМн, в которых в качестве модулирующего сигнала $u_{\Omega}(t)$ выступает некоторая дискретная функция

$$I(t) = \begin{cases} \pi, & \text{если информационный символ "1" равен 1,} \\ 0, & \text{если информационный символ "0" равен 0.} \end{cases} \quad (14)$$

Смена значений в (14) определяется скоростью передачи информации.

Следовательно, подставив (14) в (11), получим сигнал, дискретность которого определяет скорость изменения параметра $I(t)$.

Между тем, для исключения дискретных разрывов модулированного колебания $s_{\text{ФМ}}(t)$ предлагается дискретной функцией $I(t)$ модулировать модулирующее колебание $u_{\Omega}(t)$, в результате чего получим

$$\hat{u}_{\Omega}(t) = U_{\Omega} \cos(\Omega t + I(t)). \quad (15)$$

И лишь затем модулирующим сигналом (15) модулировать искомое колебание (знак $\hat{}$ указывает на дискретный характер модулирующего колебания). Результирующий сигнал целесообразно представить в терминах выражения (13):

$$s_{\text{ФМ}}(t) = U_m \cos[\omega_{\text{н}}t + \varphi_0 + \hat{u}_{\Omega}(t)] = U_m \cos[\omega_{\text{н}}t + \varphi_0 + \hat{m}_{\text{ФМ}} \cos(\Omega t)]. \quad (16)$$

В результате указанных процедур получаем сигнал ФМ, в котором индекс $m_{\text{ФМ}}$ будет без разрывности изменяться по закону $\cos(\Omega t)$ в интервале от 0 до π .

Таким образом, нам удалось сформировать непрерывное колебание, обеспечивающее перенос дискретной информации. Для удобства определим полученный указанным образом сигнал как фазодискретное непрерывное (ФДН) колебание.

Анализ спектральной эффективности фазодискретных непрерывных сигналов

В интересах практического использования сигналов ФДН целесообразно оценить их по рассмотренным выше параметрам, в частности, оценить их спектральную эффективность. Для этого было проведено имитационное моделирование.

В частности, был определен информационный сигнал $I(t)$ в виде импульсной последовательности (рис. 3).

Здесь и далее на рисунках указано не абсолютное время, а число дискретных отсчетов.

Затем на его основе в соответствии с выражениями (13), (14) и (16) сформировали сигналы ФМн2 и ФДН2 с одинаковыми несущими частотами. Предварительно для формирования сигнала ФДН2 был синтезирован модулирующий сигнал в соответствии с выражением (15). Синтез осуществлялся из расчета один период на длительности информационного символа.

Так, на рис. 4 показаны: информационный сигнал $I(t)$ на длительности трех информационных символов; сигнал $\hat{u}_{\Omega}(t)$, модулирующий результирующий сигнал ФДН2; сформированный сигнал ФМн2 $s_{\text{ФМн2}}(t)$; сформированный сигнал $s_{\text{ФДН2}}(t)$. Указаны также разрывы фазы в местах смены информационных символов.

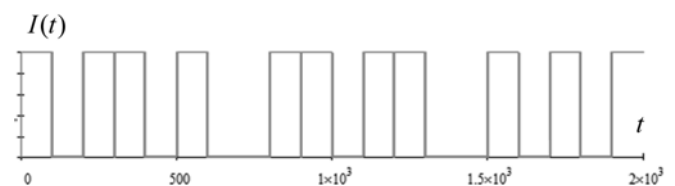


Рис. 3. Последовательность информационных импульсов

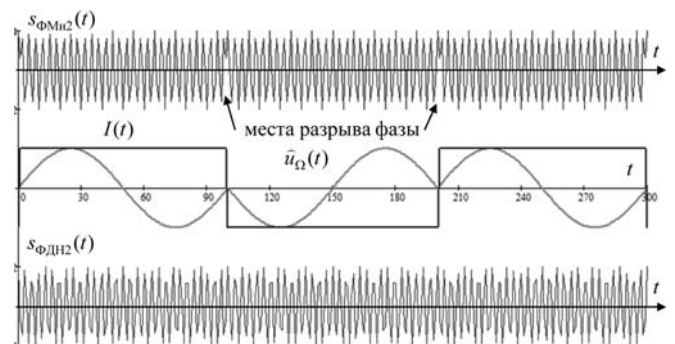


Рис. 4. Сигналы, синтезированные в процессе эксперимента

Заключение

Результаты теоретического исследования и данные практического эксперимента позволяют заключить, что разработанный подход отвечает требованиям формирования сигналов с плавным изменением фазы, обеспечивающих передачу дискретной информации.

В настоящей работе проанализирована возможность формирования сигналов для двухпозиционной фазовой манипуляции. Однако разработанный подход допускает использование в качестве модулирующего сигнала и других сигнальных конструкций, в частности, ФМн4, сигналы с минимальным частотным сдвигом, а также более сложные сигналы квадратурно-амплитудной модуляции [8] и др.

Дальнейшие исследования авторы связывают с вопросами разработки эффективных методов демодуляции сигналов ФДН [9]. Кроме того, характер спектра сигналов ФДН позволяет предположить, что гармоника основной частоты не является информационной, и это позволяет ее исключить в ходе дальнейшей обработки. Таким образом, по аналогии с амплитудной модуляцией открывается возможность синтеза ФДН сигналов с балансной модуляцией, с частично подавленной несущей и с одной боковой полосой. В то же время ярко выраженный пик в частотной области может быть использован для обнаружения излучений [10]. Но все эти вопросы требуют дальнейшей тщательной проработки.

Список литературы

1. Меркушева А. В. Элементы прикладной теории время-частотных преобразований и обработки нестационарных сигналов в информационно-измерительных системах // Информационные технологии. 2012. № 3. С. 1–32.
2. Мандрикова О. В., Полозов Ю. А. Автоматизированный способ обработки сигналов со сложной структурой // Информационные технологии. 2008. № 12. С. 15–19.
3. Табаков Ю. Г. Модель и алгоритм обработки низкочастотного сигнала для тренажера на основе вейвлет-преобразований // Информационные технологии. 2015. Т. 21. № 6. С. 464–468.
4. Дворников С. В., Дворников С. С., Спирин А. М. Синтез манипулированных сигналов на основе вейвлет-функций // Информационные технологии. 2013. № 12. С. 52–55.
5. Hong L. and Ho K. C. Classification of BPSK and QPSK signals with unknown signal level using the Bayes technique // Proc. IEEE ISCAS. 2003. P. IV.1–IV.4.
6. Panagiotou P., Anastasopoulos A., Polydoros A. Likelihood ratio tests for modulation classification // Proc. IEEE MILCOM. 2000. P. 670–674.
7. Скляр Б. Цифровая связь. Теоретические основы и практическое применение. Изд. 2-е испр.: Пер. с англ. М.: ИД "Вильямс", 2004. 1104 с.
8. Дворников С. В., Пшеничников А. В., Русин А. А., Дворников А. С. Повышение помехоустойчивости сигналов кам-16 с трансформированными созвездиями // Вопросы радиоэлектроники. Сер. "Техника телевидения". 2014. № 2. С. 51–56.
9. Дворников С. В., Устинов А. А., Пшеничников А. В., Борисов В. В., Москалец А. Г., Бурькин Д. А. Демодуляция сигналов ОФТ на основе адаптивного порога // Вопросы радиоэлектроники. Сер. "Техника телевидения". 2013. № 2. С. 90–97.
10. Дворников С. В. Метод обнаружения сигналов диапазона ВЧ на основе двухэтапного алгоритма принятия решения // Научное приборостроение. 2005. Т. 15, № 3. С. 114–119.

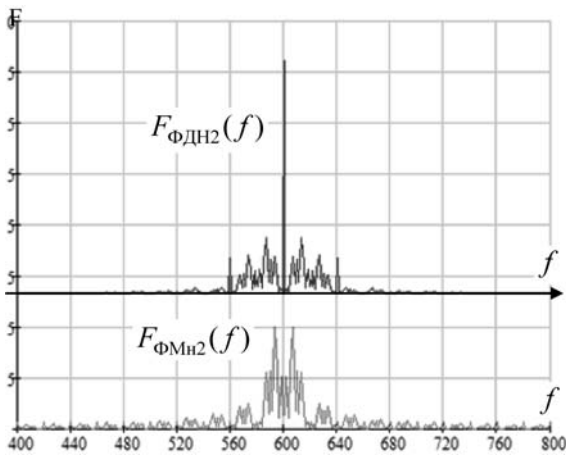


Рис. 5. Спектры исследуемых сигналов

С учетом того, что целью исследования являлась оценка спектральной эффективности сигналов ФДН по отношению к ФМн, были синтезированы спектры указанных сигналов, представленные на рис. 5. Для лучшей наглядности спектры разнесены по оси ординат, а значения частоты даны в спектральных отсчетах преобразования Фурье.

Спектральная эффективность оценивалась как отношение полос, занимаемых спектрами исследуемых сигналов:

$$\gamma^A = \frac{\Delta F_{\text{ФМн}2}^A}{\Delta F_{\text{ФДн}2}^A}, \quad (17)$$

где A — число (в процентах), обрабатываемых (учитываемых) частотных составляющих спектра.

Таким образом, значение $\gamma^{95\%} = \frac{\Delta F_{\text{ФМн}2}^{95\%}}{\Delta F_{\text{ФДн}2}^{95\%}} = 2,4$

указывает, что для учета 95 % частотных составляющих спектра полоса частот для сигнала ФМн2 должна быть в 2,4 раза шире аналогичной полосы для сигнала ФДН2.

В таблице указаны значения показателя, определяемого формулой (17), полученные в ходе исследования.

Результаты спектральной эффективности сигнала ФДН2 по отношению к сигналу ФМн2

$A, \%$	99,5	98	97	96	95	94	93	92	91	90	80
γ^A	10	4,8	3,7	3	2,4	2,1	1,8	1,7	1,3	1,25	1

Анализ результатов таблицы показывает, что по уровню 80 % учета спектральных составляющих исследуемые сигналы занимают равные полосы частот. Однако с увеличением показателя A спектральная эффективность ФДН2 существенно возрастает по отношению к ФМн2.

Forming Signals with Continuous Phase for Broadcast Discrete Information

The limitation of frequency range requires the development of new types of modulated signals. To determine the effectiveness of new types of signals are invited to identify the noise immunity of the signals and their spectral efficiency as a criteria. Protection against interference is defined by the signal / noise ratio, for predetermined value of the bit error rate. Under spectral efficiency we mean the bandwidth required to reliably transmit the signal at a predetermined rate.

Currently, the most effective criterion for noise immunity is phase shift keying. Therefore signals of PSK were taken as the basis for the study. The analysis showed that the synthesis signals of phase shift keying can be representing as a sequence of radio pulses. The RF pulse duration is determined by changed of the phase discontinuity information symbol. Consequently, the spectrums of phase-shift keyed signals are similar to the spectrum of a sequence of pulses. Provided that the sequence and signals have the same data rate.

The results allow concluding that the phase-shift keyed signals have a good the noise immunity, but they do not satisfy in terms of spectral efficiency. This fact explains the phase discontinuity when changing the information symbol.

In order to improve the spectral efficiency of the phase-shifted signals are invited to perform the following sequence of actions.

In the first phase is created a signal which is a modulating signal. The modulating signal is a signal whose phase signal is manipulated by the law of the information signal. Thus, the minimum length of the information signal is equal to one period of the frequency of the modulating oscillations. The result is a phase-shift keyed signal of low frequency.

The second stage of algorithm produces a resultant modulated signal. For this purpose, phase of high frequency signal formed by the law by which changes phase-shift keyed signal of low frequency.

The result is obtained the signal of phase-modulation, whose spectrum occupies a smaller bandwidth compared with the corresponding speed transmission PSK signal. At the same time, the resultant signal of phase-modulation allows you to transfer discrete information. This fact makes it interesting for digital information transmission systems via radio.

The experiment showed that a new class of signals is has 2–3 times better spectral efficiency level of conservation in the transmission band of 85–95 % of the energy. A new class is defined as signals with signals of phase-modulated continuous-discrete.

Keyword: phase manipulation, discrete information, spectral efficiency signals, phase manipulation with smoothly phase changes, control channels

References

1. **Merkusheva A. V.** Jelementy prikladnoj teorii vremjachastotnyh preobrazovanij i obrabotki nestacionarnyh signalov v informacionno-izmeritel'nyh sistemah, *Informacionnye tehnologii*, 2012, no. 3, pp. 1–32.
(Merkusheva A. V. Elements of applied theory of time, frequency conversion and processing of non-stationary signals in the data-measuring systems. Information Technology. 2012. N. 3. P. 1–32.)
2. **Mandrikova O. V., Polozov Ju. A.** Avtomatizirovannyj sposob obrabotki signalov so slozhnoj strukturoj, *Informacionnye tehnologii*, 2008, no. 12, pp. 15–19.
(Mandrikova O. V., Polozov Y. A. An automated method for processing signals with a complex structure. Information Technology. 2008. N. 12. P. 15–19.)
3. **Tabakov Ju. G.** Model' i algoritm obrabotki nizkochastotnogo signala dlja trenazhera na osnove vejvlet-preobrazovanij, *Informacionnye tehnologii*. 2015. vol. 21. no. 6. pp. 464–468.
(Tabakov Y. G. Model and algorithm of the baseband signal processing for the simulator based on wavelet transforms. Information Technology. 2015. T. 21. no 6. S. 464–468.)
4. **Dvornikov S. V., Dvornikov S. S., Spirin A. M.** Sintez manipulirovannyh signalov na osnove vejvlet-funkcij, *Informacionnye tehnologii*, 2013, no. 12, pp. 52–55.
(Dvornikov S. V., Dvornikov S. S., Spirin A. M. Synthesis manipulates-bathrooms signals based on wavelet functions. Information Technology. 2013. N. 12. P. 52–55.)
5. **Hong L. and Ho K. C.** Classification of BPSK and QPSK signals with unknown signal level using the Bayes technique. *Proc. IEEE ISCAS*, 2003, pp. IV.1–IV.4.
6. **Panagiotou P., Anastasopoulos A., Polydoros A.** Likelihood ratio tests for modulation classification. *Proc. IEEE MILCOM*, 2000, pp. 670–674.
7. **Sklyar B.** Cifrovaja svjaz'. Teoreticheskie osnovy i praktičeskoe primenenie. Izd. 2-e ispr.: Per. s angl. Moscow: ID "Vil'jams", 2004. 1104 p.
(B. Sklar. Digital communication. Theoretical bases and practical application. Ed. 2nd Corr.: Trans. from English. Moscow: Publishing House "Williams", 2004. 1104 p.)
8. **Dvornikov S. V., Pshenichnikov A. V., Rusin A. A., Dvornikov A. S.** Povyshenie pomehoustojčivosti signalov kam-16 s transformirovannyimi sovezdijami, *Voprosy radioelektroniki. Serija: Tehnika teledenija*, 2014, no. 2, pp. 51–56.
(Dvornikov S. V., Pshenichnikov A. V., Rusin A. A., Dvornikov A. S. In elevated noise immunity of the signals with 16QAM constellations transformed. Questions electronics. Series: Television Technology. 2014. N. 2. P. 51–56.)
9. **Dvornikov S. V., Ustinov A. A., Pshenichnikov A. V., Borisov V. V., Moskalec A. G., Burykin D. A.** Demoduljacija signalov OFT na osnove adaptivnogo poroga. *Voprosy radioelektroniki. Serija: Tehnika teledenija*, 2013, no. 2, pp. 90–97.
(Dvornikov S. V., Ustinov A. A., Pshenichnikov A. V., Borisov V. V., Moscalets A. G., Burykin D. A. Signal Demodulation OFT based on the ADAP-tive threshold. Questions electronics. Series: Television Technology. 2013. N. 2. P. 90–97.)
10. **Dvornikov S. V.** Metod obnaruzhenija signalov diapazona VCh na osnove dvuhjetapnogo algoritma prinjatija reshenija, *Nauchnoe priborostroenie*, 2005, vol. 15, no. 3, pp. 114–119.
(Dvornikov S. V. Detection Method of HF signals on the basis of a two-stage decision algorithm. Scientific instrument. 2005. Vol. 15, N. 3. P. 114–119.)

СИСТЕМЫ АВТОМАТИЗИРОВАННОГО ПРОЕКТИРОВАНИЯ CAD-SYSTEMS

УДК 338.45, 004.415.2

Ю. И. Петров, канд. экон. наук, доц., e-mail: mailtoyuri.petrov@gmail.com,
Московский технологический университет

Система поддержки принятия решений по созданию конкурентоспособной технически сложной продукции

Рассматриваются вопросы проектирования и реализации системы поддержки принятия решений по созданию конкурентоспособной технически сложной продукции. Представлена архитектура программного комплекса, а также приведены иллюстрации работы разработанного программного обеспечения.

Ключевые слова: информационная система, система поддержки принятия решений, СППР, программное обеспечение, Delphi, СУБД, Firebird, конкурентоспособность продукции, разработка сложных промышленных изделий

Введение

Несмотря на различные оценки экспертов [1] основой экономики государств с высоким уровнем технологического развития по-прежнему остается и, по-видимому, будет оставаться промышленный сектор. При этом конкурентоспособность государства и конкретной отрасли национальной экономики определяется конкурентоспособностью предприятий, которая, в свою очередь, существенно зависит от конкурентоспособности выпускаемой ими продукции [2]. Вопросы обеспечения конкурентоспособности продукции отечественных промышленных предприятий приобретают еще большую актуальность в связи с введенными антироссийскими санкциями [3] и необходимостью освоения новых рынков на фоне увеличения влияния стран БРИКС в мире [4].

Создание новой модели (ассортиментной единицы) продукции промышленного предприятия требует решения многих задач и выполнения множества работ, объединенных соответствующим бизнес-процессом. В работах [5–7] был предложен ряд бизнес-процессов, позволяющих оптимально подойти к вопросу разработки продукции с учетом специфики конкретных предприятий, однако представленные в его рамках оптимизационные модели (на основе задачи о "многомерном ранце" с булевыми переменными) требуют достаточно трудоемких расчетов. В связи с этим целесообразно создание специального программного обеспечения (ПО), функционирующего с учетом соответствующей методологии и облегчающего специалисту принятие решений относительно проектирования ассортимента продукции.

В настоящей статье приведена архитектура такого программного комплекса, а также проиллюстрированы примеры его работы.

Исходные данные

Сложное промышленное изделие можно представить в виде модели технически сложной продукции, производство которой возможно в различных вариантах исполнения, определяемых составами используемых комплектующих; при этом внешний вид, функциональные возможности и технические характеристики изделия являются показателями его конкурентоспособности [5, 8]. В зависимости от стратегии производителя (акцент на качество, снижение цены, особенности/характеристики изделия и т. д.) состав показателей конкурентоспособности продукции может быть различен. В целом, предлагается подразделять их на две группы:

- необходимые (соблюдение заданных значений которых позволит выйти на рынок — обеспечит соответствие требованиям стандартов и норм — и гарантировать сопоставимость продукции с существующими аналогами);
- достаточные (значения которых необходимо максимизировать или минимизировать для эффективного позиционирования изделия на рынке).

Выявление необходимых и достаточных показателей конкурентоспособности включает три этапа.

Во-первых, необходимо провести исследование конкурентов, рынка и особенностей продукции. В частности, на данном этапе следует выделить характерные свойства продукции, например, такие как оценка предпочтительности (отражает влияние на выбор потребителей), себестоимость воплощения в единице разрабатываемого изделия, энерго-

потребление и т. д. Значения таких параметров свойств аддитивны относительно конфигурации изделия, что позволяет контролировать и регламентировать значения его характеристик, например:

- ограничение на себестоимость позволяет учесть особенности ценообразования, в том числе при позиционировании изделия на конкретных ценовых нишах рынка;
- ограничение на энергопотребление позволяет обеспечить соответствие классу энергопотребления или заданное время автономной работы.

Вторым этапом будет являться составление экономико-математической модели, где специалист самостоятельно выделит ряд характеристик в качестве нормальных и достаточных для конкретного рынка. Первая группа сформирует ограничения, а вторая — целевую функцию оптимизационной задачи.

Заключительным третьим этапом является решение задачи, интерпретация результатов и выбор варианта исполнения изделия.

Функциональные и архитектурные особенности системы

Второй этап является вычислительно сложным и требует от специалиста значительных усилий на решение задачи. При этом его деятельность может быть автоматизирована за счет создания специализированного ПО (системы поддержки принятия решений — СППР), способствующего формированию, решению и интерпретации результатов решения задачи, в связи с чем было принято решение о необходимости его разработки.

Создаваемое ПО должно обеспечивать:

- централизованное накопление, хранение, обработку и анализ данных, необходимых для разработки сложного промышленного изделия;
- автоматизированное решение оптимизационной задачи определения оптимального состава свойств на основе вводимых данных;
- формирование отчетных документов, включающих полученную оптимальную конфигурацию изделия и оценку его положения на рынке.

Под необходимыми для разработки изделия данными понимаются:

- информация об изделиях-конкурентах, их свойствах и характеристиках;

- сведения о комплектующих, используемых в изделиях-конкурентах и доступных производителю;
- технологические требования и технологические ограничения;
- цели и интересы производителя промышленных изделий;
- предпочтения потребителей или интересы заказчика изделия.

Создаваемое ПО должно иметь:

- базу данных (БД), хранящую: данные, необходимые для решения задачи разработки сложного промышленного изделия; результаты решения задачи, свойства и характеристики разработанного изделия; информацию о пользователях, имеющих доступ к системе и их полномочиях;
- графический интерфейс пользователя, позволяющий специалисту эффективно взаимодействовать с системой и получать аналитическую информацию в виде специально оформленных отчетов.

На основании предъявляемых требований была принята архитектура СППР "Goods Builder", компоненты которой представлены на рис. 1.

Предложенная архитектура, основанная на работах [9, 10] и источниках [11, 12], хорошо зарекомендовала себя во время эксплуатации соответствующих программных средств. В качестве системы управления базами данных (СУБД), применяемой для создания и управления данными СППР, используется Firebird [13]. Данная СУБД может использоваться на разнообразных серверных и клиентских платформах Windows и Linux, а также платформах UNIX. Firebird является СУБД промышленного применения, ее возможности соответствуют стандартам SQL, кроме того, имеются некоторые мощные расширения языка процедурного программирования конкретного производителя. Для реализации клиентской части был использован язык программирования Delphi XE2 [14].

Программное обеспечение "Goods Builder"

Рассмотрим иллюстрацию работы системы на наглядном тестовом примере по разработке (сборке) персонального компьютера.

При первом запуске системы появляется окно входа, где указываются параметры учетной записи — имя пользователя и пароль, а также настройки подключения к серверу БД (IP-адрес и порт компьютера) и путь к файлу базы данных (рис. 2).

После указания имени пользователя, пароля и настроек подключения происходит соединение с базой данных, и на экране отображается стартовое окно системы (рис. 3).



Рис. 1. Общая архитектура СППР "Goods Builder"

Специалисту предлагается два режима работы с системой:

- работа с базой данных системы ("База данных");
- решение оптимизационной задачи ("Модуль решения").

В режиме работы с базой данных пользователю доступны следующие операции:

- добавление информации о новых изделиях, редактирование записей о существующих и удаление записей об устаревших изделиях;
- управление множеством возможных свойств разрабатываемого или имеющегося изделия;
- добавление, редактирование и удаление групп свойств на основании их общих характеристик;
- управление набором характеристик групп, относящихся к конкретному свойству изделия.

В режиме работы с базой данных специалист может добавлять, изменять и удалять сведения об изделиях-конкурентах, а также их характеристиках (в том числе произвольных). На рис. 4 представлена страница "Изделия" редактора БД.

На рис. 5 представлены страницы "Группы" и "Характеристики" редактора базы данных.

На рис. 6 представлена страница "Свойства" редактора базы данных.

Для управления содержимым каждой категории в редакторе предусмотрен специальный интерфейс. Пример окна добавления/редактирования изделия приведен на рис. 7. Аналогично предоставляется возможность редактирования групп, свойств и характеристик изделия.

Модуль решения содержит страницы "Проекты" и "Решение" (рис. 8).

Для формирования задачи специалист должен выполнить три шага:

1. Добавить новый проект либо отредактировать существующий нажатием соответствующих кнопок на странице "Проекты".

2. В открывшемся окне (рис. 9) указать целевую функцию, а также одно или несколько ограничений на характеристики проектируемого изделия, с использованием кнопок в правой части группы "Ограничения".

3. Нажать кнопку "ОК" для завершения добавления/редактирования проекта и его выбора для дальнейшего решения задачи.

После выбора проекта будет доступна страница "Решение". Для решения задачи требуется нажать кнопку "Выполнить решение!". По оконча-

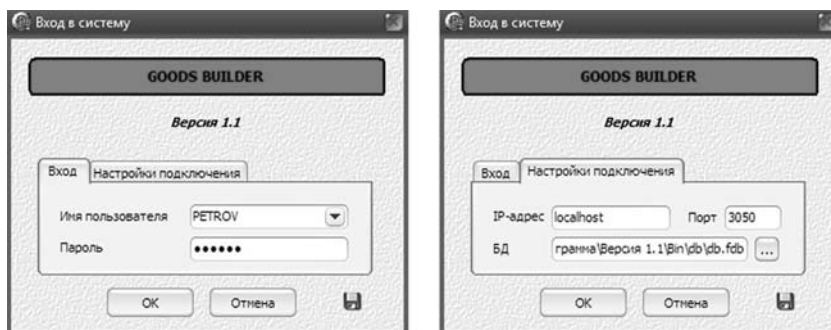


Рис. 2. Окно входа в систему

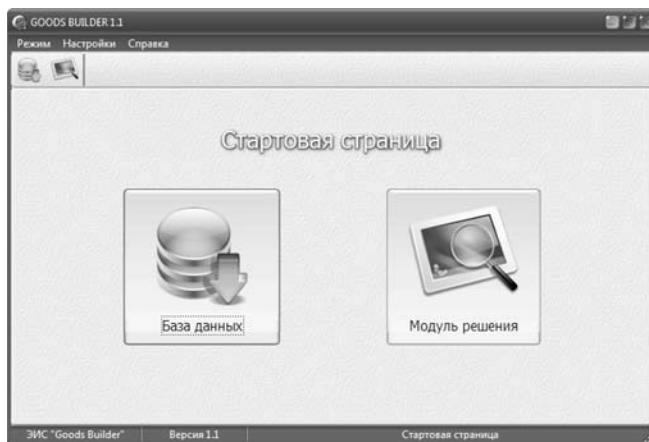


Рис. 3. Стартовая страница СППР "Goods Builder"

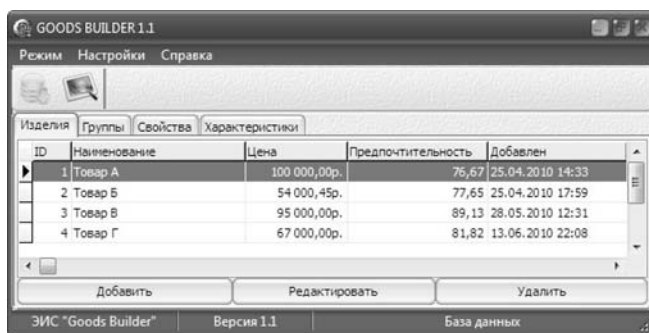


Рис. 4. Страница "Изделия" редактора базы данных

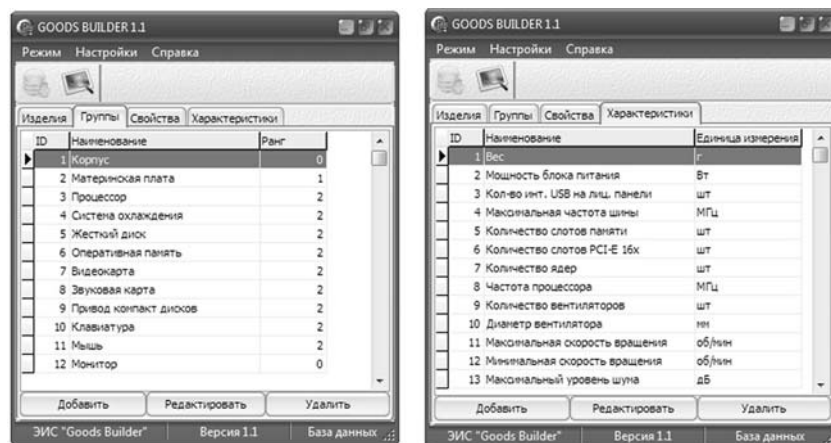


Рис. 5. Страницы "Группы" и "Характеристики" редактора базы данных

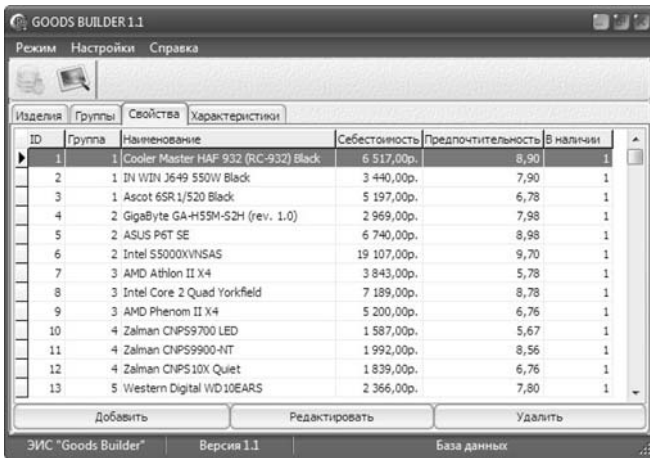


Рис. 6. Страница "Свойства" редактора базы данных

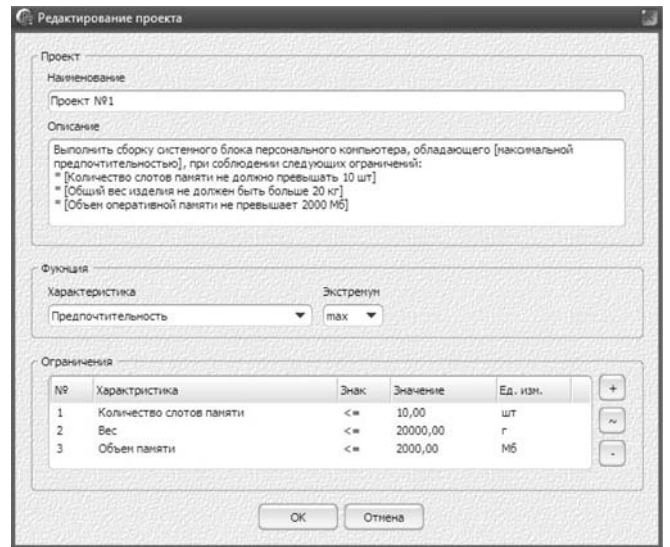


Рис. 9. Окно добавления/редактирования проекта

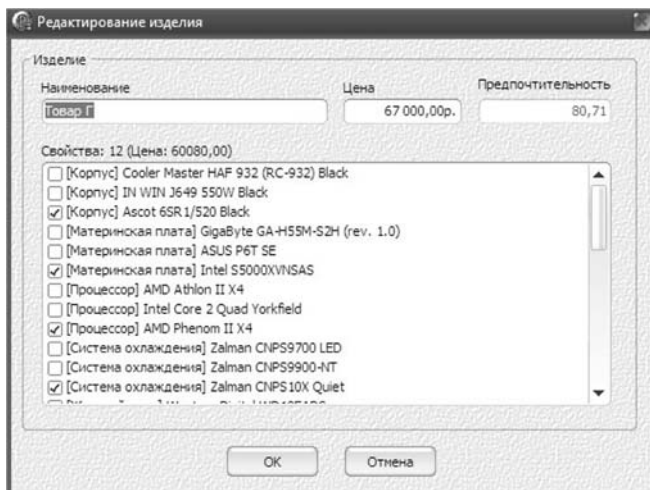


Рис. 7. Окно добавления/редактирования изделия

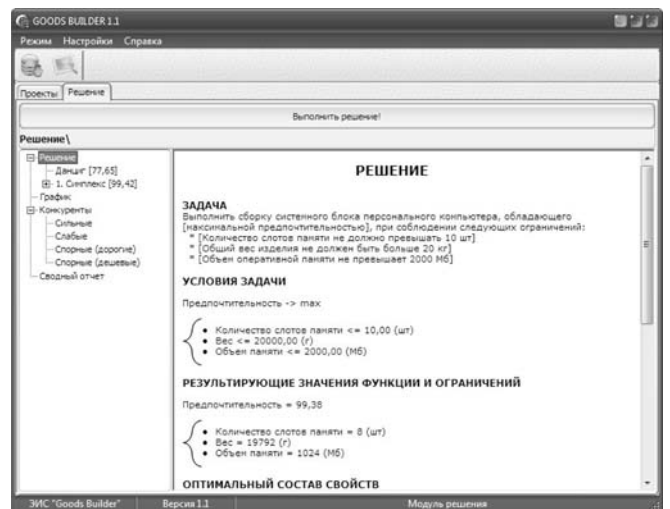


Рис. 10. Пример результатов решения задачи в СППР "Goods Builder"

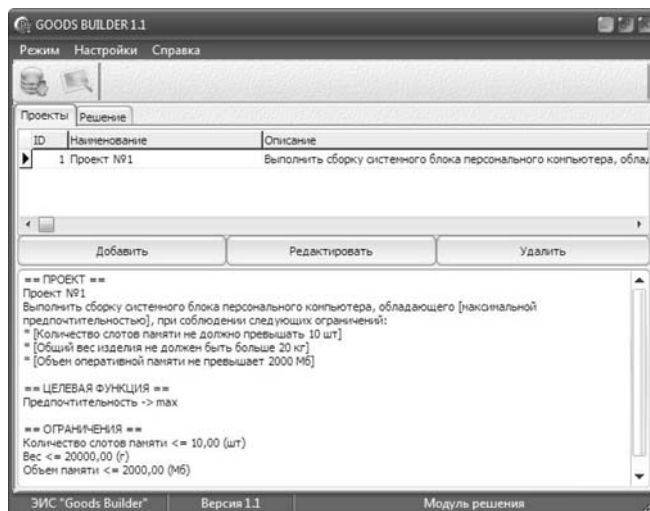


Рис. 8. Модуль решения СППР "Goods Builder"

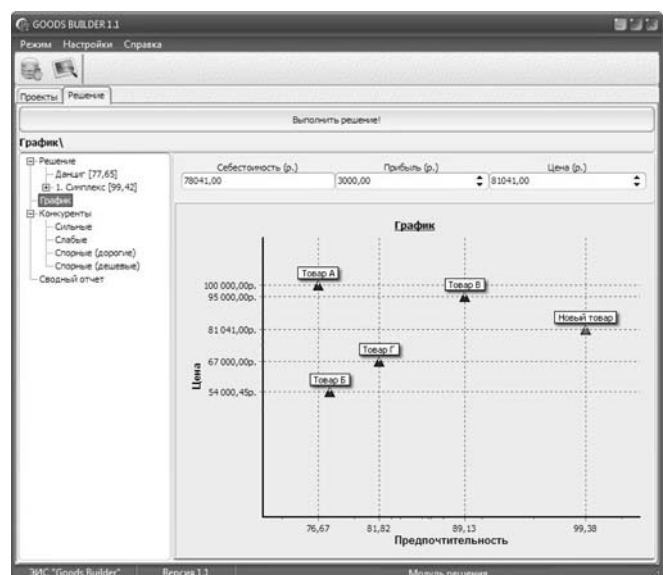


Рис. 11. Окно назначения цены полученного изделия

нии вычислений в дереве решения слева и окне описания справа будут отражены (рис. 10):

- результат решения с указанием оптимального состава свойств изделия;
- все стадии выполнения алгоритма с подробным ходом решения;
- график назначения цены полученного изделия (рис. 11);
- информация о конкурентах, их анализ и сравнительная характеристика с полученным изделием;
- сводный отчет о решении задачи;
- служебная информация системы, определяемая в ходе вычислений.

График, подобный представленному на рис. 11, позволяет наглядно сравнить разработанный товар с имеющимися в базе данных конкурентами. Назначение итоговой цены товара возможно следующими способами:

- перемещением красной отметки товара левой клавишей мыши;
- непосредственным указанием значения цены или прибыли в соответствующих полях в верхней части окна графика.

Система также позволяет сохранить и/или вывести результаты решения на печать — для этого достаточно выбрать пункт "Сводный отчет" в дереве решений.

Заключение

В статье были рассмотрены некоторые вопросы создания технически сложной продукции в современных условиях, показано, что разработка новых ассортиментных единиц является комплексным процессом, который сопровождается решением сложных экономико-математических моделей, что в свою очередь делает целесообразным разработку специальных информационных систем для специалистов (СППР).

Была рассмотрена архитектура такой СППР, включающая клиентскую и серверную части — комбинацию технологий Delphi и Firebird. Используя систему, специалист получает возможность комплексно работать с задачей разработки изделия, составлять проекты, условия задач, заполнять базу комплектующих и изделий-конкурентов, а также проводить анализ решений, имея возможность просмотреть результаты работы в электронном или печатном виде.

В дальнейшем функциональность программного обеспечения может быть расширена, например, за счет использования технологий Data Mining [15, 16].

Список литературы

1. **Что** ждет мировую промышленность в XXI веке. [Электронный ресурс] (дата обращения 15.07.2014 г.). URL: <http://www.interfax.ru/business/318506>, свободный.
2. **Петров Ю. И.** Адаптация отечественной продукции к требованиям мировых и региональных рынков // Актуальные вопросы экономических наук: Сб. матер. XXII Международной научно-практической конференции. В 2-х частях. Часть 2 / Под общ. ред. Ж. А. Мингалевой, С. С. Чернова. Новосибирск: Изд-во НГТУ, 2011. С. 202—209.
3. **Антироссийские санкции** — [Электронный ресурс] (дата обращения 22.06.2015 г.). URL: http://ruxpert.ru/Антироссийские_санкции, свободный.
4. **Расширение** влияния ШОС и БРИКС [Электронный ресурс] (дата обращения 15.07.2015 г.). URL: <http://www.youtube.com/watch?v=EaJXCnswJTQ>, свободный. — Загл. с экрана.
5. **Петров Ю. И.** Развитие бизнес-процессов создания конкурентоспособной технически сложной продукции машиностроительных предприятий: Дис. ... канд. экон. наук. Москва. 2012. 148 с.
6. **Самохин Ю. М., Петров Ю. И.** Проблемы развития бизнес-процессов создания конкурентоспособной продукции машиностроительных предприятий // Управление экономическими системами: электронный научный журнал. 2012. № 4 [Электронный ресурс]. URL: <http://www.uecs.ru/teoriya-upravleniya/item/1238-2012-04-09-08-10-34>.
7. **Петров Ю. И.** Современное состояние промышленности в РФ и необходимость освоения внутренних и внешних территориальных рынков // Известия высших учебных заведений. Проблемы полиграфии и издательского дела. 2011. № 6. С. 207—212.
8. **Петров Ю. И.** Особенности производства сложных промышленных изделий в современных условиях // Научные итоги 2011 года: достижения, проекты, гипотезы. Сб. матер. I Междунар. научно-практ. конф. В 2-х частях. Часть 2 / Под общ. ред. С. С. Чернова. Новосибирск: Изд-во НГТУ, 2011. С. 218—227.
9. **Петров Ю. И.** Архитектура и алгоритмическое обеспечение информационной системы нормоконтроля выпускных квалификационных работ // Программная инженерия. 2013. № 3. С. 26—32.
10. **Петров Ю. И., Карнаухов Ю. А.** Программное обеспечение "FK-Monitoring" как средство мониторинга бизнес-процессов // Информационные технологии. 2015. Т. 21, № 7. С. 503—510.
11. **Спинеллис Д., Гусиос Г.** Идеальная архитектура. Ведущие специалисты о красоте программных архитектур. СПб.: Символ-Плюс, 2010. 528 с.
12. **Форд Н., Найгард М.** 97 этюдов для архитекторов программных систем. СПб.: Символ-Плюс, 2010. 224 с.
13. **Firebird: The true open source database for Windows, Linux, Mac OS X and more.** — [Electronic resource] (дата обращения 15.07.2014 г.). URL: <http://www.firebirdsql.org>, свободный. — Загл. с экрана.
14. **Delphi XE2 Overview.** [Electronic resource] (дата обращения 23.07.2014 г.). URL: <http://edn.embarcadero.com/article/41593>, свободный.
15. **Использование** методов Data Mining в анализе деятельности предприятия. [Электронный ресурс] (дата обращения 12.05.2015 г.). URL: <http://cyberleninka.ru/article/n/ispolzovanie-metodov-data-mining-v-analize-deyatelnosti-predpriyatiya>, свободный. — Загл. с экрана.
16. **Петров И. П., Петров Ю. И., Петрова Н. В.** Применение технологий Data Mining в вопросах дифференциации образовательного процесса // Информатизация образования и науки. 2015, № 3 (27). С. 123—133.

Decision Support System for Development of Technically Complex Manufactured Goods

Despite different estimates of experts, industrial sector still remains and probably will be the basis of the economies of the states with the high level of technological development. Thus competitiveness of the state and specific branch of national economy is defined by competitiveness of the enterprises which, in turn, significantly depends on competitiveness of their goods. Problems of the goods competitiveness of the domestic industrial enterprises are still in big relevance in connection with the imposed anti-Russian sanctions, and need of new markets development according to increase in influence of the countries of BRICS in the world.

Creation of new model (assortment unit) of the complex industrial product requires the solution of many tasks and execution of a set of the operations integrated by the corresponding business process. In previous author's papers a business-process, that allows to approach optimum in a problem of product development and takes into account some of the enterprises specifics, was offered. However the optimization models provided in the business process (based on a Knapsack problem with boolean variables) are enough labor-consuming and leads to the idea of special software development (decision support system), that can be programmed according to the offered methodology and can facilitate making decisions on product design to the expert.

Article considers issues of design and implementation of the decision support system, that includes two components: server-side (Firebird Server) and client-side (Delphi desktop application). Using system, the specialist has an opportunity to work with a product development problem, to make projects, to fill conditions of problems, database of accessories and competitors' products, and also to carry out the analysis of the decisions, having opportunity to see results of work in an electronic or printed form.

Keywords: information system, analytical information system, software, business process, monitoring, publishing holding, OLAP, Delphi, PostgreSQL, product competitiveness, complex industrial products development.

References

1. **Что** здет мировују промышленност' в XXI веке. [Electronic resource], (data obrashheniya 15.07.2014 g.). URL: <http://www.interfax.ru/business/318506>, svobodnyj.
2. **Petrov Yu. I.** Adaptacija otechestvennoj produkcii k trebovanijam mirovyh i regional'nyh rynkov, *Aktual'nye voprosy jekonomicheskikh nauk: sbornik materialov II Mezhdunarodnoj nauchno-prakticheskoy konferencii: v 2-h chastjah. Chast' 2 / Pod obshh. red. Zh. A. Mingalevoj, S. S. Chernova.* Novosibirsk: Izdatel'stvo NGTU, 2011, pp. 202—209.
3. **Antirossijskie** sankcii. — [Electronic resource] (data obrashheniya 22.06.2015 g.). URL: http://ruXpert.ru/Antirossijskie_sankcii, svobodnyj.
4. **Rasshirenie** vlijaniya ShOS i BRIKS. [Electronic resource] (data obrashheniya 15.07.2015 g.). URL: <http://www.youtube.com/watch?v=EaJXCnswJTQ>, svobodnyj.
5. **Petrov Yu. I.** *Razvitie biznes processov sozdaniya konkurentosposobnoj tehnikeski slozhnoj produkcii mashinostroitel'nyh predpriyatij:* Dis. ... kand. jekon. nauk. Moskva. 2012. 148 p.
6. **Samohin Yu. M., Petrov Yu. I.** Problemy razvitiya biznes-processov sozdaniya konkurentosposobnoj produkcii mashinostroitel'nyh predpriyatij, *Upravlenie jekonomicheskimi sistemami: jelektronnyj nauchnyj zhurnal.* 2012, no. 4 [Electronic resource]. URL: <http://www.uecs.ru/teoriya-upravleniya/item/1238-2012-04-09-08-10-34>.
7. **Petrov Yu. I.** Sovremennoe sostojanie promyshlennosti v RF neobhodimost' osvoenija vnutrennih i vneshnih territorial'nyh rynkov, *Izvestija vysshih uchebnyh zavedenij. Problemy poligrafii i izdatel'skogo dela,* 2011, no. 6, pp. 207—212.
8. **Petrov Yu. I.** Osobennosti proizvodstva slozhnyh promyshlennyh izdelij v sovremennyh uslovijah, *Nauchnye itogi 2011 goda: dostizhenija, proekty, gipotezy: sbornik materialov I Mezhdunarodnoj nauchno-prakticheskoy konferencii: v 2-h chastjah. Chast' 2.* Pod obshh. red. S. S. Chernova. Novosibirsk: Izdatel'stvo NGTU, 2011, pp. 218—227.
9. **Petrov Yu. I.** Arhitektura i algoritmicheskoe obespechenie informacionnoj sistemy normokontrolja vypusnykh kvalifikacionnyh rabot, *Programmaja inzhenerija,* 2013, no. 3, pp. 26—32 (Architecture and Algorithmic Support of Graduation Thesis Review Information System).
10. **Petrov Yu. I., Karnauhov Yu. A.** Progammnoe obespechenie "FK-Monitoring" kak sredstvo monitoringa biznes-processov, *Informacionnye tehnologii,* 2015, vol. 21, no. 7, pp. 503—510 ("FK-Monitoring" Software as a Tool for a Business Process Monitoring).
11. **Spinellis D., Gusios G.** *Ideal'naja arhitektura. Vedushhie specialisty o krasote programmnyh arhitektur.* SPb.: Simvol-Pljus, 2010. 528 p. (Ideal Architecture. Leading Experts On Beauty Software Architecture).
12. **Ford N., Najgard M.** *97 jetjudov dlja arhitektorov programmnyh sistem.* SPb.: Simvol-Pljus, 2010, 224 p.
13. **Firebird: The true open source database for Windows, Linux, Mac OS X and more.** [Electronic resource]. URL: <http://www.firebirdsql.org/>
14. **Delphi XE2** Overview. [Electronic resource]. URL: <http://edn.embarcadero.com/article/41593>.
15. **Ispol'zovanie metodov Data Mining v analize dejatel'nosti predpriyatija.** [Electronic resource] (data obrashheniya 12.05.2015 g.). URL: <http://cyberleninka.ru/article/n/ispolzovanie-metodov-data-mining-v-analize-deyatelnosti-predpriyatija>, svobodnyj.
16. **Petrov I. P., Petrov Yu. I., Petrova N. V.** Primenenie tehnologii Data Mining v voprosah differenciacii obrazovatel'nogo processa, *Informatizacija obrazovanija i nauki,* 2015. no. 3 (27), pp. 123—133 (Application Of Data Mining Technologies In Matters Of Differentiation Of The Educational Process).

А. Н. Родионов, д-р техн. наук, вед. науч. сотр.,
Вычислительный центр ДВО РАН, ran@newmail.ru

Системная интерпретация ключевых категорий концептуальных моделей баз данных

Исследованы динамические аспекты поведения сущностей в предметной области на основе моделей их собственных жизненных циклов. В роли последних использованы адаптированные процессные модели, содержащие в качестве базовых компонентов события, условия и действия. В контексте переменных свойств, меняющих свои значения в ходе реализации сущностями своих жизненных циклов, рассмотрены, интерпретированы и увязаны между собой информационно-управленческие категории, необходимые для последующего построения структурной компоненты базы данных.

Ключевые слова: жизненный цикл сущности, роли, состояния, действия, метки времени, модель и структуры данных

Введение

Каждая из актуальных задач моделирования данных, будь то задача нарушения хронологической последовательности, адаптивности структур или разбиения множества сущностей на подклассы (типы) или какая-то другая — все они должны опираться на строгую систему взаимоисключающих и в то же время увязанных между собой однозначных понятий, не допускающих двойного толкования. В этом плане применительно к области, связанной с организацией баз данных, просматриваются определенные пробелы, сдерживающие развитие соответствующей теории.

Данные, факты, события, действия, роли, состояния, статусы, условия, метки времени, жизненные циклы, сущности, объекты и свойства — вот список категорий, широко используемых в профессиональном общении и публикациях, затрагивающих область информационных знаний, которую можно обозначить как инженерия данных. За каждой из перечисленных категорий могут скрываться близкие по смыслу, но не идентичные понятия. Например, анализ многочисленных ролевых концепций, сделанный в работе [1], показывает, что роли могут быть интерпретированы как связи между сущностями, использованы в качестве оснований для генерализации/специализации последних и выступать в качестве самостоятельных объектов баз данных.

Аналогичная картина наблюдается и в отношении ряда других категорий моделей и баз данных [2—4]. Более того, практически не ставится и не решается задача по увязке этих категорий между собой.

Постоянно растущие требования к информационным системам (ИС) [5] и в первую очередь к их функциональности, что выражается в постепенном

вытеснении фактографических систем системами управления потоками работ, влекут за собой появление новых информационных объектов и необходимость в их точных определениях.

В настоящей работе введена формальная основа в виде моделей жизненных циклов сущностей (МЖЦС), повторяющих управленческие процессы предметной области (ПО), и на ее базе уточнены и увязаны между собой ключевые понятия даталогического моделирования.

1. База данных как результат отражения взаимодействующих сущностей ПО

Что должна содержать база данных (БД), чтобы считаться полной? В зависимости от того, какой будет ответ на этот вопрос, такой окажется и формальная основа для определения минимального набора взаимосвязанных категорий структурной составляющей БД и интерпретации последних.

Организацию и функционирование произвольной ПО можно представить посредством множества целевых моделей. Структурная компонента БД (модель данных) — одна из таких целевых моделей, в которой находят отражение сущности, образующие ПО, и различные аспекты их взаимодействия между собой. (В дальнейшем будем ссылаться на реляционную модель данных как один из классов логических моделей, обладающий наилучшей формальной основой.)

Собственно самих разновидностей сущностей (типов, множеств сущностей) не так уж и много, как может показаться на первый взгляд. Во всяком случае, их гораздо меньше, чем других категорий объектов, которые также "включаются" в базу данных. Покажем это чуть позже, а пока перечислим

те из типов сущностей, наличие которых можно обнаружить в любой предметной области: личности; товарно-материальные ценности; готовая продукция; финансы. Если говорить о конкретной ПО, то в ней всегда будут присутствовать специфические, характерные только для нее типы. В системах образования это будут дисциплины, виды занятий, учебные планы, в жилищно-коммунальном секторе — оборудование, услуги и т. д.

От сущностей и будем далее отталкиваться. Пока же условимся называть сущности и другие, представляющие интерес и включаемые в базы данных категории, просто объектами. Покажем, что представляет собой ПО, классифицируемая также еще и как эргономическая система, сугубо в "сущностном" аспекте.

У всех систем есть цель. Цель эргономической системы получить какой-то конечный результат в форме готового продукта или услуги, используя (переработав) некоторые ресурсы (исходные сущности). (Экономическая, социальная и прочие составляющие оставлены в стороне, поскольку на первом плане сейчас находятся сущности). Каким образом формируется конечный результат? За счет трансформации (преобразования) исходных или промежуточных сущностей, которые были получены на одной из стадий переработки. (Термин "трансформация" далее будем применять для любых манипуляций с сущностями, в том числе и приводящих к смене их местоположений.)

Продемонстрируем в контексте данных, что происходит с сущностями как в процессе их "переработки", так и по завершении преобразования. Для этого достаточно показать, что собой представляет отдельно взятое переменное свойство сущности (которое в отличие от постоянного может изменить свое текущее значение) и что с ним происходит в результате преобразования.

Концепция свойства, как и необходимость в разбиении множеств объектов на непересекающиеся подмножества для исключения избыточности и обеспечения компактности структур, в которых будут размещены данные, являются ключевыми для понимания и формализации преобразовательных процессов.

Будем придерживаться следующих обозначений для описания соответствующих множеств и свойств элементов, входящих в эти множества. Пусть

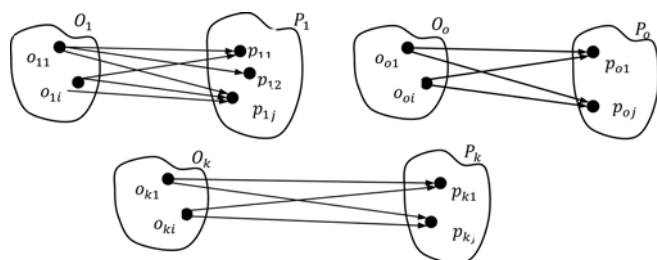


Рис. 1. Соответствия между множествами объектов и категориями свойств

$\Theta = \{O_1, O_2, O_o, \dots, O_K\}$ — совокупность K непересекающихся множеств, образованных объектами ПО. (Каждое такое множество в теории моделирования данных принято называть типом сущности (объекта)). Любому элементу $o_{oi} \in O_o$ ставятся в соответствие все элементы другого множества P_o — множества видов (категорий) свойств, которыми обладают все элементы O_o (рис. 1).

Все переменные свойства, если придерживаться классификации, приводимой в работе [6], могут быть однозначными (принимающими одно единственное значение в произвольный момент времени) или многозначными (имеющими одновременно множество значений). Например, свойство "Наименование" применительно к маркам истребителей в его однозначном и многозначном вариантах будет отличаться количеством значений. СУ-25, Frogfoot и Грач — это альтернативные названия одного и того же советского/российского штурмовика.

Следует различать классы (виды, типы) свойств и их отдельные экземпляры, которые содержат значения, принимаемые свойством в определенные периоды времени. Далее на формальном уровне с использованием небольшого примера продемонстрируем, каким образом задаются класс и экземпляр свойства.

Пусть, например, O_o — это канцелярские принадлежности (КП), o_{oi} — шариковая ручка (ШР), а p_{oj} — свойство, предназначенное для указания местоположения (МП) o_{oi} . Воспользуемся относительным способом задания местоположения, когда сущность позиционируется посредством перечисления других сущностей. В этом случае для задания p_{oj} достаточно просто перечислить элементы множества объектов $\Theta' \subseteq \Theta$, участвующих в формировании свойства, и указать семантику (*significs*) свойства, чтобы исключить потенциальные дубликаты p_{oj} , характеризующиеся тем же самым Θ' . Отсюда, $p_{oj} = (\Theta', \text{significs})$.

Согласно сказанному местоположение КП как класса можно, например, задать так: $p_{КП} \text{ МП} = (\text{Склады}, \text{МОЛ}, \text{Позиционирование})$, где *Позиционирование* — семантика свойства, *МОЛ* — материально-ответственные лица.

Чтобы перейти к значению произвольного p_{oj} потребуется воспользоваться отображением Γ либо множества O_o , либо множества, представляющего декартово произведение O_o с какими-то другими множествами, входящими в совокупность Θ , во множество Q . Последнее в общем случае также является декартовым произведением как минимум трех множеств: одного из множеств значений D_i (домен) и множеств, представляющих метки времени соответственно начала и окончания действия свойств: T_m^{start} , T_m^{finish} . Индекс m здесь указывает на разновидность используемого множества, содержащего метки времени определенного формата.

Таким образом, $\Gamma_o: O \rightarrow Q$, где $Q = D_i \times T_m^{start} \times T_m^{finish}$, $O = O_o \wedge O_o^s \times \prod_{o \in K} O_o$, где s — степень множества O_o , $s \neq 0$.

Применительно к той же ШР для задания значения, которое ее свойство местоположение может иметь в интервале $[T_m^{start} = 7, T_m^{finish} = 10]$, требуется указать экземпляры объектов множеств *Склады* и *МОЛ*, сославшись на идентификаторы последних.

Поскольку сейчас речь идет не о виде свойства, его значении, то обозначение свойства следует снабдить дополнительным индексом i , ссылающимся на объект. В итоге, значение свойств задаем как: $p_o = \text{кп } i = \text{шр } j = \text{мп} = (\text{Склад} = 10, \text{МОЛ} = 12, 7, 10)$. Далее, в целях удобства индекс o в обозначении свойства будем опускать, и трактовать p_{ij} как j -е свойство i -го объекта.

Для моделирования структур, в которых будут размещаться данные, более удобной формой представления свойства является схема реляционного отношения:

$$R(\text{Свойство})_{ij} = p_{ij} = \{O_i, \bar{O}_i, T_m^{start}, T_m^{finish}, D_i\}, (1)$$

в котором имена атрибутов ассоциируются с перечисленными ранее множествами, а \bar{O}_i — это кортеж произвольной длины, составленный из элементов множества Θ .

Из всего сказанного вытекает, что свойство объекта p_{oj} есть не что иное, как зафиксированный результат взаимодействия этого объекта с другими объектами ПО. Отсюда о траектории объекта в ПО можно судить по тому, в какие моменты времени и на сколько менялись значения его свойств.

Представленная концепция свойства была бы неполной, если бы не был рассмотрен способ фиксации экземпляров свойств в базах данных, который зачастую не отражает истинный характер изменения свойства. На схеме, приведенной на рис. 2, представлены потенциальные траектории свойства и моменты времени, когда происходит запись в БД его новых значений.

Обратим внимание на дискретный характер фиксации, а также моменты времени t_f^1, t_s^2, t_f^3 и t_f^2/t_s^3 , когда свойство мгновенно приобретает новое значение. Относительно t_f^1 и t_f^3 можно однозначно утверждать, что отображение $\Gamma_o: O \rightarrow Q$ перестало существовать. Правда, неизвестно, что предшествовало

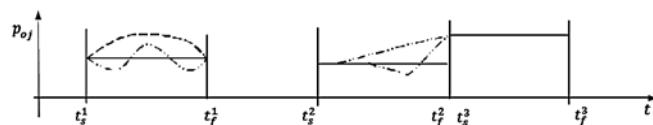


Рис. 2. Варианты динамики свойства сущности

наступлению такого события. Можно также предположить, что некоторые свойства постепенно и вероятно нелинейно меняли свои значения, о чем на рис. 2 говорят штрих-пунктирные и штриховые линии, отражающие характер этих изменений. Но это только предположения, которые если и имели место в ПО, то никакого важного значения для последней не представляли.

Несмотря на то, что соотношение (1) позволяет задать значение любого свойства сущности, требуется выработать и в дальнейшем придерживаться некоторых соглашений, когда нужно отразить факт присутствия или отсутствия сущности в ПО. Свойство "местоположение" как нельзя лучше подходит для этой цели. Достаточно, чтобы все кортежи реляционного отношения, соответствующего схеме (1), для этого свойства имели непустые значения атрибута T_m^{finish} .

Представление в базе данных сущностей предметной области и значений их свойств — это стандартный, минимальный набор требований, на основании его можно восстановить траектории, по которым любая сущность перемещалась в ПО (если траектории представить в виде последовательностей значений свойств). Но очевидно, что интерес представляют не только текущие, но и потенциальные траектории, по которым сущности могут двигаться в будущем, а также силы (механизмы), вследствие которых сущности пришли или придут в движение.

Для решения этой задачи потребуется рассмотреть, исследовать и в какой-то степени дополнить концепцию жизненного цикла сущности.

2. Интерпретация и использование процессных моделей для представления динамики (жизненных циклов) сущностей

Интуитивно (ввиду отсутствия единой точки зрения на модель жизненного цикла сущности — МЖЦС), простейшая МЖЦС должна содержать элементы, отражающие факты пересечения сущностью границ предметной области в прямом и обратном направлении (поступающие в ПО и выходящие из ПО), а также преобразования (переходы) сущности в другие сущности в результате декомпозиции или вхождения в состав таких сущностей. Но подобная модель не охватывает всего многообразия трансформаций, которые происходят или могут произойти с сущностью, и не отвечает на ключевой вопрос о причинах, приводящих к трансформациям.

Поиск сопоставимых по своему целевому назначению моделей из других областей информатики показывает, что примерно аналогичные функции поддерживаются в моделях классов eEPC [7] (событийно-управляемые процессы) и IDEF3 [8] (процессы сценариев), применяемых для описания логики и компонентов бизнес-процессов. События, условия и действия — обязательные атрибуты таких моделей, что дает основание, после незначительных модификаций, рассматривать их в качестве

Образцы ключевых характеристик реальных бизнес-процессов

Трансформируемая сущность	Иницирующие события	Название процесса	Изменяемое свойство
Пассажир Клиент	Запрос на регистрацию Запрос на стрижку	Регистрация пассажира на авиарейс Стрижка клиента в парикмахерском салоне	Регистрация Фасон прически
Покупатель Груз	Выбор приобретаемых товаров Запрос на перевозку груза	Покупка бытовой техники Доставка груза автотранспортной компанией	Приобретенные товары Пункт назначения
Измерительный прибор Упаковочная тара	Запрос на поверку Запрос на изготовление упаковок	Поверка измерительного прибора Изготовление упаковочной тары	Дата очередной поверки Появление новой сущности

МЖЦС. Сделаем выбор в пользу eEPC, как содержащих большее число разновидностей объектов по сравнению с IDEF3 и поэтому позволяющих более точно и детально описать тот или иной бизнес-процесс (БП).

Сформулируем гипотезу, согласно которой любому бизнес-процессу можно поставить в соответствие один и только один тип сущности и трактовать модель этого процесса как последовательность шагов по трансформации одного из свойств, принадлежащих этому типу.

В таблице перечислены БП, в которых специальным образом выделены: тип трансформируемой сущности; свойство, получающее новое значение; запускающее (иницирующее) процесс событие. Таким образом, с формальной точки зрения любой бизнес-процесс задает следующее ключевое преобразование: $p_{ij}(t_s) \xrightarrow{B_r} p_{ij}(t_f)$, где $p_{ij}(t_s)$ — j -е свойство i -й сущности в момент времени начала БП t_s ; $p_{ij}(t_f)$ — то же самое в момент t_f -его завершения; B_r — БП с индексом r .

Обратим внимание на то, что входы всех процессов содержат одно из свойств сущности, принадлежащей некоторому типу, и событие, запускающее процесс, а конечным результатом выполнения является новое значение, приобретаемое свойством.

Если следовать сформулированной гипотезе, то eEPC-модели в первом приближении подходят на роль МЖЦС. Далее в контексте сущностных взаимодействий продемонстрируем, что происходит с сущностями предметной области в процессе перехода от $p_{ij}(t_s)$ к $p_{ij}(t_f)$.

3. Компоненты жизненных циклов сущностей: состояния, события, действия, условия

Совокупность значений, принимаемых свойствами сущности в произвольный момент времени, определяет то, что принято считать состоянием [2]. Таким образом, состояние i -й сущности в момент времени t — $S_i(t)$ формально задается кортежем $S_i(t) = (p_{i1}(t), p_{i2}(t), \dots, p_{ij}(t))$, где $p_{ij}(t)$ — значение j -го свойства в момент t . О целесообразности отслеживания $S_i(t)$ будет сказано ниже, а пока отметим, что одной единственной характеристики, даже такой обобщенной, каковой является состояние, явно

недостаточно для описания поведения объектов, которое по своей природе динамично и непрерывно. Важно также знать, каким образом менялось то или иное свойство. Эту функцию в ИС выполнял действия. Действия, таким образом, в какой-то степени подменяют собой непрерывность. Покажем в общем виде, что происходит с сущностью в промежутке времени, когда одно из ее свойств меняет свое значение. Наглядное представление об этом дает срез фрагмента бизнес-процесса (рис. 3), на котором представлены только объекты, участвующие в серии взаимодействий, повлекших за собой изменение их свойств.

Пусть значение j -го свойства сущности с индексом i в промежутке времени $[t_s, t_f]$ изменилось на величину $\Delta p_{ij} = p_{ij}(t_s) - p_{ij}(t_f)$. Поскольку, как ранее было замечено, любое свойство — это результат взаимодействия сущности с другими сущностями (в общем случае с другими объектами), то Δp_{ij} могло произойти только, если в промежутке $[t_s, t_f]$ менялись какие-то другие свойства $j' \neq j$ i -й сущности и (или) свойства других сущностей $i' \in i, i' \in I, j, j' \in J$, где I, J — множества индексов сущностей и свойств i -й сущности соответственно. Таким образом, всякому изменению свойства p_{ij} , которое (изменение) для удобства обозначим как \vec{p}_{ij} , может быть поставлена в соответствие последовательность $\{\vec{p}_{ij}^s\}$, где s — порядковый номер элемента последовательности, образованной модифицируемыми свойствами как i -го объекта, так и свойствами других i' -х объектов. Более того, таких последовательностей, участвующих в достижении новых значений p_{ij} , может быть несколько. Например, процессы регистрации пассажиров в разных аэропортах на рейсы разных авиакомпаний могут различаться друг от друга, что, соответственно, порождает и уникаль-



Рис. 3. Последовательность модифицируемых свойств, "участвующих" в достижении $|\Delta p_{ij}|$

ные $\{\vec{p}_{ij}\}_k$. (Здесь k обозначает индекс подобной последовательности.)

Структура (содержание) конкретной последовательности зависит только от самого факта изменения значения свойства. Поэтому не имеет принципиального значения приращение свойства $[\Delta p_{ij}]$.

Среди элементов, образующих $\{\vec{p}_{ij}^s\}$ последовательности, встречаются не только свойства, но и встроенные $\{\vec{p}_{i'j}^s\}$ или $\{\vec{p}_{ij'}^s\}$. Данное обстоятельство отражает факт того, что в рамках основного процесса разворачиваются отдельные, самостоятельные подпроцессы.

Сформулируем и запишем очевидные темпоральные ограничения, которые должны выполняться для $\forall\{\vec{p}_{ij}^s\}$.

1. Обозначим через $[t_s, t_f]_{p_{ij}} = Y_{ij}$ интервал изменения p_{ij} . Первое ограничение состоит в том, что

$\forall [t_s, t_f]_{p_{ij}}^{n_{ij}} \in Y_{ij} \wedge \forall [t_s, t_f]_{p_{i'j}}^{n_{i'j}} \in Y_{ij}$, где $n_{i'j}$ и $n_{ij'}$ — порядковые номера интервалов изменения соответственно p_{ij} и $p_{i'j}$, $n_{i'j} = 1 \dots N_{i'j}$; $n_{ij'} = 1 \dots N_{ij'}$. $N_{i'j}$, $N_{ij'}$ — общее число таких интервалов. Другими словами, вес промежутки времени, в течение которых происходит изменение любого из свойства последовательности $\{\vec{p}_{ij}^s\}$, должны лежать в границах Y_{ij} .

2. Если какое-то из свойств $\{\vec{p}_{ij}^s\}$ несколько раз меняет свое значение в интервале Y_{ij} , т. е. $n \neq 1$, как,

например, p_{ij} на рис. 3, то все интервалы $Y_{ij}^{n_{ij}}$ не

должны пересекаться, $\bigcap_{n_{ij}=1}^{N_{ij}} Y_{ij}^{n_{ij}} = \emptyset$. Сказанное свидетельствует о том, что свойство не может быть подвергнуто многократному изменению в одном и том же временном диапазоне. Например, "Преподаватель не может одновременно читать две и более лекции, один и тот же самолет не в состоянии обслуживать рейсы, которые выполняются в одно и то же время."

3. Третье ограничение касается величины "технологического" разрыва Δ_{ij} (см. рис. 3), который может иметь место для отдельных свойств сущности. Если, например, занятия со студентами проводят в аудиториях, расположенных в разных корпусах, то необходимо время для перемещения студентов из одного корпуса в другой корпус. (Более детальное рассмотрение этого ограничения ввиду его очевидной значимости и сложности моделирования составляет предмет отдельного исследования.) Заметим, что последнее ограничение требует фиксации Δ_{ij} в метаструктурах базы данных.

Перечисление одних только последовательностей $\{\vec{p}_{ij}^s\}$, несмотря на то, что они показывают траекторий, по которым меняется то или иное свойство сущности, не дает ответа на вопрос о том, что происходит с сущностью в промежутке $[t_s, t_f]$. Как было отмечено ранее, эта роль возлагается на дей-

ствия. Это могут быть как действия, выполняемые самой i -й сущностью, так и действия, совершаемые другими i' -ми сущностями, воздействующими на i . Таким образом, каждое действие $a_{i'i}^h$ кроме очевидного названия, характеризуется также источником и адресатом. Здесь h — индекс действия; i — ссылка на объект, выступающий в роли источника; i' — индекс объекта, на которое направлено действие.

В отличие от показанного ранее соответствия (см. рис. 1) между множествами объектов и свойств в отношении действий уместно говорить о композиции двух отображений: $D_1: \Theta \rightarrow A$ и $D_2: A \rightarrow \Theta$, где A — множество действий. В результате получаем $D: \Theta \rightarrow \Theta$.

Подчеркнем, что действия могут совершать не все объекты, а только их часть в лице сущностей (что в общем-то очевидно, так как другие категории объектов являются в некотором роде производными от сущностей). В свою очередь, адресатами действий могут выступать все объекты, а не только сущности.

Используя концепцию действий, рассмотрим детально элементарное \vec{p}_{ij} (не включающее, как было отмечено ранее, собственные $\{\vec{p}_{ij}\}_k$). Вообще говоря, интерес представляют только несколько моментов, которые сформулируем в виде вопросов.

1. Вследствие того, что свойство некоторого объекта p_{ij} согласно (1), предполагает указание как исходного объекта, так и перечисление объектов, участвующих во взаимодействии с исходным, могут ли в изменении этого свойства принимать участие третьи объекты?

2. Может ли \vec{p}_{ij} произойти только в результате взаимодействий исключительно третьих объектов (без участия исходного и вторичных объектов)?

Ответы на оба вопроса будут отрицательными. Если предположить, что отыщутся некие третьи объекты, влияющие посредством своих действий на \vec{p}_{ij} , то они также будут вынуждены изменить какие-то свои свойства. Но в этом случае \vec{p}_{ij} не будет элементарным.

Представленные рассуждения позволяют определиться (рис. 4) с частными (а, б, в) и общим (г)

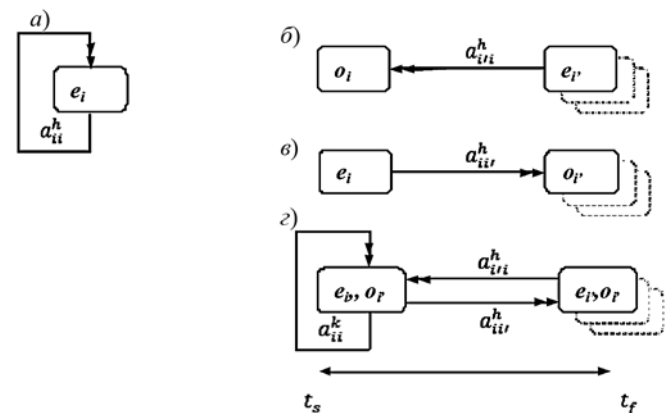


Рис. 4. Классы взаимодействий между объектами

классами взаимодействий между сущностями посредством действий, приводящих к модификации значений их свойств. Сущности здесь помечены символом e . (Далее везде для обозначения экземпляров и типов будем использовать соответственно строчные и прописные буквы.)

Увязав между собой объекты, их свойства и действия, как элементы динамики свойств, можно, опираясь на упомянутые выше eEPC-модели, приступить к последовательному поэтапному построению МЖЦС. Ограничимся на первом этапе только действиями, сделав акцент на их альтернативность и множественность в реализации \vec{p}_{ij} .

Фрагмент процесса (рис. 5) иллюстрирует две альтернативные последовательности действий для получения фотографического снимка. Для каждого действия можно указать объект-источник и объекты, на которые направлены действия. Выполнение любой их двух последовательностей $\{a^1, a^2, a^3, a^4\}$ или $\{a^1, a^2, a^3, a^5\}$, приводит к получению цифровой фотографии, что в контексте данных трактуется как изменение одного из свойств фотографии, ранее обозначенного как "местоположение".

Этот небольшой фрагмент достаточно информативен, так как дает дополнительную информацию о действиях. Помимо исходных и целевых объектов, в действиях a^4 и a^5 явно присутствуют объекты, посредством которых выполняются действия — фотоаппарат и смартфон соответственно.

Таким образом, каждое действие характеризуется своей семантикой, сущностью, которое его осуществляет, подмножеством сущностей \hat{E} , посредством которых оно выполняется, подмножеством объектов \hat{O} , на которые направлено, а также моментами времени начала и окончания действия:

$$\text{Действие} = (\text{significs}, e_i, \hat{E}, \hat{O}, T_m^{\text{start}}, T_m^{\text{finish}}). \quad (2)$$

Если посредством действий можно показать, что делает или не делает сущность в предметной области в произвольный момент времени, то закономерно возникает вопрос, касающийся того, что приводит сущность в "движение" или по-другому, что или кто активизирует каждое действие, результатом которого является новое значение свойства

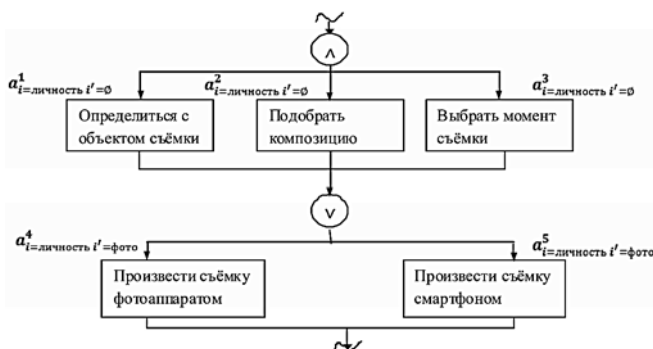


Рис. 5. Альтернативные последовательности действий

сущности. Теория моделирования бизнес-процессов [7, 9] говорит о том, что инициатором любого действия является событие или совокупность событий.

В научной литературе можно встретить множество определений события. Применительно к моделированию данных событие обычно трактуется как мгновенный факт [4]. Факты, в свою очередь, могут ассоциироваться с \vec{p}_{ij} , как это предлагается в работе [2]. Например, к мгновенным фактам (событиям) в [2] относят продажи, покупки, платежи.

В то же время события, если считать их элементами управления (а не только мгновенными фактами, что является только формой их проявления), не сводятся исключительно к изменению свойства \vec{p}_{ij} . Раскроем механизм функционирования событий, для чего исследуем отношения, возникающие между событиями и действиями, событиями и сущностями, действиями и сущностями. Как и в случае с действиями, будем также опираться на организацию моделей бизнес-процессов, принятых ранее в качестве прообразов моделей жизненных циклов сущностей.

В таких моделях могут быть выделены несколько групп событий, которые инициируют действия. Это события V^E , источники которых сущности. Например, "Клиент выразил желание заключить договор аренды помещения", "Воздушное судно запросило разрешение на посадку", "Банк получил платежное поручение, на основании которого он должен провести транзакцию". Во всех случаях явно присутствуют один или несколько сущностных адресатов (агентов).

Другую группу составляют "временные" события V^T . Например, наступило время "Ч" и активизировалось событие, запустившее процесс сдачи экзамена.

В отдельную группу V^S сводятся события, возникающие, когда сущности достигают каких-то определенных, характерных состояний. "Клиент открывает вклад на сумму 700 000 рублей", "Величина прожиточного минимума составила 7253 рубля", "Объем заказа превысил установленный максимум", "Уровень текущих запасов достиг критических (страховых) значений". (Обратим внимание на то, что здесь фигурируют состояния, вопрос о целесообразности использования которых поднимался ранее.)

Еще одну группу составляют события, сигнализирующие о том, успешно завершилось то или иное действие или нет, $V^R = \{true, false\}$.

Чтобы каким-то образом различать результирующие события, которые "привязаны" непосредственно к действиям и выступают для них одновременно в роли входных и выходных событий, и события, составляющие V^E , V^T и V^S , являющиеся по отношению к действиям всегда входными, будем последние называть иницилирующими событиями V^I , $V^I = V^E \cup V^T \cup V^S$.

На схеме (рис. 6) показаны стандартные отношения, возникающие между действиями и событиями.

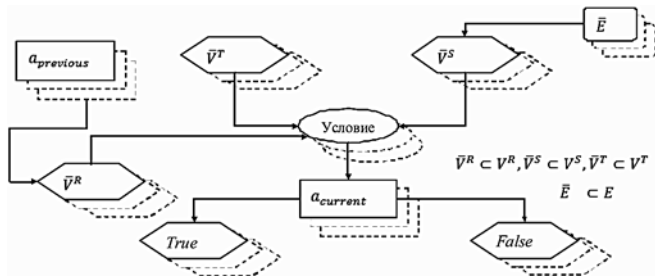


Рис. 6. Отношения между событиями и действиями



Рис. 7. Входы и выходы условия как объекта МЖЦС

В соответствии с представленными конфигурациями каждому действию предшествует одно или несколько инициирующих событий и (или) события, сгенерированные уже завершившимися действиями.

Показанные соответствия между множествами событий, действий и сущностей очевидны и вытекают из природы поведения сущностей в организационно-экономических системах, в которых ничего кроме сущностей и взаимодействий последних друг с другом обнаружить нельзя.

Осталось определиться с условиями, которые задают альтернативные траектории "движения" сущностей в рамках их жизненных циклов, а также с подмножеством "пустых" действий, которые не оказывают никакого воздействия на свойства сущностей.

"Имеется ли бронь на гостиничный номер?", "Есть ли на складе достаточный запас товара?", "Какова фактическая стоимость кредита?". Это все примеры "пустых" действий. Их объединяет то, что они сводятся к получению либо информации непосредственно о текущих свойствах сущностей, либо информации, которая может быть выведена на основании тех значений, которые принимали свойства. (Схема реляционного отношения, которая будет представлена ниже, позволяет фиксировать такие действия, не прибегая к использованию дополнительных отношений.)

Реальные сущности в границах моделей своих жизненных циклов могут "двигаться" по различным, альтернативным траекториям. "Развилочную" миссию выполняют условия, функционирование которых можно условно представить в виде "черного" ящика (рис. 7). Входы для условий — это разнообразные события, выходы — инициируемые действия.

С точки зрения математической логики любое событие представляет собой высказывание, которое всегда истинно. Тогда условие, в свою очередь, — это сложное высказывание, полученное из простых высказываний посредством логических связок "не", "и", "или". (Заметим, что импликация и эквиваленция, которые также используются для построения сложных высказываний в ePC не используются.)

4. Факты и привязка ко времени объектов, событий и действий

Моделирование событий, действий и условий было бы неполным, если бы не был затронут хотя бы в общих чертах фактор времени.

Формально метка времени — это одна или несколько единиц измерения времени (квантов времени C), поставленных в соответствие какому-то информационному объекту: свойству сущности, событию или действию. Речь идет о кортежах вида: $(o_i, c_1, c_2, \dots, c_h)$, $(v_i, c_1, c_2, \dots, c_h)$, $(a_i, c_1, c_2, \dots, c_h)$. Собственно говоря, факты — это и есть подобные кортежи.

Обратим внимание на то, что условия не требуют привязки ко времени. Будем полагать, что, во-первых, они срабатывают мгновенно, а во-вторых, их окружают события и действия, которые уже снабжены метками времени и поэтому несложно установить момент времени срабатывания каждого конкретного условия.

Очевидно, что если подходить к вопросу привязки системно, то за интегрирующую основу логично принять МЖЦС, объединяющую все перечисленные в заголовке раздела категории.

Какой бы аспект моделирования данных не затрагивался, всегда необходимо проводить различие между моделированием типов и моделированием их экземпляров, а также устанавливая отношения между получаемыми моделями. Будем и здесь придерживаться этой схемы. Отметим наиболее значимые моменты. Как и в отношении объектов, действий и событий МЖЦС также разбиваются на пересекающиеся множества, именуемые типами, а их элементы считаются экземплярами. Например, для типа процесса "Проведение медицинского освидетельствования участников дорожного движения" всегда будет иметь место множество реализаций, число которых будет увеличиваться по мере появления очередного "клиента". Далее, под МЖЦС всегда будет подразумеваться тип, а ЖЦС ассоциироваться с экземпляром.

По понятным причинам в составе МЖЦС в качестве их элементов могут присутствовать только типы (объектов, действий и условий). Соответственно ЖЦС будут всегда содержать только экземпляры объектов, действий и условий. Вероятно, что должны существовать и различия в форматах и механизмах представления и привязки ко времени объектов, входящих в состав МЖЦС и ЖЦС. В МЖЦС всегда будет фигурировать абстрактное относительное (нереальное) время ("начало учебного года", каждая вторая неделя месяца и т. п.). Тогда как в процессе отслеживания реальных траекторий сущностей (в ЖЦС) будет задействовано фактическое, астрономическое время. Наглядное представление об использовании меток времени в МЖЦС дает схема, приведенная на рис. 8. На ней присутствуют две альтернативные метки времени D_f и C_h , а также N_c , являющаяся производной от

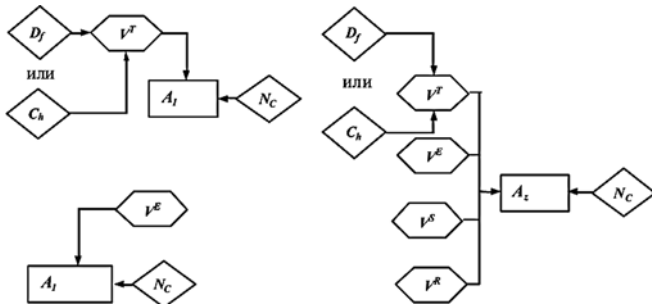


Рис. 8. Допустимые метки времени для стандартных конфигураций событий и действий в МЖЦС

них. Раскроем содержание и назначение всех перечисленных форматов.

Будем руководствоваться тем, что любая единица измерения времени всегда характеризуется продолжительностью, а совокупности таких единиц образуют последовательности (месяцы года, дни недели, недели года и т. д.). Поэтому для моделирования относительных времен достаточно сослаться на порядковый номер кванта h в такой последовательности C_h . Но можно, если база данных "поддерживает" требуемую для моделируемого процесса разновидность квантов — "временной" тип (формат) данных D , воспользоваться одним из таких форматов D_f , где f — порядковый номер формата. Предположим, что "встроенный" формат имеет размерности $hhmm$ (часы-минуты). Тогда $D_f = hhmm = 12.00$ будет означать, что событие произойдет в 12 часов дня.

Иницирующие события в зависимости от их класса (сущностей V^E , времени V^T и состояний V^S) на первый взгляд могут соотноситься как с квантами C_h , так и с временными типами данных D_f . Но для элементов V^E и V^S , ввиду случайного характера их наступления, метку времени ни в виде D_f , ни посредством C_h , задать нельзя. Что касается V^R , то оно наступает в момент завершения действия и может быть получено расчетным путем.

Вне зависимости от того, используется ли в качестве меток времени для событий D_f или C_h , продолжительность любого действия N_c всегда будет измеряться числом квантов, в течение которых это действие будет выполняться. Альтернативный для N_c вариант с указанием времени начала действия отпадает ввиду того, что формально МЖЦС — это граф, а в графе присутствует множество путей разной длины (продолжительности) достижения конкретного действия. Поэтому нельзя точно определить, когда начнется текущее действие A_z .

В отношении исходных действий A_1 также не имеет смысла задание времени их начала, поскольку оно совпадает с меткой времени события V^T .

Измерение N_c посредством квантов в случае использования D_f для событий, актуализирующих действия, приводит к необходимости решения задачи определения размера кванта для выбранного D_f . Перечислим варианты, которые могут возникнуть. Иницирующее D_f задействует одну единицу изме-

рения, а продолжительность действия — другую. Например, D_f — это часы, а действия — минуты. Поскольку подобных несоответствий в МЖЦС скорее всего окажется несколько, приемлемым решением был бы выбор наименьшего используемого в МЖЦС кванта и приведение всех D_f и C_h к этому значению. (В противном случае невозможно определить время завершения каждого действия и всего процесса в целом.) Если придерживаться этого соглашения, то вид кванта h становится одной из важнейших характеристик процесса, определяя по сути все возникающие в последнем темпоральные ограничения.

Появление в МЖЦС V^T -событий, которые активизируют промежуточные и завершающие действия A_z , при некорректном задании D_f или C_h может привести к старту текущего A_z при еще не завершившихся предшествующих действиях. Для примера возьмем процесс "Проведение научной конференции", в котором пленарные и секционные заседания начинаются в фиксированные времена и первый доклад в секционной части может начаться не раньше завершения последнего доклада в пленарной части.

Поскольку на графе МЖЦС к произвольному A_z может вести несколько путей, то длину максимального пути можно использовать для формализации ограничения, исключающего возникновение подобных ситуаций.

Пусть P_z — множество путей, ведущих к A_z , каждый из которых характеризуется длиной, измеряемой числом квантов. $L = \max P_z$ — наибольший элемент множества P_z . Тогда $C_h(V_z^T(A_z)) \geq L$.

В отличие от МЖЦС отслеживание траекторий реальных сущностей, которые протекают в границах этих МЖЦС, ведется уже в реальном масштабе времени, что исключает использование C_h . Все современные СУБД поддерживают встроенный "временной" формат вида $uuuummddhhss$ и содержат семейство "временных" функций, позволяющих получить из указанного формата широкий диапазон разновидностей квантов. Тем самым, практически исключается необходимость в поддержке каких-то уникальных квантовых форматов.

Соответствия, возникающие между метками времени d_f и событиями, d_f и действиями в ЖЦС, приведены на рис. 9.

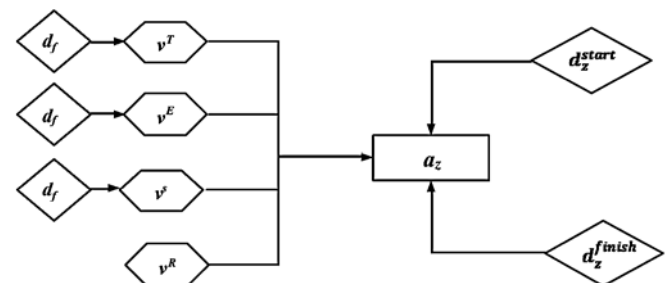


Рис. 9. Допустимые метки времени для событий и действий в ЖЦС

Представленные здесь отображения ($a_z \rightarrow d_z^{start}$ и $a_z \rightarrow d_z^{finish}$) справедливы для $\forall a_z; d_z^{start}$ и d_z^{finish} задают моменты времени соответственно начала и завершения действия a_z .

5. Условия в качестве объектов базы данных и структуры, моделирующие МЖЦС

Вопрос, на который предстоит получить ответ, формулируется следующим образом: нужно ли хранить в базах данных условия и если да, то каким образом это следует делать?

Для отслеживания траекторий реальных сущностей нет необходимости в фиксации условий, потому что интерес представляют только те факты, в какой последовательности, когда и на какое значение менялись свойства сущностей. Этот минимальный набор требований, который выдвигают учетные ИС, именуемые еще и как системы обработки транзакций. Повышение ИС в классе, например, переход их в разряд систем принятия решений, влечет за собой расширение исходного набора требований, который может быть дополнен, в том числе, и требованиями отслеживания потенциальных траекторий трансформации сущности. Такая информация может быть получена, если только в базе данных хранится непосредственно сама модель жизненного цикла, элементами которой в том числе являются и условия.

Поскольку условие в общем виде задается с помощью сложного высказывания (например, "Размер вклада лежит в пределах от 100 до 700 тысяч рублей"), оно может быть определено на домене символов (символьной строки).

Далее покажем, к какой структурной организации (модели данных) приводят результаты, полученные в ходе рассмотрения динамики сущностей в ПО. Ограничимся простейшей подсхемой (рис. 10), отражающей только МЖЦС в ее базовой конфигура-

ции. Имена атрибутов с надстрочным индексом "*" обозначают первичные ключи. Атрибуты, не содержащие индексов, соответствуют внешним ключам.

Часть структур и большинство атрибутов не вошли в модель. Например, на схеме отсутствуют отношения, в которых перечисляются роли, несмотря на то, что в определенных случаях они подменяют классы сущностей в МЖЦС. Не отражены ретроспективные и перспективные модели жизненных циклов и некоторые другие не менее значимые моменты.

Тем не менее, на схеме представлены структуры, образующие ядро МЖЦС-кластера. Посредством этих структур в базе данных может быть размещена информация об узлах и связях графа, каковым с формальной точки зрения является любая МЖЦС.

Структурный кластер, представляющий, в отличие от МЖЦС, экземпляры последних, будет иметь еще более сложную организацию, в том числе включать типы сущностей, характерные для конкретной ПО.

6. Локализация множества сущностей. Исключение ролей и услуг

В начале работы говорилось о том, что разновидностей того, что принято называть сущностями, которые часто ассоциируются с именованными структурами баз данных, в любой ПО на самом деле не так уж и много. За сущности могут приниматься объекты, выполняющие классификационные функции, "слабые сущности", которые на самом деле представляют собой свойства, документы, большинство из них содержат данные о взаимодействиях сущностей (часть документов "заменяют" композитные, составные сущности), и ряд других объектов. Все они, без сомнения, важны и необходимы, но в предметных областях демонстрируют отличное от сущностей поведение, выполняя, по сути, вспомогательную функцию.

К сущностям часто причисляют и услуги, хотя природа последних как самостоятельной категории понятна не до конца. Включение в модель данных структур, представляющих МЖЦС и их экземпляры, дает основание представить новое толкование услуги, которое, с нашей точки зрения, наиболее точно отражает ее суть.

Услуга — это именованная МЖЦС в форме бизнес-процесса. Действительно, если обратиться к приведенной выше таблице, то название каждого из процессов не что иное, как название услуги. (По крайней мере, в той общепринятой манере, как их (услуги) принято обозначать). Таким образом, услуги — это не

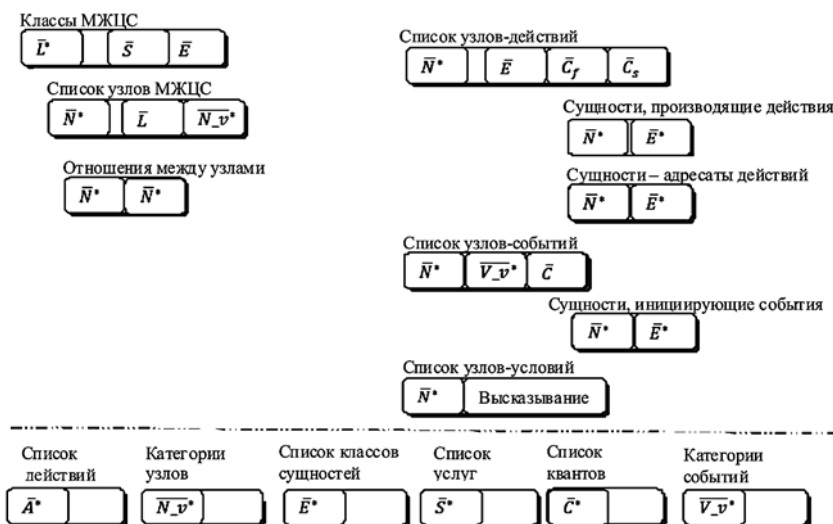


Рис. 10. Структурная организация МЖЦС (концептуальная модель)

сущности. Тем не менее, в реальных базах данных они часто включаются в состав отношений, где перечисляются сущности и задаются характеристики последних. Это своего рода компромисс, типичный для ИС, когда в границах одной и той же ИС одновременно сочетаются и классические учетные фрагменты, и компоненты систем более высокого уровня. Очевидно, что в ИС, поддерживающих исключительно бизнес-процессы, потребность в фиксации услуг отпадает сама собой.

Еще одно подмножество объектов, которые не принадлежат к сущностям, — это роли. Потребность в ролевой концепции возникает всякий раз, когда трансформация сущности в предметной области не ограничивается ее одним единственным жизненным циклом. Возьмем для примера некоего абстрактного человека, являющегося элементом типа сущности "Личности" и работающего в университете. Он может быть преподавателем и в то же время получать *n*-е высшее образование в том же университете, рецензировать научную работу, участвовать в работе диссертационного совета и т. д. Одним словом, "играть" несколько ролей в соответствии с алгоритмами, которые задаются определенными жизненными циклами.

Таким образом, в представленном контексте роли и должности — это тождественные понятия. И такого рода объекты можно обнаружить во многих предметных областях.

Если опираться на только что представленное назначение ролей, то не составляет труда установить два важных для полноты отражения предметной области соответствия: между элементами множеств бизнес-процессов и экземпляров ролей с одной стороны, и между множествами ролей и подмножеством экземпляров реальных сущностей, с другой стороны (рис. 11).

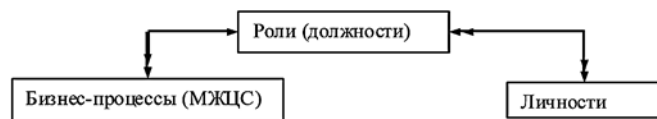


Рис. 11. Стандартные отношения между ролями, типами сущностей и МЖЦС

Тем самым удастся показать, что та или иная сущность (личность в рассматриваемом случае) может параллельно участвовать в нескольких процессах.

Использование ролей в качестве посредника между МЖЦС и экземплярами сущностей обладает еще одним положительным эффектом. Не составляет труда получить ответ на такой вопрос: может ли сущность одновременно участвовать в нескольких процессах? Постановка и решение данной задачи

выходит за рамки статьи и может составить предмет отдельного исследования.

Заключение

Установление взаимно однозначного соответствия между бизнес-процессами и жизненными циклами сущностей формирует основу для более точной интерпретации и взаимной увязки ключевых понятий, используемых в моделях данных и методах их проектирования. Исследование кинематических и динамических аспектов поведения сущностей в ПО через призму процессов раскрывает "движущие силы" в виде событий, условий и действий, приводящие к разноплановым трансформациям сущностей. Результатом таких трансформаций становятся новые значения, которые приобретают переменные свойства сущностей. При этом переход к новым значениям происходит вследствие серий взаимодействий, в том числе и среди сущностей, которые напрямую, через бизнес-процесс, не вовлечены в трансформацию.

Учет в моделях данных динамики сущностей позволяет также однозначно определиться с содержанием таких категорий, как услуги и роли. Услуги ассоциируются с названиями МЖЦС, а роли, которые привязываются к сущностям, однозначно задают подмножество МЖЦС.

Полученные результаты делают возможным решение, по крайней мере, двух актуальных задач моделирования данных: разработку концептуальных схем, исключающих нарушение темпорального порядка фактов в процессе регистрации последних в структурах базы данных и моделирования минимальных временных технологических заделов, определяющих готовность сущности участвовать в новых взаимодействиях.

Список литературы

1. **Steimann F.** On the representation of roles in object-oriented and conceptual modeling // *Data and knowledge engineering*. 2000. Vol. 35. P. 83–106.
2. **Jimenez L. J.** REEEM: Reenhancing the entity-relationship model // *Data and knowledge engineering*. 2006. Vol. 58. P. 410–435.
3. **Date C. J., Hugh Darwen, Nikos A.** Lorentzos Time and relation theory. — Morgan Kaufmann Publishers, 2014. 560 p.
4. **Gregersen H., Jensen C. S.** et al. The consensus glossary of temporal database concepts. *Temporal Databases: Research and Practice*. Springer, 1998. P. 367–405.
5. **McGinnes S., Kapros E.** Conceptual independence: A design principle for the construction of adaptive information systems // *Information system*. 2015. Vol. 47. P. 33–505.
6. **Родионов А. Н.** Качество дагалогических схем. Принцип компактности моделей данных и его приложения II // *Информатика и системы управления*. 2012. № 2. С. 14–25.
7. **Шеер А.** Моделирование бизнес-процессов. М.: Весть-МетаТехнология, 2000. 207 с.
8. **Елиферов В. Г., Репин В. В.** Бизнес-процессы: Регламентация и управление. М.: ИНФРА-М, 2011. 319 с.
9. **Вил ван дер Аалст, Кейс ванн Хей.** Управление потоками работ: модели, методы и системы. М.: ФИЗМАТЛИТ, 2007. 316 с.

System Interpretation of Conceptual Data Model Key Categories

This paper examines dynamic aspects of entities behavior at domain of interest by means of original entities lifespan models (ELM). After some supplements to extended event driven process chain (eEPC) model the last uses as a basic for ELM. Such ELM enables to determine entities trajectories and objects which compel entities to migrate with the confine of area. Dynamic of entity interprets as property changes due to interaction of an entity with another objects including entities. Any ELM associates with one and only one property which have to be modified. Actions, events and conditions as components of an entity lifespan are connected with objects and timestamps. Thereby all data model categories become interlocking and pursue the aim to derive more detail and exact database. Addition of unified structures for recording ELM and its instances in data model permits to construe some information objects as roles.

Keywords: entity lifespan, roles, states, actions, timestamps, data model, data structures

References

1. **Steimann F.** On the representation of roles in object-oriented and conceptual modeling, *Data and knowledge engineering*, 2000, vol. 35, pp. 83–106.
2. **Jimenez L. J.** REEEM: Reenhancing the entity-relationship model, *Data and knowledge engineering*, 2006, vol. 58, pp. 410–435.
3. **Date C. J., Hugh Darwen, Nikos A.** *Lorentzos Time and relation theory*, Morgan Kaufmann Publishers, 2014, 560 p.
4. **Gregersen H., Jensen C. S.** et al. The *consensus glossery of temporal database concepts*. Temporal Databases: Research and Practice, Springer, 1998, pp. 367–405.
5. **McGinnes S., Kapros E.** Conceptual independence: A design principle for the construction of adaptive information systems, *Information system*, 2015, no. 47, pp. 33–505.
6. **Rodionov A. N.** Kachestvo datalogicheskikh shem. Princip kompaktnosti modelej dannyh i ego prilozheniya II, *Informatika i sistemy upravleniya*, 2012, no. 2, pp. 14–25.
7. **Sheer A.** Modelirovanie biznes-processov, Moscow, Vest'-MetaTehnologija, 2000, 207 p.
8. **Eliferov V. G., Repin V. V.** *Biznes-processy: Reglamentacija i upravlenie*, Moscow, IN-FRA-M, 2011, 319 p.
9. **Vil van der Aalst, Kejs vann Hej.** *Upravlenie potokami robot: modeli, metody i sistemy*, Moscow, FIZMATLIT, 2007, 316 p.

УДК 519.95

М. В. Кузнецова, аспирант,
Московский физико-технический институт, Москва,
В. В. Стрижов, д-р физ.-мат. наук, вед. науч. сотр.,
Вычислительный Центр им. А. А. Дородницына ФИЦ ИУ РАН, Москва.

Локальное прогнозирование временных рядов с использованием инвариантных преобразований*

Описан метод построения прогностической модели одномерного временного ряда. Предложено найти сегменты локальной предыстории, похожие на прогнозируемый сегмент. Близкие по заданной функции расстояния сегменты объединены в кластер. Прогностическое значение вычислено путем усреднения значений сегментов кластера. Для повышения качества прогнозирования введено инвариантное преобразование сегментов — преобразование, сохраняющее эквивалентность временных рядов на кластерах. Для преобразования использована функция, построенная с помощью метода динамического выравнивания временных рядов. Предложенный алгоритм проиллюстрирован временными рядами, описывающими движение человека и содержащими показания акселерометра. Построен ретроспективный прогноз временного ряда. В вычислительном эксперименте сравниваются два алгоритма построения прогностической модели: с кластеризацией сегментов временного ряда и с использованием метода ближайшего соседа.

Ключевые слова: кластеризация, алгоритм ближайшего соседа, путь наименьшей стоимости, динамическое выравнивание, локальное прогнозирование, монотонное преобразование, ретроспективный прогноз, опорный сегмент

* Работа выполнена при поддержке РФФИ: проект 16-07-01155.

Введение

Решается задача построения прогноза временного ряда методом локального прогнозирования [1]. Пусть во временном ряду выделен опорный сегмент — некоторая последняя часть временного ряда. На основании значений этого сегмента необходимо спрогнозировать несколько отсчетов временного ряда. Идея локального метода прогнозирования заключается в следующем. Предполагается, что во временном ряду находятся сегменты, похожие на опорный. Прогноз строится экстраполяцией взвешенного среднего арифметического найденных сегментов.

Предполагается, что найденные сегменты временного ряда, похожие на опорный, имеют различные растяжения и сжатия относительно оси времени. Поэтому для улучшения качества прогноза предлагается ввести инвариантное преобразование, учитывающее трансформации сегментов. Для иллюстрации свойств предлагаемого алгоритма выбраны временные ряды, описывающие движение человека. Процедура построения прогностической модели включает следующие этапы.

1. *Кластеризация.* Сегменты временного ряда кластеризуют [2—4]. Это делается для выделения схожих групп сегментов и построения в каждой из групп инвариантного преобразования.

2. *Введение преобразования.* Для улучшения качества прогноза и учета растяжений и сдвигов сегментов относительно шкалы времени вводится преобразование сегментов. Предполагается, что во временном ряду найдется эталонный сегмент — сегмент, относительно которого можно выровнять все остальные. Задача выбора эталонного сегмента решается в работе [5]. Выравнивающую функцию строят с помощью пути наименьшей стоимости [6]. В качестве эталонного сегмента в данной работе взят центр кластера или опорный сегмент.

3. *Построение прогностической модели.* Прогноз строят как взвешенное среднее найденных преобразованных сегментов.

4. *Оценка ошибки полученного прогноза.* Для оценки ошибки используют понятие ретроспективного прогноза — прогноз строят на период времени, для которого уже имеются исторические значения. Затем полученные прогнозные значения сравнивают с историческими данными [7].

Сегменты временного ряда кластеризуют одним из методов: k -means; C -means [5, 8]; с помощью спектральных методов (вычисляется матрица Грамма) [9], EM -алгоритма и цепей Маркова [10].

Прогностическую модель строят с учетом преобразований, позволяющих выявить похожие во введенной метрике сегменты временного ряда. В работе [11] проведено сравнение различных метрик: евклидовой, взвешенной евклидовой и метрики Минковского. В работах [12, 13] используется путь наименьшей стоимости. В работе [14] рассмотрены понятие инвариантного преобразования и выбор

наиболее подходящего для решения задачи прогнозирования. В работе [15] введено понятие параметрического инвариантного преобразования. В работе [5, 16] с помощью метода динамического выравнивания введена функция сдвига между кривыми и затем рассмотрено выравнивание кривых относительно некоторой эталонной, в целях получения кривой общей формы. В работе [17] рассматривается парное выравнивание кривых относительно друг друга и получение устойчивой оценки для функции деформации.

Новизна работы заключается во введении понятия инвариантного преобразования сегментов временного ряда, построенного с помощью динамического выравнивания временных рядов.

Данная работа состоит из трех частей. Первая часть — это формальная постановка задачи. Во второй части описывается алгоритм преобразования и прогнозирования. Третья часть — вычислительный эксперимент и исследование эффективности алгоритма.

1. Постановка задачи и построение прогностической модели

Для временного ряда требуется построить ретроспективный прогноз — прогноз, который строится в прошлое, т. е. для того периода времени, для которого уже имеются исторические значения.

Временной ряд $X = \{x_t\}_{t=1}^n$ — последовательность, упорядоченная по времени, где n — длина временного ряда, t — номер отсчета.

Данные исходного временного ряда делятся на обучающую $(x_i, t_i)_{i=1}^{n-L}$ и контрольную выборки

$(x_i, t_i)_{i=n-L+1}^n$. На основании обучающей выборки требуется построить ретроспективно-прогностическую модель $\tilde{f} = [\tilde{x}_{n-L+1}, \dots, \tilde{x}_n]$, минимизирующую функцию ошибки:

$$E = \arg \min_{\tilde{x} \in \mathbb{R}^1} \frac{1}{L} \sum_{j=1}^L |x_{n-j} - \tilde{x}_{n-j}|. \quad (1)$$

Разобьем временной ряд на сегменты $\mathbf{S} = \{\mathbf{s}_\beta\}$, где у каждого сегмента будет свой номер. Пусть $\mathbf{s}_0 = \{x_t\}_{t=l}^{l+N}$ — опорный сегмент (сегмент с номером 0), последовательность N отсчетов временного ряда. Требуется построить прогноз следующих h отсчетов временного ряда. Для построения прогноза предлагается выбрать $K + 1$ сегментов таких, которые удовлетворяют требованию их компактного расположения среди всевозможных сегментов временного ряда. Для построения прогноза присоединим к каждому сегменту его продолжение, состоящее из h элементов: $(\mathbf{s} : \tilde{\mathbf{s}})_k$, где знак $:$ обозначает присоединение векторов. Вычислим прогноз сег-

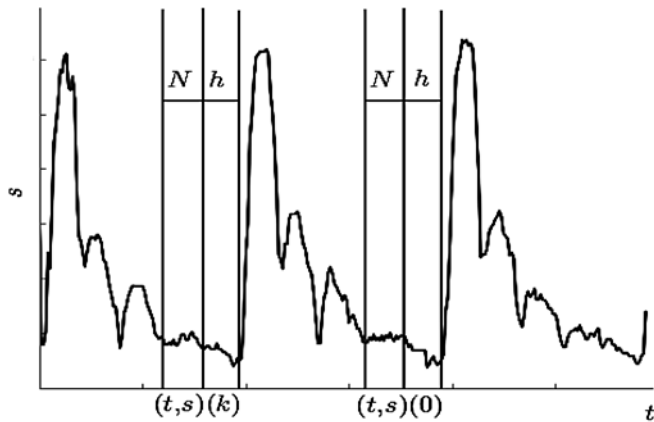


Рис. 1. График, определяемый методом локального прогнозирования

мента с номером $k = 0$ как усреднение присоединенных k сегментов:

$$\tilde{s}_0 = \frac{1}{K} \sum_{k=1}^K \tilde{s}_k.$$

В случае, когда для выполнения требования компактного расположения множества сегментов, по которым выполняется прогноз, предполагаются небольшие монотонные преобразования времени, предлагается сделать следующее. Ввести выравнивающую относительно сегмента s_0 функцию a , и выбирать K сегментов $a(s)_k$, удовлетворяющих требованиям компактности. Рассмотреть соединенные сегменты $a(s : \tilde{s})_k$. Прогноз вычислить как усреднение выровненных сегментов:

$$\tilde{s}_0 = \frac{1}{K} \sum_{k=1}^K a(\tilde{s})_k.$$

Рис. 1 иллюстрирует метод локального прогнозирования.

Опишем альтернативные алгоритмы построения прогностической модели.

1. *Многокластерная модель.* Сегменты временного ряда кластеризуются, затем определяется, к какому кластеру относится прогнозируемый сегмент. Находятся ближайшие сегменты в этом кластере. Происходит преобразование сегментов — сегменты выравниваются относительно центра кластера.

2. *Модель ближайшего соседа.* Во временном ряду ищутся сегменты, похожие на опорный сегмент. Найденные сегменты выравниваются относительно опорного сегмента.

Прогноз вычисляется как взвешенное среднее найденных сегментов.

2. Кластеризация сегментов

Теперь более подробно опишем каждый из этапов построения прогностической модели. Задано множество сегментов S и число кластеров K . Случайным образом выбираются сегменты — исходные центры кластеров $\{\mu_k^0\}$, $k = 1 \dots K$ (верхний индекс — номер итерации алгоритма). Решим задачу

$$Q = \frac{\sum_{i=1}^m [s_i \in C_k] \rho(s_i, \mu_k)}{\sum_{i=1}^m [s_i \in C_k]} \rightarrow \min,$$

где индикаторная функция $[\cdot]$ определена как

$$[s_i \in C_k] = \begin{cases} 1, & \text{если } s_i \in C_k; \\ 0, & \text{если } s_i \notin C_k. \end{cases}$$

Выполним следующие шаги: для каждого сегмента s_i вычислим расстояние ρ до центра μ_k каждого кластера, сегмент относится к тому кластеру, расстояние до центра которого оказалось наименьшим. Затем вычислим центр каждого кластера по формуле

$$\mu_k^h = \frac{1}{|C_k^h|} \sum_{s \in C_k^h} s,$$

где C_k^h — k -й кластер на шаге h , а $|C_k^h|$ — его мощность. До тех пор, пока не прекратится изменение положения центров кластеров, переносим этот центр кластера μ_k^h в найденный вектор μ_k^{h+1} . После выполнения алгоритма кластеризации множество сегментов S разбито на K кластеров.

3. Инвариантное преобразование и путь наименьшей стоимости

Инвариантным преобразованием назовем такое преобразование a сегмента s_i , которое сохраняет эквивалентность на кластерах, т. е. если $s_i \in C_k$ то и $a(s_i) \in C_k$.

Во временном ряду задан эталонный сегмент, относительно которого трансформируются остальные сегменты ряда. Сегменты трансформируются с помощью инвариантного преобразования. Выбор эталонного сегмента зависит от динамики временного ряда. Для модели ближайшего соседа эталонным сегментом является опорный. Для построения прогностической модели все сегменты трансформируются относительно опорного сегмента. В многокластерной модели похожие сегменты объединяются в кластер. Сегменты трансформируются относительно эталонного сегмента — центра кластера.

Обозначим v — эталонный сегмент, а $\{s_i\}_{i=1}^n$ — сегменты, которые требуется трансформировать относительно эталонного. Каждый сегмент s_i трансформируем относительно v . Введем понятие пути наименьшей стоимости между эталонным сегментом v и сегментом s_i :

$$s_i = [x_{i1}, \dots, x_{ij}]^T, v = [x_{v1}, \dots, x_{vj}]^T.$$

Зададим матрицу Ω с элементами-парами из $(l \times l)$ — декартова произведения, квадрата множества $L^2 = \{1, \dots, l\}^2$. Обозначим путь π в матрице Ω — последовательностью

$$\pi = (\pi_1(1), \pi_2(1)), (\pi_1(2), \pi_2(2)), \dots, (\pi_1(N), \pi_2(N)),$$

где N — длина пути, которая удовлетворяет условию $l \leq N < 2l - 1$.

Путь удовлетворяет следующим условиям.

1. *Граничные условия.* Начало и конец пути π находятся на диагонали в противоположных углах Ω , т. е.

$$1 = \pi_1(1) \leq \pi_1(2) \leq \dots \leq \pi_1(N),$$

$$1 = \pi_2(1) \leq \pi_2(2) \leq \dots \leq \pi_2(N).$$

2. *Непрерывность.* В шаге пути π участвуют только соседние элементы матрицы, включая соседние по диагонали; $\pi_1(j+1) \leq \pi_1(j) + 1$ и $\pi_2(j+1) \leq \pi_2(j) + 1$.

3. *Монотонность.* Точки π монотонно перемещаются во времени:

$$(\pi_1(j+1) - \pi_1(j)) + (\pi_2(j+1) - \pi_2(j)) \geq 1.$$

Определение 1. Стоимостью пути наименьшей стоимости $\text{Cost}(\pi, \mathbf{s}_i, \mathbf{v})$ между сегментами \mathbf{s}_i и \mathbf{v} назовем

$$\text{Cost}(\pi, \mathbf{s}_i, \mathbf{v}) = \frac{1}{N} \sum_{i, i' \in \pi} \Omega_{i, i'}$$

Знаменатель N нужен для того, чтобы учесть длину пути π .

Определение 2. Путь наименьшей стоимости $\hat{\pi}$ между сегментами \mathbf{s}_i и \mathbf{v} — это путь, имеющий наименьшую стоимость среди всех возможных путей между сегментами \mathbf{s}_i и \mathbf{v} :

$$\hat{\pi} = \arg \min_{\pi} \text{Cost}(\pi, \mathbf{s}_i, \mathbf{v}).$$

Для получения пути наименьшей стоимости в работе использован метод динамического выравнивания, который рекурсивно находит длину пути наименьшей стоимости по матрице γ , элементы которой определяются следующим образом:

$$\gamma_{i,j} = d_{L_p}(\mathbf{s}_i, \mathbf{v}) + \min(\gamma_{i,j-1}, \gamma_{i-1,j}, \gamma_{i-1,j-1}).$$

Определение 3. Назовем проекциями пути наименьшей стоимости π_x, π_y отображение подмножества $\hat{\pi}$ декартова произведения L^2 в его сомножители L .

Обозначим $L = \{1 \dots l\}$ — множество индексов сегмента \mathbf{s}_i и $P = [1 \dots N]$ — множество индексов пути наименьшей стоимости $\hat{\pi}$. Определим f и g — сюръективные отображения элементов множества P в элементы множества L :

$$f, g : P \rightarrow L.$$

Определим $F : L \rightarrow 2^P$ как функцию, возвращающую все преобразы индексов пути

$$F(j) = \{p | f(p) = j\}, \text{ для любого } j \in L: F(j) \neq \emptyset.$$

Определим сюръективное отображение $G : 2^P \rightarrow L$, $G(j) = \{g(p) | p \in j\}$.

Определим функцию усреднения по значениям полученных индексов

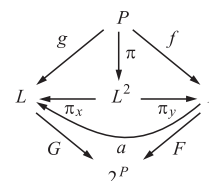
$$\text{avg} : L \rightarrow \mathbb{R},$$

$$\text{где } \text{avg}(G) = \frac{\left(\sum_{j \in G} \mathbf{s}_j \right)}{|G|}.$$

Таким образом, итоговое преобразование a сегмента \mathbf{s}_i к эталонному сегменту \mathbf{v} — композиция:

$$a = \text{avg} \circ G \circ F : \mathbf{s}_i \mapsto \hat{\mathbf{s}}_i.$$

Покажем преобразование a на следующей коммутативной диаграмме:



4. Алгоритмы построения прогностической модели

4.1. Многокластерная модель

В каждом кластере C_k прогностическую модель \tilde{f} строим путем усреднения преобразованных сегментов.

1. Найдем ближайший к опорному сегменту \mathbf{s} центр кластера μ_k :

$$k^* = \arg \min_{k \in 1 \dots K} \rho(\mu_k, \mathbf{s}).$$

Для этого сегмента требуется построить прогноз.

2. Выполним преобразование a опорного сегмента \mathbf{s} и всех сегментов $\mathbf{s}_{k_i^*}$ в кластере C_{k^*} .

3. Преобразуемый сегмент находим следующим образом:

$$\tilde{\mathbf{s}}_j = a_s(a_{\hat{\mathbf{s}}_{k_i^*}}(\mathbf{s}_{k_{i+1}^*})).$$

4. Прогноз вычислим как среднее m найденных преобразованных сегментов $\tilde{\mathbf{s}}_j$, где среднее определяем как взвешенное среднее арифметическое. Формула для вычисления весов введена в работе [18]:

$$\tilde{\mathbf{s}} = \frac{\sum_{j=1}^m w_j \mathbf{s}_j}{\sum_{j=1}^m w_j}, \text{ где } w_j = \left(1 - \frac{\rho_{ij}^2}{\rho_{i_{m+1}}^2} \right)^2, \quad (2)$$

где $\rho_{i_{m+1}}^2$ — расстояние до $(m+1)$ -го ближайшего сегмента.

4.2. Модель ближайшего соседа

1. Находим во временном ряде m сегментов \mathbf{s}_{i_m} , похожих на опорный \mathbf{s} . В качестве функции расстояния берем стоимость пути наименьшей стоимости.

2. Прогноз вычисляем по формуле (2).

5. Вычислительный эксперимент

Проиллюстрируем предложенный алгоритм построения прогностической модели на данных физической активности человека. Цель: сравнить точность построения прогноза с введением и без введения инвариантного преобразования.

Рассмотрим временные ряды, полученные с акселерометра мобильного устройства. Построим прогноз движения человека. В качестве данных возьмем временной ряд, содержащий показания акселерометра. График временного ряда изображен на рис. 2.

Временной ряд состоит из четырех последовательных действий человека: ходьба вперед, подъем, спуск и опять движение вперед. Диапазон измерений используемого акселерометра — $\pm 6g$, диапазон частот — 100 Гц. Изначально временной ряд содержал показания акселерометра в трех пространственных направлениях; для удобства была проведена нормировка ряда.

Сравним алгоритмы построения прогностической модели с кластеризацией и с использованием метода ближайшего соседа с алгоритмом, который не использует никаких преобразований. На рис. 3 (см. четвертую сторону обложки) показаны полученные прогнозы с использованием двух алгоритмов (красным и зеленым цветами) и исторические данные (синим цветом), а также прогноз, полученный без применения преобразований (розовым цветом).

Сравним результаты работы алгоритмов построения прогностической модели с использованием кластеризации и с использованием метода ближайшего соседа. На рис. 4 (см. четвертую сторону обложки) показаны полученные прогнозы с использованием двух алгоритмов (красным и зеленым цветами) и исторические данные (синим цветом).

Алгоритм построения прогностической модели с использованием метода ближайшего соседа показал себя точнее (таблица). Построим прогноз временного ряда по опорному сегменту. На рис. 5 (см. четвертую сторону обложки) зеленым выделен

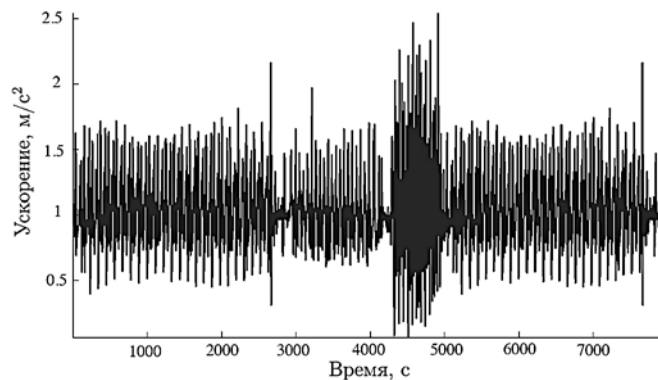


Рис. 2. Графики ходьбы вперед, подъема вверх, спуска вниз, ходьбы вперед

опорный сегмент, синим — исторические данные, красным — полученный прогноз.

На рис. 6 (см. четвертую сторону обложки) показан модуль ошибки (1) на прогнозируемом сегменте по сравнению с историческими данными.

В таблице приведены сравнительные результаты работы алгоритмов построения прогностической модели на исследуемых данных в зависимости от разных длин прогнозируемых сегментов (N — число экспериментов; L — средняя длина прогнозируемого сегмента).

Сравнение результатов работы алгоритмов на данных, полученных с акселерометра

Используемый алгоритм	Лучшая MAPE, %	Средняя MAPE, %	N	L
Многокластерная модель	12,88	15,69	7	100
Модель ближайшего соседа	2,63	5,88	7	100

Точность прогноза зависит от выбора значений параметров — числа кластеров; числа ближайших соседей, длины сегмента, длины предыстории. Оптимальная длина сегмента зависит от природы задачи, например, от частоты данных по времени и характерных времен разных типов процессов. Все эти параметры настраиваются кросс-валидацией в зависимости от данных. Приведенный выше вычислительный эксперимент может быть воспроизведен программным обеспечением [19].

Заключение

Предложен метод локального прогнозирования временных рядов. Предложено инвариантное преобразование сегментов временного ряда для повышения точности прогноза. Преобразование основано на методе динамического выравнивания временных рядов. Обоснован выбор эталонного сегмента. Реализованы различные алгоритмы построения прогностических моделей с использованием инвариантных преобразований. Алгоритмы протестированы на реальных данных движения человека, сравнительные результаты сведены в таблицу.

Список литературы

1. **McNames J.** Innovations in local modeling for time series prediction: Ph. D. thesis, Stanford University, 1999. 161 p.
2. **Hartigan J. A., Wong M. A.** Algorithm as 136: A k-means clustering algorithm // Applied statistics. 1979. Vol. 28, N. 1. P. 100—108.
3. **Loochach R., Garg K.** Effect of distance functions on simple k-means clustering problem // International Journal of Computer Applications. 2012. Vol. 49, N. 6. P. 7—9.
4. **Воронцов К. В.** Лекции по алгоритмам кластеризации и многомерного шкалирования. URL: <http://www.ccas.ru/voron/download/Clustering.pdf> (дата обращения: июнь 12, 2014)
5. **Wang K., Gasser T.** Aligment of curves by dynamic time warping // The Annals of Statistics. 1997. Vol. 25, N. 3. P. 1251—1276.
6. **Keogh E. J., Pazzani M. J.** Derivative dynamic time warping // In First SIAM International Conference on Data Mining, 2001.

7. Айвазян С. А., Мхитарян В. С. Прикладная статистика. Основы эконометрики. Т. 2. М.: ЮНИТИ-ДАТА, 2001. 432 с.
8. Abonyi J., Feil B., Nemeth S., Arva P. Fuzzy Clustering Based Segmentation of Time-Series // 5th International Symposium on Intelligent Data Analysis. Berlin. 2003. Vol. 2810. P. 275—285.
9. Jebara T., Song Y., Thadani K. Spectral Clustering and Embedding with Hidden Markov Models // Lecture Notes in Computer Science. 2007. Vol. 4701. P. 164—175.
10. Coviello E., Antoni B. Chan, Gert R. G. Lanckriet. Clustering Hidden Markov Models with Variational HEM // Journal of Machine Learning Research. 2014. Vol. 15. P. 697—747.
11. Варфоломеева А. А. Локальные методы прогнозирования с выбором метрики // Машинное обучение и анализ данных. 2012. Т. 1, №. 3. С. 367—375.
12. Iglesias F., Kastner W. Analysis of Similarity Measures in Times Series Clustering for the Discovery of Building Energy Patterns // Energies. 2013. Vol. 6. P. 579—597.
13. Романенко А. А. Выравнивание временных рядов: прогнозирование с использованием DTW. URL: <http://rlu.ru/5K7n> (дата обращения: июнь 12, 2014).
14. Цыганова С. В. Локальные методы прогнозирования с выбором преобразования // Машинное обучение и анализ данных. 2012. Т. 1, №. 3. С. 311—317.
15. Кононенко Д. С. Оценка параметров инвариантных преобразований в задачах прогнозирования временных рядов. Магистерская диссертация, Московский физико-технический институт, 2013, 24 с. URL: <http://rlu.ru/5K7g> (дата обращения: июнь 22, 2014).
16. Arribas-Gil A., Muller H.-G. Pairwise dynamic time warping for event data // Computational Statistics Data Analysis. 2014. Vol. 69. P. 255—268.
17. Tang R., Muller H.-G. Pairwise curve synchronization for functional data // Biometrika. 2008. Vol. 95, N. 4. P. 875—889.
18. Магнус Я. Р., Катышев П. К., Пересецкий А. А. Эконометрика. Т. 2. М.: Дело, 2014, 576 с.
19. Кузнецова М. В. Алгоритм локального прогнозирования временных рядов. URL: <http://rlu.ru/5K79> (дата обращения: декабрь 25, 2015).

M. V. Kuznetsova¹, Graduate Student, V. V. Strijov², Leading Researcher

¹ Moscow Institute of Physics and Technology, Moscow;

² Computing Centre of the Russian Academy of Sciences, Moscow

Local Forecasting of Time Series with Invariant Transformations

The paper describes a univariate time series forecasting model. It proposes to find — segments of local history, which are similar to the forecasted segment. A distance function is used to cluster segments. The forecast is the average of the value of time series from this cluster. To improve the quality of forecast the paper proposes an invariant transformation of segments. This transformation holds the equivalence of time series respect to clusters. The transformation is a function, constructed by the dynamic time warping procedure. The retrospective forecasting procedure calculates the accuracy of the forecasting model. Accelerometer time series of a person's motion are used in computational experiment. It compares two constructing forecasting models. The first one clusters segments, the second one uses k-nearest neighbor algorithm to select similar segments.

Keywords: clustering, k-nearest neighbor algorithm, dynamic time warping, warping path, dynamic aligning, local forecasting, monotonic transformation, retrospective forecast, supported segment

References

1. McNames J. Innovations in local modeling for time series prediction: Ph. D. thesis, Stanford University, 1999, 161 p.
2. Hartigan J. A., Wong M. A. Algorithm as 136: A k-means clustering algorithm, *Applied statistics*, 1979, vol. 28, no. 1, pp. 100—108.
3. Loochach R., Garg K. Effect of distance functions on simple k-means clustering problem, *International Journal of Computer Applications*, 2012, vol. 49, no. 6, pp. 7—9.
4. Vorontsov K. V. Lektsii po algoritmam klasterizatsii i mnogomernogo shkalirovaniya, URL: <http://www.ccas.ru/voron/download/Clustering.pdf> (accessed: June 12, 2014).
5. Wang K., Gasser T. Aligment of curves by dynamic time warping, *The Annals of Statistics*, 1997, vol. 25, no. 3, pp. 1251—1276.
6. Keogh E. J., Pazzad M. J. Derivative dynamic time warping, *In First SIAM International Conference on Data Mining*, 2001.
7. Aivazjan S. A., Mhitarjan V. S. *Prikladnaya statistika. Osnovy ekonometriki*, Т. 2, Moscow, YuNITI-DATA, 2001. 432 p.
8. Abonyi J., Feil B., Nemeth S., Arva P. Fuzzy Clustering Based Segmentation of Time-Series, *5th International Symposium on Intelligent Data Analysis*, Berlin, 2003, vol. 2810, pp. 275—285.
9. Jebara T., Song Y., Thadani K. Spectral Clustering and Embedding with Hidden Markov Models, *Lecture Notes in Computer Science*, 2007, vol. 4701, pp. 164—175.
10. Coviello E., Antoni B. Chan, Gert R. G. Lanckriet. Clustering Hidden Markov Models with Variational HEM, *Journal of Machine Learning Research*, 2014, vol. 15, pp. 697—747.
11. Varfolomeeva A. A. Lokalnye metody prognozirovaniya s vyborom metriki, *Journal of Machine Learning and Data Analysis*, 2012, vol. 1, 3, pp. 367—375.
12. Iglesias F., Kastner W. Analysis of Similarity Measures in Times Series Clustering for the Discovery of Building Energy Patterns, *Energies*, 2013, vol. 6, pp. 579—597.
13. Romanenko A. A. *Vyiravnivanie vremennykh ryadov: prognozirovanie s ispolzovaniem DTW*, URL: <http://rlu.ru/5K7n> (accessed: June 12, 2014).
14. Tsyganova S. V. Lokalnye metody prognozirovaniya s vyborom preobrazovaniya, *Journal of Machine Learning and Data Analysis*, 2012, vol. 1, 3, pp. 311—317.
15. Kononenko D. S. *Otsenka parametrov invariantnykh preobrazovaniy v zadachah prognozirovaniya vremennykh ryadov*: Masters thesis, Moscow Institute of Physics and Technology, 2013, 24 p. URL: <http://rlu.ru/5K7g> (accessed: June 22, 2014).
16. Arribas-Gil A., Muller H.-G. Pairwise dynamic time warping for event data, *Computational Statistics Data Analysis*, 2014, vol. 69, pp. 255—268.
17. Tang R., Muller H.-G. Pairwise curve synchronization for functional data, *Biometrika*, 2008, vol. 95, no. 4, pp. 875—889.
18. Magnus Ya. R., Kartushev P. K., Peresetskii A. A. *Ekonometrika*, Moscow, Delo, 2004. 576 p.
19. Kuznetsova M. V. *Algorithm of local modelling*, Available at <http://rlu.ru/5K79> (accessed: December 25, 2015).

А. Н. Четырбоцкий, д-р физ.-мат. наук, вед. науч. сотр., e-mail: Chetyrbotsky@yandex.ru
 Дальневосточный геологический институт ДВО РАН, г. Владивосток
 Дальневосточный федеральный университет, г. Владивосток

Автоматическая классификация объектов многомерной выборки рекурсивными методами построения кривых Пеано (на примере выборки Р. Фишера)

На основании рекурсивных методов построения кривых Пеано разработан алгоритм автоматической классификации объектов многомерной выборки. Сущность разработки состоит в согласовании между собой соответствующих распределений образов объектов на отрезке вещественной оси. Для оценки работоспособности алгоритма используется общедоступная для использования выборка 150 четырехмерных наблюдений трех разновидностей цветков ириса (выборка Р. Фишера). Приведены результаты работы алгоритма и выполнен их анализ.

Ключевые слова: многомерная выборка, кластерный анализ, заполняющие пространство кривые

Введение

Многообразие современных явлений и систем в различных дисциплинах актуализирует проблемы их автоматической количественной систематизации (без вмешательства человека и отсутствии обучающих выборок) многомерных объектов. Такая ситуация обусловлена гигантским ростом объемов разнородной эмпирической информации. Для их согласования и углубленной обработки требуются значительные вычислительные ресурсы. Так, в некоторых задачах астрофизики число объектов (планет, метеоритов и других небесных тел) может превышать 10^{10} соответствующих единиц [1]. Значимость результатов классификации здесь обусловлена еще и тем, что ассоциации звезд в нашей Галактике обычно незаметны на фотографиях. Они выделяются из фона лишь как сгущения/кластеры звезд определенного типа [2]. В этой ситуации под кластером/классом объектов обычно понимают изолированную в исходном признаковом пространстве компактную совокупность объектов выборки [3, 4].

Эффективный инструментальный анализа выборочного материала основан на применении соответствующих так называемых кривых Пеано, под которыми понимается любое непрерывное отображение числового отрезка на плоский квадрат [4] (в связи с ее другим определением как непрерывного образа отрезка, полностью заполняющего квадрат, в работе [5] соответствующая кривая именуется "заполняющая пространство кривая" или ЗПК). Такая кривая была впервые построена Дж. Пеано [6], а существенно более простой ее вариант — Дж. Гилбертом [7]. Р. Г. Стронгин разработал оригинальный алгоритм ее построения [8], а В. В. Александров и Н. Д. Горский выполнили авторскую компьютерную реализацию построения [9]. Кривые Пеано имеют широкую область практического применения: распознавание образов [6, 10]; поиск глобаль-

ного оптимума многоэкстремальных функций [11]; работа с базами данных [12] и т. д. Интересно заметить, отмеченный в работе [13], такой факт: согласно современным исследованиям, ДНК заполняет каждую клетку таким образом, что ее пространственная конструкция приближается к структуре кривой Пеано.

Широкие возможности кривых Пеано для решения многообразных практических задач в меньшей степени адаптированы для решения задач автоматической классификации. Целью работы является разработка алгоритма решения этой задачи на основании рекурсивных методов построения кривых Пеано. На примере общедоступной выборки Р. Фишера [14—16] разработан алгоритм автоматического выделения в признаковом пространстве многомерной выборки сравнительно таких кластеров ее объектов. Сущность предлагаемого алгоритма состоит в согласовании между собой результатов нелинейных отображений выборочных точек признакового пространства (объектов выборки) на отрезок вещественной оси с помощью рекурсивного построения кривых Пеано.

Выборка данных

Поскольку выборка Р. Фишера уже стала традиционным полигоном для разработок и тестирования различных методов анализа данных (в системе MATLAB эта выборка представлена набором iris.dat, а демопрограмма их анализа — модулем irisfcm.m [16]), то представляется полезным детальное описание ее структуры.

В работе [14] приведены результаты линейного дискриминантного анализа выборки, объектами которой выступают 150 четырехмерных наблюдений трех разновидностей цветков ириса (*Setosa*, *Versicolor* и *Virginica*), где каждый из них представлен 50 наблюдениями. Отдельное наблюдение ха-

рактически характеризуется четырьмя количественными признаками: длиной X_1 и шириной X_2 чашелистиков, длиной X_3 и шириной X_4 лепестков. Deskриптивные статистики, матрица корреляции признаков, коэффициенты корреляции переменных и главных компонент (ГК) на переменные и их значимость (графа %) приведены в табл. 1 (вычисления ГК в MATLAB выполняет программа `princomp.m`).

Наименьшая вариабельность X_2 отражается в ее низком влиянии на кластерную структуру выборки. Для X_3 имеет место обратная ситуация, иллюстрацией чему являются распределения на рис. 1 (см. третью сторону обложки).

Распределение выборочных точек на рис. 1, а показывает, что в пространстве $\{X_1, X_2, X_4\}$ имеет место значительное перекрытие выборочных объектов 2-го и 3-го видов ириса. На рис. 1, б оно практически отсутствует. В обоих случаях объекты 1-го вида образуют их компактное изолированное существование.

Вследствие высокой вариации X_3 она вносит наибольший вклад в неравномерность распределения

Таблица 1

Deskриптивные статистики выборки

	X_1	X_2	X_3	X_4	%
\bar{X}	5,833	3,061	3,752	1,197	—
σ_X	0,825	0,437	1,761	0,759	—
ν_X	0,681	0,191	3,099	0,577	—
X_1	1	-0,108	0,861	0,803	—
X_2	-0,108	1	-0,429	-0,364	—
X_3	0,861	-0,429	1	0,958	—
X_4	0,803	-0,364	0,958	1	—
C_1	0,887	-0,399	0,997	0,963	92,1
C_2	0,410	0,813	-0,051	-0,055	5,5
C_3	-0,196	0,405	0,011	0,209	2
C_4	0,063	-0,126	-0,045	0,163	0,4

Примечание: первые три строки характеризуют средние значения (измерения в см) X_i , стандартные отклонения σ_X и коэффициенты вариации ν_X ; последующие четыре строки — матрица корреляции; последние четыре строки — коэффициенты корреляции $r(X_i, C_k)$ переменной X_i и главной компоненты C_k ; последний столбец — значимость ГК (отношение собственного значения соответствующей матрицы алгоритма нахождения ГК к их сумме).

Таблица 2

Диапазоны изменения переменных, см

	1	2	3
X_1	5,006 ± 0,598	5,936 ± 0,876	6,558 ± 1,118
X_2	3,428 ± 0,643	2,770 ± 0,533	2,982 ± 0,557
X_3	1,460 ± 0,301	4,260 ± 0,797	5,536 ± 0,930
X_4	0,246 ± 0,179	1,326 ± 0,336	2,018 ± 0,467

Примечание: в столбцах характеристики отдельных видов цветков.

объектов выборки. В табл. 2 представлены диапазоны изменения переменных в формате $\langle X_{ik} \rangle \pm t_{\alpha} \sigma_{ik}$, где первый член среднее значение i -й переменной k -го вида, t_{α} — t -статистика Стьюдента уровня значимости α (следуя [17], при таком объеме выборки и $\alpha = 0,005$ она принимается равной 1,96) и σ_{ik} — соответствующее среднее квадратичное отклонение.

Сопоставление элементов столбцов показывает перекрытие диапазона изменения X_1 для всех разновидностей. Такая же ситуация отмечается и для X_2 . Вследствие чего в рамках этих переменных цветки не разделяются по их разновидностям. В переменных X_3 и X_4 значимо выделяются цветки 1-го вида ирисов, что обусловлено значимыми отличиями диапазонов их изменений.

Переменная X_2 имеет наименьшие корреляционные связи с остальными переменными, а переменная X_3 — наибольшие. При этом только она коррелирует со второй ГК (коэффициент корреляции 0,883). Такая ситуация отражается в их корреляционных связях с ГК. В частности, среди всех переменных только X_2 коррелирует со второй ГК: коэффициент корреляции $r(X_2, C_2) = 0,833$, а $r(X_2, C_1) = -0,399$. Коэффициенты корреляции C_2 с остальными переменными не превышают 0,410. Высокие значения $r(X_1, C_1) = 0,889$, $r(X_3, C_1) = 0,998$ и $r(X_4, C_1) = 0,963$ указывают на характер изменения этой ГК: ее рост обусловлен ростом этих переменных. На рис. 2 (см. третью сторону обложки) представлена конфигурация проекций объектов выборки на плоскость первых двух ГК.

Распределение маркеров объектов выборки на плоскости первых двух ГК указывает существенное различие в рассматриваемом признаковом пространстве ирисов 1-го вида (*Setosa*). Ирисы других видов (*Versicolor* и *Virginica*) сосредоточены в отдельном эллипсоидальном вытянутом достаточно плотном облаке. Конфигурация такого распределения указывает, что исходная выборка определяется двумя кластерами: первый кластер составляют ирисы *Setosa*, а второй кластер — ирисы совокупности *Versicolor* и *Virginica*.

Автоматическая классификация методами кривых Пеано

Автоматическая классификация объектов выборки есть процедура разбиения выборки на отдельные достаточно однородные (в смысле заранее заданного критерия) кластеры/классы (группы), выявлении ее структуры при отсутствии обучающих выборок [3]. Представляется, что эффективный инструментальный решения этой проблемы можно построить на основании последовательности построений кривых Пеано. С их помощью выполняется отображение исходных выборочных точек n -мерного гиперкуба на отрезок вещественной оси. Суть построений кривых состоит в равномерном регулярном разбиении гиперкуба выборочных точек

на составляющие (элементы разбиения или гиперкуванты в терминологии работы [5]) и проведение определенным образом через их центры кусочно-ломаной без самопересечений кривой. Выборочные точки элемента разбиения предписываются его центру. Поскольку число разбиений является входным параметром такой процедуры, то с помощью кривых Пеано устанавливается иерархия распределения объектов выборки. Для случая $n = 2$ приближения к так называемым разверткам Пеано (кривым Пеано) представлены на рис. 3.

При таком отображении близкие на отрезке вещественной оси образы выборочных точек близки и в исходном пространстве. Однако близким в исходном пространстве точкам могут соответствовать

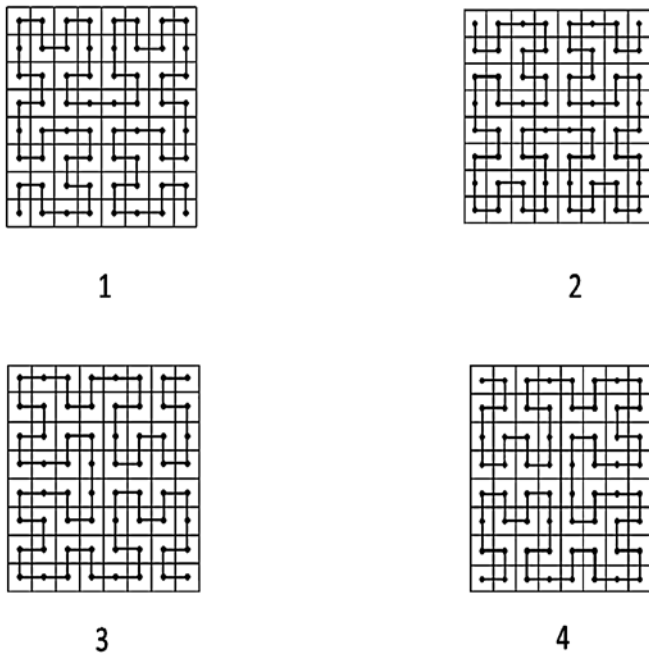


Рис. 3. Приближения к кривым Пеано (черными кружками отмечены узлы сетки, цифрами — варианты вращений)

достаточно далекие их образы. Действительно, точки отрезка имеют левого и правого соседей, а точки пространства R^n имеют соседей по 2^n направлениям. Поэтому для выявления кластерной структуры многомерной выборки следует выявить в этом пространстве смежные элементы разбиения с выборочными точками (под смежностью понимается наличие между элементами разбиения общей грани).

Для выявления смежных элементов разбиения в работе [9] выполняются отображения гиперкубов, которые получены из исходного гиперкуба некоторым сдвигом позиций выборочных точек. Недостаток такого подхода состоит в трудности формализации направления и численного значения такого сдвига, а также количества требуемых отображений. Здесь решение этой проблемы следует ортогональным вращениям исходного гиперкуба вокруг начала координат. Суммарное число поворотов развертки при отображении n -мерного гиперкуба на одномерный отрезок равно 2^n [18]. Матрица вращения определяется выражением

$$A = \begin{pmatrix} \cos\theta & \mp\sin\theta \\ \pm\sin\theta & \cos\theta \end{pmatrix},$$

где θ — угол поворота. При двумерном признаковом пространстве число таких вращений равно четырём (см. рис. 3). Поэтому здесь $\theta = 0, \pi/2, \pi, 3\pi/2$.

Распределения маркеров объектов на рис. 4 (см. третью сторону обложки) показывает, что они заполняют 12 элементов разбиения. При этом конфигурация образов выборочных точек относительно друг друга не зависит от вращения (тут уместна аналогия с распределениями точек абсолютно твердого тела при его движении [19]). Анализ случаев подтверждает достоверность наличия в исследуемой выборке двух кластеров. Один из них составляют ирисы *Setosa*, а второй — ирисы совокупности *Versicolor* и *Virginica*.

Здесь для автоматической классификации объектов предлагается такой алгоритм. При его построении учитывается тот факт, что при различных отображениях узор конфигурации и распределение выборочных точек не изменяются. Для каждого результата вращения выполняется отображение выборочных точек на отрезок вещественной оси.

Далее образы элементов разбиения упорядочиваются согласно номерам элементов разбиения первого отображения. Например, если при первом отображении интервал отрезка соответствует i -му номеру исходного гиперкуба, то образы выборочных точек этого гиперкуба при других отображениях также получают этот номер. Сведенные в единую таблицу (табл. 3) эти номера позволяют определить смежность гиперкубов. К одному кластеру приписывают те выборочные точки, которые принадлежат только смежным гиперкубам. На диагонали таблицы указываются номера кластеров.

Таблица 3

Распределение общего числа случаев смежности гиперкувантов

\	1	2	3	4	5	6	7	8	9	10	11	12
1	1	6	0	6	0	0	0	0	0	0	0	0
2	6	1	6	0	0	0	0	0	0	0	0	4
3	0	6	1	6	0	0	0	0	0	0	2	0
4	6	0	6	2	4	0	0	0	0	0	0	0
5	0	0	0	4	2	6	0	0	0	0	0	0
6	0	0	0	0	6	2	0	0	0	0	0	0
7	0	0	0	0	0	0	1	0	6	0	2	0
8	0	0	0	0	0	0	0	1	6	0	0	0
9	0	0	0	0	0	0	6	6	1	4	0	0
10	0	0	0	0	0	0	0	0	4	1	6	0
11	0	0	2	0	0	0	2	0	0	6	1	6
12	0	4	0	0	0	0	0	0	0	0	6	1

Примечание: элементы таблицы n_{ij} — суммарное число общих граней (смежность) i -го и j -го элементов разбиения при отображениях их выборочных точек; элементы диагонали — номера кластеров.

Заклучение

В работе разработан алгоритм автоматической классификации выборочного материала на основании рекурсивного метода построения кривой Пеано. Работоспособность алгоритма иллюстрируется результатом решения этой проблемы на примере общедоступной выборки Р. Фишера 150 четырехмерных наблюдений.

Список литературы

1. **Springel V.** The cosmological simulation code GADGET-2 // *Mon. Not. R. Astron. Soc.* 2005. 364, pp. 1105–1134.
2. **Ефремов Ю. Н., Чернин А. Д.** Крупномасштабное звездообразование в галактиках // *УФН.* 2003. Т. 173, № 1. С. 3–25.
3. **Фукунага К.** Введение в статистическую теорию распознавания образов. М.: Наука, 1979. — 368 с.
4. **Лузин Н. Н.** Теория функций действительного переменного. 2-е изд. М.: Учпедгиз, 1948. 321 с.
5. **Александров В. В., Горский Н. Д., Поляков А. О.** Рекурсивные алгоритмы обработки и представления данных. — Препр. Ленинград: НИВЦ АН СССР, 1979. 53 с.
6. **Peano G.** Sur une courbe, qui remplit toute une aire plane // *Math. Ann.*, 1890, 36 (1). P. 57–160.
7. **Hilbert D.** Über die stetige Abbildung einer Linie auf ein Flächenstück // *Math. Annln.*, 1891. 38. P. 459–460.
8. **Стронгин Р. Г.** Численные методы решения многоэкстремальных задач (информационно-статистические алгоритмы). М.: Наука, 1978. 240 с.
9. **Александров В. В., Горский Н. Д.** Алгоритмы и программы структурного метода обработки данных. Л.: Наука, 1983. 208 с.
10. **Четырбоцкий А. Н.** Методы и алгоритмы решения задач снижения размерности пространства описания. Владивосток: ДВО АН СССР, 1991. 95 с.
11. **Стронгин Р. Г., Гергель В. П., Баркалов К. А.** Параллельные методы решения задач глобальной оптимизации // *Известия высших учебных заведений. Приборостроение.* 2009. Т. 52, № 1. С. 25–32.
12. **Jin G., Mellor-Crummey J.** Using Space-filling Curves for Computation Reordering // *LACSI*, 2005.
13. **Бауман К. Е.** О квадратно-линейном отношении правильных кривых Пеано. Автореф. ... дис. ... канд. физ.-мат. наук. М.: МГУ, 2012. <http://mech.math.msu.su/~snarkyfiles/vak/arzh2.pdf>
14. **Fisher R. A.** The use of multiple measurements in taxonomic problems // *Ann. Eugen.* 1936-7, 179. P. 466–475.
15. **Кендалл М. Дж., Стьюарт А. Т.** Многомерный статистический анализ и временные ряды. М.: Наука, 1976. 736 с.
16. <http://matlab.exponenta.ru/fuzzylogic/book2/4/irisfcm.php>
17. **Болч Б., Хуань Дж.** Многомерные статистические методы для экономик. М.: Статистика, 1979. 317 с.
18. **Баркалов К. А., Рябов В. В., Сидоров С. В.** О некоторых способах балансировки локального и глобального поиска в параллельных алгоритмах глобальной оптимизации // *Выч. мет. и программирование*, 2010. Т. 11, вып. 4. С. 382–387.
19. **Жилин П. А.** Рациональная механика сплошных сред: учеб. пособие. СПб.: Изд-во политех. ун-та, 2012. 584 с.

A. N. Chetyrbotsky, D. Sc., Leading Researcher, Far East Geological Institute,
The Far Eastern Federal University e-mail: Chetyrbotskv@vandex.ru

Automatic Classification of Objects of Multidimensional Sampling Recursive Method of Constructing Curves of Peano (for Example, Fisher's iris data)

Based on the recursive method of constructing curves of Peano developed an algorithm of automatic classification of objects of multidimensional sampling. The essence of development is to agree among themselves the respective distributions of images of objects on the segment of the real axis. To evaluate the performance of the algorithm used publicly available for use by the sample of 150 observations of four-three species of iris flowers (Fisher's iris data). The results of the algorithm and performed their analysis.

Keywords: multidimensional sampling, cluster analysis, space-filling curves

References

1. **Springel V.** The cosmological simulation code GADGET-2, *Mon. Not. R. Astron. Soc.*, 2005, 364, pp. 1105–1134.
2. **Efremov Ju. N., Chernin A. D.** Krupnomasshtabnoe zvezdoobrazovanie v galaktikah, *UFN*, 2003, vol. 173, no. 1, pp. 3–25.
3. **Fukunaga K.** *Vvedniye statisticheskuyu teoriju raspoznavaniya obrazov*, Moscow, Nauka, 1979, 368 p.
4. **Luzin N. N.** *Teoriya funktsiy dejstvitel'nogo peremennogo*, 2-e izd, Moscow, Uchpedgiz, 1948, 321 p.
5. **Aleksandrov V. V., Gorskij N. D., Poljakov A. O.** Rekursivnye algoritmy obrabotki i predstavleniya dannyh, Prepr. Leningrad, NIVC AN SSSR, 1979, 53 p.
6. **Peano G.** Sur une courbe, qui remplit toute une aire plane, *Math. Ann.*, 1890, 36 (1), pp. 57–160.
7. **Hilbert D.** Über die stetige Abbildung einer Linie auf ein Flächenstück, *Math. Annln.*, 1891, 38, pp. 459–460.
8. **Strongin R. G.** Chislennye metody resheniya mnogoekstremnykh zadach (informacionno-statisticheskie algoritmy), Moscow, Nauka, 1978, 240 p.
9. **Aleksandrov V. V., Gorskij N. D.** Algoritmy i programmy strukturnogo metoda obrabotki dannyh. Leningrad, Nauka, 1983, 208 p.
10. **Chetyrbockij A. N.** *Metody i algoritmy resheniya zadach snizheniya razmernosti prostranstva opisaniya*, Vladivostok, DVO AN SSSR, 1991, 95 p.
11. **Strongin R. G., Gergel' V. P., Barkalov K. A.** Parallelnye metody resheniya zadach global'noj optimizacii, *Izvestiya vysshih uchebnykh zavedenij. Priborostroenie*, 2009, vol. 52, no. 10, pp. 25–32.
12. **Jin G., Mellor-Crummey J.** Using Space-filling Curves for Computation Reordering, *LACSI*, 2005.
13. **Bauman K. E.** O kvadratno-linejnom otnoshenii pravil'nykh krivykh Peano. — avtoref. ... dis. kand. fiz.-mat nauk, Moscow: MGU, 2012. <http://mech.math.msu.su/~snark/files/vak/arzh2.pdf>
14. **Fisher R. A.** The use of multiple measurements in taxonomic problems, *Ann. Eugen.*, 1936-7, 179, pp. 466–475.
15. **Kendall M. Dzh., St'juart A. T.** *Mnogomernyj statisticheskij analiz i vremennye rjady*, Moscow, Nauka, 1976, 736 p.
16. <http://matlab.exponenta.ru/fuzzylogic/book2/4/irisfcm.php>
17. **Bolch B., Huan' Dzh.** *Mnogomernye statisticheskie metody dlja jekonomiki*, Moscow, Statistika, 1979, 317 p.
18. **Barkalov K. A., Rjabov V. V., Sidorov S. V.** O nekotorykh sposobah balansirovki v lokal'nogo i global'nogo poiska parallel'nykh algoritmah global'noj optimizacii, *Vych. Met. Programirovanie*, 2010. — t. 11, vyr. 4. — s. 382–381.
19. **Zhilin P. A.** Racional'naja mehanika sploshnykh sred: ucheb. posobie, Sankt Peterburge, Izd-vo politeh. un-ta, 2012, 584 p.

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ОРГАНИЗАЦИОННЫХ И СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ СИСТЕМАХ

INFORMATION TECHNOLOGIES IN THE ORGANIZATIONAL AND SOCIO-ECONOMIC SYSTEMS

УДК 004.83

М. Г. Мамедова, д-р техн. наук, проф., зав. отделом,

З. Г. Джабраилова, канд. техн. наук, доц., гл. науч. сотр.,

Институт Информационных Технологий НАН Азербайджана, Баку, depart15@iit.ab.az, www.ikt.az

Методологический подход к многокритериальному принятию решений в задачах управления человеческими ресурсами

Разработан методологический подход к принятию решений в задачах управления человеческими ресурсами (УЧР). Приведена обобщенная концептуальная модель принятия решений в задачах УЧР. Для обеспечения адаптивности многокритериального принятия решений в задачах УЧР предложена модификация метода TOPSIS, заключающаяся во включении в алгоритм принятия решений дополнительных компонентов, позволяющих избавиться от иерархической структуры критериев и учесть коэффициенты компетентности экспертов, участвующих в процедуре оценки альтернатив. Приведена апробация метода на примере задачи трудоустройства.

Ключевые слова: управление человеческими ресурсами, принятие решений, нечеткая среда, многокритериальная оптимизация, компетентность экспертов, метод TOPSIS

Введение

В условиях перехода к экономике, основанной на знаниях, эффективная деятельность и конкурентоспособность организации (предприятия, компании, фирмы и т. д.) становится существенно зависимой от человеческого фактора и от правильного выбора политики управления человеческими ресурсами (УЧР) [1, 2]. В то же время процессы глобализации и быстрая смена технологий обуславливают изменения на рынке труда, что, в свою очередь, вызывает серьезные трансформации во взаимоотношениях с персоналом и требует разработки новых концептуальных подходов и научно обоснованных методов в политике регулирования этих отношений в зависимости от конкретной задачи УЧР. Согласно этой концепции УЧР представляет собой особый вид управленческой деятельности. Главным объектом управления в данном случае являются человек и его компетенции, включающие знания, навыки и профессиональные умения, личностные и поведенческие качества, мотивационные установки, интеллектуальный и квалификационный потенциалы последнего, а УЧР направлено на поддержку стратегии деятельности организации в контексте возрастания роли и значимости человеческого фактора [3–5]. Поэтому для принятия адекватных новым условиям решений относительно

планирования персонала, отбора, наема, адаптации к изменяющейся рыночной среде, удержания, увольнения, продвижения, развития, обучения и мотивации кадров лицо, принимающее решение (ЛПР), должно оценить и принять во внимание широкий спектр информации относительно компетенций работников, иметь возможность сравнить претендентов на основе множества разнородных признаков (критериев), осуществить выбор наилучшего решения (кандидата) с учетом множества влияний, предпочтений, интересов и возможных последствий. Все эти особенности задач УЧР определяют их многокритериальный характер. При этом следует принять во внимание также объем, количественный и качественный характер, сложность и противоречивость потока информации, поступающей к ЛПР, которые позволяют идентифицировать задачи УЧР как слабоструктурируемые, для которых построение объективных моделей принципиально невозможно или же достаточно сложно. Наряду с указанными проблемами, возникающими при генерации и выборе управленческих решений, необходимо учесть также предпочтения ЛПР, компетентность (знания, интуицию, опыт и др.) специалистов-экспертов.

В работах [6–9] приведены фундаментальные основы многокритериального выбора, указаны проблемы, с которыми сталкиваются разработчики, и

даны некоторые приложения интеллектуальных систем поддержки принятия решений в УЧР. К разряду основных проблем в задачах принятия решений в многокритериальной постановке можно отнести способы получения, характер и тип информации, методы ее представления и обработки, определение числа рассматриваемых вариантов (альтернатив) и числа описывающих их признаков, иерархическую структурированность последних, технологии представления экспертных знаний и т. п. При этом способность человека меньше ошибаться, работая с вербальными данными, требует выбора методов оперирования с лингвистическими переменными. Поэтому в задачах УЧР для оперирования такими данными целесообразно применение моделей и методов, базирующихся на теории нечетких множеств и нечеткой логики [10, 11]. Преодоление перечисленных сложностей обуславливает необходимость выбора, создания или развития методологических подходов к многокритериальному анализу и принятию решений в УЧР на базе интеллектуальных технологий, методов и компьютерных систем поддержки принятия решений [12].

Цель настоящего исследования — разработка методологического подхода к принятию управленческих решений в задачах УЧР, обладающих такими специфическими особенностями, как многокритериальность и разнородность данных, описывающих задачи УЧР, иерархичность, количественный и качественный характер, неоднозначность критериев, определяющих необходимость учета экспертной оценки их весов, а также влияние уровня компетентности самих экспертов на принятое решение.

1. Многокритериальные методы принятия решений в задачах УЧР: анализ литературы

Анализ литературы показывает, что среди требующих интеллектуальной поддержки процессов принятия решений задач УЧР, к разряду которых относятся управление трудоустройством, аттестация и организация системы вознаграждения персонала, планирование карьеры, формирование резерва, авторами основное внимание уделяется отбору и найму кадровых ресурсов, что связано с наибольшей практической применимостью последних.

На сегодня при решении задач подбора персонала разработчиками предпочтение отдается в основном методам многокритериального принятия решений, в качестве которых можно указать методы анализа дерева решений [13], методы анализа иерархий — АНР (*Analytic Hierarchy Process*) [14—16], методы предпочтения по сходству с идеальными решениями — TOPSIS (*Technique for Order of Preference by Similarity to Ideal Solution*) [17—24], экспертные системы [7, 25] и др.

Так, в работе [13] для разработки эффективного механизма формирования правил по подбору персонала в высокотехнологичные компании исполь-

зуется метод анализа данных, базирующийся на дереве решений и ассоциативных правилах. Не умаляя преимуществ данного подхода, следует отметить такие недостатки последнего, как невозможность генерации правил в слабоструктурируемых областях, где требуются экспертные знания, а также сложность построения оптимального дерева решений.

В работе [14] предложена кадровая система отбора персонала на основе нечеткого анализа иерархических процессов — ФАНР (*Fuzzy Analytic Hierarchy Process*), позволяющая провести оценку альтернатив с использованием как качественных, так и количественных критериев. В работе [15] для снижения субъективности при оценке личностных качеств и ключевых профессиональных навыков претендентов на должность также применяли нечеткую логику и метод анализа иерархических процессов. Согласно авторам работ [14—16] компьютерная система поддержки принятия решений, базирующаяся на ФАНР, устраняет ограничения на объем информации о претендентах и оказывает помощь менеджерам в принятии наилучших решений (кандидата) в нечетких условиях. Однако хотя указанные подходы обеспечивают принятие наилучшего решения среди возможных, тем не менее они не дают возможности сделать отбор альтернативы, предпочтительной по всем критериям, т. е. наиболее близкой к идеальному (оптимальному) решению. Такую возможность предоставляет метод TOPSIS, впервые предложенный в работе [17]. В дальнейшем этот метод нашел свое применение в различных задачах принятия решений, с учетом специфики которых исследователями предложены различные его модификации. Эти модификации сводятся, в основном, к введению группового принятия решений, а также новых метрик для расчета расстояния до идеально положительного и идеально отрицательного решений.

Так, в работе [18] для подбора персонала предложен метод группового принятия решений на базе нечеткого метода TOPSIS, где функции принадлежности альтернатив критериям описаны значениями нечетких лингвистических переменных, представленных в виде нечетких треугольных чисел. В работе [19] для отбора персонала введен дополнительный этап, который позволяет агрегировать разнородную информацию с использованием взвешенного усредненного оператора (*OWA — Ordered Weighted Averaging*). В работе [20] на примере подбора преподавателей университета также предложена модификация метода TOPSIS. В данном случае модификация заключается в том, что в процессе многокритериального принятия решений предлагается новая мера для вычисления расстояния до идеально положительного и идеально отрицательного решений. В работе [21] поддержка принятия решений при отборе наиболее квалифицированных человеческих ресурсов на базе метода TOPSIS

включает новую концепцию рейтинга (упорядочения) альтернатив, заключающуюся в том, что окончательное решение определяется не по близости расстояния альтернатив к идеальному решению, а на основе вето порога (*veto set*), установленного ЛПР. В работе [22] авторами предложено очередное развитие метода TOPSIS, заключающееся во введении в алгоритм принятия окончательного решения на основе вето порога, установленного ЛПР, также относительно важности последних.

Математически обоснованный и сравнительно простой алгоритм расчета интегральных оценок, возможность ввода дополнительных этапов и элементов в процесс принятия решений, предоставляемые методом TOPSIS, в качестве преимуществ последнего обеспечивают его применимость в решении широкого спектра практических задач. В работе [22] модифицированный метод TOPSIS использован для решения задачи выбора менеджеров среднего уровня в греческую ИТ-компанию, в [23] показано применение метода многокритериального нечеткого принятия решений на базе TOPSIS для подбора человеческих ресурсов в крупный греческий банк. В работе [24] для совершенствования процесса отбора и наема персонала в *Khodro company* (Иран) внедрена нечеткая модель системы поддержки принятия решений на основе метода TOPSIS.

Обзор специальной литературы позволил выделить основные направления развития метода TOPSIS. Так, этот метод не позволяет учитывать иерархическую структурированность критериев, являющуюся специфической особенностью, свойственной задачам УЧР. Один из возможных путей решения этой проблемы предлагается в работе [21]. Нами в работах [26–28] также предложены методы решения задач УЧР на основе многокритериальной скалярной оптимизации [29], позволяющие учитывать иерархическую структурированность неравнозначных критериев. Однако этот подход, базирующийся на агрегировании функций принадлежности альтернатив критериям, обладает следующим недостатком: альтернатива, имеющая нулевую принадлежность по некоторым критериям, в результате может оказаться наилучшей [7].

Далее метод TOPSIS не дает возможности оценить уровень компетентности экспертов, участвующих в процессе принятия решений. Хотя ЛПР старается отобрать экспертов примерно с одинаковой компетентностью, однако на практике выполнить это условие трудно, и предпочтения экспертов в той или иной степени воздействуют на принятое решение. Более того, хотя эксперты являются основной фигурой в системе поддержки принятия решений в УЧР и уровень их компетентности в определенной степени влияет на конечный результат, тем не менее вопросам оценки компетентности экспертов, участвующих в процессе принятия решений, в литературе не уделено должного внимания, и эта про-

блема остается одной из слабо проработанных. В обзорной статье [30] также подчеркивается важность решения этой проблемы.

В настоящей работе в качестве метода многокритериальной оптимизации в задачах УЧР выбран достаточно гибкий метод TOPSIS, предоставляющий возможность введения новых компонентов, устраняющих указанные выше его недостатки. Так, предложенная нами модификация TOPSIS заключается в интегрировании в алгоритм принятия решений дополнительных компонентов, обеспечивающих на первом этапе расчет на основе метода анализа иерархий весовых коэффициентов частных критериев, позволяющего избавиться от иерархической структуры критериев, а на втором этапе введение в алгоритм предварительно рассчитанных коэффициентов компетентности экспертов, участвующих в процедуре оценки альтернатив. В целях обеспечения устойчивости критериев к границам интервала достоверности в настоящей работе при использовании метода группового принятия решений функции принадлежности альтернатив критериям описаны значениями нечетких лингвистических переменных, представленных в виде нечетких трапециевидных чисел.

2. Специфика задач управления человеческими ресурсами и их обобщенная концептуальная модель

В работах [26, 31–33] авторами проведено исследование основных задач УЧР, выделены те из них, в решении которых требуются интеллектуальная поддержка, выявлены их специфические особенности. Так, в качестве примеров задач, решение которых сводится к принятию эффективных решений, показаны следующие задачи: управление процессами трудоустройства (отбор, оценка и прием на работу); аттестация (выявление соответствия персонала занимаемой должности); организация системы вознаграждения персонала; планирование карьеры работников (продвижение по службе); формирование резерва и т. п. Анализ перечисленных задач УЧР позволил выявить следующие характерные особенности последних:

- многокритериальность и разнородность данных, описывающих задачи УЧР;
- многоуровневая иерархическая структура критериев, выраженная в том, что каждый отдельный критерий верхнего уровня базируется на агрегировании частных критериев ближайшего нижнего уровня;
- количественный и качественный характеры критериев;
- невозможность однозначного определения критериев и изменчивость области их значений;
- различная степень влияния критериев и показателей на рассматриваемые варианты (объекты, альтернативы), а также необходимость учета различия их весов. Это определяет необходи-

мость привлечения к процессу принятия решений экспертов (носителей информации) и учета мнений последних;

- влияние уровня компетентности экспертов на качество принятого решения;
- наличие в реальных ситуациях большого числа разнородных частных критериев, затрудняющих формальное сравнение альтернатив.

Перечисленные особенности задач УЧР позволяют идентифицировать их как задачи многокритериального анализа и принятия решений в нечеткой среде. Обычно многокритериальный анализ востребован при решении таких категорий задач, как выбор, оценка, сравнение, отбор, ранжирование и классификация объектов (альтернатив) в нечеткой среде. Эти задачи относятся к категории наиболее распространенных в системах поддержки принятия решений и встречаются в различных сочетаниях [34, 35].

Анализ наработанных к настоящему времени подходов и методов к решению задач УЧР показал их большое многообразие и позволил выделить обусловленность такой ситуации следующими факторами: 1) постановка задачи принятия решений в УЧР; 2) уровень сложности поставленной задачи, т. е. степень учета в постановке задачи УЧФ их специфики (частично или полностью); 3) содержательное и количественное различие множеств характеризующих задачи УЧР критериев и частных критериев, влияющих на расчет интегрального показателя; 4) различие единиц измерения частных критериев и методов оценки их весов (субъективных, объективных); 5) способы агрегирования частных критериев; 6) использование разных методов свертки критериев; 7) необходимость участия экспертов в процессе принятия решений или наоборот.

Таким образом, при выборе из широкого спектра того или иного метода для решения вышеуказанных задач УЧР необходимо руководствоваться в максимальном учете специфическими особенностями последних, указанными выше. Наряду с этим выбираемый методологический подход должен обеспечить:

- 1) отсутствие ограничений на число альтернатив, критериев и частных критериев;
- 2) расчет уровня компетентности экспертов, участвующих в процедуре принятия решений;
- 3) учет иерархической структурированности критериев, описывающих альтернативы;
- 4) возможность распространения предлагаемого методологического подхода на все задачи УЧР, требующие интеллектуальной поддержки.

На основе комплексного подхода к учету специфики процессов УЧР обобщенная концептуальная модель принятия решений в задачах УЧР может быть представлена следующим набором информации:

$$M_{УЧР} = (X, K, Y, E, V, P, L, W),$$

где $X = \{x_1, x_2, \dots, x_n\} = \{x_i, i = \overline{1, n}\}$ — множество допустимых альтернатив; $K = \{K_1, K_2, \dots, K_m\} = \{K_j, j = \overline{1, m}\}$ — множество критериев выбора, характеризующих альтернативы; $K_j = \{k_{j1}, k_{j2}, \dots, k_{jt}\} = \{k_{jt}, t = \overline{1, T}\}$ — множество подкритериев, характеризующих каждый из критериев; Y — область определения значений каждого частного критерия; E — группа экспертов, участвующих в процедуре принятия решений; V — множество отношений между экспертами в соответствии с предпочтениями ЛПР; P — отношения между множествами X, K и E ; L — лингвистические выражения, отражающие степень удовлетворения альтернатив частным критериям (степень принадлежности); W — отношения между критериями и частными критериями.

Согласно концептуальной модели сущность принятия решений в задачах УЧР заключается в нахождении: 1) систематизированного списка альтернатив ($X \rightarrow X^*$), ранжированных от наилучшего (оптимального) к наихудшему (или наоборот); 2) наилучшего (оптимального) варианта альтернатив. Для этого необходимо сведение многокритериальной задачи оптимизации к однокритериальной, учитывающей:

- нечеткие отношения между множествами альтернатив, критериев оценивания альтернатив, группы экспертов, участвующих в оценке степени удовлетворения альтернатив критериям с учетом лингвистического характера последних;
- нечеткие отношения между критериями и частными критериями;
- нечеткие отношения компетентности экспертов в соответствии с предпочтениями ЛПР.

3. Общая постановка задачи

Принимая во внимание специфические особенности задач принятия решений в УЧР и предложенную концептуальную модель, дадим общую постановку задачи многокритериального ранжирования/выбора альтернатив. Под задачей многокритериальной оптимизации, как правило, подразумевают нахождение максимума или минимума векторного критерия на допустимом множестве альтернатив.

Пусть заданы следующие компоненты задач УЧР в организации:

1. $X = \{x_i, i = \overline{1, n}\}$ — множество альтернатив.
2. $K = \{K_j, j = \overline{1, m}\}$ — множество количественных и качественных критериев, по которым оцениваются альтернативы.
3. $K_j = \{k_{jt}, t = \overline{1, s_j}\}$ — множество частных критериев.
4. $E = \{e_l, l = \overline{1, g}\}$ — множество экспертов.

5. $w_j, j = \overline{1, m}$ — коэффициенты относительной важности критериев ($K = \{k_j, j = \overline{1, m}\}$).

6. $w_{jp}, t = \overline{1, s_j}, j = \overline{1, m}$ — коэффициенты относительной важности частных критериев ($k_j = \{k_{jp}, t = \overline{1, s_j}\}$).

7. $v_l, l = \overline{1, g}$ — коэффициенты компетентности экспертов.

Пусть $f(x)$ — целевая функция, которая обеспечивает выбор наилучших альтернатив:

1) $f(x) = \max(f(x_1), f(x_2), \dots, f(x_n))$ и $f(x) \rightarrow [0, 1]$, где $f(x_i)$ — результирующий вектор оценки альтернативы $x_i \in X$ в соответствии с интегральным критерием K , т. е. $f(x_i) \rightarrow K(x_i)$;

2) $K(x_i) = (p(x_i), w, v)$ — интегральная оценка альтернативы x_i , в соответствии с набором критериев оценки, весов частных критериев в интегральном критерии K и коэффициентом относительной важности компетентности экспертов, где $p(x_i)$ —

интегральная оценка альтернативы $x_i, i = \overline{1, n}$ в соответствии со значениями лингвистической переменной согласно предпочтению экспертов; $w = (w_1, \dots, w_z)$ — веса частных критериев в интегральном критерии $K, z = \overline{1, Z}, Z$ — общее число частных критериев; $v = (v_1, \dots, v_g)$ — коэффициент относительной важности компетентности экспертов в соответствии с предпочтениями ЛПР;

3) $f(x_i) > 0$, при условии $p(x_i) \geq 0$;

4) $g(K(x), w, v) \in G, x \in X$,

$$w_j > 0, j = \overline{1, m}, \sum_{j=1}^m w_j = 1,$$

$$w_{jt} > 0, t = \overline{1, s_j}, \sum_{t=1}^{s_j} w_{jt} = 1,$$

$$w_z > 0, z = \overline{1, Z},$$

$$v_l > 0, l = \overline{1, g}, \sum_{l=1}^g v_l = 1.$$

Требуется найти те альтернативы, которые в наибольшей степени соответствуют целевым функциям и ограничениям.

Согласно постановке задачи множество допустимых решений формируется посредством исключения из исходного множества альтернатив тех, которые не удовлетворяют поставленной цели и принятым ограничениям.

Как было показано выше, многокритериальные задачи УЧР относятся к классу слабоструктурированных проблем, содержащих большое число как качественных, так и количественных критериев оценки качества решений: При этом ЛПР (эксперт) исходит из своих субъективных предпочтений об эффективности возможных альтернатив и важности

различных критериев. Как известно, при построении модели предпочтений ЛПР вырабатывается большой объем информации. Это обуславливает сложность, а зачастую и невозможность расчета показателей эффективности и однозначного выбора наилучшего решения аналитическими методами. Поэтому основу существующих концепций оценки предпочтений составляет переход от аналитических методов к эвристическим и включение ЛПР (экспертов) в качестве одного из главных компонентов задачи принятия решений.

Как следует из постановки задачи многокритериальности принятия решений в УЧР, необходим эффективный инструментарий, позволяющий строить сложные процедуры принятия решений и оценивать широкий спектр альтернатив. В соответствии с постановкой задачи в работе предпочтение отдано одному из современных методов многокритериального выбора — технологии TOPSIS, которая модифицирована к условиям решаемой проблемы.

4. Метод TOPSIS

Основная идея метода TOPSIS заключается в том, что наиболее предпочтительная альтернатива должна иметь не только наибольшую близость к идеальному решению, но и быть дальше всех остальных альтернатив от неприемлемого решения [17]. Здесь наилучшее (оптимальное) решение представляет собой вектор, содержащий максимальные значения по каждому критерию для всех альтернатив, а неприемлемое (наихудшее) решение есть вектор, содержащий минимальные значения по каждому критерию. Как следует из сущности метода TOPSIS, с использованием последнего достаточно эффективно можно решить задачи нечеткой многокритериальной оптимизации, которые составляют математическую основу поддержки принятия решения в задачах УЧР. Под многокритериальной оптимизацией в теории принятия решений подразумевается выбор наилучшего решения среди возможных альтернатив [6, 7].

Метод TOPSIS является одним из эффективных инструментов содействия ЛПР и экспертам в формулировании их целей и субъективных предпочтений, структурировании множества критериев, оценки альтернатив в процессе принятия решений на языке нечеткой математики, лингвистических переменных, нечетких множеств и нечетких чисел.

Решение задачи оптимизации с использованием TOPSIS предполагает необходимость перевода значений качественных лингвистических переменных, выражающих степень удовлетворения той или иной альтернативы критериям, в нечеткие числа.

Нечеткое число представляет собой нечеткое подмножество универсального множества действительных чисел, имеющее нормальную и выпуклую функцию принадлежности, для которой существует такое значение носителя, где функция принадлеж-

ности равна единице, а при отступлении от своего максимума влево или вправо функция принадлежности убывает [36]. Согласно [37] нечеткие суждения экспертов, сформулированные в терминах естественного языка, могут быть описаны нечеткими треугольными и нечеткими трапециевидными (трапезоидными) числами. В настоящей работе, учитывая необходимость обеспечения устойчивости критериев к границам интервала достоверности, используется нечеткое трапециевидное число (рис. 1). Формирование суждений эксперта в виде нечеткого трапециевидного числа на практике реализуется следующим образом. Исследуемый объект по выбранному критерию оценивается экспертом четверкой чисел (n_1, n_2, n_3, n_4) , где n_1, n_2, n_3, n_4 — действительные числа. Суть этой процедуры заключается в том, что значение критерия находится в пределах от n_1 до n_4 , но вероятнее всего оно находится в пределах от n_4 до n_3 . Если в четверке $\bar{n} = (n_1, n_2, n_3, n_4)$ средние числа будут равны, т. е. $n_2 = n_3$, то нечеткое трапециевидное число \bar{n} преобразуется в нечеткое треугольное число.

Посредством операций над функциями принадлежности на основе сегментного принципа вводят операции над нечеткими числами [38].

При использовании технологии TOPSIS следует принять во внимание некоторые операции над нечеткими числами.

Пусть заданы два нечетких трапециевидных числа $\bar{n} = (n_1, n_2, n_3, n_4)$ и $\bar{m} = (m_1, m_2, m_3, m_4)$. Ниже приведены операции суммирования, разности и произведения этих чисел:

$$\begin{aligned} \bar{n} \oplus \bar{m} &= [n_1 + m_1, n_2 + m_2, n_3 + m_3, n_4 + m_4], \\ \bar{n} - \bar{m} &= [n_1 - m_4, n_2 - m_3, n_3 - m_2, n_4 - m_1], \\ \bar{n} \otimes \bar{m} &\cong [n_1 m_1, n_2 m_2, n_3 m_3, n_4 m_4], \\ \bar{n} \otimes r &= [n_1 r, n_2 r, n_3 r, n_4 r]. \end{aligned} \quad (1)$$

Расстояние между двумя нечеткими трапециевидными числами определяется из выражения [18, 39]

$$d_c(\bar{n}, \bar{m}) = \sqrt{\frac{1}{4}((n_1 - m_1)^2 + (n_2 - m_2)^2 + (n_3 - m_3)^2 + (n_4 - m_4)^2)}. \quad (2)$$

Если $\bar{n} = \bar{m}$, т. е. \bar{n} и \bar{m} равнозначны, то $d_c(n, m) = 0$.

Для реализации метода необходимо оперировать лингвистическими переменными и их значениями, выражающими вербальные шкалы оценки для измерения признаков. При этом уровни располагаются в порядке возрастания интенсивности проявления этих признаков. В данном случае число значений (градаций) лингвистических переменных равно семи. На рис. 2 приведено графическое изображение принципа преобразования лингвистических значений в числовые соответствия.

Табл. 1 демонстрирует семиуровневые значения лингвистической переменной и соответствующие им нечеткие трапециевидные числа.

В соответствии с табл. 1 для каждого значения лингвистической переменной можно найти числовое соответствие. Так, например, числовое соответствие лингвистического значения "не очень хорошо", представляющего собой одну из градаций измерения свойств, в 10-балльной системе оценки определяется как (5, 6, 7, 8).

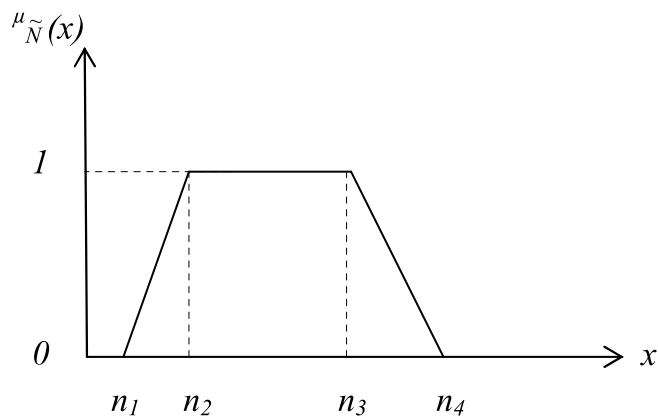


Рис. 1. Нечеткое трапециевидное число

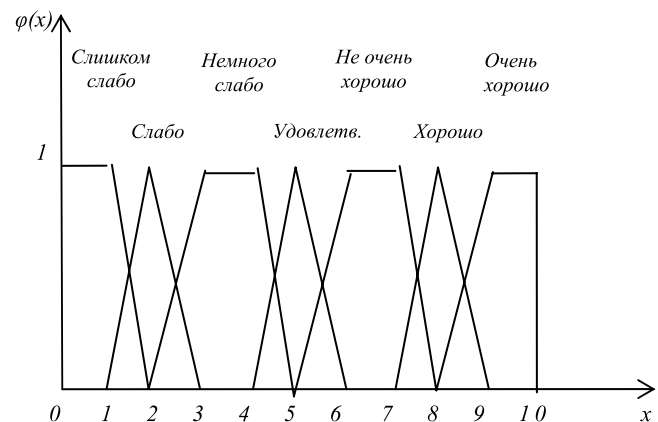


Рис. 2. Преобразование лингвистических значений в нечеткие трапециевидные числа

Таблица 1
Лингвистические значения и соответствующие им нечеткие трапециевидные числа

Лингвистические значения	Нечеткие трапециевидные числа
Слишком слабо	(0,0,1,2)
Слабо	(1,2,2,3)
Немного слабо	(2,3,4,5)
Удовлетворительно	(4,5,5,6)
Не очень хорошо	(5,6,7,8)
Хорошо	(7,8,8,9)
Очень хорошо	(8,9,10,10)

Алгоритм многокритериальной оптимизации задач УЧР на базе метода TOPSIS

Согласно постановке задачи, алгоритм многокритериальной оптимизации задач УЧР на базе метода TOPSIS предполагает выполнение следующей последовательности действий.

Шаг 1. Для проведения многокритериальной оптимизации задач УЧР на базе метода TOPSIS необходимо прежде всего избавиться от иерархической структурированности критериев (рис. 3). С этой целью на основе метода анализа иерархий (МАИ) Саати с помощью коэффициентов относительной важности критериев $\{K_j, j = \overline{1, m}\}$ и частных критериев $\{k_{jt}, t = \overline{1, s_j}\}$ определяются веса [40, 41], с которыми последние войдут в расчет интегрального критерия K . В формализованном виде произ-

ведением w_j , где $\sum_{j=1}^m w_j = 1$ и w_{jt} , где $\sum_{t=1}^{s_j} w_{jt} = 1$, определяется w_{jt}^K — вес частного критерия k_{jt} в расчете интегрального критерия $K = \{k_j, j = \overline{1, m}\}$, т. е. $w_{jt}^K = w_{jt} \cdot w_j$.

В результате двухступенчатая иерархическая структура критериев выбора $K = \{K_j, j = \overline{1, m}\}$, ха-

рактеризующих альтернативы, сводится к вычислению интегрального критерия, который учитывает веса частных критериев $\{k_{jt}, t = \overline{1, s_j}\}$, что позволяет, соответственно, избавиться от иерархической структурированности (рис. 4).

В последующих шагах для упрощения индексов все частные критерии объединяются в едином множестве G :

$$G = \{k_{jt}, j = \overline{1, m}, t = \overline{1, s_j}\} = \{k_z, z = \overline{1, Z}\},$$

$$z = s_{j-1} + t, j = \overline{1, m}, t = \overline{1, s_j}, s_0 = 0.$$

Здесь Z — общее число частных критериев, характеризующих альтернативы, т. е. $Z = \sum_{j=1}^m s_j$. В та-

ком случае $w_z = w_{jt}^K$.

Шаг 2. Степени принадлежности (соответствия) альтернатив частным критериям оцениваются лингвистическими значениями (см. табл. 1) и выражаются нечеткими трапециевидными числами $R^l = (r_{iz}^l) = (a_{iz}^l, b_{iz}^l, c_{iz}^l, d_{iz}^l)$. Так, например, если степень удовлетворения (принадлежности) альтернативы x_i частному критерию k_z экспертом l оценена значением "хорошо", то это выражается как $r_{iz}^l = (7, 8, 8, 9)$, а если экспертом дана оценка

"очень хорошо", то $r_{iz}^l = (8, 9, 10, 10)$ и т. д. В результате экспертной оценки степеней принадлежности альтернатив частным критериям получаем следующую матрицу:

$$R^l = [r_{iz}^l], l = \overline{1, g} \Leftrightarrow$$

$$\Leftrightarrow \{a_{iz}^l, b_{iz}^l, c_{iz}^l, d_{iz}^l\}, l = \overline{1, g}.$$

Шаг 3. Этот шаг предполагает предварительный расчет коэффициентов компетентности экспертов $v_l, l = \overline{1, g}$. С этой целью авторами применена модификация метода, заключающаяся в интегрировании в алгоритм предварительно рассчитанных коэффициентов компетентности экспертов, участвующих в процедуре оценки альтернатив.

С учетом коэффициентов компетентности экспертов $v_l, l = \overline{1, g}$, формируется матрица $R^{vl} = [r_{iz}^{vl}], l = \overline{1, g} \Leftrightarrow$
 $\Leftrightarrow \{a_{iz}^{v_l}, b_{iz}^{v_l}, c_{iz}^{v_l}, d_{iz}^{v_l}\}, l = \overline{1, g}$. Элементами этой матрицы являются трапециевидные числа, которые выражают

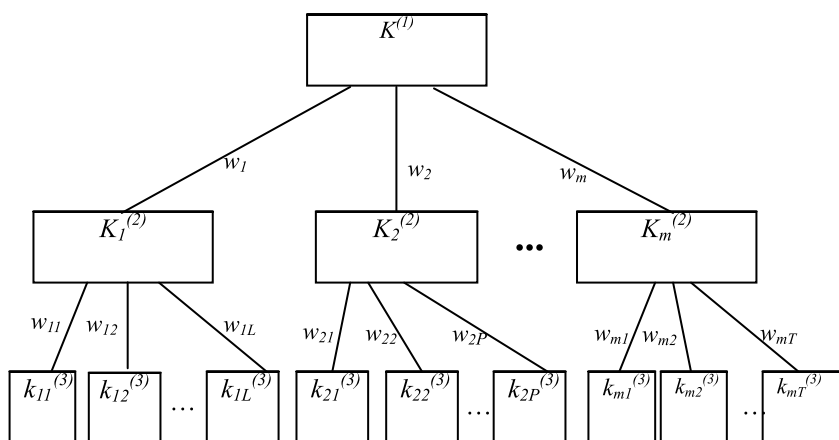


Рис. 3. Иерархическая структурированность критериев выбора, характеризующих альтернативы

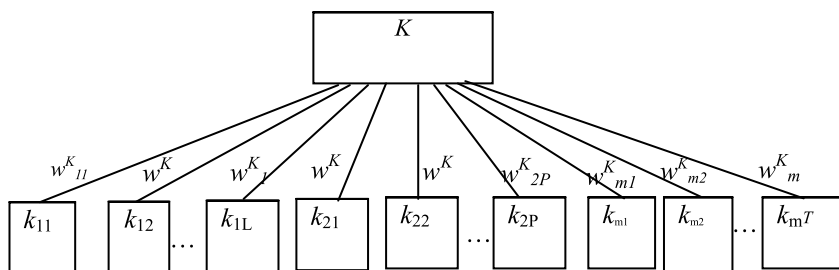


Рис. 4. Сведение иерархической структуры критерия K к интегральному вектору частных критериев

степень удовлетворения альтернативы x_i частным критериям k_z с учетом компетентности экспертов и рассчитываются следующим образом:

$$\begin{aligned} a_{iz}^{v_l} &= a_{iz}^l \cdot v_l; \\ b_{iz}^{v_l} &= b_{iz}^l \cdot v_l; \\ c_{iz}^{v_l} &= c_{iz}^l \cdot v_l; \\ d_{iz}^{v_l} &= d_{iz}^l \cdot v_l. \end{aligned} \quad (3)$$

Шаг 4. Определяется единая — агрегированная матрица

$$\begin{aligned} R^{v_l} &= [r_{iz}^{v_l}], l = \overline{1, g} \Leftrightarrow \{a_{iz}^{v_l}, b_{iz}^{v_l}, c_{iz}^{v_l}, d_{iz}^{v_l}\}, \\ l = \overline{1, g} \Rightarrow R_{iz} &= [r_{iz}] \Leftrightarrow \{a_{iz}, b_{iz}, c_{iz}, d_{iz}\}. \end{aligned}$$

Элементы этой матрицы вычисляются следующим образом:

$$\begin{aligned} a_{iz} &= \{\min a_{iz}^{v_l}, l = \overline{1, g}\}; \\ b_{iz} &= \frac{1}{g} \sum_{l=1}^g b_{iz}^{v_l}; \\ c_{iz} &= \frac{1}{g} \sum_{l=1}^g c_{iz}^{v_l}; \\ d_{iz} &= \{\max d_{iz}^{v_l}, l = \overline{1, g}\}. \end{aligned} \quad (4)$$

Шаг 5. Элементы матрицы $R_{iz} = [r_{iz}] \Leftrightarrow \{a_{iz}, b_{iz}, c_{iz}, d_{iz}\}$ умножаются на вес частных критериев. В результате этой операции строится взвешенная нечеткая матрица $R_{iz}^w = [r_{iz}^w] \Leftrightarrow \{a_{iz}^w, b_{iz}^w, c_{iz}^w, d_{iz}^w\}$. Здесь

$$\begin{aligned} a_{iz}^w &= a_{iz} \cdot w_z; \\ b_{iz}^w &= b_{iz} \cdot w_z; \\ c_{iz}^w &= c_{iz} \cdot w_z; \\ d_{iz}^w &= d_{iz} \cdot w_z. \end{aligned} \quad (5)$$

Шаг 6. Полученная матрица нормализуется. Для этого применяется метод Hsu и Sehn [42], на основе которого определяются $d_z^+ = \max d_{iz}^w, i = \overline{1, n}$. Далее на основе выражения

$$\begin{aligned} R_{iz}^N &= [r_{iz}^N] \Leftrightarrow \\ \Leftrightarrow \{a_{iz}^N, b_{iz}^N, c_{iz}^N, d_{iz}^N\} &\Leftrightarrow \left\{ \frac{a_{iz}^w}{d_z^+}, \frac{b_{iz}^w}{d_z^+}, \frac{c_{iz}^w}{d_z^+}, \frac{d_{iz}^w}{d_z^+} \right\} \end{aligned} \quad (6)$$

определяются элементы нормализованной матрицы принятия решений.

Шаг 7. На основе взвешенных значений определяется идеальное позитивное (наилучшее) решение (ИПР) X^* . С этой целью для каждого $k_z, z = \overline{1, Z}$ отбираются

$$d_z^* = \{\max d_{iz}^N, i = \overline{1, n}\}, \quad (7)$$

и формируется следующая матрица:

$$\begin{aligned} X^* &= [d_z^*] = \\ &= [(d_1^*, d_1^*, d_1^*, d_1^*), \dots, (d_Z^*, d_Z^*, d_Z^*, d_Z^*)]. \end{aligned} \quad (8)$$

Согласно выражению (6) $d_z^* = 1$ для $\forall z$, т. е. все элементы матрицы X^* равны единице.

Шаг 8. Вычисляется идеальное негативное (наихудшее) решение (ИНР) X^- . С этой целью для каждого $k_z, z = \overline{1, Z}$, отбираются

$$a_z^- = \{\min a_{iz}^N, i = \overline{1, n}\} \quad (9)$$

и формируется следующая матрица:

$$\begin{aligned} X^- &= [a_z^-] = \\ &= [(a_1^-, a_1^-, a_1^-, a_1^-), \dots, (a_Z^-, a_Z^-, a_Z^-, a_Z^-)]. \end{aligned} \quad (10)$$

Шаг 9. С использованием формулы (2) по индивидуальным значениям каждого частного критерия рассчитывается расстояние альтернатив до ИПР:

$$\begin{aligned} D_z^*(x_i, X^*) &= \\ &= \sqrt{\frac{1}{4}((a_{iz}^N - d_z^*)^2 + (b_{iz}^N - d_z^*)^2 + (c_{iz}^N - d_z^*)^2 + (d_{iz}^N - d_z^*)^2)}. \end{aligned} \quad (11)$$

На основе полученных результатов формируется вектор $[D^*] = [D_1^*, \dots, D_Z^*]$.

Шаг 10. По индивидуальным значениям каждого частного критерия рассчитывается расстояние альтернатив до ИНР:

$$\begin{aligned} D_z^-(x_i, X^-) &= \\ &= \sqrt{\frac{1}{4}((a_{iz}^N - a_z^-)^2 + (b_{iz}^N - a_z^-)^2 + (c_{iz}^N - a_z^-)^2 + (d_{iz}^N - a_z^-)^2)}. \end{aligned} \quad (12)$$

На основе полученных результатов формируется вектор $[D^-] = [D_1^-, \dots, D_Z^-]$.

Шаг 11. Определяется расстояние каждой из альтернатив до ИПР:

$$D^*(x_i) = \sqrt{\sum_{z=1}^Z (D_z^*(x_i, X^*))^2}. \quad (13)$$

Шаг 12. Определяется расстояние каждой из альтернатив до ИНР:

$$D^-(x_i) = \sqrt{\sum_{z=1}^Z (D_z^-(x_i, X^-))^2}. \quad (14)$$

Шаг 13. Рассчитывается интегральный показатель (коэффициент близости) для каждой сравниваемой альтернативы как отношение вычисленного для нее расстояния от идеально наилучшего ре-

шения к сумме расстояний до наилучшего и наихудшего решений:

$$D(x_i) = D^*(x_i) + D^-(x_i),$$

$$\varphi(x_i) = \frac{D^-(x_i)}{D(x_i)}. \quad (15)$$

В соответствии со значением коэффициента близости $\varphi(x_i)$ появляется возможность ранжирования альтернатив. Так, чем ближе к единице значение коэффициента близости $\varphi(x_i)$, тем предпочтительнее сравниваемая альтернатива. Наибольшее значение интегрального показателя $\varphi(x_i)$ определяет наилучшую альтернативу, т. е. оптимальное решение. Наименьшее значение $\varphi(x_i)$ соответствует наихудшей альтернативе.

6. Применение предложенного метода для решения задач отбора и приема на работу

В УчР задача отбора персонала имеет большое значение, так как только при наличии квалифицированного персонала организация может успешно функционировать. Поэтому вопросы принятия кадровых решений, свободных от субъективизма, являются достаточно актуальными. Трудности, с которыми сталкиваются организации в процессе определения кандидата, наиболее приемлемого для конкретной должности как в плане профессиональной пригодности, так и с позиций соответствия корпоративному стилю, обуславливают необходимость разработки и совершенствования научных подходов к управлению кадровой политикой.

Предложенный инструментальный подход был апробирован в процессе решения задач отбора и приема на работу для оценки кандидатов. Эксперименты проводили для оценки кандидатов, претендующих на должность в Отдел УчР Государственной Нефтяной Компании Азербайджанской Республики. С этой целью были реализованы следующие действия.

1. С участием четырех экспертов для приема на работу в Отдел УчР была сформирована система критериев (табл. 2). На основе метода попарного сравнения [40] определены коэффициенты относительной важности критериев и частных критериев. При этом рассмотрена также задача выявления противоречий в экспертных оценках [40, 43]. На основе полученных результатов выявлены коэффициенты относительной важности критериев и частных критериев, а также веса частных критериев, с которыми они участвуют при расчете интегральной оценки критерия K .

Согласно шагу 1 алгоритма поставленной задачи, для упрощения индексов все частные критерии объединяются в едином множестве G . Тогда $G = \{k_z, z = \overline{1, 12}\}$.

2. Полученный на основе оценки указанных частных критериев интегральный показатель (коэффициент близости сравниваемых альтернатив) $\varphi(x_i)$ выражает определенное на интервале $[0, 1]$ значение степени шанса приема на работу каждого кандидата x_i . Значения этой величины позволяют принять заключительное решение относительно каждого альтернативного претендента. В процессе проведения экспериментов по приему на работу

Таблица 2

Коэффициенты относительной важности критериев и частных критериев, веса частных критериев в K

Критерии	Коэффициенты относительной важности критериев	Частные критерии	Единое множество частных критерий G	Коэффициенты относительной важности частных критериев	Весовые коэффициенты частных критериев
K_1 — профессиональные (образование, знания, профессиональные навыки, умения и т.п.)	0,11	k_{11} — соответствие полученного образования требованиям рабочего места	k_1	0,54	0,06
		k_{12} — научно-исследовательские способности	k_2	0,46	0,05
K_2 — мотивационные	0,08	k_{21} целеустремленность	k_3	0,47	0,04
		k_{22} — ориентация на результат	k_4	0,53	0,04
K_3 — деловые	0,4	k_{31} — трудолюбие	k_5	0,2	0,08
		k_{32} — креативность	k_6	0,22	0,13
		k_{33} — инициативность	k_7	0,26	0,10
		k_{34} — самостоятельность	k_8	0,32	0,09
K_4 — личностные	0,1	k_{41} — обучаемость	k_9	0,63	0,06
		k_{42} — исполнительность	k_{10}	0,37	0,04
K_5 — индивидуально-психологические и здоровье	0,31	k_{51} — физическое здоровье	k_{11}	0,35	0,11
		k_{52} — нервно-психологическая устойчивость	k_{12}	0,65	0,20

экспертами были сформулированы следующие варианты возможных заключительных решений:

- Если $\varphi(x_j) \in [0, 0.25)$, то претендент явно не соответствует требованиям должности, т. е. кандидатура однозначно отклоняется.
- Если $\varphi(x_j) \in [0.25, 0.50)$, то претендент слабо соответствует требованиям должности, поэтому его наем представляет большой риск.
- Если $\varphi(x_j) \in [0.50, 0.65)$, то претендент частично (в некоторой степени) соответствует требованиям рабочего места. Прием кандидата на работу связан с небольшим риском, который может компенсироваться в процессе работы за счет высоких показателей по другим компетенциям.
- Если $\varphi(x_j) \in [0.65, 0.8)$, то претендент соответствует требованиям рабочего места, при том что некоторые показатели могут быть легко восполнены в процессе адаптации.
- Если $\varphi(x_j) \in [0.8, 1]$, то претендент полностью соответствует всем требованиям рабочего места.

3. Определены коэффициенты компетентности экспертов, участвующих в оценке претендентов на трудоустройство. Эти коэффициенты определены с использованием метода попарного сравнения [40, 43]. В табл. 3 представлена матрица попарного сравнения компетентности экспертов в соответствующей сфере, построенная на основе лингвистического выражения "небольшое превосходство эксперта 1 и эксперта 4 над экспертом 2 и экспертом 3".

В результате попарного сравнения получены следующие значения коэффициентов компетентности экспертов: $v_1 = 0,375$, $v_2 = 0,125$, $v_3 = 0,125$, $v_4 = 0,375$.

4. С участием четырех экспертов на основе семиуровневых лингвистических переменных проведена оценка на степень удовлетворения (принадлежность) 12 частным критериям трех кандидатов на вакантное место, прошедших требуемые этапы отбора.

5. На основе оценок четырех экспертов построена $(3 \times 12 \times 4)$ -мерная обобщенная матрица трапециевидных нечетких чисел (табл. 4).

6. С учетом компетентности экспертов на основе формулы (3) строится матрица трапециевидных нечетких чисел и определяются агрегированные в соответствии с формулой (4) трапециевидные нечеткие числа (табл. 5).

7. Элементы матрицы агрегированных трапециевидных нечетких чисел с использованием фор-

мулы (5) умножаются на веса частных критериев и результаты нормализуются (табл. 5).

8. В соответствии с выражениями (7)–(10) формируется интегральная матрица нечетких идеальных позитивных (наилучших) и нечетких идеальных негативных (наихудших) решений (табл. 6).

9. Результаты рассчитанных на основе формулы (11) по значению каждого частного критерия расстояний альтернатив до ИПР представлены в табл. 6.

Таблица 4

Матрица трапециевидных нечетких чисел, отражающих принадлежность альтернатив частным критериям

Частные критерии	Альтернативы	Эксперт 1	Эксперт 2	Эксперт 3	Эксперт 4
k_1	x_1	(7,8,8,9)	(7,8,8,9)	(7,8,8,9)	(7,8,8,9)
	x_2	(7,8,8,9)	(8,9,10,10)	(4,5,5,6)	(5,6,7,8)
	x_3	(7,8,8,9)	(7,8,8,9)	(8,9,10,10)	(7,8,8,9)
k_2	x_1	(8,9,10,10)	(8,9,10,10)	(7,8,8,9)	(8,9,10,10)
	x_2	(5,6,7,8)	(7,8,8,9)	(8,9,10,10)	(5,6,7,8)
	x_3	(7,8,8,9)	(8,9,10,10)	(8,9,10,10)	(4,5,5,6)
k_3	x_1	(7,8,8,9)	(7,8,8,9)	(4,5,5,6)	(5,6,7,8)
	x_2	(7,8,8,9)	(8,9,10,10)	(8,9,10,10)	(7,8,8,9)
	x_3	(8,9,10,10)	(7,8,8,9)	(7,8,8,9)	(8,9,10,10)
k_4	x_1	(5,6,7,8)	(7,8,8,9)	(8,9,10,10)	(7,8,8,9)
	x_2	(8,9,10,10)	(8,9,10,10)	(8,9,10,10)	(8,9,10,10)
	x_3	(7,8,8,9)	(8,9,10,10)	(7,8,8,9)	(7,8,8,9)
k_5	x_1	(7,8,8,9)	(5,6,7,8)	(8,9,10,10)	(7,8,8,9)
	x_2	(8,9,10,10)	(7,8,8,9)	(7,8,8,9)	(8,9,10,10)
	x_3	(7,8,8,9)	(8,9,10,10)	(8,9,10,10)	(7,8,8,9)
k_6	x_1	(7,8,8,9)	(7,8,8,9)	(8,9,10,10)	(7,8,8,9)
	x_2	(8,9,10,10)	(5,6,7,8)	(7,8,8,9)	(8,9,10,10)
	x_3	(5,6,7,8)	(7,8,8,9)	(8,9,10,10)	(5,6,7,8)
k_7	x_1	(8,9,10,10)	(8,9,10,10)	(8,9,10,10)	(8,9,10,10)
	x_2	(7,8,8,9)	(8,9,10,10)	(7,8,8,9)	(7,8,8,9)
	x_3	(7,8,8,9)	(7,8,8,9)	(4,5,5,6)	(5,6,7,8)
k_8	x_1	(7,8,8,9)	(8,9,10,10)	(8,9,10,10)	(4,5,5,6)
	x_2	(7,8,8,9)	(8,9,10,10)	(7,8,8,9)	(7,8,8,9)
	x_3	(5,6,7,8)	(7,8,8,9)	(8,9,10,10)	(7,8,8,9)
k_9	x_1	(8,9,10,10)	(7,8,8,9)	(7,8,8,9)	(8,9,10,10)
	x_2	(7,8,8,9)	(7,8,8,9)	(8,9,10,10)	(7,8,8,9)
	x_3	(8,9,10,10)	(5,6,7,8)	(7,8,8,9)	(8,9,10,10)
k_{10}	x_1	(5,6,7,8)	(7,8,8,9)	(8,9,10,10)	(5,6,7,8)
	x_2	(7,8,8,9)	(8,9,10,10)	(8,9,10,10)	(4,5,5,6)
	x_3	(7,8,8,9)	(8,9,10,10)	(7,8,8,9)	(7,8,8,9)
k_{11}	x_1	(8,9,10,10)	(7,8,8,9)	(7,8,8,9)	(7,8,8,9)
	x_2	(5,6,7,8)	(7,8,8,9)	(8,9,10,10)	(7,8,8,9)
	x_3	(8,9,10,10)	(8,9,10,10)	(8,9,10,10)	(8,9,10,10)
k_{12}	x_1	(7,8,8,9)	(8,9,10,10)	(7,8,8,9)	(7,8,8,9)
	x_2	(7,8,8,9)	(7,8,8,9)	(4,5,5,6)	(5,6,7,8)
	x_3	(7,8,8,9)	(8,9,10,10)	(8,9,10,10)	(7,8,8,9)

Таблица 3

Матрица попарного сравнения и коэффициенты компетентности экспертов

Попарное сравнение экспертов	E_1	E_2	E_3	E_4	Коэффициенты компетентности экспертов (v_j)
E_1	1	3	3	1	0,375
E_2	0,33	1	1	0,33	0,125
E_3	0,33	1	1	0,33	0,125
E_4	1	3	3	1	0,375

Элементы нормализованной матрицы принятия решений

Частные критерии	Альтернативы	Весовые коэффициенты частных критериев	Агрегированные трапециевидные нечеткие числа	Элементы взвешенной нечеткой матрицы (x100)	Элементы нормализованной матрицы принятия решений
k_1	x_1 x_2 x_3	0.06	(0.875,2,2,3.375) (0.5,1.715, 1.88, 3.375) (0.875,2.031,2.06, 3.375)	(5.25, 12, 12, 20.25) (3, 10.29, 11.28, 20.25) (5.25, 12.186, 12.36, 20.25)	(0.259, 0.592, 0.592, 1) (0.148, 0.508, 0.557, 1) (0.259, 0.601,0.61, 1)
k_2	x_1 x_2 x_3	0.05	(0.875,2.219,2.44,3.75) (0.875,1.586,1.926,3.375) (1.1,1.781,1.846,3.375)	(4.375, 11.08, 12.2, 18.75) (4.375, 7.93, 9.63, 16.875) (5.5, 8.905, 9.23, 16.875)	(0.233, 0.59, 0.65, 1) (0.233, 0.422, 0.513, 0.9) (0.293, 0.474, 0.492, 0.9)
k_3	x_1 x_2 x_3	0.04	(0.5, 1.684, 1.862, 3.375) (1, 2.062, 2.125, 3.375) 0.875, 2.188, 2.385, 3.75	2, 6.736, 7.448, 13.5) (4, 8.248, 8.5, 13.5) (3.5, 8.752, 9.54, 15)	(0.133, 0.449, 0.496, 0.9) (0.266, 0.549, 0.566, 0.9) (0.233, 0.583, 0.636, 1)
k_4	x_1 x_2 x_3	0.04	(0.875, 1.809, 1.969, 3.375) (1, 2.25, 2.5, 3.75) (0.875, 2.031, 2.062, 3.375)	(3.5, 7.236, 7.876, 13.5) (4, 9, 10, 15) (3.5, 8.124, 8.248, 13.51)	(0.233, 0.482, 0.525, 0.9) (0.266, 0.6, 0.667, 1) (0.233, 0.541, 0.549, 0.9)
k_5	x_1 x_2 x_3	0.08	(0.5, 1.68, 1.835, 3.375) (0.875, 2.19, 2.38,3.75) (1, 2.03, 2.08, 3.375)	(4, 13.44, 14.68, 27) (7, 17.52, 19.04, 30) (8, 16.24, 16.64, 27)	(0.133, 0.448, 0.489, 0.9) (0.233, 0.584, 0.634, 1) (0.266, 0.541, 0.554, 0.9)
k_6	x_1 x_2 x_3	0.03	(0.875, 2.03, 2.06, 3.375) (0.625, 2.219, 2.325, 3.75) (0.875, 1.565, 1.9, 3.375)	(2.625, 6.09, 6.18, 10.125) (1.825, 6.657, 6.975, 11.25) (2.625, 4.695, 5.7, 10.125)	(0.233, 0.541, 0.549, 0.9) (0.162, 0.591, 0.62, 1) (0.233, 0.417, 0.506, 0.9)
k_7	x_1 x_2 x_3	0.10	(1, 2.25, 2.5, 3.75) (0.875, 2.031, 2.06, 3.375) (0.5, 1.684, 1.86, 3.375)	(10, 22.5, 25, 37.5) (8.75, 20.31, 20.6, 33.75) (5, 16.84, 18.6, 33.75)	(0.266, 0.6, 0.667, 1) (0.233, 0.541, 0.549, 0.9) (0.133, 0.449, 0.496, 0.9)
k_8	x_1 x_2 x_3	0.09	(1, 1.781, 1.9, 3.375) (0.875, 2.031, 2.062, 3.375) (0.875, 1.836, 1.969, 3.375)	(9, 16.029, 17.1, 30.375) (7.875, 18.279, 18.54, 30.375) (7.875, 16.524, 17.721, 30.375)	(0.296, 0.527, 0.562, 1) (0.259, 0.601, 0.61, 1) (0.259, 0.544, 0.583, 1)
k_9	x_1 x_2 x_3	0.06	(0.875, 2.188, 2.385, 3.75) (0.875, 2.031, 2.062, 3.375) (0.625, 2.125, 2.343, 3.75)	(5.25, 13.128, 14.31, 22.51) (5.25, 12.186, 12.372, 20.25) (3.75, 12.75, 14.058, 22.51)	(0.233, 0.583, 0.635, 1) (0.233, 0.541, 0.549, 0.9) (0.166, 0.566, 0.624, 1)
k_{10}	x_1 x_2 x_3	0.04	(0.875, 1.586, 1.875, 3.375) (1, 1.781, 1.843, 3.375) (0.875, 2.031, 2.062, 3.375)	(3.5, 6.34, 7.5, 13.51) (4, 7.124, 7.372, 13.51) (3.5, 8.124, 8.248, 13.51)	(0.259, 0.469, 0.555, 1) (0.296, 0.527, 0.545, 1) (0.259, 0.601, 0.61, 1)
k_{11}	x_1 x_2 x_3	0.11	(0.875, 2.093, 2.187, 3.75) (0.875,1.808, 1.968, 3.375) (1, 2.25, 2.5, 3.75)	(9.625, 23.023, 24.057, 41.25) (9.625, 19.888, 21.648, 37.125) (11, 24.75, 27.5, 41.25)	(0.233, 0.558, 0.583, 1) (0.233, 0.482, 0.524, 0.9) (0.267, 0.6, 0.667, 1)
k_{10}	x_1 x_2 x_3	0.20	(0.875, 2.031, 2.063, 3.375) (0.5, 1.683, 1.812, 3.375) (1, 2.062, 2.125, 3.375)	(17.5, 40.62, 41.26, 67.5) (10, 33.66, 36.24, 67.5) (20, 41.24, 42.5, 67.5)	(0.259, 0.601, 0.611, 1) (0.148, 0.498, 0.536, 1) (0.296, 0.61, 0.629, 1)

Таблица 6

Интегральная матрица ИПР и ИНР, расстояний альтернатив до ИПР и ИНР по значению каждого частного критерия

Частные критерии	X^*	X^-	$D(x_1 X^*)$	$D(x_2 X^*)$	$D(x_3 X^*)$	$D(x_1 X^-)$	$D(x_2 X^-)$	$D(x_3 X^-)$
k_1	(1,1,1,1)	(0.148, 0.148,0.148, 0.148)	0,46	0,539	0,463	0,531	0,636	0,537
k_2	(1,1,1,1)	(0.233, 0.233, 0.233, 0.233)	0,467	0,54	0,504	0,417	0,372	0,481
k_3	(1,1,1,1)	(0.133, 0.133, 0.133, 0.133)	0,573	0,483	0,472	0,452	0,491	0,551
k_4	(1,1,1,1)	(0.233, 0.233, 0.233, 0.233)	0,51	0,449	0,501	0,383	0,476	0,398
k_5	(1,1,1,1)	(0.133, 0.133, 0.133, 0.133)	0,575	0,472	0,488	0,45	0,55	0,486
k_6	(1,1,1,1)	(0.162, 0.162, 0.162, 0.162)	0,501	0,502	0,543	0,458	0,523	0,427
k_7	(1,1,1,1)	(0.133, 0.133, 0.133, 0.133)	0,449	0,501	0,573	0,563	0,483	0,452
k_8	(1,1,1,1)	(0.259, 0.259, 0.259, 0.259)	0,476	0,463	0,481	0,421	0,443	0,428
k_9	(1,1,1,1)	(0.166, 0.166, 0.166, 0.166)	0,472	0,501	0,505	0,522	0,455	0,515
k_{10}	(1,1,1,1)	(0.259, 0.259, 0.259, 0.259)	0,506	0,48	0,463	0,412	0,419	0,443
k_{11}	(1,1,1,1)	(0.233, 0.233, 0.233, 0.233)	0,491	0,522	0,448	0,457	0,384	0,477
k_{12}	(1,1,1,1)	(0.148, 0.148, 0.148, 0.148)	0,462	0,545	0,442	0,537	0,499	0,545

Расстояние сравниваемых альтернатив до ИПР и ИНР, коэффициент их близости к идеальному решению и соответствующие ранги

Альтернативы	X^*	X^-	$X^* + X^-$	$\varphi_K(x_i)$	Ранги	Решение
x_1	1,7208	1,6287	3,3495	0,486	3	Прием кандидата на работу связан с большим риском
x_2	1,7344	1,6719	3,4063	0,491	2	Прием кандидата на работу связан с большим риском
x_3	1,6619	1,6651	3,3270	0,501	1	Прием кандидата на работу связан с небольшим риском, который может компенсироваться за счет высоких показателей по другим компетенциям

10. Результаты рассчитанных на основе формулы (12) по значению каждого частного критерия расстояний альтернатив до ИНР приведены в табл. 6.

11. В соответствии с формулами (13), (14) определены расстояния каждой из альтернатив до ИПР и ИНР, соответственно. Далее с использованием формулы (15) рассчитаны значения интегрального показателя, выражающего степень близости каждой сравниваемой альтернативы к идеальному решению, и по результатам определены ранги каждой альтернативы (табл. 7).

В соответствии с полученными результатами наилучшим (оптимальным) вариантом решения в данном случае является альтернатива x_3 .

Заключение

Предложенный методологический подход к решению задач УЧР с использованием многокритериальной оптимизации на базе метода TOPSIS позволяет повысить адекватность принятых решений за счет приоритизации по степени близости к идеальному решению, обеспечивает объективность и транспарентность принимаемых управленческих решений и предоставляет возможности для расширения сфер приложения методов многокритериальной оптимизации.

Преимущества предложенного подхода к многокритериальной оптимизации для поддержки принятия решений в УЧР сводятся к следующему:

- математическая обоснованность и относительная простота расчетов интегральных показателей, позволяющих осуществить ранжирование альтернативных решений, проводить дальнейший анализ и выбор окончательного варианта решения;
- отсутствие ограничений на число альтернатив и критериев, характеризующих объект исследования;
- учет в алгоритме принятия решений уровня компетентности экспертов, участвующих в процедуре принятия решений;
- учет иерархической структурированности критериев, описывающих альтернативы;
- описание значений нечетких лингвистических переменных в виде нечетких трапециевидных чисел, обеспечивающих устойчивость критериев к границам интервала достоверности;
- возможность распространения предлагаемого методологического подхода на все задачи УЧР, требующие интеллектуальной поддержки процессов принятия решений.

В работе приведена пошаговая демонстрация возможностей предложенного метода в рамках процесса многокритериального анализа и выбора решений на примере задачи отбора и приема на работу, представляющая методическую ценность.

Использование описанного методологического подхода в качестве математической основы компьютерной системы поддержки принятия решений в задачах УЧР может стать действенным инструментом подготовки и принятия эффективных решений в УЧР.

Список литературы

1. Cole G. A. Personnel and Human Resource Management. Wadsworth: Thomson Learning, 5 edition. 2002, 448 p.
2. Spencer L. M., Spencer S. M. Competence at work models for superior performance. Wiley India Pvt. Limited, 2008. 384 p.
3. Armstrong M. Armstrong's handbook of strategic human resource management. 10th ed. London: Kogan Page, 2006. 216 p.
4. Макарова И. К. Управление человеческими ресурсами. Пять уроков эффективного HR-менеджмента. М.: Дело, 2007. 232 с. URL: <http://www.alleng.ru/d/manag/man094.htm>
5. Иванцевич Дж. М., Лобанов А. А. Человеческие ресурсы управления. М.: Аспект Пресс. 2004. 245 с.
6. Ларичев О. И. Теория и методы принятия решений: учебник. Изд. второе, перераб. и доп. М.: Логос, 2002. 392 с.
7. Микони С. В. Многокритериальный выбор на конечном множестве альтернатив. СПб.: Лань, 2009. 272 с.
8. Орловский С. А. Проблемы принятия решений при нечеткой исходной информации. М.: Наука, 1981. 212 с.
9. Мамедова М. Г. Принятие решений на основе баз знаний с нечеткой реляционной структурой. Баку: Элм, 1997. 296 с.
10. Кофман А. Введение в теорию нечетких множеств. М.: Радио и связь, 1982. 432 с.
11. Zadeh L. A. Fuzzy Sets // Information and control. 1965. Vol. 8, № 3. P. 338—351.
12. Трахтенгерц Э. А. Возможности и реализация компьютерных систем поддержки принятия решений // Известия РАН. Теория и системы управления. 2001. № 3. С. 86—103.
13. Chien C. F., Chen L. F. Data mining to improve personnel selection and enhance human capital: A case study in high-technology industry // Expert Systems with Applications. 2008. Vol. 34, No. 2. P. 280—290.
14. Gungor Z., Serhadhoglu G., Kesen S. E. A fuzzy AHP approach to personnel selection problem // Applied Soft Computing. 2009. No. 9. P. 641—649.
15. Chen P. C. A Fuzzy Multiple criteria decision making model in employee recruitment // IJCSNS International Journal of Computer Science and Network Security. 2009. Vol. 9. No. 7. P. 113—117.
16. Mohsen Varmazyar, Behrouz Nouri. A fuzzy AHP approach for employee recruitment // Decision Science Letters. 2014. No 3. P. 27—36.
17. Hwang C. L., Yoon K. Multiple attributes decision making methods and applications. Heidelberg. Berlin: Springer. 1981. 191 p.
18. Chen C. T. Extensions of the TOPSIS for group decision-making under fuzzy environment // Fuzzy Sets and Systems. 2000. Vol. 114. No. 1. P. 1—9.
19. Dursun M., Karsak E. A fuzzy MCDM approach for personnel selection // Expert Systems with Applications. 2010. Vol. 37, No. 6. P. 4324—4330.
20. Saghafian S., Hejazi S. R. Multi-criteria group decision making using a modified fuzzy TOPSIS procedure // Computational In-

telligence for Modelling, Control and Automation, 2005 and International Conference on Intelligent Agents, Web Technologies and Internet Commerce-International Conference. Vienna. 2005. 28–30 Nov. Vol. 2. P. 215–221.

21. **Kelemenis A., Askounis D.** A new TOPSIS-based multi-criteria approach to personnel selections // Expert Systems with Applications. 2010. Vol. 37, No. 6. P. 4999–5008.

22. **Kelemenis A., Ergazakis K., Askounis D.** Support managers' selection using an extension of fuzzy TOPSIS // Expert Systems with Applications. 2011. Vol. 38, No. 3. P. 2774–2782.

23. **Polychroniou P. V., Giannikos I.** A fuzzy multicriteria decision-making methodology for selection of human resources in a Greek private bank // Career Development International. 2009. Vol. 14, No. 4. P. 372–387.

24. **Nobari S.** Design of fuzzy decision support system in employee recruitment // Journal of Basic and Applied Scientific Research. 2011. Vol. 1, No. 11. P. 1891–1903.

25. **Akhlagh E.** A rough-set based approach to design an expert system for personnel selection // World Academy of Science, Engineering and Technology. 2011. Vol. 54. P. 202–205. URL: <http://waset.org/Publications/a-rough-set-based-approach-to-design-an-expert-system-for-personnel-selection/14092>

26. **Мамедова М. Г., Джабраилова З. Г.** Нечеткий логический подход к задаче оценки кадрового потенциала // Менеджмент в России и за рубежом. 2004. № 5. С. 111–117. URL: <http://mevri.ru/annotations/2004/5/>

27. **Mammadova M. G., Jabrayilova Z. Q.** Methods of Family Income estimation in the targeting social Assistance System // Appl. Comput. Math. 2007. Vol. 6, No 1. P. 80–87. URL: <http://acmij.az/view.php?lang=az&menu=journal&id=240>

28. **Mammadova M. H., Jabrayilova Z. Q., Mammadzada F. R.** Fuzzy Decision-Making Support Methods for the Selection of IT-Professionals // International Journal of Engineering and Innovative Technology (IJEIT), 2014. Vol. 3, No. 7. P. 169–175. URL: http://ijeit.com/Vol%203/Issue%207/IJEIT1412201401_31.pdf

29. **Neumann J. V., Morgenstern O.** Theory of games and economic behavior. One of Princeton University presses. Notable Centenary Titles. 2007. 776 p.

30. **Afshari A.R., Nikolić M., Čočkalović D.** Applications of fuzzy decision making for personnel selection problem — a review // Journal of Engineering Management and Competitiveness (JEMC). 2014. Vol. 4, No. 2. P. 68–77. URL: <http://www.tfzr.rs/jemc/files/Vol4No2/V4N22014-02.pdf>

31. **Mammadova M., Jabrayilova Z., Nobari S.** Application of TOPSIS method in support of decisions made in staff management

issues / IV International Conference "Problems of Cybernetics and Informatics" (PCI–2012). Baku. 2012. Vol. 4. P. 195–198. URL: <http://www.pci2012.science.az/8/11.pdf>

32. **Mammadova M. H., Jabrayilova Z. Q., Mammadzada F. R.** Application of Fuzzy Optimization Method in Decision-making for Personnel Selection // Intelligent Control and Automation. 2014. Vol. 5, No 4. P. 190–204. URL: [file:///C:/Users/HP/Downloads/ICA_2014101511233964%20\(1\).pdf](file:///C:/Users/HP/Downloads/ICA_2014101511233964%20(1).pdf)

33. **Мамедова М. Г., Джабраилова З. Г.** Нечеткая многокритериальная модель поддержки принятия решений в задачах управления персоналом // Проблемы информационных технологий. 2012. № 2. С. 37–46. URL: <http://ipit.az/storage/files/article/1736718f0ccee21e4cbd816d9dcb28d2.pdf>

34. **Saaty T. L., Cho Y.** The decision by the US Congress on China's trade status: A multicriteria analysis // Socio-Economic Planning Sciences. 2001. Vol. 35, No. 6. P. 243–252.

35. **Belton V., Stewart T.** Multiple criteria decision analysis: An Integrated Approach. Kluwer Academic Publishers. Massachusetts. 2002. 118 p.

36. **Карелин В. П.** Модели и методы представления знаний и выработки решений в интеллектуальных информационных системах с нечеткой логикой // Вестник ТИУЭ. Таганрог. 2014. № 1. С. 75–82.

37. **Zadeh L. A.** Fuzzy Logic = Computing with Words // IEEE Transactions on Fuzzy Systems. 1996. Vol. 4, No. 2. P. 103–111.

38. **Заде Л. А.** Понятие лингвистической переменной и его применение к принятию приближенных решений. М.: Мир. 1976. 167 с.

39. **Chen C. T., Lin C. T., Huang S. F.** A fuzzy approach for supplier evaluation and selection in supply chain management // International Journal of Production Economics. 2006. Vol. 102, No. 2. P. 289–301.

40. **Саати Т. Л.** Принятие решений. Метод анализа иерархий. М.: Радио и связь, 1993. 320 с.

41. **Ногин В. Д.** Принятие решений при многих критериях. С-Пб.: 2007. 103 с.

42. **Hsu H. M., Chen C. T.** Fuzzy credibility relation method for multiple criteria decision-making problems // Information Sciences. 1997. Vol. 96, No. 1–2. P. 79–91.

43. **Jabrayilova Z. G., Nobari S. M.** Processing methods of information about the importance of the criteria in the solution of personnel management problems and contradiction detection // Problems of information technology. 2011. No. 2. P. 57–66. URL: <http://ipit.az/storage/files/article/8a78f78a95641546667c863d448bfa7d.pdf>

M. H. Mammadova, D. Sc. of eng., Prof., Head of Department IIT of ANAS

Z. G. Jabrayilova, Phd of eng., Assistant Prof., Chief Researcher IIT of ANAS

Institut of Information Technology of Azerbaijan National Academy of Science

AZ1141, Azerbaijan Republic, Baku c., B. Vahabzadeh str., 9 depart15@iit.ab.az, www.ikt.az

The Methodological Approach to Multi-Criteria Decision-making in Human Resource Management

The methodological approach to decision-making in human resource management (HRM) is developed. The paper presents a generalized conceptual decision-making model in HRM. To ensure the adaptability of multi-criteria decision-making in HRM a modified TOPSIS method is proposed. Introducing additional components into the decision-making algorithm, this modification excludes the hierarchical structure of criteria, and takes into account the competence of experts. The method is tested on the employment case study.

Keywords: human resource management, decision-making, fuzzy environment, multi-criteria optimization, experts' competence, TOPSIS method

References

1. **Cole G. A.** *Personnel and Human Resource Management*, Wadsworth, Thomson Learning, 5 ed. 2002. 448 p.

2. **Spencer L. M., Spencer S. M.** *Competence at work models for superior performance*, Wiley India Pvt. Limited. 2008. 384 p.

3. **Armstrong M.** *Armstrong's handbook of strategic human resource management*, 10th ed., London, Kogan Page, 2006, 216 p.

4. **Makarova I. K.** *Upravlenie chelovecheskimi resursami. Pyat urovnev effektivnogo HR-menedzhmenta*, Moscow, Delo, 2007, 232 p. URL: <http://www.alleng.ru/d/manag/man094.htm>

5. **Ivantsevich Dzh. M., Lobanov A. A.** *Chelovecheskie resursy upravleniya*, Moscow, Aspekt Press, 2004, 245 p.

6. **Larichev O. I.** Teoriya i metody prinyatiya resheniy, Ucheb-nik, Izd. 2-nd, Moscow, LogoS, 2002, 392 p.
7. **Mikoni S. V.** Mnogokriterialnyy vybor na konechnom mnozhestve alternativ, Sankt-Peterburg, Lan, 2009, 272 p.
8. **Orlovskiy S. A.** Problemy prinyatiya resheniy pri nechetkoy ishodnoy informatsii, Moscow, Nauka, 1981, 212 p.
9. **Mamedova M. G.** Prinyatie resheniy na osnove baz znaniy s nechetkoy relyatsionnoy strukturoy, Baku, Elm, 1997, 296 p.
10. **Kofman A.** Vvedenie v teorii nechetkikh mnozhestv, Moscow, Radio i svyaz, 1982, 432 p.
11. **Zadeh L. A.** Fuzzy Sets, *Information and control*, 1965, vol. 8, no. 3, pp. 338–351.
12. **Trahtengerts E. A.** Vozmozhnosti i reahzatsiya kompyuternykh sistem podderzhki prinyatiya resheniy, *Izvestiya RAN. Teoriya i sistemy upravleniya*, 2001, no. 3, pp. 86–103.
13. **Chien C. F., Chen L. F.** Data mining to improve personnel selection and enhance human capital: A case study in high-technology industry, *Expert Systems with Applications*, 2008, vol. 34, no. 2, pp. 280–290.
14. **Gungor Z., Serhadlioglu G., Kesen S. E.** A fuzzy AHP approach to personnel selection problem, *Applied Soft Computing*, 2009, no. 9, pp. 641–649.
15. **Chen P. C.** A Fuzzy Multiple criteria decision making model in employee recruitment, *IJCSNS International Journal of Computer Science and Network Security*, 2009, vol. 9, no. 7, pp. 113–117.
16. **Mohsen Varmazyar, Behrouz Nouri.** A fuzzy AHP approach for employee recruitment, *Decision Science Letters*, 2014, no. 3, pp. 27–36.
17. **Hwang C. L., Yoon K.** *Multiple attributes decision making methods and applications*, Heidelberg, Berlin: Springer, 1981, 191 p.
18. **Chen C. T.** Extensions of the TOPSIS for group decision-making under fuzzy environment, *Fuzzy Sets and Systems*, 2000, vol. 114, no. 1, pp. 1–9.
19. **Dursun M., Karsak E.** A fuzzy MCDM approach for personnel selection, *Expert Systems with Applications*, 2010, vol. 37, no. 6, pp. 4324–4330.
20. **Saghafian S., Hejazi S. R.** Multi-criteria group decision making using a modified fuzzy TOPSIS procedure, *Computational Intelligence for Modelling, Control and Automation, 2005 and International Conference on Intelligent Agents, Web Technologies and Internet Commerce-International Conference*, Vienna, 2005, 28–30 Nov, vol. 2, pp. 215–221.
21. **Kelemenis A., Askounis D.** A new TOPSIS-based multi-criteria approach to personnel selections, *Expert Systems with Applications*, 2010, vol. 37, no. 6, pp. 4999–5008.
22. **Kelemenis A., Ergazakis K., Askounis D.** Support managers' selection using an extension of fuzzy TOPSIS, *Expert Systems with Applications*, 2011, vol. 38, no. 3, pp. 2774–2782.
23. **Polychroniou P. V., Giannikos I.** A fuzzy multicriteria decision-making methodology for selection of human resources in a Greek private bank, *Career Development International*, 2009, vol. 14, no. 4, pp. 372–387.
24. **Nobari S.** Design of fuzzy decision support system in employee recruitment // *Journal of Basic and Applied Scientific Research*, 2011, vol. 1, no. 11, pp. 1891–1903.
25. **Akhlagh E.** A rough-set based approach to design an expert system for personnel selection, *World Academy of Science, Engineering and Technology*, 2011, vol. 54, pp. 202–205. URL: <http://waset.org/Publications/a-rough-set-based-approach-to-design-an-expert-system-for-personnel-selection/14092>
26. **Mamedova M. G., Dzhabrailova Z. G.** Nechetkiy logicheskii podhod zadache otsenki kadrovogo potentsiala, *Menedzhment v Rossii i za rubezhom*, 2004, no. 5, pp. 111–117. URL: <http://mevriz.ru/annotations/2004/5/>
27. **Mammadova M. G., Jabrayilova Z. Q.** Methods of Family Income estimation in the targeting social Assistance System- Appl. Comput. Math. vol. 6, 2007, no. 1, pp. 80–87. URL: <http://acmij.az/view.php?lang=az&menu=iournal&id=240>
28. **Mammadova M. H., Jabrayilova Z. Q., Mammadzada F. R.** Fuzzy Decision-Making Support Methods for the Selection of IT-Professionals, *International Journal of Engineering and Innovative Technology (IJEIT)*, 2014, vol. 3, no. 7, pp. 169–175. URL: http://ijeit.com/Vol%203/Issue%207/IJEIT1412201401_31.pdf
29. **Neumann J. V., Morgenstern O.** *Theory of games ana economic behavior*. One of Princeton University presses, Notable Centenary Titles, 2007, 776 p.
30. **Afshari A. R., Nikolic M., Čočkaló D.** Applications of fuzzy decision making for personnel selection problem – a review, *Journal of Engineering Management and Competitiveness (JEMC)*, 2014, vol. 4, no. 2, pp. 68–77. URL: <http://www.tfzr.rs/jemc/files/Vol4No2/V4N22014-02.pdf>
31. **Mammadova M., Jabrailova Z., Nobari S.** Application of TOPSIS method in support of decisions made in staff management issues, *IV International Conference "Problems of Cybernetics and Informatics" (PCI-2012)*, Baku, 2012, vol. 4, pp. 195–198. URL: <http://www.pci2012.science.az/8/11.pdf>
32. **Mammadova M. H., Jabrayilova Z. Q., Mammadzada F. R.** Application of Fuzzy Optimization Method in Decision-making for Personnel Selection, *Intelligent Control and Automation*, 2014, vol. 5, no. 4, pp. 190–204. URL: [file:///C:/Users/HP/Downloads/ICA_2014101511233964%20\(1\).pdf](file:///C:/Users/HP/Downloads/ICA_2014101511233964%20(1).pdf)
33. **Mamedova M. G., Dzhabrailova Z. G.** Nechetkaya mnogokriterialnaya model podderzhki prinyatiya resheniy v zadachah upravleniya personalom, *Problemy informatsionnykh tehnologiy*, 2012, no. 2, pp. 37–46. URL: <http://jpit.az/storage/files/article/1736718f0ccee21e4cbd816d9dcb28d2.pdf>
34. **Saaty T. L., Cho Y.** The decision by the US Congress on China's trade status: A multicriteria analysis, *Socio-Economic Planning Sciences*, 2001, vol. 35, no. 6, pp. 243–252.
35. **Belton V., Stewart T.** *Multiple criteria decision analysis: An Integrated Approach*, Kluwer Academic Publishers, Massachusetts, 2002, 118 p.
36. **Karelin V. P.** Modeli i metody predstavleniya znaniy i vyirabotki resheniy v intejktualnykh informatsionnykh sistemah s nechetkoy logikoy, *Vestnik TIUE, Taganrog*, 2014, no. 1, pp. 75–82.
37. **Zadeh L. A.** Fuzzy Logic = Computing with Words, *IEEE Transactions on Fuzzy Systems*, 1996, vol. 4, no. 2, pp. 103–111.
38. **Zade L. A.** *Ponyatie lingvisticheskoy peremennoy i ego primenenie k prinyatiyu priblizhennykh resheniy*, Moscow, Mir, 1976, 167 p.
39. **Chen C. T., Lin C. T., Huang S. F.** A fuzzy approach for supplier evaluation and selection in supply chain management, *International Journal of Production Economics*, 2006, vol. 102, no. 2, pp. 289–301.
40. **Saati T. L.** *Prinyatie resheniy. Metod analiza ierarhiy*, Moscow, Radio i svyaz, 1993, 320 p.
41. **Nogin V. D.** *Prinyatie resheniy pri mnogih kriteriyah*, S-Pb., 2007, 103 p.
42. **Hsu H. M., Chen C. T.** Fuzzy credibility relation method for multiple criteria decision-making problems, *Information Sciences*, 1997, vol. 96, no. 1–2, pp. 79–91.
43. **Jabrailova Z. G., Nobari S. M.** Processing methods of information about the importance of the criteria in the solution of personnel management problems and contradiction detection, *Problems of information technology*, 2011, no. 2, pp. 57–66. URL: <http://jpit.az/storage/files/article/8a78f78a95641546667c863d4448bfa7d.pdf>

Адрес редакции:

107076, Москва, Стромынский пер., 4

Телефон редакции журнала (499) 269-5510

E-mail: it@novtex.ru

Технический редактор *Е. В. Конова*.

Корректор *Т. В. Пчелкина*.

Сдано в набор 07.04.2016. Подписано в печать 23.05.2016. Формат 60×88 1/8. Бумага офсетная.

Усл. печ. л. 8,86. Заказ ПГ616. 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.