

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

4(212)
2014

ТЕОРЕТИЧЕСКИЙ И ПРИКЛАДНОЙ НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

Издается с ноября 1995 г.

УЧРЕДИТЕЛЬ

Издательство "Новые технологии"

СОДЕРЖАНИЕ

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ

Астахова И. Ф., Ушаков С. А. Применение искусственных иммунных систем для распараллеливания процесса вычисления 3

МОДЕЛИРОВАНИЕ И ОПТИМИЗАЦИЯ

Зак Ю. А. Сужение множества Парето на основе методов нечеткой логики в задачах многокритериального выбора 7

Семькина Н. А., Смирнов А. А. Применение генетического алгоритма для решения одной задачи оптимального управления 12

Мистров Л. Е. Метод оптимизации облика иерархических систем информационной безопасности 18

Маничев В. Б., Жук Д. М., Витюков Ф. Т. Достоверное и точное численное моделирование динамических систем при невысоких требованиях к математической точности конечных результатов 27

Бронштейн Е. М., Зелёв П. А. Задача маршрутизации транспортного средства с учетом зависимости стоимости перевозок от загрузки 33

ГЕОИНФОРМАЦИОННЫЕ СИСТЕМЫ

Афанасьев А. А., Замятин А. В. Оценка применимости подходов к идентификации изменений ландшафтного покрова по данным дистанционного зондирования Земли 38

ОБРАБОТКА ИЗОБРАЖЕНИЙ

Терехин А. В. Распознавание трехмерных объектов по изображениям двух проекций 43

Балонин Н. А., Балонин Ю. Н., Сергеев М. Б. О модификации метода Скарпи вычисления матриц Мерсенна для задач преобразования изображений 48

СЕТИ И СИСТЕМЫ СВЯЗИ

Перепелкин Д. А. Динамическое формирование структуры и параметров линий связи корпоративной сети на основе данных о парных перестановках маршрутов 52

БЕЗОПАСНОСТЬ ИНФОРМАЦИИ

Волосатова Т. М. Исследования стеганографических методов защиты проектной документации от несанкционированного доступа 61

ИНФОРМАЦИЯ

XVI Всероссийская научная конференция RCDL-2014 "Электронные библиотеки: перспективные методы и технологии, электронные коллекции" 72

Приложение. Нейрокомпьютеры с использованием мемристоров

Главный редактор:

СТЕМПКОВСКИЙ А. Л.,
акад. РАН, д. т. н., проф.

Зам. главного редактора:

ИВАННИКОВ А. Д., д. т. н., проф.
ФИЛИМОНОВ Н. Б., д. т. н., с.н.с.

Редакционный совет:

БЫЧКОВ И. В., акад. РАН, д. т. н.

ЖУРАВЛЕВ Ю. И.,
акад. РАН, д. ф.-м. н., проф.

КУЛЕШОВ А. П.,
акад. РАН, д. т. н., проф.

ПОПКОВ Ю. С.,
чл.-корр. РАН, д. т. н., проф.

РУСАКОВ С. Г.,
чл.-корр. РАН, д. т. н., проф.

СОЙФЕР В. А.,
чл.-корр. РАН, д. т. н., проф.

СОКОЛОВ И. А., акад.
РАН, д. т. н., проф.

СУЕТИН Н. В., д. ф.-м. н., проф.

ЧАПЛЫГИН Ю. А.,
чл.-корр. РАН, д. т. н., проф.

ШАХНОВ В. А.,
чл.-корр. РАН, д. т. н., проф.

ШОКИН Ю. И.,
акад. РАН, д. т. н., проф.

ЮСУПОВ Р. М.,
чл.-корр. РАН, д. т. н., проф.

Редакционная коллегия:

АВДОШИН С. М., к. т. н., доц.
АНТОНОВ Б. И.

БАРСКИЙ А. Б., д. т. н., проф.

ВАСЕНИН В. А., д. ф.-м. н., проф.

ГАЛУШКИН А. И., д. т. н., проф.

ДИМИТРИЕНКО Ю. И., д. ф.-м. н., проф.

ДОМРАЧЕВ В. Г., д. т. н., проф.

ЗАГИДУЛЛИН Р. Ш., к. т. н., доц.

ЗАРУБИН В. С., д. т. н., проф.

ИСАЕНКО Р. О., к. т. н., с.н.с.

КАРПЕНКО А. П., д. ф.-м. н., проф.

КОЛИН К. К., д. т. н., проф.

КУЛАГИН В. П., д. т. н., проф.

КУРЕЙЧИК В. М., д. т. н., проф.

КУХАРЕНКО Б. Г., к. ф.-м. н., доц.

ЛЬВОВИЧ Я. Е., д. т. н., проф.

МИХАЙЛОВ Б. М., д. т. н., проф.

НЕЧАЕВ В. В., к. т. н., проф.

РЯБОВ Г. Г., чл.-корр. РАН, д. т. н., проф.

СОКОЛОВ Б. В., д. т. н., проф.

УСКОВ В. Л., к. т. н. (США)

ФОМИЧЕВ В. А., д. т. н., проф.

ЧЕРМОШЕНЦЕВ С. Ф., д. т. н., проф.

ШИЛОВ В. В., к. т. н., доц.

Редакция:

БЕЗМЕНОВА М. Ю.

ГРИГОРИН-РЯБОВА Е. В.

ЛЫСЕНКО А. В.

ЧУГУНОВА А. В.

Информация о журнале доступна по сети Internet по адресу <http://novtex.ru/IT>.

Журнал включен в систему Российского индекса научного цитирования.

Журнал входит в Перечень научных журналов, в которых по рекомендации ВАК РФ должны быть опубликованы научные результаты диссертаций на соискание ученой степени доктора и кандидата наук.

CONTENTS

INTELLIGENT SYSTEMS AND TECHNOLOGIES

Астасова И. Ф., Ушаков С. А. Application Artificial Immanuel System for Converting to Parallel of Process Calculation 3

MODELING AND OPTIMIZATION

Зак Ю. А. Narrowing of the Pareto-Based Methods Fuzzy Logic in the Problems of Multi-Criteria Secetction. 7

Семькина Н. А., Смירнов А. А. The Use of Genetic Algorithm for Solution of the Problem of Optimal Control 12

Мистров Л. Е. Method of Optimization of Shape of the Hierarchical Systems of Information Security 18

Маничев В. Б., Жук Д. М., Витукوف Ф. Т. Method of Optimization of Shape of the Hierarchical Systems of Information Security 27

Бронштейн Е. М., Зельков П. А. The Vehicle Routing Problem with the Transportation Cost Dependence of Vehicle Loading 33

GEOINFORMATION SYSTEMS

Афанасьев А. А., Замыатин А. В. The Applicability Analysis of the Approaches to the Identification of Land Cover Changes by Remote Sensing Data 38

IMAGE PROCESSING

Терехин А. В. Three-Dimensional Object Recognition by Images of Two Projections . 43

Балонин Н. А., Балонин Я. Н., Сергеев М. Б. Scarpis Method Calculation of the Hadamard and Mersenne Matrices 48

NETWORKS AND COMMUNICATION SYSTEMS

Перепелкин Д. А. Dynamic Corporate Network Structure and Communication Links Loading Formation Based on Routes Pairs Permutations Data. 52

CRYPTOSAFETY INFORMATION

Волосатова Т. М. Researches of Steganographic Methods to Protect the Design Documentation from Unauthorized Access 61

Information 72

APPLICATION. Neurocomputers with Using Memristors

Editor-in-Chief:

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

Deputy Editor-in-Chief:

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

Chairman:

Bychkov I. V., Member of RAS,
Dr. Sci. (Tech.), Prof.
Zhuravljov Yu. I., Member of RAS,
Dr. Sci. (Phys.-Math.), Prof.
Kuleshov A. P., Member of RAS,
Dr. Sci. (Tech.), Prof.
Popkov Yu. S., Corresp. Member of RAS,
Dr. Sci. (Tech.), Prof.
Rusakov S. G., Corresp. Member of RAS,
Dr. Sci. (Tech.), Prof.
Soifer V. A., Corresp. Member of RAS,
Dr. Sci. (Tech.), Prof.
Sokolov I. A., Member of RAS,
Dr. Sci. (Phys.-Math.), Prof.
Suetin N. V.,
Dr. Sci. (Phys.-Math.), Prof.
Chaplygin Yu. A., Corresp. Member of RAS,
Dr. Sci. (Tech.), Prof.
Shakhnov V. A., Corresp. Member of RAS,
Dr. Sci. (Tech.), Prof.
Shokin Yu. I., Member of RAS,
Dr. Sci. (Tech.), Prof.
Yusupov R. M., Corresp. Member of RAS,
Dr. Sci. (Tech.), Prof.

Editorial Board Members:

Avdoshin S. M., Cand. Sci. (Tech.), Ass. Prof.
Antonov B. I.
Barsky A. B., Dr. Sci. (Tech.), Prof.
Vasenin V. A., Dr. Sci. (Phys.-Math.), Prof.
Galushkin A. I., Dr. Sci. (Tech.), Prof.
Dimitrienko Yu. I., Dr. Sci. (Phys.-Math.), Prof.
Domrachev V. G., Dr. Sci. (Tech.), Prof.
Zagidullin R. Sh., Cand. Sci. (Tech.), Ass. Prof.
Zarubin V. S., Dr. Sci. (Tech.), Prof.
Isaenko R. O., Cand. Sci. (Tech.)
Karpenko A. P., Dr. Sci. (Phys.-Math.), Prof.
Kolin K. K., Dr. Sci. (Tech.)
Kulagin V. P., Dr. Sci. (Tech.), Prof.
Kurechik V. M., Dr. Sci. (Tech.), Prof.
Kukhareenko B. G., Cand. Sci. (Phys.-Math.)
Ljvovich Ya. E., Dr. Sci. (Tech.), Prof.
Mikhailov B. M., Dr. Sci. (Tech.), Prof.
Nechaev V. V., Cand. Sci. (Tech.), Ass. Prof.
Ryabov G. G., Corresp. Member of RAS,
Dr. Sci. (Tech.), Prof.
Sokolov B. V., Dr. Sci. (Tech.)
Uskov V. L. (USA), Dr. Sci. (Tech.)
Fomichev V. A., Dr. Sci. (Tech.), Prof.
Chermoshentsev S. F., Dr. Sci. (Tech.), Prof.
Shilov V. V., Cand. Sci. (Tech.), Ass. Prof.

Editors:

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

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

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

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ INTELLIGENT SYSTEMS AND TECHNOLOGIES

УДК 004.8

И. Ф. Астахова, д-р техн. наук, проф., e-mail: astachova@list.ru, С. А. Ушаков, аспирант,
Воронежский государственный университет

Применение искусственных иммунных систем для распараллеливания процесса вычисления

Рассматривается новый механизм иммунных систем для распараллеливания процессов вычисления символьной регрессии.

Ключевые слова: иммунные системы, параллельные алгоритмы, иммунные сети

I. F. Astachova, S. A. Ushakov

Application Artificial Immanuel System for Converting to Parallel of Process Calculation

In this article a new mechanism immanuel system for converting to parallel the processes calculation of symbol regression is considered.

Keywords: immanuel system, parallel algorithms, immanuel net

Введение

Работа посвящена особенностям применения искусственных иммунных систем для распараллеливания процесса вычисления оптимальной структуры регрессионной модели нескольких свободных переменных.

Символьная регрессия — метод построения регрессионных моделей путем перебора различных произвольных суперпозиций функций из некоторого заданного набора. Суперпозиция функций при этом называется "программой", а стохастический оптимизационный алгоритм построения таких суперпозиций называется генетическим программированием. Однако вместо обычно используемых генетических алгоритмов будет использоваться искусственная иммунная система.

Есть множество значений свободных переменных x_1, \dots, x_r , где $x_i \in R_n$, и соответствующих им значений функции y_1, \dots, y_r . Эти два множества образуют множество исходных данных D .

Также задано множество функций, которые будут использованы при построении суперпозиции. Будем рассматривать только непрерывно дифференцируемые функции $g: R_n \rightarrow R$, такие как \sin , \cos , полиномы.

Рассмотрим произвольную суперпозицию f , состоящую из не более m функций g . Требуется найти такую суперпозицию, которая бы доставляла максимум (минимум) заданного функционала $p(f, D)$. Этот функционал и определяет целевую функцию. Его выбирают таким образом, чтобы он показывал степень приближения построенной функции к истинной. Для этого можно использовать различные метрики и расстояния, относительную или абсолютную погрешность — разницу исходных значений и значений полученной функции в этих же точках.

Таким образом, входными данными для задачи является набор точек n -мерного пространства и значений функции в этой точке. Решением задачи является функция, зависящая от n переменных, наилучшим образом аппроксимирующая исходную (в заданных точках).

Описание используемой искусственной иммунной системы

Первой задачей является определение того, как будут представлены лимфоциты. Так как решением является функция (суперпозиция заданных простых функций), то лимфоцит представляет собой функцию, закодированную каким-либо образом.

Удобным представлением как для реализации, так и для последующей работы алгоритма можно считать дерево выражения. Так как использоваться будут только бинарные и унарные операции, данное дерево будет бинарным.

Лимфоциты должны поддерживать следующие операции [1–3]:

- вычисление значения представляемой функции в точке;
- вычисление меры аффинности (рассмотрена ниже) данного лимфоцита;
- упрощение соответствующего выражения;
- возвращение строкового представления выражения (с правильно расставленными скобками).

Введем понятие *аффинности*. В биологии аффинность антител — прочность связывания активных центров молекулы антитела с детерминантными группами антигена. Антигеном является множество входных данных — значений функций в заданных точках. В данной модели аффинность характеризует степень "успешности" аппроксимации, т. е. является значением целевой функции.

Рассмотрим *алгоритм функционирования искусственной иммунной системы* [4–8]:

Шаг 1. В используемой системе первоначальный набор лимфоцитов будет сгенерирован случайным образом на основании заданных переменных, используемых функций и максимально допустимой высоты дерева выражения. Числа, используемые в дереве выражения, берутся из наперед заданного отрезка.

Шаг 2. В качестве основного алгоритма иммунной сети используется алгоритм селекции и клонирования (Clonal Selection Algorithm). Идея этого алгоритма заключается в том, что со временем лимфоциты улучшают свою меру аффинности к определенным антигенам. Алгоритм полагается на дарвиновскую теорию, в которой степень приспособленности организма — мера аффинности, репродукция осуществляется с помощью деления клеток (лимфоцитов), вариативность — соматической гипермутацией.

Рассмотрим, как в данной искусственной иммунной сети можно реализовать шаг 2 этого алгоритма.

Шаг 3. Пусть начальный набор лимфоцитов уже сформирован, для каждого из них вычисляется мера аффинности. Затем из всего набора лимфоцитов выбирается определенное число наиболее сильно реагирующих (наиболее приспособленных в терминах алгоритма селекции и клонирования). Оставшиеся не могут справиться со своей задачей, поэтому они удаляются из рассматриваемой модели. К выбранным лимфоцитам применяется процедура гипермутации для обеспечения разнообразия набора и возможно лучшего приближения к ответу. Гипермутация для лимфоцитов, представляющих собой дерево выражений, может состоять в замене

случайно выбранного действия на другое, изменении числовой константы или переменной, добавлении/удалении поддерева случайно выбранного узла. Для обеспечения постоянного числа лимфоцитов в активном наборе можно либо использовать только сохранение текущих лучших лимфоцитов вместе с их гипермутировавшими потомками, либо недостающие лимфоциты заново генерировать случайным образом ("свежая кровь" — в организме новые лимфоциты производятся костным мозгом постоянно). После того как в активном наборе лимфоцитов будет необходимое число клеток, данный шаг повторяется. Критерием остановки является либо достижение приемлемого значения целевой функции, либо проведение максимально допустимого числа итераций. Решением является функция, представляемая лимфоцитом из текущего набора с наибольшей мерой аффинности (рис. 1).

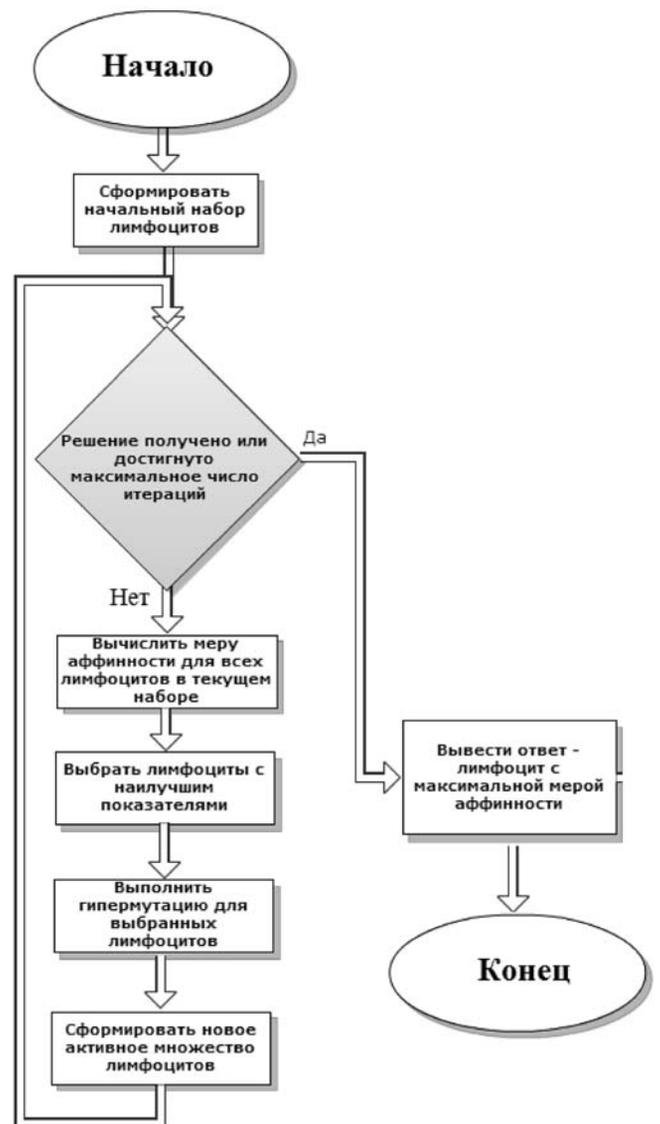


Рис. 1. Алгоритм функционирования искусственной иммунной системы

Описание распределенной версии искусственной иммунной системы

В последние годы широко применяются распределенные вычисления — способ решения трудоемких вычислительных задач с использованием нескольких компьютеров, чаще всего объединенных в параллельную вычислительную систему [wiki]. Например, Массачусетский Технологический институт создал распределенную систему FlexGP, представляющую собой генетический алгоритм, выполняющийся в облачной инфраструктуре Amazon [6]. Особенностью распределенных многопроцессорных вычислительных систем, в отличие от локальных суперкомпьютеров, является возможность неограниченного наращивания производительности за счет масштабирования. Поэтому создание распределенной версии используемой искусственной иммунной системы позволит проводить затратные вычисления намного быстрее при использовании нескольких вычислительных систем. В отличие от популярных технологий параллелизма приложений (таких как OpenMP и MPI) данная распределенная система будет работать в гетерогенной среде — вычислительной системе, состоящей из различных вычислительных блоков. В данном случае под вычислительным блоком подразумевается любой современный компьютер, поддерживающий сетевой стек TCP/IP и выбранный язык программирования — Python. Очевидно, что класс таких устройств в данный момент чрезвычайно широк.

В качестве архитектуры взаимодействия вычислительных блоков между собой была выбрана модель p2p взаимодействия. Одноранговая децентрализованная, или пиринговая (от англ. peer-to-peer, p2p — равный к равному), сеть — это компьютерная сеть, основанная на равноправии участников. В такой сети отсутствуют выделенные серверы, а каждый узел (peer) является как клиентом, так и сервером. В отличие от архитектуры клиент-сервера такая организация позволяет сохранять работоспособность сети при любом числе и любом сочетании доступных узлов.

Идея распределенной иммунной системы была навеяна тем соображением, что иммунные клетки распределены по всему организму, но в то же время информация доступна всему организму целиком. Таким образом, придем к следующей модели, алгоритм функционирования которой представлен на рис. 2.

На каждом вычислительном узле функционирует искусственная иммунная система, описанная в предыдущем разделе, т. е. сначала создается набор случайно сгенерированных лимфоцитов, затем происходят операции выбора и мутации. Однако у каждого узла есть информация по крайней мере об одном другом функционирующем узле. Поэтому один раз в определенное число итераций после шага гипермутации выбранных лимфоцитов данная иммунная система запрашивает и получает набор луч-

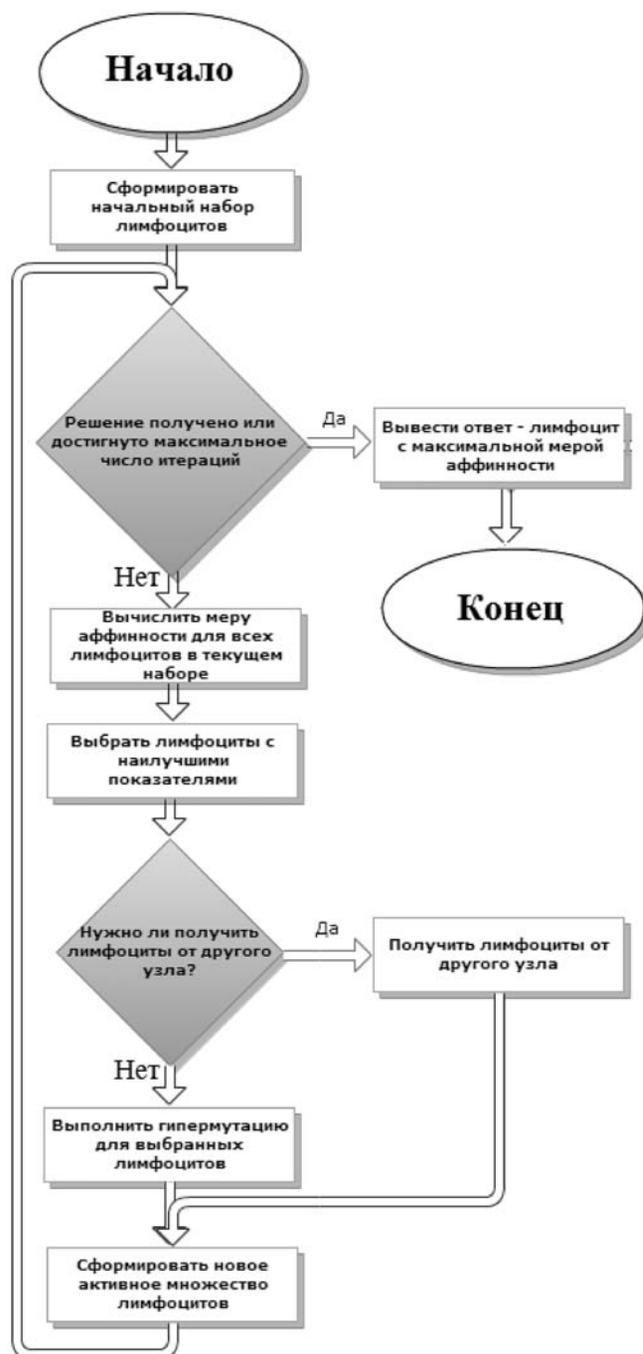


Рис. 2. Алгоритм функционирования распределенной искусственной иммунной системы

ших лимфоцитов другого узла. Полученные лимфоциты участвуют в процессе создания нового активного множества лимфоцитов. Аналогично, данный узел при обращении к нему другого узла отправит последнему свой набор лучших лимфоцитов. Таким образом, распределенная система позволяет узлам обмениваться своими самыми успешными результатами. При этом при выходе одного или нескольких узлов из строя работа сети не остановится, просто будут опрошены другие узлы, еще находящиеся в сети и работающие. Для узла, начавшего работу,

когда некоторые узлы уже добились некоторых результатов, работа начнется не с самого начала, он получит результаты от уже некоторое время работающих узлов.

Реализация распределенной искусственной иммунной системы

Для реализации был использован язык Python версии 3.3. Данный язык был выбран по следующим причинам:

- *Кросс-платформенность.* При создании приложения, работающего в гетерогенной вычислительной среде, данный фактор является очень важным. Python позволяет создать программу, не требующую модификации исходного кода и даже перекомпиляции для работы на различных платформах.
- *Быстрота разработки.* Python ориентирован на повышение производительности разработчика и читаемости кода. Также он поддерживает различные парадигмы программирования: структурное, объектно-ориентированное, функциональное.
- *Большой объем стандартной библиотеки.* Она включает большое число полезных функций и классов для различных задач, таких как сетевое взаимодействие, многопоточность, управление процессами.

При создании программного продукта использовалась система контроля версий Git, исходный код хранится на github. Программный комплекс состоит из следующих классов:

- классы для представления лимфоцитов (рис. 3, см. третью сторону обложки);
- классы для представления искусственной иммунной системы (рис. 4, см. третью сторону обложки);
- классы для представления сетевого взаимодействия (рис. 5, см. третью сторону обложки); классы для тестирования логики работы и запуска программы.

Программа main.py состоит из следующих основных частей:

- инициализация объекта ExpressionsImmuneSystemConfig;
- получение входных данных — списка переменных и значений функции в точках;
- создание объекта SimpleRandomExchanger, который будет использоваться при создании искусственной иммунной системы;
- основной цикл — создание нескольких иммунных систем с заданными параметрами и получение решения. До цикла и после вызываются функции замера времени, выводится общее время работы всех итераций. Во время работы этого цикла выводится строка состояния, показывающая, какой процент от общего числа итераций уже завершился.

```

C:\Windows\system32\cmd.exe
C:\Users\Stas\PycharmProjects\ArtificialImmuneSystem>python main.py
[#####] 100%
1429.6912097281572 seconds
(1.1893395722361453e-14, '((x * x) / (y / y)) + ((1.0 + y) * (x * sin(x)))')
(2.00258506493651, '((x * x) / (28.618 + y)) + ((26.804 - y) - (y + y)))')
(3.24518607886285, '((x * x) / (38.357 + y)) + ((sin(y) - y) + (x * x)))')
(52.401254687669244, '((x * x) - (x * sin(x))) + ((x * x) / (-33.506 / y)))')
(57.79809620139138, '((x * x) - (1.0 - cos(x))) / 0.846')
(58.37500890940489, '(cos((x * x)) + (x * x)) + sin((23.936 / y) / (y - y)))')
(58.42077546530042, '((y / 0.616) - (y + x)) - ((y - x) - (x * x)))')
(58.94511496775412, '((x * x) + -2.442) + ((y - 79.859) / (-38.009 - y)))')
(59.45162064226311, '(x * x)')
(98.33321638140461, '((y + y) / 0.273) / y')
C:\Users\Stas\PycharmProjects\ArtificialImmuneSystem>

```

Рис. 6. Интерфейс и запуск программы

Пример запуска скрипта main.py и внешний вид консоли во время его выполнения представлен на рис. 6.

Заключение

В данной работе была рассмотрена возможность применения искусственной иммунной системы для решения задачи символьной регрессии. Также было предложено улучшение используемой системы, заключающееся в создании распределенной версии алгоритма. Была предложена компьютерная реализация распределенной иммунной системы, проверена ее работоспособность.

Список литературы

1. **Искусственные** иммунные системы и их применение / Под ред. Д. Дасгупты. М.: Физматлит. 2006. 344 с.
2. **Johnson C. G.** Artificial Immune Systems Programming for Symbolic Regression / Genetic Programming: 6th European Conference. 2003. P. 345—353.
3. **Farmer J. D., Packard N. and Perelson A.** The immune system, adaptation and machine learning // Physica D. 2004. Vol. 2. P. 187—204.
4. **Artificial Immune Systems and Their Applications** / D. Dasgupta (Editor). Berlin: Springer-Verlag, Inc., January. 1999. 320 p.
5. **Garrett S.** How Do We Evaluate Artificial Immune Systems? // Evolutionary Computation. 2005. Vol. 13, N 2. P. 145—178.
6. **Cutello V. and Nicosia G.** An Immunological Approach to Combinatorial Optimization Problems. // Lecture Notes in Computer Science, Springer. 2002. Vol. 2527. P. 361—370.
7. **Сайт** Массачусетского Технологического Института. URL: http://people.csail.mit.edu/dsherry/www/papers/EvoPAR_FlexGP_pre.pdf (дата обращения 20.02.2013).
8. **GitHub:** [сайт] URL: <https://github.com/StaniislavUshakov/ArtificialImmuneSystem> (дата обращения 20.05.2013).
9. **Челнаков С. А., Ушаков С. А.** Иммунные системы для решения практических задач // Вопросы современной науки и практики. 2012. № 2 (40). С. 74—78.
10. **Bersini N., Varela F.** Hints for adaptive problem solving gleaned from immune networks // Proc. First Conf. on Parallel Problem Solving from Nature. 1990. P. 343—354.
11. **Bersini H.** Immune network and adaptive control // Proc. First European Conf on Artificial Life / Eds. F J. Varela and P. Bourgine. MIT Press, 1991. P. 461—465.
12. **Bersini N., Varela F.** The immune learning mechanisms: reinforcement, recruitment and their application // Computing with Biological Metaphors / Ed. R. Paton. 1994. P. 166—192.
13. **Hunt I. E., Cooke D. E.** Learning using an artificial immune system // J. Netw. Corp. Appl. (Special issue on Intelligent Systems: Design and Application). 1996. Vol. 19. P. 189—212.
14. **Марчук Г. И.** Математические модели в иммунологии. Вычислительные методы и эксперименты. М.: Наука, 1991. 304 с.

МОДЕЛИРОВАНИЕ И ОПТИМИЗАЦИЯ MODELING AND OPTIMIZATION

УДК 004.421.2:519.8

Ю. А. Зак, д-р техн. наук, науч. эксперт, консультант, Аахен, Германия,
e-mail: yuriy_zack@hotmail.com

Сужение множества Парето на основе методов нечеткой логики в задачах многокритериального выбора

Рассматриваются и анализируются подходы к решению проблемы сужения множества Парето на основе методов нечеткой логики, использующие как фактические значения локальных критериев оптимальности, так и лингвистические оценки их значений экспертами. Рассматриваемые модели многокритериального выбора основаны на введении различного вида компромиссных fuzzy-логических векторных критериев и ограничений, учитывающих предпочтения лица, принимающего решения.

Ключевые слова: Парето-эффективные решения, нечеткая логика, относительная важность критериев, функции принадлежности, лингвистические переменные

Yu. A. Zack

Narrowing of the Pareto-Based Methods Fuzzy Logic in the Problems of Multi-Criteria Selection

Reviews and analyzes approaches to the problem of narrowing of the Pareto-based fuzzy logic methods that use the actual values of the local optimality criteria and linguistic experts to assess their values. These models are the basis of multi-criteria selection on the introduction of various types of compromise Fuzzy-logic vector criteria and constraints, taking into account the preferences of the decision maker.

Keywords: Pareto-efficient solutions, fuzzy logic, the relative importance of the criteria, the membership functions, linguistic variables

Введение

Эффективность по Парето является одним из самых важных критериев для оценки экономических систем и выбора решений среди множества рассматриваемых альтернатив (см., например, [1, 6, 7, 9]). При большом числе частных показателей эффективности множество Парето может включать большое число альтернатив. Одной из наиболее важных проблем в задачах многокритериальной оптимизации является выбор среди альтернатив, принадлежащих множеству Парето, наилучших, наиболее эффективных с точки зрения лица, принимающего решение (ЛПР), т. е. задача сужения этого множества Парето [2–5], исключив из рассмотрения некоторые Парето-оптимальные решения. Такое решение можно получить только за счет использования дополнительной информации, применяя некоторую формализованную процедуру. За счет привлечения дополнительной субъективной информации об относительной важности критериев предложены достаточно эффективные методы решения этой задачи, основанные на назначении различных весов

частных показателей эффективности и введении комплексных компромиссных критериев в виде различного вида сверток [3, 5], выделении различных групп важности критериев и назначении определенных уступок по менее важным с точки зрения ЛПР группам критериев.

Пусть $P(A)$ — подмножество альтернатив, принадлежащих множеству Парето. Эффективные подходы, основанные на использовании квантов информации, развиты в работах В. Д. Ногина [3–5]. При использовании квантов информации формулируется некоторое решение, в процессе которого ЛПР идет на определенный компромисс. При этом отдается предпочтение некоторой группе критериев, и ЛПР готов потерять по некоторой группе критериев (т. е. для всех $i \in \tilde{I}_1$) не более чем d_1 единиц, чтобы получить улучшение на d_2 единиц по другой группе критериев (т. е. для всех $i \in \tilde{I}_2$), где $\tilde{I}_1 \cap \tilde{I}_2$, $\tilde{I}_1 \subset \tilde{I}$, $\tilde{I}_2 \subset \tilde{I}$. На основе такого компромисса происходит исключение из рассмотрения некото-

рого подмножества альтернатив. В отличие от детерминированного подхода в случае использования теории нечетких множеств степень этого предпочтения оценивается значением некоторой функции принадлежности $\mu_F \in [0, 1]$ (см. [4, 7, 9]). Использование кванта детерминированной или нечеткой информации в некоторых ситуациях не приводит к значительному сужению множества Парето, а в случае применения данных не гарантирует однозначный выбор. Поэтому В. Д. Ногин рассматривает подходы, связанные с использованием не одного, а сразу нескольких квантов.

В связи с изложенным выше представляет интерес рассмотрение и других подходов сокращения размеров множества Парето с использованием методов нечеткого логического вывода. В качестве одного из таких методов может быть взят метод, использующий информацию об "идеальной точке", лингвистические переменные и аппарат нечеткой логики, позволяющий формализовать предпочтения экспертов ЛПР. Описанию этих подходов и посвящена данная работа.

1. Постановка и математическая формулировка задачи

Идеальной точкой назовем некоторый вектор значений локальных критериев оптимальности вида

$$\begin{aligned} \Phi(X^*) &= \{F_i^-(X_i^*), i = 1, \dots, n_1; \\ &F_i^+(X_i^*), i = 1, \dots, n_1 + 1, \dots, n\}, \end{aligned} \quad (1)$$

где

$$\begin{aligned} F_i^-(X_i^*) &= \min_{X \in G} F_i(X), i = 1, \dots, n_1; \\ F_i^+(X_i^*) &= \max_{X \in G} F_i(X), i = n_1 + 1, \dots, n. \end{aligned} \quad (2)$$

Как правило, такой точки не существует. Поэтому ее называют точкой утопии.

Метод сужения множества Парето и выделения подмножества эффективных решений (альтернатив) на основе информации об идеальной точке состоит в отыскании на границе Парето точки, ближайшей в соответствии с предпочтениями ЛПР к точке утопии. Обычно ЛПР формулирует свои предпочтения в виде значений весовых коэффициентов в компромиссном критерии в форме минимизации средневзвешенной суммы значений отклонений от так называемой идеальной точки.

Множество эффективных решений. Обозначим $F_i^-(X_i^*) = \min_{X \in G} F_i(X)$, $F_i^+(X_i^*) = \max_{X \in G} F_i(X)$, $i = 1, \dots, n_1, n_1 + 1, \dots, n$, — соответственно минимальные и максимальные значения по каждому из критериев оптимальности. Разобьем весь диапазон $D_i\{F_i(X)\} = [F_i^-(X^*), F_i^+(X^*)]$ возможных значений каждого i -го локального критерия оптимальности на P ин-

тервалов, обозначим $p = 1, \dots, P$, индексы интервалов, где $P = 3, 5, 7$:

$$d_i^p\{F_i(X)\} = [F_i^-(X^*) + (p-1)\Delta_i; F_i^-(X^*) + p\Delta_i], \quad (3)$$

где $\Delta_i = \frac{1}{P}[F_i^+(X^*) - F_i^-(X^*)]$,

$$\begin{aligned} F_i(X|A^k) \in d_i^p\{F_i(X)\}, \text{ if } [F_i^-(X^*) + (p-1)\Delta_i \leq \\ \leq F_i(X|A^k) < F_i^-(X^*) + p\Delta_i]. \end{aligned} \quad (4)$$

Среди всего множества альтернатив, удовлетворяющих системе ограничений задачи G , выделим подмножество альтернатив $\Omega(A)$, обеспечивающих выполнение логических условий вида $\Omega(A) = \{A^k, k = 1, \dots, K\}$, если в случае справедливости соотношений

$$F_i(A^k) \in \bigcup_{p=1}^{P_1(i)} d_i^p\{F_i(A^k)\}, i = 1, \dots, n_1; \quad (5)$$

$$F_i(A^k) \in \bigcup_{p=P_2(i)}^{P_1(i)} d_i^p\{F_i(A^k)\}, i = n_1 + 1, \dots, n. \quad (6)$$

Здесь $P_1(i)$, $i = 1, \dots, n_1$, и $P_2(i)$, $i = n_1 + 1, \dots, n$ — соответственно число первых и последних интервалов из всего возможного диапазона изменения локального критерия оптимальности, устанавливаемых ЛПР.

Подмножество альтернатив $\Omega(A)$ назовем подмножеством эффективных решений.

Тогда наиболее эффективное компромиссное решение следует искать среди альтернатив, принадлежащих пересечению следующих подмножеств:

$$A^k \in \{G \cap \Pi(A) \cap \Omega(A)\}, \quad (7)$$

тем самым сократив число альтернатив, входящих в множество Парето-эффективных решений.

2. Алгоритм решения задачи

Пусть в задаче максимизации n локальных критериев оптимальности $f_i(A^k)$, $i = 1, \dots, n$, необходимо выбрать наилучшее компромиссное решение из K альтернатив A^k , $k = 1, \dots, K$. Найдем значения

$$f_i^-(A) = \min_{1 \leq k \leq K} f_i(A^k),$$

$$f_i^+(A) = \max_{1 \leq k \leq K} f_i(A^k), i = 1, \dots, n, \quad (8)$$

либо

$$f_i^-(A) = \min_{1 \leq k \leq K} f_i[A^k|A^k \in G \cap \Pi(A)],$$

$$f_i^+(A) = \max_{1 \leq k \leq K} f_i[A^k|A^k \in G \cap \Pi(A)]. \quad (9)$$

Введем функции принадлежности $\mu_{f_i(A)}[f_i(A^k)] = \psi[f_i(A^k)]$ для каждого из локальных критериев

оптимальности в диапазоне значений $f_i(A^k) \in [f_i^-(A), f_i^+(A)]$, удовлетворяющие соотношениям

$$0 \leq \mu_{f_i(A)}[f_i(A^k)] \leq 1, \mu_{f_i(A)}[f_i^-(A)] = \mu_0(f_i(A)) \geq 0, \\ \mu_{f_i(A)}[f_i^+(A)] = \mu_P(f_i(A)) = 1, 0, \quad (10)$$

где P — число значений рассматриваемых интервалов значений критерия $f_i(A^k)$ в диапазоне их возможных изменений (число значений соответствующей лингвистической переменной).

Виды таких функций принадлежности, которые определяет ЛПР, могут быть произвольными. Примеры таких функций принадлежности приведены на рис. 1–5.

Для каждой из альтернатив $A^k \in G \cap \Pi(A)$, удовлетворяющей заданной системе ограничений и принадлежащей множеству Парето, вычислим значения

$$\mu_{f_i(A)}[f_i(A^k)] = \psi[f_i(A^k)].$$

Множество всех локальных критериев $\tilde{I} = \{1, \dots, n\}$ оптимальности может быть разбито на несколько различных по степени важности групп

$$\tilde{I}_1 > \tilde{I}_2 > \dots > \tilde{I}_l > \dots > \tilde{I}_L,$$

$$\bigcup_{l=1}^L \tilde{I}_l = \tilde{I}, \tilde{I}_r \cap \tilde{I}_l = \emptyset, r \neq l, r, l = 1, \dots, L. \quad (11)$$

В частном случае $L = 1$.

Для каждой группы на основе операторов нечеткой логики вычислим значение показателей по одной из следующих формул:

$$\mu_{F(\tilde{I}_l|A^k)}^1 = \min_{i \in \tilde{I}_l} \mu_{f_i(A^k)}[f_i(A^k) | A^k \in \{G \cap \Pi(A)\}], \\ l = 1, \dots, L; \quad (12)$$

$$\mu_{F(\tilde{I}_l|A^k)}^2 = L_1 \sqrt{\prod_{i \in \tilde{I}_l} \mu_{f_i(A^k)}[f_i(A^k) | A^k \in \{G \cap \Pi(A)\}]}, \\ l = 1, \dots, L; \quad (13)$$

$$\mu_{F(\tilde{I}_l|A^k)}^3 = \frac{1}{L_1} \sum_{i \in \tilde{I}_l} \mu_{f_i(A^k)}[f_i(A^k) | A^k \in \{G \cap \Pi(A)\}], \\ l = 1, \dots, L. \quad (14)$$

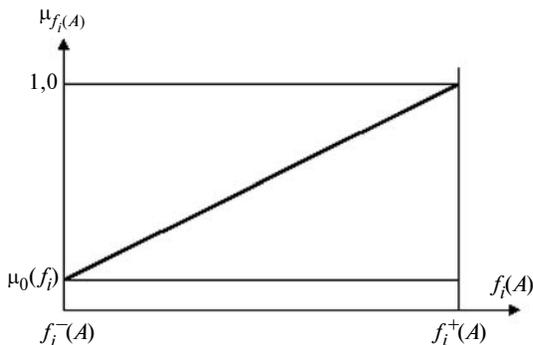


Рис. 1

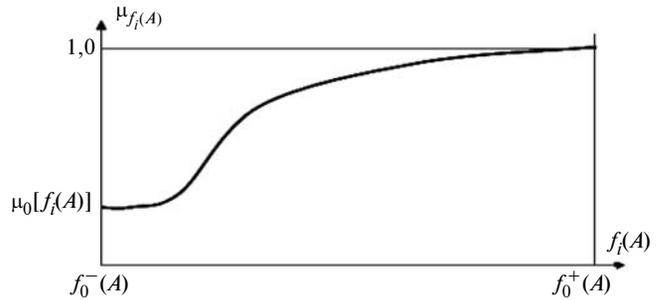


Рис. 2

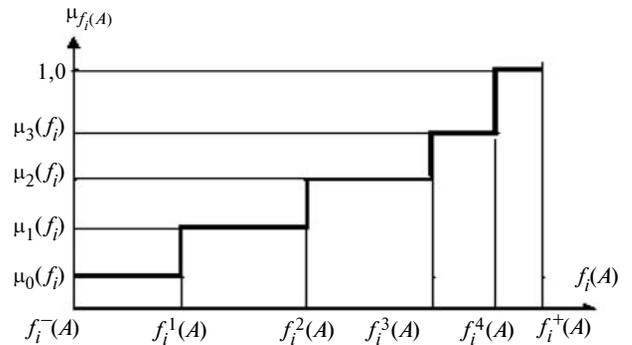


Рис. 3

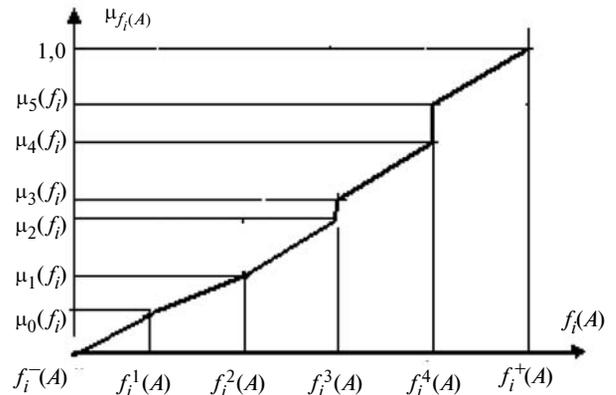


Рис. 4

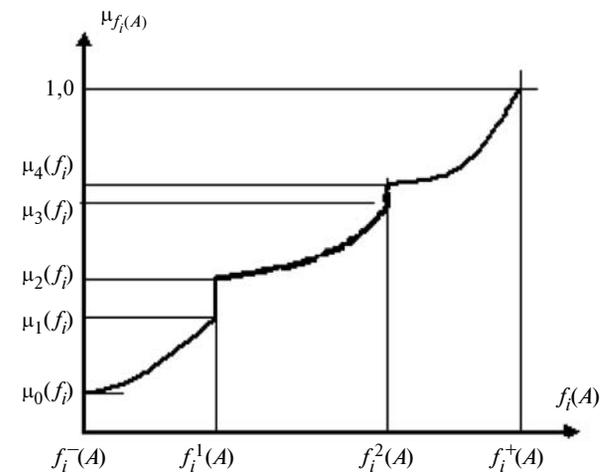


Рис. 5

Установим для каждой группы локальных критериев оптимальности некоторые предельные значения, соответственно равные

$$\bar{m}_{F(\tilde{I}_l A)}^1, \bar{m}_{F(\tilde{I}_l A)}^2, \bar{m}_{F(\tilde{I}_l A)}^3,$$

где $0 < \bar{m}_{F(\tilde{I}_l A)}^p < 1, p = 1, 2, 3$.

В качестве критериев отсева могут быть предложены следующие дополнительные требования:

1. Среди множества перспективных альтернатив $A^k \in G \cap \Pi(A)$ в качестве перспективных для дальнейшего рассмотрения оставляются такие, которые при выбранном ЛПР значении параметра p удовлетворяют одной из следующих систем неравенств:

$$\begin{aligned} \bar{\mu}_{F(\tilde{I}_l A^k)}^p &\geq \bar{m}_{F(\tilde{I}_l A)}^p, l = 1, \dots, L; \\ \sum_{l=1}^L \beta_l \bar{\mu}_{F(\tilde{I}_l A^k)}^p &\geq \bar{M}_{F(\tilde{I}_l A)}^p. \end{aligned} \quad (15)$$

Здесь $0 < \beta_l \leq 1, l = 1, \dots, L$, — весовые коэффициенты, удовлетворяющие условиям нормировки

$$\sum_{l=1}^L \beta_l = 1, 0.$$

Среди оставшихся перспективных альтернатив наиболее эффективная может быть выбрана на основе оптимизации некоторой свертки локальных критериев.

2. Ранжирование множества перспективных альтернатив $A^k \in G \cap \Pi(A)$ может быть осуществлено для какого-то конкретного $p = 1, 2, 3$ по убыванию одного из следующих показателей:

$$\Psi_1^p(A^k) = \sum_{l=1}^L \gamma_l \mu_{F(\tilde{I}_l A^k)}^p; \quad (16)$$

$$\Psi_1^p(A^k) = \sum_{l=1}^L \gamma_l \mu_{F(\tilde{I}_l A^k)}^p, \Psi_2^p(A^k) = L \sqrt[L]{\prod_{l=1}^L (\mu_{F(\tilde{I}_l A^k)}^p)^{w_l}},$$

$$\Psi_3^p(A^k) = \beta_1 \Psi_1^p(A^k) + \beta_2 \Psi_2^p(A^k). \quad (17)$$

3. Использование лингвистической информации для сужения множества Парето

В ряде случаев экспертам легче дать чисто качественную оценку значениям каждого из локальных критериев, оценив значение этого показателя некоторым термом соответствующей лингвистической переменной. Без ограничения общности и для простоты изложения предположим, что наиболее эффективными являются наибольшие значения всех локальных критериев. В случае максимизации частного критерия эффективности $F_i(X)$ фактическое значение его может быть оценено экспертом од-

ним из следующих пяти термов: *NN*-очень низкое, *N*-низкое, *Z*-среднее, *P*-высокое, *PP*-очень высокое. При выборе наиболее эффективной из множества рассматриваемых альтернатив ЛПР может быть поставлено одно из дополнительных условий:

а) значение ни одного из локальных критериев в решении не должно принадлежать областям *NN* и *N* (или *NN*, *N* и *Z*);

б) значение некоторой группы \tilde{J}_1 локальных критериев в решении должно принадлежать области *PP* (*P* или *PP*).

В данном случае представляется целесообразным использовать функции принадлежности вида, представленного на рис. 3—5, с числом разрывов, равным числу термов соответствующей лингвистической переменной, определив на основании мнений экспертов границы соответствующих интервалов. На вид этих функций в пределах каждого из выделенных термов не накладывается никаких ограничений и он может быть выбран в соответствии с решениями экспертов или ЛПР.

Пусть $f_i^l(A), l = 0, 1, \dots, L$, — граничные значения соответствующих интервалов. В случае пяти термов граничные значения этих интервалов равны $NN \in [f_i^0(A), f_i^1(A)], N \in [f_i^1(A), f_i^2(A)], Z \in [f_i^2(A), f_i^3(A)], P \in [f_i^3(A), f_i^4(A)], PP \in [f_i^4(A), f_i^5(A)]$. Заметим, что даже нормированные значения границ диапазонов $f_i^l(A)$ для различных локальных критериев могут существенно отличаться друг от друга; $0 < \mu_A(f_i^l(A)) \leq 1, 0, l = 0, 1, \dots, L$, — значения функции принадлежности на границах соответствующих лингвистических термов.

Из рассмотрения исключаются Парето-оптимальные решения, для которых:

— в случае а) не выполняются ограничения вида

$$\mu_A(f_i^l(A^k)) < \mu_A(A_i^l(A)) \quad \forall i \in \tilde{I}; k \in \tilde{K}; l = 2 \text{ или } l = 3;$$

— в случае б) не выполняются ограничения вида

$$\min_{i \in \tilde{J}} \mu_A(f_i(A^k)) \geq \mu_A(A_i^l(A)), l = 3 \text{ или } l = 4; k \in \tilde{K}.$$

Иллюстративный пример

Значения частных критериев эффективности различных альтернатив приведены в таблице. В последних двух столбцах таблицы определены и минимальные $f_i^-(A) = \min_{1 \leq k \leq 12} f_i(A^k)$ и максимальные $f_i^+(A) = \max_{1 \leq k \leq 12} f_i(A^k)$ значения каждого из критериев.

Все критерии являются максимизирующими.

Критерии	Значения частных критериев эффективности различных альтернатив												Предельные значения критериев	
	A^1	A^2	A^3	A^4	A^5	A^6	A^7	A^8	A^9	A^{10}	A^{11}	A^{12}	min	max
$f_1(A)$	10	6	12	8	4	6	15	9	7	14	10	10	6	15
$f_2(A)$	12	4	6	15	7	5	3	10	8	7	6	5	3	15
$f_3(A)$	8	9	7	3	13	9	10	5	10	6	7	8	3	15

Множество Парето включает следующие семь альтернатив:

$$\Pi(A) = \{A^1, A^2, A^3, A^4, A^5, A^6, A^7\}.$$

Диапазоны изменения каждого частного критерия эффективности разбиты на пять интервалов, т. е. определены пять лингвистических термов каждого из этих показателей:

$$D[f_1(A)] = \{NN = (6,8), N = (8,10), Z = (10,11), P = (11,13), PP = [13,15]\};$$

$$D[f_2(A)] = \{NN = (3,5), N = (5,7), Z = (7,9), P = (9,12), PP = [12,15]\};$$

$$D[f_3(A)] = \{NN = (3; 4,5), N = (4,5; 6), Z = (6,8), P = (8, 11), PP = [11, 13]\}.$$

Функции принадлежности соответствующих нечетких множеств имеют вид:

$$\mu_A(f_1(A)) = \begin{cases} 0, & \text{if } 6,0 \leq f_1(A^k) < 8,0, \\ 0,2, & \text{if } 8,0 \leq f_1(A^k) < 10,0, \\ 0,5, & \text{if } 10,0 \leq f_1(A^k) < 11,0, \\ 0,75, & \text{if } 11,0 \leq f_1(A^k) < 13,0, \\ 1,0, & \text{if } 11,0 \leq f_1(A^k) \leq 13,0; \end{cases}$$

$$\mu_A(f_2(A)) = \begin{cases} 0,5x - 1,5, & \text{if } 3,0 \leq f_2(A^k) < 5,0, \\ 0,4, & \text{if } 5,0 \leq f_2(A^k) < 7,0, \\ 0,1x - 0,3, & \text{if } 7,0 \leq f_2(A^k) < 9,0, \\ 0,9, & \text{if } 9,0 \leq f_2(A^k) < 12,0, \\ 1,0, & \text{if } 12,0 \leq f_2(A^k) \leq 15,0; \end{cases}$$

$$\mu_A(f_3(A)) = \begin{cases} 0,1, & \text{if } 3,0 \leq f_3(A^k) < 4,5, \\ 0,35, & \text{if } 4,5 \leq f_3(A^k) < 6,0, \\ 0,1x - 0,25, & \text{if } 6,0 \leq f_3(A^k) < 8,0, \\ 0,8, & \text{if } 8,0 \leq f_3(A^k) < 11,0, \\ 0,1x - 0,3, & \text{if } 11,0 \leq f_3(A^k) \leq 13,0. \end{cases}$$

Если эффективное подмножество альтернатив не должно включать альтернативы, частные критерии эффективности не должны принадлежать лингвистическим термам NN и N , то из множества Парето должны быть исключены следующие альтер-

нативы: A^3 , так как $f_2(A^2) \in N$; A^4 , так как $f_1(A^4) \in N$, $f_3(A^4) \in NN$; A^5 , так как $f_1(A^5) \in NN$; A^7 , так как $f_2(A^7) \in NN$.

Эффективное множество Парето имеет вид $\{\Pi(A) \cap \Omega(A)\} = \{A^1, A^{10}\}$, т. е. включает только две альтернативы.

Заключение

Предложенные в работе алгоритмы выбора наиболее эффективных решений среди множества альтернатив, принадлежащих множеству Парето, основаны на учете предпочтений лица, принимающего решение (ЛПР), и методов нечеткой логики. Рассматриваемые в работе методы используют информацию об "идеальной точке", лингвистические переменные и аппарат нечеткой логики, позволяющий формализовать предпочтения экспертов и ЛПР.

Сформулированы компромиссные критерии эффективности принимаемого решения в форме различного вида сверток функций принадлежности частных критериев. Полученные результаты могут быть использованы в решении прикладных задач многокритериального выбора для различных приложений в технике, экономике, логистике и организации обслуживания.

Список литературы

1. **Подиновский В. В., Ногин В. Д.** Парето-оптимальные решения многокритериальных задач. М.: Наука, 1989.
2. **Ногин В. Д., Басков О. В.** Сужение множества Парето на основе учета произвольного конечного набора числовой информации об отношении предпочтения // Доклады Академии наук. 2011. Т. 38, № 4. С. 1—4.
3. **Ногин В. Д.** Проблема сужения множества Парето: подходы к решению // Искусственный интеллект и принятие решений. 2008. № 1. С. 98—111.
4. **Ногин В. Д.** Сужение множества Парето на основе нечеткой информации // Information Technologies & Knowledge. 2012. V. 6, N 2. P. 157—168.
5. **Климова О. Н.** Сужение множества Парето на основе взаимно зависимой информации с использованием нелинейных функций. URL: <http://cyberleninka.ru/article/n/suzhenie-mnozhestva-pareto-na-osnove-vzaimno-zavisimoy-informatsii-s-ispolzovaniem-nelineynyh-funktsiy>
6. **Зак Ю. А.** Принятие многокритериальных решений. М.: Экономика, 2011. 236 с.
7. **Зак Ю. А.** Принятие решений в условиях размытых и нечетких данных. М.: Либроком, 2012. 348 с.
8. **Figueira J., Greco S., Ehgott M.** Multiple criteria decision analysis: State of the art surveys. Springer Science + Business Media, 2005.
9. **Zimmermann H.-J.** Unscharfe Entscheidungen und Multi-Criteria Analyse // Proc. Oper. Res. 1979. V. 6. S. 99—109.

Н. А. Семькина, канд. физ.-мат. наук, доц., e-mail: semykina.tversu@yandex.ru,
 А. А. Смирнов, аспирант,
 ФГБОУ ВПО "Тверской государственный университет"

Применение генетического алгоритма для решения одной задачи оптимального управления

Рассматривается многокритериальная задача оптимального управления, которая моделирует деятельность предприятия с учетом ограниченности ресурсов. Представлены результаты сравнения применения классического подхода — численных методов решения задачи оптимального управления с фазовыми ограничениями и генетического алгоритма.

Ключевые слова: математическое моделирование, численные методы оптимального управления, многокритериальная задача

N. A. Semykina, A. A. Smirnov

The Use of Genetic Algorithm for Solution of the Problem of Optimal Control

In the article the multicriterion optimal control problem is considered. The model simulates the activity of an enterprise with limited resources. Comparison of the solution by numerical methods of optimal control with the genetic algorithm is represented.

Keywords: numerical methods of optimal control, mathematical modeling, a multicriteria problem

Введение

Одним из преимуществ генетического алгоритма является его применимость для многомерных и многоэкстремальных задач. Поэтому в настоящее время он активно используется при решении многокритериальных оптимизационных задач в экономике.

Постановка задачи

Рассмотрим многокритериальную задачу управления, моделирующую оптимальное распределение денежных ресурсов предприятия. Модель строится в следующих предположениях [1–2]:

- предприятие развивается за счет внутренней инвестиционной стратегии, которая определяется долей прибыли, отчисляемой на реинвестирование, и единовременного кредитного ресурса;
- основные производственные фонды и численность наемных работников являются факторами производства;
- численность работников описывается логистическим уравнением с учетом привлечения кадров;
- часть дохода сберегается на счете в банке;
- списание задолженности осуществляется за счет перечисления с банковского счета предприятия.

Математическая модель может быть представлена в виде следующей системы уравнений:

$$\dot{K}(t) = -\mu K(t) + b(t)Y(K, L), K(0) = K_0; \quad (1)$$

$$\dot{L}(t) = n(t)L(t) \left(1 - \frac{L(t)}{L_{\max}} \right) + \gamma(t)L(t), L(0) = L_0; \quad (2)$$

$$\begin{aligned} \dot{d}(t) &= (r_1 - h(t))d(t) + (1 - v) \times \\ &\times [(1 - a - b(t))Y(K, L) - \omega L(t)], d(0) = d_0; \quad (3) \end{aligned}$$

$$\dot{s}(t) = r_2 s(t) - h(t)d(t), s(0) = \Phi_0, \quad (4)$$

где $K(t)$ — стоимость основных производственных фондов; $L(t)$ — численность наемных работников; $Y(K, L)$ — выпуск продукции в стоимостном выражении; μ — коэффициент амортизации капитала; $b(t)$, $0 < b(t) < b_{\max}$ — доля прибыли, отчисляемой на реинвестирование; $d(t)$ — прибыль, сберегаемая в банке; a — коэффициент материальных затрат; $s(t)$ — задолженность перед банком; r_1 — ставка банковского счета; $h(t)$, $0 < h(t) < h_{\max}$ — функция, характеризующая погашение кредиторской задолженности; r_2 — начисляемый процент на сумму долга; Φ_0 — единовременный кредит; v — ставка налогообложения; ω — ставка заработной платы; $\gamma(t)$ — функция миграции кадров из-за низкого уровня доходов в более развитые экономические центры.

Предприятие регулирует уровень реинвестирования для извлечения максимальной прибыли за время T :

$$\int_0^T e^{-\delta t} (1 - \nu) [(1 - a - b(t))Y(K, L) - \omega L(t)] dt \rightarrow \max \quad (5)$$

и для уменьшения кредиторской задолженности на конец рассматриваемого периода

$$s(T) \rightarrow \min. \quad (6)$$

Здесь $\delta > 0$ — коэффициент дисконтирования.

Сформулированная задача является задачей оптимального управления с двумя критериями качества (5), (6). Управлением будем считать функции, характеризующие долю прибыли, отчисляемой на реинвестирование $b(t)$, и погашение кредиторской задолженности $h(t)$.

Переменные $K(t)$, $L(t)$, $d(t)$, $s(t)$ являются фазовыми с соответствующими ограничениями:

$$K(t) \geq 0, 0 \leq L(t) \leq L_{\max}, d(t) \geq 0, 0 \leq s(t) \leq K(t).$$

Используя исследования, которые были сделаны ранее [2], будем считать, что фазовые ограничения для функции $K(t)$, $L(t)$ и $d(t)$ не активны. Тогда фазовые ограничения будут иметь вид

$$0 \leq s(t) \leq K(t). \quad (7)$$

Для дальнейшего решения задачи сделаем линейную свертку критериев (5), (6). Каждый критерий имеет свой весовой неотрицательный коэффициент, отражающий степень важности для работы предприятия ($\Lambda_1 + \Lambda_2 = 1$):

$$\Lambda_1 \int_0^T e^{-\delta t} (1 - \nu) [(1 - a - b(t))Y(K, L) - \omega L(t)] dt - \Lambda_2 s(T) \rightarrow \max. \quad (8)$$

Необходимые условия оптимальности

Выпишем необходимые условия оптимальности. Функция Понтрягина поставленной задачи (1)–(4), (7), (8) будет иметь вид (рассматриваем регулярный случай)

$$H(t) = \Lambda_1 e^{-\delta t} (1 - \nu) [(1 - a - b(t))Y(K, L) - \omega L(t)] + p_K(t) (-\mu K(t) + b(t)Y(K, L)) + p_L(t) \left(n(t)L(t) \times \left(1 - \frac{L(t)}{L_{\max}} \right) + \gamma(t)L(t) \right) + p_s(t) (r_2 s(t) - h(t)d(t)) + p_d(t) ((r_1 - h(t))d(t) + (1 - \nu) \times [(1 - a - b(t))Y(K, L) - \omega L(t)]).$$

Согласно теореме о необходимых условиях оптимальности для задач с фазовыми ограничения-

ми [3] для оптимального управляемого процесса $(\bar{K}, \bar{L}, \bar{d}, \bar{s}, \bar{b}, \bar{h})$ существуют ненулевой вектор $l^0 = (l_K^0, l_d^0, l_s^0, l_L^0)^T$ и неотрицательные регулярные меры η_i ($i = 1, 2$), сосредоточенные на множествах $D_1 = \{t \in [0, T] : s(t) = 0\}$ и $D_2 = \{t \in [0, T] : s(t) - K(t) = 0\}$, такие, что:

1) сопряженные вектор-функции являются решением системы интегральных уравнений

$$p_K(t) = \int_t^T [Y'_K(\tau) \{(\Lambda_1 e^{-\delta \tau} + p_d(\tau))(1 - \nu) \times (1 - a - b(\tau)) + p_K(\tau)b(\tau)\} - \mu p_K(\tau)] d\tau + \int_t^T d\eta_2;$$

$$p_d(t) = \int_t^T (p_d(\tau)(r_1 - h(\tau)) - p_s(\tau)h(\tau)) d\tau;$$

$$p_s(t) = -\Lambda_2 + \int_t^T p_s(\tau)r_2 d\tau + \int_t^T d\eta_1 - \int_t^T d\eta_2;$$

$$p_L(t) = \int_t^T ((\Lambda_1 e^{-\delta \tau} + p_d(\tau))(1 - \nu) \times ((1 - a - b(\tau))Y'_L(\tau) - \omega) + p_K(\tau)b(\tau)Y'_L(\tau) + p_L(\tau) \left[n(\tau) \left(1 - \frac{2L(\tau)}{L_{\max}} \right) + \gamma(\tau) \right]) d\tau, \quad (9)$$

где $\tau \in [t, T]$ — независимая переменная, с условиями на правом конце траектории

$$p_K(T) = -l_K^0 - \eta_2[T], p_d(T) = -l_d^0,$$

$$p_s(T) = -l_s^0 - \eta_1[T] + \eta_2[T], p_L(T) = -l_L^0;$$

2) почти при всех $t \in [0, T]$ функция Понтрягина достигает своего максимального значения при оптимальном управлении $(\bar{b}(t), \bar{h}(t))$:

$$H(t, \bar{K}, \bar{d}, \bar{s}, \bar{L}, \bar{b}, \bar{h}, \bar{p}) = \max_{0 \leq b \leq b_{\max}, 0 \leq h \leq h_{\max}} H(t, K, d, s, L, b, h, p).$$

Введем функции переключения:

$$\varphi_1(t) = Y(K, L) [p_K(t) - e^{-\delta t} (1 - \nu) - p_d(t) (1 - \nu)],$$

$$\varphi_2(t) = d(t) (p_d(t) + p_s(t)).$$

Из второго условия теоремы получаем следующую конечномерную задачу, из которой определяем оптимальное управление:

$$\varphi_1(t)b + \varphi_2(t)h \rightarrow \sup_{0 \leq b \leq b_{\max}, 0 \leq h \leq h_{\max}}. \quad (10)$$

Предположим, что меры η_i ($i = 1, 2$) имеют плотность $\rho_i(t) > 0$, $i = 1, 2$. При этом $d\eta_i = \rho_i dt$. Дифференцируя равенства (9), получим

$$\begin{aligned} p_K(t) &= -(\Lambda_1 e^{-\delta t} + p_d(t))(1 - \nu)(1 - a - b(t)) Y'_K(t) + \\ &+ p_K(t)(\mu - b(t) Y'_K(t)) + \rho_2; \\ p_d(t) &= -p_d(t)(r_1 - h(t)) + p_s(t)h(t); \\ p_s(t) &= -p_s(t)r_2 + \rho_1 - \rho_2; \\ p_L(t) &= -(\Lambda_1 e^{-\delta t} + p_d(t))(1 - \nu)((1 - a - b(t)) \times \\ &\times Y'_L(t) - \omega) - p_K(t)b(t) Y'_L(t) - \\ &- p_L(t) \left[n(t) \left(1 - \frac{2L(\tau)}{L_{\max}} \right) + \gamma(t) \right]. \end{aligned}$$

Неотрицательные меры η_i ($i = 1, 2$), определенные на множествах D_i , удовлетворяют условиям $sd\eta_1 = 0$, $(s - K)d\eta_2 = 0$, где $d\eta_i \geq 0$, $i = 1, 2$.

Для построения численного решения удобнее свести задачу оптимального управления с фазовыми ограничениями к последовательности задач без ограничений. В этом случае учтем условия (7) в функционале, используя внешние штрафные функции

$$\begin{aligned} J_k(t, K, L, d, s, b, h) &= \\ &= -\Lambda_1 \int_0^T e^{-\delta t} (1 - \nu) [(1 - a - b(t)) Y(K, L) - \omega L(t)] dt + \\ &+ \Lambda_2 s(T) + A_k \int_0^T [\max\{-s(t); 0\}]^2 dt + \\ &+ B_k \int_0^T [\max\{s(t) - K(t); 0\}]^2 dt \rightarrow \inf, \end{aligned} \quad (11)$$

где A_k, B_k ($k = 1, 2, \dots$) — последовательности положительных чисел.

Используя принцип максимума Понтрягина, запишем систему для сопряженных вектор-функций $p_K(t), p_L(t), p_d(t), p_s(t)$ с условиями трансверсальности на правом конце траектории:

$$\begin{aligned} p'_K(t) &= -(\Lambda_1 e^{-\delta t} + p_d(t))(1 - \nu)(1 - a - b(t)) Y'_K(t) - \\ &- p_K(t)(-\mu + b Y'_K(t)) - 2B_k \max\{s(t) - K(t), 0\}; \\ p'_d(t) &= p_s(t)h(t) - p_d(t)(r_1 - h(t)); \\ p'_s(t) &= -p_s(t)r_2 - 2A_k \max\{-s(t), 0\} + \\ &+ 2B_k \max\{s(t) - K(t), 0\}; \\ p'_L(t) &= -(\Lambda_1 e^{-\delta t} + p_d(t))(1 - \nu) \times \\ &\times ((1 - a - b(t)) Y'_L(t) - \omega) - b(t)p_K(t) Y'_L(t) - \\ &- p_L(t) \left[n \left(1 - \frac{2L(\tau)}{L_{\max}} \right) + \gamma(t) \right]; \\ p_K(T) &= 0, p_d(T) = 0, p_s(T) = -\Lambda_2, p_L(T) = 0. \end{aligned} \quad (12)$$

В результате исходная задача сводится к краевой задаче принципа максимума Понтрягина, которая включает в себя систему дифференциальных

уравнений (1)–(4), (12) и граничные условия (13), где оптимальное управление определяется из (10).

Решение нелинейной краевой задачи является неформализуемой проблемой: в каждом конкретном случае нелинейная краевая задача требует индивидуального подхода для поиска решения. Стандартный подход в разрешении проблемы заключается в построении дискретной аппроксимации для непрерывной задачи и дальнейшее ее решение численными методами.

Построение дискретной задачи

Для построения численного решения выпишем дискретную аппроксимацию задачи (1)–(4), (11). На отрезке $[0, T]$ введем равномерную сетку $\{t_i = \tau \cdot i, 0 \leq i \leq N\}$ с шагом $\tau = T/N$. Обозначим значения фазовых функций и функций управления в точках разбиения: $K(t_i) = K^i, L(t_i) = L^i, d(t_i) = d^i, s(t_i) = s^i, b(t_i) = b^i, h(t_i) = h^i$. Используя явную разностную схему и квадратурную формулу прямоугольников, получаем следующую дискретную задачу оптимального управления. Требуется минимизировать функционал

$$\begin{aligned} W_k &= -\Lambda_1 \tau \sum_{i=0}^{N-1} e^{-\delta t_i} (1 - \nu) \times \\ &\times [(1 - a - b^i) Y^i(K^i, L^i) - \omega L^i] + \Lambda_2 s^N + \\ &+ A_k \tau \sum_{i=0}^{N-1} [\max\{-s^i; 0\}]^2 + B_k \tau \sum_{i=0}^{N-1} [\max\{s^i - K^i; 0\}]^2 \end{aligned} \quad (14)$$

при ограничениях:

$$\begin{aligned} 0 &\leq i \leq N - 1; \\ K^{i+1} &= K^i - \tau(\mu K^i - b^i Y^i(K^i, L^i)), K^0 = K_0; \\ L^{i+1} &= L^i + \tau \left(n^i L^i \left(1 - \frac{L^i}{L_{\max}} \right) + \gamma^i L^i \right), L^0 = L_0; \\ d^{i+1} &= d^i + \tau((r_1 - h^i) d^i + (1 - \nu) \times \\ &\times [(1 - a - b^i) Y^i(K^i, L^i) - \omega L^i]), d^0 = d_0; \\ s^{i+1} &= s^i + \tau(r_2 s^i - h^i d^i), s^0 = \Phi_0, \\ 0 &< b^i < b_{\max}, 0 < h^i < h_{\max}. \end{aligned} \quad (15)$$

Для полученной дискретной задачи (14)–(15) применим правило множителей Лагранжа. Запишем функцию Лагранжа

$$\begin{aligned} \Delta(K, L, d, s, b, h, p) &= -\Lambda_1 \tau \sum_{i=0}^{N-1} e^{-\delta t_i} (1 - \nu) \times \\ &\times [(1 - a - b^i) Y^i(K^i, L^i) - \omega L^i] + \Lambda_2 s^N + \\ &+ A_k \tau \sum_{i=0}^{N-1} [\max\{-s^i; 0\}]^2 + B_k \tau \sum_{i=0}^{N-1} [\max\{s^i - K^i; 0\}]^2 + \\ &+ \sum_{i=0}^{N-1} p_K^{i+1} (K^{i+1} - K^i + \tau(\mu K^i - b^i Y^i(K^i, L^i))) + \\ &+ \sum_{i=0}^{N-1} p_L^{i+1} \left(L^{i+1} - L^i - \tau \left(n^i L^i \left(1 - \frac{L^i}{L_{\max}} \right) + \gamma^i L^i \right) \right) + \end{aligned}$$

$$\begin{aligned}
 & + \sum_{i=0}^{N-1} p_s^{i+1} (s^{i+1} - s^i - \tau(r_2 s^i - h^i d^i)) + \\
 & + \sum_{i=0}^{N-1} p_d^{i+1} (d^{i+1} - d^i - \tau(r_1 - h^i) d^i + \\
 & + (1 - \nu)[(1 - a - b^i) Y^i(K^i, L^i) - \omega L^i]).
 \end{aligned}$$

Пусть $[\bar{K}, \bar{L}, \bar{d}, \bar{s}, \bar{b}, \bar{h}]$ — локально-оптимальный процесс дискретной задачи, тогда существуют множители Лагранжа $p_K^i, p_L^i, p_d^i, p_s^i, i = \overline{1, N}$, не равные нулю, одновременно такие, что выполнены условия стационарности:

1) по фазовым переменным $l = \overline{1, N-1}$

$$\begin{aligned}
 \frac{\partial \Delta}{\partial K^l} &= -\tau(1 - \nu)(1 - a - \bar{b}^l) \times \\
 & \times \frac{\partial Y^l(\bar{K}^l, \bar{L}^l)}{\partial K^l} (\Lambda_1 e^{-\delta t_l} + p_d^{l+1}) - 2B_k \tau \max\{\bar{s}^l - \bar{K}^l, 0\} + \\
 & + p_K^l - p_K^{l+1} + p_K^{l+1} \tau \left(\mu - \bar{b}^l \frac{\partial Y^l(\bar{K}^l, \bar{L}^l)}{\partial K^l} \right) = 0;
 \end{aligned}$$

$$\begin{aligned}
 \frac{\partial \Delta}{\partial L^l} &= -\tau(1 - \nu) \left[(1 - a - \bar{b}^l) \frac{\partial Y^l(\bar{K}^l, \bar{L}^l)}{\partial L^l} - \omega \right] \times \\
 & \times (\Lambda_1 e^{-\delta t_l} + p_d^{l+1}) - \tau p_K^{l+1} \bar{b}^l \frac{\partial Y^l(\bar{K}^l, \bar{L}^l)}{\partial L^l} + p_L^l - \\
 & - p_L^{l+1} - p_L^{l+1} \tau \left(n^l - \frac{2\bar{L}^l}{L_{\max}} + \gamma^{l+1} \right) = 0;
 \end{aligned}$$

$$\frac{\partial \Delta}{\partial d^l} = p_d^l - p_d^{l+1} - p_d^{l+1} \tau(r_1 + \bar{h}^l) + \tau p_s^{l+1} \bar{h}^l = 0;$$

$$\begin{aligned}
 \frac{\partial \Delta}{\partial s^l} &= -2A_k \tau \max\{-\bar{s}^l, 0\} + 2B_k \tau \max\{\bar{s}^l - \bar{K}^l, 0\} + \\
 & + p_s^l - p_s^{l+1} (1 + \tau r_2) = 0;
 \end{aligned}$$

2) по переменным управления

$$\begin{aligned}
 \frac{\partial \Delta}{\partial b^l} &= \tau Y^l(\bar{K}^l, \bar{L}^l) (\Lambda_1 e^{-\delta t_l} (1 - \nu) - p_K^{l+1} + \\
 & + p_d^{l+1} (1 - \nu)) = 0, l = \overline{1, N-1};
 \end{aligned}$$

$$\frac{\partial \Delta}{\partial h^l} = \tau \bar{d}^l (p_d^{l+1} + p_s^{l+1}) = 0, l = \overline{1, N-1}.$$

Условия стационарности по фазовым переменным используем для построения рекуррентных соотношений для вычисления сопряженных вектор-функций. Далее можно применить любой стандартный алгоритм из семейства градиентных методов.

Для решения поставленной задачи с помощью генетического алгоритма будем использовать вещественное представление решений [4—6]. В силу многокритериальности рассматриваемой задачи единственное оптимальное решение определить невозможно. Результатом работы генетического алгоритма является множество эффективных (недоминируемых) решений задачи I_p , обычно называемое *популяцией*. Элементы популяции $u^k \in I_p$ принято называть *особями*.

Каждая итерация генетического алгоритма представляет собой процесс получения новой популяции I_p из популяции I_{p-1} путем применения специальных операторов *селекции*, *скрещивания* и *мутации* к особям из I_{p-1} .

В генетическом алгоритме центральным понятием является понятие функции приспособленности особи $\mu_k = \mu(u^k)$. Значения этой функции используются в операторе селекции особей. В качестве функции приспособленности для рассматриваемой задачи будем использовать функцию, основанную на понятии доминирования по Парето. Для каждой особи u^k текущей популяции I_p необходимо построить фазовые траектории $K^k(t), L^k(t), d^k(t), s^k(t)$, а затем вычислить значения критериев качества (5), (6). Пусть ψ_k — это число особей, которые доминирует особь u^k (имеется в виду доминирование в пространстве значений критериев качества задачи). Значение функции приспособленности положим равным $\mu_k(u^k) = 1/\psi_k$. Будем учитывать, что особь доминирует себя, поэтому минимально возможное значение ψ_k равно 1 (это значение соответствует недоминируемым особям). Таким образом, значения функции приспособленности μ_k для всех недоминируемых особей равно 1 и меньше 1 для всех остальных особей.

Применение функции приспособленности такого вида позволяет приближенно построить множество Парето-задач, тогда как, например, при использовании агрегированного скалярного критерия различным наборам весовых коэффициентов соответствует только одно слабоэффективное решение. Также это позволяет избежать типичных проблем, которые обычно возникают при использовании линейного агрегированного скалярного критерия (например, невозможности получения решений из невыпуклой части Парето-фронта и сложности подбора весовых коэффициентов).

Процедура поиска решений имеет следующий вид.

Шаг 1. Создание начальной популяции I_0 .

1.1. Задаем номер популяции $p = 0$, максимальное число популяций p_{\max} . Также задаем число отрезков N в разбиении временного промежутка $[0, T]$ и размер популяции m .

1.2. Случайным образом, используя равномерное распределение на множестве допустимых значений управления, формируем начальную популя-

цию I_0 допустимых векторов управления. В результате получаем m начальных допустимых векторов $u^k = (b_1^k, \dots, b_N^k, h_1^k, \dots, h_N^k)$, где $0 < b_i^k < b_{\max}$, $0 < h_i^k < h_{\max}$, $k = \overline{1, m}$.

1.3. Вычисляем значение функции приспособленности для каждой особи $u^k \in I_0$: $\mu_k = \mu(u^k)$, $k = \overline{1, m}$. Также вычисляем приспособленность популяции

$$I_0 \text{ в целом: } \mu = \sum_{k=1}^m \mu_k.$$

Шаг 2. Формирование новой популяции I_p .

2.1. Формирование пустой новой популяции I_p . На этом шаге также применяем эффективную модификацию генетического алгоритма — стратегию элитизма. Ее суть в том, что небольшое число самых приспособленных особей u^k из предыдущей популяции I_{p-1} переходит в новую популяцию I_p без изменений.

2.2. *Селекция* — это операция, которая осуществляет отбор особей u^k в соответствии со значениями функции приспособленности μ_k для последующего их скрещивания.

При решении поставленной задачи в качестве оператора селекции будем использовать оператор рулетки, который осуществляет отбор особей с помощью двух "запусков" рулетки, вероятность участия особи в селекции пропорциональна относительной приспособленности особи μ_k/μ . Результатом шага являются две особи u^* и u^{**} , выбранные в качестве родительской пары с помощью одного из операторов селекции.

2.3. *Скрещивание* — это операция, при которой из нескольких, обычно двух особей, называемых родителями, порождается одна или несколько новых, называемых потомками, путем обмена частями родительских хромосом. В простейшем случае скрещивание в генетическом алгоритме реализуется так же, как и в биологии: хромосомы разрезаются в случайной точке (точках) и обмениваются частями между собой. Для данной задачи в качестве оператора селекции будем использовать оператор дискретного кроссовера, который создает одного потомка $u^q = (b_1^q, \dots, b_N^q, h_1^q, \dots, h_N^q)$, где b_i^q и h_i^q выбираем из двух значений u_i^* и u_i^{**} случайно с вероятностью 0,5 для всех $i = \overline{1, N}$.

Результатом шага является потомок u^q , полученный путем применения одного из операторов скрещивания к родительским особям u^* и u^{**} .

2.4. *Мутация* — это преобразование хромосомы, случайно изменяющее обычно один (реже несколько) из ее генов. Оператор мутации предназначен для того, чтобы поддерживать разнообразие особей в популяции. Рассматривается потомок u^q , полученный в результате операции скрещивания. Среди генов $(b_1^q, \dots, b_N^q, h_1^q, \dots, h_N^q)$ случайно (с ве-

роятностью $\frac{1}{2N}$) выбирается один с номером l , подлежащий замене.

В рассматриваемой задаче будем использовать оператор неоднородной мутации. Новое значение u_l^q определяется следующим образом:

$$u_l^q = \begin{cases} u_l^q + \varepsilon(p, x_i - u_l^q), & z = 0; \\ u_l^q + \varepsilon(p, u_l^q - y_i), & z = 1, \end{cases}$$

где z — случайное число, выбранное с вероятностью 0,5 из множества $\{0, 1\}$; а $\varepsilon(p, u)$ — функция, определяющая значение на промежутке $[0, u]$:

$\varepsilon(p, u) = u(1 - \rho^{(1-p/p_{\max})^g})$, ρ — случайное число на отрезке $[0, 1]$, p — номер текущей популяции, g — параметр оператора.

Результатом шага является потомок — мутант u^q , полученный путем применения оператора мутации к особи u^q .

2.5. Формирование новой популяции I_p .

Если особь u^q является допустимой, то она добавляется в новую популяцию I_p .

Если размер популяции I_p остается меньше чем m , то формирование популяции I_p продолжается (переход к шагу 2.2).

Если размер популяции I_p становится равным m , то вычисляются значения функции приспособленности μ_k для каждой особи $u^k \in I_p$, а также приспособленность μ популяции I_p в целом. Затем выполняется переход на шаг 3.

Шаг 3. Проверка условия окончания работы алгоритма.

Если $p < p_{\max}$, то значение p полагается равным $p + 1$ и выполняется переход к шагу 2.

Если же сформировано требуемое число популяций, т. е. $p = p_{\max}$, то в качестве множества решений задачи выбираются особи, значение функции приспособленности которых равно наибольшему значению функции приспособленности из текущей популяции.

Анализ численных экспериментов

Построенная математическая модель была применена для исследования деятельности ОАО "Старицкий сыр" [2]. На основе статистических данных с помощью метода наименьших квадратов была получена двухфакторная производственная функция $Y(t) = 0,0000053K^{0,292}(t)L^{4,34}(t)$ и определены остальные параметры модели:

- капитал фирмы в начальный момент времени $K_0 = 46\ 824$ тыс. руб.;
- число наемных работников на начало периода $L_0 = 159$ чел.;
- предельная численность наемных работников $L_{\max} = 210$ чел.;
- на текущем счету в банке $d_0 = 2104$ тыс. руб.;
- задолженность фирмы $s_0 = 29\ 639$ тыс. руб.;

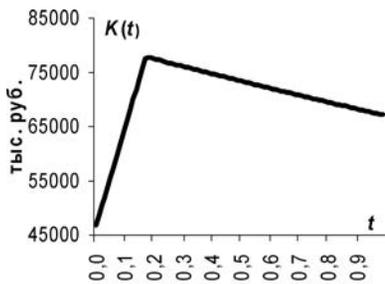


Рис. 1. Динамика стоимости $K(t)$ производственных фондов

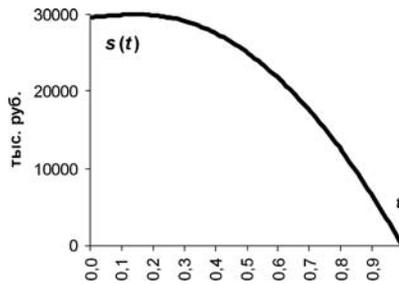


Рис. 3. Динамика величины задолженности $s(t)$ перед банком

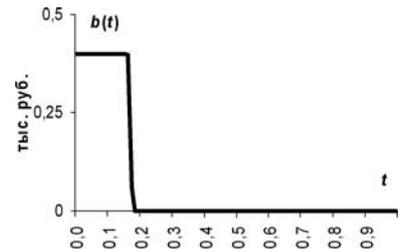


Рис. 5. Динамика доли прибыли $b(t)$, отчисляемой на реинвестирование

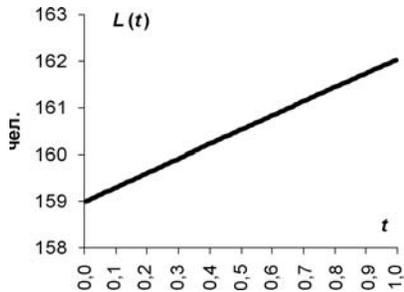


Рис. 2. Динамика численности $L(t)$ наемных работников

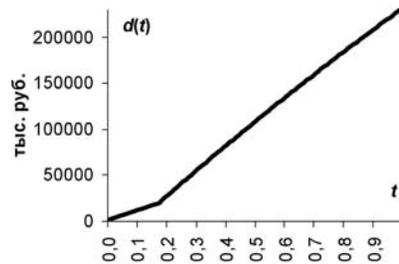


Рис. 4. Динамика величины прибыли $d(t)$, сберегаемой в банке

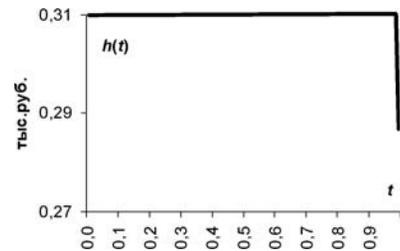


Рис. 6. Динамика погашения задолженности $h(t)$

- амортизационные отчисления $\mu = 0,18 \text{ год}^{-1}$;
- постоянные материальные затраты $a = 0,31 \text{ год}^{-1}$;
- процент на задолженность $r_2 = 0,17 \text{ год}^{-1}$;
- процент по вкладам $r_1 = 0,04 \text{ год}^{-1}$;
- ставка по налогу на прибыль $\nu = 0,2 \text{ год}^{-1}$;
- коэффициент прироста наемных работников $n = 0,08 \text{ год}^{-1}$;
- ставка заработной платы $\omega = 6 \text{ тыс. руб.}$

Для дискретной задачи (14)—(15) с использованием методов градиентного спуска и генетического алгоритма были построены численные решения. Решение, полученное с помощью генетического алгоритма, совпадает с традиционным методом оптимизации. Результаты численных экспериментов представлены в виде графиков функций на рис. 1—6.

Анализируя полученную производственную функцию $Y(t)$, заметим, что коэффициент эластичности производства относительно затрат труда намного превышает коэффициент эластичности по капиталу ($4,34 > 0,292$). Следовательно, фактор труда вносит наибольшую долю в прирост конечного продукта, т. е. в прирост прибыли. Поэтому для максимизации функционала (5) функция, характеризующая численность наемных работников $L(t)$, должна возрасти (рис. 2). Функция стоимости $K(t)$ основных производственных фондов достигает некоторого максимального значения на начальном интервале времени $[0, \tau]$, что объясняется поведением функции управления $b(t)$, а затем убывает в связи с отсутствием реинвестирования (рис. 1).

Фирма функционирует для извлечения максимальной прибыли. Реинвестирование является вложением в будущую прибыль, поэтому для рассматриваемого промежутка времени (1 год) предпри-

ятие определяет инвестиционную стратегию как минимальную долю прибыли, отчисляемой на реинвестирование. Функция управления $b(t)$ имеет ступенчатый вид: первые месяцы фирма отчисляет 0,4 прибыли на развитие предприятия, остальное время отчисления равны нулю (рис. 5). Функция управления, характеризующая погашение кредиторской задолженности $h(t)$, является постоянной (рис. 6). Это демонстрирует равномерное погашение кредита $s(t)$, которое завершается к концу рассматриваемого периода (рис. 3).

Данные результаты были получены при условии, что для работы предприятия имеют одинаковую степень важности и максимизация прибыли (5), и уменьшение кредиторской задолженности (6) на конец рассматриваемого периода. То есть весовые коэффициенты линейной свертки (8) равны ($\Lambda_1 = \Lambda_2$). Для исследования других вариантов предпочтения весовых коэффициентов требуется заново решать задачу оптимального управления (1)—(4), (7)—(8). В случае решения многокритериальной задачи с помощью генетического алгоритма мы получаем мно-

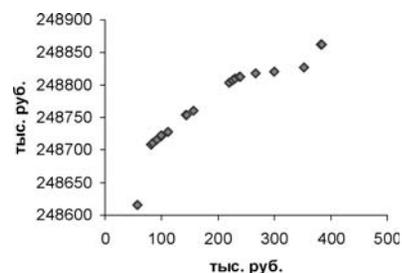


Рис. 7. Множество Парето многокритериальной задачи (1)—(4), (8)

жество решений, которое представлено в виде множества Парето. На рис. 7 по оси абсцисс отложено значение функционала (6), характеризующего кредиторскую задолженность на конец рассматриваемого периода, а по оси ординат — значение функционала (5), характеризующего прибыль за время T .

Само множество Парето иллюстрирует естественную для экономики ситуацию: увеличение прибыли неизбежно приводит к росту кредиторской задолженности.

Заключение

Данный численный эксперимент показал преимущество генетического алгоритма. Используя метод проекции градиента, можно построить одно оптимальное решение многокритериальной задачи с соответствующими весовыми коэффициентами вхождения функционалов (8). Генетический алгоритм

обеспечивает одновременную оптимизацию всех критериев качества, поэтому результатом является множество решений многокритериальной задачи.

Список литературы

1. Герасимов Б. И., Пучков Н. П., Протасов Д. Н. Дифференциальные динамические модели. Тамбов: ТГТУ, 2010.
2. Семькина Н. А., Федорова Е. А. Оптимизационная задача деятельности экономических агентов // Математические методы управления: сб. науч. тр. Тверь: Твер. гос. ун-т, 2011.
3. Андреева Е. А., Семькина Н. А. Оптимальное управление: учеб. пособие. Тверь: Твер. гос. ун-т, 2010.
4. Jain Abraham L., Goldberg R. Evolutionary multiobjective optimization: theoretical advances and applications. Verlag, New York: Springer, 2005.
5. Рутковская Д., Пилиньский М., Руковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы. М.: Горячая линия — Телеком, 2004.
6. Michalewicz Z., Janikow C. Z., Krawczyk J. B. A modified genetic algorithm for optimal control problems // Computers and mathematics with applications. 1992. Vol. 23, Is. 12.

УДК 519.856

Л. Е. Мистров, д-р техн. наук, доц., проф. кафедры,
Центральный филиал ФГБОУ ВПО "Российская академия правосудия", г. Воронеж,
e-mail: mistrov_le@mail.ru

Метод оптимизации облика иерархических систем информационной безопасности

Предлагается метод оптимизации иерархических систем информационной безопасности для обеспечения эффективного применения функциональных систем в условиях конкуренции. Метод базируется на методах теорий максимина, погрупповой оптимизации, динамического программирования и гомотопическом методе решения нелинейных оптимизационных задач с экстремальными ограничениями.

Ключевые слова: функциональная система, конкуренция, конфликт, информационная безопасность, средства, комплексы и системы информационной безопасности, модель, метод, критерий и показатели эффективности, оптимизация

L. E. Mistrov

Method of Optimization of Shape of the Hierarchical Systems of Information Security

The method of optimization of hierarchical systems of information security for ensuring effective use of functional systems in the conditions of the competition is offered. The method is based on methods of theories a maximine, pogruppovy optimization, dynamic programming and a homotopic method of the solution of nonlinear optimizing tasks with extreme restrictions.

Keywords: Functional system, competition, conflict, information security, means, complexes and systems of information security, model, method, criterion and efficiency indicators, optimization

Общие положения

В современных условиях выполнение значимых социально-экономических задач различного типа функциональными системами (ФС) осуществляется в условиях конкуренции, характеризующейся кон-

фликтом типа "соперничество" с одной или несколькими системами. Принятие ими решений осуществляется в условиях определенности, риска, неопределенности и нечеткости (в смысле Заде) на основе анализа составляющих информации, прин-

ципов действия информационных средств и способов их применения в прогнозируемой обстановке. Снижение качества принимаемых ФС решений может возникать за счет разрушения каналов и источников получения информации и искажения информации на основе методов и средств дезинформации, активными помехами, ложными целями, дезинформации, нейтрализации на время активной фазы конфликта наиболее ценных сотрудников систем, средств снижения заметности и т. п., которые могут применяться в целях обеспечения информационной безопасности (ИБ) как элементов, так и ФС в целом на различных этапах решения задач в исполнительных и защитных операциях (совокупности действий). Их целью является нарушение до требуемого уровня эффективности принятия решений в контурах управления конкурирующих ФС (для определенности, ФС $\{B\}$) и обеспечение в реальном масштабе времени обработки информации по всем циклам управления в контурах своей системы управления. Эффективность ИБ пропорциональна уровню технологического развития и масштабам использования различного назначения элементов получения информации в иерархических контурах принятия решений уровня средств, комплексов, технических, организационно-технических и организационных систем (ТС, ОТС, ОС) ФС. Произошедшие за последние годы в экономике России изменения выявили ряд дискуссионных и актуальных вопросов, носящих теоретический и прикладной характер по обеспечению эффективного развития различного типа и уровня ФС, что и обусловило поиск новых методов и средств обеспечения их устойчивого развития, в том числе на основе разработки и применения систем ИБ (СИБ). Это и определило целевую направленность и содержание предлагаемой статьи.

Под СИБ понимается совокупность объединенных единством цели подсистем ИБ уровня ОТС и ТС, а также комплексов ИБ (КИБ), предназначенных для обеспечения с заданной эффективностью функционирования ФС методами и средствами ИБ. Основу облика СИБ составляют допустимые варианты основных тактико-технических требований (ОТТТ) по составу и структуре, порядку функционирования и основным тактико-техническим характеристикам (ТТХ), определяющих ее эффективность применения. Допустимость вариантов облика СИБ определяется условиями, ограничивающими их реализацию при формировании задания на создание и применение. Такими ограничивающими условиями являются ресурсы, выделяемые на создание элементов СИБ, их размещение на исполнительных/обеспечивающих или специализированных элементах ФС (стоимостные, массогабаритные, энергетические, технологические и эксплуатационные) и пространственно-временные ограничения на ее применение (реализацию основного предназначения).

В общем случае конкуренция ФС $\{A^*\}$ и $\{B^*\}$ представляется в виде многошаговой итерационной

процедуры управления конфликтным взаимодействием их элементов, начиная с нижнего (технического) уровня и возвращаясь с каждого иерархически верхнего уровня на нижний на основе принципов математической декомпозиции и агрегатирования частных целей. Оценка эффективности СИБ осуществляется по интегральным показателям эффективности функционирования ФС применительно к типовым условиям применения — ядрам конфликта [2]. В соответствии с целью СИБ и условиями ее применения формируется система следующих показателей эффективности:

- целевые или интегральные показатели (математическое ожидание числа выполненных задач) на уровнях ядер конфликта типа двусторонних действий ФС (ее ОС) и сценариев (математическое ожидание числа выполненных задач на различных этапах действий) применительно к действиям ОТС;
- системные показатели (математическое ожидание числа исполнительных элементов, выполнивших поставленные задачи) на уровне ядер конфликта типа эпизодов применительно к действиям среднего и малого уровня ОТС и ТС;
- информационно-системные показатели (вероятности наведения исполнительного элемента на объект воздействия, нейтрализации объекта воздействия и другие) на уровне ядер конфликта типа ситуаций применительно к действиям комплексов;
- информационные показатели (вероятности обнаружения, сопровождения и другие) на уровне ядер конфликта типа дуэлей применительно к применению средств.

В свою очередь, на каждом ядре конфликта осуществляется дальнейшая декомпозиция каждой частной задачи исследования эффективности СИБ на совокупность задач допустимой сложности с введением на каждом уровне специальных принципов и аппроксимирующих аналитических зависимостей, позволяющих для принятых предположений, допущений и ограничений провести вычисление введенной системы показателей эффективности, проанализировать полученные результаты и установить присущие частному моделируемому процессу основные закономерности. При этом оценка эффективности СИБ проводится поуровнево на основе парирования межуровневой неопределенности исходных данных, вычисления иерархической системы показателей эффективности отдельных элементов и их последовательного преобразования на основе принципа гомотопической инвариантности с нижнего уровня в верхний.

Постановка задачи

Задача синтеза СИБ формулируется следующим образом: задана для разработки ФС $A(A_1, A_2, A_3)$, облик которой характеризуется множеством допустимых (в смысле ограничений) вариантов состава

и структуры (A_1) , порядком функционирования (стратегий поведения) (A_2) и ТТХ (A_3) . Ее потенциальным конкурентом является ФС $B(B_1, B_2, B_3)$, аналогично определяемая множеством вариантов (B_1, B_2, B_3) . Заданы цели и типовые условия применения ФС $\{A^*\}$ и $\{B^*\}$, определяемые их составом, порядком функционирования и ТТХ. Необходимо определить оптимальный вариант СИБ $A(A_1, A_2, A_3)$, обеспечивающий максимум значения целевой функции эффективности применения ФС:

$$U_A\{A^{opt}(A_1^{opt}, A_2^{opt}, A_3^{opt})\} = \max_{A(A_1, A_2, A_3)} \min_{B(B_1, B_2, B_3)} U_A\{A(A_1, A_2, A_3), B(B_1, B_2, B_3), A^*, B^*\} \quad (1)$$

при ограничениях

$$C_A^* \geq F_A^C(A); G_A^* \geq F_A^G(A); \mathcal{A}_A^* \geq F_A^{\mathcal{A}}(A); \\ C_B^* \geq F_B^C(B); G_B^* \geq F_B^G(B); \mathcal{A}_B^* \geq F_B^{\mathcal{A}}(B),$$

где $F_A^C(A)$, $F_A^G(A)$, $F_A^{\mathcal{A}}(A)$, $F_B^C(B)$, $F_B^G(B)$, $F_B^{\mathcal{A}}(B)$ — функции, связывающие множество параметров вариантов СИБ $A(A_1, A_2, A_3)$ и $B(B_1, B_2, B_3)$ со стоимостью, массой, габаритными размерами и энергопотреблением при разработке, производстве, применении и размещении средств ИБ на исполнительных и обеспечивающих элементах (ИЭ, ОЭ) и специальных носителях ФС, соответственно; C_A^* , G_A^* , \mathcal{A}_A^* , C_B^* , G_B^* , \mathcal{A}_B^* — основные ограничения при создании СИБ, определяемые ФС $\{A^*\}$ и $\{B^*\}$ по стоимости, массе, размерам и энергопотреблению, соответственно; $U_A\{A(A_1, A_2, A_3), B(B_1, B_2, B_3), A^*, B^*\}$ — целевая функция эффективности применения СИБ $\{A\}$ как функция параметров $A(A_1, A_2, A_3)$ и $B(B_1, B_2, B_3)$ и ФС $\{A^*\}$ и $\{B^*\}$, соответственно.

Задача (1) инвариантно может быть сформулирована для каждого элемента уровня иерархии, рассматриваемого при структуризации СИБ на основе выполнения следующих принципов: полнота рассматриваемых стратегий, незанижение возможностей конкурента и невысхождение своих возможностей.

Основным предназначением СИБ $\{A\}$ является обеспечение действий ФС $\{A^*\}$ в условиях конфликта с ФС $\{B^*\}$, в состав которой входит СИБ $\{B\}$. Конфликт между ФС $\{A^*\}$ и $\{B^*\}$ осуществляется в формах повседневной деятельности (ПД) и экстремальных действий (ЭД), чередуясь случайным образом в пространстве и во времени. Форма ПД — основная и связана с подготовкой ЭД, которые характеризуются массированным и кратковременным использованием сил и средств СИБ $\{A\}$ и $\{B\}$ при ведении защитных и исполнительных действий (ЗД, ИД) применительно к условиям активной фазы развития конфликта.

Пусть ФС $\{A^*\}$, имея начальный общий ресурс C_A^O , последовательно участвует в m , $m = \overline{1, M}$, ЭД (условно называемых $\mathcal{E}D_m$), решая в каждом из них определенное число задач ($j = \overline{1, T_m}$), и использует для этого часть ресурса $C_A^1(A_1)$. Этим действиям противодействует ФС $\{B^*\}$, используя часть своего начального ресурса $C_B^1(B_1)$ из ресурса C_B^O , снижая ресурс C_A^O . Другой частью ресурса C_B^O ФС $\{B^*\}$ последовательно осуществляет свои ЭД, обеспечиваемые СИБ $B(B_2)$, выполняя последовательно несколько ЭД B_j , решая в каждом из них определенное число задач ($j = \overline{1, J}$) (воздействуя на СИБ $A(A_2)$) ФС $\{B^*\}$ снижает часть ресурса C_A^1 . Вместе с тем СИБ $\{A\}$ противодействует СИБ $B(B_1, B_2)$, используя первоначальный ресурс $C_{\Phi A} = (C_{\Phi A}^1, C_{\Phi A}^2)$ и тем самым снижая потери элементов ФС $\{A^*\}$ за счет сохранения ресурсов C_A^1 и C_A^2 .

Цель задачи синтеза — обоснование оптимального варианта СИБ $\{A\}$ на основе реализации принципа "эффективность-стоимость", обеспечивающего заданную эффективность выполнения поставленной задачи ФС $\{A^*\}$ при заданном или оптимальном ресурсе $C_{\Phi A}^*$. В качестве основной целевой функции, удовлетворяющей общей цели синтеза, целесообразным является применение коэффициента экономической целесообразности ($K_{\mathcal{E}Ц}$) использования СИБ $\{A\}$ при решении общей и частных задач в конфликте

$$K_{\mathcal{E}Ц} = \frac{C_A^*(1 - \alpha^1) \bar{P}_A^1(C_A^1, C_B^1, \alpha^1, \beta^1)}{C_A^1 \bar{P}_A^1(C_A^1, C_B^1, \alpha^1 = 0, \beta^1)} \times \\ \times \frac{C_A^2(1 - \alpha^2) \bar{P}_A^2(C_A^2, C_B^2, \alpha^2, \beta^2)}{C_A^2 \bar{P}_A^2(C_A^2, C_B^2, \alpha^2 = 0, \beta^2)} = K_{\mathcal{E}Ц1} K_{\mathcal{E}Ц2}, \quad (2)$$

где C_A^1 , C_A^2 , C_B^1 , C_B^2 — доли ресурсов (выраженные в стоимостных единицах), определяющие состав сил конфликтующих ФС $\{A^*\}$ и $\{B^*\}$ при ведении ЗД/ИД, соответственно ($C_A^1 + C_A^2 = C_A^O$; $C_B^1 + C_B^2 = C_B^O$); α^1 , α^2 , β^1 , β^2 — относительные доли ресурсов, затрачиваемых ФС для реализации целевого предназначения СИБ $\{A\}$ и $\{B\}$ при ведении ЗД/ИД, соответственно ($\alpha^1 + \alpha^2 = \alpha^0$; $\beta^1 + \beta^2 = \beta^0$); $\bar{P}_A^1(C_A^1, C_B^1, \alpha^1, \beta^1)$, $\bar{P}_A^2(C_A^2, C_B^2, \alpha^2, \beta^2)$ — относительное математическое ожидание количества сохраненных ресурсов C_A^1 и C_A^2 ФС $\{A^*\}$ при реализации своего предназначения при ведении ЗД/ИД в конфликте с ФС $\{B^*\}$, использующей ресурсы C_B^1 и C_B^2 в условиях применения СИБ $\{A\}$ ($\alpha^1 = \alpha^2 = 0$ — условия, когда СИБ не применяется).

Основными исполнительными элементами СИБ $\{A\}$ являются разнотипные и разнородные КИБ, реализующие частные задачи обеспечения

действий элементов ФС. Они составляют техническую основу СИБ и являются в принятой иерархической структуре элементами самого нижнего уровня. КИБ формируются на основе средств ИБ, технический облик и способы применения которых предполагаются известными (заданными). По своему предназначению все их множество возможно разделить на две основные группы:

1) комплексы специального назначения, включающие комплексы информационной индивидуальной и групповой безопасности (КИИБ, КИГБ), обеспечивающие применение одиночных и групповых элементов ФС $\{A^*\}$, соответственно;

2) комплексы информационной безопасности общего назначения (КИОБ), обеспечивающие действия больших групп элементов ФС $\{A^*\}$.

В общем случае модель синтеза СИБ представляется в виде иерархической многоуровневой совокупности частных задач синтеза, формулируемых как модели распределения общего ресурса. Под ресурсом СИБ понимается совокупность разнотипных элементов ИБ, информационного обеспечения и управления (ЭИБ, ЭИО, ЭУ), выраженная в относительной доли стоимости, затрачиваемой на их производство и эксплуатацию, от общей стоимости защищаемых элементов ФС. Синтез СИБ осуществляется на основе принципа последовательного наращивания ее структуры при реализации методов ИБ элементов ФС на иерархических уровнях: комплексов — СИБ₃ на иерархически нижнем 3-м уровне (включает совокупность объединенных целью ЭУ, ЭИО и ЭИБ в составе КИИБ); малого и среднего уровня ОТС и ТС — СИБ₂ (ЭУ, ЭИО и ЭИБ в составе КИИБ и КИГБ) и ОТС — СИБ₁ (ЭУ, ЭИО и ЭИБ в составе КИИБ, КИГБ и КИОБ). Исходя из этого, многоуровневая модель распределения общего ресурса ИБ определяется иерархической структурой ФС на уровнях: общий ресурс СИБ (α), равный сумме ресурсов СИБ₁ при реализации ЗД (α^1) и ИД (α^2) и стоимости элементов АСУ ($\alpha_{ПУ}$); ресурс СИБ₁ ($\alpha_{СИБ_1}^{1,2}$), равный сумме стоимостей КИОБ ($\xi_{КИОБ}^{1,2}$), элементов пункта управления (ПУ) СИБ₁ ($\xi_{ПУ}^{1,2}$) и СИБ₂ ($\alpha_{СИБ_2}^{1,2}$); ресурс СИБ₂ ($\alpha_{СИБ_2}^{1,2}$), включающий сумму стоимостей КИГБ ($\varepsilon_{КИГБ}^{1,2}$), ПУ СИБ₂ ($\varepsilon_{ПУ}^{1,2}$) и СИБ₁ ($\varepsilon_{СИБ_1}^{1,2}$), включающих сумму стоимостей КИИБ элементов комплексов — ($\varpi_{ЭК}^{1,2}$) и ПУ СИБ₃ ($\varpi_{ПУ}^{1,2}$). В принципе, возможно рассмотрение и других схем декомпозиции, но при соблюдении основного свойства — иерархического распределения ресурса на различных уровнях.

В обоих случаях (то есть при структурно-параметрическом синтезе ОТТТ к СИБ и ее элементам) для каждого уровня задач подлежат установлению

зависимости целевых функционалов и функций ограничений от множества оптимизируемых вариантов ОТТТ. При этом функции ограничений при разных значениях ресурсов (стоимостного, массогабаритного, энергетического и т. д.) определяют область допустимых вариантов параметров СИБ, оптимизируемых на каждом иерархическом уровне синтеза. Более того, допустимое множество параметров каждого верхнего уровня включает совокупность частных также замкнутых множеств параметров нижнего уровня.

Принцип математической декомпозиции при решении задачи синтеза СИБ

В основу метода математической декомпозиции положен принцип гомотопической инвариантности или "деформации"¹ [3]. Принцип инвариантности общего функционала с изолированной экстремалью совокупности групп разнотипных параметров применительно к рассматриваемой задаче формулируется следующим образом: если в процессе деформации общего функционала его экстремаль изолирована (имеет решение), то она является точкой экстремумов всех частных функций групп параметров.

Поскольку существует, хотя и сложная детерминированная или игровая монотонная взаимосвязь критериальных функционалов каждого верхнего иерархического уровня с функциями нижних уровней, то в основе подхода к решению такого класса задач допустимо использование принципа последовательного "сужения" множества эффективных точек (локальных экстремумов) от нижних уровней к верхним.

Начальным этапом (нулевой уровень) являются задачи параметрического синтеза ЭИБ, ЭИО и ЭУ, включающие уточнение ОТТТ (по видам и параметрам средств ИБ, составу и порядку функционирования) с позиции учета ограничений их размещения на защищаемых ОЭ (ИЭ) и специализированных элементах и способов совместного применения в составе КИИБ, КИГБ и КИОБ.

Особенность формулируемой задачи обусловлена детерминированной зависимостью общего функционала от совокупности частных функциональных компонент оптимальных по Парето², т. е. $F(\Omega) \rightarrow P - \text{opt}$, где $\Omega = (\bar{x}, \bar{y}, \bar{z}) \in R^N$ (N -мерное пространство векторов $\bar{x}, \bar{y}, \bar{z}$); $F(\Omega)$ — аналитическая зависимость компонент $f_i(\Omega, \lambda)$, $0 \leq \lambda \leq 1$, $i = \overline{1, M}$, $\Omega = (\Omega_1, \dots, \Omega_N)$.

Гомотопический принцип устанавливает инвариантность оптимума по Парето, используемого для аналитического представления общего функцио-

¹ Гомотопическая инвариантность с позиции технической интерпретации определяет допустимые условия математического подхода к многоцелевой многоуровневой оптимизации и его сходимость.

² Точка $\Omega^* \in R^M$ является локально оптимальной по Парето отображения $F(f_1(\Omega), \dots, f_M(\Omega))$, если для всех Ω из окрестности V точки Ω^* выполняется условие $F(\Omega^*) - F(\Omega) \geq 0$, $\Omega \in R^N$.

нала. Пусть $(F(\Omega, \lambda) \rightarrow P - \text{opt})$ является невырожденной деформацией задачи $(F(\Omega, 0) \rightarrow P - \text{opt})$ в задачу $(F(\Omega, 1) \rightarrow P - \text{opt})$ и если точки $(\Omega^*, 0)$ и $(\Omega^*, 1)$ являются точками локального P -оптимума, а точка $(\Omega^*, 1)$ является единственной, то эта точка является локально оптимальной для любых функций ограничений параметров $\Omega \in R^N$ [3].

Другая особенность формулируемой задачи — взаимозависимость компонент $f_i(\Omega)$, $i = \overline{1, M}$, не только от группы параметров $\bar{x}, \bar{y}, \bar{z}$, которая обусловлена способами совместного применения средств ИБ, КИИБ, КИГБ и КИОБ, но и множеством параметров условий их применения на различных уровнях типовых дуэлей (уровень средств ИБ), ситуаций (уровень СИБ₃), эпизодов (уровень СИБ₂), сценариев (уровень СИБ₁) и двусторонних действий (уровень СИБ в целом). И эта зависимость обуславливает необходимость учета неопределенности этих условий.

Внешние условия связаны большей частью с возможными составами и способами применения ИЭ конкурирующих ФС и обеспечивающих их функционирование различного типа систем, комплексов и средств ИБ. Поэтому они определяют в большей мере неопределенность целей элементов нижнего уровня СИБ при реализации частных задач ИБ. Методы учета такого рода неопределенностей в моделях конфликта лежат в области теоретико-игрового подхода. Поскольку цели ФС не всегда прямо противоположны, то для определения гарантированного результата при выборе оптимальных составов комплексов и систем ИБ различных иерархических уровней является целесообразным использование биматричных игр с ненулевой суммой, при этом для решения задачи самого верхнего уровня применение многошаговой по этапам двусторонних действий биматричной игры [4].

Для более ясного представления существа метода рассмотрим решение задачи на примере учета межуровневой неопределенности между нижним и 3-м уровнями.

На нижнем уровне разрабатываются предложения по составу разнотипных КИБ для обеспечения действий отдельных элементов ФС при реализации задач ИБ в отдельных типовых ситуациях типа "ИЭ (ОЭ) ФС $\{B^*\} \leftrightarrow$ КИБ ОЭ (ИЭ) ФС $\{A^*\}$ ". При этом внутренними стратегиями максиминного критерия сторон являются: для ФС $\{A^*\}$ — варианты состава средств ИБ и их ТТХ; для ФС $\{B^*\}$ — варианты состава средств заданного типа ИЭ. Внешними условиями сторон являются: для ФС $\{A^*\}$ — варианты массогабаритных и энергетических ограничений при оснащении ОЭ ФС индивидуальными средствами ИБ и варианты способов применения групповых средств ИБ в совпадающих участках (диапазонах) действия; для ФС $\{B^*\}$ — совокупность типов средств активного и/или информационного воздействия для каждого вида ИЭ. Эти условия рассматривают-

ся в качестве стратегий ФС $\{A^*\}$ ($i = 1, \dots, I$) и ФС $\{B^*\}$ ($j = 1, \dots, J$) и используются для формирования исходных матриц показателей эффективности КИБ в интересах обеспечения применения ОЭ (ИЭ) ФС $\{A^*\}$ для расчета показателей эффективности 3-го уровня — СИБ₃.

На 3-м уровне разрабатываются предложения по оснащению элементов КИИБ в зависимости от заданных вариантов массогабаритных и энергетических ограничений. При этом определяются состав КИИБ от воздействия различного целевого предназначения средств ИЭ ФС $\{B^*\}$ в различных пространственно-временных секторах. Состав КИИБ будет, с одной стороны, зависеть от вариантов ограничений, с другой стороны, от вариантов состава совместно используемых КИГБ и КИОБ элементов ФС, а также от вида используемых ИЭ (ОЭ) ФС $\{B^*\}$.

В качестве целевой функции на данном уровне используется зависимость вероятности снижения эффективности функционирования элементов ФС $\{A^*\}$ от активного и/или информационного воздействия различных видов ИЭ (ОЭ) ФС $\{B^*\}$, которая рассчитывается с использованием математических моделей конфликта применительно к типовым эпизодам (например, "средства КИОБ \leftrightarrow средства ИЭ (ОЭ)"). В модели конфликта 3-го уровня подлагат учету задачи целераспределения, назначения (выбора) типа ИЭ (ОЭ), способы его применения и эффективности воздействия на защищаемые элементы СИБ за несколько этапов ИД (ЗД). Таким образом, при выборе состава КИИБ элементов остается неопределенность относительно условий применения различного вида ИЭ ФС $\{B^*\}$. Разрешение этой неопределенности осуществляется аналогично выбору для нижнего уровня с использованием метода биматричных игр с ненулевой суммой ($A \leftrightarrow B: \|a_{ij}^o\|, \|b_{ij}^o\|$). Аналогично решаются задачи и для 2-го (СИБ₂) и 1-го (СИБ₁) уровней.

Неопределенность относительно вариантов ограничений парируется переходом к задачам технико-экономического обоснования оптимального варианта КИИБ (x_*^{opt}), обеспечивающего $x_*^{opt} = \text{Arg}(\varpi_x(\bar{X}_*)) \rightarrow \max_{(x_*)} \min_{(\beta_x)} (1 - \varpi_x(\bar{X}_*))P(\varpi_x(\bar{X}_*), \beta_x, \bar{y}^*, \bar{z}^*)$, $0 \leq \varpi_x(x_*) < 1$; $0 < \beta_x \leq 1$, где $\varpi_x(\bar{X}_*)$, β_x — соответственно, относительная стоимость (к общей стоимости защищаемого элемента — ОЭ (ИЭ)) варианта КИИБ оптимального состава \bar{x}^* для заданных ограничений и варианта стоимости оптимальной стратегии при выборе типа ИЭ (ОЭ) ФС $\{B^*\}$ в модели конфликта уровня типовой ситуации.

Разрешение неопределенности относительно параметров (\bar{y}^*, \bar{z}^*) соответствующих КИГБ и КИОБ элементов СИБ осуществляется методом многошаговой итерационной погрупповой оптимизации при решении задач на верхних уровнях (2-м и 1-м).

Метод решения

Конкретизация целевых функций и условий конфликта ФС $\{A^*\}$ и $\{B^*\}$ в виде (1) позволяет представить конфликт конкурирующих сторон многошаговой биматричной игрой на выживание с ненулевой суммой вида

$$\begin{aligned} \max_{\{p_i\}} \min_{\{g_j\}} \left\{ \left(\sum_{i,j} a_{ij} p_i g_j \right) / \left(\sum_{i,j} b_{ij} p_i g_j \right) \right\} = \\ = \min_{\{g_j\}} \max_{\{p_i\}} \left\{ \left(\sum_{i,j} a_{ij} p_i g_j \right) / \left(\sum_{i,j} b_{ij} p_i g_j \right) \right\}, \quad (3) \end{aligned}$$

где i — стратегии поведения ФС $\{A^*\}$ (варианты СИБ на уровне ОТС, ТС, комплексов на основе КИИБ, КИГБ и КИОБ, реализующих индивидуальные, групповые, индивидуально-групповые и общие методы ИБ ("ИБ", "ГБ", "ИБ ⊕ ГБ", "ОБ") для обеспечения действий элементов ФС); j — стратегии поведения ФС $\{B^*\}$ (варианты состава, структуры и способов применения ее элементов); $\|a_{ij}\|, \|b_{ij}\|$ — платежные матрицы сторон; $a_{ij} = \mathcal{A}_{ij}^A / C_i^A$; $b_{ij} = \mathcal{A}_{ij}^B / C_j^B$; $\mathcal{A}_{ij}^A, \mathcal{A}_{ij}^B$ — выигрыш (проигрыш) ФС $\{A^*\}$ и $\{B^*\}$ при реализации ij -х стратегий, соответственно; C_i^A, C_j^B — стоимости i -го и j -го вариантов состава (стратегий поведения) ФС $\{A^*\}$ и $\{B^*\}$.

Результатом игры является средняя (в смысле математического ожидания) цена n -шаговой ($n = \overline{1, N}$) конкурентной борьбы (игры) Π^* , заключающаяся в выборе оптимальной (равновесной) или рациональной (смешанной) стратегии — $R_{i^*j^*}$ (или R^*).

Поскольку биматричные игры решаются приближенным методом последовательного исключения "доминируемых" стратегий, сводя их к игре типа "2 × 2" и выбору из них одной (или двух) стратегий ФС $\{A^*\}$ — $i_1^* = 1$ (или $i_2^* = 2$), то конечное решение является условным, которое подлежит последующему обсуждению исследователем как с заказчиком, так и с разработчиком с использованием дополнительных критериев (показателей).

Аналитическое решение игры, вследствие нелинейного характера конфликта ФС, является затруднительным, однако оно существует и может быть получено путем последовательного применения принципа оптимальности к группе постоянно усложняющихся математических моделей, начиная с наиболее простых. Наличие иерархии моделей позволяет путем перехода от уже решенных моделей более низкого уровня (в смысле сложности представления) иерархии к более высокому получить решение в обобщенной модели вида (3). Последовательное решение совокупности усложняющихся моделей позволяет на каждом уровне вести оптимизацию только от одной переменной. После того как решение всей совокупности моделей будет получено в явном аналитическом виде, можно получить

конкретные числовые результаты. При этом расчеты ведутся в обратном порядке, путем последовательного разукрупнения (декомпозиции) моделей и их платежных функций. Вначале, на самом верхнем уровне иерархии моделей, определяются доли ресурсов (в стоимостном отношении), которые следует выделить на некоторые две более простые подсистемы. Эти доли ресурсов определяются способами распределения ресурсов в моделях более низкого уровня иерархии, вплоть до самого низшего уровня (уровня физических эффектов), где основными характеристиками являются уже технические показатели конкретных средств и способы их функционирования.

Инвариантная (3) процедура оценки эффективности используется и при решении частных задач нижних уровней ФС с использованием метода групповой оптимизации [5] и гомотопического метода решения нелинейных оптимизационных многоуровневых задач со следующими экстремальными ограничениями [3]: 1) на уровне СИБ — обоснование целесообразного варианта для каждой i -й стратегии поведения ФС $\{A^*\}$ с учетом активного воздействия ее ИЭ (ОЭ) на уязвимые элементы ФС $\{B^*\}$; 2) на уровне ОТС с учетом информационного воздействия на уязвимые элементы ОТС $\{B^*\}$ для каждой i -й стратегии варианта состава СИБ₁ методами ОБ ⊕ ГБ ⊕ ИБ; 3) на уровне ТС — методами ИБ ⊕ ГБ; 4) на уровне комплексов — методами ИБ.

Для определенности рассмотрим процедуру решения одной из инвариантных задач технико-экономического обоснования состава СИБ₁ (1-й уровень оптимизации облика СИБ). Задачей такого обоснования является выбор целесообразного варианта СИБ₁ с позиции принципа оптимальности "эффективность—стоимость" в условиях ограниченного использования ресурса разнотипных средств, комплексов и технических систем ИБ, т. е. ограничений.

В целом задача решается с использованием теоретико-игрового метода, но применительно к частным внешним условиям игры ОТС $\{A_{\text{ОТС}}^*\}$ и $\{B_{\text{ОТС}}^*\}$ (далее A^* и B^*) и к каждой i -й начальной стратегии (далее для упрощения записи i опускаем). В основе игр находятся масштабные численные оценки, которые в относительных величинах позволяют сформировать значения элементов платежных матриц ОТС $\{A^*\}$ и $\{B^*\}$, $\|a_{ij}\|, \|b_{ij}\|$. Коэффициент технико-экономической целесообразности любого r -го, $r = \overline{1, R}$, способа применения ОТС применительно к любому j -му, $j = \overline{1, J}$, способу применения ОТС $\{B^*\}$ K_r^A , вычисляется в виде

$$K_r^A = \frac{C_o^A (1 - \alpha_r) \sum_{i=1}^I \gamma_i^r \left\{ \min_{\{j\}} [\bar{P}_{rj}^i(\alpha_r)] \right\}}{C_o^A \sum_{i=1}^I \gamma_i^r \left\{ \min_{\{j \in J\}} [\bar{P}_{rj}^i(\beta_r = 0)] \right\}}, \quad (4)$$

где K_r^A — отношение сохраненного ущерба (стоимости конкурентоспособных элементов) ОТС $\{A^*\}$ при использовании r -го варианта способа применения СИБ₁ для защиты (реализации эффективных ЗД и ИД) от активного и/или информационного воздействия элементов ФС $\{B^*\}$ к сохраненной стоимости элементов ОТС $\{A^*\}$, но без применения r -го варианта СИБ₁.

В выражении (4) введены следующие обозначения:

C_o^A — суммарная стоимость всех i -х элементов ОТС $\{A^*\}$; $\alpha_r = C_{\text{ОТС}}^r / C_o^A$ — относительная общая стоимость r -го варианта применения СИБ₁; $\gamma_i^r = C_i^r / C_o^A$ — относительная стоимость суммы защищаемых элементов i -го типа; $\bar{P}_{rj}^i(\alpha_r)$ — среднее относительное математическое ожидание числа сохраненных элементов i -го типа ОТС $\{A^*\}$ как функция общей стоимости r -го варианта состава СИБ₁ при j -м воздействии ОТС $\{B^*\}$; $\bar{P}_{rj}^i(\alpha_r = 0)$ — то же, но без СИБ₁ ($\alpha_r = 0$); значения $\bar{P}_{rj}^i(\alpha_r)$ рассчитывают, например, с использованием смешанной аналитическо-стохастической модели [6].

При выборе оптимального варианта СИБ₁ целесообразно использовать максиминный критерий

$$\text{Arg}\alpha^* \rightarrow \max_{\{\alpha_r\}} \left\{ K_r^A = (1 - \alpha_r) \frac{1}{f_o} \sum_{f_o}^I \gamma_i^r \left\{ \min_{\{l \in L\}} [\bar{P}_{rj}^i(\alpha_r)] \right\} \right\}; \quad (5)$$

$$f_o = \sum_{i=1}^I \gamma_i^{r=0} \left\{ \min_{\{j \in J\}} [\bar{P}_{oj}^i(\alpha_r = 0)] \right\};$$

$$\alpha_r = \alpha_1 + \alpha_2 + \alpha_3 + \alpha_{\text{ПУ}},$$

где $\alpha_1, \alpha_2, \alpha_3, \alpha_{\text{ПУ}}$ — относительная (C_A) стоимость соответственно КИИБ, КИГБ, КИОБ и ПУ в составе СИБ₁; отличаются типами, количественным составом и способами применения, их значения заданы (определены на нижних уровнях) при ограничениях:

$$1) \quad P_r = \sum_{i=1}^I \gamma_i^r \left\{ \min_{\{j \in J\}} [\bar{P}_{rj}^i(\alpha_r)] \right\} \geq P_{\text{зад}}; \quad K_r^A \geq 1;$$

$$C_o^A = \text{const};$$

$$2) \quad 0 \leq \alpha_r < 1 \text{ для всех } r \in R;$$

3) исходные условия конкуренции ОТС (многошаговой игры на выживание с ненулевой суммой).

Траектория поиска эффективности СИБ₁ осуществляется аналитической оптимизацией по Нэшу интегральных критериев эффективности отдельных элементов в иерархической структуре ОТС раздельно по составляющим частям управления, информационного обеспечения и ИБ и формированием по ее результатам области допустимых по эффективности решений. Так как выражение (5)

представляет собой максиминную дискретную стохастическую оптимизационную задачу, то для ее решения возможно применить итерационную процедуру последовательной поэтапной оптимизации.

Итерационная процедура обоснования состава и характеристик СИБ

При решении задач 2-го уровня определяется оптимальный состав СИБ₂, включающий совокупность КИИБ всех ОЭ и совокупность КИГБ элементов ТС от воздействия ИЭ ФС $\{B^*\}$ и ИЭ в ОД (ГД) по объектам ФС $\{B^*\}$. На 1-м уровне выполняется обоснование состава СИБ₁, в который могут входить СИБ₂, несколько КИОБ групп объектов и МД ОТС, а также возможно КИИБ и КИГБ отдельных элементов ОТС. Совместное применение КИИБ, КИГБ и КИОБ обуславливает необходимость перераспределения их ресурсов при реализации задач ИБ действий ОТС, а следовательно, взаимную зависимость ОТГТ к разнотипным КИБ и, в общем, нелинейную зависимость от них целевой функции. Для решения этих задач возможно применение метода погрупповой оптимизации, известного под названием метода Гаусса—Зейделя [5].

Существо метода погрупповой оптимизации состоит в использовании многошаговой итерационной процедуры. Применительно к задаче 3-го уровня интерпретацию метода можно выразить следующим образом. Пусть некоторая целевая функция $f_{1i}(\bar{x}_1, \bar{y}_1, \bar{z}_1)$ зависит от состава и характеристик КИИБ (вектор параметров \bar{x}), КИГБ (\bar{y}) и КИОБ (\bar{z}) основных элементов ФС и есть ряд ограничений типа $\varpi(\bar{x}) + \varepsilon(\bar{y}) + \xi(\bar{z}) \leq \alpha_p^*$; $0 \leq \alpha_p^* < 1$; $0 \leq \varpi(\bar{x})$; $0 \leq \varepsilon(\bar{y})$; $0 \leq \xi(\bar{z})$. При этом существует детерминированная зависимость целевой функции $f_{1i}(\bar{x}_1, \bar{y}_1, \bar{z}) = F[\{f_{2j}(\bar{x}_2, \bar{y}_2, \bar{z}_3^*), j \in J\}; \{f_{1k}(\bar{x}_1, y_k^*, z_3^*), k \in K\}]$, где $\bar{x}, \bar{y}, \bar{z}$ — группы оптимизируемых параметров, соответственно, КИИБ, КИГБ и КИОБ в составе СИБ₃ (\bar{x}_3), уровня СИБ₂ (\bar{x}_2, \bar{y}_2) и СИБ₁ ($\bar{x}_1, \bar{y}_1, \bar{z}_1$); $f_{1i}(\dots), f_{2j}(\dots)$ и $f_{1k}(\dots)$ — целевые функции для 3-го, 2-го и 1-го уровней, соответственно.

Это позволяет реализовать итеративный процесс: на первом шаге (1-я итерация) определяется условно-оптимальный вектор параметров \bar{x}_3^{-1} относительно k -х КИИБ 3-го уровня, обеспечивающий для заданных значений параметров \bar{y}_2^* и \bar{z}_1^* , соответственно КИГБ и КИОБ, максимум целевой функции 3-го уровня $f_{1k}(\bar{x}_1^{-1}, \bar{y}_2^*, \bar{z}_1^*)$ при ограничениях $0 \leq \varpi(\bar{x}_3) < 1$; 2-я итерация — определяется условно-оптимальный вектор параметров \bar{y}_2^{-1} относительно j -х КИГБ 2-го уровня СИБ₂, обеспечиваю-

шего при определенном в результате 1-й итерации условно-оптимальном векторе параметров КИИБ (\bar{x}_1^3) и заданном значении параметров КИОБ (z_1^*) максимуму целевой функции 2-го уровня $f_{2j}(\bar{x}_2^1, \bar{y}_2, z_1^*)$, при ограничениях $0 \leq \varepsilon(\bar{y}_2) < \alpha_p^*$; 3-я итерация — определяются условно-оптимальные векторы параметров \bar{z}_1^1 относительно i -х КИОБ 1-го уровня СИБ₁, обеспечивающих при определенных на 3-м и 2-м уровнях условно-оптимальных векторов параметров КИИБ (\bar{x}_1^1) и КИГБ (\bar{y}_1^1) максимум целевой функции 1-го уровня $f_{1i}(\bar{x}_1^1, \bar{y}_1^1, \bar{z}_1^1)$, при ограничениях $0 \leq \xi(\bar{z}_1) < \alpha_p^*$. Для 1-го шага запоминается условно-оптимальное значение целевой функции $f_{1i}^1(\bar{x}_1^1, \bar{y}_1^1, \bar{z}_1^1)$.

На втором и последующих шагах итеративный процесс повторяется, т. е. последовательно решаются частные задачи для трех итераций и определяются условно-оптимальные значения векторов на втором и последующих l -х шагах ($\bar{x}_3^l, \bar{y}_3^l, z_3^l$) и соответствующие им значения целевой функции $f_1^l(\bar{x}_1^l, \bar{y}_1^l, z_1^l)$, $l = 1, \dots, L$.

Считается, что многошаговый процесс при определенных $f_1^l(\bar{x}_1^l, \bar{y}_1^l, z_1^l)$ сходится по Нэшу за конечное число итерационных шагов. При выполнении условия $(f_{1i}^{l*}(\dots) - f_{1i}^{(l-1)*}(\dots)) \leq \gamma_{\text{зад}} \rightarrow 0$ ($l = 1, \dots, l^*$; $l^* = L$) процесс останавливается, за оптимальное принимается решение, соответствующее l^* -му шагу ($\bar{x}_1^{l*}, \bar{y}_1^{l*}, z_1^{l*}$).

Необходимыми условиями сходимости являются непрерывность целевых функций всех уровней и монотонная зависимость функций верхнего уровня от целевых функций нижних уровней. Известно, что данный метод оптимизации неустойчив в отношении целевых функций, имеющих несколько экстремальных точек (экстремалей). Вместе с тем имеется ряд подходов, учитывающих эту неопределенность при использовании некоторой модификации метода случайного поиска при выборе начальных условий на первом шаге. Кроме того, для ускорения сходимости при решении конкретных системотехнических задач возможно использование сочетания методов случайного и целенаправленного поиска.

Следует отметить, что при решении задач на 2-м и 1-м уровнях остается межуровневая неопределенность, обусловленная типовыми условиями частных конфликтов на уровне типовых эпизодов (уровень ТС) и сценариев (уровень ОТС). Разрешение этой неопределенности осуществляется в соответствии с [4] на основе теории многошаговых биматричных игр с ненулевой суммой (матриц $\|a_{ij}^2\| - \|b_{ij}^2\|$ и $\|a_{ij}^3\| - \|b_{ij}^3\|$) при выборе оптимальных составов КИГБ (на уровне эпизодов) и КИОБ (на уровне

сценариев). Кроме того, на 1-м уровне остается неопределенность относительно целей ИБ ОЭ от ИЭ и ИЭ от средств ФС $\{B^*\}$. Необходимость разрешения этой неопределенности обусловлена взаимосвязью ресурса КИОБ при реализации одновременно этих двух целей, т. е. ресурс одного КИОБ может одновременно использоваться для решения задач ИБ как от элементов в ИД, так и ЗД ФС $\{B^*\}$.

Разрешение этой неопределенности осуществляется при решении задачи самого верхнего (в нашем случае 0-го) уровня — выбора варианта СИБ.

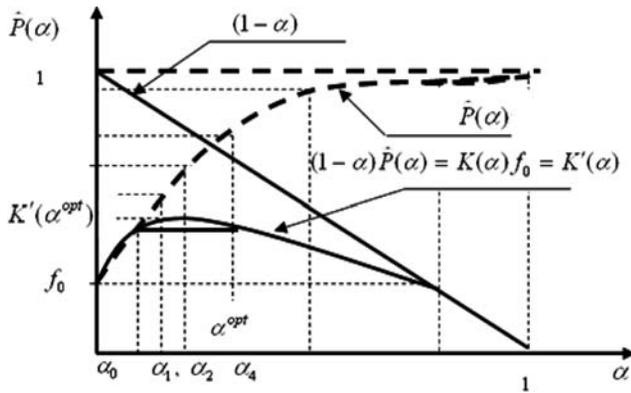
Аналитическое представление решения задачи синтеза СИБ₁

Структурно задача синтеза СИБ решается поуровнево. В качестве исходных данных (нулевой уровень иерархии СИБ₁) используются данные результатов элементарного взаимодействия между средствами ОТС $\{B\}$ и различными методами и средствами информационного воздействия. На этом уровне, исходя из информации о свойствах, структуре и характеристиках информационного пространства исходной размерности, формулируется возможное множество вариантов последовательностей информационных воздействий, подлежащих исследованию.

На 1-м уровне иерархии определяются базисные свойства информационного пространства взаимодействия элементов КИБ с информационными средствами ИЭ (ОЭ) конкурирующих ОТС. Цель моделирования состоит в формировании полной модели фазового пространства состояний элементов конкурирующих ОТС. Оптимизация параметров фазового пространства информационного взаимодействия элементов ОТС осуществляется на основе расчета гарантированных вероятностных оценок функционирования ее элементов применительно к условиям типового ядра конфликта уровня "типовая ситуация". В качестве параметров рассматриваются исходные данные по характеристикам взаимодействующих элементов ОТС, уравнения для расчета параметров пространства фазового состояния ОТС, различные целевые функции и вид параметров моделирующих информационных воздействий. Из всех результатов расчетов выбираются те, для которых: а) экстремальные точки значений целевых функций представлены в максимальном диапазоне значений; б) при наличии экстремальных значений выделяются области, в которых они расположены. Оптимизация результатов на данном уровне основывается на получении фазового пространства взаимодействия элементов конкурирующих ОТС в условиях применения элементов СИБ₁.

На первом этапе данного уровня:

а) с использованием полученных решений на нижних уровнях, относительно возможных оптимизационных решений при обосновании оптимальных составов, структуры, порядка функционирования и характеристик СИБ₂ (КИИБ, КИГБ, КИОБ



Ряд характерных зависимостей при синтезе СИБ

и ПУ), т. е. условно-оптимальных значений множества $(\alpha_1^*, \alpha_2^*, \alpha_3^*, \alpha_{ПУ}^*)$ формируются варианты состава СИБ₁ по принципу "0" или "1", т. е.: 0001, 0010, 0011, 0100, 0101, ..., 1111;

б) все принятые решения нумеруются по возрастанию номеров r от 0001 до 1111, рассчитываются значения α_r , для $r = \overline{1, R}$ ($R = 8$) и отбрасываются все значения α_r , удовлетворяющие условию $0 \leq \alpha_r < 1$.

На *втором этапе*: для каждого r -го, $r = \overline{1, R}$, варианта состава СИБ₁ (следовательно $|\alpha_r| < 1$) рассчитываются с использованием математической модели [например, 6] значения $|P_r^*|$, $r \in R$. Модель учитывает все возможные стратегии поведения ОТС $\{B\}$ ($j \in J$) за счет оптимального целераспределения применительно к условиям каждой ее стратегии. По вычисленным дискретным значениям формируется методами кусочно-линейной аппроксимации непрерывная функция $\hat{P}(\alpha) \approx P(\alpha)$, обобщенный вид которой приведен на рисунке.

На *2-м уровне*, по данным, полученным с предыдущего уровня, определяются экстремальные значения целевых функций элементов СИБ₁ в N -элементном пространстве информационного взаимодействия ОТС и применительно к элементам меньшей размерности в структуре условий стохастического конфликтного взаимодействия ОТС проверяется соблюдение в них принципа гомотопической инвариантности на основе:

а) определения значений α^* из условия $\max_{0 \leq \alpha \leq 1} \{ \hat{K}(\alpha) = K(\alpha)f_0 = (1 - \alpha)\hat{P}(\alpha) \}$ при следую-

щих ограничениях: $0 \leq \alpha < 1; f_0 < \hat{P}(\alpha) \leq 1$, а также расчет оптимальных значений $\hat{K}(\alpha = \alpha^{opt})$ и значения $\hat{P}(\alpha = \alpha^{opt})$;

б) проверки условия $\hat{P}(\alpha = \alpha^{opt}) \geq P_{тр}$; если условие выполняется, то это является достаточным для вычеркивания всех значений $\alpha < \alpha^{opt}$ до значения α' , при котором $\hat{P}(\alpha') \geq P(\alpha)$; если условия не выполняются, требуется проведение анализа дальнейших решений ... →.

На *3-м уровне* (уровень аналитической оптимизации) анализируется внутренняя структура фазового пространства состояний, строятся семейства изомногообразий, характеризующие результаты информационного взаимодействия ОТС и определяется принадлежность экстремальных значений целевых функций вариантов облика СИБ₁ области допустимых решений.

На *4-м уровне* иерархии (поиск окончательного решения) выявляются особенности фазового пространства информационного взаимодействия элементов ОТС и обоснование по Парето предпочтительного варианта облика СИБ₁. На этом уровне иерархии исследуются уже детерминированные процессы информационного взаимодействия ОТС, для которых влияние случайных возмущений незначительно:

а) поскольку решается частная (по условию) задача на основе гомотопического метода исследования нелинейных оптимизационных задач [3], возможен переход от абсолютного решения, т. е. $\alpha = \alpha^{opt}$, к решению оптимальности по Парето. Это обусловлено тем, что выбор в допустимой области предпочтительного варианта СИБ₁, исходя из требования минимизации стоимости реализации ее облика в виде (1), осуществляется на основе методов векторной оптимизации, сужающих область возможного компромисса за счет ослабления требований к критериям эффективности введением допустимого предела их изменений;

б) решение оптимальности по Парето позволяет выделить и сузить область возможного компромисса ОТТТ к СИБ на основе ослабления исходных требований к критерию "эффективность—стоимость" (в случае $\alpha^{opt} \neq \Delta\alpha^П$), где $\Delta\alpha^П$ есть некоторый заданный предел (например: 5 % от $\hat{P}(\alpha = \alpha^{opt})$ в условиях $|\Delta\alpha^П| \rightarrow 0$). Введение условия оптимизации по Парето при обосновании предпочтительных (уже не оптимальных) вариантов СИБ является существенным, поскольку может значительно улучшить решение на вышестоящем уровне, снизить меру требований к СИБ₁ за счет совместного использования в ее составе разнотипных КСБ и приблизиться к принципам оптимальности (модели предпочтений) заказчика СИБ₁;

в) в этом случае целесообразным вариантом состава СИБ является тот, который обеспечивает минимальное значение $(\alpha^{opt} - \Delta\alpha^П)$ стоимости ее вариантов.

Оптимизационные задачи применительно к нижним уровням иерархии СИБ₁ могут быть поставлены и решены аналогичным образом.

Сходимость вариантов облика СИБ в метрическом пространстве параметров элементов управления, информационного обеспечения и ИБ ФС, исходя из принципа сжимающих отображений [6], обеспечивается нахождением точки пересечения множеств вариантов на различных аспектах синтеза.

Данная точка и определяет предпочтительный облик СИБ [7].

Реализация приведенного метода синтеза облика СИБ обеспечивает эффективное функционирование ФС в условиях оптимального конкурентного активного/информационного противодействия.

Список литературы

1. Мистров Л. Е. Методы и средства информационной безопасности организационно-технических систем // Информационная безопасность регионов. 2010. № 1 (6). С. 22–32.

2. Мистров Л. Е., Сербулов Ю. С. Методический подход к системному представлению многоцелевого конфликта функ-

циональных иерархических систем // Информационно-измерительные и управляющие системы. 2005. № 3. Т. 3. С. 26–30.

3. Бобылев Н. А., Коровин С. К., Скалыга В. И. О гомотопическом методе исследования многокритериальных задач // Автоматика и телемеханика. 1996. № 10.

4. Вентцель Е. С. Исследование операций. М.: Сов. радио, 1972. 550 с.

5. Гурич Л. С., Дымарский Я. С., Меркулов А. Д. Задачи и методы оптимального распределения ресурсов. М.: Сов. радио, 1968. 463 с.

6. Мистров Л. Е. Математическая модель эффективности обеспечивающей организационно-технической системы // Приборы и системы. Управление, контроль, диагностика. 2007. № 1. С. 64–70.

7. Колмогоров А. Н., Фомин С. В. Элементы теории функций и функционального анализа. М.: Наука, 1981.

УДК 519.622.2

В. Б. Маничев, канд. техн. наук, доц., e-mail: manichev@bmstu.ru,

Д. М. Жук, канд. техн. наук, доц., e-mail: zhuk@bmstu.ru,

Ф. Т. Витюков, аспирант, e-mail: vietfed7@gmail.com,

МГТУ имени Н. Э. Баумана

Достоверное и точное численное моделирование динамических систем при невысоких требованиях к математической точности конечных результатов

Основной недостаток известных программных комплексов численного моделирования динамических систем, например MATLAB-SIMULINK, состоит в получении неверного результата численного моделирования динамических систем при невысоких заданных требованиях к математической точности конечных результатов такого моделирования. Невысокие требования к математической точности решения математических моделей динамических систем объясняются тем, что исходные внутренние параметры этих систем и, как следствие, соответствующие значения коэффициентов дифференциальных уравнений, как правило, получены с невысокой математической точностью. В данной работе рассмотрены методы, алгоритмы и программа ПА10 (SADEL-PA10), устраняющие этот недостаток. Приведены результаты соответствующих численных экспериментов.

Ключевые слова: математическое моделирование, численное моделирование, динамические системы, обыкновенные дифференциальные уравнения (ОДУ), дифференциально-алгебраические уравнения (ДАУ), линейные алгебраические уравнения (ЛАУ), методы интегрирования

V. B. Manichev, D. M. Zhuk, F. T. Vitukov

Reliable and Accurate Numerical Modeling of Dynamic Systems at Low Requirements to the Mathematical Accuracy of The End Results

The main shortcoming of known software package for dynamic systems simulation, for example, MATLAB-SIMULINK, consists in receiving incorrect result of dynamic systems simulation at the low requirements to the mathematical accuracy for the end results of simulation. The low requirements to the mathematical accuracy of the end results are explained by that internal parameters of dynamic systems and, as a result, corresponding values of the differential equations coefficients, are received as a rule with a low mathematical accuracy. In this article methods, algorithms and program SADEL-PA10 eliminating this shortcoming are considered. Results of the corresponding numerical experiments are given.

Keywords: mathematical modelling, simulation, dynamical systems, ordinary differential equations (ODE), differential algebraic equations (DAE), linear algebraic equations (LAE), integration methods

Введение

В статье используется следующая терминология.

Математическая модель (*mathematical model*) — система ОДУ—ДАУ, отображающая функционирование моделируемой динамической системы с требуемой достоверностью и точностью.

Математическое моделирование (*mathematical modeling*) — процесс получения математических моделей динамических систем на основе фундаментальных законов физики, химии и т. п.

Компьютерная модель — алгоритмическая и программная реализация математической модели.

Численное (компьютерное) моделирование (*simulation*) — процесс получения результатов математического моделирования динамических систем с помощью компьютерных моделей.

Жесткие системы ОДУ—ДАУ (*stiff ODE—DAE systems (stiff set of ODE—DAE)*) — системы ОДУ—ДАУ со степенью жесткости более 10^6 .

Плохо обусловленные системы ЛАУ (*badly conditioned (ill-conditioned) LAE systems (ill-conditioned set of LAE)*) — системы ЛАУ с числом обусловленности более 10^6 , к решению которых часто сводится решение жестких и сверхжестких систем ОДУ—ДАУ неявными методами интегрирования.

Достоверность численного моделирования (*reliability of simulation*) — отклонение результатов компьютерного моделирования динамических систем от заведомо точных результатов не более чем на 20 % (точные результаты получают аналитически или экспериментально, как правило, оценивая относительную погрешность численного моделирования в процентах).

Точность численного моделирования (*accuracy of simulation*) — фактическая степень отклонения результатов достоверного численного моделирования от заведомо точных результатов.

Математическая точность (*mathematical accuracy*) решения систем ОДУ—ДАУ — заданная относительная погрешность нормы векторов решения систем ОДУ—ДАУ (параметр *eps*) по сравнению с заведомо точным аналитическим решением в относительных единицах (в MATLAB по умолчанию *eps* = 0,001).

Компьютерная точность (precision) решения систем ОДУ—ДАУ и ЛАУ — количество полученных верных значащих цифр в каждом элементе векторов решения систем ОДУ—ДАУ и ЛАУ по сравнению с заведомо точным аналитическим (идеальным для систем ЛАУ) решением.

Решение систем ЛАУ с повышенной точностью вычислений (*LAE systems solution with extra precision calculations*) — применение методов получения решений с повышенной точностью вычислений (*extra precision calculations*), реализованных в пакетах прикладных математических программ, а также в библиотеках математических программ и описанных в документации на соответствующие пакеты и библиотечки.

Пакеты прикладных математических программ для получения решений с повышенной точностью вычислений: Maple (<http://maplesoft.com> — реализован метод Software Floating Point; MATLAB (<http://mathworks.com>) — реализован метод Variable Precision Arithmetic; Mathematica (<http://wolfram.com>) — реализован метод Arbitrary Precision Arithmetic.

Библиотеки программ для получения решений с повышенной точностью вычислений: библиотека Intel для десятичных вычислений с повышенной точностью (стандарт IEEE 754—2008 Decimal Floating-Point for Intel® Architecture Processors, IEEE Standard for Floating-Point Arithmetic 754-2008, IEEE, 2008: <http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=5223319&url> [1]), а также многочисленные Си (Си++) библиотеки для вычислений с повышенной точностью. Возможно также использование переменных типа Decimal на языке C# или класса java.math.BigDecimal на языке Java.

Научно-исследовательские работы по математическому и численному моделированию динамических систем выполняются преподавателями, аспирантами, сотрудниками и студентами кафедры САПР МГТУ имени Н. Э. Баумана с 1974 года (первоначально эти работы выполнялись под руководством профессора И. П. Норенкова для математического и компьютерного моделирования электронных схем). Были разработаны программные комплексы (ПК) ПА1—ПА9 для численного моделирования динамических систем. При программной реализации численных методов решения систем ОДУ—ДАУ (при разработке математических ядер этих ПК) следует выделить два основных типа погрешностей численного решения систем ОДУ—ДАУ — глобальная погрешность (или качественная ошибка) и локальная погрешность (или количественная ошибка). Математики-программисты основное внимание уделяют оценке математической локальной погрешности интегрирования (обычно используя метод Рунге или методы, основанные на предсказании-коррекции (*prediction-correction methods*)), хотя главное — это получение качественно достоверного и корректного численного решения систем ОДУ—ДАУ.

В работе [2] было показано, что для гарантии получения качественно корректного решения разнообразных систем ОДУ при контроле только локальной погрешности интегрирования численный метод решения систем ОДУ должен быть *AL*-устойчивым (или $A(\pi/2)$ -устойчивым), т. е. абсолютно (*A*-) устойчивым строго в левой (*L*-) полуплоскости комплексной плоскости устойчивости методов численного решения систем ОДУ—ДАУ.

В работе [3] на основе Паде-аппроксимации экспоненты в конечную дробь и использования многочленов Якоби были получены коэффициенты *AL*-устойчивых неявных методов интегрирования

систем ОДУ—ДАУ 2-го, 4-го и 6-го порядков точности. Программная реализация этих методов сводится к многократному получению матриц Якоби для решаемой системы ОДУ—ДАУ и решению соответствующих систем нелинейных алгебраических уравнений (НАУ), которые для получения компьютерной точности сводятся к решению соответствующих систем ЛАУ на каждом шаге интегрирования, что, как правило, приводит к нескольким тысячам и более обращений к программе-решателю систем ЛАУ (*LAE solver*) на всем заданном отрезке интегрирования системы ОДУ—ДАУ. Эти *AL*-устойчивые неявные методы интегрирования систем ОДУ—ДАУ 2-го и 4-го порядков точности были реализованы в составе библиотеки SADEL при использовании удвоенной точности вычислений на языке Си [4].

Конечная математическая и компьютерная точность численного моделирования динамических систем может быть невысокой, так как исходные внутренние параметры моделируемых систем и, как следствие, соответствующие значения коэффициентов систем ОДУ—ДАУ, получают с невысокой математической и компьютерной точностью, поэтому, по умолчанию, математическая точность решения систем ОДУ—ДАУ задается невысокой ($eps = 0,001$ в ПК MATLAB). Решение с данным параметром eps большого числа тестовых и практических задач математического и численного моделирования динамических систем с помощью библиотеки SADEL и других библиотек и ПК показало, что для получения качественно достоверного и корректного решения жестких и сверхжестких систем ОДУ—ДАУ необходимо на всех шагах численного интегрирования обеспечивать численное решение соответствующих систем ЛАУ с повышенной точностью, т. е. с точностью в 15 верных значащих цифр для всех значений элементов вектора решения систем ЛАУ [7]. Также был определен основной недостаток известных программных комплексов моделирования динамических систем (на примере ПК MATLAB-SIMULINK): возможное получение неверного результата (без предупреждения пользователя) численного моделирования динамической системы при невысоких заданных требованиях к математической точности интегрирования систем ОДУ—ДАУ (параметр eps , по умолчанию $eps = 0,001$ в ПК MATLAB-SIMULINK).

В настоящее время разрабатывается ПК ПА10 (SADEL-PA10), в котором планируется устранение этого недостатка. Библиотека SADEL (*Sets of Algebraic and Differential Equations solvers Library*) — библиотека решателей для систем алгебраических и дифференциальных уравнений, предназначенная для решения систем ЛАУ и систем ОДУ—ДАУ. Библиотека SADEL на языке Си является математическим ядром платформы компьютерного моделирования динамических процессов для разнородных (*multiphysics, multidiscipline*) технических систем и

объектов — ПК ПА10 (SADEL-PA10)), превосходящем по достоверности и точности результатов математические ядра подобных зарубежных ПК — MATLAB-SIMULINK, Maple-MapleSim, Mathematica-SystemModeler, C-Library NAG в части решения жестких систем ОДУ и плохо обусловленных систем ЛАУ при невысоких требованиях к математической точности результатов численного моделирования [5, 6].

Обоснование принципов разработки ПК ПА10 (SADEL-PA10)

Для достоверного и точного численного моделирования динамических систем необходимо обеспечить достоверность и точность численного решения систем ОДУ—ДАУ.

В известных программах-решателях систем ОДУ—ДАУ (ODE—DAE Solvers) используются три классические постановки задач решения систем ОДУ—ДАУ.

1. Нормальная форма Коши в координатном базисе дифференциальных переменных состояния (явная форма представления систем ОДУ—ДАУ):

$$dX/dt = F(X, t),$$

где X — вектор координатного базиса дифференциальных переменных состояния размерностью m ; F — вектор-функция правых частей системы ОДУ размерностью m ; t — независимая переменная (обычно время). Заданы начальные условия $X_0 = X(0)$ и отрезок интегрирования $t = [T0, TK]$, $T0$ — заданное время начала интегрирования; TK — заданное время окончания интегрирования.

2. Дифференциально-алгебраическая форма разных индексов в полном координатном базисе дифференциально-алгебраических переменных (полужавная форма):

$$\begin{cases} dX/dt = F(X, Y, t) \\ G(X, Y) = 0 \end{cases} \text{ или } M dX/dt = F(X, t),$$

где Y — вектор алгебраических переменных полного координатного базиса дифференциально-алгебраических переменных размерностью k ; G — вектор-функция размерностью k ; M — сингулярная матрица ранга m размером $(m + k)^2$. Заданы согласованные начальные условия: $X_0 = X(0)$, $Y_0 = Y(0)$ и отрезок интегрирования $t = [T0, TK]$.

3. Дифференциально-алгебраическая форма в полном координатном базисе дифференциально-алгебраических переменных (неявная форма):

$$G(X, dX/dt, Y, t) = 0; dX/dt = F(X, t),$$

где G — вектор-функция размерностью $m + k$. Заданы согласованные начальные условия и отрезок интегрирования. Задача решения систем ОДУ—ДАУ в данной постановке была поставлена в полном координатном базисе дифференциально-алгебраических переменных Линдой Петзолд и ее научным руководителем, известным математиком Гиром в 1982 г., была разработана программа DASS (*Diffe-*

rential Algebraic Systems Solver) на основе метода BDF (*Backward Differential Formula*), заменой в вышеприведенной формуле вектора производных по формулам Гира $dX/dt = BDF(X, h)$ и решением на каждом шаге интегрирования замкнутой системы нелинейных алгебраических уравнений (НАУ) $G(X, Y) = 0$ относительно дифференциально-алгебраических переменных X и Y [8].

Программа-решатель систем ДАУ (DAE Solver) в ПК ПА10 (SADEL-PA10), рассматривает системы ДАУ в еще более **расширенном координатном пространстве переменных**, в котором к дифференциально-алгебраическим переменным добавлены производные дифференциальных переменных. Задача решения систем ДАУ в расширенном координатном пространстве переменных была поставлена в работе [9]:

$$G(PX, X, Y, t) = 0, \quad (1)$$

где $PX = dX/dt$ — вектор производных дифференциальных переменных по времени для расширенного координатного пространства переменных размерностью m ; G — вектор-функция размерностью $m + k$. Заданы начальные условия $X_0 = X(0)$ и отрезок интегрирования $t = [T0, TK]$.

Все численные методы решения систем ОДУ—ДАУ основаны на последовательном нахождении значений векторов $X(t)$ и $PX(t)$, $Y(t)$, начиная с нулевых начальных условий, т. е. $X_0, PX_0, Y_0, X_1, PX_1, Y_1, \dots, X_{n-1}, PX_{n-1}, Y_{n-1}, X_n, PX_n, Y_n, \dots$, где n — номер шага интегрирования. Переход от каждого предыдущего значения $X_{n-1}, PX_{n-1}, Y_{n-1}$ к последующему X_n, PX_n, Y_n осуществляется по формулам численного интегрирования, составляющим основу тех или иных методов численного решения систем ОДУ. На рис. 1. сплошной линией показано точное, а жирными точками — численное решения для i -й дифференциальной переменной системы (1). Также на рис. 1. использованы следующие обозначения: EAt_n — локальная погрешность (погреш-

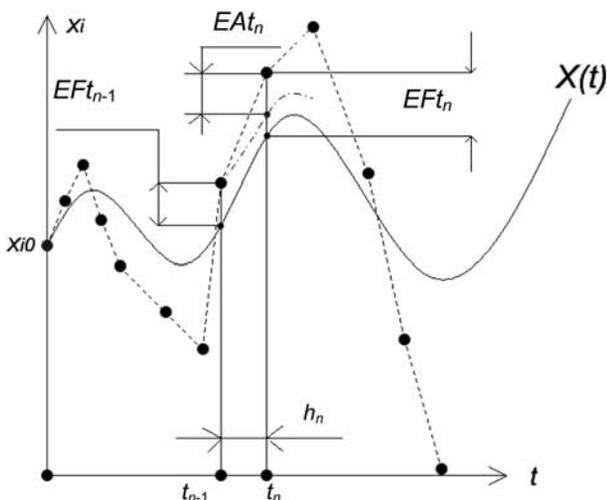


Рис. 1. Полная и локальная погрешность численного решения систем ОДУ—ДАУ

ность аппроксимации) метода интегрирования на текущем шаге интегрирования h_n , равная разности между численным решением, полученным в конце шага интегрирования, и точным решением, полученным из численного решения в начале шага интегрирования (показано штрих-пунктирной линией на рис. 1); EFt_{n-1} — это полная, накопленная к началу шага интегрирования h_n погрешность численного интегрирования; EFt_n — это полная погрешность численного интегрирования после выполнения шага интегрирования h_n .

При программной реализации численных методов решения систем ОДУ—ДАУ, как правило, контролируется только локальная погрешность интегрирования (параметр eps в программах численного решения систем ОДУ—ДАУ), но при невысоких требованиях к точности интегрирования ($eps = 0,001$) возможно получение недостоверного численного решения, так как не локальная, а полная погрешность определяет устойчивость получаемого численного решения системы ОДУ—ДАУ, и отсутствие контроля полной погрешности является причиной получения возможных недостоверных решений.

Новые алгоритмы для решения системы (1) основаны на реализации одношаговых методов интегрирования (многошаговые методы с порядком точности выше 2-го не являются *AL*-устойчивыми) и совместном решении систем НАУ $G(X, Y, PX, t_i) = 0$ и систем из m алгебраических уравнений численного интегрирования $H_i(X, PX, h) = 0$, которые доопределяют систему (1) на соответствующих стадиях одношаговых методов интегрирования, относительно векторов дифференциально-алгебраических переменных X, Y и производных дифференциальных переменных PX [6]. Параметры алгебраических уравнений численного интегрирования $H_i(X, PX, h) = 0$ были получены, исходя из требования получения при любых шагах интегрирования устойчивых решений для пяти базовых фундаментальных решений линейных неоднородных систем ОДУ [11]. При невысокой математической точности интегрирования шаг интегрирования может стать таким большим, что **не *AL*-устойчивые методы интегрирования могут потерять устойчивость для некоторых фундаментальных решений**. Область абсолютной устойчивости *AL*-устойчивых методов интегрирования показана на рис. 2.

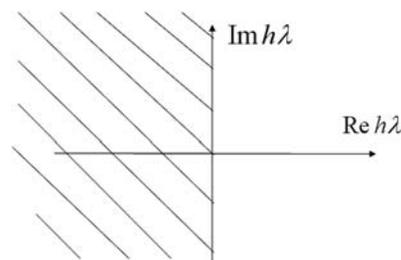


Рис. 2. Область абсолютной устойчивости (заштрихована) *AL*-устойчивых методов интегрирования

Были разработаны системы из m линейных алгебраических уравнений численного интегрирования $H_i(X, PX, h) = 0$ S -стадийных неявных методов интегрирования систем ДАУ вида (1) в расширенном координатном пространстве дифференциально-алгебраических переменных и производных дифференциальных переменных X, Y и PX (DABC методы интегрирования систем ДАУ) [3]:

$$H_i(PX_i, X_i, X_{n-1}, PX_{n-1}, h_n) = h_n \sum_{j=1}^s d_{ij} P_j X - \sum_{j=1}^s a_{ij} X_j - b_i X_{n-1} - h_n c_i P X_{n-1} = 0;$$

$$X_n = X_s, PX_n = P X_s, i = 1, \dots, s, t_n = t_s;$$

s — число стадий;

$h_n = (t_n - t_{n-1})$ — n -й шаг интегрирования;

d_{ij}, a_{ij}, b_i, c_i — параметры метода.

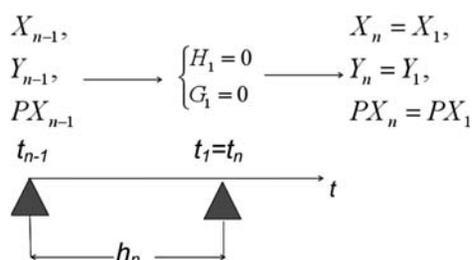


Рис. 3. Схема алгоритма неявных одностадийных одношаговых методов Эйлера и трапеций (для М1 и М2 — число стадий $S = 1$)

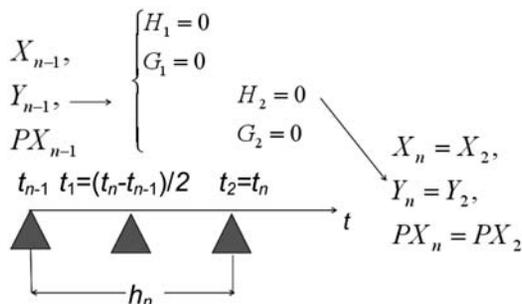


Рис. 4. Схема алгоритма неявного двухстадийного одношагового метода (для М3 — число стадий $S = 2$)

На основе этих алгебраических уравнений численного решения систем ОДУ—ДАУ были определены параметры и реализованы три метода интегрирования [6]:

М1 — A -устойчивый неявный метод первого порядка точности (совпадает по параметрам с неявным методом Эйлера);

М2 — AL -устойчивый неявный метод второго порядка точности (совпадает по параметрам с неявным методом трапеций);

М3 — AL -устойчивый неявный метод четвертого порядка точности (совпадает по параметрам с неявным методом Лобатто IIIA [10])

На рис. 3 и рис. 4 показаны схемы работы алгоритмов для реализации этих методов.

Точность решения систем ОДУ—ДАУ в методах М2, М3 обеспечивается контролем локальной погрешности используемого метода интегрирования, а достоверность обеспечивается AL -устойчивостью данных методов интегрирования.

Результаты численных экспериментов

Сравнение программ-решателей систем ОДУ—ДАУ с решателем М3 библиотеки SADEL при невысоких требованиях к точности интегрирования ($eps = 0,001$) было проведено для тестовых задач с известным аналитическим решением [5, 6]. В табл. 1 и 2 приведены обобщенные результаты сравнения.

Знак (–) означает недостоверное решение теста без предупреждения.

В табл. 1 и 2 сравнивались только методы интегрирования программ-решателей систем ОДУ—ДАУ, рекомендуемые в соответствующих ПК и библиотеках математических программ для решения жестких систем ОДУ—ДАУ. Знак минус означает невозможность получения решения или (в большинстве случаев) качественно неверный результат без всякого предупреждения пользователя о возможных ошибках. В качестве примера на рис. 5 представлен результат качественно неверного решения тестовой задачи с известным аналитическим реше-

Таблица 1

Программа-решатель систем ОДУ—ДАУ	ТЕСТЫ					
	ТЕСТ 1	ТЕСТ 1	ТЕСТ 2	ТЕСТ 2	ТЕСТ 3	ТЕСТ 4
	Жесткость MU=1e6	Жесткость MU=1e9	Коэффициенты kt=1, ki=1, ku=0,01	Коэффициенты ku=1, ki=1, kt=1e-104	Жесткость MU=1e6	Параметры реального лазера
SADEL C-Library 2010 Метод М3	+	+	+	+	+	+
NAG C-Library 2012 Метод BDF	+	–	–	–	–	+
IMSL C-Library 2011 Метод BDF	+	–	–	–	–	–

Знак (–) означает недостоверное решение теста без предупреждения. Си-библиотека Intel MKL не имеет программ-решателей систем ОДУ—ДАУ

Программа-решатель систем ОДУ—ДАУ	ТЕСТЫ					
	ТЕСТ 1	ТЕСТ 1	ТЕСТ 2	ТЕСТ 2	ТЕСТ 3	ТЕСТ 4
	Жесткость MU=1e6	Жесткость MU=1e9	Коэффициенты kt=1, ki=1, ku=0,01	Коэффициенты ku=1, ki=1, kt=1e-104	Жесткость MU=1e6	Параметры реального лазера
SADEL C-Library 2010 Метод M3	+	+	+	+	+	+
Mathcad 2010 Метод Radau	+	-	-	-	+	+
MATLAB 2012 Метод ODE15s	+	-	-	-	-	+
Maple 2008 Метод Rosenbrock	-	-	+	-	-	-
Mathematica 2012 Метод BDF	-	-	+	+	+	+

Знак (-) означает недостоверное решение теста без предупреждения

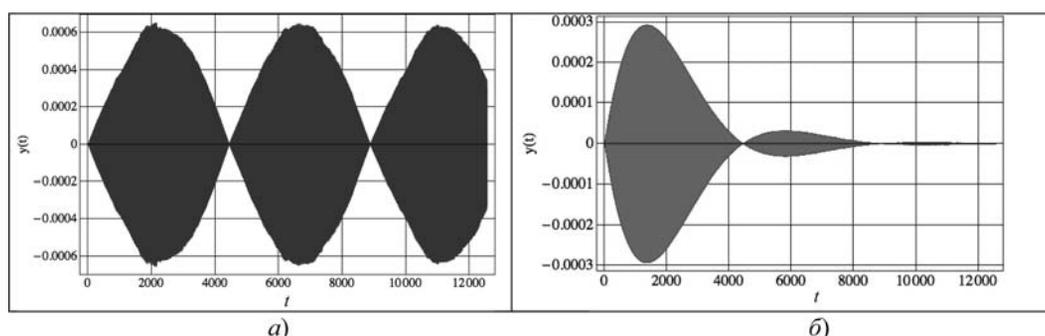


Рис. 5. Решение тестовой задачи ТЕСТ 2 для $eps = 0,001$:

a — метод BDF из библиотеки C-Library NAG; *б* — метод M3 из Си-библиотеки SADEL

нием (ТЕСТ 2) [6] с помощью программы-решателя систем ОДУ—ДАУ неявным методом BDF из библиотеки C-Library NAG (*a*), а также правильное решение этого теста с помощью методов M2 и M3 из библиотеки SADEL (*б*). Предупреждений о возможном не достоверном решении в программе-решателе библиотеки C-Library NAG не было.

Заключение

Разработка программ-решателей систем ОДУ—ДАУ и ЛАУ на языке Си для библиотеки SADEL и разработка ПК ПА10 (SADEL-PA10)) позволяют сделать следующие выводы.

1. Для достоверного и точного решения систем ДАУ при невысоких требованиях к математической точности результатов следует применять только **AL-устойчивые методы интегрирования**, при использовании других методов интегрирования необходимо гарантировать устойчивость численного решения при увеличении шага интегрирования в случае невысоких значений параметра eps .

2. Любые, даже правдоподобные, результаты моделирования динамических систем, полученные с помощью соответствующих программных комплексов, в которых используются не AL-устойчивые

методы интегрирования, следует многократно перепроверять для более высоких требований к точности интегрирования решаемых систем ОДУ—ДАУ, ограничив максимальный размер шага интегрирования.

3. Для получения качественно достоверного и корректного решения жестких и сверхжестких систем ОДУ—ДАУ необходимо на всех шагах численного интегрирования обеспечивать численное решение соответствующих систем ЛАУ с повышенной точностью, т. е. с точностью в 15 верных значащих цифр для всех значений элементов вектора решения систем ЛАУ.

Новые научные результаты предполагается получить в направлениях разработки новых высокоэффективных неявных DABC-методов интегрирования с полностью заполненными матрицами D и A , разработки новых методов и алгоритмов автоматического выбора шага для AL-устойчивых неявных методов интегрирования, новых алгоритмов численно-аналитического почленного вычисления матриц Якоби, а также эффективных алгоритмов и программ решения систем ЛАУ высокой и сверхвысокой размерности с повышенной точностью вычислений.

Работа выполнена при поддержке РФФИ, грант 12-07-00324/12.

Список литературы

1. IEEE Standard for Floating-Point Arithmetic 754-2008, IEEE, 2008 // http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=5223319&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxppls%2Fabs_all.jsp%3Farnumber%3D5223319
2. Евстифеев Ю. А., Маничев В. Б. Эффективный А-устойчивый метод интегрирования обыкновенных дифференциальных уравнений для программ анализа электронных схем // Изв. вузов. Радиоэлектроника. 1986. Т. 29, № 11. С. 31–35.
3. Маничев В. Б., Глазкова В. Н. Методы интегрирования систем ОДУ для адаптируемых программных комплексов анализа РЭС // Радиотехника. 1988. № 4. С. 88–91.
4. Андронов А. В., Жук Д. М., Кожевников Д. Ю., Маничев В. Б. Библиотека математических программ-решателей на языке Си: SADEL. URL: <http://pa10.ru>
5. Жук Д. М., Маничев В. Б., Сахаров М. К. SADEL — библиотека "сверхточных" решателей для программного комплекса ПА10 (SADEL-PA10) // В сб. научных трудов МЭС — 2012. М.: ИППМ РАН, 2012. С. 147–153.
6. Маничев В. Б., Жук Д. М., Сахаров М. К. SADEL — Си библиотека для решения алгебраических и дифференциальных

уравнений с максимально возможной компьютерной точностью // Информационные технологии. 2012. № 10. С. 7–14.

7. Маничев В. Б., Глазкова В. Н., Кожевников Д. Ю., Кирьянов Д. А., Сахаров М. К. Решение систем линейных алгебраических уравнений с удвоенной точностью вычислений на языке Си // Вестник МГТУ. Сер. Приборостроение. 2011. Вып. 4. С. 25–36.

8. Guiyou Mao, Linda R. Petzold. Efficient integration over discontinuities for differential-algebraic systems // Computers & Mathematics with Applications. January 2002. Vol. 43, Iss. 1–2. P. 65–79.

9. Норенков И. П., Трудоношин В. А., Федорук В. Г. Метод формирования математических моделей для адаптируемых программных комплексов анализа радиоэлектронных схем // Радиотехника. 1986. № 9. С. 67–72.

10. Хайрер Э., Ваннер Г. Решение обыкновенных дифференциальных уравнений. Жесткие задачи: пер. с англ. М.: Мир, 2001. 498 с.

11. Жук Д. М., Маничев В. Б., Ильницкий А. О. Методы и алгоритмы решения дифференциально-алгебраических уравнений для моделирования систем и объектов во временной области // Информационные технологии. 2010. № 7. С. 16–24.

УДК 656.073

Е. М. Бронштейн, д-р физ.-мат. наук, проф., проф. кафедры, П. А. Зелёв, аспирант, e-mail: pz1988@ya.ru, Уфимский государственный авиационный технический университет

Задача маршрутизации транспортного средства с учетом зависимости стоимости перевозок от загрузки

Рассматривается задача построения маршрута доставки грузов от одного производителя нескольким потребителям одним транспортным средством (ТС) по круговому маршруту с минимальными расходами. При этом стоимость транспортировки является функцией от загрузки ТС и качества дороги. Построена соответствующая математическая модель, для случая линейной зависимости стоимости проезда от загрузки получена линейная целочисленная модель. Для решения задачи наряду с точным предложен модифицированный эвристический алгоритм Кларка—Райта. Проведен вычислительный эксперимент.

Ключевые слова: маршрутизация, перевозка, эвристический алгоритм, загрузка транспортного средства

Е. М. Bronshtein, P. A. Zeljov

The Vehicle Routing Problem with the Transportation Cost Dependence of Vehicle Loading

The problem of a route creation of delivery of goods from one supplier to several consumers by one vehicle (V) on a circular route with the minimum expenses is considered. Thus the cost of transportation is function from loading of the vehicle and quality of the road. The corresponding mathematical model is constructed, for a case of linear dependence of fare from loading the linear integer model is received. For the solution of a task along with the exact Clark—Right's modified heuristic algorithm is offered. Computing experiment is made.

Keywords: routing, transportation, heuristic algorithm, vehicle loading

Введение

Систематическое изучение оптимизационных задач транспортной логистики (Vehicle Routing Problem, VRP) началось с работы [1]. За последние полвека поставлено множество задач такого типа, развиты как точные, так и эвристические методы

решения. Частным случаем подобных задач является задача коммивояжера.

Классификация оптимизационных задач транспортной логистики приведена, например, в работе [2]. На сайте [3] аккумулируется информация об этой области исследований.

Задача, рассматриваемая в данной работе, переключается с CVRP (Capacitated VRP), в этих задачах учитываются ограничения на вместимость транспортного средства (ТС) (см., например, [4]). В большинстве случаев при постановке задач маршрутизации стоимость транспортировки является функцией, зависящей только от расстояния между городами.

В работе [5] предлагается при постановке задач маршрутизации учитывать ряд дополнительных факторов, важных для практики, некоторые из которых рассматриваются в настоящей статье.

На практике стоимость транспортировки зависит не только от протяженности путей между потребителями. Предприятия, занимающиеся транспортировкой грузов между удаленными потребителями, при расчете стоимости перевозки должны учитывать также и влияние загрузки ТС, состояния дороги и т. п.

В работах [6, 7] предложена классификация задач маршрутизации в зависимости от соотношения величин суммарного отгружаемого груза q и грузоподъемности Q ТС. Случай $Q \geq q$ соответствует традиционной задаче коммивояжера, а для случая $Q < q$ используется термин "задача развозки". Во втором случае груз доставляется более чем за один рейс, т. е. требуется составить сеть маршрутов, исходящих от единственного отправителя. Очевидно, с практической точки зрения задача развозки имеет большую ценность как более отвечающая реальным условиям перевозок.

В работе [8] предлагается подход к решению задачи развозки с пересечением маршрутов, где в качестве целевой функции выбран расход топлива, зависящий в свою очередь от следующих факторов:

- нормативного расхода топлива;
- пробега ТС;
- поправочного коэффициента к норме;
- нормы расхода топлива на пробег ТС в снаряженном состоянии без груза;
- базовой нормы расхода топлива на пробег ТС в снаряженном состоянии;
- нормы расхода топлива на дополнительную массу прицепа или полуприцепа;
- собственной массы прицепа или полуприцепа;
- нормы расхода топлива на транспортную работу;
- объема транспортной работы;
- массы груза;
- пробега ТС с грузом.

Рассматриваемая задача является NP-трудной, в связи с чем точные методы применимы лишь при малых размерностях задач (число пунктов менее 10), и актуальной является разработка эвристических методов ее решения.

Постановка задачи

Груз следует доставить из пункта производства (базы, склада) потребителям по круговому маршруту транспортным средством, грузоподъемность кото-

рого $Q \geq \sum_{i=1}^n q_i$, где q_i — потребность в грузе в i -м пункте потребления; n — число пунктов потребления. Для единообразия будем считать пункт производства нулевым пунктом.

Дороги между пунктами характеризуются двумя показателями: расстоянием l_{ij} ($i, j = 0, 1, 2, \dots, n$) и коэффициентом сложности дороги k_{ij} между пунктами i, j . Расстояние l_{ij} может зависеть от направления движения (пример: есть дороги с односторонним движением), коэффициент сложности k_{ij} дороги влияет на расход топлива, его значение также может зависеть от направления маршрута (пример: подъем в гору или спуск с горы).

В перевозках участвует одно ТС с заданной грузоподъемностью Q и функцией расхода топлива $f(q)$ на единицу пути дороги стандартного качества.

Функция расхода топлива в зависимости от массы перевозимого груза является возрастающей и, как правило, вогнутой на всем интервале области определения функции $[0; Q]$ (см. рисунок).

Введем булевы переменные X_{ij} ($i, j = 0, \dots, n$), равные 1 тогда и только тогда, когда следующим после i -го пункта на пути следования ТС является j -й пункт, $i, j = 0, \dots, n$. Необходимо найти булевы переменные X_{ij} такие, что:

$$\sum_{i=0}^n X_{ij} = 1, j = \overline{0, n}; \quad (1)$$

$$\sum_{j=0}^n X_{ij} = 1, i = \overline{0, n}; \quad (2)$$

$$X_{00} = 0. \quad (3)$$

Ограничение (1) показывает, что в каждый пункт доставки ТС прибывает в точности один раз; ограничение (2) показывает, что из каждого пункта доставки ТС выезжает один раз. Ограничение (3) означает, что из нулевого пункта (базы, склада) ТС выезжает

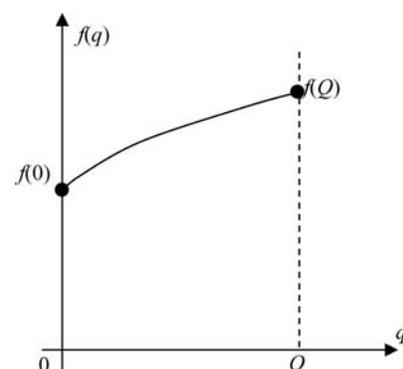


График функции стоимости транспортировки груза в зависимости от его массы

в какой-нибудь другой. Для других пунктов потребления данное условие будет следовать из последующих ограничений. Введем целочисленные переменные v_i ($i = 1, \dots, n$) такие, что:

$$1 \leq v_i \leq n; \quad (4)$$

$$(v_i - v_j) + nX_{ij} \leq n - 1. \quad (5)$$

Переменные v_i имеют смысл номеров потребителей в порядке прохождения в цепи, содержащей все пункты, за исключением базы. Из условий (4) и (5) следует отсутствие подциклов.

Введем булевы переменные Z_{is} , равные 1 тогда и только тогда, когда $v_s > v_i$ ($i, s = 1, \dots, n$).

Справедливы ограничения:

$$\sum_{i=0}^n \sum_{s=0}^n Z_{is} = n(n+1)/2; \quad (6)$$

$$nZ_{is} \geq v_s - v_i. \quad (7)$$

Условие (7) обеспечивает выполнение равенства $Z_{is} = 1$ при $v_s > v_i$. Заметим, что для обеспечения равенства $Z_{is} = 0$ при $v_s \leq v_i$ по смыслу величин v_s число величин v_p , меньших v_s , равно $v_s - 1$. Таким образом, нужное свойство обеспечивается равенством

$$\sum_{i=1}^n Z_{is} = v_s - 1. \quad (8)$$

Стоимость доставки грузов

$$R(X) = \sum_{i=0}^n \sum_{j=0}^n X_{ij} f \left(\sum_{s=0}^n Z_{is} q_s \right) l_{ij} k_{ij}. \quad (9)$$

Необходимо минимизировать расходы на транспортировку:

$$R(X) \rightarrow \min. \quad (10)$$

Линейная целочисленная модель

Примем, что стоимость транспортировки является линейной функцией от массы перевозимого груза: $f(q) = aq + b$, где a и b — коэффициенты, определяемые эмпирически. Тогда целевая функция примет вид

$$R(X) = a \sum_{s=0}^n \sum_{i=0}^n \sum_{j=0}^n X_{ij} Z_{is} q_s l_{ij} k_{ij} + b \sum_{i=0}^n \sum_{j=0}^n X_{ij} l_{ij} k_{ij}. \quad (11)$$

Введем булевы переменные T_{ijs} ($i, j, s = 0, \dots, n$) следующим образом:

$$T_{ijs} = X_{ij} Z_{is}. \quad (12)$$

Справедливы ограничения:

$$T_{ijs} \geq X_{ij} + Z_{is} - 1, \quad T_{ijs} \leq X_{ij}, \quad T_{ijs} \leq Z_{is}. \quad (13)$$

Целевая функция примет вид

$$R(X) = a \sum_{s=0}^n \sum_{i=0}^n \sum_{j=0}^n T_{ijs} q_s l_{ij} k_{ij} + b \sum_{i=0}^n \sum_{j=0}^n X_{ij} l_{ij} k_{ij}. \quad (14)$$

Таким образом, переменные поставленной задачи: X_{ij} — булевы, их число $(n+1)^2$; Z_{is} — булевы, их число $(n+1)^2$; T_{ijs} — булевы, их число $(n+1)^3$; v_i — целочисленные, их n . Общее число ограничений имеет порядок $O(n^2)$.

Модифицированный алгоритм Кларка—Райта

Для решения задачи (1)–(14) был модифицирован эвристический алгоритм Кларка—Райта [9, 10]. Достоинствами метода являются его простота, надежность и гибкость.

Опишем основные идеи алгоритма.

Метод Кларка—Райта относится к числу приближенных итерационных методов, причем на каждой итерации осуществляется попытка сращивания двух маршрутов по определенным правилам. Предлагаемая модификация метода Кларка—Райта заключается в возможности варьирования глубины сращивания маршрутов p , а также в учете зависимости стоимости перевозки от направления прохождения цикла. Кроме того, в модификации метода Кларка—Райта стоимость грузоперевозки является функцией от загрузки ТС. В классическом алгоритме Кларка—Райта используется глубина сращивания $p = 0$.

Рассмотрим на небольшом примере идею сращивания циклов. Изначально генерируются n циклов (маршрутов) вида $0 - i - 0$ ($i = 1, 2, \dots, n$). На каждом из последующих шагов рассматриваются всевозможные варианты сращивания двух циклов, из которых выбирается вариант сращивания, удовлетворяющий условиям грузоподъемности ТС, а также имеющий минимальную из стоимостей перевозок.

Допустим, на некотором шаге имеется маршрут α :

$$0 - A_1 - A_2 - A_3 - A_4 - A_5 - A_6 - 0,$$

а также маршрут β :

$$0 - B - 0,$$

где $A_1, A_2, A_3, A_4, A_5, A_6$ — посещаемые в маршрутах пункты α (города), B — некоторая последовательность пунктов, посещаемых в маршруте β . Согласно требованию задачи о единственности посещения, маршруты α и β не могут содержать одинаковых пунктов. Обозначим $\bar{\alpha}$ следующий маршрут:

$$0 - A_6 - A_5 - A_4 - A_3 - A_2 - A_1 - 0,$$

полученный инвертированием направления маршрута α . При сращивании маршрутов последова-

тельность пунктов одного маршрута целиком вставляется между двумя соседними пунктами другого маршрута, при этом рассматриваются прямой и инвертированный варианты обоих маршрутов. Тогда вариантами сращивания маршрутов α и β на очередном шаге алгоритма будут следующие:

$$\begin{aligned}
 p = 0: \\
 & 0 - \beta - A_1 - A_2 - A_3 - A_4 - A_5 - A_6 - 0, \\
 & 0 - \bar{\beta} - A_1 - A_2 - A_3 - A_4 - A_5 - A_6 - 0, \\
 & 0 - A_1 - A_2 - A_3 - A_4 - A_5 - A_6 - \beta - 0, \\
 & 0 - A_1 - A_2 - A_3 - A_4 - A_5 - A_6 - \bar{\beta} - 0, \\
 & 0 - \beta - A_6 - A_5 - A_4 - A_3 - A_2 - A_1 - 0, \\
 & 0 - \bar{\beta} - A_6 - A_5 - A_4 - A_3 - A_2 - A_1 - 0, \\
 & 0 - A_6 - A_5 - A_4 - A_3 - A_2 - A_1 - \beta - 0, \\
 & 0 - A_6 - A_5 - A_4 - A_3 - A_2 - A_1 - \bar{\beta} - 0;
 \end{aligned}$$

$$\begin{aligned}
 p = 1: \\
 & 0 - A_1 - \beta - A_2 - A_3 - A_4 - A_5 - A_6 - 0, \\
 & 0 - A_1 - \bar{\beta} - A_2 - A_3 - A_4 - A_5 - A_6 - 0, \\
 & 0 - A_1 - A_2 - A_3 - A_4 - A_5 - \beta - A_6 - 0, \\
 & 0 - A_1 - A_2 - A_3 - A_4 - A_5 - \bar{\beta} - A_6 - 0, \\
 & 0 - A_6 - \beta - A_5 - A_4 - A_3 - A_2 - A_1 - 0, \\
 & 0 - A_6 - \bar{\beta} - A_5 - A_4 - A_3 - A_2 - A_1 - 0, \\
 & 0 - A_6 - A_5 - A_4 - A_3 - A_2 - \beta - A_1 - 0, \\
 & 0 - A_6 - A_5 - A_4 - A_3 - A_2 - \bar{\beta} - A_1 - 0;
 \end{aligned}$$

$$\begin{aligned}
 p = 2: \\
 & 0 - A_1 - A_2 - \beta - A_3 - A_4 - A_5 - A_6 - 0, \\
 & 0 - A_1 - A_2 - \bar{\beta} - A_3 - A_4 - A_5 - A_6 - 0, \\
 & 0 - A_1 - A_2 - A_3 - A_4 - \beta - A_5 - A_6 - 0, \\
 & 0 - A_1 - A_2 - A_3 - A_4 - \bar{\beta} - A_5 - A_6 - 0, \\
 & 0 - A_6 - A_5 - \beta - A_4 - A_3 - A_2 - A_1 - 0, \\
 & 0 - A_6 - A_5 - \bar{\beta} - A_4 - A_3 - A_2 - A_1 - 0, \\
 & 0 - A_6 - A_5 - A_4 - A_3 - \beta - A_2 - A_1 - 0, \\
 & 0 - A_6 - A_5 - A_4 - A_3 - \bar{\beta} - A_2 - A_1 - 0;
 \end{aligned}$$

$$\begin{aligned}
 p = 3: \\
 & 0 - A_1 - A_2 - A_3 - \beta - A_4 - A_5 - A_6 - 0, \\
 & 0 - A_1 - A_2 - A_3 - \bar{\beta} - A_4 - A_5 - A_6 - 0.
 \end{aligned}$$

В данном случае при $p = 3$ всего два варианта сращивания маршрутов α и β .

Говоря обобщенно, на каждом из последующих шагов рассматриваются всевозможные пары циклов, для каждой из которых при $p = 0$ строится не более восьми циклов (при $p \geq 1$ число циклов равно $16p$) следующим образом: каждый из новых циклов состоит из начального или конечного отрезка (длиной не более p) одного из циклов, затем проходится второй цикл (без нулевой вершины), затем продол-

жается движение по первому циклу. При этом учитывается возможность изменения направления движения по каждому из циклов на противоположное, поскольку задача не предполагается симметричной и изменение загрузки ТС зависит от порядка прохождения пунктов.

Например, сращивание циклов $a_1^1, a_2^1, \dots, a_{k_1}^1$ и $a_1^2, a_2^2, \dots, a_{k_2}^2$ (с исключенной нулевой вершиной) при единичных начальном и конечном отрезках дает следующие циклы:

$$\begin{aligned}
 & 0, a_1^1, a_1^2, a_2^2, \dots, a_{k_2}^2, a_2^1, \dots, a_{k_1}^1, 0; \\
 & 0, a_1^1, a_{k_2}^2, a_{k_2-1}^2, \dots, a_1^2, a_2^1, \dots, a_{k_1}^1, 0; \\
 & 0, a_{k_1}^1, a_{k_1-1}^1, \dots, a_2^1, a_1^2, a_2^2, \dots, a_{k_2}^2, a_1^1, 0; \\
 & 0, a_{k_1}^1, a_{k_1-1}^1, \dots, a_2^1, a_{k_2}^2, a_{k_2-1}^2, \dots, a_1^2, a_1^1, 0.
 \end{aligned}$$

Еще четыре цикла возникнут при добавлении части второго цикла между пунктами $a_{k_1}^1$ и $a_{k_1-1}^1$. Затем первый и второй циклы можно поменять местами, тогда получим еще восемь циклов.

В каждом случае вычисляется разность между суммой расходов на доставку грузов по отдельным циклам и расходов на доставку грузов по их объединению и выбирается объединенный цикл, для которого эта разность максимальная. При этом в модификации алгоритма Кларка—Райта для каждой пары циклов перебираются все возможные варианты их сращивания при всех значениях глубины сращивания p от 0 до заданного пользователем.

Вычислительный эксперимент

Задача (1)–(7), (9), (10) решалась двумя способами: точным методом и модифицированным алгоритмом Кларка—Райта, которые были реализованы в среде Scilab-4.1.2. Поскольку на задачах с размерностями $n \geq 8$ время точного решения линейной целочисленной задачи оказалось весьма продолжительным (в некоторых случаях более 45 ч), сравнение алгоритмов проводилось для случайно сгенерированных задач при размерностях $n = \overline{6, 9}$. Для каждой размерности было решено по 20 задач каждым из алгоритмов.

Генерация тестовых примеров осуществлялась для значений грузоподъемности ТС $Q \in [1, 5 \dots 5]$ т; элементы матрицы L протяженности дорог между пунктами генерировались из диапазона $L_{ij} \in [1 \dots 100]$, $i, j = \overline{1, n}$ км; элементы матрицы коэффициентов сложности дорог между пунктами генерировались из диапазона $k_{ij} \in [1 \dots 5]$, $i, j = \overline{1, n}$; вектор потреб-

Результаты численного эксперимента

n	Точный алгоритм	Модифицированный алгоритм Кларка—Райта	
	Среднее время работы алгоритма, с	Среднее время работы алгоритма, с	Среднее отклонение стоимости транспортировки, %
6	6,91	12,99	4,01
7	65,26	30,43	18,86
8	538,93	46,32	19,36
9	5495,81	78,11	26,74

ностей пунктов в грузе q содержит компоненты из диапазона $q_i \in [300...1500]$ кг, при этом учитывалось условие вместимости всех грузов в одно ТС, поскольку рассматривается задача построения одного кругового маршрута. Функция расхода топлива задана в виде $f(q) = 15 + 3q$.

Все сгенерированные задачи были решены с помощью точного алгоритма, а также с помощью модифицированного алгоритма Кларка—Райта при значении параметра $p = 3$. Для каждого значения $n = \overline{6, 9}$ в таблице приведены результаты численного эксперимента — средние значения стоимости перевозок, отклонения от оптимального значения стоимости, вычислительного времени для точного алгоритма и модифицированного алгоритма Кларка—Райта.

Как видно из таблицы, алгоритм Кларка—Райта на небольших размерностях показывает хорошие результаты. Вычислительное время, затрачиваемое точным алгоритмом, оказывается неприемлемо велико (более 90 мин уже при $n = 9$), в то время как алгоритм Кларка—Райта при увеличении размерности требует несущественно больше времени на проведение расчетов.

Заключение

Рассматриваемая проблема является актуальной прикладной задачей. Полученные результаты позволяют сделать вывод о целесообразности разработки модификации алгоритма Кларка—Райта или адаптации другого эвристического метода для решения задачи транспортировки грузов с учетом загрузки транспортных средств для снижения отклонения получаемых результатов от точно найденного оптимума.

Работа выполнена при поддержке РФФИ (проект 13-01-00005).

Список литературы

1. **Dantzig G. B., Ramser J. H.** The Truck Dispatching Problem // Management Science. 1959. V. 6 (1). P. 80—91.
2. **Бронштейн Е. М., Зайко Т. А.** Детерминированные оптимизационные задачи транспортной логистики // Автоматика и телемеханика, 2010. № 10. С. 133—147.
3. **Ralphs T. K., Kopman L., Pulleyblank W. R., and Trotter L. E. Jr.** On the Capacitated Vehicle Routing Problem // Math. Program. Ser. B 94. 2003. P. 343—359.
4. **Kara I., Kara B. Y., Yetis M. K.** Energy Minimizing Vehicle Routing Problem // Proc. Combinatorial Optimization and Applications, First International Conference, COCOA 2007, Xi'an, China, August 14—16, 2007. V. 4616.
5. **Clarke G. and Right J. W.** Scheduling of Vehicles from a Central Depot to a Number of Delivery Points // Operations Research. 1963. V. 11. P. 568—581.
6. **Житков В. А., Ким К. В.** Методы оперативного планирования грузовых автомобильных перевозок. М.: Транспорт, 1984. 218 с.
7. **Житков В. А.** Планирование автомобильных перевозок грузов мелкими партиями. М.: Транспорт, 1976. 112 с.
8. **Буйлова М. В.** Маршрутизация в задаче развозки с пересечением маршрутов. URL: <http://www.sworld.com.ua/konfer26/799.pdf>
9. **Хаггарти Р.** Дискретная математика для программистов. М.: Техносфера, 2005. 400 с.
10. **Программа TopLogistic Prof** — профессиональная автоматизация транспортной логистики. URL: <http://www.toplogistic.ru/313.html>
11. **Андерсон Дж.** Дискретная математика и комбинаторика: Пер. с англ. М.: Издат. дом "Вильямс", 2003.
12. **Григорьев М. Н., Уваров С. А.** Логистика. Краткий курс лекций. М.: Юрайт, 2012. 207 с.
13. **Бауэрсокс Д. Дж., Клосс Д. Дж.** Логистика: интегрированная цепь поставок. М.: ЗАО "Олимп-бизнес". 2010. 640 с.
14. **Зак Ю. А.** Прикладные задачи теории расписаний и маршрутизации перевозок. М.: Либроком, 2012. 394 с.

УДК 528.854.2

А. А. Афанасьев, аспирант, e-mail: afanasyevaa@tpu.ru,

А. В. Замятин, канд. техн. наук, проректор-директор, e-mail: zamyatin@tpu.ru,

Институт кибернетики Национального исследовательского Томского политехнического университета

Оценка применимости подходов к идентификации изменений ландшафтного покрова по данным дистанционного зондирования Земли

Рассматривается актуальная проблема выбора оптимального подхода к задаче идентификации изменений ландшафтного покрова по данным дистанционного зондирования Земли. Описана стандартная процедура идентификации изменений ландшафтного покрова и наиболее популярные подходы к анализу изменений. Приведены результаты численных экспериментов по оценке качества идентификации изменений с использованием данных с различной ландшафтно-классовой структурой и разной степенью изменений.

Ключевые слова: дистанционное зондирование Земли, идентификация изменений

A. A. Afanasyev, A. V. Zamyatin

The Applicability Analysis of the Approaches to the Identification of Land Cover Changes by Remote Sensing Data

The article considering actual problem of choosing the optimal approach to the land cover changes identification using the remote sensing data. The standard procedure of land cover changes identification and the most popular approaches to the changes analysis are described. The results of numerical experiments to assess the quality of the changes identification using different landscape-class structure data and varying degrees of change is adduced.

Keywords: remote sensing, change detection

Введение

Идентификация изменений ландшафтного покрова по аэрокосмическим изображениям (АИ) находит применение в таких задачах, как мониторинг экологической обстановки, лесопользования, сельскохозяйственных угодий, динамики урбанизации [1]. Крайне полезной идентификация изменений может быть при мониторинге протекания чрезвычайных ситуаций, а также раннем (превентивном) обнаружении причин их возникновения. При этом поиск и выбор подходящих в каждом конкретном случае средств и методов идентификации изменений не является тривиальным. Эта задача осложняется широким набором потенциально доступных средств идентификации изменений, основанных на различном математическом аппарате, в условиях отсутствия общепринятых формализованных процедур такого поиска и выбора. Сегодня она решается эмпирически без значительной проработки и обоснования [2, 3].

Данная работа посвящена анализу распространенных методов идентификации изменений ландшафтного покрова по данным дистанционного зондирования Земли (ДЗЗ) для определения области их практической применимости и в целях повышения эффективности решения задачи идентификации изменений.

Процедура идентификации изменений

Исходными данными для типовой задачи идентификации изменений являются два разновременных АИ одного и того же фрагмента ландшафта, представленных в виде трехмерных матриц $\mathbf{I}_1 = \{i_{1_{xyz}}\}$, $x = 1 \dots H$, $y = 1 \dots W$, $z = 1 \dots M$ и $\mathbf{I}_2 = \{i_{2_{xyz}}\}$, $x = 1 \dots H$, $y = 1 \dots W$, $z = 1 \dots M$, где H и W — число элементов в строках и столбцах исходных АИ, а M — число диапазонов/каналов изображения, причем при $M = 1$ АИ относят к панхроматическим, при $M > 1$ к мульти- и гиперспектральным.

Процесс идентификации изменений в общем случае, как правило, проходит в четыре этапа.

1. Формирование I_1 и I_2 с помощью предварительной обработки исходных АИ (геометрической, радиометрической, атмосферной и топографической (если анализируется горная местность) коррекций) [4].

2. Формирование разностного изображения D применением одного из методов разностного анализа. В результате его применения получаем матрицу $D = \{d_{xyz}\}$, в каждой ячейке которой находится значение, показывающее степень изменения. В случае $M > 1$ потребуются дополнительная операция объединения каналов, которую можно проводить, например, по принципу максимального значения по каналам в каждой точке [5]. В результате получаем двумерную матрицу разности $D' = \{d_{xy}\}$.

3. Формирование итоговой матрицы изменений $V = \{b_{xy}\}$, где $b_{xy} = 1$ означает наличие значимого изменения (в смысле некоторого порогового значения или функции) в данной точке, а $b_{xy} = 0$ — отсутствие такового. Матрица изменений формируется в результате применения пороговой функции, где $b_{xy} = 1$, если $|D_{xy}| > \tau$. Значение порога τ определяется либо вручную, либо с помощью специальных дополнительных операций.

Некоторые искажения при обнаружении значимых изменений возникают вследствие влияния на результат разности качества исходных данных и условий съемки, таких как состояние атмосферы, азимут и высота солнца над горизонтом, влажность почвы и других факторов [6, 7]. Частично эту проблему можно решить предварительной обработкой (этап 1). Учитывая указанные выше сложности, а также мульти- или гиперспектральность АИ, при которых изменения ландшафтного покрова отражаются неодинаково на снимках в разных спектральных диапазонах, разработаны несколько основных подходов (алгебраический, трансформационный, классификационный) к идентификации изменений ландшафтного покрова по данным ДЗЗ.

Наиболее простой подход — алгебраический. Он либо не предусматривает модификации данных, либо предполагает их простую модификацию. К алгебраическим методам относят методы *разности*, *отношения*, *регрессии*, *векторного анализа*, *вычитания фона*, а также *вегетационных индексов* [3]. Гораздо более сложный подход — трансформационный [3]. В этом случае исходные данные подвергают достаточно сложным математическим преобразованиям, а лишь затем сопоставляют. Наиболее часто применяют преобразования Карунена—Лозва (*метод главных компонент*), Кауга—Томаса (метод "Колпачка с кисточкой"), Грэмма—Шмидта, а также статистику "хи-квадрат" [8, 9]. Целью преобразования Кархунена—Лозва является нахождение новой системы координат для матриц I_1 и I_2 по отдельности с ортогональными осями, отсортированными

ми в порядке убывания дисперсии вдоль осей. В этом случае матрицы I_1 и I_2 рассматривают как таблицу точек в M -мерном пространстве [10]. Результаты преобразований Кауга—Томаса и Грэмма—Шмидта похожи, но в первом случае матрицу преобразования вычисляют особым образом индивидуально для каждого спутникового съемочного аппарата, а во втором случае алгоритм преобразовывает множество векторов исходных данных к множеству ортонормированных векторов [11]. Статистику "хи-квадрат" используют для применения критерия Пирсона [13].

Классификационный подход предполагает применение к исходным АИ различных алгоритмов контролируемой или неконтролируемой классификации и затем сопоставление результатов классификации АИ.

Как отмечено выше, не существует строго определенного набора методов разностного анализа, но есть ряд методов, более распространенных на практике [3,12—15]. В нашем исследовании используются восемь методов разностного анализа, условно названных *разность* (*Image Difference, ID*), *отношение* (*Image Rationing, IR*), *векторный анализ* (*Change Vector Analysis, CVA*), *неконтролируемая кластеризация* (*Unsupervised Clustering, UC*), *главных компонент* (*Principal Component Analysis, PCA*), *хи-квадрат* (*Chi-Square, CS*), *линейная регрессия* (*Linear Regression, LR*), *квадратичная регрессия* (*Square Regression, SR*) ввиду отсутствия устоявшейся русскоязычной терминологии для этих методов.

Методы идентификации изменений

Метод *разность* применяется чаще других ввиду своей простоты в реализации и высокой производительности [3, 4, 14]. Каждая ячейка D содержит абсолютное значение разности соответствующих ячеек I_1 и I_2 , т. е. $d_{xyz} = |i_{1_{xyz}} - i_{2_{xyz}}|$. Соответственно, чем меньше значение конкретного элемента матрицы, тем меньше вероятность того, что в соответствующей области произошло значимое изменение.

Метод *отношение* также сравнительно широко применяемый отличается от предыдущего тем, что вместо разности выполняется деление: $d_{xyz} = |i_{1_{xyz}} / i_{2_{xyz}}|$. В этом случае $d_{xyz} \in (0, +\infty)$, и чем ближе значение d_{xyz} к 1, тем меньше вероятность того, что в соответствующей ее области произошло значимое изменение.

В методе *векторный анализ* рассматриваются значения ячеек матриц I_1 и I_2 с фиксированными координатами x и y как координаты векторов $V_{1_{xy}} = \{i_{1_{xy1}}, i_{1_{xy2}}, \dots, i_{1_{xyM}}\}$ и $V_{2_{xy}} = \{i_{2_{xy1}}, i_{2_{xy2}}, \dots, i_{2_{xyM}}\}$. Таким образом, значения ячеек матрицы разности D можно получать как евклидово расстояние между соответствующими векторами. Достоинство данного метода в том, что он позволяет сразу по-

лучать двумерную матрицу \mathbf{D} без дополнительного слияния каналов. Также этот метод иногда применяют для классификации типов изменений, так как в этом случае кроме значения изменений возможно нахождение M -мерного вектора изменения [3,16,17].

В методе **главных компонент** используется статистическое преобразование Карунена—Лоэва для уменьшения размерности [10]. В результате получаем преобразованные матрицы \mathbf{I}'_1 и \mathbf{I}'_2 . Чаще всего после преобразования оставляют только один или два первых слоя [3]. Данный метод также относится к наиболее эффективным и часто используемым, несмотря на низкую вычислительную эффективность.

В методе **хи-квадрат** используется критерий согласия Пирсона для проверки отклонения значений второго изображения от значений первого. Рассчитывается статистика $\chi^2 = \sum_{z=1}^M (i_{1_{xyz}} - i_{2_{xyz}})^2 / i_{2_{xyz}}$. Далее

это значение сравнивается с табличным значением, выбирающимся в зависимости от значения M .

В методах с использованием регрессии (**линейная регрессия** и **квадратичная регрессия**) разность определяется как разница между значениями ячеек \mathbf{I}_2 и вычисленного значения на основе регрессионной функции (значения $i_{1_{xyz}}$ рассматриваются как аргументы, а соответствующие им $i_{2_{xyz}}$ как значения функции) [13]. Сначала каким либо методом (например методом **наименьших квадратов**) подбирается вид регрессионной функции $f(x)$ и ее коэффициенты. Далее, как описано выше, применяем формулу $d_{xyz} = |f(i_{1_{xyz}}) - i_{2_{xyz}}|$.

Метод **неконтролируемой кластеризации** предусматривает предварительную кластеризацию изображений \mathbf{I}_1 и \mathbf{I}_2 с получением \mathbf{I}_1^c и \mathbf{I}_2^c , а затем нахождение разности $d_{xyz} = |i_{1_{xyz}}^c - i_{2_{xyz}}^c|$. Сложность этого метода заключается в подборе вида и параметров алгоритма кластеризации. В нашей работе применяется широко используемый на практике алгоритм k -средних.

Экспериментальные исследования применимости методов идентификации изменений

Постановка задачи. Проведенные эксперименты должны выявить, какие методы идентификации изменений эффективнее использовать для различных типов ландшафтного покрова и степени изменения. Для этого используем набор АИ с различной ландшафтно-классовой структурой и преднамеренно внесенными в каждое АИ изменениями известной степени. В качестве тестовых данных будем использовать АИ с различной ландшафтно-классовой структурой. Источником данных был выбран космический аппарат Landsat ETM+ с пространственным разрешением 30 м. Размер исходных данных

1700 × 1666 пикселей, число каналов — 7. Для оценки ландшафтно-классовой структуры исходных АИ используем такие статистические параметры, как среднее значение ячеек матрицы $\mathbf{I}_1 (m_x)$, стандартное отклонение (D_x), а также среднюю межканальную корреляцию (K). Формирование тестовых исходных пар данных позволит точно оценить качество выполненного разностного анализа, так как эталонная маска изменений будет известна априори. Применим к полученным данным рассматриваемые методы идентификации изменений и оценим их точность с помощью широко применяемого параметра AUC (*area under ROC curve*) — площадь под ROC-кривой (*receiver operating characteristic* — операционная характеристика приемника) [18, 19]. Данный параметр помогает оценить принципиальную возможность эффективной бинарной классификации конкретного метода, поскольку при его расчете рассматриваются все возможные значения порога.

Для формирования экспериментального массива данных разновременных тестовых АИ с известной степенью изменений и различной ландшафтно-классовой структурой предложена следующая методика. На каждое исходное АИ искусственно вносят изменения. Сначала оно сливается со своей смещенной копией $i'_{1_{xyz}} = i_{1_{(x+n1)(y+n2)}}$, где $n1 \ll H$ и $n2 \ll W$ — случайные целые числа, по мажоритарному принципу: $i''_{1_{xyz}} = \max(i_{1_{xyz}}, i'_{1_{xyz}})$, чем имитируется поступательное изменение границ объектов. Затем случайным образом добавляется некоторое количество градиентных пятен: $i'''_{1_{xyz}} = i''_{1_{xyz}} + s_{xyz}$, где $\mathbf{S} = \{s_{xyz}\}$ — матрица, содержащая градиентные пятна, для обеспечения различных по интенсивности изменений. Для одного исходного АИ \mathbf{I}_1 формируется несколько АИ \mathbf{I}_2 с различной степенью изменения ($\mathbf{I}'''_{1\alpha_1}, \mathbf{I}'''_{1\alpha_2}, \dots, \mathbf{I}'''_{1\alpha_n}$, где α — доля измененных пикселей). К ним применяется один из методов идентификации изменений для получения эталонных масок изменения \mathbf{V}_α^e . Далее на измененные АИ накладывается градиент изменения интенсивности для имитации разных условий освещенности ($i_{2_{xyz}} = i'''_{1_{xyz}} + g_{xyz}$, $\mathbf{G} = \{g_{xyz}\}$ — матрица, содержащая градиент). Для каждого эксперимента входящей парой изображений будут исходная АИ и каждая из сформированных \mathbf{I}_1 и $\mathbf{I}_{2\alpha_i}$, для которых находим разностные изображения $\mathbf{D}_{\alpha_i}^m$ каждым из алгоритмов идентификации изменений.

Методика проведения экспериментов

Ко всем полученным парам данных применяли все указанные выше методы идентификации изменений. Входными данными для анализа качества по параметру AUC будут $\mathbf{D}_{\alpha_i}^m$ и $\mathbf{V}_{\alpha_i}^e$. Также для каждого эксперимента проведем замеры времени выполнения метода. Для установления степени влияния параметра оценки ландшафтно-классовой структуры используем ранговый метод.

Результаты экспериментов

Как видно из приведенных результатов экспериментов, в большинстве случаев наиболее качественный результат дают методы *разность* (ID) и *векторный анализ* (CVA). Результаты остальных методов отличаются незначительно (рис. 1).

При различных степенях изменений α методы проявляют себя по-разному (рис. 2). Метод *неконтролируемая кластеризация* (UC) показывает лучший результат при $15\% < \alpha < 30\%$, а также мало уступает методу *разность* (ID) при значении $\alpha < 15\%$. Однако он показывает худший результат при $\alpha > 45\%$. Также следует отметить, что качество методов *отношение* (IR) и *хи-квадрат* (CS) заметно возрастает при увеличении доли изменений α .

С увеличением m_x возрастает качество методов *отношение* (IR) и *хи-квадрат* (CS), а также падает качество метода *неконтролируемая кластеризация* (UC).

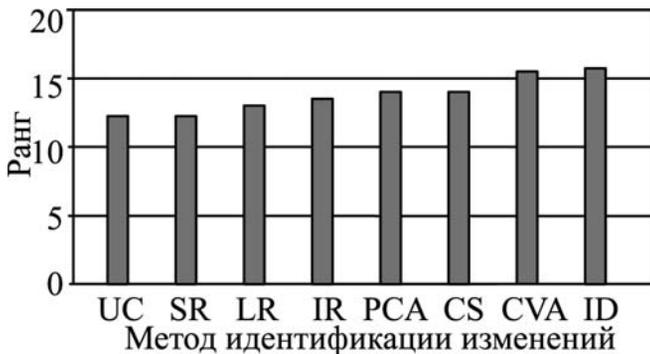


Рис. 1. Общий рейтинг методов идентификации изменений

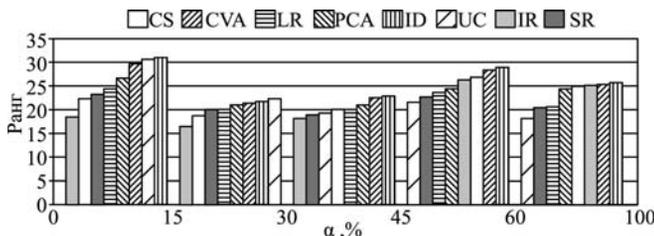


Рис. 2. Эффективность методов в зависимости от степени произошедших изменений α

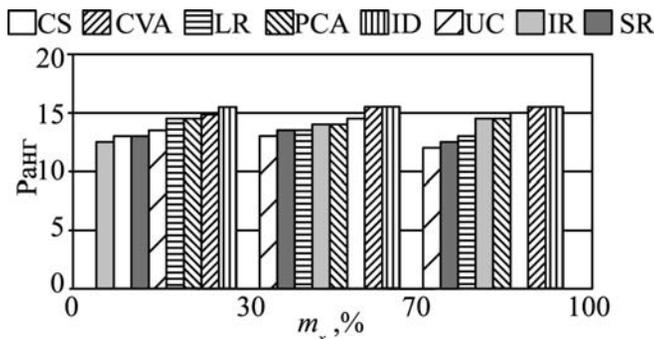


Рис. 3. Эффективность методов в зависимости от значения m_x

Другие методы при этом не показывают значительных изменений (рис. 3).

При малом D_x метод *отношение* (IR) показывает результат, очень близкий к *разности* (ID), но при возрастании этого параметра качество данного метода сильно падает. В то же время метод *главных компонент* (PCA) показывает положительную динамику (рис. 4).

При возрастании межканальной корреляции сильно возрастает качество метода *главных компонент* (PCA). Другие методы мало зависят от этого параметра (рис. 5).

Эксперименты по измерению времени работы алгоритмов разностного анализа (рис. 6) выполнены на ПЭВМ с процессором Intel Pentium IV 3 ГГц и объемом оперативной памяти 4 Гбайт под управлением операционной системы Windows 7 (пакет обновления 3). Типичная диаграмма распределе-

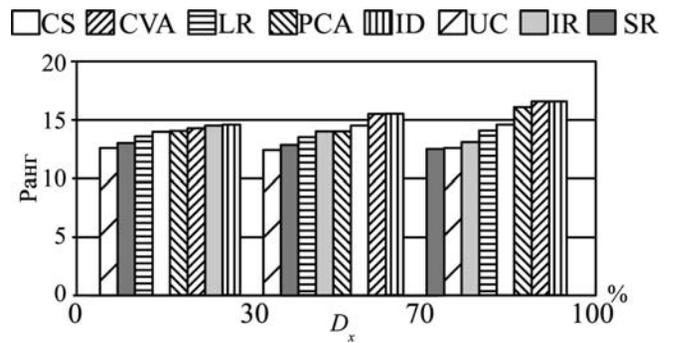


Рис. 4. Рейтинг качества методов в группах по стандартному отклонению

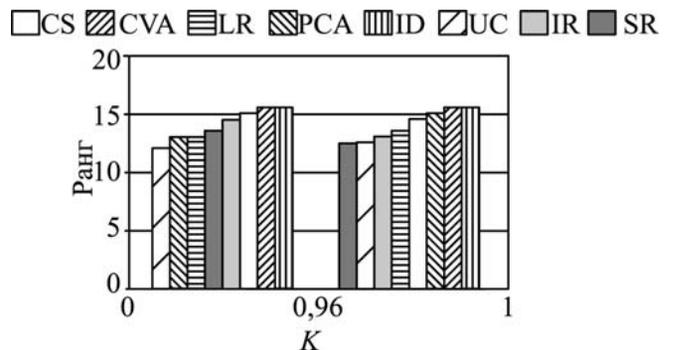


Рис. 5. Рейтинг качества методов в группах по средней межканальной корреляции

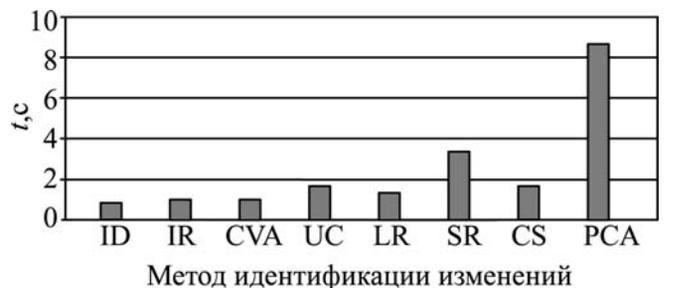


Рис. 6. Время работы алгоритмов разностного анализа

ния времени работы алгоритмов разностного анализа (рис. 6) дает наглядное представление об их вычислительной сложности. Несмотря на высокие временные затраты некоторых методов (в особенности метода *главных компонент* (РСА) их применение оправдано в связи с устойчивостью результатов и их сравнительно высоким качеством.

Заключение

Идентификация изменений ландшафтного покрова в большом массиве данных требует формализованного подхода к выбору методов разностного анализа для обеспечения качества идентификации изменений. Однако такой выбор не является тривиальной задачей и до сих пор ведутся исследования по поиску критериев его качества. По приведенным результатам исследований найдена определенная взаимосвязь между эффективностью методов идентификации изменений и статистическими параметрами исходных данных (математическое ожидание, дисперсия, межканальная корреляция, степень изменения между разновременными данными), позволяющая в некоторой степени упростить задачу поиска подходящего метода (методов) при решении практических задач.

В общем случае, без возможности проведения глубокого предварительного анализа исходных данных, наиболее целесообразно применение методов *разность* и *векторный анализ* с устойчивыми и качественными результатами. Метод *неконтролируемой кластеризации* применим при малой степени изменений ($\alpha < 30\%$) и малой дисперсии разновременных данных. Метод *главных компонент* отличаются высокие результаты при большой дисперсии и значительной корреляции между каналами АИ. Метод *отношения* применим при высокой степени изменений ($\alpha > 45\%$) разновременных данных, высоким математическом ожидании и незначительной дисперсии данных. Метод *хи-квадрат* может быть применим в случае высокой степени изменений ($\alpha < 45\%$), среднего или высокого математического ожидания и низкой дисперсии. Широкое применение на практике методов с использованием *регрессии* показало неплохие результаты лишь в некоторых случаях с низкой степенью изменений ($\alpha < 15\%$) разновременных данных, а также для данных с небольшим математическим ожиданием и дисперсией, что осложняет их непосредственное использование на практике. Оценка вычислительной эффективности реализованных методов показала целесообразность применимости всех методов вследствие не критичных показателей по этому параметру.

Предложенный подход предполагает предварительный анализ данных и применение наиболее подходящего метода разностного анализа. Особенно подход оправдывает себя при единовременном анализе большого количества разновременных АИ, так как в этом случае экспертный выбор подходящего

метода может быть крайне трудоемким. Помимо этого возможен "каскадный" анализ данных, при котором большая сцена АИ с выявленными изменениями редуцируется и анализируется другим, более чувствительным методом идентификации изменений, при улучшенном качестве результатов.

Работа выполнялась при поддержке РФФИ (грант № 11-07-00027а) и Госзадания "Наука" (№ 8.8113.2013).

Список литературы

1. **Васильев К. К., Крашениников В. Р., Ташлинский А. Г.** Статистический анализ последовательностей многомерных изображений // *Научно-технические технологии*. 2013. № 5. С. 5–11.
2. **Almutairi A., Warner T. A.** Change detection accuracy and image properties: a study using simulated data // *Remote Sensing*. 2010. N 2. P. 1508–1529.
3. **Lu D., Mausel P.** Change detection techniques // *Remote Sensing*. 2004. V. 25. N 20. P. 2365–2407.
4. **Radke R. J.** Image Change Detection Algorithms: A Systematic Survey // *IEEE Trans. Image Process.* 2005. V. 14. N 3. P. 294–307.
5. **Gong P.** Change detection using principal component analysis and fuzzy set theory // *Canadian Journal of Remote Sensing*. 1993. N 19. P. 22–29.
6. **Jianguo G., Haigang S., Guorui M., Qiming Z.** A review of multitemporal remote sensing data change detection algorithms // *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. 2008. V. 37. P. 757–762.
7. **Гуенко А. Я., Гуенко Г. А., Говоров М. О.** Геоинформационное картографирование и мониторинг изменения окружающей среды // *СибБезопасность-СпасСиб*. 2012. № 1. С. 150–157.
8. **Collins J. B., Woodcock C. E.** An Assessment of Several Linear Change Detection Techniques for Mapping Forest Mortality Using Multitemporal Landsat TM Data // *Remote Sensing of Environment*. 1996. N 56. P. 66–77.
9. **Dhakal A.S., Amada T., Aniya M., Sharma R. R.** Detection of areas associated with flood and erosion caused by a heavy rainfall using multitemporal Landsat TM data // *Photogrammetric Engineering and Remote Sensing*. 2002. N 68. P. 233–240.
10. **Shlens J.** A Tutorial on Principal Component Analysis. 2009. 12 p. [Электронный ресурс]. URL: <http://www.snl.salk.edu/~shlens/pca.pdf> (22.08.2013), свободный.
11. **Kauth R. J., Thomas G. S.** The Tasseled Cap — A Graphic Description of the Spectral-Temporal Development of Agricultural Crops as Seen by LANDSAT // *LARS Symposia*. 1976. P. 41–51.
12. **Pacifici F.** Change detection algorithms: State of the art. [Электронный ресурс]. URL: http://www.disp.uniroma2.it/earth_observation/pdf/CD-Algorithms.pdf (22.08.2013), свободный.
13. **Dianat R., Kasaei S.** On Automatic Threshold Selection in Regression Method for Change Detection in Remote Sensing Images [Электронный ресурс]. URL: http://sharif.edu/~skasaei/Papers/Ist_08_RD.pdf (22.08.2013), свободный.
14. **İlsever M., Ünsalan C.** Two-Dimensional Change Detection Methods. Springer, 2012. 72 p.
15. **Ridd M. K., Liu J.** A comparison of four algorithms for change detection in an urban environment // *Remote Sens. Environ.* 1998. V. 68. P. 95–100.
16. **Sohl T.** Change analysis in the United Arab Emirates: an investigation of techniques // *Photogrammetric Engineering and Remote Sensing*. 2010. N 65. P. 475–484.
17. **Lambin E. F., Strahlers A. H.** Change-vector analysis in multitemporal space: A tool to detect and categorize land-cover change processes using high temporal-resolution satellite data // *Remote Sensing of Environment*. 1994. V. 48. P. 231–244.
18. **Fawcett T.** ROC Graphs: Notes and Practical Considerations for Researchers. [Электронный ресурс]. URL: <http://binf.gmu.edu/mmaso/ROC101.pdf> (22.08.2013), свободный.
19. **Pontius R. G. Jr., Schneider L. C.** Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA // *Agriculture, Ecosystems and Environment*. 2001. N 85. P. 239–248.

ОБРАБОТКА ИЗОБРАЖЕНИЙ IMAGE PROCESSING

УДК 004.932.2

А. В. Терехин, аспирант, e-mail: terehin_murom@mail.ru,
Муромский институт (филиал) ГОУ ВПО "Владимирский государственный университет
имени Александра Григорьевича и Николая Григорьевича Столетовых"

Распознавание трехмерных объектов по изображениям двух проекций

Предлагается подход к распознаванию трехмерных объектов методом вычисления оценок по признакам формы с применением моделей окто-деревьев.

Ключевые слова: распознавание, трехмерный объект, проекция, алгоритм вычисления оценок

A. V. Terekhin

Three-Dimensional Object Recognition by Images of Two Projections

In this paper, proposes an approach to the recognition of three-dimensional objects by computing estimates based on features of forms and octree models.

Keywords: recognition, three-dimensional object, projection, algorithm of computing estimates

Введение

Современная техника развивается очень быстро, но ее совершенствование не может происходить без улучшения программного обеспечения. Сами по себе различные механизмы не могут думать или видеть. Для реализации этих функций необходимо соответствующее программное обеспечение, в котором в зависимости от конкретной задачи распознавания используются различные алгоритмы. Среди решаемых задач выделяются следующие:

- идентификация объектов;
- определение их местоположения;
- выявление дефектов.

Существует множество подходов, алгоритмов и методов, позволяющих распознавать объекты на изображениях. Некоторые из них больше используют аппаратные возможности систем, другие — опираются на программные комплексы. В данной статье речь пойдет о задаче распознавания трехмерных объектов и о подходе, который позволяет ее решить.

На данный момент эта задача в основном решается средствами плоской геометрии с использованием признаков контура, кривизны, формы, отверстий, текстурных признаков [1—4].

Алгоритмы и методы, использующие перспективные преобразования для восстановления формы трехмерного объекта по одному изображению, не

могут решить задачу распознавания в ситуациях, когда у одного объекта главный вид идентичен по форме, размеру, числу отверстий и другим признакам какому-нибудь из видов другого объекта. Примером могут служить цилиндр и сфера одного диаметра. Если цилиндр расположен основанием к камере, ни один из этих признаков не отличит его от сферы. Сразу возникает проблема распознавания трехмерных объектов в подобных случаях [5, 6].

В данной статье предлагается подход к ее решению на основе использования двух камер и специального набора эталонов, хранящего шесть ортогональных проекций каждого из рассматриваемых объектов [8] и модели окто-деревьев [9—11]. В качестве характеристик путем проведенных ранее исследований были выбраны признаки формы и отверстий [4].

Метод вычисления оценок признаков формы

Современные роботы, снабженные телевизионными камерами, способны достаточно хорошо видеть, чтобы работать с реальным миром. Они могут определять тип объектов, в каких они находятся отношениях между собой, какие группы образуют, какой текст содержат и т. д. Однако сложные задачи распознавания, например распознавание похожих

трехмерных быстро движущихся объектов, требуют совершенствования методов и средств для своего решения.

Классический метод вычисления оценок описывается в работе [11]. Ниже предложена его модификация для распознавания трехмерных объектов.

Пусть дано множество M объектов. На этом множестве существует разбиение на конечное число подмножеств (классов) $\Omega_i, i = 1, \dots, m, M = \cup \Omega_i$. Объекты ω задаются значениями некоторых признаков $x_j, j = \{1, \dots, n\}$.

Любой трехмерный объект отображается на двумерной плоскости в виде проекции. Одним из способов проецирования является прямоугольное (ортогональное) проецирование). В нем совместно изображаются виды сверху, спереди и слева. На практике, при производстве, различные объекты, требующие идентификации, обычно лежат в произвольном положении, поэтому необходимо рассмотреть не только эти три проекции, но и их противоположные (снизу, сзади, справа). Предложенный подход к распознаванию объектов основан на описании шести ортогональных проекций трехмерного объекта и формировании по ним трехмерной модели в виде окто-дерева. Неизвестный объект фиксируется с помощью двух камер. Первая камера (основная) — располагается над сценой и делает снимок каждой ортогональной проекции объекта. Вторая камера (дополнительная) располагается под заданным углом к центру сцены и производит снимок косоугольной проекции объекта.

Задача распознавания трехмерного объекта заключается в том, чтобы для заданного экземпляра проекции φ объекта ω и набора классов $\Omega_1, \dots, \Omega_m$ по обучающей информации о классах $I_0 = (\Omega_1, \dots, \Omega_m)$ и описанию $I(\omega)$ (табл. 1) вычислить предикаты:

$$P_i(\omega \in \Omega_i) = \{1(\varphi \in \Omega_i), 0(\varphi \notin \Omega_i)\},$$

где $i = 1, \dots, m$.

Метод вычисления оценок для признаков формы трехмерных объектов сводится к построению двух бинарных матриц оценок. Для построения используются оценки $\varepsilon_{11}, \dots, \varepsilon_{m6}$ и $\gamma_k, \dots, \gamma_l$ (где m — число объектов, $1, \dots, 6$ — номера проекций). Две проекции считаются похожими, если сумма значений их оценок является максимальной.

Решение о том, к какому классу отнести объект, выносится на основе вычисления степени сходства распознавания объекта (строки) со строками, принадлежность которых к заданным классам известна:

$$\begin{cases} W_{1i} = \sum_j \varepsilon_j \\ W_{2i} = \sum_j \gamma_j \end{cases}$$

где W_{1i}, W_{2i} — оценки сходства ортогональной и косоугольной проекций неизвестного объекта с i -м классом; ε, γ — оценки j -го признака.

Таблица обучения для ортогональной проекции трехмерного объекта

Объект	Проекция	Признаки и их значения				
		x_1	...	x_n	Оценка	Класс
ω_1	φ_1	α_{111}	...	α_{11n}	ε_{11}	Ω_1
	
	φ_6	α_{161}	...	α_{16n}	ε_{16}	
...
ω_m	φ_1	α_{m11}	...	α_{m1n}	ε_{m1}	Ω_m
	
	φ_6	α_{m61}	...	α_{m6n}	ε_{m6}	

Таблица 2

Таблица обучения для косоугольной проекции трехмерного объекта

Объект	Проекция	Признаки и их значения				
		x_1	...	x_n	Оценка	Класс
ω_k	q_k	β_{k1}	...	β_{kn}	γ_k	Ω_k
...
ω_l	q_l	β_{l1}	...	β_{ln}	γ_l	Ω_m

В табл. 2 хранятся значения признаков формы и их оценки только для тех эталонных объектов, которые были распознаны с применением данных из табл. 1.

Оценка сходства ортогональной проекции W_{1i} вычисляется по данным из табл. 1, соответственно оценка сходства косоугольной проекции W_{2i} вычисляется по данным табл. 2.

Искомый эталон определяется в два этапа по следующему правилу:

$$1) Y = \max(W_{1i}),$$

где Y — множество эталонов, у которых одна из проекций "похожа" на ортогональную проекцию неизвестного объекта;

$$2) S = \max(W_{2i}),$$

где S — искомый эталон.

Вектор признаков

В данном подходе используются признаки формы, пропорций (рис. 1) и отверстий [5]. Признаки разбиты на три группы:

- признаки пропорций объекта;
- признаки формы;
- признаки отверстий.

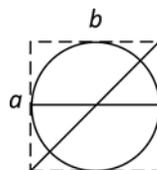


Рис. 1. Схематическое представление объекта № 1, вписанного в прямоугольник; a, b — стороны описанного прямоугольника

Для вычисления выбранных признаков вокруг объекта строится описанный прямоугольник минимально возможной площади. Это нужно для получения постоянных величин при вычислении признаков проекции трехмерного объекта на плоскость, не зависящих от его ориентации и расположения.

Признаки пропорций

Прямоугольный коэффициент формы (ПКФ) — отношение ширины к длине описанного вокруг проекции объекта прямоугольника (диапазон значений от 0 до 1). Данный признак позволяет вычислить пропорции и представить их в виде коэффициента с диапазоном значений от 0 до 1:

$$k_{\text{пр}} = \frac{l_a}{l_b},$$

где l_a, l_b — соответственно длины сторон описываемого вокруг объекта прямоугольника.

Коэффициент периметр — площадь (КПП) — отношение периметра внешнего контура объекта к его площади:

$$k_{\text{пп}} = \frac{P_o}{S_o},$$

где P_o, S_o — соответственно периметр и площадь объекта.

Признаки формы

Диагональный коэффициент формы (ДКФ) — отношение длины объекта (длина отрезка максимальной длины между двумя точками контура объекта) к длине диагонали описанного прямоугольника (диапазон значений от 0 до 1) [4]. Для объекта, изображенного на рис. 2, диагональный коэффициент:

$$k_d = \frac{l_e}{l_c},$$

где l_c — длина диагонали описанного прямоугольника; l_e — длина объекта.

Коэффициенты диагоналей объекта (КД). Диагонали объекта — отрезки, лежащие на диагоналях описанного вокруг него прямоугольника и соединяющие две точки контура объекта. Координаты точек пересечения диагонали и контура объекта находятся по общему уравнению прямой. На рис. 2 диагоналями объекта являются FG' и $F'G$. Обозначим их длины l_{d1} и l_{d2} .

Коэффициент диагонали объекта — отношение длины диагонали объекта к длине объекта:

$$k_{d1} = \frac{l_{d1}}{l_e};$$

$$k_{d2} = \frac{l_{d2}}{l_e}.$$

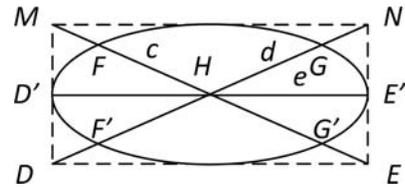


Рис. 2. Схематическое представление объекта № 2, вписанного в прямоугольник:

c, d — диагонали описанного прямоугольника; e — длина отрезка максимальной длины между двумя точками контура объекта, F, F', G, G' — точки пересечения диагоналей описанного прямоугольника с контуром объекта

Коэффициенты диагональных отрезков (КДО). Вводятся понятия диагональных отрезков и коэффициентов диагональных отрезков. Диагональные отрезки описанного прямоугольника соединяют его вершины с серединой максимального отрезка объекта (рис. 2 — HN, HM, HD, HE). Диагональные отрезки объекта лежат на диагональных отрезках описанного прямоугольника и соединяют середину максимального отрезка объекта с точкой пересечения ее контура.

Коэффициенты диагональных отрезков — отношения длин соответствующих диагональных отрезков объекта к диагональным отрезкам описанного вокруг него прямоугольника.

Для объекта, изображенного на рис. 2, диагональными отрезками будут: HG, HF, HF', HG' ; для описанного прямоугольника: HN, HM, HD, HE . Их длины можно условно обозначить как l с индексом соответствующего отрезка:

$$k_{\text{до1}} = \frac{l_{HG}}{l_{HN}}, k_{\text{до2}} = \frac{l_{HF}}{l_{HM}}, k_{\text{до3}} = \frac{l_{HF'}}{l_{HM}}, k_{\text{до4}} = \frac{l_{HG'}}{l_{HE}}$$

где $k_{\text{до1}}, k_{\text{до2}}, k_{\text{до3}}, k_{\text{до4}}$ — коэффициенты диагональных отрезков объекта.

Признаки отверстий

Признаки отверстий (ПО) подразделяются на коэффициент числа отверстий, коэффициенты расстояний:

Коэффициент числа отверстий

$$k_{\text{по}} = \frac{1}{n+1},$$

где n — число отверстий.

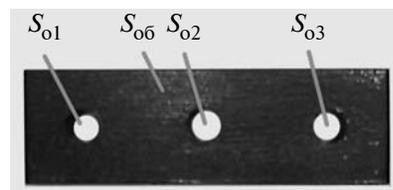


Рис. 3. Пример проекции реального трехмерного объекта

При $n = 0$ значение $k_{\text{по}}$ будет равно 1; $k_{\text{по}} \rightarrow 0$ при $n \rightarrow \infty$.

Площадной коэффициент отверстий

$$k_{S_o} = \frac{\sum_{i=1}^n S_{o_i}}{S_{\text{об}}},$$

где S_{o_i} — площадь i -го отверстия; $S_{\text{об}}$ — площадь объекта (рис. 3).

Коэффициенты расстояний. Для вычисления данных признаков находятся координаты центра объекта и отверстий, а также расстояния от центра объекта до середины каждого из них, выбирается минимальное и максимальное.

Коэффициент минимального расстояния между центром отверстия и центром объекта

$$k_{l_{\min}} = \frac{l_{\min}}{l_{\text{об}}},$$

где l_{\min} — расстояние от центра объекта до середины ближайшего отверстия.

Коэффициент максимального расстояния между центром отверстия и центром объекта:

$$k_{l_{\max}} = \frac{l_{\max}}{l_{\text{об}}},$$

где l_{\max} — расстояние от центра объекта до середины максимально удаленного отверстия.

В случае отсутствия отверстий на проекции объекта значения $k_{l_{\min}}$, $k_{l_{\max}}$, k_{S_o} будут равны 0, и объект будет распознаваться по остальным признакам формы.

Отклонения признаков

Так как алгоритм вычисления оценок работает с бинарными таблицами, для его применения необходимо вычислить диапазоны отклонений значений признаков. Для этого было разработано тестовое приложение с генератором объектов, выполняющее следующую последовательность действий:

1. По выбранным из базы эталонов объектов генерируются выборки тестов.

2. Вычисляются признаки экземпляра каждого теста с помощью выборки сгенерированных объектов одного типа.

Столбцами матрицы являются векторы признаков:

$$\begin{aligned} x_1 &= \{a_{11}, a_{12}, \dots, a_{1n}\}; \\ &\dots; \\ x_n &= \{a_{1n}, a_{2n}, \dots, a_{mn}\}. \end{aligned}$$

Матрица A выборки сгенерированных объектов одного типа выглядит следующим образом:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix},$$

где n — число признаков, m — число генерируемых объектов одной выборки (настраивается пользователем); a_{ij} — значение признака j генерируемого объекта i ($i = 1, \dots, n, j = 1, \dots, m$).

3. Вычисляются отклонения каждого признака по каждому столбцу матрицы A :

$$\Delta k_i = \max(a_i) - \min(a_i),$$

где Δk_i — отклонение i -го признака объекта.

Пункты 1, 2 выполняются для каждой выборки. Формируются векторы отклонения признаков:

$$\Delta K_l = \{\Delta k_1, \Delta k_2, \dots, \Delta k_n\},$$

где l — тип объекта; ΔK_l — вектор отклонений объекта типа l .

4. Отклонения по каждой выборке объектов сохраняются в памяти ЭВМ.

5. Из полученных векторов ΔK_l строится матрица K , содержащая отклонения значений всех признаков для каждого типа объектов:

$$K = \begin{bmatrix} \Delta k_{11} & \Delta k_{12} & \dots & \Delta k_{1n} \\ \Delta k_{21} & \Delta k_{22} & \dots & \Delta k_{2n} \\ \dots & \dots & \dots & \dots \\ \Delta k_{l1} & \Delta k_{l2} & \dots & \Delta k_{ln} \end{bmatrix}.$$

Столбцами матрицы K являются векторы отклонений по признакам x :

$$\begin{aligned} x_1 &= \{\Delta k_{11}, \Delta k_{21}, \dots, \Delta k_{l1}\}; \\ &\dots; \\ x_n &= \{\Delta k_{1n}, \Delta k_{2n}, \dots, \Delta k_{ln}\}. \end{aligned}$$

По матрице K определяются максимальные отклонения по столбцам:

$$\delta_i = \max(\Delta k_{ij}),$$

где δ_i — отклонение i -го признака.

Вектор отклонений признаков для любого типа объектов из заданного набора D вычисляется по формуле:

$$D = \{\delta_1, \delta_2, \dots, \delta_n\}.$$

Модель окто-дерева

Окто-деревья — это иерархические древовидные структуры, используемые в некоторых графических системах для представления объемных объектов. Их основное достоинство заключается в удобстве хранения трехмерных моделей объектов и возможности вариаций хранения данных в зависимости от требуемой детализации [9—11]. Древовидная структура организована так, что каждый узел соответствует области трехмерного пространства. Это представление объемных тел использует пространственную когерентность, чтобы снизить требования к памяти для хранения трехмерных объектов. Кроме того, это представление удобно для хранения информации о внутренних областях объектов.

В данном подходе модели окто-дерева используются для построения косоугольной проекции эталона для сравнения с косоугольной проекцией неизвестного объекта.

Эксперименты

Для проверки предложенного подхода была разработана программа на ЭВМ и проведены эксперименты на тестовых сгенерированных и реальных объектах. В качестве технического обеспечения использовались две HD web-камеры Logitech C310 с установленным разрешением изображения 1280×960 пикселей. Объем экспериментальной выборки составил 2000 экземпляров. В экспериментах использовались десять различных трехмерных объектов, у каждого из которых рассматривалось шесть ортогональных и одна косоугольная проекция. При проведении экспериментов использовались различные комбинации признаков из представленного вектора.

На рис. 4 изображен график вероятности распознавания трехмерного объекта по одной ортогональной

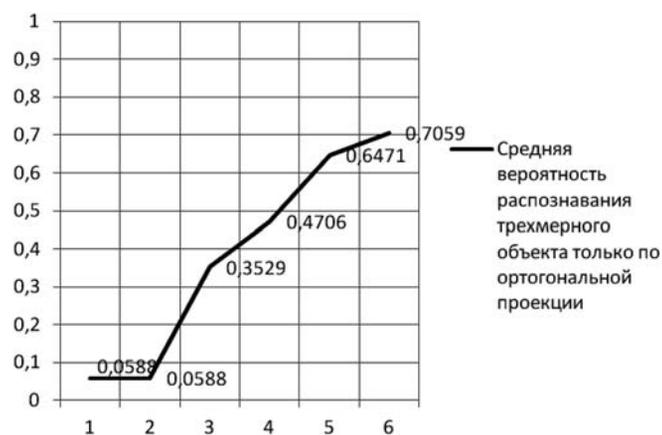


Рис. 4. График вероятности распознавания трехмерного объекта только по одной ортогональной проекции с различным числом используемых признаков (по горизонтальной оси — число используемых признаков формы)

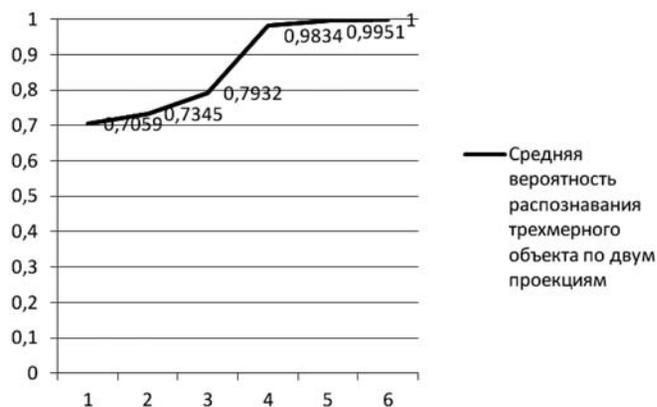


Рис. 5. График вероятности распознавания трехмерного объекта по двум проекциям с различным числом используемых признаков (по горизонтальной оси — число используемых признаков формы)

ной проекции неизвестного объекта при использовании различного числа признаков. Здесь 1 — ПКФ; 2 — ПКФ, ПО; 3 — ПКФ, отверстия, КПП; 4 — ПКФ, ПО, КПП, КДО; 5 — ПКФ, ПО, КПП, КДО, КД; 6 — ПКФ, ПО, КПП, КДО, КД, ДКФ. Из графика следует, что даже при использовании всех признаков формы объекта средняя вероятность распознавания не превышает 0,7059, т. е. составляет 70,59 %. Это связано с тем, что у разных трехмерных объектов могут совпадать различные ортогональные проекции, в связи с этим и возникают ошибки распознавания. Но, тем не менее, данный этап позволяет сократить число рассматриваемых эталонов на 71 %. Это дает возможность сократить число вычислений при анализе косоугольной проекции, если объект не был распознан по ортогональной.

Рис. 5 иллюстрирует среднюю вероятность распознавания трехмерного объекта по ортогональной и косоугольной проекции с использованием различного числа признаков формы.

Заключение

Исследования показали, что даже независимо от набора признаков при распознавании трехмерных объектов по их проекции не достаточно одного изображения. Применение моделей окто-деревьев позволяет сократить число требуемых изображений неизвестного объекта до двух. При этом достигается 100 %-ная вероятность распознавания.

Из 2000 сгенерированных испытаний все 2000 объектов по двум изображениям и трем группам признаков были распознаны верно.

Данный подход не требует применения дополнительных лазерных подсветок или датчиков расстояний и может быть использован в производстве на сборочном конвейере для обеспечения зрения робота манипулятора — определения типа и обнаружения местоположения неизвестного объекта на сцене.

Список литературы

1. Хорн Б. К. П. Зрение роботов: Пер. с англ. М.: Мир, 1989. 487 с.
2. Гонсалес Р., Вудс Р. Цифровая обработка изображений. М.: Техносфера, 2005. 1072 с.
3. Розенфельд А. Распознавание и обработка изображений с помощью вычислительных машин. М.: Мир, 1972. 232 с.
4. Терехин А. В. Метод формирования вектора признаков для идентификации проекций реальных трехмерных объектов // Наука и современность — 2013. Сб. матер. XX междунар. научно-практ. конф. Новосибирск, 20 февраля 2013 г. / Под общей редакцией С. С. Чеснокова. 2013. 300 с.
5. Sensoren. URL: <http://www.sensoren.ru/>
6. Sick sensor intelligence. URL: <http://www.sick-automation.ru/>
7. Терехин А. В. Метод описания эталонов трехмерных объектов по форме их проекций и признакам отверстий // АМиСОД — электронный журнал / Под ред. С. С. Садыкова, Д. Е. Андрианова. Муром: Муромский институт (филиал) ВлГУ, 2013.
8. Jackins C. L., Tanimoto S. L. Octrees and Their Use in Representing Three-dimensional Objects. // CGIP (Computer Graphics Image Processing). 1980. Vol. 14, N 3. P. 249—270.
9. Noborio H., Fukuda S., Arimoto S. Construction of the Octree Approximating Three-dimensional Objects by Using Multiple Views // IEEE Trans. PAMI (IEEE Trans. Pattern Analysis and Machine Intelligence). 1988. V. 10, N 6. P. 769—782.
10. Chien C. H., Aggarwat J. K. Volume/Surface Octrees for the Representation of 3-D Objects // CGIP (Computer Graphics Image Processing). 1986. Vol. 36. P. 100—113.
11. Журавлев Ю. И., Никифоров В. В. Алгоритмы распознавания, основанные на вычислении оценок // Кибернетика. 1971. Т. 3. С. 1—11.

УДК 519.61:511-33

Н. А. Балонин, д-р техн. наук, проф., **Ю. Н. Балонин**, программист, e-mail: yuraball@mail.ru, Санкт-Петербургский государственный университет аэрокосмического приборостроения,
М. Б. Сергеев, д-р техн. наук, проф.,
Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики

О модификации метода Скарпи вычисления матриц Мерсенна для задач преобразования изображений

Рассматриваются матрицы Мерсенна, двухуровневые по модулю. Дано определение слоя и сечения семейства квазиортогональных матриц, приведены алгоритмы вычисления матриц соседних слоев и сечений по матрицам меньшего порядка. Приводится новая формулировка алгоритма квадрирования матриц Скарпи. Показывается способ вычисления приближения к матрицам Адамара порядка 668 и других порядков.

Ключевые слова: ортогональные матрицы, квазиортогональные матрицы, матрицы Адамара, матрицы Белевича, матрицы Мерсенна, числа Мерсенна, метод Скарпи, массив Вильямсона, защита видеоданных

N. A. Balonin, Yu. N. Balonin, M. B. Sergeev

Scarpis Method Calculation of the Hadamard and Mersenne Matrices

Two-level in modulus Mersenne matrices are examined in the article. The definitions of a layer and a cross-section of quasi-orthogonal matrices family are given. Algorithms for calculating adjacent layers and cross-sections matrices by lesser degree matrices are presented along with a new form of Scarpis matrices squaring algorithm. There is also a way of calculation of approximation to the Hadamard matrices of 668 and other orders is shown in the article.

Keywords: orthogonal matrices, quasi-orthogonal matrices, Hadamard matrices, Belevitch matrices, Mersenne matrices, Mersenne prime, Scarpis method, Williamson array, video data protection

Введение

Ортогональные и квазиортогональные матрицы нашли применение в задачах обработки и преобразования изображений в целях сжатия, маскирования, помехоустойчивого кодирования [1—3]. Анализ показывает, что для ряда приложений целочисленность значений элементов таких матриц не столь

важна, как их высокая размерность и наличие экстремальных качеств [1, 3, 4].

В работе [5] приведены алгоритм и программа поиска и исследования **М**-матриц [4], представителями которых на порядках $n = 4k - d$ являются при $d = 0$ матрицы Адамара [6], при $d = 1$ — матрицы Мерсенна [7], при $d = 2$ — матрицы Эйлера [8] и при $d = 3$ — матрицы Ферма [9].

Первый оригинальный метод нахождения матриц Адамара, сводившийся к вычислению матрицы высоких порядков на основе матриц более низких порядков, опубликовал итальянский математик У. Скарпи [10]. В отношении некоторых матриц результат оказался не превзойденным даже при использовании более позднего метода Пэли [11], нашедшего применение теории конечных полей Галуа для построения матриц Адамара с помощью символов Лежандра.

Следует отметить, что порядки матриц, найденных Скарпи, $(n - 1)n$ для $n = 4k$, где $k = 1, 2, 3, \dots$, кратны порядкам матриц Мерсенна [7]. В настоящей работе рассматривается модификация метода Скарпи как новый способ вычисления матриц Мерсенна и Адамара, а также их приближений.

Слои и сечения матриц семейства Адамара

Матрица Мерсенна M_n [7] — квадратная матрица порядка $n = 4k - 1$ с элементами $\{1, -b\}$ такая, что $M_n^T M_n = \mu I_n$. Здесь I_n — единичная матрица,

$$\mu = \frac{(n+1) + (n-1)b^2}{2}, \text{ причем } b = \frac{1}{2} \text{ при } n = 3, \text{ в ос-}$$

тальных случаях $b = \frac{q - \sqrt{4q}}{q - 4}$, где $q = n + 1$. Число элементов уровня b [8] в каждом столбце матрицы на единицу меньше числа единичных элементов.

На рис. 1 приведена матрица третьего порядка, построенная для первого числа Мерсенна $n = 2^k - 1 = 3$ при $k = 1$. Черный цвет элемента портрета матрицы соответствует $-b$ (или -1 у матриц Адамара), а белый — единице.

Определение 1. Слоем матриц семейства Адамара будем называть совокупность квазиортогональных матриц с известными функциями зависимости значений элементов (уровней) матриц слоя от значений порядков, заданных показателем d .

Определение 2. Сечением будем называть совокупность матриц соседних слоев для значений порядков, заданных показателем k .

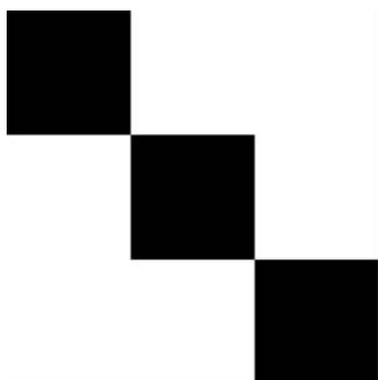


Рис. 1. Портрет матрицы M_3

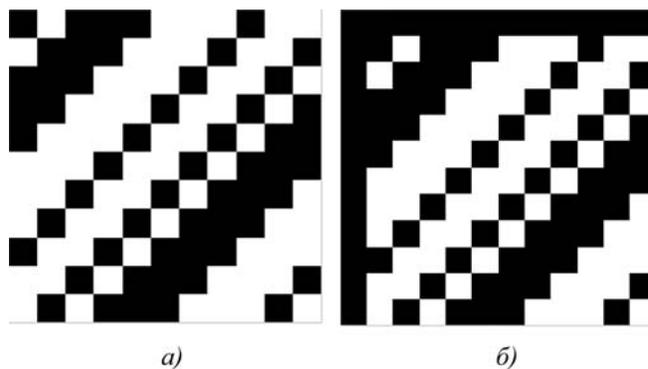


Рис. 2. Портреты матриц:
а — Мерсенна (M_{11}); б — Адамара (H_{12})

Матрицы Адамара, Мерсенна и Эйлера, согласно [6–8], образуют слои для $n = 4k - d$ для $d = 0, 1, 2$ соответственно. Матрицы Ферма такого непрерывного слоя не образуют, поскольку для них функции уровня определены на узком множестве значений $n = 2^k + 1$ при четных k [9].

Такая классификация матриц семейства Адамара позволяет рассматривать названные выше объекты как проявление одного математического объекта, заданного совокупностью слоев и сечений.

Нахождение любой матрицы сечения автоматически влечет за собой нахождение всех остальных — они взаимно зависимы. Например, матрица Адамара H_4 получается из приведенной выше матрицы Мерсенна M_3 округлением ее отрицательных элементов до значения -1 с добавлением каймы в виде строки и столбца с отрицательными элементами для соблюдения баланса положительных и отрицательных элементов.

Этот алгоритм построения матриц Адамара напрямую использует свойство общности, вытекающие из близости матриц сечения соседствующих слоев. На рис. 2 это продемонстрировано нахождением матрицы Адамара двенадцатого порядка (впервые она предложена в работе [6]), но на основе матрицы Мерсенна M_{11} .

Приближение матриц Адамара матрицами Мерсенна

Проблемные порядки матриц Адамара, для которых производительность современных вычислительных машин не позволяет пока их вычислить, хорошо известны и они вчетверо больше порядков матриц Мерсенна.

Иными словами, здесь возможно применение разновидности метода Скарпи с использованием блочно-составных структур четвертого порядка. Ограничение на число блоков (четыре) препятствует точной аппроксимации матриц Адамара матрицами, меньшими их по размерам, но это путь получения точного или достаточно близкого к ним приближе-

ния. Среди ортогональных структур четвертого порядка хорошо известен массив Вильямсона [12]

$$W = \begin{pmatrix} w_1 & w_2 & w_3 & w_4 \\ -w_2 & w_1 & -w_4 & w_3 \\ -w_3 & w_4 & w_1 & -w_2 \\ -w_4 & -w_3 & w_2 & w_1 \end{pmatrix}.$$

Замещение элементов этого массива четырьмя матрицами Вильямсона позволило найти перебором элементов все предыдущие столь же проблемные матрицы Адамара. К сожалению, число операций перестановок их элементов слишком велико (даже при упрощении структур матриц Вильямсона до некоторых циклических, связанных с задающими их векторами), а, главное, это не имеет серьезной перспективы, поскольку не дает принципиального решения — обнаружение очередной матрицы отодвигает проблемный порядок, и только.

Имея только одну двухуровневую матрицу Мерсенна вчетверо меньшего порядка, используем ее следующим образом: в качестве первой матрицы Вильямсона используем удвоенную матрицу Мерсенна, а в качестве остальных трех — ее же, но с округленными до значений $\{1, -1\}$ коэффициентами. В результирующей матрице число вещественных коэффициентов уровня b уменьшается вчетверо, но остается возможность обеспечить ортогональность варьированием его значения.

Для получения симметричной структуры приравняем друг к другу последние три элемента массива Вильямсона, инвертируем по знаку три нижних его строки и переставим местами второй и третий столбец. Такую структуру будем называть Пропусом (Propuses).

Работоспособность подхода Скарпи подтверждается тем, что стартовые матрицы Мерсенна M_3 и Эйлера E_6 сразу переходят в матрицы Адамара

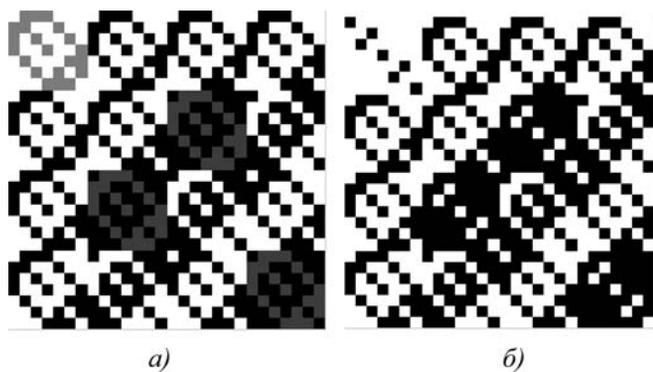


Рис. 3. Портреты матриц: а — Вильямсона (W_{28}); б — Адамара (H_{28})

$H_{12} = W_{12}$ и $H_{24} = W_{24}$. Отрыв от аппроксимируемой таким способом структуры матриц Адамара начинается только с применения матрицы M_7 , да и то, без существенного изменения итоговой матрицы, предложенный в работе [4] алгоритм сжатия ее адямаровой нормы (h -нормы) позволяет свести итог к H_{28} , что демонстрируется на рис. 3.

Задача нахождения матрицы Адамара решается редукцией структуры: измененная сжатием одна из четырех матриц Вильямсона содержит в явном виде матрицу Мерсенна M_3 . То, что за матрицами Вильямсона стоят матрицы Мерсенна, ранее в литературе не обсуждалось.

Учитывая, что матрицы Мерсенна с ростом порядка стремятся к матрицам Адамара, массив Вильямсона с матрицей M_{167} даже без адаптации структуры на порядке 668 близок к матрице H_{668} (значение $b \approx 0,7$).

Модификация метода Скарпи

Наиболее интересный результат можно получить, находя с помощью матриц Мерсенна низкого порядка такие же матрицы более высокого порядка. Для учета условий ортогональности применим циклическую перестановку столбцов блоков. Поскольку в оригинале [10] еще не существует понятия матриц Мерсенна с их естественным разделением множества всех элементов на целые (1) и вещественные ($-b$), идея прежнего метода Скарпи не может быть выражена в столь простой формулировке, как приводимая ниже.

Модифицированный алгоритм. Любую матрицу Мерсенна порядка n , где n — простое число, можно вставить саму в себя с циклическим сдвигом, пропорциональным ее положению, используя в качестве каймы замещаемый элемент матрицы: после нормализации и усечения каймы получим снова матрицу Мерсенна.

Циклический сдвиг столбцов определяется произведением индексов элементов расширяемой матрицы, отсчет их начинается с 0. Стартовый элемент каймы блока (пересечение его первых строки и столбца) выбирается отрицательным. Под нормализацией подразумевается выравнивание знаков первого столбца и строки итоговой матрицы так, чтобы они были отрицательными — усечение каймы воспроизводит их количественный дефицит на единицу в матрице Мерсенна, но теперь уже порядка $n^2 + n - 1$. Коэффициент $-b$ матрицы рассчитывается заново, согласно определению. Отметим, что полином $n^2 + n - 1 = 0$ имеет корни, равные по модулю числам золотого сечения, т. е. известная в математике пропорция так или иначе лежит в основе этого построения.

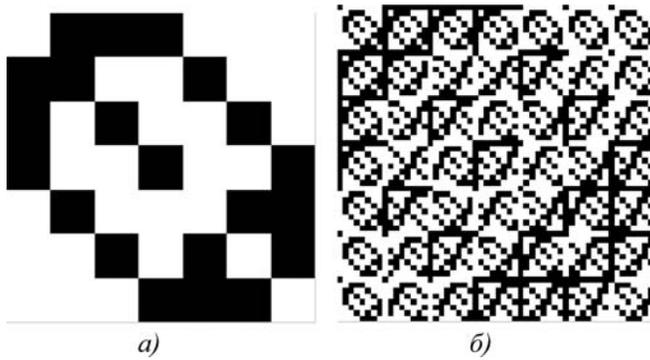


Рис. 4. Портреты матриц:
 а — Мерсенна (M_7); б — Адамара (H_{56})

В качестве примера применим матрицу Мерсенна M_7 не для построения матрицы Вильямсона, а для увеличения числа блоков составной матрицы до значения ее порядка, вставляя матрицу саму в себя согласно алгоритму. После нормализации и усечения общей каймы получим матрицу Мерсенна M_{55} . Побочный продукт этого алгоритма — промежуточная матрица Адамара, поскольку это матрицы одного сечения.

На рис. 4 показана исходная матрица M_7 и результирующая матрица Адамара H_{56} до ее нормализации, что позволяет наблюдать структуру и характер сдвигов столбцов размножаемой матрицы.

Заключение

Для вычисления матриц Мерсенна, Адамара и близких к ним матриц семейства высоких порядков с экстремальными свойствами, важных для задач помехоустойчивого кодирования информации и защитного маскирования изображений, предложен новый алгоритм их вычисления.

Матрицы Адамара не всегда являются конечной целью подобных алгоритмов, поскольку известны приложения, в которых содержательная сторона задачи связывается с модульно двухуровневыми мат-

рицами Мерсенна. По простоте и надежности поиска матрицы Мерсенна предпочтительнее матриц Адамара. Справа и слева от матриц Мерсенна по значениям их порядков находятся связанные с ними взаимно однозначным соответствием матрицы Адамара и Эйлера.

Модификация метода Скарпи позволяет непосредственно вычислять матрицы Мерсенна и Адамара высоких порядков, недостижимых при использовании метода Пэли [11]. Уменьшение числа блоков и отказ от перестановок элементов матрицы Мерсенна позволяют эффективно находить и приближения матриц Адамара высоких проблемных порядков.

Список литературы

1. **Мироновский Л. А., Слаев В. А.** Стрип-метод преобразования изображений и сигналов. СПб: Политехника, 2006. 163 с.
2. **Ерош И. Л., Сергеев А. М., Филатов Г. П.** О защите цифровых изображений при передаче по каналам связи // Информационно-управляющие системы. 2007. № 5. С. 20—22.
3. **Балонин Ю. Н., Востриков А. А., Сергеев М. Б.** О прикладных аспектах применения M -матриц // Информационно-управляющие системы. 2012. № 1. С. 92—93.
4. **Балонин Н. А., Сергеев М. Б.** M -матрицы // Информационно-управляющие системы. 2011. № 1. С. 14—21.
5. **Балонин Ю. Н., Сергеев М. Б.** Алгоритм и программа поиска и исследования M -матриц // Научно-технический вестник информационных технологий, механики и оптики. 2013. № 3. С. 82—86.
6. **Hadamard J.** Résolution d'une question relative aux déterminants. Bulletin des Sciences Mathématiques 17. 1893. P. 240—246.
7. **Балонин Н. А., Сергеев М. Б., Мироновский Л. А.** Вычисление матриц Адамара—Мерсенна // Информационно-управляющие системы. 2012. № 5. С. 92—94.
8. **Балонин Н. А., Сергеев М. Б.** О двух способах построения матриц Адамара—Эйлера // Информационно-управляющие системы. 2013. № 1. С. 7—10.
9. **Балонин Н. А., Сергеев М. Б., Мироновский Л. А.** Вычисление матриц Адамара—Ферма // Информационно-управляющие системы. 2012. № 6. С. 90—93.
10. **Scarpis U.** Sui determinanti di valore massimo, Rendiconti della R. Istituto Lombardo di Scienze e Lettere 31. 1898. P. 1441—1446.
11. **Paley R. E. A. C.** On orthogonal matrices // Journal of Mathematics and Physics. 1933. Vol. 12. P. 311—320.
12. **Williamson J.** Hadamard's Determinant Theorem and the Sum of Four Squares // Duke Math. J. 1944. Vol. 11. P. 65—81.

СЕТИ И СИСТЕМЫ СВЯЗИ

NETWORKS AND COMMUNICATION SYSTEMS

УДК 004.73:519.2

Д. А. Перепелкин, канд. техн. наук, доц.
ФГБОУ ВПО "Рязанский государственный радиотехнический университет"
E-mail: dmitryperpelkin@mail.ru

Динамическое формирование структуры и параметров линий связи корпоративной сети на основе данных о парных перестановках маршрутов

Предложены математическая модель и алгоритм адаптивной ускоренной маршрутизации, позволяющие динамически сформировать структуру и параметры линий связи корпоративной сети на основе данных о парных перестановках маршрутов и повысить эффективность ее функционирования.

Ключевые слова: адаптивная ускоренная маршрутизация, алгоритмы маршрутизации, алгоритм парных перестановок маршрутов, динамические изменения, динамические корпоративные сети

D. A. Perepelkin

Dynamic Corporate Network Structure and Communication Links Loading Formation Based on Routes Pairs Permutations Data

Mathematical model and adaptive accelerated routing algorithm dynamic forming corporate network structure and communication links loading based on routes pairs permutation data and increases the efficiency of its operation is offered.

Keywords: adaptive accelerated routing, routing algorithms, routes pair's permutations algorithm, dynamic changes, dynamic corporate networks

Введение

Быстрый рост числа компьютерных сетей сопровождается непрерывной сменой сетевых технологий, направленной на повышение быстродействия и надежности сетей, возможностей интегрированной передачи данных, голоса и видеоинформации.

В современных корпоративных сетях обеспечение высокоскоростного и надежного обмена информацией между узлами сети при жестких требованиях к задержкам информации, при возможных подключениях и отказах узлов и линий связи в сети является одной из важнейших проблем. Модификация структуры сети, подключение или отказ узлов и линий связи приводят к полному пересчету таблиц маршрутизации. Использование традиционных методов маршрутизации в этих условиях оказывается неэффективным. Разработка новых перспективных подходов для решения задачи маршрутизации позволяет повысить эффективность функциони-

рования корпоративных сетей за счет уменьшения трудоемкости построения оптимальных маршрутов и обеспечить надежность и быстродействие передачи данных.

Постановка задачи

Выбор маршрутов в узлах связи телекоммуникационной системы проводится в соответствии с реализуемым алгоритмом маршрутизации.

В настоящее время наибольшее распространение получили алгоритмы адаптивной маршрутизации. Они обеспечивают автоматическое обновление таблиц маршрутизации после изменения конфигурации сети. Используя протоколы адаптивных алгоритмов, маршрутизаторы могут собирать информацию о топологии связей в сети и оперативно реагировать на все изменения конфигурации связей.

Анализ применяемых в современных корпоративных сетях алгоритмов адаптивной маршрутиза-

ции показывает, что для построения таблиц маршрутизации используют два известных алгоритма: Беллмана—Форда с трудоемкостью порядка $O(N^3)$ и Дейкстры с трудоемкостью $O(N^2)$, где N — число маршрутизаторов в корпоративной сети [1—3].

Применение этих алгоритмов в условиях динамического формирования структуры и параметров линий связи при построении дерева оптимальных маршрутов в современных корпоративных сетях является неэффективным вследствие высокой трудоемкости поиска оптимальных маршрутов и необходимости полного пересчета таблиц маршрутизации.

Развитие в последнее время технологии программно-конфигурируемых сетей [4] позволяет формулировать различные задачи оптимальной маршрутизации и балансировки сетевого трафика [5, 6].

В работах [7—9] разработан алгоритм парных переходов, позволяющий за счет сбора дополнительной информации учесть возможные изменения параметров линий связи корпоративной сети и не проводить полный пересчет маршрутных таблиц. Это позволило снизить трудоемкость расчета таблиц маршрутизации до значения порядка $O(kN)$, где k — число фактически выполненных парных переходов.

В работах [2, 10—16] на основе алгоритма парных переходов предложены эффективные алгоритмы адаптивной ускоренной маршрутизации при динамическом подключении узлов и линий связи корпоративной сети, которые также позволили снизить трудоемкость построения оптимальных маршрутов передачи данных до значения $O(kN)$. Недостатком предложенных алгоритмов является то, что при динамических изменениях параметров линий связи и в структуре корпоративной сети после каждого парного перехода необходимо рассчитывать дополнительную информацию для того, чтобы определить оптимальный маршрут до других узлов в сети.

В работах [3, 17—21] предложен алгоритм парных перестановок маршрутов при динамических изменениях параметров линий связи корпоративной сети, что позволило снизить трудоемкость построения оптимальных маршрутов передачи данных до значения $O(N)$. Однако данный алгоритм не учитывает возможность динамического формирования дерева оптимальных маршрутов корпоративной сети при подключении узлов и линий связи в ее структуру. В связи с этим трудоемкость построения оптимальных маршрутов передачи данных оказывается выше трудоемкости существующих алгоритмов.

Разработка новых, более эффективных моделей и алгоритмов адаптивной ускоренной маршрутизации при динамическом формировании структуры и параметров линий связи корпоративной сети на основе данных о парных перестановках маршрутов позволяет уменьшить трудоемкость построения таблиц маршрутизации до значения $O(N)$.

Для повышения эффективности функционирования корпоративных сетей предложены математическая модель и алгоритм адаптивной ускоренной маршрутизации при динамическом формировании структуры и параметров линий связи корпоративной сети на основе данных о парных перестановках маршрутов, позволяющие уменьшить трудоемкость построения таблиц маршрутизации до значения $O(N)$ по сравнению с известными алгоритмами с трудоемкостью $O(N^2)$ и $O(kN)$.

В общем случае для решения данной задачи применяется графовая модель корпоративной сети, в которой множество вершин графа соответствует множеству узлов связи или маршрутизаторов в сети, а множество ребер соответствует возможным каналам связи между этими узлами. Каждое ребро, соответствующее каналу связи, имеет свой вес. На практике весу ребра могут соответствовать стоимость аренды канала связи, затраты на оплату единицы трафика, передаваемого по каналу связи, соответствующему данному ребру либо более сложная функция, учитывающая большее число параметров корпоративной сети.

Математическую модель корпоративной сети представим в виде неориентированного взвешенного связного графа $G = (V, E, W)$, где V — множество вершин (узлов связи), $|V| = N$, E — множество ребер (каналов или линий связи), $|E| = M$, W — множество весов ребер (стоимость каналов или линий связи).

Пусть на графе G в некоторый момент времени уже решена задача поиска оптимальных маршрутов до всех узлов множества $V_s = V \setminus \{v_s\}$ из начального узла v_s , т. е. построено дерево оптимальных маршрутов с корнем в узле v_s . Обозначим это дерево как T_g . Рассмотрим множество каналов E графа G . По признаку вхождения каналов в дерево T_g можно разделить исходное множество E на два подмножества: $E_T \in T_g$ и $E_R \notin T_g$, $E_T \cup E_R = E$.

Множество каналов дерева E_T — множество каналов дерева T_g для графа G . Для заданного графа G согласно свойству дерева мощность множества E_T будет равна мощности множества V минус единица $|E_T| = |V| - 1$.

Множество каналов замены для дерева E_R — множество каналов графа G , не вошедших в дерево T_g . При соответствующих условиях некоторый канал $e_{i,j} \in E_R$, инцидентный узлам v_i и v_j , может перейти в множество каналов дерева E_T , заменив собой некоторый канал $e_{k,p} \in E_T$. При этом инцидентность канала $e_{k,p}$ узлу v_i или v_j является обязательным условием. В свою очередь, канал $e_{i,j}$ перейдет в множество E_R .

Будем называть такие переходы парными переходами и обозначать $e_{i,j} - e_{k,p}$.

В множестве E_R можно выделить два подмножества.

Множество каналов замены E_S для дерева — это такое подмножество множества E_R , элементы-каналы которого участвуют, по крайней мере, в одном отношении парного перехода.

Множество непарных каналов E_P — это такое подмножество множества E_R , элементы-каналы которого не участвуют ни в одном отношении из множества R .

В общем случае множество E_P может быть пустым $|E_P| = 0$. Множество E_S будет пустым только при условии, что исходный связный граф G является деревом, и задача поиска оптимальных маршрутов в этом случае лишена смысла.

Для каждого канала связи $e_{i,j} \in E$ на шкале значений весов определены точка вхождения в дерево $w_{i,j}^t$ и точка вхождения в множество замены $w_{i,j}^s$, причем $w_{i,j}^t \leq w_{i,j}^s$, под которыми понимается максимально возможный вес канала $e_{i,j}$ при его вхождении в множество каналов дерева $E_T \in T_g$ и в множество каналов замены для дерева $E_S \notin T_g$ соответственно.

Обозначим $w_{i,j}$ — вес канала, соединяющего узлы v_i и v_j . Узел v_i располагается ниже по иерархии в дереве оптимальных маршрутов относительно v_j . Множество E_T — множество каналов, каждый элемент которого входит, по крайней мере, в один оптимальный маршрут из начального узла связи, E_R — множество остальных каналов. $E_R \cup E_T = E$, $E_R \cap E_T = \emptyset$. Обозначим V_T — множество узлов, до которых найден оптимальный маршрут из начального узла связи; V_R — множество остальных узлов; $V_R \cup V_T = V$; $V_R \cap V_T = \emptyset$.

Будем называть R_k — совокупность подмножества $V^{(V_k)} \subseteq V$ узлов, через которые проходит оптимальный маршрут до узла v_k из исходного узла v_s , и подмножества $E^{(V_k)} \subseteq E$ каналов, составляющих этот маршрут.

Назовем V_k — дерево T_k или совокупность подмножества $V_T^{(V_k)} \subseteq V$, состоящего из всех узлов, оптимальные маршруты до которых из исходного узла содержат узел v_k и подмножества $E_T^{(V_k)} \subseteq E$ каналов, составляющих эти маршруты после v_k при движении от узла v_s .

Обозначим множество маршрутов до узла v_i из исходного узла v_s через Π_i , где элемент множества $\pi_{i,k} \in \Pi_i$ будет множеством не повторяющихся каналов $e_{i,j} \in E$, образующих вместе маршрут, соединяющий v_s и v_i . Каждому $\pi_{i,k} \in \Pi_i$ поставим в соответствие число, равное сумме весов входящих в него каналов, т. е. длину маршрута $d_{i,k} \in D_i$, где D_i представляет собой множество оценок оптимальных маршрутов до узла v_i из исходного узла v_s . На множестве Π_i задан селектор H , возвращающий оптимальный маршрут из множества Π_i . В том случае, если существует несколько маршрутов в Π_i с минимальной длиной, то выбирается один из них. Оп-

тимальный маршрут до узла v_i будем обозначать $\pi_i = H(\Pi_i)$, оценку его длины — d_i .

Математическую модель динамического формирования структуры и параметров линий связи на основе данных о парных перестановках маршрутов при динамических подключениях и отказах узлов и линий связи корпоративной сети можно представить в виде следующих теорем.

Теорема 1. При подключении канала связи $e_{i,j}$, инцидентного узлам связи V_i и V_j , причем V_i лежит ниже чем V_j по дереву иерархий, с весом $w_{i,j}$, без изменения окажутся оптимальные маршруты и их оценки для узлов множества $V^{(V_i)}$.

Доказательство. Пусть канал связи $e_{i,j}$ входит в маршрут $\pi_{i,k} \neq \pi_i$ и $\pi_{j,p} \neq \pi_j$. Если данный канал не уменьшает оценок обоих инцидентных ему узлов V_i и V_j , то есть $d_{i,k} \geq d_i$ и $d_{j,p} \geq d_j$, то оптимальные маршруты до всех узлов не изменятся, так как канал $e_{i,j}$ оказывает влияние, прежде всего, на инцидентные ему узлы множества V . Так как существовавшие до изменения маршруты до узлов V_i и V_j имели меньшую оценку, то канал $e_{i,j}$ не включается, и оптимальные маршруты до всех узлов не изменятся. Если уменьшилась оценка какого-либо инцидентного узла, например V_j , то эта оценка $d_{j,p}$ будет оценкой оптимального маршрута до узла V_j , и канал $e_{i,j}$ войдет в состав этого маршрута, так как не существует другого оптимального маршрута π_j до узла V_j , кроме маршрута $\pi_{j,p}$, содержащего канал $e_{i,j}$. Этот оптимальный маршрут $\pi_{j,p}$ не будет существовать, если не будут существовать оптимальные маршруты до всех промежуточных узлов $V_k \in V^{(V_i)}$ этого маршрута. Невозможно будет сказать останутся ли неизменными оптимальные маршруты до остальных узлов графа. Теорема доказана.

Следствие 1. При подключении канала связи $e_{i,j}$, инцидентного узлам связи V_i и V_j , с весом $w_{i,j}$, при не изменении оценок d_i и d_j оптимальные маршруты до всех узлов не изменятся. Если изменится оценка какого-либо узла V_i и V_j , то необходимо определить новые оптимальные маршруты и их маршруты замены для множества узлов, не принадлежащих множеству $V^{(V_i)}$.

Следствие 2. При подключении канала связи $e_{i,j}$, инцидентного узлам связи V_i и V_j , причем V_i лежит ниже чем V_j по дереву иерархий, с весом $w_{i,j}$, то необходимо определить новые маршруты замены (с учетом канала $e_{i,j}$) для множества узлов, не принадлежащих множеству $V^{(V_i)}$.

Теорема 2. При подключении некоторого узла связи V_i для всех узлов, не инцидентных узлу V_i (то есть не имеющих канала связи $e_{i,k}$), оптимальные маршруты и их оценки окажутся без изменения.

Доказательство. Новый узел связи V_i не входит в оптимальный маршрут к некоторому узлу V_i . Следовательно, и в маршруте к этому узлу $\pi_{v,j}$ отсутствует канал связи $e_{i,k}$. Поэтому новый маршрут к узлу V_j , содержащий данный канал, будет иметь

оценку $d_{j,i} > d_j$ и оптимальные маршруты не изменятся для всех узлов, не инцидентных узлу V_i . Теорема доказана.

Следствие 1. При подключении некоторого узла связи V_k , имеющего каналы связи с узлами V_i и V_j , если $d_i < d_j$, оптимальные маршруты и их маршруты замены до узла V_i не изменятся.

Следствие 2. При подключении некоторого узла связи V_k , имеющего каналы связи с узлами V_i и V_j , если $d_i < d_j$, то необходимо определить новые оптимальные маршруты и их маршруты замены для множества узлов, инцидентных узлу V_k , кроме узла V_i .

Следствие 3. При подключении некоторого узла связи V_k , имеющего каналы связи с узлами V_i и V_j , необходимо определить оптимальные маршруты и их маршруты замены до узла V_k с учетом инцидентных ему каналов $e_{i,k}$ и $e_{k,j}$.

Теорема 3. При отказе некоторого канала связи $e_{i,j}$ инцидентного узлам связи V_i и V_j , входящего в дерево оптимальных маршрутов, причем $d_i < d_j$, оптимальные маршруты и их оценки до узла V_i окажутся без изменения.

Доказательство. Узлы связи V_i и V_j входят в дерево оптимальных маршрутов. Так как $d_i < d_j$, то $d_i + w_{i,j} = d_j$ и канал $e_{i,j}$ не входит в маршрут π_i к узлу V_i . Поэтому при отказе канала $e_{i,j}$ оптимальные маршруты и их оценки до узла V_i не изменятся. Теорема доказана.

Следствие 1. При отказе некоторого канала связи $e_{i,j}$ инцидентного узлам связи V_i и V_j , входящего в дерево оптимальных маршрутов, причем $d_i < d_j$, необходимо определить новые оптимальные маршруты для множества узлов, инцидентных узлу V_j .

Следствие 2. При отказе некоторого канала связи $e_{i,j}$ инцидентного узлам связи V_i и V_j , входящего в дерево оптимальных маршрутов, причем $d_i < d_j$, необходимо в списках маршрутов до каждого узла исключить те маршруты, которые включали канал $e_{i,j}$.

Теорема 4. При отказе некоторого узла связи V_i для всех узлов, не инцидентных узлу V_i (то есть не имеющих канала связи $e_{i,k}$), оптимальные маршруты и их оценки окажутся без изменения.

Доказательство. Пусть узел связи V_i не входит в оптимальный маршрут к некоторому узлу V_j . Тогда и в маршруте к этому узлу $\pi_{v,j}$ отсутствует канал $e_{i,k}$. При отказе узла V_i будут исключены все каналы $e_{i,k}$, которые не входят в маршрут $\pi_{v,j}$, следовательно, оптимальный маршрут и его оценка окажутся без изменения. Теорема доказана.

Следствие 1. При отказе некоторого узла связи V_i необходимо определить новые оптимальные маршруты для множества узлов, инцидентных узлу V_i , имеющих в своем оптимальном маршруте канал $e_{i,k}$.

Следствие 2. При отказе некоторого узла связи V_i необходимо в списках маршрутов исключить все маршруты до узла V_i , а также те маршруты, кото-

рые включали инцидентные узлу V_i каналы (то есть имели каналы $e_{i,k}$).

На основе сформулированных и доказанных выше теорем и предложенной математической модели разработан алгоритм парных перестановок маршрутов при динамическом формировании дерева оптимальных маршрутов корпоративной сети в условиях подключений и отказов узлов и линий связи в ее структуру. Укрупненно алгоритм имеет следующий вид.

Шаг 1. Выполнить первоначальную инициализацию исходных данных и определить веса линий связи $w_{i,j}$.

Шаг 2. Построить дерево оптимальных маршрутов корпоративной сети.

Шаг 3. Для узла связи, являющегося листом дерева, провести поиск всех парных переходов без ограничений. Эти списки для удобства дальнейшей работы привязываются к узлу, инцидентному рассматриваемому каналу связи и расположенному ниже по иерархии.

Шаг 4. Если узел связи не является листом дерева, то вычислить парные переходы для этого узла и выбрать лучшие значения потенциалов парных переходов для потомков узла и собственных парных переходов. Подобную процедуру выполнить для формирования списков парных переходов в случае динамического подключения узлов и линий связи корпоративной сети.

Шаг 5. Для каждого узла сформировать полный список парных переходов. Число элементов в каждом из этих списков не превышает числа узлов графа. Такое решение позволяет отказаться от предварительной сортировки потенциалов или приращений для парных переходов без значительного усложнения алгоритма обработки изменения.

Шаг 6. Для каждого канала связи корпоративной сети определить точку вхождения в дерево оптимальных маршрутов и точку вхождения в множество замены.

Шаг 7. Для каждого узла связи корпоративной сети сформировать полный список возможных маршрутов, проходящих через каналы, состоящие в отношении парного перехода, включая и каналы, входящие в дерево оптимальных маршрутов.

Шаг 8. Анализируя полученную используемым протоколом маршрутизации информацию, определить, произошло ли динамическое подключение или отказ узлов и линий связи в структуру корпоративной сети:

- а) если да — перейти к шагу 9;
- б) иначе — к шагу 13.

Шаг 9. Используя список возможных маршрутов замены определить, требуется ли сделать парную перестановку маршрутов:

- а) если да — перейти к шагу 10;
- б) иначе — к шагу 13.

Шаг 10. Для каждого узла связи, у которого в списках возможных маршрутов произошли изменения, определить путь минимальной длины и поместить его в дерево оптимальных маршрутов.

Шаг 11. Построить новое дерево оптимальных маршрутов с учетом изменений.

Шаг 12. Сформировать таблицы маршрутизации.

Шаг 13. Передать пакеты по доступным эквивалентным маршрутам и установить флаг передачи.

Шаг 14. Пересчитать точки вхождения в дерево и переформировать список маршрутов замены для каждого узла связи, у которого произошли изменения.

Шаг 15. Перейти к шагу 8.

Работа составных частей алгоритма основывается на использовании доказанных выше теорем, следовательно, можно сделать вывод о корректности работы всего алгоритма в целом.

Используя сформулированные и доказанные теоремы, удастся рассчитать дерево оптимальных маршрутов за линейное время. Такой результат получается за счет использования дополнительной информации о возможных маршрутах замены.

Оценка трудоемкости алгоритма парных перестановок маршрутов при динамическом формировании структуры и параметров линий связи корпоративной сети показывает, что верхняя оценка трудоемкости составляет $\Omega(N)$ и нижняя оценка $\Theta(N)$.

Таким образом, разработанный алгоритм парных перестановок маршрутов позволяет повысить эффективность функционирования корпоративной сети и динамически сформировать дерево оптимальных маршрутов при подключениях и отказах узлов и линий связи в ее структуру за счет использования дополнительной информации о парных перестановках маршрутов.

Пример работы алгоритма

Рассмотрим пример работы алгоритма парных перестановок маршрутов при динамическом формировании дерева оптимальных маршрутов при возможных подключениях и отказах узлов и линий связи корпоративной сети. В качестве примера

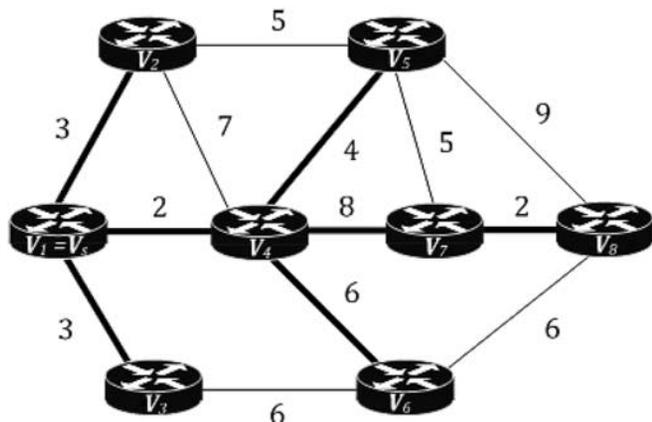


Рис. 1. Граф G корпоративной сети

рассмотрим модель корпоративной сети, представленную на рис. 1.

На рис. 1 жирными линиями обозначено построенное дерево оптимальных маршрутов. Множество каналов дерева составляет $E_T = \{e_{1,2}; e_{1,3}; e_{1,4}; e_{4,5}; e_{4,6}; e_{4,7}; e_{7,8}\}$; множество каналов замены $E_S = \{e_{2,4}; e_{2,5}; e_{3,6}; e_{5,7}; e_{6,8}\}$; множеством непарных каналов будет $E_P = \{e_{5,8}\}$. Если рассмотреть канал связи $e_{4,5}$, то для него точка вхождения в дерево будет составлять 6, а точка вхождения в множество замены — 13. При этом данный канал связи находится в отношении парного перехода с каналом $e_{2,5}$, который, в свою очередь, находится в отношении парного перехода множества замены с каналом $e_{5,7}$. После попадания $e_{4,5}$ в множество непарных каналов эта парная перестановка примет вид: $e_{2,5} - e_{5,7}$.

Рассмотрим работу алгоритма парных перестановок маршрутов на примере графа G корпоративной сети, показанного на рис. 1, в котором уже решена задача поиска оптимальных маршрутов и построено соответствующее дерево. После того как сформирован список парных переходов, дополнительно рассчитываем список оптимальных маршрутов и их маршрутов замены из исходного узла связи до каждого узла сети.

Таким образом, списки оптимальных маршрутов и их маршрутов замены до всех узлов данной модели корпоративной сети из исходного узла v_5 будут сформированы следующим образом.

Узел связи V_2 .

Маршрут $\pi_2 = \{e_{1,2}\}$ с оценкой $d_2 = 3$ — оптимальный маршрут.

Маршрут $\pi_2^{(1)} = \{e_{1,4}; e_{4,2}\}$ с оценкой $d_2^{(1)} = 2 + 7 = 9$ — маршрут замены.

Маршрут $\pi_2^{(2)} = \{e_{1,4}; e_{4,5}; e_{5,2}\}$ с оценкой $d_2^{(2)} = 2 + 4 + 5 = 11$ — маршрут замены.

Узел связи V_3 .

Маршрут $\pi_3 = \{e_{1,3}\}$ с оценкой $d_3 = 3$ — оптимальный маршрут.

Маршрут $\pi_3^{(1)} = \{e_{1,4}; e_{4,6}; e_{6,3}\}$ с оценкой $d_3^{(1)} = 2 + 6 + 6 = 14$ — маршрут замены.

Узел связи V_4 .

Маршрут $\pi_4 = \{e_{1,4}\}$ с оценкой $d_4 = 2$ — оптимальный маршрут.

Маршрут $\pi_4^{(1)} = \{e_{1,2}; e_{2,4}\}$ с оценкой $d_4^{(1)} = 3 + 7 = 10$ — маршрут замены.

Маршрут $\pi_4^{(2)} = \{e_{1,2}; e_{2,5}; e_{5,4}\}$ с оценкой $d_4^{(2)} = 3 + 5 + 4 = 12$ — маршрут замены.

Маршрут $\pi_4^{(3)} = \{e_{1,3}; e_{3,6}; e_{6,4}\}$ с оценкой $d_4^{(3)} = 3 + 6 + 6 = 15$ — маршрут замены.

Узел связи V_5 .

Маршрут $\pi_5 = \{e_{1,4}; e_{4,5}\}$ с оценкой $d_5 = 2 + 4 = 6$ — оптимальный маршрут.

Маршрут $\pi_5^{(1)} = \{e_{1,2}; e_{2,5}\}$ с оценкой $d_5^{(1)} = 3 + 5 = 8$ — маршрут замены.

Маршрут $\pi_5^{(2)} = \{e_{1,4}; e_{4,7}; e_{7,5}\}$ с оценкой $d_5^{(2)} = 2 + 8 + 5 = 15$ — маршрут замены.

Маршрут $\pi_5^{(3)} = \{e_{1,2}; e_{2,4}; e_{4,5}\}$ с оценкой $d_5^{(3)} = 3 + 7 + 4 = 14$ — маршрут замены.

Узел связи V_6 .

Маршрут $\pi_6 = \{e_{1,4}; e_{4,6}\}$ с оценкой $d_6 = 2 + 6 = 8$ — оптимальный маршрут.

Маршрут $\pi_6^{(1)} = \{e_{1,3}; e_{3,6}\}$ с оценкой $d_6^{(1)} = 3 + 6 = 9$ — маршрут замены.

Узел связи V_7 .

Маршрут $\pi_7 = \{e_{1,4}; e_{4,7}\}$ с оценкой $d_7 = 2 + 8 = 10$ — оптимальный маршрут.

Маршрут $\pi_7^{(1)} = \{e_{1,4}; e_{4,5}; e_{5,7}\}$ с оценкой $d_7^{(1)} = 2 + 4 + 5 = 11$ — маршрут замены.

Маршрут $\pi_7^{(2)} = \{e_{1,2}; e_{2,5}; e_{5,7}\}$ с оценкой $d_7^{(2)} = 3 + 5 + 5 = 13$ — маршрут замены.

Узел связи V_8 .

Маршрут $\pi_8 = \{e_{1,4}; e_{4,7}; e_{7,8}\}$ с оценкой $d_8 = 2 + 8 + 2 = 12$ — оптимальный маршрут.

Маршрут $\pi_8^{(1)} = \{e_{1,4}; e_{4,6}; e_{6,8}\}$ с оценкой $d_8^{(1)} = 2 + 6 + 6 = 14$ — маршрут замены.

Маршрут $\pi_8^{(2)} = \{e_{1,3}; e_{3,6}; e_{6,8}\}$ с оценкой $d_8^{(2)} = 3 + 6 + 6 = 15$ — маршрут замены.

Маршрут $\pi_8^{(3)} = \{e_{1,4}; e_{4,5}; e_{5,7}; e_{7,8}\}$ с оценкой $d_8^{(3)} = 2 + 4 + 5 + 2 = 13$ — маршрут замены.

Рассмотрим на данном примере работу алгоритма парных перестановок маршрутов при динамическом формировании дерева оптимальных маршрутов в условии изменения структуры и параметров линий связи корпоративной сети.

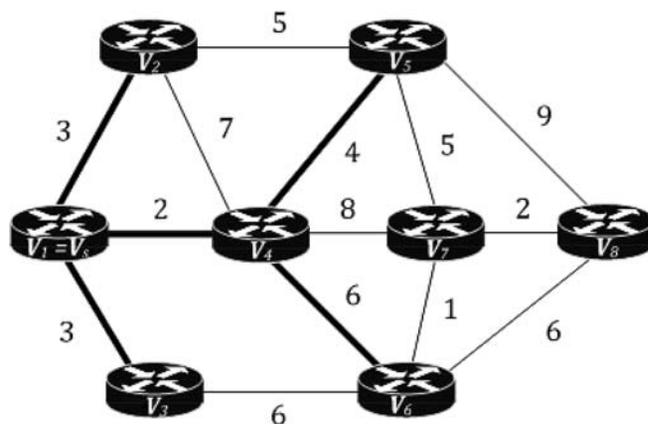


Рис. 2. Подключение канала связи $e_{6,7}$

При подключении канала связи $e_{6,7}$ структура графа G корпоративной сети примет вид, показанный на рис. 2.

На рис. 2 жирными линиями обозначено дерево оптимальных маршрутов, которое не требует изменения.

То есть при подключении канала связи $e_{6,7}$ дерево оптимальных маршрутов до узлов V_2, V_3, V_4, V_5 и V_6 согласно теореме 1 не изменится. Для узлов V_7 и V_8 необходимо переформировать список парных переходов, пересчитать точки вхождения в дерево и в множество замены, а также переформировать списки оптимальных маршрутов и их маршрутов замены с учетом подключенного канала связи $e_{6,7}$. Так для узла связи V_7 будут добавлены маршруты $\pi_7^{(1)} = \{e_{1,4}; e_{4,6}; e_{6,7}\}$ с оценкой $d_7^{(1)} = 2 + 6 + 1 = 9$ и $\pi_7^{(2)} = \{e_{1,3}; e_{3,6}; e_{6,7}\}$ с оценкой $d_7^{(2)} = 3 + 6 + 1 = 10$. Для узла связи V_8 будут добавлены маршруты $\pi_8^{(1)} = \{e_{1,4}; e_{4,6}; e_{6,7}; e_{7,8}\}$ с оценкой $d_8^{(1)} = 2 + 6 + 1 + 2 = 11$ и $\pi_8^{(2)} = \{e_{1,3}; e_{3,6}; e_{6,7}; e_{7,8}\}$ с оценкой $d_8^{(2)} = 3 + 6 + 1 + 2 = 12$. После того как новые списки маршрутов для узлов V_7 и V_8 сформированы, из них необходимо выбрать маршрут с минимальной оценкой. Оптимальный маршрут до узла связи V_7 составит $\pi_7 = \{e_{1,4}; e_{4,6}; e_{6,7}\}$ с оценкой $d_7 = 9$. Оптимальный маршрут до узла связи V_8 составит $\pi_8 = \{e_{1,4}; e_{4,6}; e_{6,7}; e_{7,8}\}$ с оценкой $d_8 = 11$.

Таким образом, для графовой модели корпоративной сети, представленной на рис. 2, множество каналов дерева составит $E_T = \{e_{1,2}; e_{1,3}; e_{1,4}; e_{4,5}; e_{4,6}; e_{6,7}; e_{7,8}\}$; множество каналов замены $E_S = \{e_{2,4}; e_{2,5}; e_{3,6}; e_{4,7}; e_{5,7}; e_{6,8}\}$; множество непарных каналов $E_P = \{e_{5,8}\}$.

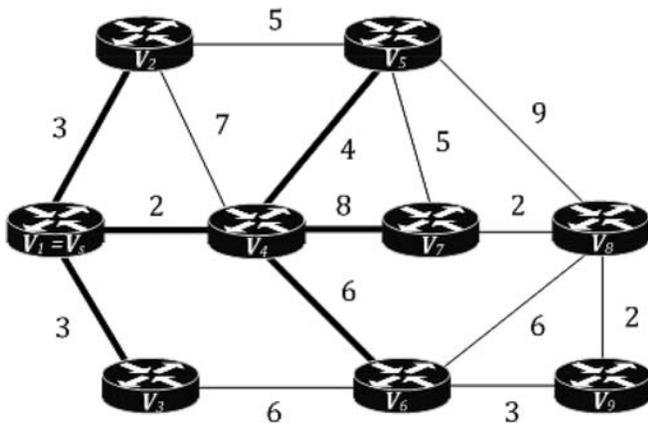


Рис. 3. Подключение узла связи V_9

При подключении узла связи V_9 в исходную структуру графа G , корпоративная сеть примет вид, показанный на рис. 3.

На рис. 3 жирными линиями обозначено дерево оптимальных маршрутов, которое не требует изменения.

То есть при подключении узла связи V_9 дерево оптимальных маршрутов до вершин V_2, V_3, V_4, V_5, V_6 и V_7 согласно теореме 2 не изменится. Для узлов связи V_8 и V_9 необходимо переформировать список парных переходов, пересчитать точки вхождения в дерево и в множество замены, а также переформировать списки оптимальных маршрутов и их маршрутов замены с учетом подключенного узла связи V_9 . Так для узла связи V_8 будут добавлены маршруты $\pi_8^{(1)} = \{e_{1,4}; e_{4,6}; e_{6,9}; e_{9,8}\}$ с оценкой $d_8^{(1)} = 2 + 6 + 3 + 2 = 13$ и $\pi_8^{(2)} = \{e_{1,3}; e_{3,6}; e_{6,9}; e_{9,8}\}$ с оценкой $d_8^{(2)} = 3 + 6 + 3 + 2 = 14$. Для узла связи V_9 будут добавлены маршруты $\pi_9^{(1)} = \{e_{1,4}; e_{4,7}; e_{7,8}; e_{8,9}\}$ с оценкой $d_9^{(1)} = 2 + 8 + 2 + 2 = 14$, $\pi_9^{(2)} = \{e_{1,4}; e_{4,6}; e_{6,9}\}$ с оценкой $d_9^{(2)} = 2 + 6 + 3 = 11$ и $\pi_9^{(3)} = \{e_{1,3}; e_{3,6}; e_{6,9}\}$ с оценкой $d_9^{(3)} = 3 + 6 + 3 = 12$. После того как новые списки маршрутов для узлов связи V_8 и V_9 сформированы, из них необходимо выбрать маршрут с минимальной оценкой. Оптимальный маршрут до узла V_8 составит $\pi_8 = \{e_{1,4}; e_{4,7}; e_{7,8}\}$ с оценкой $d_8 = 12$. Оптимальный маршрут до узла V_9 составит $\pi_9 = \{e_{1,4}; e_{4,6}; e_{6,9}\}$ с оценкой $d_9 = 11$.

Таким образом, для графовой модели корпоративной сети, представленной на рис. 3, множество каналов дерева составит $E_T = \{e_{1,2}; e_{1,3}; e_{1,4}; e_{4,5}; e_{4,6}; e_{4,7}; e_{7,8}; e_{6,9}\}$; множество каналов замены

$E_S = \{e_{2,4}; e_{2,5}; e_{3,6}; e_{5,7}; e_{8,9}\}$; множество непарных каналов $E_P = \{e_{5,8}; e_{6,8}\}$.

При отказе канала связи $e_{4,7}$ в исходной структуре графа G корпоративная сеть примет вид, показанный на рис. 4.

На рис. 4 жирными линиями обозначено дерево оптимальных маршрутов, которое не требует изменения.

То есть при отказе канала связи $e_{4,7}$ дерево оптимальных маршрутов до узлов V_2, V_3, V_4, V_5 и V_6 согласно теореме 3 не изменится. Для узлов V_7 и V_8 необходимо пересчитать точки вхождения в дерево и во множество замены, а также переформировать списки оптимальных маршрутов и их маршрутов замены с учетом отказанного канала связи $e_{4,7}$. Оптимальный маршрут до узла V_7 составит $\pi_7 = \{e_{1,4}; e_{4,5}; e_{5,7}\}$ с оценкой $d_7 = 2 + 4 + 5 = 11$. Оптимальный маршрут до узла V_8 составит $\pi_8 = \{e_{1,4}; e_{4,5}; e_{5,7}; e_{7,8}\}$ с оценкой $d_8 = 2 + 4 + 5 + 2 = 13$.

Таким образом, для графовой модели корпоративной сети, представленной на рис. 4, множество каналов дерева составит $E_T = \{e_{1,2}; e_{1,3}; e_{1,4}; e_{4,5}; e_{4,6}; e_{5,7}; e_{7,8}\}$; множество каналов замены $E_S = \{e_{2,4}; e_{2,5}; e_{3,6}; e_{6,8}\}$; множество непарных каналов $E_P = \{e_{5,8}\}$.

При отказе узла связи V_7 в исходной структуре графа G корпоративная сеть примет вид, показанный на рис. 5.

На рис. 5 жирными линиями обозначено дерево оптимальных маршрутов, которое не требует изменения.

То есть при отказе узла связи V_7 дерево оптимальных маршрутов до узлов V_2, V_3, V_4, V_5 и V_6 согласно теореме 4 не изменится. Для узла V_8 необходимо пересчитать точки вхождения в дерево и в множество замены, а также переформировать списки оптимальных маршрутов и их маршрутов замены с учетом отказанного узла связи V_7 . Оптимальный маршрут до узла V_8 составит $\pi_8 = \{e_{1,4}; e_{4,6}; e_{6,8}\}$ с оценкой $d_8 = 2 + 6 + 6 = 14$.

Таким образом, для графовой модели корпоративной сети, представленной на рис. 5, множество каналов дерева составит $E_T = \{e_{1,2}; e_{1,3}; e_{1,4}; e_{4,5}; e_{4,6}; e_{6,8}\}$; множество каналов замены $E_S = \{e_{2,4}; e_{2,5}; e_{3,6}; e_{5,8}\}$.

Экспериментальная часть

Для подтверждения правильности математической модели и алгоритма парных перестановок маршрутов при динамических подключениях и отказах узлов и линий связи корпоративной сети разработано программное обеспечение моделирования процессов маршрутизации.

При разработке основное внимание уделялось корректности предлагаемого алгоритма и размерности решаемой задачи.

На рис. 6–8 представлены результаты моделирования алгоритма парных перестановок маршру-

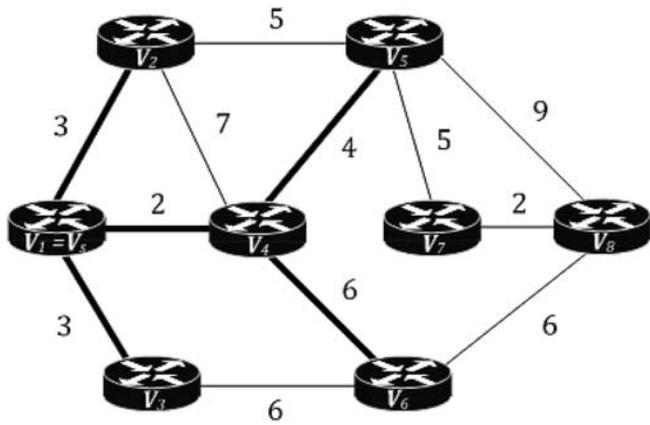


Рис. 4. Отказ канала связи $e_{4,7}$

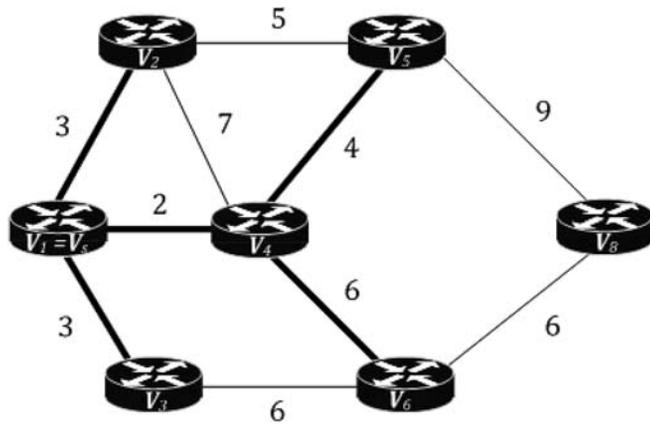


Рис. 5. Отказ узла связи V_7

тов при динамических подключениях и отказах узлов и линий связи корпоративной сети.

Для каждого испытания на множестве обработанных изменений выбирали минимальное, максимальное и среднее значения размерности задачи, выраженные через число узлов связи, для которых необходим поиск оптимальных маршрутов. Для каждого эксперимента были найдены значения оценок математического ожидания и среднего квадратичного отклонения числа изменений. Для алгоритма парных перестановок маршрутов определялось число фактически выполненных перестановок маршрутов при динамических подключениях и отказах узлов и линий связи в структуру корпоративной сети.

Были проведены исследования графовых моделей корпоративных сетей, состоящих из 10, 100 и 200 узлов связи. Исследование разработанного алгоритма парных перестановок маршрутов показало, что максимальное значение числа изменений существенно меньше размерности для каждой из рассмотренных моделей сети, а значение оценки математического ожидания не превышает единицы. Более того, обнаружена тенденция уменьшения значения оценки математического ожидания числа изменений дерева оптимальных маршрутов с увеличением числа узлов связи в сети.

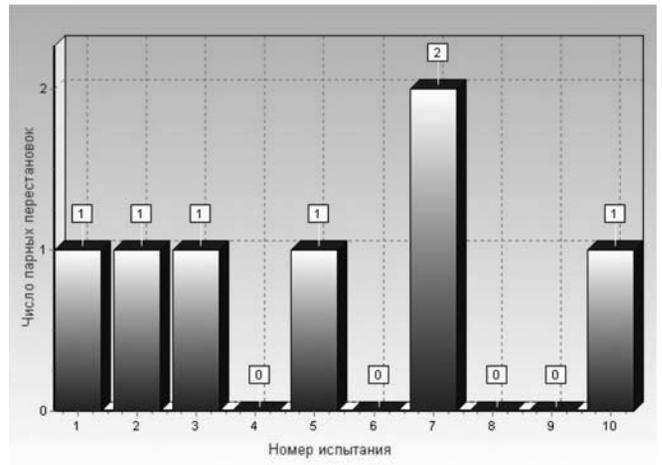


Рис. 6. Число изменений корпоративной сети из 10 узлов

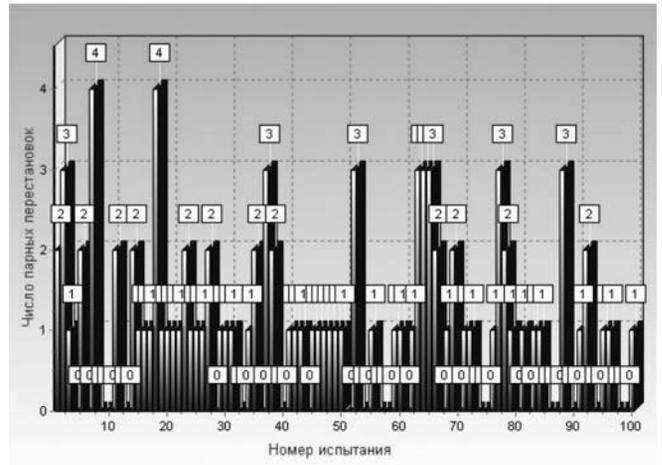


Рис. 7. Число изменений корпоративной сети из 100 узлов

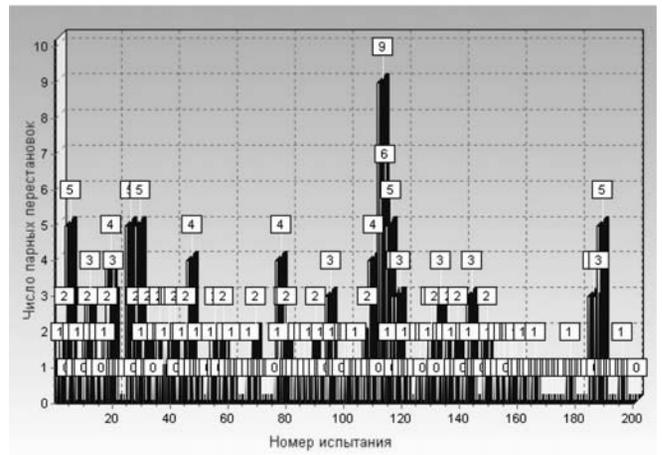


Рис. 8. Число изменений корпоративной сети из 200 узлов

В таблице для различного числа узлов связи N графа корпоративной сети приведены обобщенные статистические характеристики числа изменений дерева маршрутов: минимальное (Min) и максимальное (Max) значения, значения оценок математиче-

N	Min	Max	МО	СКО
10	0	0,2	0,07	0,0463
100	0	0,04	0,0096	0,0081
200	0	0,045	0,0044	0,0059

ского ожидания (МО) и среднего квадратичного отклонения (СКО).

При динамических подключениях и отказах узлов и линий связи корпоративной сети разработанный алгоритм парных перестановок маршрутов в отличие от существующих алгоритмов позволяет проводить перестроение таблицы маршрутизации не полностью, а только той ее части, в которой произошли изменения. Для этого необходимо один раз просмотреть списки оптимальных маршрутов и их маршрутов замены только для тех узлов сети, оценки длин которых изменились при подключениях и отказах узлов и линий связи корпоративной сети, и на основе предварительно собранной информации о списках маршрутов выполнить изменения для динамического формирования нового дерева оптимальных маршрутов. При этом трудоемкость изменения дерева оптимальных маршрутов является линейной функцией от числа узлов сети и определяется выражением $O(N)$.

На основе этого можно сделать вывод, что предложенный алгоритм парных перестановок маршрутов является эффективным в условиях динамического формирования структуры и параметров линий связи корпоративной сети при построении дерева оптимальных маршрутов за счет использования дополнительной информации о возможных маршрутах замены.

Заключение

Разработанные математическая модель и алгоритм парных перестановок маршрутов позволяют повысить эффективность функционирования корпоративной сети за счет уменьшения трудоемкости построения таблиц маршрутизации до значения порядка $O(N)$ в условиях динамического формирования структуры и параметров линий связи при построении дерева оптимальных маршрутов.

Работа выполнена при финансовой поддержке гранта Президента РФ для молодых ученых кандидатов наук МК-819.2014.9.

Список литературы

1. Олифер В. Г., Олифер Н. А. Основы компьютерных сетей. СПб.: Питер, 2009. 352 с.
2. Корячко В. П., Перепелкин Д. А. Корпоративные сети: технологии, протоколы, алгоритмы. М.: Горячая линия — Телеком, 2011. 219 с.
3. Корячко В. П., Перепелкин Д. А. Анализ и проектирование маршрутов передачи данных в корпоративных сетях. М.: Горячая линия — Телеком, 2012. 235 с.

4. Turner J. Openflow: Enabling innovation in campus networks // SIGCOMM Computer Communication Review. 2008. Vol. 38, N 2. P. 69–74.
5. Ижванов Ю. Л. Динамическая оценка состояния компьютерных сетей на основе метода сетевой томографии и задачи балансировки трафика // Информатизация образования и науки. 2013. № 3 (19). С. 35–40.
6. Шибанов А. П., Корячко В. П., Ижванов Ю. Л. Моделирование агрегированного телекоммуникационного канала с технологией открытых потоков // Радиотехника. 2012. № 3. С. 109–112.
7. Уваров Д. В., Перепелкин А. И. Алгоритм динамической маршрутизации в вычислительной сети // Вестник Рязанской государственной радиотехнической академии. 2003. № 12. С. 77–80.
8. Уваров Д. В., Перепелкин А. И., Корячко В. П. Построение дерева кратчайших путей в графе на основе данных о парных переходах // Системы управления и информационные технологии. 2004. № 4 (16). С. 93–96.
9. Корячко В. П., Перепелкин Д. А. Построение дерева оптимальных маршрутов корпоративной сети в условиях динамического изменения нагрузки на ее линиях связи // Информационные технологии: межвуз. сб. науч. тр. Рязань: Изд-во РГРТУ. 2011. С. 7–18.
10. Перепелкин Д. А., Перепелкин А. И. Разработка алгоритмов адаптивной маршрутизации в корпоративных вычислительных сетях // Вестник Рязанского государственного радиотехнического университета. 2006. № 19. С. 114–116.
11. Перепелкин А. И., Перепелкин Д. А. Разработка алгоритма динамической маршрутизации на базе протокола OSPF в корпоративных вычислительных сетях // Вестник Рязанского государственного радиотехнического университета. 2009. № 28. С. 68–72.
12. Корячко В. П., Перепелкин Д. А., Перепелкин А. И. Алгоритм динамической маршрутизации на базе протокола IGRP в корпоративных вычислительных сетях // Вестник Рязанского государственного радиотехнического университета. 2010. № 32. С. 94–98.
13. Перепелкин Д. А. Алгоритм адаптивной ускоренной маршрутизации на базе протокола OSPF при динамическом добавлении элементов корпоративной сети // Вестник Рязанского государственного радиотехнического университета. 2010. № 34. С. 65–71.
14. Перепелкин Д. А. Алгоритм адаптивной ускоренной маршрутизации на базе протокола OSPF при динамическом отказе элементов корпоративной сети // Вестник Рязанского государственного радиотехнического университета. 2011. № 37. С. 53–58.
15. Корячко В. П., Перепелкин Д. А. Алгоритм адаптивной ускоренной маршрутизации на базе протокола IGRP при динамическом добавлении элементов корпоративной сети // Вестник Рязанского государственного радиотехнического университета. 2012. № 40. С. 47–54.
16. Перепелкин Д. А. Алгоритм адаптивной ускоренной маршрутизации на базе протокола IGRP при динамическом отказе узлов и линий связи корпоративной сети // Вестник Рязанского государственного радиотехнического университета. 2012. № 42–1. С. 33–38.
17. Корячко В. П., Перепелкин Д. А., Перепелкин А. И. Алгоритм парных перестановок маршрутов в корпоративных сетях // Системы управления и информационные технологии. 2010. Т. 40. № 2. С. 51–56.
18. Корячко В. П., Перепелкин Д. А., Перепелкин А. И. Повышение эффективности функционирования корпоративных сетей при динамических изменениях в их структуре и нагрузках на линии связи // Вестник Рязанского государственного радиотехнического университета. 2010. № 33. С. 49–55.
19. Перепелкин Д. А., Перепелкин А. И. Повышение качества функционирования корпоративных сетей на базе протокола IGRP // Качество. Инновации. Образование. 2010. № 8. С. 56–60.
20. Перепелкин Д. А., Перепелкин А. И. Повышение качества функционирования корпоративных сетей на базе протокола OSPF // Качество. Инновации. Образование. 2010. № 12. С. 51–56.
21. Перепелкин Д. А., Перепелкин А. И. Алгоритм адаптивной ускоренной маршрутизации в условиях динамически изменяющихся нагрузок на линиях связи корпоративной сети // Информационные технологии. 2011. № 3. С. 2–7.

БЕЗОПАСНОСТЬ ИНФОРМАЦИИ CRYPTOSAFETY INFORMATION

УДК 004.056:003.26

Т. М. Волосатова, канд. техн. наук, доц.,
МВТУ имени Н. Э. Баумана, e-mail: tamaravo@gmail.ru

Исследования стеганографических методов защиты проектной документации от несанкционированного доступа

Основой публикации являются результаты проведенных исследований по разработке метода и средств стеганографической защиты проектной документации, продуцируемой в САПР, от несанкционированного доступа. Приведены результаты теоретико-экспериментального анализа возможностей основных алгоритмов цифровой стеганографии, предназначенных для сокрытия данных в различных контейнерах. При анализе учтены существенные особенности структуры проектной документации, определяющие выбор, либо синтез того или иного стеганографического алгоритма.

Ключевые слова: проектная документация, средства стеганографической защиты, цифровая стеганография, алгоритмы цифровой стеганографии, структура проектной документации

T. M. Volosatova

Researches of Steganographic Methods to Protect the Design Documentation from Unauthorized Access

The basis of publications are the results of researches on the development of the method and means of protection steganography CAD project documentation from unauthorized access. The results of theoretical and experimental analysis of the characteristic features of digital steganography algorithms designed to hide data in a variety of containers. Taken into account in the analysis of the essential features of the structure of the project documentation, defining the selection or synthesis of a steganographic algorithm.

Keywords: design documentation, tools steganographic security, digital steganography, digital steganography algorithms, the structure of the project documentation

Введение

Основой публикации являются результаты проведенных исследований по разработке метода и средств стеганографической защиты проектной документации, продуцируемой в САПР, от несанкционированного доступа. Проведен теоретико-экспериментальный анализ возможностей основных алгоритмов цифровой стеганографии, предназначенных для сокрытия данных в различных контейнерах. В настоящей работе выбирается и обосновывается стеганографический метод сокрытия проектных данных для защиты от несанкционированного доступа (НСД). Для читателей, впервые столкнувшихся со стеганографией, раскрываем принятые в работе следующие основные положения стеганографии [1].

• **Стеганографическая система** (стегосистема) — объединение методов и средств, используемых

для создания скрытого канала для передачи информации.

- **Сообщение (стега)** — это термин, используемый для общего названия передаваемой скрытой информации, будь то лист с надписями молоком, голова раба или цифровой файл.
- **Контейнер** — любая информация, используемая для сокрытия тайного сообщения.
- **Пустой контейнер** — контейнер, не содержащий секретного послания.
- **Заполненный контейнер** (стегоконтейнер) — контейнер, содержащий секретное послание.
- **Стеганографический канал** (стегоканал) — канал передачи стегоконтейнера.
- **Ключ** (стегоключ) — секретный ключ, нужный для сокрытия стегоконтейнера. Ключи в стегосистемах бывают двух типов: секретные и открытые.

Если стегосистема использует секретный ключ, то он должен быть создан или до начала обмена сообщениями, или передан по защищенному каналу. Стегосистема, использующая открытый ключ, должна быть устроена таким образом, чтобы невозможно было получить из открытого закрытый ключ. В этом случае открытый ключ мы можем передавать по незащищенному каналу.

- **Стегоанализ (стеганоанализ)** — раздел стеганографии; наука о выявлении факта передачи скрытой информации в анализируемом сообщении.
- **Робастность** — устойчивость к атакам и помехам.
- **Объем контейнера (стега)** измеряется в байтах.
- **Бит** — по Шеннону единица количества информации, по Хартли — структурная мера (один двоичный разряд).
- **Байт** — 8 бит.
- **Информация** — по Шеннону мера неопределенности в пространстве событий.

Как известно, для защиты проектной документации возможно применение одного из двух методов: стеганографического и криптографического (рис. 1).

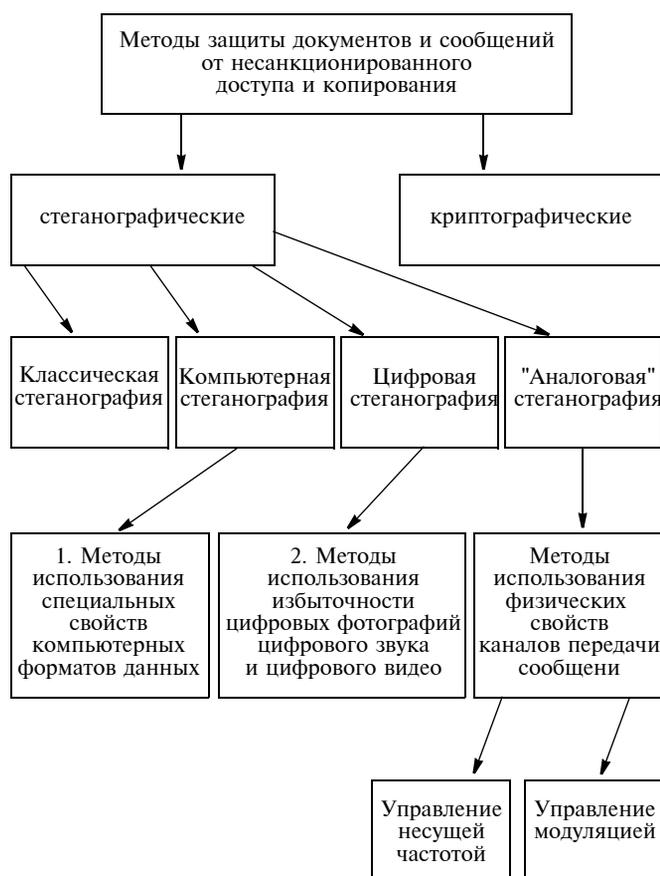


Рис. 1. Классификация методов сокрытия данных и сообщений

Для читателя, не знакомого с криптографией, далее приводятся принятые в работе основные понятия криптографии.

- **Открытый (исходный) текст** — данные (не обязательно текстовые), передаваемые без использования криптографии.
- **Шифрованный (закрытый) текст** — данные, полученные после применения криптосистемы с указанным ключом.
- **Криптосистема** — семейство обратимых преобразований открытого текста в шифрованный.
- **Ключ** — параметр шифра, определяющий выбор конкретного преобразования данного текста. В современных шифрах алгоритм шифрования известен, и криптографическая стойкость шифра целиком определяется секретностью ключа.
- **Криптоанализ** — наука, изучающая математические методы нарушения конфиденциальности и целостности информации.
- **Криптоаналитик** — человек, создающий и применяющий методы криптоанализа.
- **Криптографическая атака** — попытка криптоаналитика вызвать отклонения в атакуемой защищенной системе обмена информацией.
- **Взлом или вскрытие** — успешная криптографическая атака.
- **Шифрование** — процесс нормального применения криптографического преобразования открытого текста на основе алгоритма и ключа, в результате которого возникает шифрованный текст.
- **Расшифровывание** — процесс нормального применения криптографического преобразования шифрованного текста в открытый.
- **Дешифрование (дешифровка)** — процесс извлечения открытого текста без знания криптографического ключа на основе известного шифрованного. Термин дешифрование обычно применяют по отношению к процессу криптоанализа шифротекста (криптоанализ сам по себе, вообще говоря, может заключаться и в анализе шифросистемы, а не только зашифрованного ею открытого сообщения).

Криптографические методы для решения задачи защиты проектной документации и сообщений от несанкционированного доступа представляются не всегда применимыми по следующим причинам:

- применение криптографических методов весьма строго регламентировано ФСТЭК и аналогичными контролирующими организациями и не всегда приемлемо в коммерческих проектных организациях;
- собственно факт защиты документации от НСД очевиден;
- использование криптографии для шифрования чертежей и схем весьма проблематично, если не невозможно.

В настоящей работе предлагается применять стеганографические алгоритмы для шифрования проектной документации в целях защиты от несанкционированного доступа.

Рассмотрим наиболее распространенные стеганографические методы и алгоритмы. Для защиты проектной документации, хранимой в информационной подсистеме САПР крупной организации либо передаваемой по внутренним каналам передачи сообщений, наиболее приемлемой представляется группа методов компьютерной стеганографии, приведенная на рис. 1.

В работе проведено исследование основных стеганографических методов и алгоритмов сокрытия данных от несанкционированного доступа [1–4] применительно к задачам информационной безопасности (ИБ) САПР. Далее приводятся и обсуждаются результаты исследований, проведенных автором, либо с ее участием, часть которых опубликованы в работах [5–20].

Основные методы компьютерной стеганографии можно разбить на шесть групп:

- первая группа — это методы, использующие резервированные поля расширения компьютерных форматов данных, например, *bad trake*;
- вторая группа — это методы специального форматирования текстовых файлов;
- третья группа — методы сокрытия в неиспользуемых местах гибких дисков. При использовании методов третьей группы информация записывается в обычно неиспользуемых местах дисков (например, в нулевой дорожке);
- четвертая группа — методы, использующие имитирующие функции (*mimic-function*). Эти методы основаны на генерации текстов и являются обобщением акростиха. Для тайного сообщения генерируется осмысленный текст, скрывающий само сообщение;
- пятая группа — методы удаления идентифицирующего файл заголовка. В этих методах скрываемое сообщение шифруется, а в результирующем удаляется идентифицирующий заголовок, оставляя только зашифрованные данные. Получатель заранее знает о передаче сообщения и имеет недостающий заголовок;
- шестая группа — это методы создания специальных файловых систем ("*FATnew*").

Во второй группе можно выделить три подгруппы методов:

- методы использования известного смещения слов, предложений, абзацев. Они основаны на изменении положения строк и расстановки слов в предложении, что обеспечивается вставкой дополнительных пробелов между словами;

- методы выбора определенных позиций букв (нулевой шифр). Акростих — частный случай этого метода (например, начальные буквы каждой строки образуют сообщение);
- методы использования специальных свойств полей форматов, не отображаемых на экране. Методы основаны на использовании специальных "невидимых", скрытых полей для организации сносок и ссылок (например, использование черного шрифта на черном фоне).

Эти методы, за исключением методов сокрытия части данных на экране дисплея рабочей станции АРМ, целесообразно применять для обмена данными с использованием мобильных носителей (FLASH, CD, DVD).

Для защиты данных от НСД в базах архивов и при передаче по открытым каналам предлагается применять методы цифровой стеганографии, среди которых можно выделить четыре группы.

1. Методы использования избыточности аудио- и визуальной информации. Это методы использования избыточности цифровых фотографии, звука и видео. Они используют младшие разряды цифровых отсчетов, содержат мало полезной информации. Их заполнение дополнительной информацией практически не влияет на качество восприятия, что и дает возможность сокрытия конфиденциальной информации.

2. Методы, учитывающие особенности зрения человека, а именно:

- разрешающая способность;
- спектральная чувствительность.

3. Методы, учитывающие особенности слуха человека:

- частотная характеристика;
- нелинейные свойства (на фоне шума).

4. Методы использования особенностей модуляции в каналах передачи аудиосообщений.

1. Анализ алгоритмов встраивания сообщений в изображения

В приведенных ниже алгоритмах и основных результатах проектная документация рассматривается как цифровые водяные знаки и условно обозначается аббревиатурой ЦВЗ. Все современные алгоритмы встраивания информации в контейнеры должны обеспечивать робастность встраиваемых ЦВЗ. В частности, изображения-контейнеры со встроенными ЦВЗ должны предусматривать возможность сжатия контейнера любым из известных методов.

Для оценки качества восстановленного изображения при проведении исследований методов цифровой стеганографии, опубликованных в работах [6–20], использована мера среднего квадратиче-

ского искажения, определяемая как среднее квадратическое отклонение (СКО):

$$\text{СКО} = \frac{1}{N} \sum_{i=1}^N (x_i - \hat{x}_i)^2,$$

где N — число пикселей в изображении; x_i, \hat{x}_i — значения пикселей исходного и восстановленного изображений.

Применена также модификация этой меры — пиковое отношение сигнал/шум (ПОСШ), определяемое как

$$\text{ПОСШ} = 10 \log_2 \frac{N \cdot 255^2}{\sum_{i=1}^N (x_i - \hat{x}_i)^2},$$

где 255 — максимальное значение яркости полутонового изображения (т. е. 8 бит/пиксель).

Оба из перечисленных критериев использованы при обработке экспериментальных данных, результаты которых приведены в настоящей работе.

По способу встраивания информации стегоалгоритмы можно разделить на линейные (аддитивные), нелинейные и другие. Алгоритмы аддитивного внедрения информации заключаются в линейной модификации исходного изображения, а ее извлечение в декодере выполняется корреляционными методами. При этом ЦВЗ обычно складывается с изображением-контейнером, либо "вплавляется" (*fusion*) в него. В нелинейных методах встраивания информации используется скалярное либо векторное квантование. Среди других методов определенный интерес представляют методы, использующие идеи фрактального кодирования изображений.

1.1. Алгоритмы встраивания данных в пространственной области

Преимуществом алгоритмов встраивания ЦВЗ в пространственной области [5–11] является то, что ЦВЗ внедряется в области исходного изображения и нет необходимости выполнять вычислительно громоздкие линейные преобразования изображений. ЦВЗ внедряется за счет манипуляций яркостью или цветовыми составляющими встраиваемого изображения

$$l(x, y) \in \{1, \dots, L\} = (r(x, y), g(x, y), b(x, y)),$$

$r(x, y)$ — яркость красной цветовой составляющей в точке с координатами x, y ; $g(x, y)$ — яркость зеленой цветовой составляющей в точке с координатами x, y ; $b(x, y)$ — яркость синей цветовой составляющей в точке с координатами x, y .

Большинство алгоритмов встраивания ЦВЗ в пространственную область аудиофайлов основаны

на использовании широкополосных сигналов (ШПС). Основной идеей применения ШПС в стеганографии является то, что данные внедряются в шумовой сигнал малой мощности. Так как используется сигнал малой мощности, то для защиты ЦВЗ применяют помехоустойчивые коды.

Алгоритм Катера (Kutter). Этот алгоритм [5, 6] предполагает, что изображение имеет RGB-кодировку. Встраивание выполняется в канал синего цвета, так как к синему цвету система человеческого зрения наименее чувствительна. Далее принято, что S_i — встраиваемый бит, $I = \{R, G, B\}$ — контейнер, $p = (x, y)$ — псевдослучайная позиция, в которой выполняется вложение. Секретный бит встраивается в канал синего цвета путем модификации яркости:

$$L(p) = L_r R(p) + L_g G(p) + [L_b \& \Delta L_b] B(p),$$

где $[L_b \& \Delta L_b]$ означает модификацию исходного значения яркости синей компоненты путем сложения либо изменения части битов S_i .

По полученным результатам можно сделать следующий вывод: чем больше в изображении зеленой компоненты и меньше синей и красной, тем меньше вероятность ошибки при разных коэффициентах робастности. При этом, как показали эксперименты [5], максимум вероятности составляет 0,57.

Алгоритм Брундокса (Bruyndonckx). В этом алгоритме Брундокса [4, 6] ЦВЗ представляет собой строку бит. Для повышения помехоустойчивости применяется код Боуза—Чоудхури—Хоквингема (БЧХ). Внедрение осуществляется за счет модификации яркости блока 8×8 пикселей. Процесс встраивания осуществляется в три этапа:

- первый этап — классификация пикселей внутри блока на две группы (A и B) с примерно однородными яркостями;
- второй этап — разбиение каждой группы на категории, определяемые данной сеткой;
- третий этап — модификация средних значений яркости каждой категории в каждой группе.

В процессе исследований [11] выделены два типа блоков: блоки с шумовым контрастом и блоки с резко выраженными перепадами яркости. В блоках второго типа зоны с отличающейся яркостью не обязательно должны располагаться вплотную друг к другу, не обязательно должны содержать равное число пикселей. Более того, некоторые пиксели вообще могут не принадлежать ни одной из зон. В блоках первого типа классификация особенно затруднена.

Для выполнения первого этапа, а именно классификации пикселей, значения их яркости сортируются по возрастанию. Далее находится точка, в которой наклон касательной к получившейся гистограмме максимален. Эта точка является грани-

цей, разделяющей две зоны в том случае, если наклон больше некоторого порога. В противном случае пиксели делятся между зонами поровну.

Во время второго этапа при сортировке пикселей на блоки накладывают маски, разные для каждой зоны и каждого блока. Назначение масок состоит в обеспечении секретности внедрения.

На третьем этапе во время проведения модификации ЦВЗ, представленные как множество пикселей, оказываются разделенными на пять подмножеств: две зоны (A и B), две категории, и пиксели, не принадлежащие какой-либо зоне (для блоков первого типа). Пусть $l_{1A}, l_{2A}, l_{1B}, l_{2B}$ — средние значения яркости для пикселей двух зон и категорий соответственно, причем $l_{1A} < l_{2A}, l_{1B} < l_{2B}$. Встраивание бита ЦВЗ S_i (модификация) и равенство значений яркостей в каждой зоне i осуществляются (например, в $i = 1$ с весами n_i, K) по формулам

$$s = \begin{cases} 1, & \left\{ \begin{array}{l} l'_{1A} > l'_{1B} \\ l'_{2A} > l'_{2B} \end{array} \right\} \\ 0, & \left\{ \begin{array}{l} l'_{1A} < l'_{1B} \\ l'_{2A} < l'_{2B} \end{array} \right\} \end{cases},$$

$$\frac{n_{1A}l'_{1A} + n_{1B}l'_{1B}}{n_{1A} + n_{1B}} = l_1 \text{ и } \frac{n_{2A}l'_{2A} + n_{2B}l'_{2B}}{n_{2A} + n_{2B}} = l_2,$$

где $l_{1A}, l_{2A}, l_{1B}, l_{2B}$ — среднее значение яркости для пикселей двух зон A, B ; s — значение бита.

Алгоритм извлечения ЦВЗ является обратным алгоритму внедрения. При этом вычисляются средние значения яркостей и находятся разности

$$s'' = \begin{cases} 0, & \text{если } l''_{1A} - l''_{1B} < 0 \text{ и } l''_{2A} - l''_{2B} < 0 \\ 1, & \text{если } l''_{1A} - l''_{1B} > 0 \text{ и } l''_{2A} - l''_{2B} > 0 \end{cases},$$

где $l_{1A}, l_{2A}, l_{1B}, l_{2B}$ — средние значения яркости для пикселей двух зон A, B ; s'' — значение восстановленного бита.

Как показали эксперименты [11], алгоритм позволяет устойчиво скрывать сообщение, объем которого не превышает 10 % объема контейнера.

Алгоритм Ленгелаара (Langelaar). Данный алгоритм [2, 6, 8, 9] также работает с блоками 8×8 . Вначале создается псевдослучайная маска нулей и единиц такого же размера $pat(x, y) \in \{0, 1\}$. Далее каждый блок V делится на два субблока V_0 и V_1 , в зависимости от значения маски. Для каждого субблока вычисляется среднее значение яркости, l_0 и l_1 . Далее

выбирается некоторый порог α , и бит ЦВЗ встраивается следующим образом:

$$S = \begin{cases} 1, & l_0 - l_1 > +\alpha \\ 0, & l_0 - l_1 < -\alpha \end{cases},$$

где l_0 и l_1 — среднее значение яркости для пикселей каждого субблока; S — значение бита.

Если это условие не выполняется, необходимо изменять значения яркости пикселей субблока V_1 . Для извлечения бита ЦВЗ вычисляются средние значения яркости субблоков — l_0 и l_1 . Разница между ними позволяет определить искомым бит:

$$S = \begin{cases} 1, & l_0 - l_1 > 0 \\ 0, & l_0 - l_1 < 0 \end{cases}.$$

Исследование эффективности метода показало, что увеличение порога α непосредственным образом влияет на качество сокрытия — чем больше порог, тем большие искажения стего вносит в изображение-контейнер. Эффект наблюдался как визуально, так и на построенных гистограммах. Увеличение порогового значения помимо того, что вносит искажения, также увеличивает неустойчивость изображения к шуму. При восстановлении сообщения из контейнера с шумом зафиксированы искажения (до 10 %). Исследования показали, что предельное значение заполняемости контейнера составляет 8 %.

Алгоритм Питаса (Pitas). В данном алгоритме [6] ЦВЗ представляет собой двумерный массив битов размером, равным изображению, причем число единиц в нем равно числу нулей. Существует несколько версий алгоритма, предложенного Питасом. Вначале предлагалось встраивать бит ЦВЗ в каждый пиксель изображения, но позже было решено использовать для этой цели блоки размером 2×2 или 3×3 пикселя, что делает алгоритм более робастным к сжатию или фильтрации. ЦВЗ складывается с изображением:

$$I'(x, y) = I(x, y) + \alpha s(x, y).$$

Здесь обозначения соответствуют ранее принятым, т. е. I — яркость; $\alpha \neq \text{const}$ — произвольный порог; S — значение бита.

В случае использования для внедрения блоков детектор ЦВЗ вычисляет среднее значение яркости этого блока. Отсюда появляется возможность неравномерного внедрения ЦВЗ в пиксели, т. е. величина $\alpha \neq \text{const}$. Таким образом можно получить ЦВЗ, оптимизированный по критерию робастности к процедуре сжатия алгоритмом JPEG. Для этого в блоке 8×8 элементов заранее вычисляют "емкость" каждого пикселя (с учетом дискретного

косинусного преобразования (ДКП) и матрицы квантования JPEG). Затем ЦВЗ внедряют в соответствии с вычисленной емкостью. Эта оптимизация выполняется раз и навсегда, и найденная маска применяется для любого изображения. Это как раз и является недостатком метода.

Алгоритм Роджена (Rongen). В этом алгоритме [6] также, как и в алгоритме Питаса, ЦВЗ представляет собой двумерную матрицу единиц и нулей с примерно равным их количеством. Пиксели, в которые можно внедрять единицы (то есть робастные к искажениям), определяются на основе некоторой характеристической функции (характеристические пиксели). Эта функция вычисляется локально, на основе анализа соседних пикселей. Характеристические пиксели составляют примерно 1/100 от общего числа, так что не все единицы ЦВЗ встраиваются именно в эти позиции. Детектор находит значения характеристических пикселей и сравнивает с имеющимся у него ЦВЗ. Если в изображении ЦВЗ не содержится, то в характеристических пикселях число единиц и нулей будет примерно поровну.

Достоинства этого алгоритма учтены авторами работы [5] при разработке комбинированного метода сокрытия изображения.

Алгоритм PatchWork. В основе этого алгоритма [6, 13] лежит статистический подход. Вначале псевдослучайным образом на основе ключа выбираются два пикселя изображения. Затем значение яркости одного из них увеличивается на некоторое значение (от 1 до 5), значение яркости другого — уменьшается на то же значение. Далее этот процесс повторяется большое число раз ($n \approx 10\,000$) и находится сумма значений всех разностей. По значению этой суммы судят о наличии или отсутствии ЦВЗ в изображении.

Если a_i и b_i — значения выбираемых пикселей, δ — значение приращения, то сумма разностей значений пикселей равна

$$S_n = \sum_{i=1}^n [(a_i + \delta) - (b_i - \delta)] = 2\delta n + \sum_{i=1}^n (a_i - b_i).$$

Математическое ожидание величины $\sum_{i=1}^n (a_i - b_i)$

(суммы разности значений пикселей в незаполненном контейнере) близко к нулю при достаточно большом n . Математическое ожидание величины S_n будет больше 2δ ; S_n имеет гауссовское распределение. В стегодетекторе в соответствии с ключом проверяется значение S_n , и в том случае, если оно значительно отличается от нуля, выносится решение о наличии ЦВЗ. Для повышения робастности

алгоритма вместо отдельных пикселей можно использовать блоки или *patches* (отсюда и название алгоритма). Использование блоков различного размера может рассматриваться как формирование спектра вносимого ЦВЗ шума (шейпинг), аналогично тому, как это применяется в современных модемах.

Алгоритм PatchWork является достаточно стойким к операциям сжатия изображения, его усечения, изменения контрастности. Основным недостатком алгоритма является его неустойчивость к аффинным преобразованиям, т. е. поворотам, сдвигу, масштабированию. Другой недостаток заключается в малой пропускной способности. Так, в базовой версии алгоритма для передачи 1 бита скрытого сообщения требуется 20 000 пикселей. Проведенные в работах [5, 13] исследования показали, что фоновая компонента (шум) существенно ограничивает возможности восстановления текста из стегоконтейнера. Когда человеческому глазу заметен эффект "размытия", процент восстанавливаемого текста стремится к нулю.

Метод LSB (Least Significant Bit — наименьший значащий бит). Суть этого метода заключается в замене последних значащих битов в контейнере (изображения, аудио- или видеозаписи) на биты скрываемого сообщения [6, 7]. Разница между пустым и заполненным контейнерами должна быть неощутима для органов восприятия человека. Градации, обусловленные подменой младших битов, незаметны для человека. Кроме того, эти градации могут вообще не отобразиться при использовании низкокачественных устройств вывода.

Методы LSB являются неустойчивыми ко всем видам атак и могут быть использованы только при отсутствии шума в канале передачи данных. Практическим преимуществом метода является возможность сокрытия данных как в изображениях, так и аудиофайлах. Ограничения по объему контейнера позволяют отдавать предпочтение сокрытию данных именно в аудиофайлах.

Обнаружение LSB-кодированного стего осуществляется по аномальным характеристикам распределения значений диапазона младших битов отсчетов цифрового сигнала.

1.2. Алгоритмы встраивания ЦВЗ в области преобразования

При проведении исследований алгоритмов встраивания ЦВЗ учтено, что реальные изображения не являются реализациями случайного процесса с равномерно распределенными значениями величин. Хорошо известно, и это используется в алгоритмах сжатия, что большая часть энергии изображений сосредоточена в низкочастотной части спектра. Отсюда и потребность в осуществлении

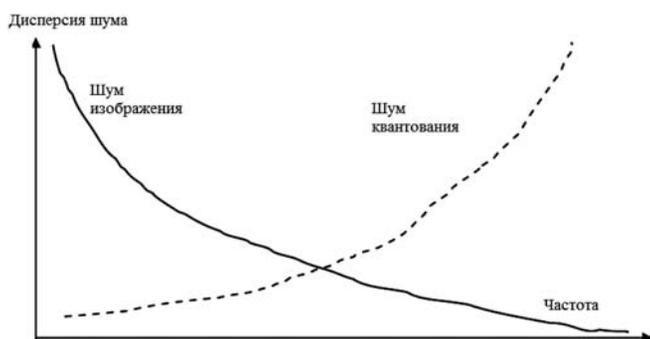


Рис. 2. Качественные графики зависимости дисперсий шума изображения и шума обработки (квантования) от частоты

декомпозиции изображения на субполосы. Стегосообщение добавляется к субполосам изображения. Низкочастотные субполосы содержат подавляющую часть энергии изображения и, следовательно, носят шумовой характер. Высокочастотные субполосы наиболее подвержены воздействию со стороны различных алгоритмов обработки, будь то сжатие или НЧ фильтрация. Таким образом, для вложения сообщения наиболее подходящими кандидатами являются среднечастотные субполосы спектра изображения.

На рис. 2 показано типичное распределение дисперсии шума изображения и результата его обработки по спектру частот [14]. Шум обработки появляется в результате квантования коэффициентов трансформанты. Его можно рассматривать как уменьшение корреляции между коэффициентами трансформанты исходного изображения и квантованными коэффициентами. Например, при высоких степенях сжатия может возникнуть ситуация, когда будут отброшены целые субполосы. То есть дисперсия шума в этих субполосах может достигать значительного значения. Налицо уменьшение корреляции между коэффициентами субполосы до квантования и после. Для получения приемлемых результатов необходимо усреднить значение шума обработки по многим изображениям.

Преобразования можно упорядочить по достигаемому выигрышу от кодирования [13]: единичное, Адамара, Хаара, ДКП, вейвлет, Карунена—Лоэва (ПКЛ). Под выигрышем от кодирования понимается степень перераспределения дисперсий коэффициентов преобразования. Наибольший выигрыш дает преобразование ПКЛ, наименьший — разложение по базису единичного импульса (то есть отсутствие преобразования). Преобразования, имеющие высокие значения выигрыша от кодирования, такие как ДКП и вейвлет-преобразование, характеризуются резко неравномерным распределением дисперсий коэффициентов субполос. Высокочастотные субполосы не подходят для вложения ввиду

большого шума обработки, а низкочастотные субполосы — ввиду высокого шума изображения. Поэтому приходится ограничиваться среднечастотными полосами, в которых шум изображения примерно равен шуму обработки. Поскольку таких полос немного, пропускная способность стегоканала невелика. В случае применения преобразования с более низким выигрышем от кодирования, например, преобразования Адамара или Фурье, имеется больше блоков, в которых шум изображения примерно равен шуму обработки. Для повышения пропускной способности стеганографического канала лучше применять преобразования с меньшими выигрышами от кодирования, хотя и плохо подходящие для сжатия.

Встраивание ЦВЗ в коэффициенты дискретного косинусного преобразования Фурье (ДКП). При использовании данного метода [5] контейнер разбивается на блоки размером 8×8 пикселей. ДКП применяется к каждому блоку, в результате чего получаются матрицы коэффициентов ДКП, также размером 8×8 . Коэффициенты ДКП обозначаются через $c_b(j, k)$, где b — номер блока; (j, k) — позиция коэффициента внутри блока. Если блок сканируется в зигзагообразном порядке (как это имеет место в JPEG), то коэффициенты обозначаются через $c_{b,j}$. Коэффициент в левом верхнем углу $c_b(0, 0)$ обычно называется DC-коэффициентом. Он содержит информацию о яркости всего блока. Остальные коэффициенты называются AC-коэффициентами. Иногда выполняется ДКП всего изображения, а не отдельных блоков. Далее рассмотрим некоторые из алгоритмов внедрения ЦВЗ в области ДКП.

Алгоритм Коча (Koch). В данном алгоритме [5, 13] в блок размером 8×8 осуществляется встраивание 1 бита ЦВЗ. Встраивание информации осуществляется следующим образом: для передачи бита 0 добиваются того, чтобы разность абсолютных значений коэффициентов была больше некоторой положительной величины, а для передачи бита 1 эта разность делается меньше некоторой отрицательной величины ε :

$$|c_b(j_{i,j}, k_{i,1})| - |c_b(j_{i,2}, k_{i,2})| > \varepsilon, \text{ если } s_i = 0;$$

$$|c_b(j_{i,j}, k_{i,1})| - |c_b(j_{i,2}, k_{i,2})| < -\varepsilon, \text{ если } s_i = 1.$$

Здесь j, k — адрес блока, подстрочные i, j — адрес внутри блока.

Алгоритм Бенхама (Benham). Этот алгоритм [5, 10] является улучшенной версией предыдущего алгоритма "Встраивание ЦВЗ в коэффициенты дискретного косинусного преобразования Фурье (ДКП)". Улучшения проведены по двум направлениям: для встраивания используются не все блоки, а лишь "пригодные" для этого; внутри блока для встраивания выбираются не два, а три коэффициента, что

уменьшает искажения. Пригодными для встраивания информации считаются блоки изображения с большими перепадами контраста. Для одних блоков характерно равенство нулю высокочастотных коэффициентов, для других блоков — очень большие значения нескольких низкочастотных коэффициентов. Эти особенности и являются критериями отсека не пригодных блоков.

При встраивании бита ЦВЗ псевдослучайно выбирают три коэффициента ДКП блока. Если нужно вложить 1, коэффициенты изменяют так, чтобы третий коэффициент стал меньше каждого из первых двух; если нужно встроить 0, его значение делают больше других. Если такая модификация приведет к слишком большой деградации изображения, коэффициенты не изменяют, и этот блок просто не используют. Изменение трех коэффициентов вместо двух, а тем более отказ от изменений в случае неприемлемых искажений уменьшает вносимые ЦВЗ погрешности. Декодер всегда сможет определить блоки, в которые ЦВЗ не встроено, повторив анализ, выполненный в кодере.

Исследования двух описанных выше методов показали ограниченные возможности по объему встраиваемого ЦВЗ (до 8 %) [9].

Алгоритм Подилчука (Podilchuk). При обнаружении ЦВЗ этот алгоритм [6] требует наличия у детектора исходного изображения. Встраиваемые данные моделируются вещественным случайным процессом с нормальным распределением, единичной дисперсией и нулевым средним. Для каждого коэффициента ДКП определяется значение порога, изменение которого может привести к деградации изображения. Этот порог зависит от позиции коэффициента в матрице (то есть частотного диапазона, за который он отвечает). Кроме того, порог обуславливается и свойствами самого изображения, т. е. контрастностью и яркостью блока.

Встраивание осуществляется следующим образом. Если абсолютное значение коэффициента меньше порога, то он не изменяется. В противном случае к нему прибавляется произведение значения порога и значения ЦВЗ. При обнаружении ЦВЗ вначале коэффициенты исходного изображения вычитаются из соответствующих коэффициентов модифицированного изображения. Затем вычисляется коэффициент корреляции и устанавливается факт наличия ЦВЗ. Таким образом, скрытность сокрытия весьма невысока.

Алгоритм Хсю (Hsu). В данном алгоритме [6, 14] декодеру ЦВЗ также требуется исходное изображение. Однако декодер определяет не факт наличия ЦВЗ, а выделяет встроенные данные. В качестве ЦВЗ выступает черно-белое изображение размером вдвое меньше контейнера. Перед встраиванием это изображение подвергается случайным перестановкам.

ЦВЗ встраивается в среднечастотные коэффициенты ДКП (четвертая часть от общего количества). Эти коэффициенты расположены вдоль второй диагонали матрицы ДКП.

Для внедрения бита ЦВЗ в коэффициент $c_b(j, k)$ определяют знак разности коэффициента текущего блока и соответствующего ему коэффициента из предыдущего блока. Если надо встроить 1, то коэффициент $c_b(j, k)$ меняют так, чтобы знак разности стал положительным, если надо встроить 0 — отрицательным.

Имеется ряд улучшений основного алгоритма. Во-первых, вместо алгебраических значений коэффициентов можно использовать их абсолютные значения. Во-вторых, вместо коэффициента из предыдущего блока можно использовать *DC*-коэффициент текущего блока. Также учитывается процесс квантования коэффициентов:

$$d_2(i) = \text{sign} \left(\left\lfloor \frac{|c_b(j, k)|}{Q(j, k)} \right\rfloor Q(j, k) - \left\lfloor \frac{|c_b(0, 0)|}{Q(0, 0)} \right\rfloor Q(0, 0) \right).$$

Здесь обозначения аналогичны тем, что приняты выше, а Q — номер блока.

Еще одним усовершенствованием этого алгоритма является порядок сортировки, при котором блоки ЦВЗ упорядочиваются по убыванию в них числа единиц. Блоки исходного изображения-контейнера также упорядочиваются по убыванию дисперсий. После этого выполняется соответствующее вложение ЦВЗ. Исследования, результаты которых опубликованы в работах [5, 14], показали, что данный алгоритм не является робастным по отношению к JPEG-компрессии.

Алгоритм Кокса (Cox). Этот алгоритм [6] является робастным ко многим операциям обработки сигнала. Обнаружение встроенного ЦВЗ в нем выполняется с использованием исходного изображения. Внедряемые данные представляют собой последовательность вещественных чисел с нулевым средним и единичной дисперсией. Для вложения информации используются несколько *AC*-коэффициентов ДКП всего изображения с наибольшей энергией.

Первый вариант можно использовать в случае, когда энергия ЦВЗ сравнима с энергией модифицируемого коэффициента. В противном случае либо ЦВЗ будет не робастным, либо искажения слишком большими. Поэтому так встраивать информацию можно лишь при незначительном диапазоне изменения значений энергии коэффициентов.

При обнаружении ЦВЗ выполняют обратные операции: вычисляют ДКП исходного и модифицированного изображений; находят разности между соответствующими коэффициентами наибольшего значения. Алгоритм принят авторами настоящей

работы во внимание при разработке комбинированного метода [14].

Алгоритм Барни (Barni). Этот алгоритм [6] является улучшением алгоритма Кокса, и в нем также выполняется ДКП всего изображения. В этом случае детектору уже не требуется исходное изображение, т. е. схема является "слепой" (без априорных данных [14]). Для встраивания ЦВЗ используют не наибольшие АС-коэффициенты, а средние по значению АС-коэффициенты. В качестве ЦВЗ выступает произвольная строка битов.

Выбранные коэффициенты модифицируют, далее выполняется обратное ДКП и проводится дополнительный шаг обработки — исходное и модифицированное изображения складывают с весовыми коэффициентами:

$$l''(x, y) = \beta(x, y)l'(x, y) + (1 - \beta)l(x, y),$$

где l'' , l' , l — соответствующие значения вышеописанных яркостей. В качестве коэффициентов β целесообразно использовать нормализованную дисперсию блоков. Здесь $\beta \approx 1$ для текстурированных областей (в которых человеческий глаз мало чувствителен к добавленному шуму) и $\beta \approx 0$ в однородных областях. Значение β находят не для каждого пикселя в отдельности, а для неперекрывающихся блоков фиксированного размера. В детекторе ЦВЗ вычисляется значение коэффициента корреляции между модифицированным изображением и ЦВЗ.

1.3. Аддитивные алгоритмы

Алгоритмы аддитивного внедрения информации заключаются в линейной модификации исходного изображения, а ее извлечение в декодере проводится корреляционными методами. При этом ЦВЗ обычно складывается с изображением-контейнером либо "вплавляется" (*fusion*) в него.

Алгоритмы на основе линейного встраивания ЦВЗ. В аддитивных методах внедрения [14] ЦВЗ представляет собой последовательность чисел w_i длины N , которая внедряется в выбранное подмножество отсчетов исходного изображения f . Основное и наиболее часто используемое выражение для встраивания информации в этом случае имеет вид

$$f'(m, n) = f(m, n)(1 + \alpha w_i),$$

здесь α — весовой коэффициент, а f' — модифицированный пиксель изображения. Другой способ встраивания водяного знака был предложен И. Коксом [14] и имеет вид

$$f'(m, n) = f(m, n) + \alpha w_i,$$

или, при использовании логарифмов коэффициентов, вид

$$f'(m, n) = f(m, n)e^{\alpha w_i}.$$

При встраивании в соответствии с первой формулой ЦВЗ в декодере определяется следующим образом:

$$w_i^* = \frac{f^*(m, n) - f(m, n)}{\alpha f(m, n)}.$$

Здесь под f^* понимаются отсчеты полученного изображения, содержащего или не содержащего ЦВЗ w . После извлечения w_i^* сравнивается с подлинным ЦВЗ, причем в качестве меры идентичности водяных знаков используется значение коэффициента корреляции последовательностей:

$$\delta = \frac{w^* w}{\|w^*\| \|w\|}.$$

Это значение варьируется в интервале $[-1; 1]$. Значения, близкие к единице, свидетельствуют о том, что извлеченная последовательность с большей вероятностью может соответствовать встроеному ЦВЗ. Следовательно, в этом случае делается заключение, что анализируемое изображение содержит водяной знак.

В декодере может быть установлен некоторый порог

$$\tau = \frac{\alpha}{SN} \sum |f'|$$

(здесь S — стандартное среднее квадратическое отклонение), который определяет вероятности ошибок первого и второго рода при обнаружении ЦВЗ. При этом коэффициент α может не быть постоянным, а может адаптивно изменяться в соответствии с локальными свойствами исходного изображения. Это позволяет сделать водяной знак более робастным.

Для увеличения робастности внедрения во многих алгоритмах применяют широкополосные сигналы. При этом информационные биты могут быть многократно повторены, закодированы с применением корректирующего кода либо к ним может быть применено какое-либо другое преобразование. Затем их модулируют с помощью псевдослучайной гауссовской последовательности. Такая последовательность является хорошей моделью шума, присутствующего в реальных изображениях. В то же время синтетические изображения (созданные на компьютере) не содержат шумов, и в них труднее незаметно встроить такую последовательность.

Для извлечения внедренной информации в аддитивной схеме встраивания ЦВЗ обычно необходимо иметь исходное изображение, что достаточно сильно ограничивает область применения подобных методов.

Также существуют слепые методы извлечения ЦВЗ [15], вычисляющие корреляцию последовательности w со всеми N коэффициентами полученного изображения f^* по формуле

$$\delta = \frac{\sum_N f^*(m, n)w_i}{N},$$

где δ — коэффициент корреляции.

Затем полученное значение коэффициента корреляции δ сравнивается с некоторым порогом обнаружения

$$\tau = \frac{\alpha}{3N} \sum_N |f^*(m, n)|.$$

Основным недостатком этого метода является то, что само изображение в этом случае рассматривается как шумовой сигнал. Существует гибридный подход (полуслепые схемы), когда часть информации об исходном изображении доступна в ходе извлечения информации, но неизвестно собственно исходное изображение.

Исследования подтвердили, что корреляционный метод стегоанализа позволяет легко обнаружить наличие или отсутствие ЦВЗ. Для получения же всех информационных битов нужно протестировать все возможные последовательности, что является предметом дополнительного исследования [14, 15].

2. Алгоритмы встраивания данных в аудиоконтейнеры

2.1. Эхо-кодирование

Данный алгоритм [12, 18] использует неравномерные промежутки между эхо-сигналами для кодирования последовательности значений. При наложении ряда ограничений соблюдается условие незаметности встраивания информации для человеческого восприятия. Эхо характеризуется тремя параметрами: начальная амплитуда, степень затухания, задержка. При достижении некоего порога между сигналом и эхом они смешиваются. Человек не может отличить эти два сигнала. В данном методе используется тот факт, что человеческое ухо воспринимает не столько форму, сколько энергию шума. Так как в формате MP3 полностью кодируется форма шумных частотных подполос, данные подполосы могут быть использованы для сокрытия данных. Входной сигнал преобразуется в частотную

область с помощью используемого в формате MP3 модифицированного дискретного косинусного преобразования (МДКП). Эксперименты показали [18—21], что при увеличении размера стего график спектра сигнала менялся незначительно, это подтверждает факт устойчивости эхо-метода к частотным атакам, сжатию. Метод эхо-кодирования не является универсальным. На качество кодирования влияет множество факторов: сам аудиосигнал, его амплитудные характеристики, слуховые качества распознающего и др. Также, при приближении к границе пропускной способности увеличивается вероятность появления одиночных ошибок. По сравнению с другими методами эхо-кодирование характеризуется весьма небольшим соотношением сигнал/шум (примерно четыре) и высоким параметром средней квадратической ошибки (порядка 10^5). Однако данный метод выигрывает за счет большой пропускной способности и устойчивости к амплитудным и частотным атакам.

2.2. Метод LSB

Этот метод в равной степени применим при использовании в качестве контейнера как с изображениями, так и аудиофайлами. В ходе экспериментов [12] со звуковыми файлами при использовании вейвлетов Добеши для выявления скрытых данных установлено, что если степень сжатия контейнера, проверяемого на наличие в нем скрытых данных файла, будет меньше некоторого порогового значения, то с большой вероятностью можно судить о присутствии в файле дополнительной информации. В противном случае контейнер можно считать пустым, т. е. не содержащим скрытого сообщения.

Заключение

Проведенный анализ и экспериментальные исследования показывают, что, в целом, стеганографические методы применимы для защиты проектной документации, продуцируемой в САПР, от несанкционированного доступа. Наиболее предпочтительно внедрение информации в спектральную область изображения. Если при этом используются линейные методы, то встраивание ЦВЗ проводят в средние полосы частот. Это объясняется тем, что энергия изображения сосредоточена в основном в низкочастотной области. Для преодоления воздействий фотопечати и сканирования наиболее успешными оказались методы, получившие название методов "модуляции" изображения-контейнера, причем модуляция может осуществляться как в частотных, так и в пространственных областях изображения. В случае внедрения ЦВЗ в частотной области модуляции подвергаются амплитудные составляющие комплексного спектра изображения-

контейнера. Для этого предварительно осуществляется вычисление амплитудной и фазовой составляющих компонентов преобразования Фурье. Для оценки последствий геометрических искажений, связанных со случайным поворотом, смещением или изменением масштаба, в изображение-контейнер, кроме ЦВЗ, встраивается изображение-шаблон. В случае внедрения ЦВЗ в пространственной области его сигнал встраивается путем модуляции исходного изображения-контейнера, а извлечение ЦВЗ (демодуляция) выполняется с помощью линейной фильтрации изображения. Если изображение цветное, сигнал ЦВЗ внедряется путем модификации значений пикселей в Blue-канале RGB-изображения. Однако комбинированные стеганографические методы, предложенные автором в работах [14–16], имеют значительные преимущества перед методами, рассмотренными выше. Моделирование, результаты которого приведены в работе [14], показало, что применение корреляционного метода распознавания, наиболее распространенного в стеганоанализе, к данным стегоалгоритмам малоэффективно. Так, при дефокусировке регистратора, изменении размеров или повороте искомого объекта в поле анализа не обнаруживается. Применение корреляционного стеганоанализа при реализации комбинированного метода стеганографии, сочетающего искусственную дефокусировку с последующим сокрытием сообщения, неэффективно. Аналогичные результаты получены при проведении теоретико-экспериментальных оценок метода прямого сравнения.

Список литературы

1. **Генне О. В.** Основные положения стеганографии // Защита информации. Конфидент, 2000. № 3. С. 5–25.
2. **Real-time Watermarking Techniques for Compressed Video Data** // Langelaar, Gerrit Cornelis — Thesis Delft University of Technology. (Veenendaal, 2000). V.: Universal Press. 136 p.
3. **Бендер В., Моримото Н., Лу А.** Методы сокрытия данных // IBM System Journal, 1996. July. P. 25–33.
4. **Bruyndonckx O., Quisquater J.-J., Macq B.** Spatial method of copyright labeling of digital images // IEEE Workshop on Non-linear Images/Signal Processing. Thessal. 1995. June. P. 19–27.
5. **Хузина Э. И.** Экспериментальные исследования алгоритма стеганографического сокрытия данных методом Катера // Сб. трудов Третьей всероссийской научно-технической конференции "Безопасные информационные технологии". (Москва, 2012). М.: Изд-во МГТУ им. Н. Э. Баумана, 2012. С. 169–172.
6. **Чичварин Н. В.** Сопоставительный анализ областей применения и граничных возможностей характерных стеганографических алгоритмов // Сб. трудов Третьей всероссийской научно-технической конференции "Безопасные информационные технологии". (Москва, 2012). М.: Изд-во МГТУ им. Н. Э. Баумана. С. 174–179.
7. **Ларионцева Е. Л., Стельмашук Н. Н.** Экспериментальные исследования эффективности стеганографического алгоритма, реализующего метод lsb // Сб. докладов Третьей всероссийской научно-технической конференции "Безопасные информационные технологии". (Москва, 2012). М.: Изд-во МГТУ им. Н. Э. Баумана. С. 94–96.

8. **Логинов К. Е.** Экспериментальные исследования устойчивости алгоритма стеганографического сокрытия данных методом Langelaar при воздействиях на стегоконтейнер // Сб. докладов Третьей всероссийской научно-технической конференции "Безопасные информационные технологии". (Москва, 2012). М.: Изд-во МГТУ им. Н. Э. Баумана. С. 99–101.

9. **Сиволапов А. С.** Исследование влияния контейнера на качество сокрытия сообщений методом Langelaar // Сб. докладов Третьей всероссийской научно-технической конференции "Безопасные информационные технологии". (Москва, 2012). М.: Изд-во МГТУ им. Н. Э. Баумана, 2012. С. 153–155.

10. **Круглая Е. И., Пилипенко А. В.** Защита данных в САПР: анализ стеганографических алгоритмов коча (koch) и бенхама (benham) // Сб. докладов Третьей всероссийской научно-технической конференции "Безопасные информационные технологии". (Москва, 2012). М.: Изд-во МГТУ им. Н. Э. Баумана, 2012. С. 87–90.

11. **Максимов Р. Л.** Экспериментальное исследование эффективности стеганографического алгоритма, реализующего метод Брундонкса (Bruyndonckx) // Сб. докладов Третьей всероссийской научно-технической конференции "Безопасные информационные технологии". (Москва, 2012). М.: Изд-во МГТУ им. Н. Э. Баумана, 2012. С. 101–105.

12. **Гончаров И. О., Заикин М. А.** Экспериментальные исследования стеганографического метода эхо-кодирования // Сб. докладов Третьей всероссийской научно-технической конференции "Безопасные информационные технологии". (Москва, 2012). М.: Изд-во МГТУ им. Н. Э. Баумана, 2012. С. 45–48.

13. **Иванова Е. Ю.** Обзор атак на стегоалгоритм PatchWork и методов противодействия // Сб. докладов Третьей всероссийской научно-технической конференции "Безопасные информационные технологии". (Москва, 2012). М.: Изд-во МГТУ им. Н. Э. Баумана, 2012. С. 66–69.

14. **Волосатова Т. М., Денисов А. В., Чичварин Н. В.** Комбинированные методы защиты данных в САПР // Информационные технологии. Приложение. 2012. № 5. С. 1–32.

15. **Волосатова Т. М., Денисов А. В., Чичварин Н. В.** Защита проектной документации от несанкционированного доступа // Сб. докладов 9-й Международной конференции "Эффективные методы автоматизации подготовки и планирования производства". (Москва, 2012). М.: Изд-во МГТУ им. Н. Э. Баумана, 2012. С. 141–144.

16. **Волосатова Т. М., Денисов А. В., Чичварин Н. В.** Метод сохранения данных с использованием искусственной дефокусировки // Сб. докладов 9-й Международной конференции "Эффективные методы автоматизации подготовки и планирования производства". (Москва, 2012). М.: Изд-во МГТУ им. Н. Э. Баумана, 2012. С. 145–148.

17. **Чичварин Н. В.** Стеганографический метод маскирования данных с использованием цифровых голограмм // Сб. докладов Всероссийской НТК "Безопасные информационные технологии". (Москва, 2011). М.: Изд-во МГТУ им. Н. Э. Баумана, 2012. С. 87–91.

18. **Ремизов А. В., Филиппов М. В., Чичварин Н. В.** Методы защиты информации в звуковых файлах // Информационные технологии. 2009. № 10. С. 10–19.

19. **Ремизов А. В., Филиппов М. В.** Сокрытие информации с использованием стеганографической файловой системы // Сб. докладов Второй всероссийской конференции "Безопасные информационные технологии". (Москва, 2011). М.: Изд-во МГТУ им. Н. Э. Баумана. С. 113–116.

20. **Ремизов А. В., Филиппов М. В.** Оценка необнаружимости стеганографических алгоритмов // Наука и образование. МГТУ им. Н. Э. Баумана. Электрон. Журн. 2012. № 3. Режим доступа: <http://technomag.edu.ru/doc/77-30569/359383.html> (дата обращения 08.05.13).

21. **Ремизов А. В., Филиппов М. В.** Необнаружимость методов сокрытия данных // Труды Десятого международного симпозиума "Интеллектуальные системы" (Вологда, 2012). Вологда: Изд. ВГУ. С. 487–491.

XVI Всероссийская научная конференция RCDL-2014 "Электронные библиотеки: перспективные методы и технологии, электронные коллекции"

13–16 октября 2014 года в Объединенном институте ядерных исследований (г. Дубна) при поддержке Российского фонда фундаментальных исследований, Института проблем информатики РАН, Московской секции ACM SIGMOD будет проводиться очередная XVI Всероссийская научная конференция с международным участием RCDL-2014 — "Электронные библиотеки: перспективные методы и технологии, электронные коллекции".

Электронные библиотеки являются объектом исследований и разработок, направленных на развитие теории и практики обработки, распространения, хранения, поиска и анализа цифровых объектов различной природы. Эти современные репозитории знаний представляют собой сложные онлайн-информационные системы.

Серия Всероссийских научных конференций RCDL, труды которых представлены на постоянном сайте <http://rcdl.ru>, нацелена на формирование российского корпуса международного сообщества ученых, развивающих это научное направление.

Научные электронные библиотеки играют важную роль в реализации активно поддерживаемых научным сообществом инициатив открытого доступа к результатам исследований, в формировании открытой науки.

В рамках конференции RCDL-2014 планируется проведение традиционного семинара молодых ученых "Диссертационные исследования по тематике информационных технологий, связанных с электронными библиотеками", на котором авторам работ, отобранных на основе предварительного рецензирования, будет дана возможность представить текущие результаты своих исследований, а также обсудить их сильные и слабые стороны с более опытными коллегами.

Рабочие языки конференции — русский и английский, представление расширенных тезисов и доклада возможно на любом из них. Программный комитет ожидает представления не опубликованных ранее работ, имеющих преимущественно *исследовательский характер* и содержащих *оригинальные* результаты.

Труды конференции будут опубликованы в виде сборника текстов принятых докладов, а также в электронном виде на постоянном сайте конференций RCDL (<http://rcdl.ru>) и в европейском репозитории трудов научных конференций CEUR Workshop Proceedings (<http://ceur-ws.org/>). Тексты лучших докладов, представленных на конференцию, будут рекомендованы к публикации (в доработанном виде) в периодических изданиях, включенных в список ВАК таких, как "Информатика и ее применения", "Системы и средства информатики", "Программная инженерия".

Расширенные тезисы докладов объемом не менее 2000 слов с библиографией принимаются Программным комитетом в электронном виде до 15 мая 2014 года, заявки на проведение тьюториалов и сопутствующих семинаров — до 20 мая 2014 года. Заявленные доклады будут рецензироваться членами Программного комитета. По результатам рецензирования Программный комитет примет решение о принятии или отклонении заявленных докладов и известит о нем авторов до 15 июня 2014 года. Готовый для публикации в трудах конференции текст принятого доклада должен быть представлен в Программный комитет не позднее 1 августа 2014 года.

Более подробная информация о конференции содержится на ее сайте (<http://rcdl2014.jinr.ru>).

Адрес редакции:

107076, Москва, Стромынский пер., 4

Телефон редакции журнала (499) 269-5510

E-mail: it@novtex.ru

Дизайнер Т.Н. Погорелова. Технический редактор Е.В. Конова.

Корректор Е.В. Комиссарова.

Сдано в набор 06.02.2014. Подписано в печать 19.03.2014. Формат 60×88 1/8. Бумага офсетная.

Усл. печ. л. 8,86. Заказ IT414. 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.